HANDBOOK OF
Energy Efficiency and Renewable Energy

Mechanical Engineering Series

Frank Kreith & Roop Mahajan - Series Editors

Published Titles

Distributed Generation: The Power Paradigm for the New Millennium
Anne-Marie Borbely & Jan F. Kreider
Elastoplasticity Theory
Vlado A. Lubarda
Energy Audit of Building Systems: An Engineering Approach
Moncef Krarti
Finite Element Method Using MATLAB, 2nd Edition
Young W. Kwon & Hyochoong Bang
Fluid Power Circuits and Controls: Fundamentals and Applications
John S. Cundiff
Fundamentals of Environmental Discharge Modeling
Lorin R. Davis
Handbook of Energy Efficiency and Renewable Energy
Frank Kreith and D. Yogi Goswami
Heat Transfer in Single and Multiphase Systems
Greg F. Naterer
Introductory Finite Element Method
Chandrakant S. Desai & Tribikram Kundu
Intelligent Transportation Systems: New Principles and Architectures
Sumit Ghosh & Tony Lee
Mathematical & Physical Modeling of Materials Processing Operations
Olusegun Johnson Ilegbusi, Manabu Iguchi & Walter E. Wahnsiedler
Mechanics of Composite Materials
Autar K. Kaw
Mechanics of Fatigue
Vladimir V. Bolotin
Mechanism Design: Enumeration of Kinematic Structures According to Function
Lung-Wen Tsai
MEMS: Applications
Mohamed Gad-el-Hak
MEMS: Design and Fabrication
Mohamed Gad-el-Hak
The MEMS Handbook, Second Edition
Mohamed Gad-el-Hak
MEMS: Introduction and Fundamentals
Mohamed Gad-el-Hak
Multiphase Flow Handbook
Clayton T. Crowe
Optomechatronics: Fusion of Optical and Mechatronic Engineering
Hyungsuck Cho
Practical Inverse Analysis in Engineering
David M. Trujillo & Henry R. Busby
Pressure Vessels: Design and Practice
Somnath Chattopadhyay
Principles of Solid Mechanics
Rowland Richards, Jr.
Thermodynamics for Engineers
Kau-Fui Wong

Vibration and Shock Handbook
Clarence W. de Silva
Viscoelastic Solids
Roderic S. Lakes

Handbook of
Energy Efficiency and Renewable Energy

Edited by

Frank Kreith

D. Yogi Goswami

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an informa business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-10: 0-8493-1730-4 (Hardcover)
International Standard Book Number-13: 978-0-8493-1730-9 (Hardcover)

Library of Congress Cataloging-in-Publication Data

Handbook of energy efficiency and renewable energy / edited by Frank Kreith and D. Yogi Goswami.
p. cm. -- (Mechanical engineering series)
Includes bibliographical references and index.
ISBN 0-8493-1730-4 (978-0-8493-1730-9 : alk. paper)
1. Energy conservation--Handbooks, manuals, etc. 2. Renewable energy sources--Handbooks, manuals, etc. I. Kreith, Frank. II. Goswami, D. Yogi. III. Title. IV. Series.

TJ163.3.H356 2007
333.79--dc22 2006024895

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

"With the exception of preventing war, this (the energy crisis) is the greatest challenge our country will face during our lifetime."

—President Jimmy Carter, April 18, 1977

Preface

Purpose

The goal of this handbook is to provide information necessary for engineers, energy professionals, and policy makers to plan a secure energy future. The time horizon of the handbook is limited to approximately 25 years because environmental conditions vary, new technologies emerge, and priorities of society continuously change. It is therefore not possible to make reliable projections beyond that period. Given this time horizon, the book deals only with technologies that are currently available or are expected to be ready for implementation in the near future.

Overview

Energy is a mainstay of an industrial society. As the population of the world increases and people strive for a higher standard of living, the amount of energy necessary to sustain our society is ever increasing. At the same time, the availability of nonrenewable sources, particularly liquid fuels, is rapidly shrinking. Therefore, there is general agreement that to avoid an energy crisis, the amount of energy needed to sustain society will have to be contained and, to the extent possible, renewable sources will have to be used. As a consequence, conservation and renewable energy technologies are going to increase in importance and reliable, up-to-date information about their availability, efficiency, and cost is necessary for planning a secure energy future.

The timing of this handbook also coincides with a new impetus for the use of renewable energy. This impetus comes from the emergence of renewable portfolio standards [RPS] in many states of the U.S. and renewable energy policies in Europe, Japan, Israel, China, Brazil, and India. An RPS requires that a certain percentage of energy used be derived from renewable resources. RPSs and other incentives for renewable energy are currently in place in 20 of the 50 states of the U.S. and more states are expected to soon follow. The details of the RPS for renewable energy and conservation instituted by state governments vary, but all of them essentially offer an opportunity for industry to compete for the new markets. Thus, to be successful, renewable technologies will have to become more efficient and cost-effective. Although RPSs are a relatively new development, it has already been demonstrated that they can reduce market barriers and stimulate the development of renewable energy. Use of conservation and renewable energy can help meet critical national goals for fuel diversity, price stability, economic development, environmental protection, and energy security, and thereby play a vital role in national energy policy.

The expected growth rate of renewable energy from portfolio standards and other stimulants in the U.S. is impressive. If current policies continue to be implemented, by the year 2017 almost 26,000 megawatts of new renewable energy will be in place in the U.S. alone. In 2005, photovoltaic production in the world has already topped 1000 MW per year and is increasing at a rate of over 30%. In Germany, the electricity feed-in laws that value electricity produced from renewable energy resources much higher than that from conventional resources, have created demand for photovoltaic and wind power. As a result, over the last three years the photovoltaic power has grown at a rate of more than 51% per year and wind power has grown at a rate of more than 37% in Germany. Recently, a number of other European countries have adopted feed-in laws similar to Germany. In fact, growth of both photovoltaic and wind power has averaged in the range of 35% in European countries. Similar policy initiatives in Japan indicate that renewable energy technologies will play an increasingly important role in fulfilling future energy needs.

Organization and Layout

The book is essentially divided into three parts:

- General overviews and economics
- Energy conservation
- Energy generation technologies

The first chapter is a survey of current and future worldwide energy issues. The current status of energy policies and stimulants for conservation and renewable energy is treated in Chapter 2 for the U.S. as well as several other countries. Economic assessment methods for conservation and generation technologies are covered in Chapter 3, and the environmental costs of various energy generation technologies are discussed in Chapter 4. Use of renewables and conservation will initiate a paradigm shift towards distributed generation and demand-side management procedures that are covered in Chapter 5.

Although renewables, once in place, produce energy from natural resources and cause very little environmental damage, energy is required in their initial construction. One measure of the energy effectiveness of a renewable technology is the length of time required, after the system begins operation, to repay the energy used in its construction, called the energy payback period. Another measure is the energy return on energy investment ratio. The larger the amount of energy a renewable technology delivers during its lifetime compared to the amount of energy necessary for its construction, the more favorable its economic return on the investment will be and the less its adverse environmental impact. But during the transition to renewable sources a robust energy production and transmission system from fossil and nuclear technologies is required to build the systems. Moreover, because there is a limit to how much of our total energy needs can be met economically in the near future, renewables will have to coexist with fossil and nuclear fuels for some time. Futhermore, the supply of all fossil and nuclear fuel sources is finite and their efficient use in meeting our energy needs should be a part of an energy and CO_2 reduction strategy. Therefore, Chapter 6 gives a perspective on the efficiencies, economics, and environmental costs of the key fossil and nuclear technologies. Finally, Chapter 7 provides projections for energy supply, demand, and prices of energy through the year 2025.

The U.S. transportation system relies 97% on oil and more than 60% of it is imported. Petroleum engineers predict that worldwide oil production will reach its peak within the next 10 years and then begin to decline. At the same time, demand for liquid fuel by an ever-increasing number of vehicles, particularly in China and India, is expected to increase significantly. As a result, gasoline prices will increase precipitously unless we reduce gasoline consumption by increasing the mileage of the vehicle fleet, reducing the number of vehicles on the road by using mass transport, and producing synthetic fuels from biomass and coal. The options to prevent an energy crisis in transportation include: plug-in hybrid vehicles, biofuels, diesel engines, city planning, and mass-transport systems. These are treated in Chapter 8; biofuels and fuel cells are treated in Chapter 24 and Chapter 27, respectively.

It is an unfortunate fact of life that the security of the energy supply and transmission system has recently been placed in jeopardy from various sources, including natural disasters and worldwide terrorism. Consequently, energy infrastructure security and risk analysis are an important aspect of planning future energy transmission and storage systems, and these topics are covered in Chapter 9.

Energy efficiency is defined as the ratio of energy required to perform a specific task or service to the amount of energy used for the process. Improving energy efficiency increases the productivity of basic energy resources by providing the needs of society with less energy. Improving the efficiency across all sectors of the economy is therefore an important objective. The least expensive and most efficient means in this endeavor is energy conservation, rather than more energy production. Moreover, energy conservation is also the best way to protect the environment and reduce global warming.

Recognizing that energy conservation in its various forms is the cornerstone of a successful national energy strategy, eight chapters (10 through 17) are devoted to conservation. The topics covered include energy management strategies for industry and buildings, HVAC controls, cogeneration, and advances in specific technologies, such as motors, lighting, appliances, and heat pumps.

In addition to the energy conservation topics covered in this handbook, there is another area for reducing energy consumption: waste management. As materials are mined, manufactured, and used to produce consumer goods, energy is required and wastes are generated. The utilization of the raw materials and the disposal of waste is a complicated process that offers opportunities for energy conservation. The U.S. Environmental Protection Agency (EPA) has identified four basic waste management options: (1) source reduction, (2) recycling and composting, (3) combustion (waste to energy incineration) and (4) landfilling. These four options are interactive—not hierarchical—because measures taken in one sector can influence another. For example, removing metals in the recycling step will increase the efficiency of combustion and facilitate landfilling.

In the course of waste management, many opportunities exist for conservation and reduction in energy consumption. The most obvious one is source reduction. Energy is saved if consumers buy less or use products more efficiently. Recycling and composting will also reduce energy requirements and preserve natural resources. From our perspective, however, combustion of waste and extraction of energy from the municipal waste stream are technologies that generate energy without requiring new primary sources and therefore fall within the purview of this handbook. The reader interested in energy-efficient integrated waste management of municipal wastes is referred to an extensive treatment in *The Handbook of Solid Waste Management*, 2nd ed., edited by G. Tchobanoglous and F. Kreith (2002), and for management of industrial solid wastes to *The Handbook of Industrial and Hazardous Waste Treatment*, 2nd ed., edited by L. K. Wang et al. (2004).

The third part of the book deals with energy storage and energy generation from renewable sources. One of the most challenging tasks, especially for renewable energy systems that cannot operate continuously, is energy storage. The main reason why fossil fuels, such as petroleum and coal, are such convenient energy sources is probably their easy storage. But there are new and relatively undeveloped storage technologies available that are discussed and evaluated in Chapter 18, which also includes the design of a robust and intelligent electric energy transmission system. Chapter 19 presents the availability of renewable sources: solar, wind, municipal waste, and biomass. The renewable generation technologies for solar thermal, wind power, photovoltaics, biomass, and geothermal are then covered in Chapters 20 through 25.

At this time, it is not clear whether hydrogen will play a major role in the national energy structure within the next 25 years, but there is ongoing discussion about the feasibility and cost of what is called "the hydrogen economy." Energy experts recognize that the generation and use of hydrogen have a critical inefficiency problem that is rooted in basic thermodynamics. This inefficiency renders hydrogen impractical and uneconomical as an energy carrier compared to electricity. There are also ground transportation options that are less expensive than using hydrogen vehicles powered by fuel cells. But there is substantial support for continuing research to eventually develop a viable place for hydrogen in a future energy structure. Therefore, the topics of hydrogen energy and fuel cells are included in Chapters 26 and 27, respectively. This information should be useful background for comparing competing options for energy generation, storage, and distribution.

We hope that this handbook will serve as a useful reference to all engineers in the energy field and pave the way for a paradigm shift from fossil fuels to a sustainable energy systems based on conservation and renewable technologies. But we also recognize the complexity of this task, and we invite readers to comment on the scope and the topics covered. A handbook such as this needs to be updated every 5 to 10 years and we will respond to readers' comments and suggestions in the next edition.

Frank Kreith
D. Yogi Goswami
Editors-in-Chief

Acknowledgment

The editors would like to acknowledge the assistance of Beverly Weiler, Jessica Todd, Barbara Graham, and Chetan P. Malhotra, PhD in the preparation of the manuscript for this book. Their help was crucial in bringing this work to fruition. The editors would also like to acknowledge the contributing authors for their forbearance and dedication to this project.

Frank Kreith
D. Yogi Goswami

Editors-in-Chief

Frank Kreith, Department of Mechanical Engineering, University of Colorado, Boulder, Colorado

D. Yogi Goswami, Clean Energy Research Center, University of South Florida, Tampa, Florida

Contributors

Aníbal T. de Almeida
Departamento de Engenharia Electrotécnica—Polo II
Universidade de Coimbra
Coimbra, Portugal

Massoud Amin
Department Electrical and Computer Engineering
University of Minnesota
Minneapolis, Minnesota

Anthony F. Armor
Electric Power Research Institute
Palo Alto, California

Barbara Atkinson
Energy Analysis Department
Lawrence Berkeley National Laboratory
Berkeley, California

Bilal M. Ayyub
Center for Technology and Systems Management
Department of Civil and Environmental Engineering
University of Maryland
College Park, Maryland

F. Barbir
Connecticut Global Fuel Cell Center
University of Connecticut
Storrs, Connecticut

Riccardo Battisti
University of Rome "La Sapienza"
Rome, Italy

Jeff Bell
World Alliance for Decentralized Energy
Edinburgh, Scotland

Dale E. Berg
Sandia National Laboratories
Albuquerque, New Mexico

Peter Biermayer
Lawrence Berkeley National Laboratory
Berkeley, California

Robert C. Brown
Center for Sustainable Environmental Technologies
Iowa State University
Ames, Iowa

Barney L. Capehart
Industrial and Systems Engineering
University of Florida
Gainesville, Florida

David E. Claridge
Mechanical Engineering Department
Texas A&M University
College Station, Texas

Robert Clear
Building Technologies Department
Lawrence Berkeley National Laboratory
Berkeley, California

Charles H. Culp
Energy Systems Laboratory
Texas A&M University
College Station, Texas

Andrea Denver
Energy Analysis Department
Lawrence Berkeley National Laboratory
Berkeley, California

Michael Durstewitz
Institut für Solare Energieversorgungstechnik (ISET)
Division Information and Energy Economy
Kassel, Germany

Ed Fouche
Global Energy Partners, LLC
Raleigh, North Carolina

Marjorie A. Franklin
Franklin Associates
Prairie Village, Kansas

Clark W. Gellings
Electric Power Research Institute
Palo Alto, California

Nitin Goel
Intel Technology India Pvt. Ltd.
Bangalore, India

D. Yogi Goswami
Clean Energy Research Center
University of South Florida
Tampa, Florida

Steve E. Greenberg
Lawrence Berkeley National Laboratory
Berkeley, California

Leonard M. Grillo
Grillo Engineering Company
Hollis, New Hampshire

Gershon Grossman
Technion—Israel
Institute of Technology
Haifa, Israel

Roel Hammerschlag
Institute for Lifecycle
Environmental Assessment
Seattle, Washington

Edwin A. Harvego
Idaho National Laboratory
Idaho Falls, Idaho

Beth Isler
Resource Systems Group, Inc.
Burlington, Vermont

Katherine Johnson
Market Development Group
Montrose, Colorado

Massoud Kayhanian
Center for Environmental and
Water Resources Engineering
Department of Civil and
Environmental Engineering
University of California at Davis
Davis, California

Kevin Kitz
U.S. Geothermal Inc.
Boise, Idaho

Kenneth D. Kok
WSMS Mid-America
Oak Ridge, Tennessee

Moncef Krarti
Joint Center for Energy
Management
Department of Civil, Environmental
and Architectural Engineering
University of Colorado
Boulder, Colorado

Jan F. Kreider
Joint Center for Energy
Management
University of Colorado
Boulder, Colorado

Frank Kreith
Department of Mechanical
Engineering
University of Colorado
Boulder, Colorado

Andy S. Kydes
Energy Information
Administration
United States Department of Energy
Washington, DC

Alex Lekov
Lawrence Berkeley National
Laboratory
Berkeley, California

Debra Lew
National Renewable Energy
Laboratory
Golden, Colorado

Xianguo Li
University of Waterloo
Waterloo, Ontario, Canada

James Lutz
Lawrence Berkeley National
Laboratory
Berkeley, California

M. Mahishi
Department of Mechanical
and Aerospace Engineering
University of Florida
Gainesville, Florida

William Marion
National Renewable
Energy Laboratory
Golden, Colorado

Gene L. Maxwell
Retired Solar Expert
Ft. Collins, Colorado

Robert McConnell
National Center for Photovoltaics
National Renewable
Energy Laboratory
Golden, Colorado

James E. McMahon
Energy Analysis Department
Lawrence Berkeley National
Laboratory
Berkeley, California

Roger Messenger
Department of
Electrical Engineering
Florida Atlantic University
Boca Raton, Florida

Stephen Meyers
Lawrence Berkeley National
Laboratory
Berkeley, California

Anil Misra
Solar Energy Society of India
New Delhi, India

Jeffrey H. Morehouse
Department of Mechanical
Engineering
University of South Carolina
Columbia, South Carolina

Pedro S. Moura
Instituto de Sistemas e Robótica
Departamento de Engenharia
Electrotécnica—Polo II
Universidade de Coimbra
Coimbra, Portugal

Christopher Namovicz
Energy Information
Administration
United States Department
of Energy
Washington, DC

Kelly E. Parmenter
Global Energy Partners, LLC
Lafayette, California

Julie Phillips
Julie Phillips and Associates
Boulder, Colorado

Robert Pratt
Pacific Northwest
National Laboratory
Richland, Washington

Ari Rabl
École des Mines
Centre d'Energetique
Paris, France

Takhir M. Razykov
Physical Technical Institute
Uzbek Academy of Sciences
Tashkent, Uzbekistan

T. Agami Reddy
Department of Civil,
Architectural and
Environmental Engineering
Drexel University
Philadephia, Pennsylvania

David S. Renné
National Renewable
Energy Laboratory
Golden, Colorado

Wesley M. Rohrer, Jr.
(Deceased)

Manuel Romero-Alvarez
Plataforma Solar de
Almeria-CIEMAT
Madrid, Spain

Greg Rosenquist
Lawrence Berkeley National
Laboratory
Berkeley, California

Rosalie Ruegg
(Retired)
National Institute of Standards and
Technology
Gaithersburg, Maryland

Ricardo Rüther
Universidade Federal
de Santa Catarina
Florianopolis, Brazil

Martin Rymes
National Renewable Energy
Laboratory
Golden, Colorado

Koichi Sakuta
Research Center for Photovoltaics
National Institute of Advanced
Industrial Science and
Technology (AIST)
Tsukuba, Japan

Christopher P. Schaber
Institute for Lifecycle
Environmental Assessment
Seattle, Washington

Hans Schweiger
AIGUASOL Engineering
Active Solar Systems Group
Barcelona, Spain

Ramesh K. Shah
Subros Limited
Noida, India

S. A. Sherif
Department of Mechanical
and Aerospace Engineering
University of Florida
Gainesville, Florida

Walter Short
National Renewable
Energy Laboratory
Golden, Colorado

Craig B. Smith
Dockside Consultants, Inc.
Newport Beach, California

Joseph V. Spadaro
École des Mines
Centre d'Energetique
Paris, France

S. S. Srinivasan
Clean Energy Center
University of South Florida
Tampa, Florida

George Tchobanoglous
Department of Civil and
Environmental Engineering
University of California at Davis
Davis, California

Ayodhya N. Tiwari
Centre for Renewable Energy
Systems Technology (CREST)
Department of Electronic
and Electrical Engineering
Loughborough University
Loughborough, Leicestershire,
United Kingdom

Kirtan K. Trivedi
Parsons Energy and
Chemicals Group, Inc.
Pasadena, California

W. Dan Turner
Wisenbaker Engineering
Research Center
Texas A&M University
College Station, Texas

Hari M. Upadhyaya
Centre for Renewable Energy
Systems Technology (CREST)
Department of Electronic
and Electrical Engineering
Loughborough University
Loughborough, Leicestershire,
United Kingdom

Charles O. Velzy
Private Consultant
White Haven, Pennsylvania

T. N. Veziroglu
Clean Energy Research Institute
University of Miami
Coral Gables, Florida

Sanjay Vijayaraghavan
Intel Technology India Pvt. Ltd.
Bangalore, India

Werner Weiss
AEE INTEC
Feldgasse, Austria

Steve Widergren
Pacific Northwest National
Laboratory
Richland, Washington

Stephen Wilcox
National Renewable
Energy Laboratory
Golden, Colorado

Roland Winston
University of California
Merced, California

Eduardo Zarza
Plataforma Solar de
Almeria-CIEMAT
Madrid, Spain

Contents

1

Global Energy System

D. Yogi Goswami
University of South Florida

Frank Kreith
University of Colorado

Global energy consumption in the last half century has increased very rapidly and is expected to continue to grow over the next 50 years. However, we expect to see significant differences between the last 50 years and the next. The past increase was stimulated by relatively "cheap" fossil fuels and increased rates of industrialization in North America, Europe, and Japan; yet while energy consumption in these countries continues to increase, additional factors are making the picture for the next 50 years more complex. These additional complicating factors include the very rapid increase in energy use in China and India (countries representing about a third of the world's population); the expected depletion of oil resources in the not-too-distant future; and the effect of human activities on global climate change. On the positive side, the renewable energy (RE) technologies of wind, biofuels, solar thermal, and photovoltaics (PV) are finally showing maturity and the ultimate promise of cost competitiveness.

Statistics from the International Energy Agency (IEA) World Energy Outlook 2004 show that the total primary energy demand in the world increased from 5536 GTOE in 1971 to 10,345 GTOE in 2002, representing an average annual increase of 2% (see Table 1.1 and Figure 1.1).[1]

Of the total primary energy demand in 2002, fossil fuels accounted for about 80%, with oil, coal, and natural gas accounting for 35.5, 23, and 21.2%, respectively. Biomass accounted for 11% of all the primary energy in the world, with almost all of it being traditional biomass for cooking and heating in developing countries; biomass is used very inefficiently in these applications.

[1]The energy data for this chapter came from many sources, which use different units of energy, making it difficult to compare the numbers. The conversion factors are given here for a quick reference: MTOE = Mega tons of oil equivalent; 1 MTOE = 4.1868×10^4 TJ (Terra Joules) = 3.968×10^{13} Btu., GTOE = Giga tons of oil equivalent; 1 GTOE = 1000 MTOE; Quadrillion Btu, also known as Quad: 10^{15} British Thermal Units or Btu; 1 Btu = 1055 J, 1 TWh = 10^9 kWh, 1 kWh = 3.6×10^6 J.

TABLE 1.1 World Total Energy Demand (MTOE)

Energy Source/Type	1971	2002	Annual Percentage of Change 1971–2002
Coal	1407	2389	1.7
Oil	2413	3676	1.4
Gas	892	2190	2.9
Nuclear	29	892	11.6
Hydro	104	224	2.5
Biomass and Waste	687	1119	1.6
Other Renewables	4	55	8.8
Total	5536	10,345	2.0

Source: From IEA, *World Energy Outlook*, International Energy Agency, Paris, 2004. With permission.

The last 10 years of data for energy consumption from British Petroleum Corp. (BP) also shows that the average increase per year was 2%. However, it is important to note (from Table 1.2) that the average worldwide growth from 2001 to 2004 was 3.7% with the increase from 2003 to 2004 being 4.3%. The rate of growth is rising mainly due to the very rapid growth in Pacific Asia, which recorded an average increase from 2001 to 2004 of 8.6%.

More specifically, China increased its primary energy consumption by 15% from 2003 to 2004. Unconfirmed data show similar increases continuing in China, followed by increases in India. Fueled by high increases in China and India, worldwide energy consumption may continue to increase at rates between 3 and 5% for at least a few more years. However, such high rates of increase cannot continue for long. Various sources estimate that the worldwide average annual increase in energy consumption will be 1.6%–2.5% (IEA 2004; IAEA 2005). Based on a 2% increase per year (average of the estimates from other sources), the primary energy demand of 10,345 GTOE in 2002 would double by 2037 and triple by 2057. With such high energy demand expected 50 years from now, it is important to look at all of the available strategies to fulfill the future demand, especially for electricity and transportation.

Although not a technical issue in the conventional sense, no matter what types of engineering scenarios are proposed to meet the rising energy demands world population, as long as exponential growth in world population continues, the attendant problems of energy and food consumption, as well as environmental degradation, may have no long-term solution (Bartlett 2004). Under current demographic trends, the United Nations forecasts a rise in the global population to around 9 billion in the year 2050. This increase in 2.5 billion people will occur mostly in developing countries with aspirations for a higher standard of living. Thus, population growth should be considered as a part of the overall supply and demand picture to assure the success of future global energy and pollution strategy.

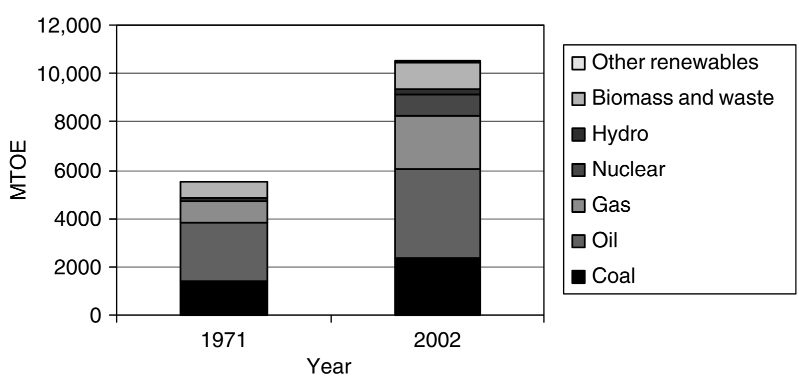


FIGURE 1.1 **(See color insert following page 774.)** World primary energy demand (MTOE). (Data from IEA, *World Energy Outlook*, International Energy Agency, IEA, Paris, 2004. With permission.)

TABLE 1.2 Primary Energy Consumption (MTOE)

Region	2001	2002	2003	2004	Average Increase/Year (%)	2004 Change Over 2003 (%)
North America including U.S.A.	2681.5	2721.1	2741.3	2784.4	1.3	1.6
U.S.A.	2256.3	2289.1	2298.7	2331.6	1.1	1.4
South and Central America	452.0	454.4	460.2	483.1	2.2	5.0
Europe and Euro–Asia	2855.5	2851.5	2908.0	2964.0	1.3	1.9
Middle East	413.2	438.7	454.2	481.9	5.3	6.1
Africa	280.0	287.2	300.1	312.1	3.7	4.0
Asia Pacific	2497.0	2734.9	2937.0	3198.8	8.6	8.9
World	9179.3	9487.9	9800.8	10,224.4	3.7	4.3

This data does not include traditional biomass, which was 2229 MTOE in 2002 according to IEA data.
Source: From British Petroleum Corporation, *BP Statistical Review of World Energy, 2006*, British Petroleum, London, 2006, http://www.bp.com/statisticalreview/.

1.1 Major Sectors of Primary Energy Use

The major sectors using primary energy sources include electrical power, transportation, heating, industrial, and others, such as cooking. The IEA data shows that the electricity demand almost tripled from 1971 to 2002. This is not unexpected because electricity is a very convenient form of energy to transport and use. Although primary energy use in all sectors has increased, their relative shares except for transportation and electricity have decreased (Figure 1.2). Figure 1.2 shows that the relative share of primary energy for electricity production in the world increased from about 20% in 1971 to about 30% in 2002. This is because electricity is becoming the preferred form of energy for all applications.

Figure 1.3 shows that coal is presently the largest source of electricity in the world. Consequently, the power sector accounted for 40% of all CO_2 emissions in 2002. Emissions could be reduced by increased

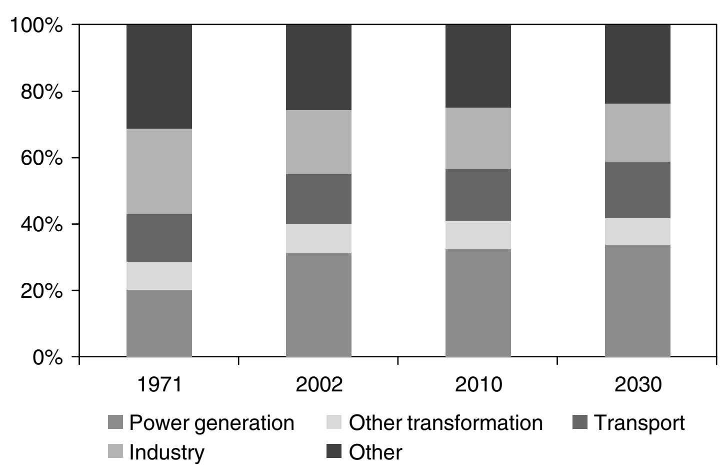


FIGURE 1.2 **(See color insert following page 774.)** Sectoral shares in world primary energy demand. (Data and forecast from IEA, *World Energy Outlook*, IEA, Paris, 2004. With permission.)

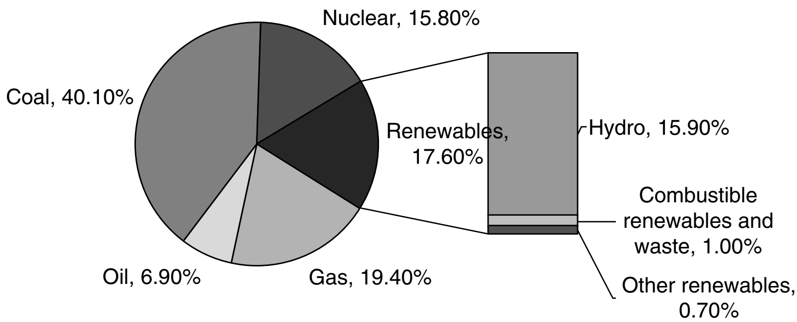


FIGURE 1.3 World electricity production by fuel in 2003. (Data from IEA, *Renewables Information*, IEA, Paris, 2005. With permission.)

use of RE sources. All RE sources combined accounted for only 17.6% share of electricity production in the world, with hydroelectric power providing almost 90% of it. All other RE sources provided only 1.7% of electricity in the world. However, the RE technologies of wind power and solar energy have vastly improved in the last two decades and are becoming more cost effective. As these technologies mature and become even more cost competitive in the future they may be in a position to replace a major fraction of fossil fuels for electricity generation. Therefore, substituting fossil fuels with RE for electricity generation must be an important part of any strategy of reducing CO_2 emissions into the atmosphere and combating global climate change.

1.2 Electrical Capacity Additions to 2030

Figure 1.4 shows the additional electrical capacity forecast by IEA for different regions in the world. The overall increase in the electrical capacity is in general agreement with the estimates from IAEA (2005)

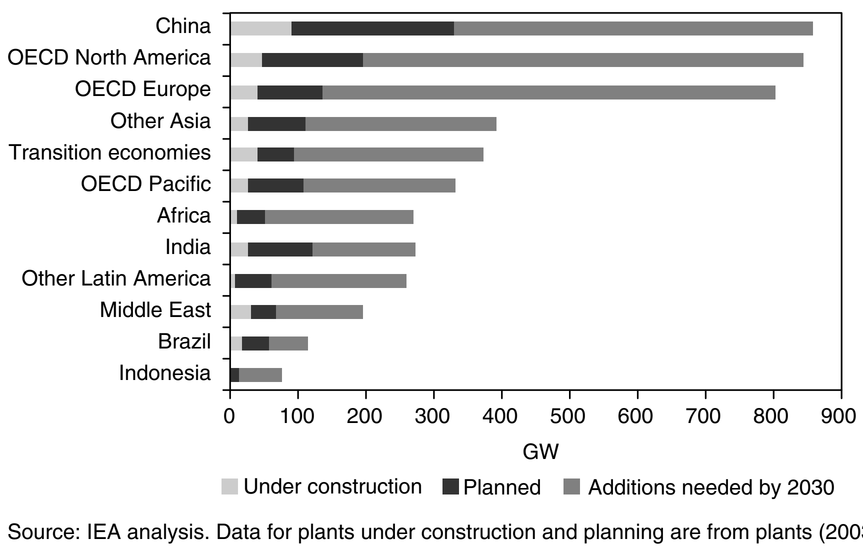


FIGURE 1.4 Electrical capacity requirements by region. (From IEA, *World Energy Outlook*, IEA, Paris, 2004. With permission.)

which project an average annual growth of about 2%–2.5% up to 2030. It is clear that of all countries, China will add the largest capacity with its projected electrical needs accounting for about 30% of the world energy forecast. China and India combined will add about 40% of all the new capacity of the rest of the world. Therefore, what happens in these two countries will have important consequences on the worldwide energy and environmental situation. If coal provides as much as 70% of China's electricity in 2030, as forecasted by IEA (IEA 2004), it will certainly increase worldwide CO_2 emissions, and further increase global warming.

1.3 Transportation

Transportation is another sector that has increased its relative share of primary energy. This sector has serious concerns as it is a significant source of CO_2 emissions and other airborne pollutants, and it is almost totally based on oil as its energy source (Figure 1.5; Kreith, West, and Isler 2002). In 2002, the transportation sector accounted for 21% of all CO_2 emissions worldwide. An important aspect of future changes in transportation depends on what happens to the available oil resources, production and prices. At present, 95% of all energy for transportation comes from oil.

As explained later in this chapter, irrespective of the actual amount of oil remaining in the ground, oil production will peak soon. Therefore, the need for careful planning for an orderly transition away from oil as the primary transportation fuel is urgent. An obvious replacement for oil would be biofuels such as ethanol, methanol, biodiesel, and biogases. Some believe that hydrogen is another alternative, because if it could be produced economically from RE sources or nuclear energy, it could provide a clean transportation alternative for the future. Some have claimed hydrogen to be a "wonder fuel" and have proposed a "hydrogen-based economy" to replace the present carbon-based economy (Veziroglu and Barbir 1992). However, others (Shinnar 2003; Kreith and West 2004; Mazza and Hammerschlag 2005) dispute this claim based on the lack of infrastructure, problems with storage and safety, and the lower efficiency of hydrogen vehicles as compared to plug-in hybrid or fully electric vehicles (West and Kreith 2006). Already hybrid-electric automobiles are becoming popular around the world as petroleum becomes more expensive.

The environmental benefits of renewable biofuels could be increased by using plug-in hybrid electric vehicles (PHEVs). These cars and trucks combine internal combustion engines with electric motors to

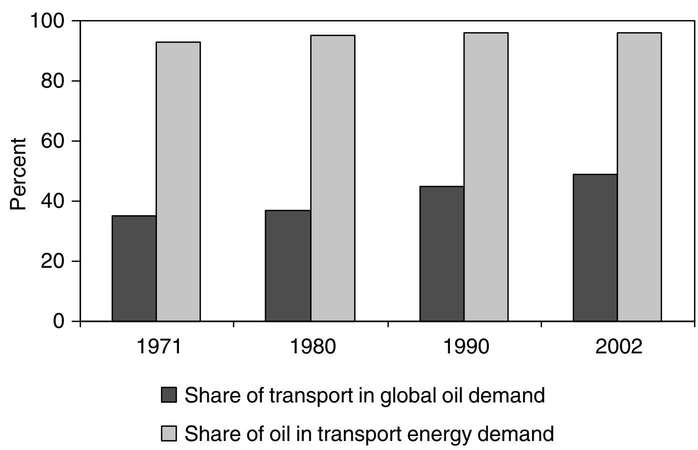


FIGURE 1.5 Share of transport in global oil demand and share of oil in transport energy demand. (Data and forecast from IEA, *World Energy Outlook*, IEA, Paris, 2004. With permission.)

maximize fuel efficiency. PHEVs have more battery capacity that can be recharged by plugging it into a regular electric outlet. Then these vehicles can run on electricity alone for relatively short trips. The electric-only trip length is denoted by a number, e.g., PHEV 20 can run on battery charge for 20 miles. When the battery charge is used up, the engine begins to power the vehicle. The hybrid combination reduces gasoline consumption appreciably. Whereas the conventional vehicle fleet has a fuel economy of about 22 mpg, hybrids such as the Toyota Prius can attain about 50 mpg. PHEV 20s have been shown to attain as much as 100 mpg. Gasoline use can be decreased even further if the combustion engine runs on biofuel blends, such as E85, a mixture of 15% gasoline and 85% ethanol (Kreith 2006; West and Kreith 2006).

Plug-in hybrid electric technology is already available and could be realized immediately without further R&D. Furthermore, a large portion of the electric generation infrastructure, particularly in developed countries, is needed only at the time of peak demand (60% in the United States), and the rest is available at other times. Hence, if batteries of PHEVs were charged during off-peak hours, no new generation capacity would be required. Moreover, this approach would levelize the electric load and reduce the average cost of electricity, according to a study by the Electric Power Research Institute (EPRI) (Sanna 2005).

Given the potential of PHEVs, EPRI (EPRI 2004) conducted a large-scale analysis of the cost, battery requirements, economic competitiveness of plug-in vehicles today and in the future. As shown by West and Kreith, the net present value of lifecycle costs over 10 years for PHEVs with a 20-mile electric-only range (PHEV20) is less than that of a similar conventional vehicle (West and Kreith 2006). Furthermore, currently available nickel metal hydride (NiMH) batteries are already able to meet required cost and performance specifications. More advanced batteries, such as lithium-ion (Li-ion) batteries, may improve the economics of PHEVs even further in the future.

1.4 World Energy Resources

With a view to meet the future demand of primary energy in 2050 and beyond, it is important to understand the available reserves of conventional energy resources including fossil fuels and uranium, and the limitations posed on them due to environmental considerations.

1.4.1 Conventional Oil

There is a considerable debate and disagreement on the estimates of "ultimate recoverable oil reserves." However, there seems to be a good agreement on the amount of "proven oil reserves" in the world. According to British Petroleum (2006), total identified or proven world oil reserves at the end of 2005 were 1200.7 billion barrels. This estimate is close to the reserves of 1266 billion barrels from other sources listed by IEA (IEA 2004). The differences among them lie in their accounting for the Canadian tar sands. Considering the present production rate of over 80 million barrels per day, these reserves will last for about 41 years if there is no increase in production. Of course, there may be additional reserves that may be discovered in the future. A recent analysis by the U.S. Energy Information Agency (2006) estimates the ultimately recoverable world oil reserves (including resources not yet discovered) at between 2.2×10^{12} barrels (bbl) and 3.9×10^{12} bbl with a mean estimate of the USGS at 3×10^{12} bbl.

Ever since petroleum geologist M. King Hubbert correctly predicted in l956 that U.S. oil production would reach a peak in l973 and then decline (Hubbert 1974), scientists and engineers have known that worldwide oil production would follow a similar trend. Today, the only question is when the world peak will occur. Bartlett (2002) has developed a predictive model based on a Gaussian curve similar in shape to the data used by Hubbert as shown in Figure 1.6. The predictive peak in world oil production depends on the assumed total amount of recoverable reserves.

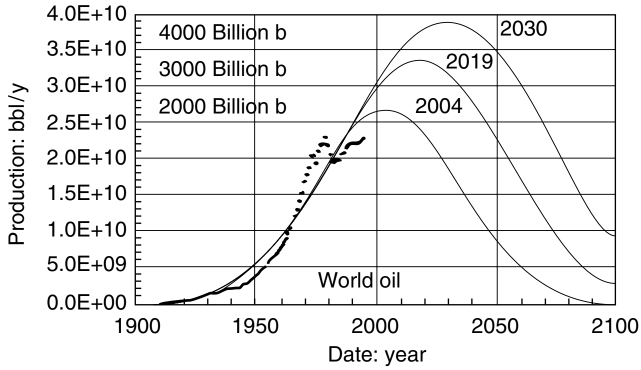


FIGURE 1.6 World oil production vs. time for various amounts of ultimate recoverable resource. (From Bartlett, A. A., *Mathematical Geology*, 32, 2002. With permission.)

If BP's estimations of oil reserves are correct, we may have already peaked in world oil production. If, however, estimates of the ultimate reserves (discovered and undiscovered) are used, we may expect the oil production to increase a little longer before it peaks. However, changing the total available reserves from 3×10^{12} bbl to 4×10^{12} bbl increases the predicted time of peak production by merely 11 years, from 2019 to 2030. There is no question that after the world peak is reached and oil production begins to drop, either alternative fuels will have to be supplied to make up the difference between demand and supply, or the cost of fuel will increase precipitously and create an unprecedented social and economic crisis for our entire transportation system.

The present trend of yearly increases in oil consumption, especially in China and India, shortens the window of opportunity for a managed transition to alternative fuels even further. Hence, irrespective of the actual amount of oil remaining in the ground, peak production will occur soon. Therefore, the need for starting to supplement oil as the primary transportation fuel is urgent because an orderly transition to develop petroleum substitutes will take time and careful planning.

1.4.2 Natural Gas

According to BP (2006), the total proven world natural gas reserves at the end of 2004 were 179.5 trillion m^3. Considering the production rate of gas in 2004, with no increase in production thereafter, these reserves would last for 67 years. However, production of natural gas has been rising at an average rate of 2.5% over the past four years. If production continues to rise because of additional use of CNG for transportation and increased power production from natural gas, the reserves would last for fewer years. Of course, there could be additional new discoveries. However, even with additional discoveries, it is reasonable to expect that all the available natural gas resources may last from about 50 to 80 years, with a peak in production occurring much earlier.

1.4.3 Coal

Coal is the largest fossil resource available to us and the most problematic from an environmental standpoint. From all indications, coal use will continue to grow for power production around the world because of expected increases in China, India, Australia, and other countries. From an environmental point of view, this would be unsustainable unless advanced "clean coal technology" (CCT) with carbon sequestration is deployed.

Clean coal technology is based on an integrated gasification combined-cycle (IGCC) that converts coal to a gas that is used in a turbine to provide electricity with CO_2 and pollutant removal before the fuel is burned (Hawkins, Lashof, and Williams 2006). According to R.C. Kelly, President and Chief Executive Officer of Minneapolis-based Xcel Energy, the company is about to build such a plant in Colorado, U.S.A. The plant will capture CO_2 and inject it underground, possibly in depleted oil fields. According to Kelly, an IGCC plant can cost 20% more to build than a conventional coal plant, but is more efficient to operate (Associated Press 2006). According to an Australian study (Sadler 2004), no carbon capture and storage system is yet operating on a commercial scale, but may become an attractive technology to achieve atmospheric CO_2 stabilization.

According to BP, the proven recoverable world coal resources were estimated to be 909 billion tons at the end of 2004 with a reserve to production ratio (R/P) of 164 years. The BP data also shows that coal use increased at an average rate of 6% from 2002 to 2005, the largest increase of all fossil resources. Because China and India are continuing to build new coal power plants, it is reasonable to assume that coal use will continue to increase for at least some years in the future. Therefore, the R/P ratio will decrease from the present value of 164 years. This R/P ratio will decrease even more rapidly when clean coal technologies such as coal gasification and liquification are utilized instead of direct combustion.

1.4.4 Summary of Fossil Fuel Reserves

Even though there are widely differing views and estimates of the ultimately recoverable resources of fossil fuels, it is fair to say that they may last for around 50–150 years with a peak in production occurring much earlier. However, a big concern is the climatic threat of additional carbon that will be released into the atmosphere. According to the estimates from the IEA, if the present shares of fossil fuels are maintained up to 2030 without any carbon sequestration, a cumulative amount of approximately 1000 gigatons of carbon will be released into the atmosphere (based on Figure 1.7). This is especially troublesome in view of the fact that the present total cumulative emissions of about 300 gigatons of carbon have already raised serious concerns about global climate change.

1.4.5 Nuclear Resources

Increased use of nuclear power presents the possibility of additional carbon-free energy use and its consequent benefit for the environment. However, there are significant concerns about nuclear waste and other environmental impacts, the security of the fuel and the waste, and the possibility of their diversion for weapon production.

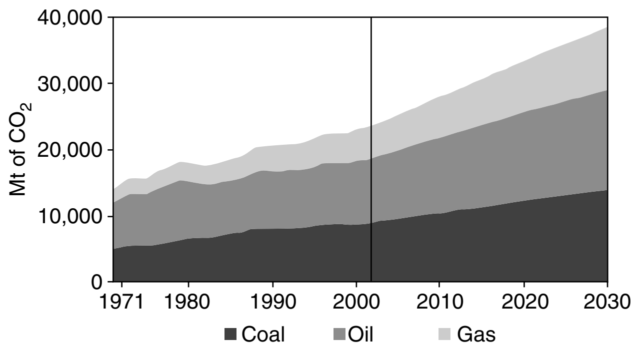


FIGURE 1.7 World energy-related CO_2 emissions by fuel. (Data and forecast from IEA, *World Energy Outlook*, IEA, Paris, 2004. With permission.)

According to the IAEA (2005) nuclear fission provided 16% of the electricity in the world in 2004, with a worldwide capacity of 368 GW. An additional 20 GW of nuclear power capacity was under construction during the same year. The IAEA also estimates that the worldwide nuclear power capacity will increase at an average rate of 0.5%–2.2% until 2030 (IAEA 2005). At present, uranium is used as the fissile material for nuclear power production. Thorium could also be used for nuclear fission; however, to date nobody has developed a commercial nuclear power plant based on thorium. Terrestrial deposits of both uranium and thorium are limited and concentrated in a few countries of the world. Estimates from the International Atomic Energy Agency (IAEA) and other sources show that the recoverable assured uranium reserves in the world are about 2.3 million tonnes to as much as 3.2 million tons (UNDP 2004). If additional estimated resources (not yet discovered) are also included, then the total resources become 5.1 million tons (UNDP 2004). Additionally, there are nonconventional uranium resources, such as sea water which contains about 3 parts per billion uranium and some phosphate deposits (more than half of them in Morocco) which contain about 100 parts per million uranium. These resources are potentially huge; however, their cost effective recovery is not certain.

Generating 1 TWh of electricity from nuclear fission requires approximately 22 tonnes of uranium (UNDP 2004). Based on the 2004 world capacity of 368 GW and an average annual growth rate of 2%, the present known uranium reserves of 2.3–3.2 million tonnes will last until 2030–2037. If all of the estimated (discovered and undiscovered) reserves of 5.1 million tonnes are considered, they will be used up by 2050. This estimate does not consider regeneration of spent fuel. At present, nuclear fuel regeneration is not allowed in the United States. However, that law could be changed in the future. Development of breeder reactors could increase the time period much further. Nuclear fusion could potentially provide a virtually inexhaustible energy supply; however, it is not expected to be commercially available in the foreseeable future.

1.4.6 Present Status and Potential of Renewable Energy (RE)

According to the data in Table 1.3, 13.3% of the world's total primary energy supply came from RE in 2003. However, almost 80% of the RE supply was from biomass (Figure 1.8), and in developing countries it is mostly converted by traditional open combustion, which is very inefficient. Because of its inefficient use, biomass resources presently supply only about 20% of what they could if converted by more efficient, already available technologies. As it stands, biomass provides only 11% of the world total primary energy, which is much less than its real potential. The total technologically sustainable biomass energy potential for the world is 3–4 TW_e (UNDP 2004), which is more than the entire present global electrical generating capacity of about 3 TW_e.

In 2003, shares of biomass and hydropower in the total primary energy mix of the world were about 11% and 2%, respectively. All of the other renewables, including solar thermal, solar PV, wind, geothermal, and ocean combined, provided only about 0.5% of the total primary energy. During the same year, biomass combined with hydroelectric resources provided more than 50% of all the primary

TABLE 1.3 2003 Fuel Shares in World Total Primary Energy Supply

Source	Share (%)
Oil	34.4
Natural Gas	21.2
Coal	24.4
Nuclear	6.5
Renewables	13.3

Source: Data from IEA, *World Energy Outlook*, IEA, Paris, 2004. With permission.

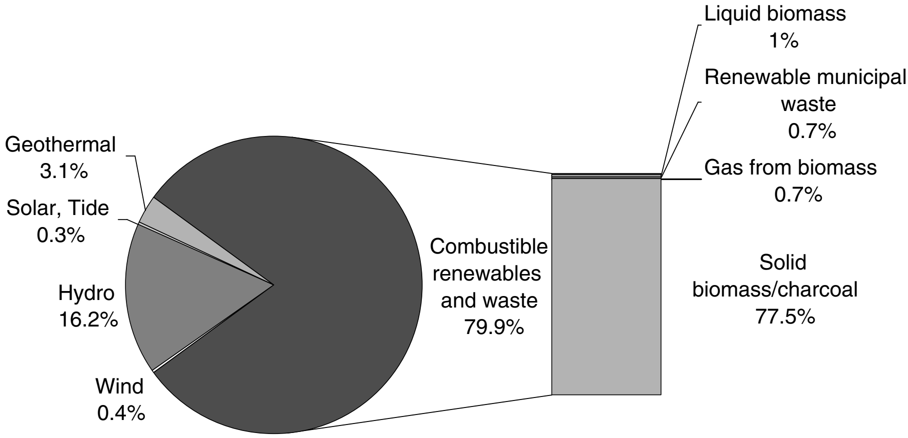


FIGURE 1.8 2003 resource shares in world renewable energy supply. (Data from IEA, *World Energy Outlook*, IEA, Paris, 2004. With permission.)

TABLE 1.4 Share of Renewable Energy (RE) in 2003 Total Primary Energy Supply (TPES) on a Regional Basis

Region	TPES	Renewables	Share of Renewables in TPES (%)
	MTOE	MTOE	
Africa	558.9	279.9	50.1
Latin America	463.9	135.5	29.2
Asia	1224.4	400	32.7
India	553.4	218	39.4
China	1425.9	243.4	17.1
Non-OECD[a] Europe	103.5	9.7	9.4
Former U.S.S.R.	961.7	27.5	2.9
Middle East	445.7	3.2	0.7
OECD	5394.7	304.7	5.6
U.S.A.	2280.8	95.3	4.2
World	10,578.7	1403.7	13.3

[a] OECD, Organization for Economic Cooperation and Development.
Source: From IEA, *Renewables Information*, IEA, Paris, 2005. With permission.

TABLE 1.5 Electricity from RE in 2002

Energy Source	2002	
	TWh	(%)
Hydropower	2610	89
Biomass	207	7
Wind	52	2
Geothermal	57	2
Solar	1	0
Tide/Wave	1	0
Total	2927	100

Source: Data from IEA, *World Energy Outlook*, IEA, Paris, 2004. With permission.

energy in Africa, 29.2% in Latin America, and 32.7% in Asia (Table 1.4; IEA 2005). However, biomass is used very inefficiently for cooking in these countries. Such use has also resulted in significant health problems, especially for women.

The total share of all renewables for electricity production in 2002 was about 17%, a vast majority (89%) of it being from hydroelectric power (Table 1.5).

1.4.7 Wind Power

Wind-energy technology has progressed significantly over the last two decades. The technology has been vastly improved and capital costs have come down to as low as $1000 per kW. At this level of capital costs, wind power is already economical at locations with fairly good wind resources. Therefore, the average annual growth in worldwide wind energy capacity from 2000 to 2003 was over 30% (Figure 1.9) and it continued to grow at that rate in 2004 and 2005. The average growth in the United States over the same period was 37.7%. The total worldwide installed wind power capacity that was 39 GW in 2003 (Figure 1.9), reached a level of 59 GW in 2005 (WWEA 2006). The total theoretical potential for onshore wind power for the world is around 55 TW with a practical potential of at least 2 TW (UNDP 2004), which is about two-thirds of the entire present worldwide generating capacity. The offshore wind energy potential is even larger.

1.4.8 Solar Energy

The amount of sunlight striking the Earth's atmosphere continuously is 1.75×10^5 TW. Considering a 60% transmittance through the atmospheric cloud cover, 1.05×10^5 TW reaches the Earth's surface continuously. If the irradiance on only 1% of the Earth's surface could be converted into electric energy with a 10% efficiency, it would provide a resource base of 105 TW, whereas the total global energy needs for 2050 are projected to be about 25–30 TW. The present state of solar energy technologies is such that solar-cell efficiencies have reached over 20% and solar thermal systems provide efficiencies of 40%–60%. With the present rate of technological development, these solar technologies will continue improving, thus bringing the costs down, especially with the economies of scale.

Solar PV panels have come down in cost from about $30/W to about $3/W in the last three decades. At $3/W panel cost, the overall system cost is around $6/W, which is still too high to compete with other resources for grid electricity. However, there are many off-grid applications where solar PV is already cost-effective. With net metering and governmental incentives, such as feed-in laws and other policies,

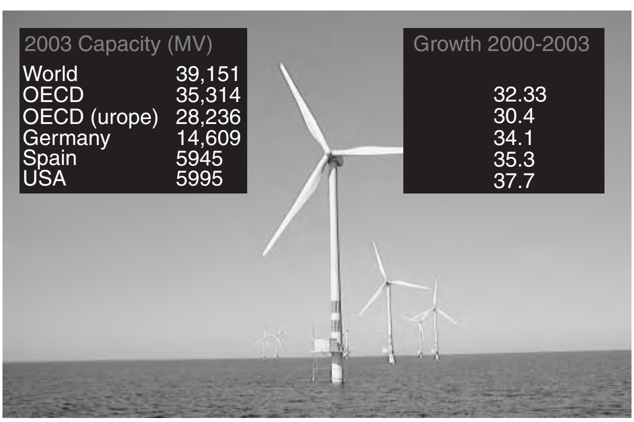


FIGURE 1.9 World wind-energy installed capacity and growth rates. (Data from IEA, *World Energy Outlook*, IEA, Paris, 2004. With permission.)

TABLE 1.6 Growth in Photovoltaics (PV) Demand (2000–2003)

Region	Percent Increase 2000–2003 (%)
OECD	32
OECD (Europe)	41.1
Germany	51.1

even grid-connected applications such as building-integrated PV (BIPV) have become cost-effective. As a result, the worldwide growth in PV production has averaged over 30% per year from 2000 to 2003, with Germany showing the maximum growth of over 51% (Table 1.6; Figure 1.10).

Solar thermal power using concentrating solar collectors was the first solar technology that demonstrated its grid power potential. A 354 MW_e solar thermal power plant has been operating continuously in California since 1988. Progress in solar thermal power stalled after that time because of poor policy and lack of R&D. However, the last five years have seen a resurgence of interest in this area and a number of solar thermal power plants around the world are under construction. The cost of power from these plants (which is so far in the range of 12–16 U.S. cents/kWh) has the potential to go down to 5 U.S. cents/kWh with scale-up and creation of a mass market. An advantage of solar thermal power is that thermal energy can be stored efficiently and fuels such as natural gas or biogas may be used as backup to ensure continuous operation. If this technology is combined with power plants operating on fossil fuels, it has the potential to extend the time frame of the existing fossil fuels.

Low temperature solar thermal systems and applications have been well developed for quite some time. They are being actively installed wherever policies favor their deployment. Figure 1.11 gives an idea of the rate of growth of solar thermal systems in the world (ESTIF 2000). Just in 2003, over 10 MW_{th} solar collectors were deployed around the world, a vast majority of those being in China (Figure 1.12).

1.4.9 Biomass

Although theoretically harvestable biomass energy potential is on the order of 90 TW, the technical potential on a sustainable basis is on the order of 8–13 TW or 270–450 exajoules/year (UNDP 2004). This potential is 3–4 times the present electrical generation capacity of the world. It is estimated that by 2025, even municipal solid waste (MSW) alone could generate up to 6 exajoules/year (UNDP 2004).

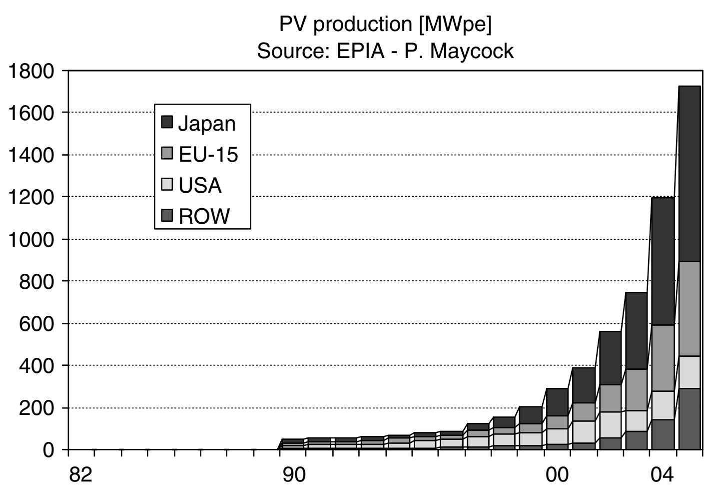


FIGURE 1.10 World solar PV production, 1990–2005 (MWp). (From Maycock, P., PV News Annual Review of the PV Market, 2006, http://www.epia.org. With permission.)

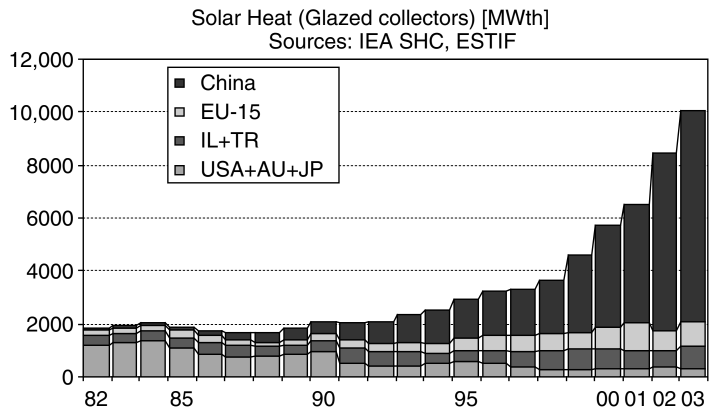


FIGURE 1.11 Deployment of solar heat (glazed) collectors (MW_{th}).

The biggest advantage of biomass as an energy resource is its relatively straightforward transformation into transportation fuels. Biofuels have the potential to replace as much as 75% of the petroleum fuels in use for transportation in the U.S.A. today (Worldwatch Institute 2006). This is especially important in view of the declining oil supplies worldwide. Biofuels will not require additional infrastructure development. Therefore, development of biofuels is being viewed very favorably by governments around the world. Biofuels, along with other transportation options such as electric vehicles and hydrogen, will help diversify the fuel base for future transportation. Figure 1.13 shows that global ethanol production more than doubled between 2000 and 2005. Biodiesel production grew almost fourfold, although it started from a much smaller base. In 2005, the world ethanol production had reached about 36 billion liters per year, whereas biodiesel production topped 3.5 billion liters during the same year. Table 1.7 shows the top five countries producing these fuels.

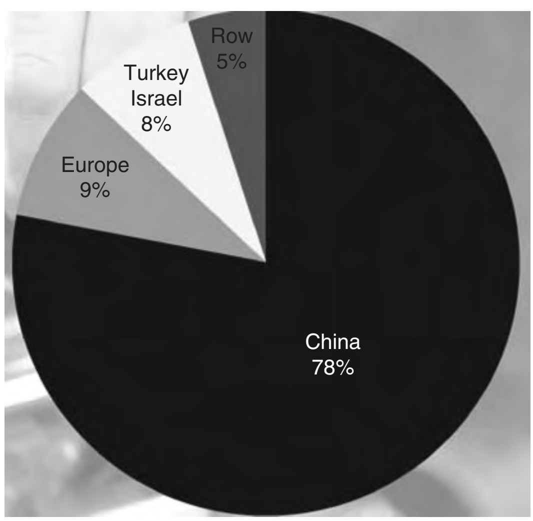


FIGURE 1.12 Worldwide distribution of solar thermal collector markets. (From ESTIF, *Solar Thermal Markets in Europe—Trends and Statistics for 2004*, European Solar Thermal Industry Federation, Brussels, Belgium, 2005. With permission.)

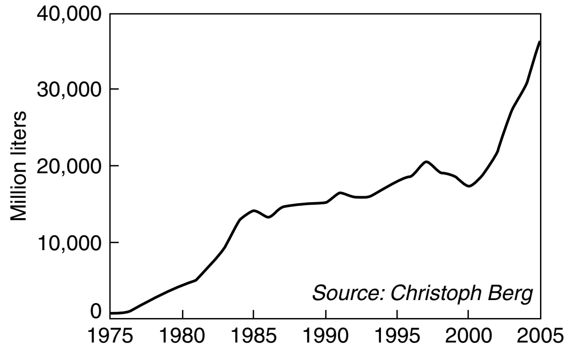


FIGURE 1.13 World fuel ethanol production, 1975–2005. (From Worldwatch, Biofuels for Transportation—Global Potential and Implications for Sustainable and Energy in the 21st Century, Report prepared for the German Federal Ministry for Food, Agriculture and Consumer Protection, Worldwatch Institute, Washington, DC, 2006. With permission.)

The present cost of ethanol production ranges from about 25 Euro cents to about 1 Euro per gasoline equivalent liter, as compared to the wholesale price of gasoline that is between 40 and 60 Euro cents per liter (Figure 1.14). Biodiesel costs, on the other hand, range between 20 Euro cents to 65 Euro cents per liter of diesel equivalent (Figure 1.15). Figure 1.16 shows the feedstocks used for these biofuels. An important consideration for biofuels is that the fuel not be produced at the expense of food while there are people going hungry in the world. This would not be of concern if biofuels were produced from municipal solid waste (MSW).

According to the Worldwatch report, a city of 1 million people produces about 1800 tonnes of MSW and 1300 tonnes of organic waste every day that, using the present-day technology, could produce enough fuel to meet the needs of 58,000 persons in the United States, 360,000 in France, and nearly 2.6 million in China at current rates of per capita fuel use (Worldwatch Institute 2006).

1.4.10 Summary of RE Resources

By definition, the term *reserves* does not apply to renewable resources. Therefore, we need to look at the annual potential of each resource. Table 1.8 summarizes the resource potential and the present costs and the potential future costs for the most important renewable resources.

TABLE 1.7 Top Five Ethanol and Biodiesel Producers in the World

	Ethanol Production (million liters/year)
Brazil	16,500
U.S.A.	16,230
China	2000
European Union	900
India	300
	Diesel Production (million liters/year)
Germany	1920
France	511
U.S.A.	290
Italy	227
Austria	83

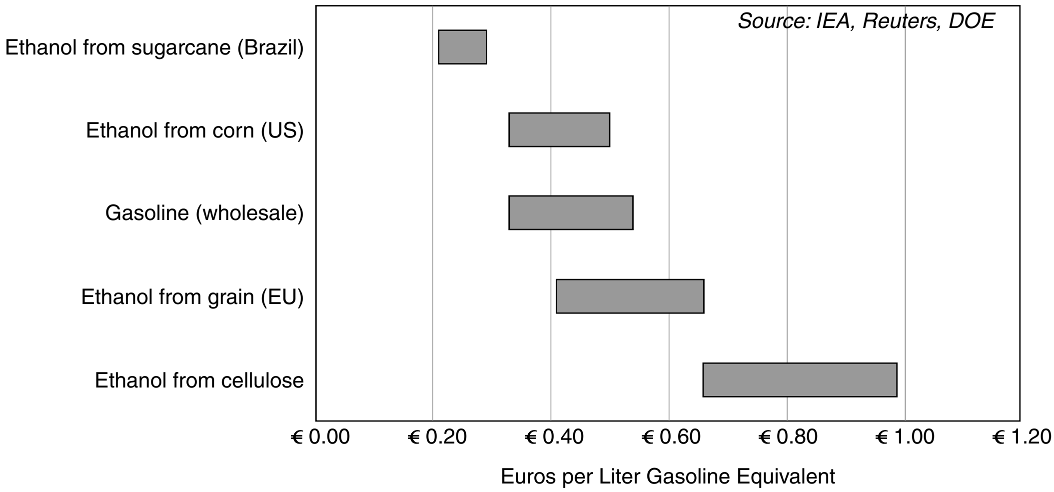


FIGURE 1.14 Cost ranges for ethanol and gasoline production, 2006. (From Worldwatch, Biofuels for Transportation—Global Potential and Implications for Sustainable Energy in the 21st Century, Report prepared for the German Federal Ministry for Food, Agriculture and Consumer Protection, Worldwatch Institute, Washington, DC, 2006. With permission.)

As in the case of other new technologies, it is expected that cost competitiveness of the RE technologies will be achieved with R&D, scale-up, commercial experience and mass production. The experience curves in Figure 1.17 show industry-wide cost reductions in the range of 10%–20% for each cumulative doubling of production for wind power, PV, ethanol, and gas turbines (UNDP 2004). Similar declines can be expected in solar thermal power and other renewable technologies. As seen from Figure 1.17, wind energy technologies have already achieved market maturity, and PV technologies are well on their way. Even though concentrating solar thermal power (CSP) is not shown in this figure, a GEF report estimates

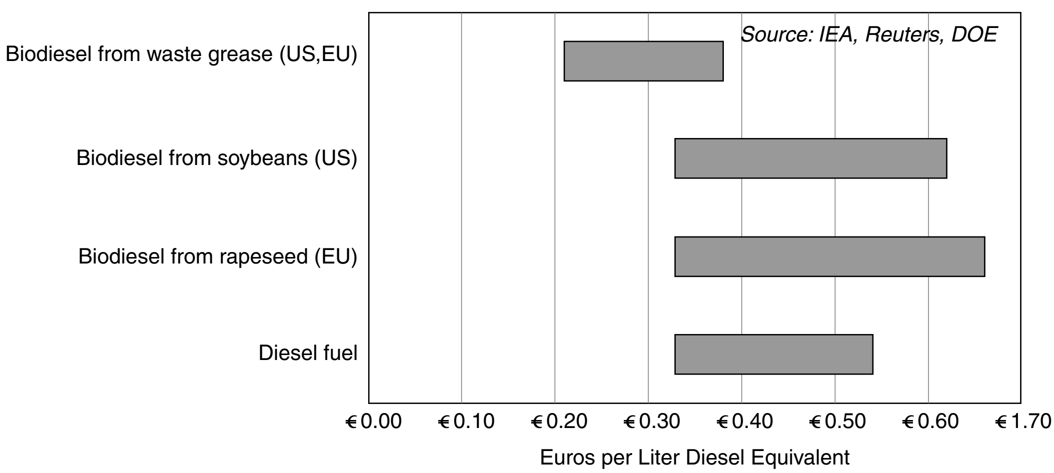


FIGURE 1.15 Cost ranges for biodiesel and diesel production, 2006. (From Worldwatch, Biofuels for Transportation—Global Potential and Implications for Sustainable and Energy in the 21st Century, Report prepared for the German Federal Ministry for Food, Agriculture and Consumer Protection, Worldwatch Institute, Washington, DC, 2006. With permission.)

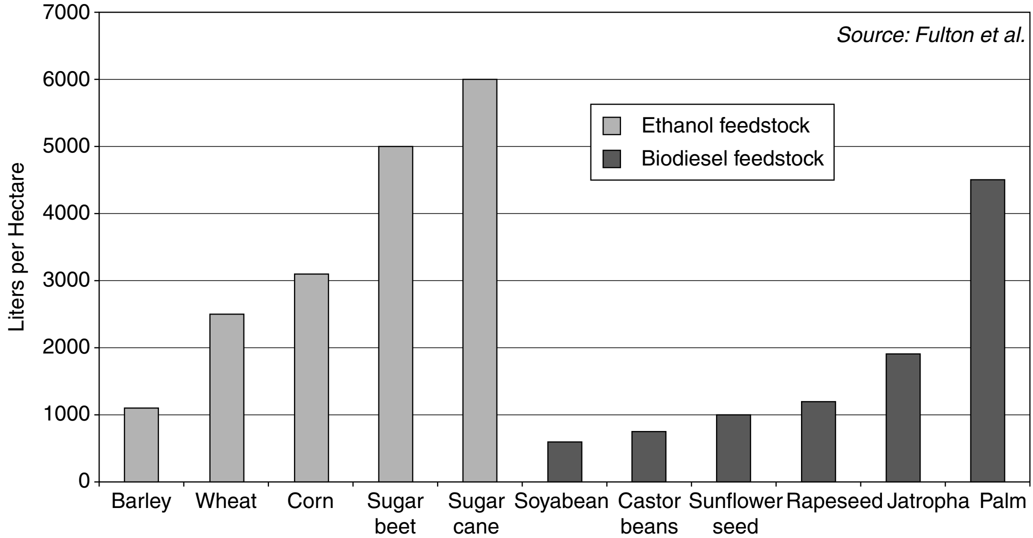


FIGURE 1.16 Biofuel yields of selected ethanol and biodiesel feedstocks. (From Worldwatch, Biofuels for Transportation—Global Potential and Implications for Sustainable and Energy in the 21st Century, Report prepared for the German Federal Ministry for Food, Agriculture and Consumer Protection, Worldwatch Institute, Washington, DC, 2006. With permission.)

that CSP will achieve the cost target of about $0.05/kWh by the time it has an installed capacity of about 40 GW (GEF 2005). As a reference point, wind power achieved that capacity milestone in 2003.

1.5 Role of Energy Conservation

Energy conservation can and must play an important role in future energy strategy, because it can ameliorate adverse impacts on the environment rapidly and economically. Figure 1.18 and Figure 1.19 give an idea of the potential of energy efficiency improvements. Figure 1.18 shows that per capita energy consumption varies by as much as a factor of three between the U.S.A. and some European countries with almost the same level of human development index (HDI). Even taking just the OECD European countries combined, the per capita energy consumption in the U.S.A. is twice as much. It is fair to assume that the per capita energy of the United States could be reduced to the level of OECD Europe of 4.2 kW by a combination of energy efficiency improvements and changes in the transportation infrastructure. This is significant because the U.S.A. uses about 25% of the energy of the whole world. The present per capita energy consumption in the U.S.A. is 284 GJ, which is equivalent to about 9 kW per person, whereas the average for the whole world is 2 kW. The Board of Swiss Federal Institutes of Technology has developed a vision of a 2-kW-per-capita society by the middle of the century (UNDP 2004). The vision is technically feasible. However, achieving this vision will require a combination of increased R&D on energy efficiency and policies that encourage conservation and use of high efficiency systems. It will also require some structural changes in the transportation systems. According to the 2004 World Energy Assessment by UNDP, a reduction of 25%–35% in primary energy in the industrialized countries is achievable cost effectively in the next 20 years, without sacrificing the level of energy services. The report also concluded that similar reductions of up to 40% are cost effectively achievable in the transitional economies and more than 45% in developing economies. As a combined result of efficiency improvements and structural changes such as increased recycling, substitution of energy intensive materials, etc., energy intensity could decline at a rate of 2.5% per year over the next 20 years (UNDP 2004).

TABLE 1.8 Potential and Status of RE Technologies

Technology	Annual Potential	Operating Capacity 2005[a,b]	Investment Costs U.S.$ per kW[b]	Current Energy Cost	Potential Future Energy cost
Biomass Energy	276–446 EJ total or 8–13 TW MSW ~ 6 EJ				
Electricity		~ 44 GW	500–6000/kW$_e$	3–12 c/kWh	3–10 c/kWh
Heat		~ 225 GW$_{th}$	170–1000/kW$_{th}$	1–6 c/kWh	1–5 c/kWh
Ethanol		~ 36 bln L	170–350/kW$_{th}$	25–75 c/L(ge)[c]	6–10 $/GJ
Biodiesel		~ 3.5 bln L	500–1000/kW$_{th}$	25–85 c/L(de)[d]	10–15 $/GJ
Wind Power	55 TW theoretical 2 TW practical	59 GW	850–1700	4–8 c/kWh	3–8 c/kWh
Solar Energy	> 100 TW				
Photovoltaics		5.6 GW	5000–10,000	25–160 c/kWh	5–25 c/kWh
Thermal power		0.4 GW	2500–6000	12–34 c/kWh	4–20 c/kWh
Heat			300–1700	2–25 c/kWh	2–10 c/kWh
Geothermal	600,000 EJ useful resource base 5000 EJ economical in 40–50 years				
Electricity		9 GW	800–3000	2–10 c/kWh	1–8 c/kWh
Heat		11 GW$_{th}$	200–2000	0.5–5 c/kWh	0.5–5 c/kWh
Ocean Energy					
Tidal	2.5 TW	0.3 GW	1700–2500	8–15 c/kWh	8–15 c/kWh
Wave	2.0 TW		2000–5000	10–30 c/kWh	5–10 c/kWh
OTEC	228 TW		8000–20,000	15–40 c/kWh	7–20 c/kWh
Hydroelectric	1.63 TW theoretical				
Large	0.92 TW econ.	690 GW	1000–3500	2–10 c/kWh	2–10 c/kWh
Small		25 GW	700–8000	2–12 c/kWh	2–10 c/kWh

[a] GW$_e$, gigawatt electrical power.
[b] GW$_{th}$, gigawatt thermal power.
[c] ge, gasoline equivalent liter.
[d] de, diesel equivalent liter.

Sources: Data from UNDP, *World Energy Assessment: Energy and the Challenge of Sustainability*, 2004, updated from other sources: Worldwatch Institute. 2006. *Biofuels for Transportation—Global Potential and Implications for Sustainable and Energy in the 21st Century*. Worldwatch Institute. Washington, DC, " Biofuels for Transportation—Global Potential and Implications for Sustainable andenergy in the 21st Century", Report prepared for the German Federal Ministry for Food, Agriculture and Consumer Protection, Worldwatch Institute, Wash., DC; World Wind Energy Association Bulletin, 2006, http://www.wwindea.org; EPIA, Photovoltoaic Barometer, http://www.epia.org; World Geothermal Power Generation 2001-[blc1]2005, GRC Bulletin, International Energy Annual, USEIA, 2006.

The summary of the economic energy efficiency potentials in North America up to the year 2010 are shown in Table 1.9. It is apparent that the greatest energy savings potential is in the transportation industry, followed by residential heating. The sources in the right-hand column refer to references in the United Nations study. In addition to the item cited in Table 1.9, it is believed that large energy savings are possible in office equipment, such as computers and communication. A similar estimate for the economic energy efficiency potential for Western Europe for the years 2010 and 2020 is presented in Table 1.10, where the resource references refer to the bibliography in Jochem (2000). Similar estimates for the energy saving potential in Japan, Asia and Latin America are presented in the same reference.

Improving energy efficiency across all sectors of the economy should become a worldwide objective (Energy Commission 2004). It should be noted, however, that free market price signals may not always be sufficient to effect energy efficiency. Hence, legislation on the state and/or national level for energy efficiency standards for equipment in the residential and commercial sector may be necessary. There is considerable debate whether incentives or mandates are the preferred way to improved energy efficiency.

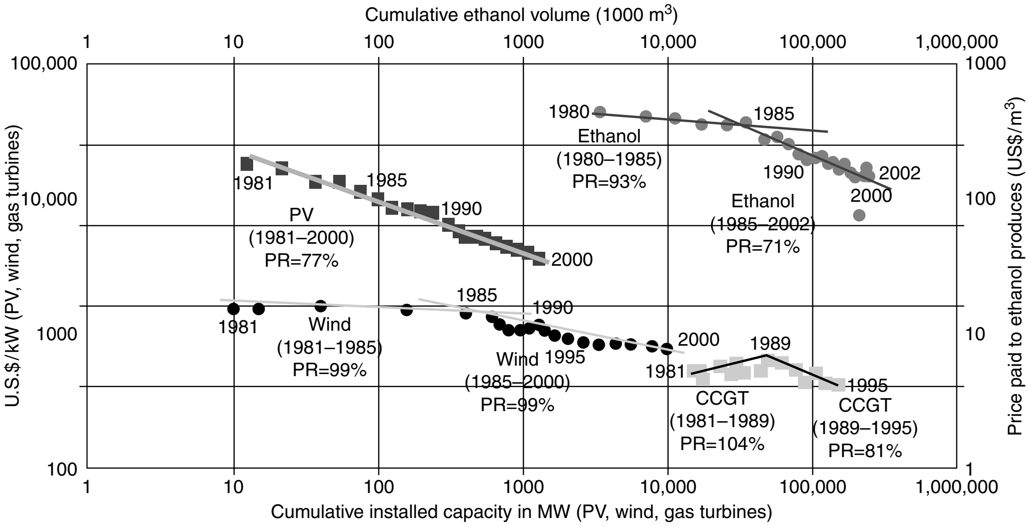


FIGURE 1.17 Experience curves for wind, PV, ethanol, and gas turbines. (Adapted from UNDP, *World Energy Assessment: Energy and the Challenge of Sustainability*, UNDP, 2004.)

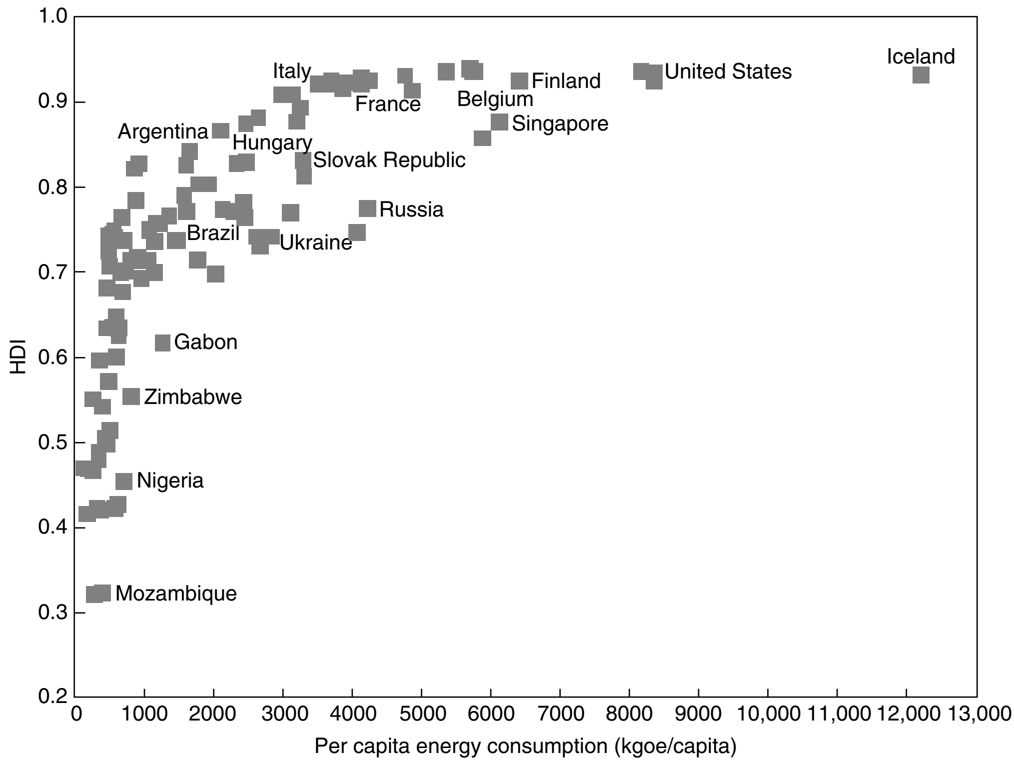


FIGURE 1.18 Relationship between Human Development Index (HDI) and per capita Energy Use, 1999–2000. (From UNDP, *World Energy Assessment: Energy and the Challenge of Sustainability*, UNDP, 2004. With permission.)

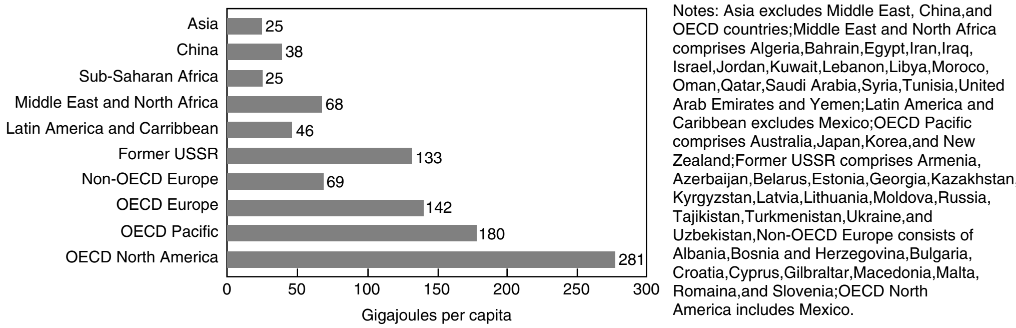


FIGURE 1.19 Per capita energy use by region (commercial and noncommercial), 2000. (From UNDP, *World Energy Assessment: Energy and the Challenge of Sustainability*, UNDP, 2004. With permission.)

Such measures may be necessary because surveys indicate that consumers consistently rank energy use and operating costs quite low on the lists of attributes they consider when purchasing an appliance or construct a building. Incentives may be the preferred option provided they induce decision makers to take appropriate action.

In the United States, by the year 2004, national efficiency standards were in effect for a variety of residential and commercial appliances. Updated standards will take effect in the next few years for several more products. Figure 1.20 shows the projected energy savings from upgraded standards for products installed in the years 2010–2020. Outside the United States, over 30 countries have also adopted

TABLE 1.9 Economic Energy Efficiency Potentials in North America, 2010

Sector and Area	Economic Potential (percent)		Energy Price Level Assumed	Base Year
	United States[a]	Canada		
Industry				
Iron and steel	4–8	29	United States: scenario for price developments[b]	United States: 1995
Aluminium (primary)	2–4			
Cement	4–8			
Glass production	4–8			Canada: 1990
Refineries	4–8	23	Canada: price scenario by province[c]	
Bulk chemicals	4–9	18		
Pulp and paper	4–8	9		
Light manufacturing	10–18			
Mining	n.a.	7		
Industrial minerals	n.a.	9		
Residential				
Lighting	53		United States: scenario for price developments	United States: 1995
Space heating	11–25			

(continued)

TABLE 1.9 (*Continued*)

Sector and Area	Economic Potential (percent)		Energy Price Level Assumed	Base Year
	United States[a]	Canada		
Space cooling	16			
Water heating	28–29			Canada: 1990
Appliances	10–33		Canada: price scenario	
Overall		13		
Commercial and public				
Space heating	48		United States: scenario for price developments	United States: 1995
Space cooling	48			
Lighting	25			
Water heating	10–20			Canada:1990
Refrigeration	31		Canada: price scenario	
Miscellaneous	10–33			
Overall	n.a.	9		
Transportation				
Passenger cars	11–17		United States: scenario for price developments	United States: 1997
Freight trucks	8–9			
Railways	16–25			
Aeroplanes	6–11			Canada: 1990
Overall	10–14	3	Canada: price scenario	

[a] Industrial energy efficiency potentials in the United States reflect an estimated penetration potential under different conditions based on the Interlaboratory Working Group on Energy Efficient and Low-Carbon Technologies (1997). There are no separate estimates available for the economic potential. The economic potential under business-as-usual fuel price developments is estimated at 7 percent in energy-intensive industries and 16 percent in light industries.

[b] The Inter-Laboratory Working Group study (1997) used price scenarios for 1997–2010 to estimate the potential for energy efficiency improvement, based on the *Annual Energy Outlook 1997* scenario (EIA, 1996). The scenario assumes a 1.2 percent annual increase in oil prices from 1997 levels.

[c] For comparison; in 2010 light fuel oil price are $6–8 a gigajoule at the 1999 exchange rate (Jaccard and Willis Energy Services, 1996).

Source: From UNDP (United Nations Development Programme), World energy assessment: Energy and the challenge of sustainability, New York, 2004.

minimum energy performance standards. These measures have been shown to be economically attractive and can provide an appreciable reduction in adverse environmental impacts.

This handbook describes energy efficiency improvements achievable with available technologies. The challenge is to adopt policies that accelerate the adoption of these technologies all over the world.

1.6 Forecast of Future Energy Mix

As explained above, it is clear that oil production will peak in the near future and will start declining thereafter. Since oil comprises the largest share of world energy consumption, a reduction in availability of oil will cause a major disruption unless other resources can fill the gap. Natural gas and coal production may be increased to fill the gap, with the natural gas supply increasing more rapidly than coal.

TABLE 1.10 Economic Energy Efficiency Potentials in Western Europe, 2010 and 2020

Sector and Technological Area	Economic Potential (percent)[a]		Energy Price Level Assumed	Base Year
	2010	2020		
Industry				
Iron and steel, coke ovens	9–15	13–20	1994	1995
Construction materials	5–10	8–15	1997	1997
Glass production	10–15	15–25	1997	1997
Refineries	5–8	7–10	1995	1997
Basic organic chemicals	5–10		1997	1996
Pulp and paper		50	1996	1997
Investment and consumer goods	10–20	15–25	1994	1995
Food	10–15		1997	1997
Cogeneration in industry		10–20	1997	1997
Residential				
Existing buildings				
Boilers and burners	15–20	20–25	Today's prices	1997
Building envelopes	8–12	10–20	Today's prices	1995
New buildings		20–30	Today's prices	1995
Electric appliances	20–30	35–45	1997	1997
Commercial, public, and agriculture				
Commercial buildings	10–20	30	8–13 c/kWh	1995
Electricity	10–25	20–37	4–10 c/kWh	1997
Heat		15–25	Today's prices	1998
Public buildings		30–40	7–15 c/kWh	1992
Agriculture and forestry		15–20	Today's prices	
Horticulture		20–30	Today's prices	
Decentralised cogeneration		20–30	Today's prices	1995
Office equipment		40–50	1995	1995
Transportation				
Cars	25		Today's prices	1995
Door-to-door integration	4			1995
Modal split of freight transport		3[b]		1995
Trains and railways		20	Today's prices	1999
Aircraft, logistics	15–20	25–30	Today's prices	1998

[a] Assumes a constant structure or use of the sector or technology considered.
[b] Refers to the final energy use of the entire sector.

However, that will hasten the time when natural gas production also peaks. Additionally, any increase in coal consumption will worsen the global climate change situation. Although CO_2 sequestration is feasible, it is doubtful that there will be any large-scale application of this technology for existing plants. However, all possible measures should be taken to sequester CO_2 from new coal-fired power plants. Presently, there is a resurgence of interest in nuclear power. However, it is doubtful that nuclear power alone will be able to fill the gap. Forecasts from IAEA show that nuclear power around the world will grow at a rate of 0.5%–2.2% over the next 25 years (IAEA 2005). This estimate is in the same range as that of IEA.

Based on this information it seems logical that the RE technologies of solar, wind and biomass will not only be essential but will hopefully be able to fill the gap and provide a clean and sustainable energy future. Although, wind and photovoltaic power have grown at rates of over 30%–35% per year over the last few years, this growth rate is based on very small existing capacities for these sources. There are many

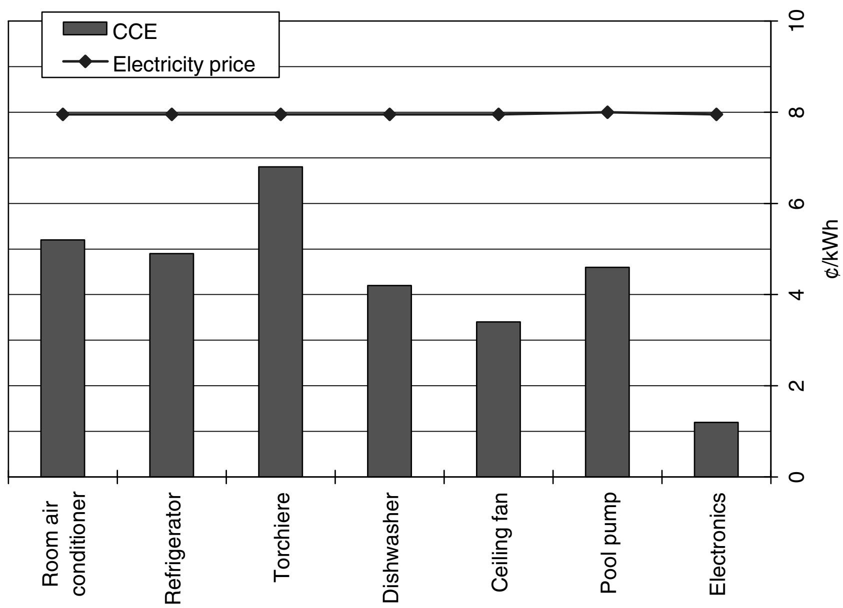


FIGURE 1.20 Comparison of cost of conserved energy for 2010 standards to projected electricity price in the residential sector.

differing views on the future energy mix. The IEA estimates (Figure 1.21) that the present mix will continue until 2030 (IEA 2004).

On the other hand, the German Advisory Council on Global Change (WBGU) estimates that as much as 50% of the world's primary energy in 2050 will come from RE, increasing to 80% by 2100 (Figure 1.22; WBGU 2003). However to achieve that level of RE use by 2050 and beyond will require worldwide effort on the scale of a global Apollo Project.

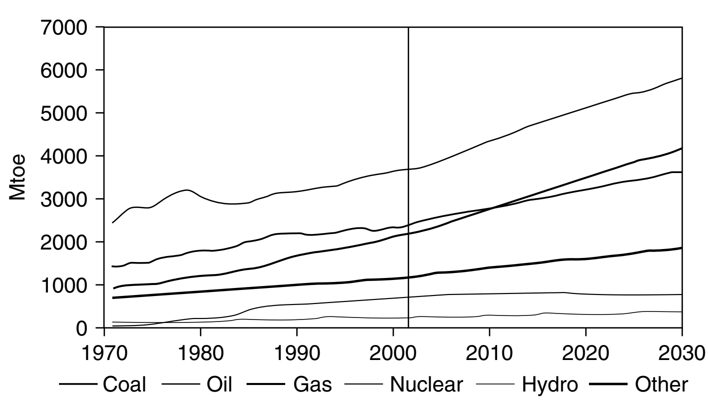


FIGURE 1.21 **(See color insert following page 774.)** World primary energy demand by fuel types according to IEA. (From IEA, *World Energy Outlook*, IEA, Paris, 2004.)

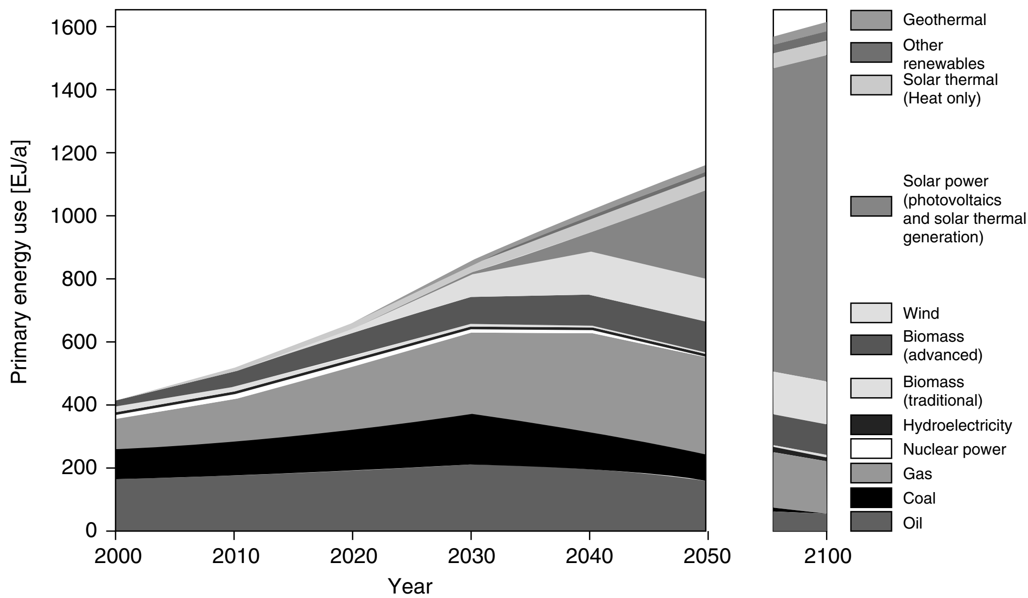


FIGURE 1.22 **(See color insert following page 774.)** The global energy mix for year 2050 and 2100 according to WBGU. (From WBGU, World in Transition—Towards Sustainable Energy Systems, German Advisory Council on Global Change, Berlin, 2003.)

References

Associated Press. 2006. *Xcel pursues 'clean coal' plant in Colo.* (August).

Bartlett, A. A. 2002. An analysis of U.S. and world oil production patterns using Hubbert-style curves, *Mathematical Geology*, 32, 1–17.

Barlett, A. A. 2004. Thoughts on long-term energy supplies: scientists and the silent lie, American Institute of Physics, 57(7), 53–55.

British Petroleum Corp. (BP). 2006. BP statistical review of world energy 2006. British Petroleum, London, UK. http://www.bp.com/statisticalreview/ (accessed on August 1, 2006).

Energy Commission. 2004. Ending the energy stalemate: A bipartisan strategy to meet America's energy challenges. *The National Commission on Energy Policy*, http://www.energycommission.org (accessed on October 12, 2006).

EPIA. 2006. *Photovoltaic Barometer*. European Photovoltaic Industry Assoc., Brussels. http://www.epia.org (accessed on October 12, 2006)

EPRI (Electric Power Research Institute). 2004. Advanced batteries for electric drive vehicles: A technology and cost-effectiveness assessment for battery electric vehicles, power-Assist hybrid electric vehicles, and plug-in hybrid electric vehicles, EPRI Technical Report 1009299, EPRI, Palo Alto, CA.

ESTIF (European Solar Thermal Industry Federation). 2005. *Solar thermal markets in Europe—Trends and statistics for 2004*, ESTIF, Brussels, Belgium.

GEF (Global Environmental Facility). 2005. Assessment of the World Bank/GEF strategy for the market development of concentrating solar thermal power, GEF report GEF/c.25/Inf.11, GEF, Washington, DC.

Hammererschlag, R. and Mazza, P. 2005. Questioning Hydrogen. *Energy Policy*, 33, 2039–2043.

Hawkins, D. G., Lashof, D. A., and Williams, R. H. 2006. What to do about coal—Carbon capture and storage, *Scientific American*, September, 68–75.

Hubbert, M. K. 1974. U.S. energy resources: A review as of 1972, In *A National Fuels and Energy Policy Study*, Serial No. 93–40 (92–75), Part 1. US Government Printing Office, Washington, DC.

IAEA (International Atomic Energy Agency). 2005. Energy, electricity and nuclear power estimates to 2030, Reference data Series No. 1, Vienna, Austria.

IEA (International Energy Agency). 2004. *World energy outlook*, IEA, Paris.

IEA (International Energy Agency). 2005. *Renewables information 2005*, IEA, Paris.

Jochem, E. 2000. Energy end-use efficiency, In *World Energy Assessment*, J. Goldenberg, ed., pp. 73–217, United Nations Development Project, New York.

Kreith, F. 2006. Plenary lecture, 2006 ASES Conference, Denver, CO. ASES, Boulder, CO.

Kreith, F. and West, R. E. 2004. Fallacies of a hydrogen economy, *JERT*, 126, 249–257.

Kreith, F., West, R. E., and Isler, B. 2002. Legislative and technical perspectives for advanced ground transportation systems, *Transportation Quarterly*, 56, 51–73.

Maycock, P. 2006. PV News annual review of the PV market, Available at http://www.epia.org (accessed on August 1, 2006).

Sadler H., Riedy, C., and Passey, R. 2004. *Geosequestration*, Discussion Paper Number 72, The Australia Institute, Sydney, Australia.

Sanna, L. 2005. Driving the solution—Plug-in hybrid vehicles, *EPRI Journal*, Fall 2005.

Shinnar, R. 2003. The hydrogen economy, fuel cells, and electric cars, *Technology in Society*, 25, 453–576.

UNDP (United Nations Development Programme). 2004. World energy assessment: Energy and the challenge of sustainability, New York.

Veziroglu, T. N. and Barbir, F. 1992. Hydrogen: The wonder fuel, *International Journal of Hydrogen Energy*, 17, 391–404.

WBGU (Germany Advisory Council on Global Change). 2003. World in transition—Towards sustainable energy systems, WBGU, Berlin. http://www.wbgu.de (accessed on August 1, 2006).

West, R. E. and Kreith, F. 2006. A vision for a secure transportation system without hydrogen, or oil *ASME Journal of Energy Resources Technologies*, JERT-06-1009-1, 128, 236–243.

Worldwatch Institute 2006. *Biofuels for Transportation—Global Potential and Implications for Sustainable and Energy in the 21st Century,* Worldwatch Institute, Washington, DC.

World Wind Energy Association. 2006. *Worldwide wind energy Bulletin, 2006*, Bonn, Germany, Available at http://www.wwindea.org (accessed on October 12, 2006).

US Energy Information 2006. *International Energy Outlook 2006*, EIA, Washington, DC.

web sites: http://pubs.usgs.gov/dds/dds-060/sum1.html; http://www.eia.doe.gov/emeu/international/oilreserves.html; http://www.eia.doe.gov/oiaf/ieo/oil.html (accessed on August 1, 2006).

2

Energy Policy

Christopher Namovicz
United States Department of Energy

Michael Durstewitz
Institut für Solare Energieversorgungstechnik (ISET)

Gershon Grossman
Technion—Israel Institute of Technology

Anil Misra
Solar Energy Society of India

Ricardo Rüther
Universidade Federal de Santa Catarina

Koichi Sakuta
National Institute of Advanced Industrial Science and Technology (AIST)

Jeff Bell
World Alliance for Decentralized Energy

Debra Lew
National Renewable Energy Laboratory

2.1 U.S. State and Federal Policies for Renewables

Christopher Namovicz

2.1.1 Tax Incentives

Tax incentives have provided a key form of direct subsidy to renewable energy and energy efficiency in the U.S. at both the state and Federal levels. These incentives can take several forms, including deductions from taxable income or a credit against tax liability. In addition, tax credits can be applied to the initial purchase or investment in a particular technology, or to the ongoing utilization of a technology or production of a covered commodity.

2.1.1.1 Investment

Renewable energy and energy-efficient technologies are typically characterized by higher up-front costs resulting in significantly reduced fuel and/or operating costs (although not all technologies fit this characterization, for example, biomass energy can involve substantial ongoing costs for fuel and operations). Many early policy incentives at both the Federal and state levels were intended to reduce the acquisition cost of these technologies, frequently through the use of tax credits proportional to capital investment costs. In some cases, such as some of the early deployment of wind generating technology in California during the 1980s, it was believed that investment incentives provided insufficient incentive for high-quality technology or projects that would continue to operate once the initial incentive had been completely realized by the project owner. However, such failures may also be attributed to insufficient technology qualification measures, such as technology criteria or screening [1]. Despite the apparent shortcomings of investment incentives in the early U.S. wind industry, these continue to see widespread use in both Federal and state policy for other renewable energy and energy efficiency technologies.

Tax incentives applied against the investment in renewable energy technologies played a significant role in the early adoption of these technologies during the 1980s. Originally adopted as part of the Energy Tax Act of 1978, and most recently set at 10% for solar and geothermal facilities by the Energy Policy Act of 1992 (EPACT 92), the Investment Tax Credit directly offsets Federal corporate income tax liability in proportion to the initial investment cost of the covered technology [2]. In addition, most renewable electricity generating technologies are also able to benefit from preferential Federal tax depreciation allowance schedules [3]. The Modified Accelerated Cost Recovery Schedule allows much faster depreciation of renewable generation investment costs than that is allowed for other generation technologies, using a 5-year schedule rather than a 15- or 20-year schedule for combustion turbines or other thermal plants [4].

Some states also have or have had tax incentives on the investment in renewable energy or energy efficiency. Additional investment tax credits in California during the 1980s, along with other policies discussed elsewhere in this section, helped spur the early adoption of wind and solar thermal generating capacity in that state [5]. California, among other states, presently offers substantial investment tax credits to preferred renewable energy technologies, such as photovoltaic (PV) systems [6]. However, these credits are not uniformly offered; vary significantly among states that do offer them, and may apply to electric generating technologies or facilities that produce renewable fuels such as ethanol. Rebates or exemptions from state-imposed sales taxes on both renewable technologies and energy-efficient appliances and equipment, also offer a mechanism to reduce the first-cost of adopting these technologies by the end user. Availability of such programs varies significantly among states, as do the sales-tax rates and the value of a rebate or exemption where offered [7]. Sales tax rebates may also (or instead) apply to a renewable fuel, such as biomass fuel. In this context, such a program may have an effect closer to that of a production incentive rather than an investment incentive.

2.1.1.2 Production/Utilization

Production-based tax incentives provide a tax credit proportional to the quantity of commodity, such as electricity generation, produced or sold in a given year. Since production-based incentives reward project performance, they should tend to transfer project performance risk to the project owner, rather than the taxing authority, and without the need for extensive qualification criteria or screening of each project or technology. However, technologies that do not produce easily marketable (and hence taxable) output, such as most energy-efficiency technologies, or where the output is generally consumed on-site (without a third-party transaction), such as on-site PV, may be not be amenable to a production-based incentive, as there may not be a sufficiently auditable record of production or the establishment of such an auditable record (such as internal metering of PV output) may add unwanted cost to a project.

The production tax credit for renewable electricity (PTC), Section 45 of the U.S. Internal Revenue Code, established by EPACT 92, provides an inflation-adjusted payment, 1.9 cents/kwh in 2005, for the third-party electricity sales from the plant during the first 10 years of operation. The range of technologies eligible for the tax credit has been expanded since its inception, and now includes wind, several types of biomass resources, solar, geothermal, and landfill-gas. However, some technologies do not receive the full credit amount or the same 10-year claim period [8]. The PTC has generally been credited with contributing to the significant growth in U.S. wind power since 1998. Having been allowed to expire and subsequently extended several times, the credit is presently set to expire at the end of 2005 [9]. A number of states also offer tax credits on the production of preferred renewable energy sources [10].

The production of certain alternative fuels, including landfill-gas, which is frequently classified as a renewable resource, was eligible for a tax credit under Section 29 of the U.S. Internal Revenue Code. Landfill-gas facilities placed in service between 1980 and July 1, 1998 are presently eligible to receive an inflation-adjusted $3 tax credit on each unit of production equivalent in energy content to a barrel of oil. Although extension of the in-service eligibility date has been proposed as part of various comprehensive energy bills, it has not been enacted into law [11].

Presently, there is a Federal tax credit of $0.52 per gallon of ethanol blended into gasoline, which will be reduced to $0.51 per gallon for 2005 and 2006 and expire in 2007. The Federal tax credit for ethanol has been extended several times in the past [12]. A number of states also have tax credits for the production of ethanol or other renewable fuels. These credits may reduce income tax liability or, like the Federal credit, be applied to a motor fuels tax (in effect, a sales tax rebate). State programs vary by credit amount as well as by restrictions on local origin of the fuel [13].

2.1.2 Regulatory

Regulatory mechanisms generally establish restrictions on market activity that are intended to result in increased adoption of policy-preferred technologies or limitation on policy-undesired technologies. Costs are typically born directly by market participants; either energy producers or consumers or both. Although regulatory policy may affect markets in many ways, this section will examine three major types of regulatory intervention: technology specification standards, target-based standards, and market facilitation or limitation policies. Table 2.1 contains a summary of several types of regulatory policy by state.

2.1.2.1 Target-Based Standards

Target-based standards establish a target metric of renewable energy or energy efficiency achievement and require regulated industry to achieve the goal. The two most important types of goal-based standards in U.S. energy policy are renewable electricity targets established by the various states and automotive fuel efficiency standards established by the Federal government (Table 2.2).

Renewable electricity targets can take the form of absolute levels of capacity (or generation) or a specified fraction of some future level of total generation (or capacity). The former, sometimes called renewable mandates, and the latter, generally called renewable portfolio standards (RPS), can be targeted for single future year or can be based on a gradually increasing compliance schedule. Renewable energy goals—found in a few states—can mimic either mandates or RPS programs, but generally lack enforceability provisions, and thus cannot be considered as regulatory policy [14]. Mandates or RPS policies can require absolute compliance by affected utilities, or, as frequently occurs, can allow the

TABLE 2.1 Database of State Incentives for Renewable Energy

	Systems Benefit Charge	Generation Fuels Disclosure	RPS/Renewable Target	Net Metering	Interconnection Standards	Resource Access/Easement Laws[a]	Green Power[b]	Other[c]
Alabama								
Alaska						✓		
Arizona	✓	✓	✓	✓	✓	✓		✓
Arkansas				✓	✓			✓
California	✓	✓	✓	✓	✓	✓		✓
Colorado		✓	✓	✓	✓	✓		✓
Connecticut	✓	✓	✓	✓	✓			✓
Delaware	✓	✓	✓	✓	✓			✓
Florida		✓	✓	✓	✓	✓		✓
Georgia				✓	✓	✓		
Hawaii			✓	✓	✓	✓		✓
Idaho				✓	✓	✓		
Illinois	✓	✓	✓	✓	✓			✓
Indiana				✓	✓	✓		
Iowa		✓	✓	✓	✓	✓	✓	✓
Kansas					✓	✓		
Kentucky				✓	✓	✓		
Louisiana				✓	✓			✓
Maine	✓	✓	✓	✓		✓		✓
Maryland		✓	✓	✓	✓	✓		✓
Massachusetts	✓	✓	✓	✓	✓	✓		✓
Michigan		✓		✓	✓			✓
Minnesota	✓	✓	✓	✓	✓	✓	✓	✓
Mississippi								
Missouri			✓		✓	✓		
Montana	✓	✓	✓	✓	✓	✓	✓	
Nebraska						✓		
Nevada		✓	✓	✓	✓	✓		✓
New Hampshire				✓	✓	✓		
New Jersey	✓	✓	✓	✓	✓	✓		✓
New Mexico			✓	✓	✓	✓	✓	✓
New York	✓	✓	✓	✓	✓	✓		✓
North Carolina					✓	✓		✓
North Dakota				✓		✓		

Ohio	✓	✓		✓	✓	✓		✓
Oklahoma				✓				
Oregon	✓	✓		✓	✓	✓		✓
Pennsylvania	✓	✓	✓	✓				✓
Rhode Island	✓		✓	✓	✓	✓		✓
S. Carolina								✓
South Dakota								
Tennessee						✓		
Texas		✓	✓	✓	✓			✓
Utah				✓	✓	✓		✓
Vermont		✓	✓	✓	✓			
Virginia		✓		✓	✓	✓		
Washington		✓		✓	✓	✓	✓	✓
West Virginia								
Wisconsin	✓		✓	✓	✓	✓		✓
Wyoming				✓	✓			
D.C	✓	✓	✓	✓	✓			
Palau								
Guam								✓
Puerto Rico								✓
Virgin Island								
N. Mariana Islands								
Americans Samoa								

[a] Limits ability of adjoining property owners from interfering with access to renewable energy resource, such as shading the sun
[b] Requires utilities to offer green power
[c] Includes contractor licensing, equipment certification, construction and design standards, and renewable energy purchase or aggregation by state or local government
Source: From http://www.dsireusa.org/summarytables/reg1.cfm?&CurrentPageID=7, accessed November 2, 2005.

TABLE 2.2 Database of State Incentives for Renewable Energy Financial Incentives

State/Territory	Personal Tax	Corporate Tax	Sales Tax	Property Tax	Rebates	Grants	Loans	Leasing/Sales	Production Incentive
Alabama	✓					✓			
Alaska							✓		
Arizona	✓		✓		✓				
Arkansas									
California	✓	✓		✓	✓	✓	✓	✓	✓
Colorado					✓		✓		✓
Connecticut				✓	✓	✓	✓		
Delaware					✓	✓			
Florida			✓		✓				
Georgia									✓
Hawaii	✓	✓			✓		✓		
Idaho	✓		✓			✓	✓		
Illinois				✓	✓	✓			
Indiana				✓		✓			
Iowa	✓	✓	✓	✓		✓	✓		
Kansas				✓		✓			
Kentucky							✓		✓
Louisiana				✓					
Maine					✓	✓			
Maryland	✓	✓	✓	✓	✓		✓		
Massachusetts	✓	✓	✓	✓	✓	✓			✓
Michigan				✓	✓	✓	✓		
Minnesota			✓	✓	✓	✓	✓		✓
Mississippi							✓		✓
Missouri		✓				✓	✓		
Montana	✓	✓		✓		✓	✓		
Nebraska							✓		
Nevada			✓	✓	✓				✓
New Hampshire				✓					
New Jersey			✓		✓	✓	✓		✓
New Mexico		✓	✓			✓			
New York	✓		✓	✓	✓	✓	✓		
North Carolina	✓	✓		✓			✓		✓
North Dakota	✓	✓	✓	✓					
Ohio		✓	✓	✓		✓	✓		
Oklahoma		✓							

Oregon	✓	✓		✓	✓?	✓	✓		✓
Pennsylvania					✓	✓	✓		✓
Rhode Island	✓		✓	✓	✓	✓			✓
South Carolina					✓				
South Dakota				✓					
Tennessee				✓			✓		✓
Texas		✓		✓	✓			✓	
Utah	✓	✓	✓						
Vermont			✓		✓	✓			✓
Virginia				✓		✓			
Washington			✓		✓	✓	✓		✓
West Virginia		✓		✓					
Wisconsin				✓	✓	✓	✓		
Wyoming			✓		✓			✓	
District of Columbia						✓			
Palau									
Guam									
Puerto Rico	✓		✓						
Virgin Islands									
N. Mariana Islands									
American Samoa									

Source: From http://www.dsireusa.org/summarytables/financial.cfm?&CurrentPageID=7, accessed November 2, 2005.

accumulation of "renewable energy credits" (REC) that can facilitate either intertemporal compliance "banking" (i.e., using RECs earned in a year to meet compliance targets in another year) and/or inter-utility or interstate credit trading (whereby, a utility that overcomplies may sell RECs to a utility that cannot meet targets with native resources). Most states with mandates or RPS policies limit the geographic source of compliance to in-state resources, resources within the electric power pool(s) that service the state, or resources that can be "delivered" to the state or state power pool. The prevailing selling price of RECs may also be used to target a penalty or alternative compliance payment, typically in the form of a price ceiling at which the state will provide RECs (without actual renewable capacity or generation) or otherwise waive actual compliance. Such "safety-valve" prices are generally intended to provide a clear maximum impact on general electricity prices. Compliance in various states may also be waived or delayed for other, statutorily sanctioned reasons, such as protecting the financial solvency of affected utilities. Policies among states also show significant variation in resource eligibility, "grand-fathering" of existing capacity, and mechanisms to show preferences among eligible technologies (such as awarding "bonus" credits or having differentiated targets for preferred technologies) [15].

In 1975, the Federal government established a target of doubling automotive fuel efficiency within 10 years. To implement this target, the aggregate sales of each manufacturer selling cars in the U.S. market had to achieve a set schedule for Corporate Average Fuel Economy (CAFE) [16]. Presently, the CAFE for each manufacturer must be 27.5 mpg (miles per gallon) for passenger cars. Lesser standards are presently set for light-duty trucks, 21 mpg for the 2005 model year, increasing to 22.2 mpg for the 2007 model year. Compliance credits allow the "banking" of mileage shortfall/overage within 3 years of a compliance year, but presently there is no mechanism in place for trading compliance shortfall/overage among manufacturers. In addition to supporting more efficient cars, present CAFE regulations also support alternative fueled vehicles, including ethanol. Vehicles that operate on an alternative fuel, or can switch between a conventional fuel (gasoline or diesel) and an alternative fuel (known as "flex-fuel" vehicles), receive additional compliance credits, based on a multiplier of the alternative fuel portion of the gas mileage of the vehicle. Although this has resulted in a significant fleet of flex-fuel vehicles in the U.S., it has not resulted in a refueling infrastructure for alternative fuels or in significant sales of alternative fuels.

2.1.2.2 Market Facilitation or Restriction

Regulatory policy can also be used to facilitate or hinder a preferred or undesirable renewable energy or energy efficiency technology from participating in the market. Facilitation can take many forms, including the target-based and technology-specification approaches discussed elsewhere in this section. Other types of market facilitation can require nondiscriminatory or even preferential market treatment of preferred technologies. Such policies operating at the Federal or state level can include "feed-in" laws, net-metering requirements, and interconnection standards.

In 1978, the Congress passed the Public Utility Regulatory Policy Act (PURPA), which established the requirement that electric utilities must interconnect (i.e., accept generation feed from) small qualifying facilities that either cogenerate process heat and electricity (combined heat and power or CHP) or utilize certain renewable resources [17]. Furthermore, PURPA established a price floor for the power, known as "avoided cost," subsequently defined to mean the cost of electricity that the utility otherwise would have purchased. In theory, PURPA established a nondiscriminatory framework for adoption of efficient industrial CHP and renewable electricity, established by the Federal government, but largely implemented by state regulatory authorities. Some of the nondiscriminatory market features that PURPA specifically applied to renewable and CHP facilities were subsequently applied to the broad class of all power generation technologies as Federal electricity policy moved toward deregulation of the wholesale power market [18].

Many states have adopted regulations at the retail/distribution level to require the acceptance of some renewable electricity feeds at an established price floor [19]. Such net-metering laws typically require load serving utilities to facilitate end user connection of renewable distributed generation technologies (especially solar, but sometimes wind or other renewable or nonrenewable technologies) on the customer side of the meter. When instantaneous generation from the local resource exceeds instantaneous customer demand, the meter is allowed to "run backward," effectively causing the utility to purchase the excess generation at the prevailing retail rate. Most states limit the size of the distributed resource, sometimes by customer class, and may also provide limits on the total generation off-set allowed (e.g., the monthly or net annual bill may not be less than zero). Some states have also established limits on the number of customers or level of installed distributed capacity that may participate in net-metering.

2.1.2.3 Technology Specification Standards

Another common form of regulatory intervention for renewable and energy-efficient technologies is the establishment of minimum product specifications, either as voluntary targets or mandatory limits on product performance. Such standards are seen as an effective approach to improving energy efficiency among individual consumers. Commercial and industrial consumers presumably have significant incentive to optimize energy efficiency for their operations to maintain or improve profitability. However, individuals, while still sensitive to energy prices, may have less motivation to seek out products with higher up-front costs to achieve lower ongoing energy costs. In some cases, market structures may affect consumer decision-making with respect to energy efficiency.

The Federal Energy Star program allows qualifying products—ranging from computer equipment to household appliances, to commercial building equipment—to display the "Energy Star" logo on product advertising and packaging [20]. This serves as a proxy for disclosure, in that the consumer is thus aware that the product is "best-in-class" for energy efficiency (although for products not displaying the logo, the consumer cannot tell if this is because the product did not meet the specification or because the manufacturer did not participate in the program). Through the Energy Policy and Conservation Act and its various amendments, the Federal government also establishes mandatory energy efficiency specifications, such as minimum levels of energy efficiency, for a wide array of consumer appliances, such as furnaces, air conditioners, light fixtures, and kitchen appliances [21]. At the state and local level, energy efficiency standards may also be incorporated into building codes.

There are both Federal and state regulations regarding transportation fuel composition that either directly or indirectly provide incentive for renewable fuels. The Federal CAFE program is discussed in the previous section. In addition, the Clean Air Act Amendments of 1990 (CAA) established a number of fuel specifications, including oxygenation that varies by region and/or season [22]. Ethanol has emerged as a preferred oxygenate, especially in states with additional ethanol incentives or that have restricted the use of alternatives, such as MTBE [23]. Restrictions on the sulfur content in diesel fuels may also encourage the use of "biodiesel" fuels derived from plant oils, if such fuels can be economically produced. Federal legislation to require minimum renewable energy content in highway fuels (sometimes referred to as a "renewable fuels standard") has previously been proposed, but has not become law as of this writing [24]

While regulatory action can encourage development of renewable and energy efficiency technologies, it can also discourage this development. Some regulatory actions are directly aimed at limiting the use of selective technologies in some locations. A few states, for example, have proposed or enacted moratoria on the development of utility-scale wind power plants;[25] used zoning or permitting processes to limit, delay, or stop development of renewable energy projects in undesired locations;[26] or established technical specification for interconnection to the power system that limits the adoption of some renewable technologies [27]. At the local level, homeowners may find deed covenants from quasi-governmental homeowner associations that restrict external modifications to a house, such as the addition of solar collectors (for PV or hot water), or may specifically restrict these technologies [28].

2.1.3 Research and Development

Government research and development (R&D) funding for renewable and energy efficiency technologies can support the adoption of these technologies by facilitating cost reductions, higher efficiency, and improved utilization. R&D funding may occur at all stages of the technology development cycle, including basic science, bench-scale technology development, proof-of-concept demonstration, and pilot applications [29]. Government funds may be directed toward government-owned research laboratories, academic institutions, or industry participants. For many projects, especially those developing technologies closer to commercialization, the government will leverage its contributions by requiring substantial cost-sharing (either financial or in-kind) with industry participants.

2.1.4 Other Direct Policy

Other common programs at the stated and Federal level, include direct payments (such as through grants or awards) and government purchase of these technologies. These mechanisms generally require continuing budgetary support, which may be provided from a dedicated revenue source, or may require periodic affirmation in appropriations process.

Direct payments, such as the Federal Renewable Electricity Production Incentive (REPI), are useful to publicly owned or not-for-profit (such as rural electric cooperatives) utilities interested in developing renewable energy resources, but which do not pay taxes. Presently, authority to provide REPI payments to new sources has expired [30]. A number of states have established system benefit funds dedicated to supporting renewable energy and energy efficiency projects and technologies. Although varying greatly by state, these programs are typically structured to collect revenue based on an additional fee on retail generation or billing, commonly referred to as a systems benefit charge or public benefit fund [31]. As a result of EPACT and subsequent Presidential orders, the various agencies of the Federal government are required to obtain a share of their energy from renewable sources and reduce their consumption of energy per square foot of facility [32]. Finally, the Federal-owned fleet of cars and other vehicles is required to meet requirements for both fuel economy and use of alternative fuels. Several state local governments have also established similar purchase or efficiency requirements for electricity or motor fuels [33].

2.1.5 Indirect Policy

Numerous other policies at the state and Federal level, while not designed specifically to address renewable energy and energy efficiency markets, may have significant or notable impact on these markets. Perhaps most significant among this broad category are efforts to regulate energy or other markets, manage government- or privately owned lands, and protect the environment.

Efforts at the Federal level to introduce competition in wholesale electricity generation markets, as well as in a number of states to introduce competitive retail electricity supply, have created the opportunity for electricity suppliers to sell "green" power—typically electricity produced from renewable, low-emission, or high efficiency technologies [34]. Such programs include competitive supply of clean or renewable power, special pricing for green power by regulated utilities, or the sale of the environmental attributes of renewable power apart from sale of electricity. In addition, the specific design of competitive wholesale markets for generation and transmission can impact the competitiveness of some renewable, especially intermittent resources like wind. FERC has recently begun an effort to address some of these intermittency issues, such as those related to schedule imbalance charges and nonfirm transmission rights [35].

Environmental regulation at the Federal or state level, for air quality, water quality, solid waste disposal, land use, and other pollution problems, can have substantial impact on both the cost and value of renewable energy and energy efficiency. The CAA of 1990 provides the foundation for cap-and-trade regulation of sulfur dioxide, and recently enacted cap-and-trade programs for nitrogen oxides and mercury [36]. While the cap-and-trade structure of these programs does not preclude renewable energy or energy efficiency from displacing covered emissions at some plants, present regulations do not assign allowances to nonfossil resources. Other CAA impacts on renewable energy and energy efficiency, include reformulated gasoline requirements discussed above, which have interacted with state-level groundwater protection efforts to provide a preference (in some states) for ethanol as a preferred fuel additive for CAA compliance. As a result of the Resource Conservation and Recovery Act, some landfill operations have been required to install collection and flaring systems to prevent the dangerous buildup of methane-rich gas that results from the decomposition of organic matter in the landfills [37] These systems have significantly reduced the cost of deploying small generators fueled by this off-gas. Impacts of land management policy at both the Federal and state level can be significant factors in renewable energy policy, either to encourage or preclude its development on government owned land.

References

1. Geilecki, M. and Poling, J. 2005. *Policies to Promote Non-hydro Renewable Energy in the United States and Selected Countries*. Energy Information Administration. p. 16. http://www.eia.doe.gov/cneaf/solar.renewables/page/non_hydro/nonhydrorenewablespaper_final.pdf#page=1 (accessed 5/31/05).
2. Geilecki, M., Mayes, F., and Prete, L. 2001. *Incentives, Mandates, and Government Programs for Promoting Renewable Energy*. Energy Information Administration. http://www.eia.doe.gov/cneaf/solar.renewables/rea_issues/incent.html (accessed 5/31/05).
3. See Geilecki, Mayes, Prete above.
4. Internal Revenue Service. *How To Depreciate Property*. Publication 946 (2004). http://www.irs.gov/publications/p946/index.html (accessed 5/31/05).
5. See pp. 8–10 of Geilecki and Poling above.
6. Interstate Renewable Energy Council. *Database of State Incentives for Renewable Energy*. http://www.dsireusa.org/ (accessed 5/31/05).
7. See Interstate Renewable Energy Council above.
8. Energy Information Administration. Issues in Focus: Production Tax Credit for Renewable Electricity Generation. pp. 58–62, *Annual Energy Outlook 2005*. DOE/EIA-0383(2005). 2005. http://www.eia.doe.gov/oiaf/aeo/pdf/issues.pdf (accessed 5/31/05).
9. See Tax Relief Extension Act of 1999, Public Law 106–170 and Job Creation and Worker Assistance Act of 2002, Public Law 107–147.
10. See Interstate Renewable Energy Council above.
11. Energy Information Administration. *Analysis of Five Selected Tax Provisions of the Conference Energy Bill of 2003*. http://www.eia.doe.gov/oiaf/servicerpt/ceb/fuel.html
12. Energy Information Administration. *Federal Financial Interventions and Subsidies in Energy Markets 1999: Primary Energy. Appendix B*. pp. 78–80. SR/OAIF/99-03. 1999. http://www.eia.doe.gov/oiaf/servicerpt/subsidy/pdf/sroiaf(99)03.pdf
13. See Interstate Renewable Energy Council above.
14. Petersik, T. *State Renewable Energy Requirements and Goals: Status Through 2003*. Energy Information Administration. 2004. http://www.eia.doe.gov/oiaf/analysispaper/rps/index.html (accessed 5/31/05).
15. See Petersik and Interstate Renewable Energy Council above.
16. National Highway Traffic Safety Administration (NHTSA). Corporate Average Fuel Economy. http://www.nhtsa.dot.gov/cars/rules/cafe/index.htm. (accessed 5/31/05).
17. See pp. 6–7 of Geilecki and Poling above.
18. Federal Energy Regulatory Commission. *Order No. 888*. April 1996. See http://www.ferc.gov/legal/ferc-regs/land-docs/order888.asp
19. See Interstate Renewable Energy Council above.
20. U.S. EPA and U.S. DOE, http://www.energystar.gov/index.cfm?c=about.ab_index
21. U.S. DOE, Office of Energy Efficiency and Renewable Energy. See http://www.eere.energy.gov/consumerinfo/factsheets/ee8.html
22. U.S. EPA, Office of Transportation and Air Quality. See http://www.epa.gov/oms/gasoline.htm

23. Energy Information Administration. *Status and Impact of State MTBE Ban*. http://www.eia.doe.gov/oiaf/servicerpt/mtbeban/index.html. Updated 2003.
24. See for example, Bill H.R.6, as passed by the House of Representatives on 2005.
25. For example, see Power, N.J. Places Moratorium on Offshore Wind. Platts, 2005. At http://www.platts.com/Magazines/POWER/Power%20News/2005/012005_11.xml
26. For example, see Scanlin, Dennis Laws, Regulations, Permits, Policies, Positions and Wind Turbines. Presentation to the 2002 Wind Summit. 2002. See http://www.wind.appstate.edu/reports/laws&regs.ppt
27. Starrs, T. and Wenger, H. 1998. *Policies to Support a Distributed Energy System*. Renewable Energy Policy Project. See http://www.repp.org/repp_pubs/articles/index_pv.html
28. See Starrs and Wenger above.
29. U.S. DOE, Office of Energy Efficiency and Renewable Energy. *Strategic Plan*. October 2002. See http://www.eere.energy.gov/office_eere/pdfs/fy02_strategic_plan.pdf
30. U.S. DOE, Office of Energy Efficiency and Renewable Energy. See http://www.eere.energy.gov/wip/program/repi.html
31. See Interstate and Renewable Energy Council above.
32. U.S. DOE, Office of Energy Efficiency and Renewable Energy. http://www.eere.energy.gov/femp/about/about.cfm
33. See Interstate and Renewable Energy Council above.
34. U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy. See http://www.eere.energy.gov/greenpower/markets/index.shtml
35. Federal Energy Regulatory Comission. Staff Paper, *Assessing the State of Wind Energy in Wholesale Electricity Markets*. 2004. See http://www.ferc.gov/legal/ferc-regs/land-docs/11-04-wind-report.pdf
36. U.S. EPA. *Clean Air Interstate Rule and Clean Air Mercury Rule*. 2005. See http://www.epa.gov/cair/rule.html and http://www.epa.gov/air/mercuryrule/rule.htm
37. Energy Information Administration. *Renewable Energy Annual 1996*. Chapter 10. Growth of the U.S. Landfill Gas Industry. 1997. See http://www.eia.doe.gov/cneaf/solar.renewables/renewable.energy.annual/

2.2 International Policies for Renewable Energy

Michael Durstewitz

2.2.1 Introduction

The relevant key issues and condition which influence any individual country and its specific policy for promoting energy conservation and deployment of renewable energy technologies are determined by resources, targets, and constraints., the resources of renewable energies and their technical and economic exploitable potentials are one important issue. These conditions and the sources of renewable energies differ in a wide range between continents, countries and even on a regional scale inside individual countries. Even if the resources for a certain region or country are favorable, there often exist diverse constraints which can limit the use and the level of exploitation of the resources of renewable energies. The limitations can result from the constraints, such as missing infrastructure or lack of financial resources for projects. Policy targets are the most relevant issues and always reflect the general attitude of governments or policy makers towards the promotion of renewable or conventional energy systems. In general these key factors are identical in each regional, national or international context but their magnitude may vary significantly on a regional, national and international scale. There are many ways to support renewable energies with special policies depending on technology, resources and policy targets. After all, there exist too many to discuss all of them within the scope of this handbook. For a specific and detailed research on policies, the newly released "Global Renewable Energy Policies and Measures Database" [1] provides information for more than 100 countries worldwide with respect to policy types, technologies, and renewable energy targets. Examples for policy targets are defined in the Kyoto-protocol or in the "White Book" of the European Union as multinational agreements, as well as several national goals and action plans to increase the level of energy supply using renewable energies or to reach a certain level of renewable energy supply within a certain time scale.

2.2.2 Overview Policy Instruments

This chapter will give an overview on the present policies for the promotion of renewable energies for different regions in our world. The spectrum of available political instruments and measures which can be used for the introduction and promotion of any technology are principally identical. However, practically policies always depend on the stage of technology development, the availability of resources and their economic competitiveness. Generally, we can distinguish two main policy pathways. The first set of measures is often described as "technology push policies," the other guiding principle is known as "market pull" or "demand pull" measures. Technology push policies have to be considered as the support of fundamental research and development which aims to improve the technologies by means of reliability, performance, and cost reduction. The intention of market pull policies is to create new markets or support market expansion in order to increase the demand. Market pull measures policies can also lead to further cost reductions by increasing the demand which effects the market situation by stimulating the competition between manufacturers and also by economy of scale effects. Additionally, "other measures" and incentives have to be mentioned which are often initiated by local or state governments. They can be allocated to either technology push or market pull measures. These can also contribute to stimulate the demand, cost reduction, and employment on a regional scale according to the targets and possibilities of these regional and local players.

Typical technology push policies are programs in research and development, the support of research centers, and facilities for testing, etc. Policies to stimulate the market pull and the demand for the initially more expensive new technologies than traditional and conventional methods, are often investment subsidies, taxation driven incentives, loans, regulations for special feed-in tariffs, grid access, information campaigns, etc [2].

In 2002, the share of renewable energies to the global total electricity generation of 16,054 TWh was estimated to be 305 TWh or less than 2% of the total generation. The official declaration of these fuel types is "other fuels" which includes geothermal, solar, wind, combustible, renewable and waste energy. In comparison to data from three decades ago the generation of these "other fuels" has increased eightfold (1973, 36 TWh), while the total electricity generation has increased "only" by the factor 1.6 (1973, 6,011 TWh). Looking at world regions, countries, and technologies, we can observe different shares of application of renewable energy sources on a regional and a technological point of view. Hydro power and wind energy have the biggest share of penetration; especially wind power has shown two digit growing rates in the last decades. As a matter of fact, wind energy meanwhile is the most competitive alternative source to the traditional energy supply in many regions of the world. According to the paradigm, to "pick the low-hanging fruits first" policies for the promotion of renewable energies have to be seen in the context with the deployment of wind energy, at least in many European countries, especially in Denmark, Germany, and Spain. Many legal instruments presented here originate from the target to promote wind energy. However, the practical application of these policy tools is also used with other sources of renewable energies. Due to this reason, many policies measures and their effectiveness can be observed and evaluated quite well when looking at these technologies. The following analysis will focus on the development of renewable energies especially wind power in a European context.

2.2.3 Policies in Europe

The situation of energy supply in Europe is highly dependent on the import of energy products. Today more than 50% of its requirements have to be imported. If the present trends continue, this figure will increase to about 70% in 2030, with a growing dependency on oil and gas imports. In 1999, the cost of energy imports was more than €240 billion Euro for the European Union (EU-15), or about 6% of its total imports and 1.2% of its gross domestic product [3]. These figures show that the European Union has a vital interest in reducing its degree of dependence on energy imports, and on the other hand to improve the development and increase the share of domestic and sustainable energy resources. In comparison with the strategy of the United States to respond to the increasing demand of energy with an increasing supply, the strategic plan of the European Commission is to implement measures which promote the efficiency of energy use with respect to the reduction of energy demand without reducing the comfort of the customers. According to the point of view of the European Commission, the reduction of energy demand can get influenced more than increasing the supply. Besides these reasons, there are many more strategic motivations and benefits to support the expansion of renewable energies and energy conservation technologies. From an environmental point of view, the increased application of renewable energy sources, energy efficiency, and energy conservation reduces the effects of climate change by the reduction of greenhouse gases and pollutant emissions from conventional power plants. Looking from the economic point of view, the expansion of energy conservation and renewable energy technologies allows a sustainable development, creates knowledge, improves the job situation by creating local employment, and will lead to the development of new industry branches in Europe.

Directive #2001/77/EC of the European Parliament and the Council for the "promotion of electricity produced from renewable energy sources" [4] gives the legal framework for the exploitation of the renewable energy sources in the European Union in the internal market. The directive sets a community target of sourcing 12% of gross inland energy consumption from renewable energy systems by 2010, with an indicative figure of 22% for electricity. It also sets indicative targets for each member state. Furthermore, the directive obliges member states to report on the progress of meeting the national indicative targets, issue guarantees of origin of electricity produced from renewable energy sources, guarantee grid access for electricity produced from renewable energy sources, and assure transparent and nondiscriminating pricing for the grid connection of producers of electricity from renewable energy sources.

Although the European Commission has set general targets for all member states, it does not stipulate them to any mandatory and legally binding policies in order to reach the goals of increasing the share of renewable energy sources. In fact, the individual member states have their full sovereignty and are free to decide which policies they prefer as steering instruments of their choice. As a consequence, different policy mechanisms are applied on a national level in Europe. Based on reports of individual national governments, the report of the European Commission will evaluate the success and the cost-effectiveness of the diverse support systems with respect to specific national targets. If necessary, the Commission can make a proposal for one general and union-wide support policy for renewable energy sources. The Commission will present its evaluations in a report to the European Parliament with detailed information on the experience acquired with the application of the different mechanisms in its member states. The report is expected to be published by the end of 2005.

According to the subsidiary principle, the EU Directive does not require a community-wide framework for all the Member States. Thus, several support mechanisms and policies are applied in different EU member states in order to promote the development of renewable energies, especially in domestic markets. The hierarchy of national policies in context with the European and international policies is shown in Figure 2.1.

Other important directives and legal instruments that have been adopted by the European Parliament and Council in order to achieve the common goal of 12% share of renewable energies of EU-15 by 2010 and to support renewable energies and energy efficiency are, for instance, in chronological order: [5]

- Directive #2000/55/EC on energy efficiency requirements for ballasts for fluorescent lighting [6]
- Directive #2001/77/EC on the promotion of electricity produced from renewable energy sources [7]
- Regulation #2422/2001/EC on Energy Star labeling for office equipment [8]
- Directive #2002/31/EC on labeling of air conditioners [9]

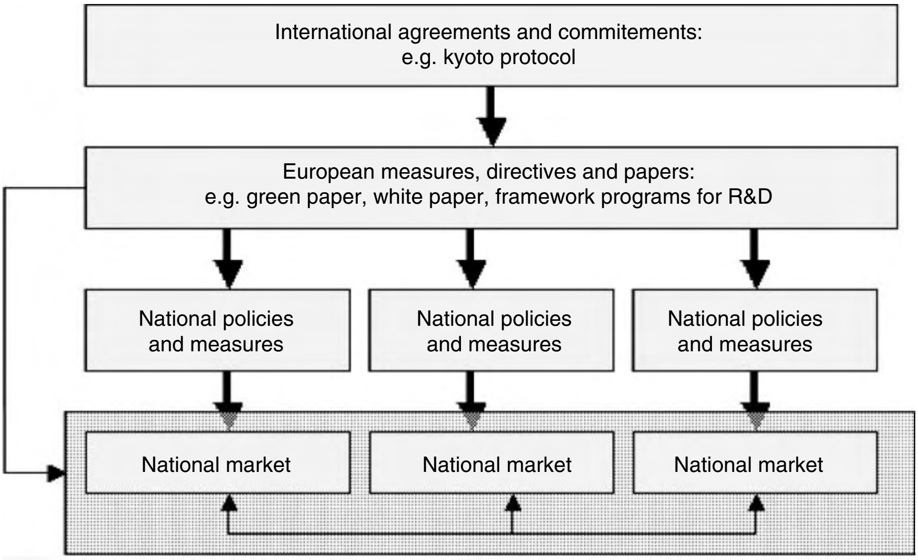


FIGURE 2.1 Hierarchy of national and European policies. (From Schaeffer, Alsema, Durstewitz, et al., Learing from the sun — Final report of the PHOTEX project, petten, NL, 2004.)

- Directive #2002/40/EC on labeling of electric ovens [10]
- Directive #2002/91/EC on energy performance of buildings [11]
- Directive #2003/30/EC on the promotion of biofuels [12]
- Directive #2003/66/EC on labeling of refrigerators [13]
- Directive #2003/96/EC for the taxation of energy products and electricity [14]
- Directive #2004/8/EC on the promotion of cogeneration [15]

With respect to the directive on the promotion of electricity by renewable energy sources, (2001/77/EC) the most common policy options are premium feed-in tariffs, quota systems, tendering, investment support, fiscal support, and green pricing. These mechanisms can be distinguished into policies which aim at the support of the supply and demand of renewable energy. The supply of renewable energy is stimulated by mechanisms supporting investments and production. The demand for the electricity from renewable energy systems is promoted by mechanisms that aim at consumption. The principles of these policy measures are particularly:

Feed-in tariffs typically support the supply side. Feed-intariffs guarantee a minimum price or an additional premium to market electricity prices per unit of electricity paid to the producers of electricity generated by renewable energy sources. This mechanism is usually accompanied by the obligation to grid operators to take and refund the electricity with a special tariff. The fixed price or fixed premiums are set for a number of years to maintain investor confidence. They can be revised and adjusted by governments to reflect long-term marginal generation costs and to stimulate further costs reduction. The costs are distributed to final customers. Feed-in tariffs are in use in several EU countries, notably in Germany, Spain, France, and Denmark.

Quota-based systems can support both, the supply and the demand side. On the supply side, quota-based systems are forced by regulations on producers and importers of electricity to produce a certain share i.e., a quota of electricity by renewable energy sources. In order to fulfill the quota obligation, producers can either install new capacity or buy an equivalent amount of electricity from renewable energy sources by means of tradable green certificates. Certificates provide an instrument for production monitoring, accounting, and transfer of electricity from renewable energy systems. In case of incompliance with the minimum quota settings, the producer will have to pay a penalty. This mechanism is presently applied in Italy. On the demand side, the quota system is applied to distributors, suppliers, and consumers in the electricity supply chain with the obligation to use a certain quota of electricity generated by renewable energy sources. This system is now introduced in the U.K., Belgium, and Sweden.

Investment support is the classical mechanism used to bring new technologies to the market by giving a certain amount of compensation of the capital cost for new installations. The mechanism has a number of advantages—it is simple, market compliant, easy adjustable to applied technology, capacity, as well as geographical and market conditions. Investment subsidies reduce the risk of operators in the investment with new technologies. However, the support of capital investment, as well as other capacity-based subsidies, is not necessarily advantageous for the cost-effective operation of renewable energy installations.

Fiscal support for renewable energy systems can be applied in different forms as direct or nondirect measures in order to stimulate the supply. An example for fiscal support measures is the exemption from energy taxes for renewable energy systems. A production-dependent tax relief per kilowatt-hour electricity stimulates the supply from new as well as old installations. Other support options, such as reduced VAT rates can be similar as investment subsidies. They affect only investments into new installations. On the demand side, energy taxes and the taxation of CO_2 emissions increase the price of conventional electricity, which indirectly favor and stimulate the demand for renewable energies because they reduce the cost shear between conventional and renewable production. Since 2002, the Dutch taxation system on energy was combined with a stimulation of renewable energy consumption. This measure has allowed prices from renewable electricity to come close or match those of conventional electricity in the private sector. As a result, a large demand for energy from renewable energies has been created, but due to the uncertainty of continuity of the instrument and international competition, it did not lead to a substantial increase in installed capacity.

Tendering systems involve a competition of proposals of potential investors who compete on a bid price per kilowatt or megawatt of electricity generation. Successful bidders earn long-term power purchase agreements. Additional costs of power purchase from renewable energy projects are financed by a compulsory levy to be paid from all customers. The tendering system mechanism is favored to be one of the most cost-effective options; on the other side, the necessary administrative overhead costs tend to be rather high as well. This mechanism is in use, in the Republic of Ireland, however, it was announced recently that Ireland will replace its support mechanism of competitive bidding for government contracts with a fixed price support mechanism [16].

Green pricing or Green electricity is an instrument based on the voluntary willingness to pay additional charges on the electricity bill of private individuals and commercial companies. Today "green tariffs" are the most common voluntary instrument to promote electricity from renewable energy sources. It mainly attracts private individuals with its opportunity to sponsor the further development of electricity from renewable energy sources by their willingness to pay additional charges per kilowatt to the regular tariff rate. However, till now, not many customers choose this support option. Table 2.3 gives an overview of the presently applied mechanisms in Europe (marked with (✓), pointing out dominant ones with an additional exclamation mark (✓!)).

2.2.4 The Case of Germany

Similar to many other European countries, the energy situation in Germany is dominated by the lack of significant fossil resources, oil and gas. The exploitation of domestic mineral coal is decreasing continuously in recent decades and mines which are still in operation receive billions of Euros fiscal support in order to survive against low priced coal on the world market. In some regions in Germany, lignite is exploited, mostly used for the firing of thermal power plants. The geographic and climatic conditions in Germany for the use of renewable energy sources cannot be considered to be optimal. In comparison with other European regions, U.K., Ireland, France, and Denmark, the wind potential of

TABLE 2.3 Scheme of Support Mechanisms for Renewable Energies in the EU

Country	Supply Support					Demand Support		
	FIT	Quota	Investment Support	Fiscal Support	Tendering	Quota	Fiscal Support	Green Pricing
Austria	✓!			✓				✓
Belgium	✓			✓		✓!		
Denmark	✓!		✓	✓		✓		
Finland	✓!		✓	✓				✓
France	✓!				✓			
Germany	✓!			✓				✓
Greece	✓!							
Ireland	✓				✓!			
Italy	✓	✓!	✓					
Luxembourg	✓!		✓	✓				
Netherlands	✓		✓				✓!	✓
Portugal	✓!		✓					
Spain	✓!							
Sweden				✓		✓!		
Turkey	✓!							
U.K.			✓	✓		✓!	✓	✓

Source: From Badelin et al. 2004. Current experience with supporting wind power in European electricity markets, Global Windpower Conference 2004, Chicago.

Germany is only moderate, except for sites close to Germany's coastlines. Insulation conditions for solar applications are also not ideal due to the unsettled climatologic conditions in contrast with countries, in the Mediterranean region. However, the goal of the German government to double the share of energy supply from renewable energy sources by the year 2010 in comparison with figures from 2000 are quite ambitious; this means that at least 4.2% of Germany's primary energy consumption and 12.5% of the electricity supply shall be produced by renewable energy sources. In the medium and long term, the targets are to achieve a share of 20% by 2020 and 50% by 2050 of the primary energy consumption. The results reached by now are very promising. Germany counts the world's highest number of wind turbine installations. In 2004, some 16,000 MW wind capacity fed 26 TWh or approximately 6% of Germany's total electricity supply into the public grids. The cumulated capacity of PV installations surpassed the 700 MW_p level. The market for solar thermal absorbers is by far the largest in Europe; more than 6 million square meters equivalent to 4,400 MW are mounted to the roofs of German homes. Biomass in solid, liquid, or gaseous form is used for many applications, domestic heating, cogeneration, and transportation; nowhere else more biodiesel is sold to fuel cars and farming vehicles. In total, 5.1% or 131 TWh of Germany's end energy supply for electricity, heat, and transportation was provided with renewable energy sources. The structure of different energy sources for the supply with renewable energies is shown in Figure 2.2.

These promising achievements have been possible only with the assistance and support of policies, guidelines, and regulations in favor of renewable energy and energy conservation. Looking back three decades in history, the activities in the field of research in renewable energies and energy conservation came into focus again after the first oil crisis back in 1973. Several and continuous measures for technology push and market pull activities were initiated by the German Federal and State governments in order to pave the pathway towards a new and sustainable energy future. Some milestones in German energy policy measures for the promotion of renewable energies have been from the 1970s until 1998.

- The demonstration program "250 MW Wind" (1989…ongoing) to gain practical and long-term experience with a large number of wind turbines by different manufacturers,

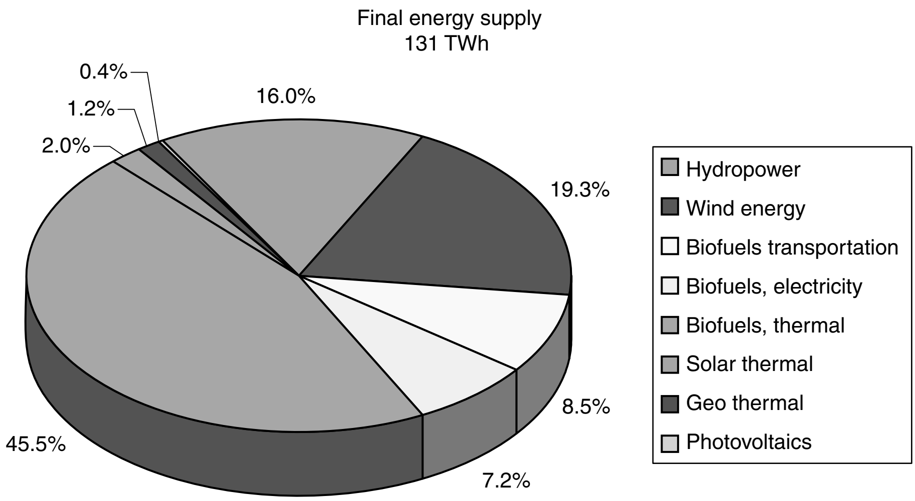


FIGURE 2.2 Distribution of renewable energy sources to end energy supply.

technologies and at various site conditions in Germany. Most of the participating operators received their annual support by production (per produced kilowatt) instead of an initial investment subsidy. This has proved to be a very simple but efficient regulation to motivate operators to keep their turbines operating in a good condition.

- Special Loans with reduced interest rates by the German State Bank who put up nearly 3 billion Euros between 1990 and 1998 for financing renewable energy projects.
- The Electricity Feed Law (EFL, 1991–2000) which guaranteed grid access and a fixed premium tariff for electricity produced by renewable energies which is fed into the public grid. The EFL in combination with the "250 MW Wind" program and the special loans can be seen as the main drivers for the boom of wind energy applications in Germany.
- More than 1 billion Euros of support for energy research programs of the Federal government between 1990 and 1998.
- The 100,000-roofs-PV-program which boosted the installation of more than 2000 PV systems up to 5 kW peak power.
- The Federal market stimulation program for the promotion of renewable energies.
- Subsidies for solar heating and heat pumps for private house owners.
- Additional measures for education, information, and dissemination.
- Support to the European Union for research activities and the development of renewable energies and energy conservation in a European context.
- Further support of about 1 billion euro was granted by state and regional governments, communities and utilities.

In autumn 1998, the conservative government was replaced by a government led by the "Red" and "Green" parties. One of their targets and major projects is the implementation of a new era of energy policy. The Federal government points out their target to reduce the CO_2 emissions by 25% in comparison with 1990 until the year 2005. The main aims are to strengthen renewable energies for a sustainable energy supply, to reduce the demand of energy and to back out of the nuclear energy program. The achievements that have been reached since 1991 and the savings which will be due in the forthcoming years are shown in Figure 2.3.

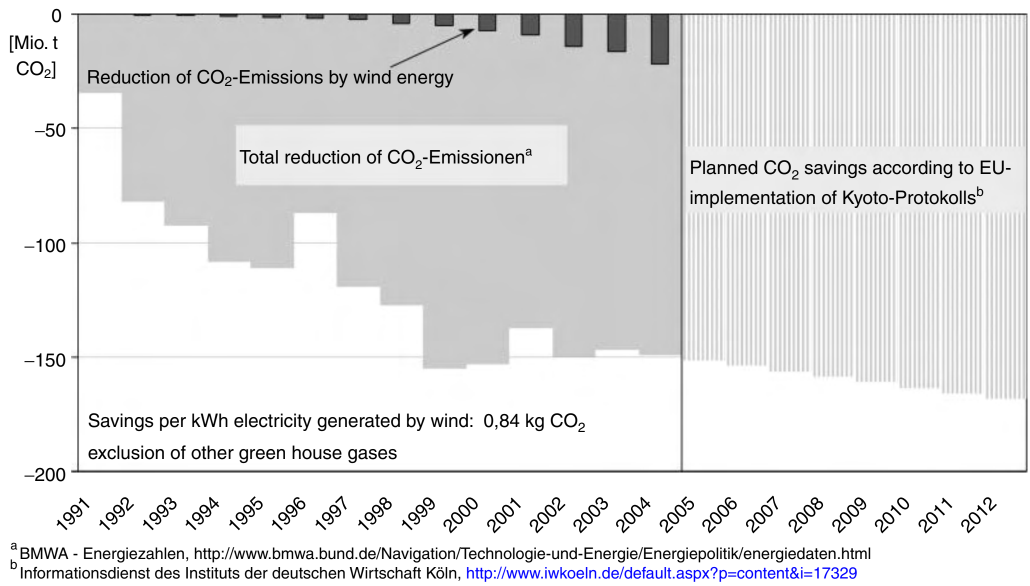


FIGURE 2.3 Total and share of wind energy to reduction of CO_2 emissions in Germany. (From Durstewitz, et al., Wind Energy Report Germany, 2005, ISET, Germany, 2005.)

The main policy measures which should pave the way towards a sustainable energy future were:

- The adoption of the "Renewable Energy Act" (REA). This law intends to guarantee fair market chances for renewable energies and remove obstacles and bottlenecks which hinder an intensive use of renewable energies. A revised version of the REA came into force in August 2004.
- The establishment of an "energy consensus" with the energy industry to search for a broadly accepted way towards a sustainable energy mix without nuclear energy.

An intensified promotion for the production and market introduction of renewable primary energy (market stimulation program). This subsidy program supports the installation of technology for the generation of heat and/or electricity. Depending on the kind of technology investment subsidies or special loan programs with options for partial abatement of debt can be granted. Since the start of the Market Stimulation Program in 1999, more than 400,000 projects were funded with more than 500 million Euros. These subsidies or loans have triggered a total investment volume of more than 4 billion Euros.

The implementation of a 100,000-roofs-PV-program (1999–2003) aimed to increase the number of PV installations in Germany. In 1999, about 50 MW_p capacity was online in Germany, until 2003 additional 100,000 roofs or 300 MW_p should be equipped with photovoltaic systems, most of them on rooftops of private houses. The program was very efficient and contributed significantly to market penetration and cost reduction of PV systems. One stimulus for the private investments of total 2.4 billion Euros was the very convenient conditions for the loans, 1.9% annual interest, term 10 years, 2 years free of redemption, and possibilities for a partial abatement of the debt).

Loans with favorable interest rate for investments in environmental technologies from the German state bank are preferred by investors for financing their projects. The KfW Förderbank, a 2003 merger of Deutsche Ausgleichsbank (DtA) and Kreditanstalt für Wiederaufbau (KfW) supports financing of investments into renewable and energy-saving technologies with different programs since many years. From 1991 until 2003 KfW-Förderbank respectively its predecessors put on the following support measures and gave loan agreements for investments into renewable energies and energy-saving projects in a total of 11.1 billion Euros.:KfW-CO_2 reduction program (1997–2003, 422.3 million Euros)

KfW-CO_2 building redevelopment program (2001–2003, 20.5 million Euros)
DtA-ERP-environment and energy-saving program (1991–2003, 6,942 million Euros)
DtA-environment program (1991–2003, 3,679 million Euros)

The backbone of the loan programs is the DtA-ERP-environment and energy-saving program, where ERP stands for "European Recovery Program." As a matter of fact, this source of financing was originally granted to the German Federal Republic after the Second World War by means of the Marshall-Plan (1948). This fund was primarily used to rebuild the industry and economy in Western Germany; however, the capital is still circulating and partially used to finance renewably and energy saving projects.

References

1. The Global Renewable Energy Policies and Measures Database, initiative led by the International Energy Agency(IEA). http://renewables.iea.org/, visited November 2005.
2. Neij, Andersen, Durstewitz. 2003. Experience Curves: A Tool for Energy Policy Assessment. Final Report of the EXTOOL project, Lund, SE.
3. European Commission: Green Paper towards a European strategy for the security of energy supply.
4. #2001/77/EC.
5. Commission Of The European Communities. 2004. The share of renewable energy in the EU, COM(2004) 366 final, Brussels, 26.5.2004.
6. http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do?uri=CELEX:32000-L0055:EN:HTML. Visited August 2005.
7. http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do?uri=OJ:L:2001:283:0033:0040:EN:PDF. Visited August 2005.
8. http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do?uri=OJ:L:2001:332:0001:0006:EN:PDF. Visited August 2005.
9. http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do?uri=OJ:L:2002:086:0026:0041:EN:PDF. Visited August 2005.
10. http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do?uri=OJ:L:2002:128:0045:0056:EN:PDF. Visited August 2005.
11. http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do?uri=OJ:L:2003:001:0065:0071:EN:PDF. Visited August 2005.
12. http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do?uri=OJ:L:2003:123:0042:0046:EN:PDF. Visited August 2005.
13. http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do?uri=OJ:L:2003:170:0010:0014:EN:PDF. Visited August 2005..
14. http://europa.eu.int/eur-lex/lex/LexUriServ/LexUriServ.do?uri=OJ:L:2003:283:0051:0070:EN:PDF. Visited August 2005.
15. http://europa.eu.int/eur-lex/pri/en/oj/dat/2004/l_052/l_05220040221en00500060.pdf. Visited August 2005.
16. Massy. 2005. Ireland goes for fixed purchase prices. Wind Power Monthly, Vol 21.

2.3 Energy Policies in India

Anil Misra

2.3.1 Energy Conservation and Renewable Energy in India

Energy conservation received greater attention in India since the mid 1970s. Structural changes in the economy during the last few years have led to the expansion of the industrial base and infrastructure in the country, and subsequently to increase in demand for energy. Any effort to enhance energy generation brings issues of available energy sources and systems. India recognizes, as anywhere else, a need to reduce the dependence on fossil fuels and transition to an era where many cost-effective and efficient energy choices are available. There has been a vigorous search during the last three decades for alternatives to fossil fuels that would ensure energy security and eco-friendly sustainable development.

This has brought the focus on new and renewable sources of energy that have emerged as a viable option to supplement energy demand in various sectors. The Government of India has been making concerted efforts to develop and promote renewable energy applications, ranging from power generation to meeting cooking energy needs of the people through biogas plants.

The energy conservation efforts in India have to be viewed in terms of coal and lignite being the long-term sustainable local energy resources. Small resources of petroleum and natural gas may be exhausted shortly and in the medium and long term, the import of oil will increase. We look briefly at the energy scene first.

The Indian energy sector has been regulated and owned by Government agencies and organizations. The apex body for the power sector is the MoP (Ministry of Power) in the Government of India. The ministry is concerned with perspective planning, policy formulation, processing of projects for investment decision, monitoring of the implementation of power projects, training and manpower development, and the administration and enactment of legislation in regard to thermal and hydro power generation, transmission, and distribution. The Ministry is assisted by the Central Electricity Authority (CEA) in all technical and economic matters, and many Public Sector Undertakings, Statutory Corporations, and Autonomous Bodies are functioning under the Ministry.

Energy consumption by different economic sectors—agriculture, industry, transport, and residential—reveals the composition of fuels used in meeting energy needs in these sectors (Figure 2.4) [1].

The industrial sector continues to be the single largest commercial energy-consuming sector using up about 50% of the total commercial energy in the country, although its share is declining gradually.

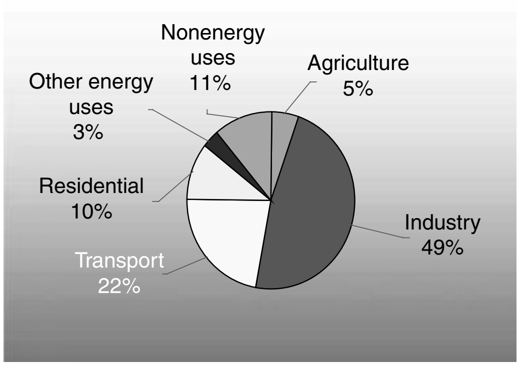


FIGURE 2.4 Commercial energy balance. (From Annual report of the Ministry of Nonconventional Energy Sources, Government of India, www.mnes.nic.in).

The commercial energy intensity of this sector has declined over the past due largely to a relatively rapid expansion of nonenergy intensive industries, adoption of modern energy efficient technologies and successful implementation of energy conservation measures. Sufficient energy savings are being achieved in energy intensive aluminum, iron and steel, textiles, chemicals, and paper and pulp industries through better housekeeping, improved capacity utilization, development of cogeneration facilities, industrial heat and waste management, and arrangements for improving the quality of electric supply.

Transport infrastructure has expanded considerably and its energy intensity has grown gradually. Rapid urbanization along with the conglomeration of industrial and commercial activities has consequently increased the demand for transport. Uncontrolled expansion of cities coupled with inadequate public transport has contributed to a phenomenal growth in the number of private enterprises (providing transport facilities) leading to energy inefficiency and severe pollution problem. The strategies to check vehicular pollution, lately in a couple of metros, are now showing significantly positive results. Strategies for pollution control in other sectors are also being worked out.

The domestic sector is the largest consumer of energy in India accounting for 40–50% of the total energy consumption, but the bulk of it consists of traditional fuels in the rural household. The commercial sector is not a major energy consumer, but energy intensity has been increasing possibly due to the growth of this sector. The electricity consumption in various sectors (only utilities) shown in Figure 2.5 indicates highest consumption of electricity in industrial and agricultural sectors.

The Government of India has set up an agenda of electrifying all villages by 2007 and providing power to all by 2012. It is following a comprehensive and holistic approach to power sector development envisaging a six level intervention strategy at the National, State, Electricity Boards, Feeder, and Consumer levels. It is proposed to electrify all the unelectrified villages through grid connectivity and the remaining 25,000 remote villages through the use of renewable energy technologies. The importance of optimum and economic utilization of power has been realized lately, and the present capacity of 8,000 MW for interregional transfer is being enhanced to 23,500 MW by the end of 2007 [2].

The new electricity bill of 2003 (Electricity Act 2003) has paved the way for delicensing of generation, nondiscriminatory open access in transmission, power trading, rural electrification, mandatory requirements of State Electricity Regulatory Commissions, mandatory metering, and stringent provisions against theft of power and so on. The trading activities of the Power Trading Corporation have already commenced from surplus to deficit regions.

In the process of tariff setting, transparency has been brought into tariff fixation, which was traditionally a closed-door exercise. The State Electricity Regulatory Commissions have instituted

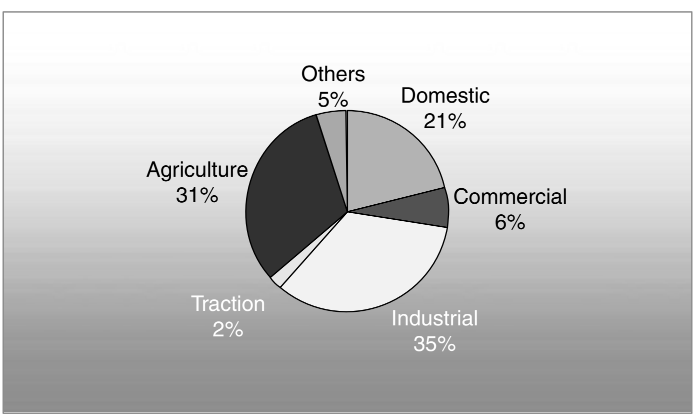


FIGURE 2.5 Electricity consumption in various sectors (utilities). (From Annual reports of the Planning Commission, India, www.planningcommission.nic.in).

measures to allocate revenue requirement in an economically efficient manner by reducing the extent of cross-subsidies. This has primarily been achieved by increasing the low-tension tariff[1] to a greater extent as compared to high-tension tariff[2]. The increase in the agricultural tariff is a bold measure initiated by a number of state commissions. The commissions have issued suitable directions to the utilities along with the tariff orders in response to various issues such as quality of supply, energy auditing, and transformer failure that were faced during the process of tariff setting.

2.3.2 Energy Conservation

Considering the vast potential of energy savings and benefits of energy efficiency, the Government of India enacted the Energy Conservation Act. The Act provides legal framework, institutional arrangement, and a regulatory mechanism at the Central and State levels to embark upon energy efficiency drive in the country. The Energy Conservation Act of 2001 facilitated the creation of the Bureau of Energy Efficiency (BEE) that has served to appropriately elevate the importance of energy efficiency issues.

The strategy developed to make power available to all by 2012, includes promotion of energy efficiency and its conservation in the country, which is found to be the least cost option to augment the gap between demand and supply. Nearly 25,000 MW capacity creation through energy efficiency in the electricity sector alone has been estimated in India. Energy conservation potential for the economy as a whole has been assessed as 23% with maximum potential in industrial and agricultural sectors.

BEE has successfully launched a complete pilot phase of program for energy efficiency in government buildings and prepared an action plan for its wider dissemination and implementation. BEE would train core group members in various departments to build capacities enabling them take up energy efficiency programs. To give national recognition through awards to industrial units for the efforts undertaken by them to reduce energy consumption in their respective units, the Ministry of Power launched the National Energy Conservation Awards in 1991. BEE provides technical and administrative support for the Awards Scheme.

The thrust areas under the long-term measures, include industry-specific task forces, notifying more industries as designated consumers, conduct of energy audit amongst notified designated consumers, recording and publication of best practices in each sector, development of energy consumption norms, and monitoring of compliance with mandated provision by designated consumers. A few examples characterizing the new initiatives are given below.

Standards and labeling program, identified as one of the key activities for energy-efficiency improvements, is expected to ensure the availability of only energy-efficient equipment and appliances. The equipment to be initially covered under S&L program are household refrigerators, air-conditioners, water heater, electric motors, agriculture pump sets, electric lamps and fixtures, industrial fans and blowers, and air-compressors.

Demand side management (DSM) together with increased electricity end use efficiency is expected to mitigate power shortages to a certain extent and drastically reduce capital needs for power capacity expansion. BEE is assisting five electric utilities to set up DSM Cell and has undertaken the preparation of investment grade feasibility reports on agricultural DSM, municipal water pumping, and domestic lighting.

Energy Conservation and Commercialization project (ECO) with USAID (U.S. Agency for International Development) aims to design and test DSM projects in the states of Karnataka and Maharashtra. It is promoting market-oriented solutions to energy sector problems, particularly the need for increased energy efficiency, assisting in creating a policy, market and financing environment in India that would advance commercialization of energy efficiency technologies and services. If further

[1]Electricity supply at 220 or 440 volts.

[2]Electricity supply at 11,000 volts or above.

endeavours to implementing pilot scale projects to test and validate DSM/EE design, planning approaches, and technologies in the Indian context.

Energy conservation building codes are under development for new commercial buildings. The Bureau has begun conducting energy audits in many prestigious state buildings in Delhi (such as Rashtrapathi Bhawan, Parliament House, South Block, North Block, Shram Shakti Bhawan, All India Institute of Medical Sciences, Safdarjung Hospital, Indira Gandhi International Airport, Sanchar Bhawan, and Rail Bhawan).

Professional certification and accreditation procedures have been established by BEE, and relevant curriculum is being introduced in the school education programs.

A number of Pilot Projects/Demonstration Projects have been taken up for load management and energy conservation through reduction of T&D losses in the System such as Distribution network of West Bengal State Electricity Board; Distribution network of Kerala State Electricity Board; Installation of 2,370 LT Switched Capacitors in Andhra Pradesh, Haryana, Punjab and Tamil Nadu; Installation of 3,000 Amorphous Core Transformers in the distribution network of various State Electricity Boards.

The BESCOM Efficient Lighting Program is the first large-scale efficient lighting project in India and South Asia. An unprecedented partnership of stakeholders—lighting vendors, consumers, utilities, retailers, financial institutions, energy efficiency service providers—have agreed to collaborate to demonstrate the technical and commercial viability of efficient lighting systems.

2.3.3 Renewable Energy Use in India

In the light of global developments, India has taken the decisive policy steps to move ahead and tap the immense potential for renewable energy (RE) sources such as solar, wind, biomass, small hydro, etc., and build the necessary skills and manpower to favourably use these resources.

The Ministry for Non-Conventional Energy Sources (MNES) is the nodal ministry for development and promotion of RE. The Ministry has nine Regional Offices, three specialized Technical Institutions, and one Financing Agency under it, which function to promote the policy and program initiatives. The States in turn have been advised to initiate conducive policies for commercial development of the RE sector.

Renewable energy devices and systems have become increasingly more visible during the last two decades, and power generation from renewable sources is also increasing. The estimated potential and the extent of exploitation so far are given in Table 2.4, and efforts have been stepped up to achieve the full potential of the use of renewable energy sources. The medium term goal is to ensure electrification of 25,000 remote and unelectrified villages, and achieves a minimum 10% share or 10,000 MW (of the estimated 100,000 MW), from renewable energy in the power generation capacity by the year 2012 [3].

Major programs in India facilitated by MNES for power generation include wind, biomass (cogeneration and gasifiers), small hydro, solar, and energy from wastes. Cumulative installed capacity

TABLE 2.4 Potential, Achievements and Targets for RE Use in India (as on 31 December 2005)

S. No.	Resource/System	Potential	Achievements
1	Solar water heating	140 mil m^2 Collector Area	1.0 mil m^2 Collector Area
2	Solar PV	20 MW[a] per Sq. km	191 MW[a]
3	Wind energy	45,000 MW	2,980 MW
4	Small hydro	15,000 MW	1,693 MW
5	Waste-to-Energy	1,700 MW	46.5 MW
6	Biomass gasification	—	62 MW
7	Biomass power/ cogeneration	19,500 MW	727 MW
8	Biogas plants	12.0 mil	3.67 mil
9	Improved cook stoves	120.0 mil	33.9 mil

[a] Of this 105 MW_p SPV products have been exported.

of RE power sources presently totals to 5,700 MW, representing about 5% of the total installed capacity in India. The Ministry supports research and development with close involvement of the industrial sector. It is hoped that there will be increased interaction and close cooperation between the research and teaching institutions of the country—which are reservoirs of knowledge and experience, and the Indian industry, which has the requisite entrepreneurship and market-orientation.

A large domestic manufacturing base has been established for renewable energy systems and products. The annual turnover of the renewable energy industry in the country, including the power generating technologies for Wind and other sources, has reached a level of over U.S. $ 640 million. Important achievements include cumulative capacity of solar PV modules production has reached 191 MWp, annual production of wind turbine industry is 1,000 MW. India is the third largest producer of silicon solar cells, and has the fifth largest wind power capacity in the world. In SPV technology, 8 companies make solar cells, 14 companies make PV modules, and 45 companies make variety of PV systems. There are 30 manufacturers of box-type solar cookers, and a few of them make indoor community cookers and solar steam cooking systems. There are 83 approved manufacturers of solar flat plate collectors for water heating, and 10 companies make equipment for small hydro projects.

2.3.4 Policy for All-Round Development of Renewable Energy

Fiscal incentives are being offered to increase the viability of RE projects; the main incentive is 60% accelerated depreciation. Other incentives include a tax holiday, lower customs duty, sales tax, and excise tax exemption for RE projects. The Indian Renewable Energy Development Agency is the main financing institution for renewable energy projects. It offers financing of the renewable projects with lower interest rates, which vary with the technology, depending on its commercial viability. Though interest rates are falling in India, they are not in the renewables sector for various reasons, but mainly due to perceived high risk. The interest rates vary from 11 (for biomass cogeneration) to 14.5% (for wind). The Clean Development Mechanism of the Kyoto Protocol has started offering additional stream of revenue making renewable energy projects more attractive.

Policy measures aim at overall development and promotion of RE technologies and applications. Policy initiatives encourage private as well as foreign direct investment including provision of fiscal and financial incentives for a wide range of RE programs.

Foreign investors can enter into a joint venture with an Indian partner for financial and/or technical collaboration and setting up of RE-based power generation projects. Proposals up to 100% foreign equity participation in a joint venture are automatically approved under the RE policy. The Government of India also encourages foreign investors to set up RE-based power generation projects on BOO (Build, Operate, Own) basis. Various Chambers of Commerce and industry associations in India provide guidance to the investors in finding appropriate partners, and foreign investors can set up a liaison office in India

MNES is promoting medium-, small-, mini- and micro enterprises for manufacturing and servicing of various types of RE systems and devices, and industrial clearances are not required for setting-up of an RE industry. CEA permits power generation projects up to IR 1,000 million (~ U.S. $ 22.7 million). Other promotional incentives include: A 5-year tax holiday for RE power generation projects, Private sector companies allowed to set up enterprises to operate as licensee or generating companies, and Customs duty concession offered to RE spares and equipment, including those for machinery required for renovation and modernization of power plants. Excise duty on a number of capital goods and instruments in the RE sector has been reduced or exempted.

A number of states in India have announced policy packages including wheeling, banking, third-party sale, and buyback, and a few of them are providing concessions or exemption in state sales tax and octroi. These rates vary from state to state for different technologies and devices, and in periodicity. So far, 14 states have announced policies for purchase, wheeling and banking of electrical energy generated from various RE sources.

References

1. TEDDY 2002/03: TERI Energy Data Directory & Yearbook. The Energy and Resource Institute (TERI), Darbari Seth Block, India Habitat Centre Complex, Lodhi Road, New Delhi 110003 (www.teriin.org).
2. Annual reports of the Ministry of Power (www.powermin.nic.in/report).
3. Annual report of the Ministry of Non-conventional Energy Sources, Government of India (www.mnes.nic.in).
4. Schaeffer, Alsema, Durstewitz et al. 2004. Learning from the sun—Final report of the PHOTEX project, Petten, NL.
5. Durstewitz et al. 2005. Wind Energy Report Germany 2005, ISET, Germany.
6. Annual reports of the Planning Commission, India(www.planningcommission.nic.in).
7. Badelin et al. 2004. Current experience with supporting wind power in European electricity markets, Global Windpower Conference 2004, Chicago.

2.4 Renewable Energy Policies in Israel

Gershon Grossman

2.4.1 Introduction

Israel is a country of 6.3 million inhabitants in an area about 20,000 square kilometers located on the eastern Mediterranean coast. Israel's electrical grid and overall energy economy are isolated from those of its neighboring countries. The precarious political situation has led in the past to the country being subjected to energy and other economic boycotts. In the absence of indigenous fuel resources, Israel's economy depends almost entirely on imported coal and petroleum. Despite the political difficulties, Israel's economy is the fastest growing in the Middle East, leading to an ever-increasing demand for energy. Israel's total energy consumption has increased by about 50% in the last 10 years, and electric power production and consumption have doubled in that period of time [1].

Israel has learnt to use the one abundant and inexhaustible energy resource it has—solar energy. Israel has been a pioneer in developing solar technology, and is leading the world in per capita utilization of solar energy. According to the Ministry of National Infrastructures (MNI) Division of Energy Conservation [2], about 85% of households (i.e.,about 1,650,000 households) use domestic solar water heating, accounting for 3% of the country's primary energy consumption. Government encouragement and regulations, along with over 50 years of proven experience, have made this application widespread. Other forms of solar energy utilization are considered, with the potential of supplying a much larger portion of the country's energy demand.

Israel faces a number of choices in securing reliable, clean, and efficient sources of energy over the next 25 years. Recent discoveries of offshore natural gas have generated interest in developing this resource for domestic consumption to reduce dependence on imported coal and petroleum, which involves both economic and environmental risks. According to a MNI-commissioned forecasting study on Israel's energy demand [3], the overall merits of developing these gas reserves entail a number of trade-offs involving energy security, environment, and cost. Development of renewable sources is free of these concerns and holds a great promise to supply the country with long-term clean and secure energy.

2.4.2 Israel's Energy Economy

The Israeli economy is largely fueled by imported coal and petroleum. Prior to 1982, only a small amount of coal was consumed and the country's primary energy source was oil and its derivatives. At present, coal accounts for about 30% of Israel's primary energy. It is almost entirely used in electric power generation.

Central Bureau of Statistics (CBS) data [1] show that the total primary energy supply to the country for the year 2004 was about 21 million tons oil equivalent (MTOE)—a 50% increase from 1994. Total end use was about 12 MTOE, consisting roughly of 4 MTOE petroleum products and 8 MTOE electric power. Among petroleum products, gas oil and diesel oil comprise the largest category. Electric consumption is about 1/3 for residential and 2/3 for industrial use (goods and services), the latter comprising a significant portion for agriculture and water pumping. Fuel consumption other than for electric power generation is split at 60, 36, and 4% among the transportation, industrial and residential sectors, respectively. The same consumption pattern has been in effect for the last 10 years. According to forecast [3], it is probable to change in the coming years, with the introduction of natural gas, which will replace heavy oil in the electric power and industrial sectors.

Another component of the energy economy—relatively small in quantity but important in significance—is solar energy, utilized mainly for domestic water heating. As mentioned earlier, it accounts for about 3% of the country's energy demand and saves about 5% of the electric power. Israel is presently the largest per capita user of solar energy in the world, challenging the notion that solar energy is not yet economical.

The Government MNI's stated policy and objectives with regard to Energy are [2]:

- The Energy Policy will guarantee, at any time in the future, the complete supply of energy demand in the country's economy. The price to the consumer will reflect and include all components of the cost for energy use: economic—direct/monetary, environmental, strategic, and that associated with land use.
- The Energy Policy will strive for maximum market competitiveness, reducing the Government's involvement to sectors where market failure exists or where involvement is necessary for security and safety considerations.
- The Energy Policy will endeavor to improve energy efficiency in electric power generation, transportation, and the residential, commercial/public and industrial sectors.
- The Energy Policy will strive to reduce environmental damage resulting from energy production and use. In particular, fuel transmission and storage infrastructures will be controlled to prevent leaks into the ground, sea or ground water, in accordance with international standards, and avoid health hazards or damage to natural resources. Greenhouse gas emissions will be controlled to a level not to exceed, in the long term, similar per capita emissions in OECD countries.
- The Energy infrastructure in Israel will be based on intelligent use of land resources, and exploit land designated for energy infrastructures for future expansion. In particular, sites with existing generation, transmission and/or storage infrastructures are to be utilized to their full operational and expansion capacity, with due consideration to strategic, safety and security issues; technological, physical, and political opportunities are to be explored to the fullest toward utilizing external land resources such as artificial islands. The Policy calls for cooperation with neighboring countries in commercial joint exploitation of land resources.

A number of measures have been initiated to promote renewable energies in electric power generation. A Government resolution dated November 2002 set an objective to produce at least 2% of the total electric power from the renewables by the year 2007. This amount is to increase by 1% every 3 years up to 5% of total electricity production by 2016. Accordingly, the Public Utilities Authority has instituted a set of premiums for licensed electricity producers employing renewables, proportional to the achieved reduction in pollution. Another Government resolution dated March 2002 promotes cogeneration, with the national potential estimated at 1200 MW. Energy conservation by the public is encouraged by market transformation—increasing public awareness of energy efficiency in appliances and the like.

A number of laws, regulations, directives, and legal instruments have been introduced to implement the Energy policy. Most of them regulate the use of petroleum and petroleum products, natural gas and electric power, and deal with licensing and safety issues. Some deal with energy efficiency, minimum standards and labeling. Recent regulations encourage cogeneration and distributed generation. Noteworthy is Article 9 of the Law for Planning and Building (1970), mandating the installation of solar water heating systems in all new constructions [4]. This Article has contributed a great deal to the advancement of solar energy utilization and will be discussed further in the next chapter.

2.4.3 Solar Energy Utilization in Israel

A visitor to Israel will unavoidably notice the urban landscape bursting with solar collectors and hot water storage tanks covering the roofs of buildings. Almost all residences in Israel are equipped with solar water heaters. The most common are the thermosyphonic system, a completely passive, stand-alone unit consisting of one or two flat plate solar collectors and an insulated storage tank. Large multi-story apartment buildings often use a central system with a collector array on the roof and a storage tank in the basement, employing a pump controlled by a differential thermostat. Other arrangements are also available. In most of the country, the solar system will supply the full demand for hot water during 9–10 months per year, with an electric resistance backup employed the rest of the time. Freeze protection is never required, except in some isolated locations. The economics: the installed cost of a typical single-family system comprising a 150 l storage tank and 2–3 m^2 flat plate collectors is about \$700; an equivalent electric-powered system costs about \$300. The difference of \$400 is recovered by the owner in about 4 years (on a simple-payback basis); these systems carry a manufacturer's warrantee for 6–8 years, and if properly maintained can last over 12 years. Several decades of nation-wide experience have generated consumer confidence and acceptance to the point that a domestic solar water heater is perceived as a common, reliable household appliance.

There is no single legislation concerning solar energy utilization in Israel. The above-mentioned Article 9 of the Law for Planning and Building (1970) [4] is probably the most important solar legislation, and has been the government's predominant contribution to Israel's success in the solar area. The law requires the builder (not the homeowner!), since 1980, to install a solar water heating system in every new building. Other laws and regulations describe in detail the size of the installation required for the various types of buildings, set minimum standards for the quality of the solar equipment and installation, and provide the regulations for retrofit installation of solar water heaters in existing multi-apartment buildings. Based on government data [5] an average single-family domestic solar water heater saves 1250 kWh electric power per year; the total contribution to the country is about 1.6 billion kWh per year, 21% of the electricity for the domestic sector or 5.2% of the national electricity consumption, providing for 3% savings on the primary energy consumption. This amounts to about 270 kWh per year per capita—the highest in the world.

Israel's example in domestic solar water heating provides an impressive demonstration of what can be achieved (in countries with similar or even more favorable climates), if the government makes a commitment to clean and environment-friendly technologies [5]. However, while solar for residential use has become an everyday reality in Israel, the much larger industrial/commercial sector uses very little solar energy, despite the fact that the industrial user is much better suited to do so than a homeowner. Some key considerations are: industry works mostly during the day, requiring little storage relative to a residence; the economy of scale provides a significant capital advantage to large industrial installations; industry generally has plenty of roof area in single-story buildings located in areas where architectural considerations do not hinder the installation of solar collectors; the industrial user is used to perform small maintenance jobs, thus eliminating the need for a fool-proof system and reducing first cost. While some industries require high-temperature process heat, there are many who need the same temperature range as the domestic user; these include textile, food, pharmaceutical, chemical, and many more. The same applies, of course, to the commercial sector. It is estimated that widespread solar energy utilization in industry for process heat and the like, and in commercial applications, could increase the country's utilization by a factor of five, if not more.

Unfortunately, present tax considerations create a negative incentive for businesses to use solar energy. An industry burning polluting fuel can write the cost off as a business expense, thus reducing its tax liability, whereas an investment in a solar heating system can only be amortized over 12–15 years, making it considerably less attractive economically. Moreover, the law presently exempts industrial plants, shops, hospitals, and high-rise buildings (height over 27 m) from the requirement to install a solar water heating system in new buildings [5]. The government could play an important role in changing this situation, by introducing appropriate measures, closing tax loopholes, and creating positive incentives for renewable energy. This can be achieved within a short time—there is no need for long-term investments and development of new technologies. Solar energy is a reality here and now, as already demonstrated by the country's residential sector.

References

1. Israel Central Bureau of Statistics web site: http://www.cbs.gov.il.
2. Israel Ministry of National Infrastructures web site: http://www.mni.gov.il/ (English version: http://eng.mni.gov.il/english/).
3. Considine, T. J. 2003. An energy demand forecasting model for Israel and the market potential for natural gas. Draft Final Report, Pennsylvania State University, May 19, 2003.
4. Planning and Building Law—application for permit, conditions thereof, fees and regulations, 5730/1970. System Article 9: Installation of a solar water heating system.
5. Nowarski, J. 2000. Solar Israel—a practical and legislative model. Renewable Energy World, March–April 2000, pp. 92–99.

2.5 Renewable Energy Policies for China

Debra Lew

2.5.1 Introduction

China has long been a world leader in renewable energy development and utilization, especially decentralized small-scale renewable energy technologies, such as small hydropower, solar water heaters, biogas digesters, and small wind turbines. Today, China is becoming a leader in sophisticated, high-technology renewables, like PV, and is positioning itself for significant growth in other sectors, such as utility-scale wind power, biopower, and biofuels. Helping to drive this transformation is the consolidation of renewable energy policy, programs, and planning into a single organization, the National Development and Reform Commission (NDRC) and the formal approval by the National People's Congress of China's first comprehensive renewable energy legislation, the Renewable Energy Law.

China recently announced an increase in their renewable energy share (including large hydro) of primary energy (excluding traditional biomass) from 7 to 15% by 2020. Realizing this target would require an estimated 130 GW of renewable energy capacity with an investment of up to US$184 billion [1].

China has multiple drivers for renewable energy development and utilization: energy diversification, energy security, environmental issues, and sustainable development. China also seeks to be a leading manufacturer of renewable energy technologies and actively supports domestic manufacturing.

2.5.2 Market Overview

China is the largest solar water heater market and producer in the world, with three-fourth of the world capacity and some 2000 manufacturers, many of which are small, cottage-type enterprises.

Over 80 million square meters of installed capacity provide 10% of China's hot water needs [2,3]. Excluding large hydropower, small hydropower remains China's largest renewable energy sector with 34 GW currently installed. China has approximately half of the world's capacity of small hydropower and a mature industry [4].

China is quickly ramping up their solar PV manufacturing base. They are now the third largest PV producer, after the European Union and Japan, and have 400 MW of manufacturing capacity. This is expected to double by the end of year 2007 [5]. Most of this PV is currently exported, but the off-grid market continues to be a key sector. Of the ambitious 1.8 GW target for 2020, about half of this is expected to be for PV rooftops, with 25% for off-grid and 25% for centralized plants [6].

Although China has world-class wind resources, they have only recently begun to overcome internal barriers and ramp up capacity. They exceeded 1 GW in 2005, a goal which had originally been targeted for the year 2000. New targets are aggressive, but China may finally have the momentum to reach them. There are several wind turbine manufacturers in China, most of which produce 600–750 kW turbines, with pilot production of megawatt-scale turbines. Additionally, China leads the world with the dissemination of some 200,000 small-scale wind turbines that provide power in remote areas, like Inner Mongolia [2].

Most of China's 2300 MW of biomass power capacity is in the sugarcane bagasse sector. China has significant untapped biomass potential, especially in agricultural and forest residues. However, China lacks mature conversion technologies. The government has set ambitious targets for biomass power generation and is offering a significant financial incentive as discussed below. The biofuels sector is likely to see rapid growth, as China is now the world's second largest oil consumer and importer. China is already the third largest biofuels and third largest ethanol producer in the world and has set a target of 14 billion liters of ethanol production in 2020.

Table 2.5 shows the current installed capacity or production from renewables as well as targets for 2010 and 2020. By 2010, renewable energy power generation should total 191 GW, or approximately 30% of China's total installed power generation capacity. By 2020, this should reach 362 GW or 34% of the total. With associated increases in renewably generated heat and fuels, China's total use of renewables should increase to 540 million tonnes of coal equivalent or 16% of China's primary energy consumption by 2020, thus exceeding their recent announcement.

2.5.3 The Renewable Energy Law

For many years, China's national and local governments have encouraged the use of renewables. Over the last decade, many provinces and autonomous regions established local subsidies and tax credits [7]. The national government established ambitious targets. Although small-scale technologies flourished under this support, utility-scale technologies like grid-connected wind power did not. For example, in 1994, a national regulation required utilities to allow interconnection of wind power and to purchase the power at cost plus a "reasonable profit," with the incremental costs shared by the grid. However,

TABLE 2.5 Current Renewable Energy Capacity/Production and 2010 and 2020 Targets

Technology	Current Installed Capacity	2010 Target	2020 Target
Wind power [2]	1260 MW	5 GW	30 GW
Solar PV [2]	70 MW	300 MW	1.8 GW
Biomass power [12,2]	2300 MW	5.5 GW	30 GW
Hydropower [2]	110 GW (35 GW is small hydro)	180 GW	300 GW
Solar water heaters [2]	80 million square meters		300 million square meters
Bioethanol [3]	3650 million liters		14 billion liters

"reasonable profits" were not defined; and it was not clear whether sharing of incremental costs should be at a regional or national level. This resulted in an ineffective policy attempt to promote wind power.

The new Law, which was developed at the beginning of 2003 and came into force at the beginning of 2006, appears to overcome these difficulties. Implementing regulations have begun to be detailed and already electricity prices for some technologies and a mechanism for sharing incremental costs at a national level have been defined. It should be noted that while the Law follows earlier policies about the importance of renewables and how renewables foster socio-economic development, this Law initiates a more market-oriented approach, addressing financial issues and how the public good of renewable energy use can be cost-shared among all end users.

Until this Law, few nationwide financial incentives for renewable energy existed. A notable exception was the reduction in value-added taxes (VAT) and income taxes for some renewables, such as biogas, small hydropower, and wind power. For example, VAT on wind power was reduced from 17 to 8.5% in 2001 [8]. This new Law provides consistent financial incentives at a national level for renewable energy power generation. The Law requires grid companies (at the end of 2002, the former State Power Corporation was restructured into five generation and two grid companies) to allow interconnection of renewable energy generators and to purchase their electricity at either "fixed" or "guided" prices, to be determined later.

2.5.4 Implementing Regulations

The Law outlines general principles which are then detailed in implementing regulations. The Chinese government recognizes that policy-making is an iterative process and has established "trial" measures for pricing that may be reviewed and revised.

In early 2006, implementing regulations regarding administrative and pricing provisions were issued. These include project approvals by NDRC for large hydropower (for 250 MW facilities and larger) and wind power plants (for 50 MW facilities and larger). Power grids are responsible for interconnection of renewable energy power plants. Two forms of pricing are proposed: fixed pricing and pricing based on a public tendering process. Biomass will receive a US$0.03/kWh subsidy for 15 years over and above the 2005 desulfurized coal power price of the province or region. Solar, ocean, and geothermal power will also receive fixed pricing based on "reasonable profits" which have not yet defined.

Wind power will receive a price based on public tendering. This is based on NDRC's wind concessions program which started in 2002 and has resulted in the awarding of long-term power purchase agreements to develop 1300 MW of wind power [9]. Winning bid prices have been in the range of US$0.05–0.06/kWh. Although tendering has proven successful in awarding contracts, it remains to be seen whether it will be successful in generating renewable energy, or whether China will follow the way of the U.K.'s Non-Fossil Fuel Obligation of the 1990s in which tendering resulted in many awarded projects that were never commissioned, likely due to underbidding [10]. As in other countries, such as the U.S., regional and local policies may be more progressive than national policies. In the case of Guangdong province, a winning wind concessions bid of US$0.061/kWh helped the local government justify and establish a feed-in tariff of US$0.064/kWh for all nonconcession wind projects in the province.

The Law also specifically supports the use of renewable energy for remote rural areas. In 2002, NDRC initiated the Township Electrification program to electrify those remaining 1061 townships that lack electricity with PV, wind, and small hydropower [11]. Approximately, US$340 million was spent on the PV- and wind-powered townships. NDRC is now expanding this program to the village level and will electrify those remaining 25–30 million people in the 20,000 villages which do not yet have electricity. The Village Electrification program will cost an estimated US$5 billion and will be implemented over the next ten years. These off-grid users will be cross-subsidized by the grid-connected users so that they will pay a tariff equivalent to the provincial grid price.

To fund the incremental costs of renewably generated power (above the provincial coal power price) as well as the operations and maintenance of the off-grid systems, a surcharge will be applied to all end users (except for those in Tibet and agricultural users) served by the grid.

2.5.5 Conclusion

After the decades of policy attempts, pilot programs, and learning from their own and others' experiences, China seems finally poised to expand their renewable energy development on a gigawatt-scale. There are still barriers to be overcome, and China is using some approaches, such as the wind concessions program that have proved unsuccessful elsewhere, but China's growing renewable energy technical capabilities and the government's solid commitment to renewables have a good chance of weathering the challenges that come their way.

References

1. Fu, J. 2005. Renewable Energy Gets huge Outlay. China Daily, p. 2 , November 8.
2. Li Junfeng, Shi Lishan, Wang Zhongying, et al. 2006. *Overview of Renewable Energy Development in China*, Beijing, China.
3. Worldwatch Institute 2006. *State of the World 2006*. W.W. Norton & Company, NY.
4. http://www.iash.info/shpchina.htm
5. Jager-Waldau, A. 2005. PV Status Report 2005: Research, Solar cell Production, and Market Implementation of Photovoltaics, European Commission Joint Research Centre, August.
6. Personal communication, Jean Ku 2006.
7. Li, Z. 2005. China's Renewables Law. *Renewable Energy World*, Feburary 7.
8. NREL, 2004. *Renewable Energy in China: Renewable Energy Policy in China: Financial Incentives* NREL/FS-710-36045, April.
9. Ku, J., Lew, D., Shi, P., and Wallace, W. 2005. The Future Is Now: Accelerating Wind Development in China, *Renewable Energy World*, pp 212–223. July–August.
10. Wiser, R., Hamrin, J., and Wingate, M. 2002. Renewable Energy Policy Options for China: A Comparison of Renewable Portfolio Standards, Feed-in Tariffs, and Tendering Policies, Center for Resource Solutions, June.
11. Ku, J., Lew, D., and Ma, S. 2003. Sending Electricity to Townships, *Renewable Energy World*, pp. 56–67, September–October.
12. Installed capacity at end of 2004. Eric Martinot, 2006. http://martinot.info/china.htm

2.6 Japanese Policies on Energy Conversation and Renewable Energy

Koichi Sakuta

2.6.1 Introduction

The current basic energy policy in Japan aims at both achieving a stable energy supply and preserving the environment. Even though Japan has achieved considerable success in energy conservation during the last two oil crises, the country's total energy demand has continued to increase. This is especially true for energy consumption in the residential/commercial and transportation sectors which has increased significantly since 1990. To realize further development in the future, it is essential to secure additional energy supplies; however, we must also reduce CO_2 emissions to prevent global warming. As the chair country of the Kyoto Conference on Climate Change (COP3), Japan is strongly committed to achieve the target of Kyoto Protocol.

The Advisory Committee for Natural Resources and Energy under the Agency for Natural Resources and Energy, the Ministry of Economy Trade and Industry (METI) intensively reviews and discusses future energy demand policies. After more than a year of study, their report "Prospect for Supply and Demand in 2030" [1] drew the following conclusions:

- Increase in total energy demand will be saturated around 2020.
- Potential for energy conservation is quite large, up to 50 million kiloliters of oil.
- Distributed power generation has a potential to produce up to 20% of total electricity.
- Renewable energy can produce up to 10% of primary energy demand in 2030.
- Technology can enable compatibility of economy and environment.

2.6.2 Energy Conservation

Through the energy efficiency efforts of the public and the government, Japan has achieved the highest level of energy efficiency anywhere in the world since the oil crises[2]. However, the weakness of Japanese energy supply structure remains unchanged, and the level of dependency on Middle-East crude oil is higher now than at the time of the oil crises. With increasing energy consumption in commercial/residential and transportation sectors in recent years, it is essential to promote energy efficiency measures for the future.

In 1998, aiming to achieve Japan's COP3 energy efficiency goal, a conservation target of 56 million kiloliters (crude oil equivalent) was formulated. This target value was revised in 2001 to 57 million kiloliters including an additional 7 million kiloliters for the rapidly increasing demand in residential/commercial and transportation sectors.

Future measures for energy conservation include:

a. Industry sector
   - Steady implementation of voluntary action plan.
   - Voluntary management system based on "The Law Concerning the Rational Use of Energy."

b. Commercial/residential sector
   - Efficiency improvement of appliance use.
   - Improved energy efficiency of houses and buildings.
   - Complete management of energy demand.

c. Transportation sector
   - Acceleration of measures to improve the fuel efficiency of cars.
   - Execution of measures to appeal to personal selection.
   - Smoothing out of traffic/traffic flow and managing automobile traffic.

d. Cross-sector measures
   - Decision on technological strategy.
   - Promotion of energy efficiency education.
   - Promotion of high efficiency cogeneration.
   - Leading the way with implementation in the public sector.

One of the unique measures is the "Top Runner Program" in the commercial and residential sector under the Energy Conservation Law. All the manufacturers are required to meet the target energy conservation standards, which are set by the best products ("top runners") presently available in the market. At present, 18 product categories are covered by the standards, including air conditioners, refrigerators, etc.

2.6.3 Renewable Energy

2.6.3.1 Outline

In the METI renewable energy policy, new energy sources are defined and promoted—which include photovoltaic (PV) power generation, wind power generation, solar thermal energy, ocean thermal energy, waste power generation, waste thermal energy, waste fuel production, biomass power generation, biomass thermal energy, biomass fuel production, cool energy of snow and ice, clean energy vehicles, natural gas cogeneration, and fuel cells[2].

Both R&D and policy measures to introduce renewable energy are strongly promoted because of their effectiveness in realizing sustainable energy systems in the future. In the 2005 METI report "Prospect for Supply and Demand in 2030," national targets for introducing various new energy sources by 2010 are presented in Table 2.6.

METI's current policy for new energy can be summarized as:

a. Law concerning the development and promotion of Oil Alternative Energy (Alternative Energy Law).
b. Long-term energy supply/demand outlook.
c. Enactment of "Law Concerning Special Measures for Promotion of New Energy Use, etc., (New Energy Law)" enacted in April 1997.
d. Government support system for the domestic introduction of new energy.
   - Technological development;
   - Validation tests;
   - Implementation promotion;
e. Promotion of international cooperation related to new energy.

TABLE 2.6 Present Status and Targets for Introduction of New Energy Sources

	FY 1999	FY 2002	FY 2010			
			Reference Case	Current Measures Case	Additional Measures Case	Under Current Goals Case
Solar power	53,000 kl	0.156 million kl	0.62 million kl	1.18 million kl	Same as at right	1.18 million kl
	0.209 million kW	0.637 million kW	2.54 million kW	4.82 million kW		4.92 million kW
Wind power	35,000 kl	0.189 million kl	0.32 million kl	1.34 million kl		1.34 million kl
	83,000 kW	0.463 million kW	0.78 million kW	3.00 million kW		3.00 million kW
Waste-fired power	1.15 million kl	1.52 million kl	2.08 million kl	5.52 million kl		5.52 million kl
	0.9 million kW	1.40 million kW	1.75 million kW	4.17 million kW		4.17 million kW
Biomass power	5,400 kl	0.226 million kl	0.226 million kl	0.34 million kl		0.34 million kl
	80,000 kW	0.218 million kW	0.218 million kW	0.33 million kW		0.33 million kW
Use of solar heat	0.98 million kl	0.74 million kl	0.74 million kl	0.74 million kl	10.72 million kl	4.39 million kl
Use of waste	44,000 kl	36,000 kl	44,000 kl	0.14 million kl		0.14 million kl
Use of biomass heat	—	—	—	0.67 million kl		0.67 million kl
Unused energy[a]	41,000 kl	60,000 kl	93,000 kl	058 million kl		0.58 million kl
Black liquor, waste wood, etc.[a]	4.57 million kl	4.71 million kl	4.87 million kl	4.87 million kl		4.94 million kl
Total (Proportion of total primary energy supply)	6.93 million kl (1.2%)	7.64 million kl (1.3%)	8.99 million kl (1.4%)	15.38 million kl (2.6%)	19.10 million kl (Approx. 3%)	19.10 million kl (Approx. 3%)

[a] Unused energy includes ice thermal energy storage; Black liquor and waste wood are a kind of biomass and includes the portion used for power generation; Because the amount of black liquor and waste material introduced depends upon the level of paper pulp produced in the energy field, it is calculated internally in the model.

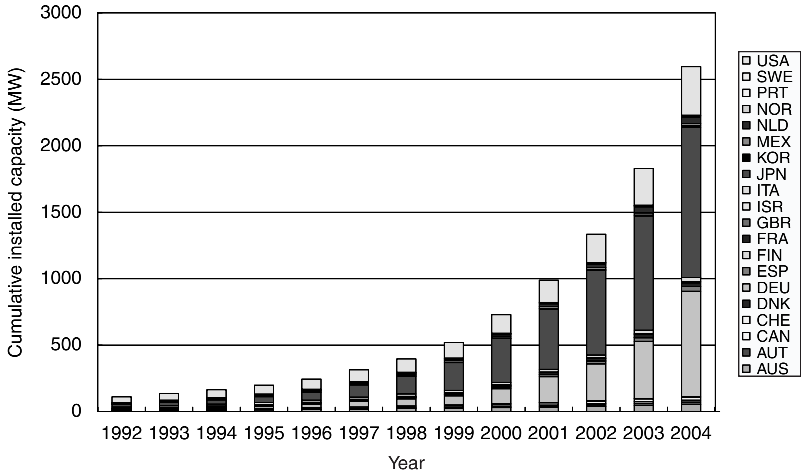


FIGURE 2.6 Cumulative installed capacity of PV systems in selected IEA countries. (From IEA/PVPS Trends in photovoltaic applications in selected IEA countries between 1992 and 2004; Hamelinck, C. N. and Faaij, A. P. C. Article in press. *Outlook for Advanced Biofuels*. Energy policy.)

2.6.3.2 Photovoltaics

Among the various renewable energies, PV power generation is one of the top priorities in Japan. As seen in Figure 2.6, about half of the cumulative PV capacity introduced in the world is currently in Japan [4]. Figure 2.7 shows the long-term R&D and promotion road map towards 2030 prepared by NEDO (New Energy and Industrial Technology Development Organization) [5]. Our target is the

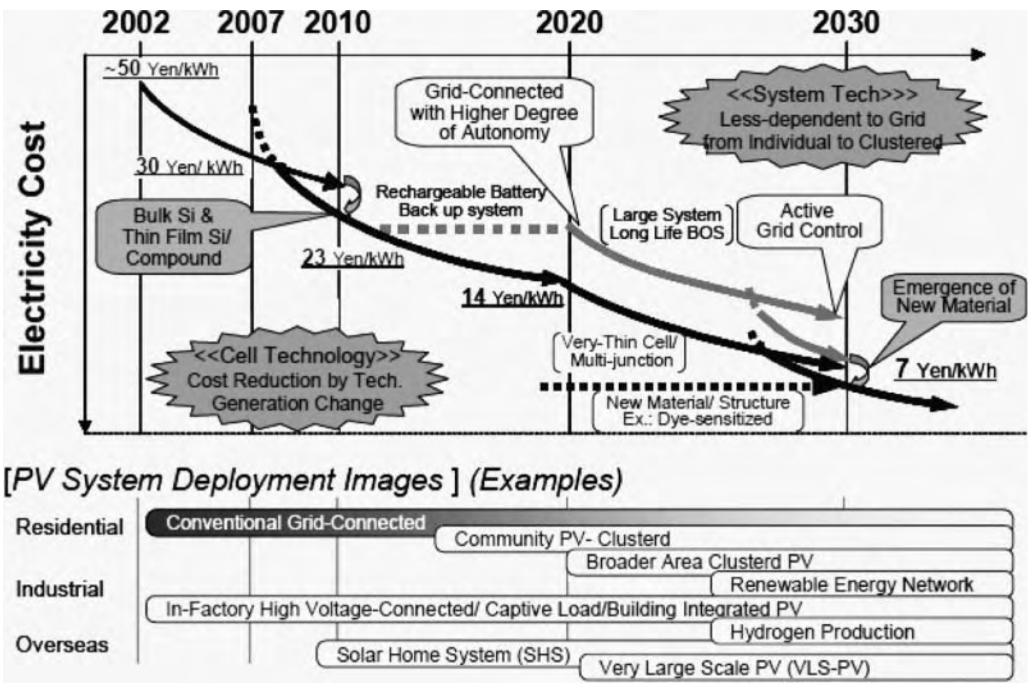


FIGURE 2.7 Scenario for improving economic efficiency of PV power generation. (From Overview of PV Roadmap Toward 2030, NEDO, 2004; Ranney, J. W. and Mann, L. K., *Biomass and Bioenergy*, 6(3), 221–228, 1994.)

introduction of 4.8 GW from PV by 2010 and 102 GW by 2030. The latter target corresponds to about 10% of total electricity demand. To achieve this goal, it is necessary to reduce the power generation cost to about 7 yen/kWh which corresponds to the wholesale cost of electricity.

2.6.3.3 Wind Energy

Introduction of wind power generation systems is progressing rapidly in areas with good wind conditions. The total capacity of around 700 MW was achieved in FY2003 and the target for 2010 is 3 GW.

2.6.3.4 Fuel Cells

High priority is also placed on the development of fuel cells. Aiming at the early introduction of fuel cell vehicles, much effort is being done in R&D.

2.6.3.5 Biomass

Biomass energy has a huge potential as future renewable energy source because of its variety and many different applications. A new initiative called "Biomass Japan" was established in 2002 by six ministries to realize a sustainable society by comprehensively utilizing various biomass products as energy resources.

2.6.3.6 Geothermal Energy

Although not included in "new energy," 19 geothermal power plants with more than 530 MW total capacity are under operation in Japan. Geothermal resources can become competitive with fossil fuel power generation in the long run, but it will be necessary to reduce the development cost and to introduce various measures to encourage further development of geothermal power.

References

1. Advisory Committee for Natural Resources and Energy. Prospect for Supply and Demand in 2030, March 2005 (in Japanese).
2. Website of METI (http://www.enecho.meti.go.jp/english/policy/index.html).
3. ANRE, METI, Energy in Japan 2005, pdf file available at http://www.enecho.meti.go.jp/english/index.htm
4. IEA/PVPS Task1. Trends in photovoltaic applications. Survey report of selected IEA countries between 1992 and 2004. Sept, 2005.
5. NEDO. Overview of PV Roadmap Toward 2030, June 2004.

2.7 Renewable Energy Policies in Brazil

Ricardo Rüther

2.7.1 Bioenergy and Transportation

Bioenergy is regarded as one of the key options to mitigate greenhouse gas (GHG) emissions and to substitute fossil fuels [1,2]. Although there are many renewable energy technologies emerging for large-scale electricity production with low or no GHG emissions, the transport sector is almost entirely based on fossil fuels, and has fewer alternatives.

Transportation represents some 27% of the world's secondary energy consumption (21% of primary), and is almost exclusively fueled by petroleum oil [3,4]. Biofuels can play an important role in addressing both the GHG emissions of transport and the dependency on petroleum oil.

A few main routes can be distinguished to produce biofuels: extraction of vegetable oils, fermentation of sugars to alcohol, gasification and chemical synthesis, and direct liquefaction. These can lead to a variety of fuels: methanol, ethanol, hydrogen, synthetic diesel, biodiesel, and bio-oil, all with very different properties, as described in [4].

With the exception of sugarcane ethanol, the traditional biofuels have a number of severe disadvantages that are related to the feedstock. The current costs of rapeseed biodiesel and ethanol from cereals or beets are much higher than the costs of gasoline and diesel, and substantial subsidies are needed to make them competitive. These high costs are a result of the low net energy yield of most annual crops (100–200 GJ/ha yr in the long term [4]), the high quality agricultural land required, and the intensive management. The lower productivity per hectare and high fertilizer requirement also limit the well-to-wheel reduction of fossil energy use, which limit the environmental benefits [5,6]. The net energy of perennial crops (220–550 GJ/ha yr), grasses (220–260 GJ/ha yr), and sugar cane (400–500 GJ/ha yr) is considerably higher, and Brazil has been a world leader in promoting biofuels for 30 years under its *Proálcool* program.

2.7.2 The *Proálcool* Program in Brazil

Since 1975, Brazil has mandated that ethanol be blended with all gasoline sold in the country. Although the required blend level is adjusted frequently, it has been in the range of 20%–25%, and all filling stations are required to sell gasohol (E25) and pure ethanol (E100). Tax preferences have been given to vehicles that run on pure ethanol, and more recently, the introduction of the so-called flex fuel vehicles by most of the local automakers allows for any proportion between ethanol and gasoline (E0–E100) to be used any time. This has led to the present (and growing) 73% share that flex fuel cars hold in the local market [7,8], which should reach 85% by the end of 2006 [9]. In 10 years, 16 billion US$ were invested in genetic research for sugar cane improvement, alcohol subsidies and agrobusiness, and policies have included blending mandates, retail distribution requirements, production subsidies, and other measures [10].

After a rocky period in the 1980s and 1990s, when high demand resulted in a sellers market which led to high alcohol prices and delivery shortages [10]. The escalating oil prices and the event of the flex fuel car have brought attention back to ethanol. Advances in electronic fuel injection technology solved many of the technological problems associated with ethanol engines (in cool mornings, drivers had to spent their precious minutes for their car engines to warm up before they could drive). In 1985, with falling oil prices, the Brazilian government was not able to keep up with the required subsidies, and in 1989 ethanol shortages and high prices resulted in very disappointed drivers and the near collapse of this promising technology. In 1990 the sector was deregulated, and the *Alcohol and Sugar Institute* (IAA—Instituto do Açúcar e do Álcool), which regulated export quotas and subsidies, was closed down. In a free, deregulated market, ethanol producers chose to turn back to sugar when the international price of that commodity recovered, and the automakers reduced alcohol driven car production to negligible levels. More recently, with the boom in flex fuel car sales, ethanol prices soared once again. This time, however, with flex fuel cars came the choice for consumers to avoid abusive price increases from alcohol producers, and whenever ethanol prices go over the mark of 70% of gasoline price, flex fuel car drivers fill up with the traditional fossil fuel. Although at present, seven out of ten cars sold in Brazil run on both fuels, only 2% of the Brazilian fleet are flex fuel cars, with an aged 16% share running exclusively on ethanol, 10% on diesel, and 72% on E25 gasoline. These shares, however, are changing fast. The first flex fuel model was released in 2003, when Volkswagen sold 48,000 units. Cumulative sales have recently reached the 1.2 million mark [9].

Having produced 36% of the 42.2 billion liters of ethanol consumed worldwide in 2005 [8], Brazil produces the lowest cost ethanol worldwide, thanks to genetic R&D. This has led to a more robust sugar cane variety that is also richer in saccharose, and the country expects to produce another 16 billion liters of ethanol in 2006. Brazil is now producing ethanol at 6550 liters per hectare, having nearly doubled the productivity from 1975 when the yield was at 3500 liters of ethanol per hectare. RD&D efforts continue, with the release of new generation of flex fuel cars, and increased productivity levels in the sugar cane industry.

2.7.3 Bio-oil

Brazil has more recently begun to target the increased use of biodiesel fuels, derived primarily from domestically produced soybean oil, with recent legislation allowing for blends of 2% biodiesel in diesel fuels (B2), which may be increased to 5% (B5) in the near future if the market responds favorably. There is a real possibility of setting up a national biodiesel program similar in scope and magnitude to *Proalcool*. In 2003, Brazil consumed approximately 38 billion liters of diesel, and 6 billion liters were imported; a large biodiesel program could replace most of the imported diesel, while creating some 200,000 new jobs in rural areas [11]. Brazil has the potential to be a world leader in biodiesel production, as it will be able to produce bio-oil from many different sources.

2.7.4 Prospects

In addition to Brazil, mandates for blending biofuels into vehicle fuels have been appearing in several other countries in recent years, mostly driven by Kyoto targets. Because ethanol can reduce carbon emissions up to 80% in comparison with petrol [8], many countries have already approved gasohol blends. In India, the government mandated 10% ethanol blending (E10) in most sugar cane-producing areas. In China, some provinces mandate E10 blending. In the U.S., a few states have adopted E10 and B2, and many states are considering similar mandates. Canada will start E5 by 2007, Columbia is adopting E10, and the Dominican Republic will have E15 and B2 by 2015 [7]. Japan is considering an E5-blending mandate based on imports from Brazil. In most of these countries, distribution infrastructure is the major drawback and takes time and investment to be established. In Brazil, half of the 30,000 service stations are equipped with ethanol dispensers.

The adoption of E5 worldwide would demand some 100 billion liters per year, more than doubling the current world demand.

References

1. Hall, D. O., Rosillo-Calle, F., Williams, R. H., and Woods, J. 1993. Biomass for energy: Supply prospects. In *Renewable Energy, Sources for Fuels and Electricity*, T. B. Johansson, H. Kelly, K. N. R. Amulya, and R. H. Williams, eds., pp. 593–653. Island Press, Washington, DC.
2. Goldemberg, J. 2000. *World Energy Assessment*. Preface. United Nations Development Programme, New York, NY, U.S.A.
3. EIA. 2003. *International Energy Outlook 2003* (www.eia.doe.gov/oiaf/ieo.index.html). Energy Information Administration, Office of Integrated Analysis and Forecasting, p. 272. Washington, DC, U.S.A.
4. Hamelinck, C.N. and Faaij, A.P.C. Article in press. *Outlook for Advanced Biofuels*. Energy Policy.
5. Ranney, J. W. and Mann, L. K. 1994. Environmental considerations in energy crop production. *Biomass and Bioenergy*, 6 (3), 221–228.
6. Berndes, G., Hoogwijk, M., and Van den Broek, R. 2003. The contribution of biomass in the future global energy supply: A review of 17 studies. *Biomass and Bioenergy*, 25 1–28.
7. Martinot, E. 2005. *REN21 Renewable Energy Policy Network—Renewables 2005 Global Status Report*, Worldwatch Institute, Washington, DC, U.S.A.
8. Guandalini, G. and Silva, C. 2006. *A Dupla Conquista*, Veja, 01/02/2006, pp. 90–94.
9. Nascimento, G. 2006. *A Febre dos Carros Flex*, Isto É, 01/03/2006, pp. 72–73.
10. Rosillo-Calle, F. and Cortez, L. A. B. 1998. Towards Proalcool II—A review of the Brazilian bioethanol programme. *Biomass and Bioenergy*, 14 (2), 115–124.
11. www.procana.com.br, www.bioenergy-lamnet.org

2.8 Policies for Distributed Generation

Jeff Bell

2.8.1 Introduction

Decentralized energy[1] is showing ever-increasing promise as a cost-effective method of electricity supply resulting in significant environmental benefits. Though every sector of the DE industry is showing impressive growth, the total share of DE in overall global electricity capacity remains only about 7.2%.[2] Lack of clear policy in the energy sector is one of the main barriers that prevents more DE investment. It is clear that adopting policies, which ensure an independent regulator combined with open markets that reward investors for noneconomic benefits of DE will be essential in driving future growth in the sector.[3]

In June 2004, the largest renewables energy event to date was held in Bonn, Germany. More than 3000 attendants from 154 nations attended. Although the conference dealt only with renewables, to a large extent the recommendations would also hold true for all forms of DE, including cogeneration. Arising from that event were some guidelines for policy makers as to what must be put in place in order for DE policy to be effective. Therefore, in order for DE to become an integral part of the energy mix, policymakers must:[4]

1. Integrate DE into overall energy policy
2. Establish clear goals
3. Establish transparent market conditions
4. Reduce or eliminate subsidies for conventional centralized energy
5. Address cost issues
6. Create incentives that are eventually phased out
7. Ensure energy issues form part of the basis of decisions in "nonenergy" sectors, such as urban planning and infrastructure, etc.
8. Educate the public on the benefits and costs of DE
9. Promote DE jobs and training
10. Develop public organizations to promote DE
11. Strengthen regional and international cooperation on DE matters
12. Secure grid access for DE
13. Not omit thermal energy in decision-making
14. Support research and development in DE
15. Involve a range of actors in promoting DE, including local governments, financiers, international bodies, and multi and bilateral banks and institutions
16. Harness the power of public procurement

[1]DE technologies consist of the following forms of power generation systems that produce electricity at or close to the point of consumption:

- High efficiency cogeneration/combined heat and power
- On-site renewable energy systems
- Energy recycling systems, including the use of waste gases, waste heat, and pressure drops to generate electricity on-site

WADE classifies such systems as DE regardless of project size, fuel or technology, or whether the system is on-grid or off-grid.

[2]World Survey of Decentralized Energy. 2005. WADE.

[3]Seven Guiding Principles for Effective Electricity Market Regulation. 2003. WADE.

[4]http://www.renewables2004.de/pdf/policy_recommendations_final.pdf

In general, it is important that a policy be stable enough in the long term to foster confidence to everyone that the rules will not change after they have committed to one plan of action. Involving locals in the establishment of DE policies is the other key to both garnering public support and establishing a basis of shared experience on which future policy revisions can be made.

The above list illustrates some of the goals of policies to promote DE. There are a multitude of approaches that can be used to achieve these goals and the following section provides many examples from all over the world. For each of the nations covered in this chapter, there will be a short summary paragraph followed by a table with three major headings: Technical, Financial, and Other.

The Technical heading will discuss policies that address technical issues, such as mandated technologies, interconnection procedures, manufacturing and interconnection standards, and safety rules. Technical policies if designed carefully can be important drivers for DE investment just as poorly designed policies or lack of policies can stifle investment in DE.

The Financial heading will cover policies that exist in the area, which are based on economic incentives, such as rebates for investing in certain technologies, tax breaks for eligible investments and pricing arrangements for power produced by onsite generators. Although project economics is affected by a multitude of factors (including fuel prices, capital expenditures, opportunity costs, etc.), financial policies can often make the difference between a project that is feasible and one that is not.

The Other section will discuss all remaining issues relating to DE policy, including more general policy guidance and legislation mandating DE use or energy efficiency/renewables.

The general organization of each section will take the form of the table below.

In some cases, no examples for the specific category of policy/technology were found for a given jurisdiction, in which case the given cell was left blank (Table 2.7).

It is very difficult to separate policies that aim to promote renewables and efficiency in general from those aimed at promoting DE. For example, many policies aim to promote renewables in general including large-scale wind farms, which should not necessarily be considered DE. Likewise, some policies are aimed at only cogeneration and do not affect renewables. In other cases, for example, in the case of rooftop PV, renewable and DE policies are one and the same. In most cases, however policies

TABLE 2.7 General Organizational Structure for Country Policy Summaries

Technical
DE in general
Large-scale cogeneration
Domestic cogeneration
PV
Wind
Hydro
Financial
DE in general
Large-scale cogeneration
Domestic cogeneration
Onsite PV
Onsite wind
Small hydro
Other
DE in general
Large-scale cogeneration
Domestic cogeneration
PV
Wind
Hydro

that affect DE will be framed in the context of policies to promote renewables or general efficiency (in the case of cogeneration). Currently, there remains a shortage of concrete policy examples that target onsite power or DE. In order for DE to truly reach its potential laws and policies will have to be increasingly aimed at DE specifically.

2.8.2 Canada

With a low population density and wide climatic variations, Canada has one of the world's highest energy use per capita. The Canadian government, still publicly committed to the Kyoto protocol, has taken some steps to promote DE using financial mechanisms, but coordinated policy cooperation between the various provincial and the Federal government may prove most effective for promoting DE (Table 2.8).

2.8.3 Mexico

Mexico is resource rich for many DE sources, including natural gas, sunshine, wind and biomass, but little of that potential has been realized. To date there has been limited success in developing a strong policy foundation for DE. Garnering support from Mexico's publicly owned energy giants will be key to promote DE technologies there (Table 2.9).

2.8.4 USA

The United States remains highly influential in international energy policy development and remains one of the few nations that have not ratified the Kyoto protocol. As the single largest energy-using nation, policy to promote DE is clearly needed to develop capacity to meet the USA's rising demand. Many innovative policy developments have been introduced at the State level that have effectively buoyed DE markets and the USA remains one of the world leaders in renewable energy development. Steps to harmonize State level policy with policy at the Federal level will be the main factor for successfully advancing DE (Table 2.10).

2.8.5 Argentina

The energy sector in Argentina, though subject to wholesale competition, remains largely dominated by a few large players. In recent years, the additional investment required to keep the power on has been shortcoming as result of market uncertainties but plentiful natural gas as well as some new policy developments promise to buoy DE investments in the near future (Table 2.11).

2.8.6 Brazil

Brazil has for a long time been a world leader in alcohol fuels and some recent policy developments suggest that it may be well on its way to showing similar leadership in development of cogeneration. The plentiful biomass resources along with newly accessible supplies of on and offshore natural gas, combined with feed-in laws for renewables and gas-fried cogeneration are likely to drive significant investment in DE in the near future (Table 2.12).

2.8.7 Chile

Chile was the first nation in the world to break up vertically integrated monopolies into separate markets for generation and wires. Strong policy is still lacking to directly promote DE, which partly explains the lack of investment in the area. Recent gas shortages in Chile have been creating somewhat of an energy crisis that should encourage gas conservation (and thus efficient gas cogeneration), but will likely reduce interest in gas-fired cogeneration in the near future. Other DE sources, such as

TABLE 2.8 Canadian Distributed Generation Policy by Technology and Type

	Technical
DE in general	The Federal government in Canada has control over interprovincial gas pipelines and transmission lines crossing international borders. Because the Canadian constitution enshrines energy as a provincial jurisdiction in all other cases; each utility tends to have its own interconnection policy. Interconnection policies tend not to be technology specific. The government funds the "micropower connect" initiative, which aims to harmonize interconnection rules in the provinces[a]
Large-scale cogeneration	Interconnection rules vary by province
Domestic cogeneration	Interconnection rules vary by province
PV	Interconnection rules vary by province
Wind	Interconnection rules vary by province
Hydro	Interconnection rules vary by province
	Financial
DE in general	There is no Federal law, which guarantees retail electricity prices. In most cases, onsite generators feeding excess power to the grid obtain wholesale electricity price or less and in some cases, retail price is earned. In only one case (a paper mill in Quebec using cogeneration), does onsite generation obtain a premium price to reflect the additional value it offers. Various funding schemes exist to promote energy innovation including: the technology Early Action Measure program, the Climate Change Technology and Innovation Program, the Federation of Canadian Municipalities Green Fund, and a program to promote onsite generation at government facilities.[b] In many cases, the utility companies of a province have much influence over provincial electricity policy so that policy is often determined by the private sector
Large-scale cogeneration	Biomass or landfill gas may qualify for Renewable Power Production Incentive: CAN¢1/kWh final program details and eligibility criteria will be announced by April 1, 2006 Cogeneration equipment is listed as Class 43.1 assets meaning they are eligible for accelerated depreciation rates for capital costs (30%). Systems with efficiency of more than 72% and all systems using biomass will be eligible for 50% capital cost allowance. Start-up expenses, such as feasibility studies and regulatory approval expenses, etc., are tax deductible under the Canadian Renewable and Conservation Expenses rule. In Quebec, Hydro Quebec will now permit independent power producers to sign long-term power supply contracts at avoided cost
Domestic cogeneration	Cogeneration equipment is listed as Class 43.1 assets meaning they are eligible for accelerated depreciation rates for capital costs (30%)
Onsite PV	May qualify for renewable power production incentive: CAN¢1/kWh final program details and eligibility criteria will be announced by April 1, 2006. Eligible for 50% capital cost allowance
Onsite wind	Wind power production incentive (WPPI) will be increased by CAN\$200 million over 5 years and CAN\$920 million over 15. CAN¢1/kWh will be paid for all eligible wind projects.[c] Eligible for 50% capital cost allowance. Siting studies and test turbines are tax deductible under the Canadian Renewable and Conservation Expenses rule
Small hydro	May qualify for renewable power production incentive: CAN¢1/kWh final program details and eligibility criteria will be announced by April 1, 2006. Eligible for 50% capital cost allowance
	Other
DE in general	The Canadian Government has a goal of supplying 20% of its electricity from renewable sources by 2010.[d] Various provinces also have similar goals including Price Edward Island, which has announced plans to become 100% renewable by 2012, Quebec, which aims to have 1000 MW wind by 2010, and Ontario which plans to install smart meters in every house by 2010

(continued)

TABLE 2.8 ***(Continued)***

	Canada is bound by the Kyoto protocol and aims at reducing its carbon emissions by 5.8% below 1990 levels by 2012
Large-scale cogeneration	
Domestic cogeneration	
PV	
Wind	
Hydro	

[a] http://www.micropowerconnect.org
[b] http://www.oja-services.nl/iea-pvps/countries/canada/index.htm
[c] Energy Law Bulletin. March 2005. McMillan Binch, LLP.
[d] Government of Canada. 2002. Climate change: achieving our commitments together. *Climate Change Plan for Canada*. En56-183/2002E, p. 33.

TABLE 2.9 Mexican Distributed Generation Policy by Technology and Type

	Technical
DE in general	In 1992, a law was established which permitted nonutility generation and onsite generation for the first time. The "Ley del Servicio Público de Energía Eléctrica" (LSPEE) forbids bilateral trading of electricity between individuals, but allows companies to generate for their own demand and cogeneration with a permit from the Comision Regulatorio de Energía (CRE)
Large-scale cogeneration	The Mexican government defined cogeneration in the 1992 law permitting private sector involvement in the electricity sector,[a] but stated that the power must be used onsite. Interconnection, wheeling, and backup services are allowed but require permits from the CRE and the energy secretariat (SENER). Interconnection standards for cogeneration plants are based on the US IEEE 1547
Domestic cogeneration	
PV	
Wind	
Hydro	
	Financial
DE in general	Private generators can sell to either of the two utilities on the following conditions[a] Cogeneration plants can sell their total surplus production Self-supply plants can sell up to 20 MWe of supply provided the plant is no bigger than 40 MWe For self-supply plants bigger than 40 MWe, up to 50% of the capacity can be sold Sale agreements cover dispatch arrangements, pricing, capacity payments, and energy payments. Prices are very low that sales rarely happen in practice. New rules to improve the situation may be passed by the Senate in early 2005 In 2001, the Mexican Energy Regulatory commission established a framework for the interconnection of intermittent sources of generation, such as solar, wind, and microhydro. The framework created incentives to feed power generated from such sources into the grid and created tariff regimes based on actual generation transmitted as opposed to nameplate capacity.[b] In effect the arrangement works like a net metering arrangement, allowing small generators to feed power into the grid in times of surplus generation, and draw power from the grid when power generation is insufficient to meet onsite demand
Large-scale cogeneration	A proposal to base surplus sales of electricity to Comision Federal de Electricidad (CFE)/Luz y Fuerza del Centro (LFC) on utility sales prices, rather than the current (early 2005) rates of US$2–3 ct/kWh may be approved by the senate in early 2005. New plant depreciation rules may also be introduced in 2005.[a] The government has revised electricity tariffs in a way that has made peak shaving attractive to independent owners of diesel fired power plants, reciprocating natural gas engines, 30–75 kW microturbines in commercial applications and 5–250 MW gas turbine plants for up to 4 h/day[c]

(continued)

TABLE 2.9 ***(Continued)***

Domestic cogeneration	
Onsite PV	
Onsite wind	
Small hydro	
	Other
DE in general	Investors have to secure a wide array of permits, which can be very costly. In some cases, more than 100 permits are required, including local/municipal planning permission, environmental permitting, gas supply and connection permits, power generation authorization and network interconnection permits, etc. The 2001–2006 National Energy Sector Plan states a goal of 1000 MW of new renewable energy capacity (including hydro)[b]
Large-scale cogeneration	CONAE provides technical assistance, including prefeasibility and online economic analysis as well as limited financial support for projects, workshops, educational programs, and conferences. A proposal to simplify permitting and authorization procedures for private cogeneration plants may be approved by the senate in early 2005. Other incentives that are being considered include: The provision of assured long-term pricing of gas and oil from Pemex together with incentive fuel pricing for cogeneration projects that meet certain criteria The introduction of improved dispatch arrangements, simplified interconnection/wheeling arrangements, and reduced backup charges from CFE The provision of incentives for cogeneration plants based on their energy savings, reduced network losses, and carbon emission reductions Specific recognition of cogeneration/DE in new electricity market reform laws At Bonn Mexico committed to a "Renewable Energy Initiative"[d] but the details of the plan remain uncertain
Domestic cogeneration	
PV	
Wind	
Hydro	

[a] WADE Market Analysis 2005.

[b] Renewable Energies in Mexico 2004. www.energia.gob.mx/work/resources/LocalContent/1835/1/e_renovables_mexico.pdf

[c] Jorge Hernandez Soulayrac, Tecnoelectric Power Consulting.

[d] Renewables 2004. *International Conference for Renewable Energies*, 1–4 June 2004, Bonn. Conference Report: Outcomes and Documentation—Political Declaration/International Action Programme/Policy Recommendations for Renewable Energies.

TABLE 2.10 USA Distributed Generation Policy by Technology and Type

	Technical
DE in general	Section 1254 of the Energy Policy Act of 2005 requires that electric utilities offer grid interconnection based on a nationwide standard. Currently to connect DE units must be IEEE P1547 compliant, which requires "automatic and rapid disconnection in the event of a fault"
Large-scale cogeneration	
Domestic cogeneration	
PV	
Wind	
Hydro	

(continued)

TABLE 2.10 ***(Continued)***

	Financial
DE in general	At a State level, the United States has many promising policies in support of renewables and DG. Grant programs, financing, R&D funds, net metering, feed-in tariffs, and tax breaks are some of the financial measures that have proven successful for promoting DE technologies at a state level.[a] A tool that has proven especially successful is the "system benefits charge" are being employed in at least 16 US states.[a] The Energy Act of 2005 grants tax credits for developers who build energy efficient buildings that could have ramifications for DE especially micro-cogeneration. It also extends the existing tax credit for production of electricity from renewable resources, such as wind, biomass, and landfill gas, and creates for the first time a tax credit for residential solar energy systems (section 1301)[b] As specific scheme exists to provide grants for rural and remote communities' electricity development projects (section 209) In 2004, about 100 million dollars were spent in the USA on R&D related to renewables[c]
Large-scale cogeneration	The new Federal US Energy bill contains a 10% tax incentive for microturbine generators and also offers a per kilowatt-hour incentive for generation from microturbines paid directly from the Department of Energy. Biogas fired turbines will receive an additional incentive of 1.5 ct/kWh (section 1336)
Domestic cogeneration	The new Energy bill provides a rebate of up to $150 for installation of a highly efficient furnace or boiler; presumably, micro-cogeneration systems would also be eligible for this[b]
Onsite PV	Under the recently passed energy bill consumers can receive a credit of up to 30% of the cost (to a maximum of $2,000), for installing solar-powered hot-water systems except for purposes other than heating swimming pools and hot tubs.[b] The law also applies to solar PV
Onsite wind	
Small hydro	
	Other
DE in general	PURPA of 1978 provided the foundation for competition in the electricity sector allowing DE investors to participate in the market for the first time. The Energy Act of 2005 makes several amendments to PURPA including the retraction of the mandate that utilities must buy power form "qualifying facilities" (which included many cogeneration facilities). Section 1253 states that as long as the utility can prove that the facility has fair access to the grid, there is no longer may obligation to buy power (nondiscriminatory access to the grid is guaranteed in section 1231). The Bill also requires all utilities to offer net metering to small onsite generators (section 1251) and offer time-differentiated rate schedules upon request so that customers can take advantage of technologies such as "smart metering" which also must be supplied by the utility (section 1252). And, without proving any clear guidance, the utilities are disallowed from relying on a single fuel, i.e., there is a Federal mandate to diversify energy portfolios At least 28 states in the U.S. have now enacted a Renewable Portfolio Standards (RPS), including[a] various states that include cogeneration in the mandate, and at least one (Pennsylvania) that includes conservation (DSM) in its mandate. Several initiatives arose from the Bonn 2004 Renewable energy conference, including a renewable energy production incentive, wind power and solar PV, and geothermal energy cost targets.[d] According to section 202 of the Energy Act of 2005, the Federal government must ensure no less than 3% of its total energy consumption is derived from renewables rising by 2.5% every 3 years to 2013 There is a Federal mandate to study the potential from renewable energy (section 201). There is a goal to develop at least 10,000 MW of nonhydro renewables on public lands by 2015
Large-scale cogeneration	The USDOE and the EPA have set ambitious targets for cogeneration. The USDOE aims to double current cogeneration capacity to 92 GW by 2010.[e] Clean standards as well as tax credits are being considered by the legislature. Some states (Nevada, the Dakotas, and Pennsylvania) include cogeneration in their RPS. The EPA has recently proposed output-based standards for air permits for new cogeneration applications that would spur cogeneration investment via reduced compliance costs

(continued)

TABLE 2.10 *(Continued)*

	Section 1817 of the Act calls for a federal study into the benefits of "distributed generation" to be undertaken and the results presented to the senate no later than 2007[b]
Domestic cogeneration	
PV	The much-publicized million solar roof initiative remains in force and is designed to build a network of local organizations and partners to develop solar energy demonstration projects throughout the country.[f] Most US states have some form of incentive at a state level to promote the use of PV. The Federal government has a goal of 20,000 solar roofs by 2010 (section 204)
Wind	
Hydro	

[a] Database of State Incentives for Renewable Energy 2005.
[b] Energy Policy Act of 2005. http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=109_cong_bills&docid=f:h6enr.txt.pdf
[c] Renewables 2005. Global Status Report, p. 21.
[d] Renewables 2004. *International Conference for Renewable Energies*, 1–4 June 2004, Bonn. Conference Report: Outcomes and Documentation—Political Declaration/International Action Programme/Policy Recommendations for Renewable Energies.
[e] USA cogeneration Roadmap.
[f] http://www.millionsolarroofs.com/

TABLE 2.11 Argentinean Distributed Generation Policy by Technology and Type[a]

	Technical
DE in general	Ley 24065 guarantees access to the grid for generators including investors in DE. All generators must meet the same technical standards to participate in the power pool, including reactive power, frequency control, etc., and must bear the cost of real time metering equipment[b]
Large-scale cogeneration	Rules have been established to allow larger DE investors to take advantage of seasonal, weekly, and daily electricity markets (including spot markets) with the argentine power pool, the Mercado Electrico Mayorista MEM, required to take and provide fair compensation for power fed into the grid
Domestic cogeneration	
PV	
Wind	
Hydro	
	Financial
DE in general	
Large-scale cogeneration	The electricity pool price is a fixed price reset every 6 months by CAMMESA and approved by the Energy Secretary to supply, demand, available capacity, and other factors. The seasonal price is maintained for at least 90 days. Generators are paid for dispatched energy and also for the capacity made available and approved by CAMMESA. Locational pricing rewards generators in areas of high congestion.[a] CHP investors are charged a fixed rate based on capacity to participate in the power pool, but can opt out of the charge if they do not intend to sell excess power. The proposed Salvatori's law (article 9) grants income tax exemption for the total capital cost of "renewable energy projects" as well as value-added tax exemption and import duty exemption.[c] Also accelerated amortization is established for renewable energy investments and investors are guaranteed a stable investment climate for a minimum of 20 years (article 10)

(continued)

TABLE 2.11 ***(Continued)***

Domestic cogeneration	
Onsite PV	The PERMER project, cofinanced by World Bank, GEF, and local governments (national and provinces) promote investment in solar panels for rural households and small public buildings[d]
Onsite wind	
Small hydro	
	Other
DE in general	The Energy Act (Law No. 24065) of January 1992 provided the legal foundation for separating the electricity industry into three sectors: generation, transmission, and distribution. Transmission companies may charge a transmission tariff, but cannot own generation of distribution assets.[a] Generation is fully competitive with each company paying a tariff for T&D. Distribution companies operate as geographic monopolies and are regulated. They may buy the energy needed to meet supply contracts from either the wholesale markets or bilateral contracts with generators. The wholesale spot market is managed by Wholesale Electricity Market Operator Company (CAMMESA).[a] DE owners of more than 1 MW capacity can apply to participate in the power pool market[b]
	The Energy Secretariat (SE) has established a fund to support emergency expansion of the transmission system to relieve transmission constraints, it is unclear if the fund would be made available also to support strategic investment in DE to reduce constraints[a]
	There is currently a law that has been approved by the Senate and is awaiting approval from the Lower Chamber called the Law of Promotion of Renewable Energy Sources (Salvatori's Law). It declares that "the generation of electricity from renewable sources is the national interest" and once it is approved, it will require 8% of the overall national demand of electricity in 2013 to be met using renewable sources.[c] It aims also to encourage national manufacturing capacity and use renewables as an economic driver for the country. Increasing human resources in the renewable energy field is an explicit goal of the legislation. Approvals for renewable energy projects from the National Executive Authority (PDE) are waived. Argentina has used feed-in tariffs and renewable portfolio standards[e]
Large-scale cogeneration	There are three types of large users as defined by the 1992 law: major, minor, and private large users[a] all are guaranteed grid access
Domestic cogeneration	
PV	
Wind	
Hydro	

[a] The New Electricity Supply Industry in Argentina and Chile, Mario Vignolo Facultad de Ingeniería—IIE Montevideo—Uruguay.
[b] Annex 12 Autogeneradores y cogeneradores version 06/Oct/1999, CAMMESA.
[c] Bill no. S-1221/03.
[d] Personal communication with Jorge Bauer.
[e] http://www.renewables2004.de/pdf/policy_recommendations_final.pdf

biomass and onsite renewables will likely gain added attention and perhaps policy support as a result (Table 2.13).

2.8.8 Czech Republic

The Czech Republic's DE heritage remains largely as a leftover from the Soviet era. There is much existing cogeneration, especially in district heating applications, but much of it will be in need of refurbishment in the following years. In the early 1990s, Czech saw a boom in the development of smaller cogeneration plants and policy is now playing an important part in once again fostering

TABLE 2.12 Brazilian Distributed Generation Policy by Technology and Type

	Technical
DE in general	All DE plants, including gas and biomass are guaranteed a market for surplus power based on law 10848
Large-scale cogeneration	
Domestic cogeneration	
PV	
Wind	
Hydro	
	Financial
DE in general	Law 10848 guarantees DE systems access to the grid and are guaranteed the average price of the larger market
	The PROINFA (Alternative Sources of Energy Incentive Programme) program seeks to establish 1100 MW from each wind, biomass and small hydro. Purchase contracts for qualifying facilities would have their contracts guaranteed by Electrobras for 20 years.[a] Qualifying projects are also eligible for grants from the National Economic and Social Development Bank (BNDES)
	The "Luz para Todos" program aims to subsidize the electrification of some 2.5 million households in Brazil by 2008 and there is a target to employ renewable energy for at least 200,000 of those households[b]
Large-scale cogeneration	Biomass CHP plants are expected to benefit from the PROINFA program
Domestic cogeneration	
Onsite PV	
Onsite wind	
Small Hydro	
	Other
DE in general	In 2004, the Brazilian government passed law 10848 with regards to electricity sector reform.[c] The law in effect created two separate market structures for electricity exchange: the bilateral contract environment (ACL), where individual generators and consumers can negotiate power purchase agreements (PPAs), and the regulated contract environment (ACR) where distribution companies must purchase the power they need to meet their contract from public auctions. The Decree 5163 dealt specifically with distributed generation, and states that in the ACR scenario distribution companies must also buy power from "alternative sources" at prices set by the government and are permitted to buy up to 10% of their required supply. The clause was designed to create incentive for CHP and other designated alternative sources. The decree states specifically that cogeneration plants with an efficiency of more than 75%, as well as hydro plants with a capacity below 30,000 kW and other renewable sources all qualify as "alternative energy" for the purposes of the law. The additional costs incurs by the grid owners are, by law, built into the formula used to calculate end user tariffs[d]
	The new law requires each distribution company to disclose 5 year plans for power supply and up to 10% of the supply needed to meet total loads can come from direct bilateral agreements with cogeneration plants instead of the typical power auctions[c]
	In addition, the revised auction structure is likely to favor cogeneration plants because winning tenders for supply must have supply on line within 3 years too short a lead time for most central plants
	Brazil has set a target of 3.3 GW of capacity from wind, biomass, and small hydro as a national target for 2006[b]
Large-scale cogeneration	Under the New Electricity Law,[c] gas-fired cogeneration would be eligible for incentives if owners could prove that overall efficiency was 75% or greater. Efficiency would be based on an efficiency index created for the purpose[d]

(continued)

TABLE 2.12 ***(Continued)***

Domestic cogeneration	
PV	
Wind	
Hydro	Reduced transmission wheeling fees have been credited as a major cause for a successful small hydro industry,[e] Brazil's Hydropower Program was mentioned as a commitment in the Bonn 2004 Renewable Conference proceedings[f]
Domestic cogeneration	
PV	
Wind	
Hydro	

[a] Brazil Weighs up New Potential. Antonio Maia Cogeneration and Onsite Power Production magazine, July–August 2005.
[b] Renewables 2005. Global Status Report, Eric Martinot, p. 25.
[c] WADE Market Analysis, Brazil. 2005.
[d] Brazil Weighs up New Potential. Antonio Maia Cogeneration and Onsite Power Production magazine, July–August 2005.
[e] Grid-Based Renewable Energy in Developing Countries: Policies, Strategies, and Lessons from the GEF World Renewable Energy Policy and Strategy Forum June 13–15, 2002, Berlin, Germany. Eric Martinot Global Environment Facility, Washington, DC.
[f] Renewables 2004. *International Conference for Renewable Energies*, 1–4 June 2004, Bonn. Conference Report: Outcomes and Documentation—Political Declaration/International Action Programme/Policy Recommendations for Renewable Energies.

TABLE 2.13 Chilean Distributed Generation Policy by Technology and Type

	Technical
DE in general	The "Superintendencia de Electricidad y Combustibles" (Secretariat of Electricity and Fuels) sets and enforces the technical standards of the system. The Chilean Electricity Law guarantees open access to the transmission system, but must contribute to investment in the "system." It is unclear if DE is included in the system
Large-scale cogeneration	
Domestic cogeneration	
PV	
Wind	
Hydro	
	Financial
DE in general	According to the "Ley Corta"[a] of 2004 distribution operators are obliged to purchase all power from such systems at a "fair price." The National Energy Commission (CNE) sets tariffs and node prices and to prepares the 10-year expansion strategy for the electric system. Prices are partially based on geographic location of generators as reflected in the methodology for calculating the Nodal Price for Electricity[b]
Large-scale cogeneration	One proposal suggested rewarding gas-fired plants capable of fuel switching higher prices for power than those that only use gas[c]
Domestic cogeneration	
Onsite PV	The government sponsored Rural Electrification Program, in cooperation with private utility Emelectric runs a program that pays up to 90% of the capital cost of the installation PV in rural areas, the remaining 10% being paid by the end users and Emelectric. The end users will pay a fixed tariff of US$8/month for a 160 Wp system that provides 220VAC/50 Hz electricity[d]
Onsite wind	
Small hydro	
	Other
DE in general	There is currently a proposal before the Chilean Senate to mandate that 5% of all power be generated via renewable sources. In the past, the Chilean government has financed various pilot projects in renewable energy and is currently a partner with the UNDP in the "Renewable Technologies for Rural Electrification project." The 2004 "Ley Corta" (an amendment to the Chilean Electricity Law) was designed to remove barriers to the incorporation of renewable electricity in the system[a]

(continued)

TABLE 2.13 *(Continued)*

	The 1982 Energy Act law DFL No. 1 provided the legal foundation for separate generation, transmission, and distribution, including wholesale competition in the generation side Smaller generators are not allowed representation on the Centro de Despacho Económico de Carga (Centre for Economic Load Dispatch) committees which has raised some questions of fairness[b] The Chilean National Energy Commission (CNE) is responsible for implementing regulations outlined in Chile's General Electricity Services Law. It proposes regulated tariffs twice a year in April and October, and prepares the overall plan for new generation capacity. The SEC is responsible for ensuring all generators, distributors, and transmission operators are in compliance with the law[a]
Large-scale cogeneration	The "Ley Corta" law also sets new efficiency standards which may benefit cogeneration[a]
Domestic cogeneration	
PV	
Wind	
Hydro	

[a] Law no. LEY-19940, De los Sistemas de Transporte de Energía Eléctrica, March 2004.
[b] The New Electricity Supply Industry in Argentina and Chile, Mario Vignolo, Facultad de Ingeniería—IIE Montevideo—Uruguay.
[c] http://www.eia.doe.gov/emeu/cabs/chile.html
[d] Non-grid Renewable Energy Policies: International Case Studies, August 16 2001. JP Ross, Center for Resource Solutions.

cogeneration investment in the country. Like other EU accession countries, efforts to comply with EU environmental and competitiveness standards will be key drivers for DE in the Czech Republic (Table 2.14).

2.8.9 Denmark

Denmark is one of the world's leading renewable energy champions and has had strong policy in support of renewables longer than most countries. Furthermore, Denmark is one of the few countries with policies established to specifically encourage wind cooperatives. Denmark is also one of the world's leaders in cogeneration. Despite its strong track record in DE, there is still much room for improvement and the evolution of Denmark's policy on DE reflects this. Economic competitiveness and energy security could prove to be the most important drivers for DE along with environmental concerns (Table 2.15).

2.8.10 Germany

Germany is one of the world's leaders in renewables with especially strong growth over the last decade. There is evidence that the German policy makers continue to work at creating the right policy environment for DE to thrive. Nevertheless, there is still much unrealized potential especially for small onsite renewable applications and cogeneration, whereas most policies to date have focused on larger scale, often remote, renewables. Climate change negotiations and security of supply concerns will continue to be key drivers for DE in the following years (Table 2.16).

2.8.11 Hungary

The majority of Hungary's existing DE capacity is in the form of district heating applications, though many have already reached the end of their life. Many promising policies have recently been adopted in Hungary, yet it remains to be seen how successful they will prove for increasing investment in DE. Hungary still has the lowest target for renewable energy of all the European Union member states. Hungary's accession to the EU will continue to prove the most important driving force in energy policy (Table 2.17).

TABLE 2.14 Czech Distributed Generation Policy by Technology and Type[a]

	Technical
DE in general	
Large-scale cogeneration	All DE systems are guaranteed "nondiscriminatory" access to the grid
Domestic cogeneration	A European standard for micro-cogeneration is currently being drafted under the auspices of Cenelec technical committee TC8X WG2, and coordinated by the Dutch National Standards Body, NEN. It will be based on the CEN Workshop Agreement CWA 14642
PV	
Wind	
Hydro	
	Financial
DE in general	All DE is guaranteed to be purchased by the utility. The rates depend on technology. In addition, distributed technologies are awarded an extra amount depending on what voltage level they are interconnected to: low voltage = 2.13 €/MWh, medium voltage = 0.90 €/MWh, and high voltage = 0.66 €/MWh
Large-scale cogeneration	The State Program to Support Energy Saving and Use of Renewable and Secondary Sources was established in 1991 to develop efficiency including cogeneration.[b] The program is still going and includes education programs and training. In 2000, the Czech parliament passed a law (458/2000 Coll.) requiring all boilers above a certain to be cogeneration. Chapter IV of a 2001 law (Act on Energy Management of the Czech Republic No. 406/2000) deals specifically with cogeneration and undertakes some measures to promote it.[b] Ministerial Decree (No. 252/2001 Coll.) of 2001 also provided a framework for power administration and tariff structures for cogeneration and renewables. The decree at that time required network owners to purchase power from cogeneration plants, but the tariff was not set but had to be negotiated. Heat had also to be purchased under the decree[b]
	Utilities must purchase electricity generated from cogeneration units. The EU cogeneration directive has been implemented in the new energy law of January 1, 2005. The new law guarantees a tariff for only surplus power from cogeneration, in other words an amount of power equal to onsite loads cannot be fed into the grid. The regulatory authority (ERU) decides what is a reasonable load for the cogeneration facility and then any surplus earns the feed-in rate. This, in part, guarantees a better price from cogeneration-generated electricity. Purchase price different for up to 1 MW and between 1 and 5 MW capacity and can be on a single price per 24 h/day basis or a different price for 8 h and the rest of the day. This law is expected to evolve in the near future
	EU emissions trading scheme will have an uncertain effect on cogeneration investment
Domestic cogeneration	In theory, domestic cogeneration would fall under 1 MW category and therefore receive a higher guaranteed price
Onsite PV	A feed-in tariff of 201 €/MWh[a] is guaranteed for PV power in the Czech Republic. Tax breaks and grant program for private systems up to 2 kWp and legal entities up to 20 kWp systems[c]
Onsite wind	A feed-in tariff of 86 €/MWh is guaranteed for wind power in the Czech Republic[a]
Small hydro	A feed-in tariff of 68 €/MWh is guaranteed for small hydro power in the Czech Republic[a]
	Other
DE in general	
Renewables in general	March/April 2005 law passed to promote renewables. Sets out an 8% target of total gross consumption by 2010. Guarantees 15 years of feed-in tariff for renewables, a premium for green electricity will also be paid. Prices are not set in law but are to be determined by the Energy Regulatory office.[d] The Law[e] has been approved by the senate but does not yet have a number and will not come into effect until it has been endorsed by the president. Currently, the utility CEZ holds about two-thirds of the generation market in the republic and also owns a majority stake in six out of the eight existing distribution companies. The government owned CEPS operates the transmission system

(continued)

TABLE 2.14 *(Continued)*

	As a whole, European policy calls for 12% of energy and 21% of electricity supply from renewables by 2010 including 1,000,000 PV systems, 10,000 MW wind and 10,000 MW biomass[f]
	At Bonn Czech Republic committed to the development of a Renewable Energy Plan[g]
Large-scale cogeneration	EU has a goal to double cogeneration use by 2010[h]
	The Czech Republic is bound by the EU cogeneration Directive 2004/8/EC which states that member states are obliged to address key market barriers such as ensuring grid access for cogeneration. The directive also states that in order for cogeneration to be considered high efficiency, it must provide an energy saving of at least 10% compared to separate production of heat and power. Harmonized efficiency reference values are to be released no later than February 2006. The directive further states that member states must establish guarantee of origin standards no less than 6 months after the harmonized reference values are released
Domestic cogeneration	EU cogeneration directive
PV	
Wind	
Hydro	

[a] Position of DE and cogeneration in the Czech Market. Former, Present and Future. Josef Jelecek Cogen Europe Conference 2005.

[b] IEA energy efficiency database: http://www.iea.org/textbase/pamsdb/detail.aspx?mode=ee&id=218

[c] Jager-Waldau, A., Ossenbrink, H., Scholz, H., Bloem, H., and Werring, L. 2004. JRC Ispra & DG TREN, *19th European Photovoltaic Solar Energy Conference and Exhibition*, Paris, June 2004.

[d] http://www.edie.net/news/news_story.asp?id=9710&channel=6

[e] Podpora výroby elektřriny z obnovitelných zdrojů energie.

[f] Directive 2001/77/EC of the European Parliament and of the Council of 27 September 2001 on the promotion of electricity produced from renewable energy sources in the internal electricity market. Energy for the Future: Renewable Sources of Energy White Paper for a Community Strategy and Action Plan COM(97)599 final (26/11/1997).

[g] Renewables 2004. *International Conference for Renewable Energies*, 1–4 June 2004, Bonn. Conference Report: Outcomes and Documentation—Political Declaration/International Action Programme/Policy Recommendations for Renewable Energies.

[h] A community strategy to promote combined heat and power (CHP) and to dismantle barriers to its development. European Commission Communiqué COM (97) 514 final.

TABLE 2.15 Danish Distributed Generation Policy by Technology and Type

	Technical
DE in general	
Large-scale cogeneration	
Domestic cogeneration	A European standard for micro-cogeneration is currently being drafted under the auspices of Cenelec technical committee TC8X WG2, and coordinated by the Dutch National Standards Body, NEN. It will be based on the CEN Workshop Agreement CWA 14642
PV	
Wind	
Hydro	
	Financial
DE in general	Denmark established the first feed-in law requiring utilities to buy electricity from onsite renewable generators at a set price in 1999.[a] A separate program, the Energy Research Program, supports actual implementation of innovative energy projects by providing grants for up to 100% of project cost. Projects involving oil and natural gas, environmentally benign heat and power production, wind, green buildings, solar energy, and energy efficiency are eligible source IEA database. Net metering is in affect for small onsite power systems, and the policy is to be reviewed in 2006[a]

(continued)

TABLE 2.15 ***(Continued)***

Large-scale cogeneration	A state subsidy for small cogeneration was introduced in 1992 for all plants fueled by waste incineration, natural gas, and renewables. At the start of the program, the tariff was 10 øre/kWh but was subsequently reduced to 7 øre/kWh, for all but plants smaller than 3 MW[b] Denmark has recently updated the way it provides incentive for CHP. Whereas before distribution companies were obligated to purchase all power from CHP facilities via a feed-in tariff, the feed-in tariff has been phased out with the 2005 market reforms to better reflect free market principles. Now, CHP plants are guaranteed the right to sell power to the grid but only if they can find interested buyers.[c] CHP plant operators are then awarded a fixed subsidy (1 €ct/kWh) per month based on a 20 operating life. Higher level of support for installations using biomass and biogas. Plants of more than 20 MW are eligible for EU C02 quotas as of 2005. In response to the EU Emissions Trading Directive and the Cogeneration Directive, Denmark has allocated allowances based on historic production for electricity: $1.68 \times 125/3 = 70$ allowances For heat = 20 allowances; total per year = 90 allowances.[c] This has the effect of making CHP more economically attractive
Domestic cogeneration	
Onsite PV	
Onsite wind	Denmark has for some time encouraged wind cooperatives with its "wind energy cooperative tax incentive." Benefits of the program include tax exemption for the first €400 in a year and 60% of the regular rate for anything thereafter. New windmills are guaranteed a fixed settlement price (feed-in tariff) of DKK 0.33/kWh for the first 22,000 full-load hours, about 10 years
Small hydro	
DE in general	Energy liberalization began in 1996 in Denmark and was as of 2003, all electricity customers are free to choose their supplier.[a] Denmark's overall goal is to increase its share of renewables from 8% in 1996 to 30% in 2025.[a] Municipal waste and geothermal energy receive special mention in the plan. Denmark has introduced a demand-side RPS meaning that all customers must source at least 20% of their electricity from certified renewables sources.[a] The system awards producers "green certificate" for every unit of electricity they generate using renewable resources and all renewable electricity is guaranteed a set feed-in price. As a whole, European policy calls for 12% of energy and 21% of electricity supply from renewables by 2010, including 1,000,000 PV systems, 10,000 MW wind, and 10,000 MW biomass[d]
	Other
Large-scale cogeneration	The Danish Energy Authority, and two main utilities have developed a research and development strategy for a range of renewable energies, including fuel cells, biomass, wind energy, and photovoltaics. Strategies for biofuels, wave energy, hydrogen, and system integration are being developed[b] The 1988 Heat Supply Act encourages renewable fueled district heating using by prohibiting electric heating in specified residential areas[b] EU has a goal to double cogeneration use by 2010[e] Denmark is bound by the EU cogeneration Directive 2004/8/EC which states that member states are obliged to address key market barriers, such as ensuring grid access for cogeneration. The directive also states that in order for cogeneration to be considered high efficiency, it must provide an energy saving of at least 10% compared to separate production of heat and power. Harmonized efficiency reference values are to be released no later than February 2006. The directive further states that member states must establish guarantee of origin standards no less than 6 months after the harmonized reference values are released In 1993, the Danish government established a goal of 75 PJ of biomass energy by 2000 (what equaled approximately 10% of total energy demand). The objectives were not met, but were reconfirmed in 2000 to include a feed-in tariff of DKK 0.30/kWh for 10 years
Domestic cogeneration	
PV	A program has been established, funded by a carbon tax on electricity, which pays up to 40% of the total cost of a PV system, including materials cost and installation[a]

(continued)

TABLE 2.15 *(Continued)*

Wind	The Danish "Agreement on wind turbines" agreed in 1996 set a goal of 200 MW of non-utility owned wind capacity and an additional 1500 MW owned by private and public utilities.[a] The goal was met 5 years ahead of schedule and the wind market remains strong
Hydro	

[a] Policy Review of Renewable Energy Sources and Energy Efficiency in the European Union and its Member States, June 2005. Reference #EuropeAid/116832/D/G/Multi.

[b] IEA policy database: http://www.iea.org/textbase/pamsdb/search.aspx?mode=re

[c] Supporting Clean Heat and Power on the Internal Energy Market, Jesper Lorentzen. The Danish Energy Authority, Cogen Europe 2005.

[d] Directive 2001/77/EC of the European Parliament and of the Council of 27 September 2001 on the promotion of electricity produced from renewable energy sources in the internal electricity market. Energy for the Future: Renewable Sources of Energy White Paper for a Community Strategy and Action Plan COM(97)599 final (26/11/1997).

[e] A community strategy to promote combined heat and power (CHP) and to dismantle barriers to its development. European Commission Communiqué COM (97) 514 final.

TABLE 2.16 German Distributed Generation Policy by Technology and Type

	Technical
DE in general	The Federal Ministry of Education and Research heads up a program aimed at integrating DE successfully into networks.[a] Federal states also undertake their own research funding
Large-scale cogeneration	
Domestic cogeneration	A European standard for micro-cogeneration is currently being drafted under the auspices of Cenelec technical committee TC8X WG2, and coordinated by the Dutch National Standards Body, NEN. It will be based on the CEN Workshop Agreement CWA 14642 Germany is also developing their own regulations for the installation of micro-cogeneration
PV	
Wind	
Hydro	
Biomass	
	Financial
DE in general	A green tax introduced in 1999 increased the price of fuel oil, natural gas, gasoline/diesel, and electrical power (including renewables) every year between 1999 and 2003[a]
Large-scale cogeneration	Incentives exist for municipal cogeneration and sub 2 MWe cogeneration plants.[b] A law was passed in 2001 which provide a financial bonus on top of feed-in tariffs to preexisting CHP plants and new plants under 2 MW capacity.[c] In Germany, CHP plants are exempt from certain fuel taxes, including the "ecotax" and mineral oil tax and the environmental tax for biofuels in heating.[c] There is a special feed-in rate set for CHP in Germany, including a premium payment incrementally decreasing overtime. There is also an additional bonus for biomass fired CHP. In addition, there is an investment subsidy for biomass fired plants.[c] The EU Emissions trading scheme will have an uncertain effect on cogeneration investment
Domestic cogeneration	
Renewables in general	Germany's 1991 "Electricity Feed-In Law (EFL)" required utilities to purchase electricity generated by renewable-energy systems with a fixed tariff. For wind and PV the tariffs were set to 90% of average electricity retail prices and adjusted every year according to price changes. In 2000 the EFL was replaced by the Renewable Energy ACT (REA), a system of guaranteed 20-year production incentives for renewables based on the cost of the type of system and its operation. So far, the tariff has benefited mostly wind and PV but biomass is also to benefit from the tariff

(continued)

TABLE 2.16 ***(Continued)***

Onsite PV	Feed-in rate guaranteed for 20 years. The feed-in tariff for PV in Germany began at 0.51 €/kWh in 2000 and 2001 and is decreasing annually by 5% (it was 0.457 €/kWh in 2003).[d] There are additional feed-in benefits for PV applications that are building integrated or used as Sound barriers of more than 0.1 €/kWh.[e] The 100,000 solar roofs initiative of the Federal government was launched in 1999 and ended in 2003, the program provided soft loans for grid-connected systems
Onsite wind	Feed-in rate currently 6–9 €ct/kWh
Small hydro	Feed-in rate set at 6.6–7.7 €ct/kWh
	Other
DE in general	The German government has played a lead role in the renewable energy policy network, an international policy network arising from the Bonn 2004 renewable energy conference.[f] Disclosure law states that all generators must state generation portfolio, including cogeneration and renewables. A target of 25% of electricity from wind by 2025 has been set, and another half of its total energy needs met via renewables by 2050. Germany plans to close all 19 of the country's nuclear-power plants by around 2025. The renewable energy law states that twice as much electricity must be generated by renewable sources by 2010 than in 2000 (i.e., about 12.5% compared to about 6.3% in 2000)[a]
	As a whole, European policy calls for 12% of energy and 21% of electricity supply from renewables by 2010 including 1,000,000 PV systems, 10,000 MW wind, and 10,000 MW biomass[g]
Large-scale cogeneration	During the process of electricity market restructuring in Germany, the main goal with respect to CHP was to maintain the existing capacity and renovate existing plant. Although Germany has more CHP than any other European nation very few policy incentives exist for new plant. EU has a goal to double cogeneration use by 2010.[h] In 1999, Germany established a quota system initially requiring 10% of total electricity production to come from CHP increasing to 20%
	Germany is bound by the EU cogeneration Directive 2004/8/EC, which states that member states are obliged to address key market barriers, such as ensuring grid access for cogeneration. The directive also states that in order for cogeneration to be considered high efficiency, it must provide an energy saving of at least 10% compared to separate production of heat and power. Harmonized efficiency reference values are to be released no later than February 2006. The directive further states that member states must establish guarantee of origin standards no less than 6 months after the harmonized reference values are released
Domestic cogeneration	
PV	
Wind	
Hydro	At the 2004 renewable energy conference in Bonn, Germany announced its Mini-Hydro Programme[i]

[a] http://www.oja-services.nl/iea-pvps/countries/germany/index.htm
[b] WADE World Survey for Decentralized Energy 2005, p. 18.
[c] Cogen Europe. International Benchmarking Report on the Status of CHP in other EU Member States. August 2004.
[d] http://www.oja-services.nl/iea-pvps/countries/germany/index.htm
[e] Jager-Waldau, A., Ossenbrink, H., Scholz, H., Bloem, H., and Werring, L. 2004. JRC Ispra & DG TREN, *19th European Photovoltaic Solar Energy Conference and Exhibition*, Paris, June 2004.
[f] Anniversary of renewables 2004. Status of the Implementation of the Conference Results. BMU.
[g] Directive 2001/77/EC of the European Parliament and of the Council of 27 September 2001 on the promotion of electricity produced from renewable energy sources in the internal electricity market. Energy for the Future: Renewable Sources of Energy White Paper for a Community Strategy and Action Plan COM(97)599 final (26/11/1997).
[h] A community strategy to promote combined heat and power (CHP) and to dismantle barriers to its development. European Commission Communiqué COM (97) 514 final.
[i] Renewables 2004. *International Conference for Renewable Energies*, 1–4 June 2004, Bonn. Conference Report: Outcomes and Documentation—Political Declaration/International Action Programme/Policy Recommendations for Renewable Energies.

TABLE 2.17 Hungarian Distributed Generation Policy by Technology and Type

	Technical
DE in general	
Large-scale cogeneration	
Domestic cogeneration	
PV	
Wind	
Hydro	
	Financial
DE in general	The 2001 Electricity Act established feed-in tariffs for renewable generators more than 100 kW capacity. There is a fixed feed-in tariff of 0.1/0.062 €/kWh for all technologies though it is adjusted annually for inflation. If the project connects at the transmission level, the major generator (MVM) pays the tariff whereas the local wires owner pays if the project is connected at the distribution level. Subsidies also exist for renewable projects.[a] The Ministry of Economy and Transport's Szëchenyi Plan provides grants for renewable energy projects up to 30% of the total project cost. The grants are available through competitive bids and have an upper limit for project size.[b] The Environment Protection and Infrastructure Operational Programme also makes grants for renewable energy projects
Large-scale cogeneration	In January 2003, a feed-in tariff was established for power generated from cogeneration facilities between 100 kW and 50 MW with an efficiency of more than 65% (Economy and Transport Ministry order 56/2002-29.XII).[b] For gas engines, the efficiency must be higher than 75%. EU emissions trading scheme will have an uncertain effect on cogeneration investment
Domestic cogeneration	The SZT-LA 2 Programme[b] which provides grants to improve the energy efficiency of residential buildings could in theory apply to micro-cogeneration investments
Onsite PV	
Onsite wind	
Small hydro	
	Other
DE in general	In 1999, the Hungarian government implemented a wide range of policy reforms under the title Hungarian Energy Policy Principles.[c] The main thrust of the reforms was to create competitive gas and electricity markets in order to meet EU directives and increase energy security and reliability. DE was also identified as desirable in the policy for its environmental benefits. As a whole, European policy calls for 12% of energy and 21% of electricity supply from renewables by 2010 including 1,000,000 PV systems, 10,000 MW wind, and 10,000 MW biomass.[b] Hungary has a national indicative target of 3.6% by 2010; there is also a government goal of increasing the percentage to 7% by the same year. The Energy Saving and Energy Efficiency Action Programme, established in 1999 also sets concrete goals of increasing renewable energy production from 28 PJ/year (in 1999) to 50 PJ/year, the plan focuses on biomass, waste, geothermal, and solar.[b] Hungary is a signatory of the Kyoto protocol and committed to reduce its carbon emissions by 6% by 2012
Large-scale cogeneration	EU has a goal to double cogeneration use by 2010[d]
	Hungary is bound by the EU cogeneration Directive 2004/8/EC which states that member states are obliged to address key market barriers, such as ensuring grid access for cogeneration. The directive also states that in order for cogeneration to be considered high efficiency, it must provide an energy saving of at least 10% compared to separate production of heat and power. Harmonized efficiency reference values are to be released no later than February 2006. The directive further states that member states must establish guarantee of origin standards no less than 6 months after the harmonized reference values are released

(continued)

TABLE 2.17 ***(Continued)***

Domestic cogeneration
PV
Wind
Hydro

[a] Jager-Waldau, A., Ossenbrink, H., Scholz, H., Bloem, H., and Werring, L. 2004. JRC Ispra & DG TREN, *19th European Photovoltaic Solar Energy Conference and Exhibition*, Paris, June 2004.

[b] Directive 2001/77/EC of the European Parliament and of the Council of 27 September 2001 on the promotion of electricity produced from renewable energy sources in the internal electricity market. Energy for the Future: Renewable Sources of Energy White Paper for a Community Strategy and Action Plan COM(97)599 final (26/11/1997).

[c] Policy Review of Renewable Energy Sources and Energy Efficiency in the European Union and its Member States, June 2005. Reference #EuropeAid/116832/D/G/Multi.

[d] A community strategy to promote combined heat and power (CHP) and to dismantle barriers to its development. European Commission Communiqué COM (97) 514 final.

2.8.12 Portugal

Since joining the EU in 1986, Portugal economy has been growing steadily and a recently renewed policy focus on energy conservation promises to increase Portugal's competitiveness and further spur growth. Improving diversity has been a key objective of energy policy in the following years. Drivers in the future are likely to be attaining European energy standards, mitigating environmental impact, and economic growth (Table 2.18).

2.8.13 United Kingdom

The United Kingdom was the first in Europe to experiment with competitive electricity markets and some DE investment arose as a direct result of that policy. Like other European nations key DE drivers in the near future for the UK remain climate change and energy security. UK has pioneered some interesting policy in defining high quality cogeneration but there is much more room for policy reform to spur investment in DE to meet environmental and national security objectives (Table 2.19).

2.8.14 China

Market potential for DE in China is enormous but clear policies will be required in order for it to be realized. Demand for electricity has been growing at rates and scales that dwarf those of Europe and North America, with capacity installed in 2004 exceeding the total capacity of the United Kingdom. Increasingly high electricity prices, rotating blackouts, coal shortages, emerging gas availability, and the growing appeal of renewables for their environmental benefits are all factors that will drive Chinese investment in DE. Policies to promote renewables have recently had a high profile but other forms of DE still require attention. International Climate Change negotiations may also rise in influence in Chinese energy policy (Table 2.20).

2.8.15 India

India is sure to be a major market for DE in the following years given its large population and the urgent need for investment in the power sector in order to meet growing demand for power. Good progress has been made in renewable energy: India leads the developing world in wind capacity. Policy with a focus on DE, however, has been more limited but major progress has been made in recent years in developing onsite domestic biogas electricity, PV, and industrial biomass cogeneration. Strong policy signals will be required to ensure nascent gas supplies are used optimally in DE applications (Table 2.21).

TABLE 2.18 Portuguese Distributed Generation Policy by Technology and Type

	Technical
DE in general	Any system up to 150 kWe can be connected to the distribution level grid (<1 kV) and simplified interconnection rules exist. Technical requirements consist mainly in the compliance with short-circuit current/power of the connection installation, maximum and minimum values for frequency and voltage, and maximum time for disconnection from the grid in case of abnormal operation. Utility workers must be able to disconnect the unit using an outside switch for safety reasons. Regional authorities grant licenses for low voltage interconnection as applications are received. Even more simplified rules are being developed to include systems up to 10 kW capacity. Procedures for connecting systems up to 150 kW capacity are much more complicated and require an engineering study of how the DE resource will affect grid operation. Such applications must be approved by the Direcção Geral de Geologia e Energia (DGE), and proposals can only be put forward upon invitation from the DGE
Large-scale cogeneration	
Domestic cogeneration	A European standard for micro-cogeneration is currently being drafted under the auspices of Cenelec technical committee TC8X WG2, and coordinated by the Dutch National Standards Body, NEN. It will be based on the CEN Workshop Agreement CWA 14642
PV	
Wind	
Hydro	
	Financial
DE in general	Many costs incurred by investors in renewable energy investments and a wide array of efficiency measures, including cogeneration, are eligible for up to 50% refund under Portugal's PRIME program.[a] Eligible costs include materials, feasibility studies/research, field trials, land, and transport of materials, among others. Individuals are ineligible for the PRIME program but are eligible for tax breaks for solar and biomass projects Renewables and cogeneration of any scale are eligible for generous feed-in tariffs. Both technologies fall into the "special regime" category of the national electricity system (established under decree-law No. 339-C/2001) Under the scope of the National Plan for Climate Change, the government is considering a new tax regime that may favor DE applications[b]
Large-scale cogeneration	Decree 168/99 of May 1999 guaranteed attractive tariffs for cogeneration investors. Tariff rates were further refined in decrees 30/2000 and 31/2000 for units smaller than or greater than 10 MWe. For units larger than 10 MWe, payment is guaranteed for 120 months, whereas units smaller than 10 MWe are offered a more generous rate: in such cases, the feed-in price is multiplied by a "loss factor" of 1.04 for units between 5 and 10 MWe and 1.02 for units less than 5 MWe. Decree law 525/2001 further improved the situation for cogenerators that annually used more than 50% renewable fuels in a 2001 decree.[c] In addition, Decree Law 188/88 states that cogeneration plants receive assistance between 15 and 25% of the capital investment in the form of grant incentives Portugal is bound by the EU cogeneration Directive 2004/8/EC which states that member states are obliged to address key market barriers, such as ensuring grid access for cogeneration. The directive also states that in order for cogeneration to be considered high efficiency, it must provide an energy saving of at least 10% compared to separate production of heat and power. Harmonized efficiency reference values are to be released no later than February 2006. The directive further states that member states must establish guarantee of origin standards no less than 6 months after the harmonized reference values are released EU emissions trading scheme will have an uncertain effect on cogeneration investment
Domestic cogeneration	Until 2003, micro-cogeneration investors received a public grant for their system corresponding to up to 20% of total purchase and installation. This has now been phased out

(continued)

TABLE 2.18 ***(Continued)***

Onsite PV	Current buyback rates for PV in IPP applications are guaranteed for the lifetime of the system ∼0.30 €/kWh (systems more than 5 kWe) and 0.51 €/kWh (systems less than 5 kWe).[b] For non-IPP systems (household or businesses) up to 5 kWe, the buyback rate is 0.25 €/kWh. Ministerial resolution no. 63/2003 set a target of 150 MW PV capacity by 2010
Onsite wind	
Small hydro	
	Other
DE in general	Decree-Law DL 68/2002 defined the legal foundation for onsite generation in Portugal and established a definition of "microgeneration" (<150 kW). Onsite microgenerators are not permitted to export more than 50% of the total power their system generates in a year. The law also defines the tariff regime for microgeneration. Different premiums are paid depending on technology and they are reviewed on an annual basis[d]
	As a whole, European policy calls for 12% of energy and 21% of electricity supply from renewables by 2010 including 1,000,000 PV systems, 10,000 MW wind, and 10,000 MW biomass[e]
Large-scale cogeneration	Decree law 313/2002 states that cogeneration investors must be guaranteed fair access to the grid and that they need not shoulder any grid upgrades. EU has a goal to double cogeneration use by 2010[f]

[a] http://www.prime.min-economia.pt/
[b] http://www.oja-services.nl/iea-pvps/countries/portugal/index.htm
[c] http://www.lexportugal.com/LexPortugal/
[d] Personal communication with Joao Montez, April 2005.
[e] Directive 2001/77/EC of the European Parliament and of the Council of 27 September 2001 on the promotion of electricity produced from renewable energy sources in the internal electricity market. Energy for the Future: Renewable Sources of Energy White Paper for a Community Strategy and Action Plan COM(97)599 final (26/11/1997).
[f] A community strategy to promote combined heat and power (CHP) and to dismantle barriers to its development. European Commission Communiqué COM (97) 514 final.

TABLE 2.19 British Distributed Generation Policy by Technology and Type

	Technical
DE in general	Only licensed electricity suppliers can register metering systems for microgeneration
	The requirements for developers looking to connect small scale embedded generation (SSEG) rated at up to and including 16 A per phase are defined in Engineering Recommendation G83/1.[a] The Distributed Generation Coordinating Group is charged with studying and making recommendations to overcome technical barriers to micropower
Large-scale cogeneration	
Small-scale	Amendment P81 of the electricity trading rules allows small generators (16 A per phase on the low voltage 230 V single phase or multiphase 400 V supply) to use existing meters rather than having to use otherwise mandated half hour interval meters
Domestic cogeneration	A European standard for micro-cogeneration is currently being drafted under the auspices of Cenelec technical committee TC8X WG2, and coordinated by the Dutch National Standards Body, NEN. It will be based on the CEN Workshop Agreement CWA 14642. The UK has also developed changes in legislation, and published Engineering Recommendations in this area[b]
PV	Grid interconnection of PV systems rated above 5 kW is governed by G.59/1 last amended in 1995

(continued)

TABLE 2.19 *(Continued)*

Wind	
Hydro	
	Financial
DE in general	The climate change levy is a tax on energy introduced in 2001. All businesses have to pay 0.0043 GBP/kWh unless their supply comes from certified renewables or certified cogeneration facilities
	There are grant programs for renewable electricity[c]
Large-scale cogeneration	EU emissions trading scheme will have an uncertain effect on cogeneration investment. Cogeneration schemes that have been certified as "good quality" (see below) are exempt from the climate change levy, basically a carbon tax
Domestic cogeneration	In the 2005 budget, sales tax on domestic cogeneration units was reduced from 17.5 to 5%, in other words, it will get the same favorable tax treatment as other renewable technologies[d]
Onsite PV	Exempt from climate change levy. Each megawatt-hour generated via renewable energy also creates a tradable "renewable obligation credit" which can be sold on the market to generators who have not reached their legal obligation to generate 10% of their output from renewable. VAT on PV installations systems has been set at the reduced rate of 5% since April 2000
Onsite wind	Exempt from climate change levy. Each megawatt-hour generated via renewable energy also creates a tradable "renewable obligation credit" which can be sold on the market to generators who have not reached their legal obligation to generate 10% of their output from renewable
Small hydro	Exempt from climate change levy. Each megawatt-hour generated via renewable energy also creates a tradable "renewable obligation credit" which can be sold on the market to generators who have not reached their legal obligation to generate 10% of their output from renewable
	Other
DE in general	Utilities Act 2000 gives DE investors the right to participate in the market using cogeneration and/or renewable energy, including private wire networks. Energy Service Companies (ESCOs) are authorized to generate, distribute, and supply electricity under The Electricity (Class Exemptions from the Requirement for a License)
	Order 2001.[e] As a whole, European policy calls for 12% of energy and 21% of electricity supply from renewables by 2010 including 1,000,000 PV systems, 10,000 MW wind, and 10,000 MW biomass[f]
Large-scale cogeneration	The UK is bound by the EU cogeneration Directive 2004/8/EC which states that member states are obliged to address key market barriers such as ensuring grid access for cogeneration. The directive also states that in order for cogeneration to be considered high efficiency, it must provide an energy saving of at least 10% compared to separate production of heat and power. Harmonized efficiency reference values are to be released no later than February 2006. The directive further states that member states must establish guarantee of origin standards no less than 6 months after the harmonized reference values are released. Nevertheless, the UK's own model for defining cogeneration deserves mention. The UK has established a cogeneration Quality Assurance scheme;[g] a voluntary methodology for determining which arrangements or projects can be defined as cogeneration and which are therefore eligible for cogeneration incentives and financial support. The scheme assesses cogeneration projects using two thresholds. The power efficiency threshold states that when a project's power efficiency is greater than 20%, then all fuel used in the project is "good quality." The second threshold, the quality index threshold, considers both power and heat efficiency which encourages good environmental practice. Though complex, the cogeneration QA scheme is robust in its consideration of technologies and fuels. The UK government has a target to obtain 10 GWe of its power by cogeneration by 2010.[h] Currently, cogeneration developers are under obligation to obtain 10% of power from renewables under the renewable obligation standards, as are any generators. This requirement is currently under review and the affect of the renewable obligation on cogeneration may soon change
	The EU has a goal to double cogeneration use by 2010[i]

(continued)

TABLE 2.19 ***(Continued)***

	Energy Efficiency Action Plan, Renewables Obligations extended from 2010/11 to 2015/16, Renewable Obligation certificates introduced[j]
	The UK has launched a micropower consultation to get public and stakeholder input into strategies for increasing the use of micropower in the country
Domestic cogeneration	
PV	
Wind	
Hydro	

[a] Technical Guide to the Connection of Generation to the Distribution Network. Distributed Generation Co-ordinating Group Technical Steering Group. February 2004.

[b] Micro-cogeneration needs specific treatment in the European Directive on Cogeneration, Cogen Europe.

[c] WADE Survey of Decentralized Energy 2005, p. 26.

[d] http://www.micropower.co.uk/content1.cfm?pageid=142

[e] www.woking.gov.uk/council/planning/publications/climateneutral2/energy.pdf

[f] Directive 2001/77/EC of the European Parliament and of the Council of 27 September 2001 on the promotion of electricity produced from renewable energy sources in the internal electricity market. Energy for the Future: Renewable Sources of Energy White Paper for a Community Strategy and Action Plan COM(97)599 final (26/11/1997).

[g] http://www.cogenerationqa.com

[h] UK White paper Our energy future—creating a low carbon economy.

[i] A community strategy to promote combined heat and power (CHP) and to dismantle barriers to its development. European Commission Communiqué COM (97) 514 final.

[j] Renewables 2004. *International Conference for Renewable Energies*, 1–4 June 2004, Bonn. Conference Report: Outcomes and Documentation—Political Declaration/International Action Programme/Policy Recommendations for Renewable Energies.

TABLE 2.20 Chinese Distributed Generation Policy by Technology and Type

	Technical
DE in general	
Large-scale cogeneration	
Domestic cogeneration	
PV	
Wind	
Hydro	
	Financial
DE in general	The Chinese Renewable Energy Law sets a feed-in tariff from registered renewable energy producers. The price is to be established and periodically adjusted by the National Development and Reform Commission. The law also established a national fund to promote renewable energy projects, and created favorable lending and tax schemes for renewable energy projects. Further policies are expected, including further tax breaks and incentives
	The Chinese Municipal Development and Reform Commission has increased the price of power at peak periods in an effort to reduce transmission congestion and encourage conservation though it is unclear which customers will be effected.[a] Such an action could potentially drive DE investment
	A separate local law, the Regenerating Power Law, introduced in the Northwest autonomous regions of Xinjiang and Inner Mongolia is also designed to promote renewable energy[b]
Large-scale cogeneration	Throughout the 1990s, the government provides capital grants and tax benefits for cogeneration but the program has since been withdrawn[a]

(continued)

TABLE 2.20 ***(Continued)***

Domestic cogeneration	
Onsite PV	
Onsite wind	Wind projects currently receive a 50% in sales tax as well as corporate and income tax[b]
Small hydro	
	Other
DE in general	According to the recent Renewable Energy Law, the Chinese National Development and Reform Commission must publish renewable energy target by January 2006. The draft Energy development plan contains various targets for renewable energy including a 12.5% target for total electric power capacity by 2020 (about 125 GW out of 1000 GW)[c] China has used feed-in tariffs and renewable portfolio standards.[d] The Chinese Renewable Energy Law was launched in March 2005 and is due to take effect on January 1, 2006. It aims to ensure 10% of energy to come from renewables by 2010.[a] This builds on the goal to have 5% of annual new generation being added to the system by 2010.[e] Specific guideline targets will be set by the National Development and Reform Commission and provincial planning agencies will then develop the specific strategies for each jurisdiction. The law sets penalties for jurisdictions that do not comply. Incentives for DE so far have tended to be financial and have not sufficiently focused on grid access issues China committed at Bonn in 2004 to introduce a National Renewable Law (which has been done) and also to formulate a National Renewable Energy Development Strategy, which is in the process of doing[f]
Large-scale cogeneration	1998 "Energy Conservation Law of PRC" promotes energy conservation including cogeneration.[b] In 2000, the government released the "Principle on Developing cogeneration" document which emphasized that plants should be sized according to heat load rather than electrical capacity, recommended guidelines for project approval and established targets for efficiency in heat/power ratios.[a] In 2002 Policy paper number 1268 again emphasized the benefits of cogeneration but included small-scale cogeneration for the first time In June 2002, the National People's Congress approved the *Cleaner Production Promotion Law*. Coming into effect on January 1, 2003 the law is expected to be a driver for cogeneration in industrial applications The latest 5-year plan aims to develop local cogeneration capacity in order to meet targets of 40% of heat supplied by cogeneration systems by 2010 and to install an additional 40 GWe of cogeneration plants by the same year
Domestic cogeneration	
PV	
Wind	20 GW of wind by 2020
Hydro	

[a] WADE Market Analysis China 2005.
[b] ChinElec China Power Monitor, Issue 02, 26th April 2005 Week 2002, NewsBase, UK.
[c] Renewables 2005. Global Status Report, World Watch Institute.
[d] http://www.renewables2004.de/pdf/policy_recommendations_final.pdf
[e] Eric Martinot 2002.
[f] Renewables 2004. *International Conference for Renewable Energies*, 1–4 June 2004, Bonn. Conference Report: Outcomes and Documentation—Political Declaration/International Action Programme/Policy Recommendations for Renewable Energies.

TABLE 2.21 Indian Distributed Generation Policy by Technology and Type

	Technical
DE in general	Grid interconnection of onsite power is "facilitated" in section 30 of the 2003 Electricity Act The onsite power section of the law was meant to help secure reliable, quality and cost effective power, and create employment opportunities. The act provided also aimed to guarantee access to the grid for onsite generators[a]
Large-scale cogeneration	
Domestic cogeneration	
PV	
Wind	
Hydro	
	Financial
DE in general	The Regulatory Commission oversees commercial arrangements between onsite generators and grid operators and determines tariffs for onsite plants. "captive power plants" are guaranteed access to the grid under Indian legislation (section 30) Energy Conservation Act states that energy intensive industries must undergo compulsory Periodic energy audits[a] According to the Electricity Act 2003, each State Electricity Regulatory Commission must set a target for a percentage of power to be purchased from nonconventional sources by each distribution company. The percentage must increase progressively as prescribed by State Electricity Regulatory Commissions. A competitive bidding process is to be employed. Differential tariffs may also be set to encourage investment in nonconventional energy (renewables and cogeneration) India has a target that 10% of new electric capacity added between 2003 and 2012 be renewable[b]
Large-scale cogeneration	New rules introduced in Maharashtra state in 2002 have created incentive for cogeneration plant investors using biomass fuel, to upgrade and improve the efficiency of their cogeneration plants. The key change is the introduction of attractive buyback rates for any surplus electricity produced by the mill. The feed-in tariff for all biomass-based cogeneration is Rs. 3.05/kWh for the first year of operation of the cogeneration project and it will rise at the rate of 2% per annum on compounded basis. The tariff should provide a stable investment climate until at least March 2007 where it will be reviewed. The Order also clearly defines what fuels and technologies qualify, and lays down clear expectations as to cost sharing of T&D infrastructure and clarifies contractual obligations in terms of power supply, power quality, and payments[c] Similar laws have also been adopted in Andhra Pradesh and Madhya Pradesh[b] The Indian government provides low interest (4%) loans to sugar industry players to develop cogeneration facilities[d]
Domestic cogeneration	
Onsite PV	
Onsite wind	Favorable investment tax credits have proven successful in encouraging a vibrant wind industry[e]
Small hydro	
	Other
DE in general	Electricity Act of 2003 includes favorable provisions for DE[f] and guarantees access to all generators regardless of size.[g] The Act has moved India toward an electricity market with separate generation, transmission and distribution of power, with increasing potential for competition. A milestone in this progress is the introduction of the "Availability Based Tariff," which is an intermediate step in the effort to develop a true spot market for electricity. An ABT has been introduced at all five electrical regions of the country at the interstate level[h] Section 30 of the 2003 Electricity Act is designed to facilitate onsite generation. The Bill Section 38, 2(d) of the 2003 Electricity Act makes it obligatory to provide "non-discriminatory open access." There is a similar obligation on the State Transmission Utility, Section 39-2(d) to provide nondiscriminatory open access[g]

(continued)

TABLE 2.21 *(Continued)*

	India's 5-year plan contains proposals for 10,000 MWe of new renewables by 2012.[f] The government goal has also stated that renewables should account 10% new generation being added to the system by 2012.[e] A percentage of power to be generated from DE in each distribution license area must be declared by the State Commission in accordance with section 86 (1) (e) of Electricity Act 2003[g]
	The National Electricity Policy requires large generation utilities to set aside a small percentage of revenue to be used for research and development related to modernization of the power sector. Whether these funds would be earmarked for DE or more centralized technology remains to be seen[a]
Large-scale cogeneration	
Domestic cogeneration	
PV	
Wind	Wind power producers may transfer the electricity generated over the transmission system to use in their own facilities or to a third party for sale[e]
Hydro	

[a] Resolution No. 23/40/2004-R&R (Vol. II) http://powermin.nic.in/indian_electricity_scenario/national_electricity_policy.htm

[b] Renewables 2005. Global Status Report, Wordwatch Insitute, p. 25.

[c] The Maharashtra Electricity Regulatory Commission 2001. Order for Non-Fossil Fuel based Co-generation Projects 2002. http://www.mercindia.com/pdf/16082002.zip

[d] Balrampur Chini Mills Limited, Sugar Scenarios 2005. www.chini.com/bcml.ppt

[e] Grid-Based Renewable Energy in Developing Countries: Policies, Strategies, and Lessons from the GEF World Renewable Energy Policy And Strategy Forum June 13–15, 2002, Berlin, Germany Eric Martinot Global Environment Facility, Washington, DC.

[f] WADE World Survey for Decentralized Energy 2005, p. 19.

[g] Electricity Act 2003 http://powermin.nic.in/acts_notification/pdf/The%20Electricity%20Act_2003.pdf

[h] Distributed Generation—A Strategy for Optimal Future Power Generation in India. Ajit Kapadia and K.N. Naik Centre for Fuel Studies and Research, Pune India.

TABLE 2.22 Japanese Distributed Generation Policy by Technology and Type

	Technical
DE in general	Technical guidelines for grid-interconnection of small DE applications have been established[a]
Large-scale cogeneration	
Domestic cogeneration	
PV	
Wind	
Hydro	
	Financial
DE in general	Japan has been traditionally used a different approach to renewable energy and energy efficiency policy tying more closely energy use to economic production[b]
	Purchasers of cogeneration and renewable energy equipment are eligible for low interest loans and are eligible for: either a deduction from its corporate or income tax equal to 7% of all costs, up to 20% of corporate or income tax to be paid in the same period—or a 30% depreciation rate equal to the cost in addition to normal depreciation
	The "Project for Promoting the Local Introduction of New Energy" provides subsidies to local governments who invest in the whole range of DE projects, including renewables, fuel cells, and gas-fired CHP.[c] The "Project for Supporting New Energy Operators" aims to support entrepreneurs who become involved in DE. Upstart firms are eligible for guaranteed debt or subsidization. Up to a third of installation costs for projects is subsidized and 90% of the debts are guaranteed[c]

(continued)

TABLE 2.22 ***(Continued)***

Large-scale cogeneration	
Domestic cogeneration	
Onsite PV	The 1992 net-metering policy required utilities to purchase excess power from PV systems at the retail rate the same year ambitious targets for PV were set. In 1994, the "70,000 Roofs" program was launched. Incentives included low interest loans, a comprehensive education and outreach program, and declining rebates for interconnected residential systems. Rebates were initially 50% of the installed cost for end users but declined annually and were phased out completely in 2002. In 1997 the program was altered to also extend rebates to landlords and housing developers. The government promoted PV with print and television advertising campaigns. By the end of the program in 2002, it had exceeded its goals.[c] The feed-in tariff originally introduced in 1997 is still in existence but has been reduced to 45,000 JPY/kW in 2004 from 90,000 JPY/kW in 2003
Onsite wind	
Small hydro	
	Other
DE in general	Japan's Law Concerning the Rational Use of Energy (Amended in 1993) states (Chapter 2, Article 4) that large energy users consuming more than 3000 kl crude oil equivalents per year or more than 12 million kWh electricity per year, are deemed "energy manager designated factories." The ministry has the power to designate certain industries as triggers to be examined closely for their performance in the above criteria (Article 6). Examples of industries that have been designated include pulp and paper, automotive factories, etc. Each designated factory must appoint an energy manager trained in energy conservation and the manager is responsible for optimizing the factory performance according to seven criteria The manager must consider seven efficiency factors in the operation of designated plants 1. Rationalization of fuel combustion 2. Rationalization of heating, cooling, and heat transfer 3. Prevention of heat loss by radiation, conduction, etc. 4. Recovery of waste heat 5. Rationalization of heat conversion into power, etc. 6. Prevention of electricity loss by resistance, etc. 7. Rationalization of conversion of electricity into power, heat, etc. The owners of designated factories are legally bound to comply with any recommendations the energy specialists may make. If it is deemed that the designated factories have not met their obligations under the law the ministry may intervene.[b] Japan's 2003 "Energy Master plan" sees DE and centralized generation technologies coexisting in the future.[a] It described the importance of investing in DE fuel cells, cogeneration, PV, wind, biomass, and waste 2200 MWe of fuel cells target by 2010[a] 4170 MWe waste by 2010[a] 330 MWe biomass by 2010[a] The government has also funded an agency to research new energy technologies, Japan Committed in Bonn to introducing a Renewable Portfolio Standard Law.[d] The current target is that 1.35% of total electricity should be from renewables by 2010 (not including geothermal or large hydro)[e] The "Law on Promotion of Green Purchasing" was passed in 2001 which requires
Large-scale cogeneration	Target of 10 GWe of reciprocating engine cogeneration by 2010[a]
Domestic cogeneration	
PV	4820 MWe PV by 2010[a]
Wind	
Hydro	

[a] Shinichi Nakane, Japan Cogeneration Center.

[b] Energy Policy Instruments—Description of Selected Countries Trine Pipi Kræmer and Morten Grauballe, AKF Forlaget http://www.akf.dk/eng/udland8.htm

[c] IEA PVPS Annual report 2004.

[d] Renewables 2004. *International Conference for Renewable Energies*, 1–4 June 2004, Bonn. Conference Report: Outcomes and Documentation—Political Declaration/International Action Programme/Policy Recommendations for Renewable Energies.

[e] Renewables 2005. Global Status Report, Worldwatch Institute, p. 25.

TABLE 2.23 Korean Distributed Generation Policy by Technology and Type

	Technical
DE in general	
Large-scale cogeneration	
Domestic cogeneration	
PV	No specific standards or technical regulations available for grid-interconnection of small onsite systems but KEMCO among others are currently working on this
Wind	
Hydro	
	Financial
DE in general	Companies that employ renewable energy technologies, processes, and equipment are eligible for government backed low-interest loans as well as a 10% investment tax credit if the project is a research and development project. Also there are plans to strengthen mandatory feed-in tariff for renewable electricity[a] to include biomass. A favorable loan program is in place including a 10-year payback at 3.5% with up to a 5-year grace period[b]
Large-scale cogeneration	The Integrated Energy Policy requires energy audits to be undertaken on all major energy users and then provides third party loans to ESCOs to cover 50% of project finance to implement efficiency upgrades including cogeneration.[c] Plans are being considered to introduce a feed-in tariff for IGCC, but it is unclear if heat recovery will be a requisite for eligible schemes[b]
Domestic cogeneration	
Onsite PV	Up to a 70% grant for the capital/installation costs of PV is in place.[b] 746.4 won/kWh feed-in tariff[b]
Onsite wind	Plans are being discussed which would provide up to a 25% grant for the capital cost of wind.[a] 107.66 won/kWh feed-in tariff[b]
Small hydro	Plans for a program similar to the PV installation grant program are being developed. 73.69 won/kWh feed-in tariff[b]
	Other
DE in general	The national utility KEPCO divided up its vertically integrated systems into T&D and generation in 2001.[a] The 1987 Promotion Act for New and Renewable Energy Development established the legal foundation for R&D into DE. The Act was amended in 1997 to try to promote wider use and dissemination. In 2002, the Act was further amended this time introducing a public obligation and certification scheme. In 2003, the 10-year "National Basic Plan for New and Renewable Energy Technology" was adopted calling for a 3% of all electricity from renewables by 2006 and 5% by 2011.[b] All new public buildings must allocate at least 5% of their construction cost to new and renewable energy. There is a large-scale T&D program in place with 9.1 trillion won budgeted between now and 2011
Large-scale cogeneration	
Domestic cogeneration	
PV	Korea has introduced a new R&D program in 2002 called "Solar Land 2010 Program" also there is a government goal of establishing 30,000 3 kW residential systems 40,000 10 kW systems for public buildings, and 30,000 20 kW systems for commercial buildings
Wind	
Hydro	

[a] http://www.oja-services.nl/iea-pvps/countries/korea/index.htm

[b] Electricity Industry Policy Research Group Korea Electrotechnology Research Institute, Sung-chul Yang presentation, January 2005.

[c] IEA policies and measures database: http://www.iea.org/textbase/pamsdb/search.aspx?mode=re

2.8.16 Japan

Japan has a long history of DE leadership and boasts the world's lowest energy consumed per GDP ratios. Still Japan is highly dependent on energy imports and additional investment in DE may go a long way in addressing this. Japan's PV manufacturing infrastructure is the world's most developed and there appears to be ever increasing policy support for DE. There remains much room for improvement and additional policy support will be key in realizing the potential of DE. Major drivers for DE growth will continue to be climate change and energy security (Table 2.22).

2.8.17 Korea

Like many OECD countries, Korea is almost totally reliant on energy imports to meet domestic demand. Energy independence, reliability, and economic competitiveness will continue to be an important driver for DE for sometime to come. It is clear that policy will play an important role in realizing Korea's ambitious goals in developing DE (Table 2.23).

3

Economics Methods

Rosalie Ruegg

National Institute of Standards and Technology

Walter Short

National Renewable Energy Laboratory

3.1 Introduction

Economic-evaluation methods facilitate comparisons among energy technology investments. Generally, the same methods can be used to compare investments in energy supply or energy efficiency. All sectors of the energy community need guidelines for making economically efficient energy-related decisions.

This chapter provides an introduction to some basic methods that are helpful in designing and sizing cost-effective systems, and in determining whether it is economically efficient to invest in specific energy efficiency or renewable energy projects. The targeted audience includes analysts, architects, engineers, designers, builders, codes and standards writers, and government policy makers—collectively referred to as the *design community*.

The focus is on microeconomic methods for measuring cost-effectiveness of individual projects or groups of projects, with explicit treatment of uncertainty. The chapter does not treat macroeconomic methods and national market-penetration models for measuring economic impacts of energy efficiency and renewable energy investments on the national economy. It provides sufficient guidance for computing measures of economic performance for relatively simple investment choices, and it provides the fundamentals for dealing with complex investment decisions.

3.2 Making Economically Efficient Choices[1]

Economic-evaluation methods can be used in a number of ways to increase the economic efficiency of energy-related decisions. There are methods that can be used to obtain the largest possible savings in energy costs for a given energy budget; there are methods that can be used to achieve a targeted reduction in energy costs for the lowest possible efficiency/renewable energy investment; and there are methods that can be used to determine how much it pays to spend on energy efficiency and renewable energy to lower total lifetime costs, including both investment costs and energy cost savings.

The first two ways of using economic-evaluation methods (i.e., to obtain the largest savings for a fixed budget and to obtain a targeted savings for the lowest budget) are more limited applications than the third, which aims to minimize total costs or maximize NB (net savings) from expenditure on energy efficiency and renewables. As an example of the first, a plant owner may budget a specific sum of money for the purpose of retrofitting the plant for energy efficiency. As an example of the second, designers may be required by state or federal building standards and/or codes to reduce the design energy loads of new buildings below some specified level. As an example of the third, engineers may be required by their clients to include, in a production plant, those energy efficiency and renewable energy features that will pay off in terms of lower overall production costs over the long run.

Note that economic efficiency is not necessarily the same as engineering thermal efficiency. For example, one furnace may be more "efficient" than another in the engineering technical sense, if it delivers more units of heat for a given quantity of fuel than another. Yet it may not be economically efficient if the first cost of the higher-output furnace outweighs its fuel savings. The focus in this chapter is on economic efficiency, not technical efficiency.

Economic efficiency is illustrated conceptually in Figure 3.1 through Figure 3.3 with an investment in energy efficiency. Figure 3.1 shows the level of energy conservation, Q_c, that maximizes NB from energy conservation, i.e., the level that is most profitable over the long run. Note that it is the point at which the curves are most distant from one another.

Figure 3.2 shows how "marginal analysis" can be used to find the same level of conservation, Q_c, that will yield the largest NB. It depicts changes in the total benefits and cost curves (i.e., the derivatives of the curves in Figure 3.1) as the level of energy conservation is increased. The point of intersection of the marginal curves coincides with the most profitable level of energy conservation indicated in Figure 3.1. This is the point at which the cost of adding one more unit of conservation is just equal to the corresponding benefits in terms of energy savings (i.e., the point at which "marginal costs" and "marginal benefits" are equal). To the left of the point of intersection, the additional benefits from increasing the level of conservation by another unit are greater than the additional costs, and it pays to invest more. To the right of the point of intersection, the costs of an addition to the level of conservation exceed the benefits and the level of total NB begins to fall, as shown in Figure 3.1. Figure 3.3 shows that the most economically efficient level of energy conservation, Q_c, is that for which the total cost curve is at a minimum.

The most economically efficient level of conservation, Q_c, is the same in Figure 3.1 through Figure 3.3. Three different approaches to finding Q_c are illustrated: (1) finding the maximum difference between

[1]This section is based on a treatment of these concepts provided by Marshall and Ruegg in *Economics of Solar Energy and Conservation Systems*, Kreith and West, ed., CRC Press, 1980.

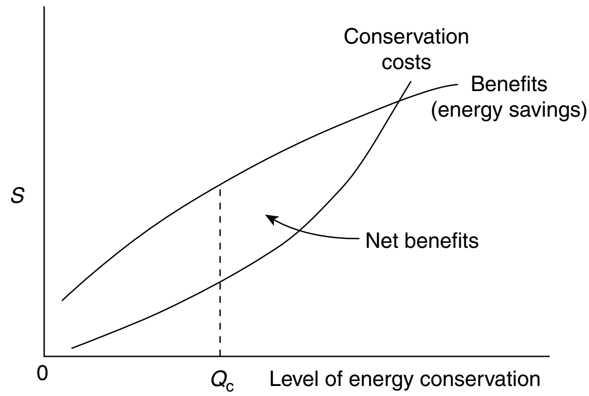


FIGURE 3.1 Maximizing net benefits.

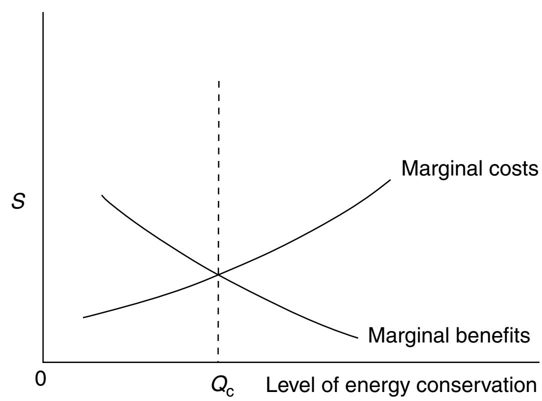


FIGURE 3.2 Equating marginal benefits and marginal costs.

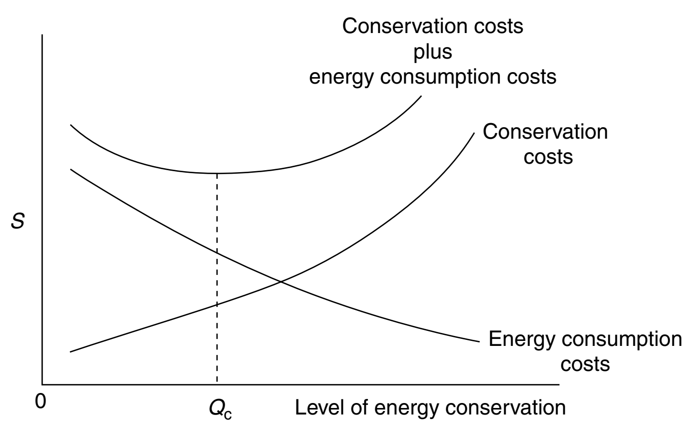


FIGURE 3.3 Minimizing life-cycle costs.

benefits and costs, (2) finding the point where marginal benefits equal marginal costs, and (3) finding the lowest life-cycle costs (LCCs).

3.3 Economic-Evaluation Methods[2]

There are a number of closely related, commonly used methods for evaluating economic performance. These include the LCC method, levelized cost of energy, NB (net present worth) method, benefit-cost (or savings-to-investment) ratio method, internal rate-of-return method, overall rate-of-return method, and payback method. All of these methods are used when the important effects can be measured in dollars. If incommensurable effects are critical to the decision, it is important that they also be taken into account. But, because only quantified effects are included in the calculations for these economic methods, unquantified effects must be treated outside the models. Brief treatments of the methods are provided; some additional methods are identified but not treated. For more comprehensive treatments, see Ruegg and Marshall (1990) and other sources listed at the end of the chapter.

3.3.1 Life-Cycle Cost Method

The life-cycle costing (LCC) method sums, for each investment alternative, the costs of acquisition, maintenance, repair, replacement, energy, and any other monetary costs (less any income amounts, such as salvage value) that are affected by the investment decision. The time value of money must be taken into account for all amounts, and the amounts must be considered over the relevant period. All amounts are usually measured either in present value or annual value dollars. This is discussed later under "Discounting" and "Discount Rate." At a minimum, for comparison, the investment alternatives should include a "base-case" alternative of not making the energy efficiency or renewable investment, and at least one case of an investment in a specific efficiency or renewable system. Numerous alternatives may be compared. The alternative with the lowest LCC that meets the investor's objective and constraints is the preferred investment.

Following is a formula for finding the LCCs of each alternative:

$$\mathrm{LCC}_{\mathrm{A1}} = I_{\mathrm{A1}} + E_{\mathrm{A1}} + M_{\mathrm{A1}} + R_{\mathrm{A1}} - S_{\mathrm{A1}}, \tag{3.1}$$

where $\mathrm{LCC}_{\mathrm{A1}}$ is the life-cycle cost of alternative Al, I_{A1} is the present-value investment costs of alternative A1, E_{A1} is the present-value energy costs associated with alternative A1, M_{A1} is the present-value nonfuel operating and maintenance cost of A1, R_{A1} is the present-value repair and replacement costs of A1, and S_{A1} is the present-value resale (or salvage) value less disposal cost associated with alternative A1.

The LCC method is particularly useful for decisions that are made primarily on the basis of cost effectiveness, such as whether a given energy efficiency or renewable energy investment will lower total cost, (e.g., the sum of investment and operating costs). It can be used to compare alternative designs or sizes of systems, as long as the systems provide the same service. The method, if used correctly, can be used to find the overall cost-minimizing combination of energy efficiency investments and energy supply investments within a given facility. However, in general, it cannot be used to find the best investment, because totally different investments do not provide the same service.

3.3.2 Levelized Cost of Energy

The levelized cost of energy (LCOE) is similar to the LCC method, in that it considers all the costs associated with an investment alternative and takes into account the time value of money for the analysis period. However, it is generally used to compare two alternative energy supply technologies or systems,

[2]These methods are treated in detail in Ruegg and Marshall *Building Economics: Theory and Practice*, Chapman and Hall, New York, NY, 1990.

e.g., two electricity production technologies. It differs from the LCC in that it usually considers taxes, but like LCC, it frequently ignores financing costs.

The LCOE is the value that must be received for each unit of energy produced to ensure that all costs and a reasonable profit are made. Profit is ensured by discounting future revenues at a discount rate that equals the rate of return that might be gained on other investments of comparable risk, i.e., the opportunity cost of capital. In equation form, this is represented as:

$$\sum_{t=1}^{t=N} \mathrm{LCOE} * Q_t/(1+d')^t = \sum_{t=0}^{t=N} C_t/(1+d)^t, \tag{3.2}$$

where N is the analysis period, Q_t is the amount of energy production in period t, C_t is the cost incurred in period t, d' is the discount rate or opportunity cost of capital. If d' is a real discount rate (excludes inflation) then the LCOE will be in real (constant) dollar terms, whereas the LCOE will be in nominal (current) dollar terms if d' is a nominal discount rate. The discount rate, d, is used to bring future costs back to their present value. If those costs are expressed in real dollars, then the discount rate should be a real discount rate; if they are in nominal dollars, the discount rate should be a nominal discount rate.

3.3.3 Net Present Value or Net Benefits Method

The net present value (NPV) method finds the excess of benefits over costs, where all amounts are discounted for their time value. (If costs exceed benefits, net losses result.)

The NPV method is also often called the *net present worth* or *net savings* method. When this method is used for evaluating a cost-reducing investment, the cost savings are the benefits, and it is often called the *net savings* (NS) method.

Following is a formula for finding the NPV from an investment, such as an investment in energy efficiency or renewable energy systems:

$$\mathrm{NPV}_{\mathrm{A1:A2}} = \sum_{t=0}^{N} \frac{(B_t - C_t)}{(1+d)^t}, \tag{3.3}$$

where $\mathrm{NPV}_{\mathrm{A1:A2}}$ is NB, i.e., present value benefits (savings) net of present value costs for alternative A1 as compared with alternative A2, B_t is benefits in year t, which may be defined to include energy savings, C_t is costs in year t associated with alternative A1 as compared with a mutually exclusive alternative A2, and d is the discount rate.

The NPV method is particularly suitable for decisions made on the basis of long-run profitability. The NPV (NB) method is also useful for deciding whether to make a given investment and for designing and sizing systems. It is not very useful for comparing investments that provide different services.

3.3.4 Benefit-to-Cost Ratio or Savings-to-Investment Ratio Method

The benefit-to-cost ratio (BCR) or savings-to-investment ratio (SIR) method divides benefits by costs or by savings by investment. When used to evaluate energy efficiency and renewable energy systems, benefits are in terms of energy cost savings. The numerator of the SIR ratio is usually constructed as energy savings, net of maintenance and repair costs; and the denominator as the sum of investment costs and replacement costs less salvage value (capital cost items). However, depending on the objective, sometimes only initial investment costs are placed in the denominator and the other costs are subtracted in the numerator—or sometimes only the investor's equity capital is placed in the denominator. As with the three preceding methods, this method is based on discounted cash flows.

Unlike the three preceding methods that provided a performance measure in dollars, this method gives the measure as a dimensionless number. The higher the ratio, the more dollar savings realized per dollar of investment.

Following is a commonly used formula for computing the ratio of savings-to-investment costs:

$$\mathrm{SIR}_{\mathrm{A1:A2}} = \frac{\sum_{t=0}^{N}[\mathrm{CS}_t(1+d)^{-t}]}{\sum_{t=0}^{N}(I_t(1+d)^{-t})}, \tag{3.4}$$

where $\mathrm{SIR}_{\mathrm{A1:A2}}$ is the SIR for alternative A1 relative to mutually exclusive alternative A2, CS_t is the cost savings (excluding those investment costs in the denominator) plus any positive benefits of alternative A1 as compared with mutually exclusive alternative A2, and I_t is the additional investment costs for alternative A1 relative to A2.

Note that the particular formulation of the ratio with respect to the placement of items in the numerator or denominator can affect the outcome. One should use a formulation appropriate to the decision maker's objectives.

The ratio method can be used to determine whether or not to accept or reject a given investment on economic grounds. It also can be used for design and size decisions and other choices among mutually exclusive alternatives, if applied incrementally (i.e., the investment and savings are the difference between the two mutually exclusive alternatives). A primary application of the ratio method is to set funding priorities among projects competing for a limited budget. When it is used in this way—and when project costs are "lumpy" (making it impossible to fully allocate the budget by taking projects in order according to the size of their ratios)—SIR should be supplemented with the evaluation of alternative sets of projects using the NPV or NB method.

3.3.5 Internal Rate-of-Return Method

The internal rate-of-return (IRR) method solves for the discount rate for which dollar savings are just equal to dollar costs over the analysis period; that is, the rate for which the NPV is zero. This discount rate is the rate of return on the investment. It is compared to the investor's minimum acceptable rate of return to determine whether the investment is desirable. Unlike the preceding three techniques, the internal rate of return does not call for the inclusion of a prespecified discount rate in the computation; rather, it solves for a discount rate.

The rate of return is typically calculated by a process of trial and error, by which various compound rates of interest are used to discount cash flows until a rate is found for which the NPV of the investment is zero. The approach is the following: (1) Compute *NPV* using Equation 3.3, except substitute a trial interest rate for the discount rate, d, in the equation. A positive NPV means that the IRR is greater than the trial rate; a negative NPV means that the IRR is less than the trial rate. (2) Based on the information, try another rate. (3) By a series of iterations, find the rate at which NPV equals zero.

Computer algorithms, graphical techniques, and—for simple cases—discount-factor tabular approaches, are often used to facilitate IRR solutions (Ruegg and Marshall 1990:7172). Expressing economic performance as a rate of return can be desirable for ease in comparing the returns on a variety of investment opportunities, because returns are often expressed in terms of annual rates of return. The IRR method is useful for accepting or rejecting individual investments or for allocating a budget. For designing or sizing projects, the IRR method, like the SIR, must be applied incrementally. It is not recommended for selecting between mutually exclusive investments with significantly different lifetimes (e.g., a project with a 35% return for 20 years is a much better investment than a project with the same 35% return for only two years).

It is a widely used method, but it is often misused, largely due to shortcomings that include the possibility of:

- No solution (the sum of all nondiscounted returns within the analysis period are less than the investment costs)
- Multiple solution values (some costs occur later than some of the returns)

- Failure to give a measure of overall return associated with the project over the analysis period (returns occurring before the end of the analysis are implicitly assumed to be reinvested at the same rate of return as the calculated IRR; this may or may not be possible)

3.3.6 Overall Rate-of-Return Method

The overall rate-of-return (ORR) method corrects for the last two shortcomings expressed above for the IRR. Like the IRR, the ORR expresses economic performance in terms of an annual rate of return over the analysis period. But unlike the IRR, the ORR requires, as input, an explicit reinvestment rate on interim receipts and produces a unique solution value.[3] The explicit reinvestment rate makes it possible to express net cash flows (excluding investment costs) in terms of their future value at the end of the analysis period. The ORR is then easily computed with a closed-form solution as shown in Equation 3.5:

$$\mathrm{ORR}_{\mathrm{A1:A2}} = \frac{\left[\sum_{t=0}^{N} (B_t - C_t)(1+r)^{N-t}\right]^{1/N}}{\left[\sum_{t=0}^{N} I_t/(1+r)^t\right]^{1/N}} - 1, \tag{3.5}$$

where $\mathrm{ORR}_{\mathrm{A1:A2}}$ is the overall rate of return on a given investment alternative A1 relative to a mutually exclusive alternative over a designated study period, B_t represents the benefits from a given alternative relative to a mutually exclusive alternative A2 over time period t, C_t is the costs (excluding that part of investment costs on which the return is to be maximized) associated with a given alternative relative to a mutually exclusive alternative A2 over time t, r is the reinvestment rate at which net returns can be reinvested, usually set equal to the discount rate, N is the length of the study period, and I_t represents the investment costs in time t on which the return is to be maximized.

The ORR is recommended as a substitute for the IRR, because it avoids some of the limitations and problems of the IRR. It can be used for deciding whether or not to fund a given project, for designing or sizing projects (if it is used incrementally), and for budget-allocation decisions.

3.3.7 Discounted Payback Method

This evaluation method measures the elapsed time between the time of an initial investment and the point in time at which accumulated discounted savings or benefits—net of other accumulated discounted costs—are sufficient to offset the initial investment, taking into account the time value of money. (If costs and savings are not discounted, the technique is called "simple payback.") For the investor who requires a rapid return of investment funds, the shorter the length of time until the investment pays off, the more desirable the investment.

To determine the discounted payback method (DPB) period, find the minimum value of Y (year in which payback occurs) such that the following equality is satisfied.

$$\sum_{t=1}^{Y} (B_t - C'_t)/(1+d)^t = I_0, \tag{3.6}$$

where B_t represents the benefits associated in period t with one alternative as compared with a mutually exclusive alternative, C'_t is the costs in period t (not including initial investment costs) associated with an alternative as compared with a mutually exclusive alternative in period t, and I_0 is the initial investment

[3]As shown in Equation 3.5, the reinvestment rate is also used to bring all investments back to their present value. Alternatively, investments after time zero can be discounted by the overall growth rate. In this case, a unique solution is not guaranteed, and the ORR must be found iteratively (Stermole and Stermole 2000).

costs of an alternative as compared with a mutually exclusive alternative, where the initial investment cost comprises total investment costs.

DPB is often (correctly) used as a supplementary measure when project life is uncertain. It is used to identify feasible projects when the investor's time horizon is constrained. It is used as a supplementary measure in the face of uncertainty to indicate how long capital is at risk. It is a rough guide for accept/reject decisions. It is also overused and misused. Because it indicates the time at which the investment just breaks even, it is not a reliable guide for choosing the most profitable investment alternative, as savings or benefits after the payback time could be significant.

3.3.8 Other Economic-Evaluation Methods

A variety of other methods have been used to evaluate the economic performance of energy systems, but these tend to be hybrids of those presented here. One of these is the required revenue method that computes a measure of the before-tax revenue in present or annual-value dollars required to cover the costs on an after-tax basis of an energy system (Ruegg and Short 1988:2223). Mathematical programming methods have been used to evaluate the optimal size or design of projects, as well as other mathematical and statistical techniques.

3.4 Risk Assessment

Many of the inputs to the above evaluation methods will be highly uncertain at the time an investment decision must be made. To make the most informed decision possible, an investor should employ these methods within a framework that explicitly accounts for risk and uncertainty.

Risk assessment provides decision makers with information about the "risk exposure" inherent in a given decision, i.e., the probability that the outcome will be different from the "best-guess" estimate. Risk assessment is also concerned with the *risk attitude* of the decision maker that describes his/her willingness to take a chance on an investment of uncertain outcome. Risk assessment techniques are typically used in conjunction with the evaluation methods outlined earlier; and not as stand-alone evaluation techniques.

The risk assessment techniques range from simple and partial to complex and comprehensive. Though none takes the risk out of making decisions, the techniques—if used correctly—can help the decision maker make more informed choices in the face of uncertainty.

This chapter provides an overview of the following probability-based risk assessment techniques:

- Expected value analysis
- Mean-variance criterion and coefficient of variation
- Risk-adjusted discount rate technique
- Certainty equivalent technique
- Monte Carlo simulation
- Decision analysis
- Real options
- Sensitivity analysis

There are other techniques that are used to assess risks and uncertainty (e.g., CAP_M, and break-even analysis), but those are not treated here.

3.4.1 Expected Value Analysis

Expected value (EV) analysis provides a simple way of taking into account uncertainty about input values, but it does not provide an explicit measure of risk in the outcome. It is helpful in explaining and illustrating risk attitudes.

3.4.1.1 How to Calculate EV

An *expected value* is the sum of the products of the dollar value of alternative outcomes and their probabilities of occurrence. That is, where a_i $(i=1,\ldots,n)$ indicates the value associated with alternative outcomes of a decision, and p_i indicates the probability of occurrence of alternative a_i, the EV of the decision is calculated as follows:

$$\mathrm{EV} = a1p1 + a2p2 + \ldots + a_n p_n. \tag{3.7}$$

3.4.1.2 Example of EV Analysis

The following simplified example illustrates the combining of EV analysis and NPV analysis to support a purchase decision.

Assume that a not-for-profit organization must decide whether to buy a given piece of energy-saving equipment. Assume that the unit purchase price of the equipment is $100,000, the yearly operating cost is $5,000 (obtained by a fixed-price contract), and both costs are known with certainty. The annual energy cost savings, on the other hand, are uncertain, but can be estimated in probabilistic terms as shown in Table 3.1 in the columns headed a_1, p_1, a_2, and p_2. The present-value calculations are also given in Table 3.1.

If the equipment decision were based only on NPV, calculated with the "best-guess" energy savings (column a_1), the equipment purchase would be found to be uneconomic. But if the possibility of greater energy savings is taken into account by using the EV of savings rather than the best guess, the conclusion is that, over repeated applications, the equipment is expected to be cost effective. The expected NPV of the energy-saving equipment is $25,000 per unit.

3.4.1.3 Advantages and Disadvantages of the EV Technique

An advantage of the technique is that it predicts a value that tends to be closer to the actual value than a simple "best-guess" estimate over repeated instances of the same event, provided, of course, that the input probabilities can be estimated with some accuracy.

A disadvantage of the EV technique is that it expresses the outcome as a single-value measure, such that there is no explicit measure of risk. Another is that the estimated outcome is predicated on many replications of the event, with the EV, in effect, a weighted average of the outcome over many like events. But the EV is unlikely to occur for a single instance of an event. This is analogous to a single coin toss: the outcome will be either heads or tails, not the probabilistic-based weighted average of both.

TABLE 3.1 Expected Value Example

Year	Equipment Purchase ($1000)	Operating Costs ($1000)	Energy Savings					
			A_1 ($1000)	P_1	A_2 ($1000)	P_2	PV[a] Factor	PV ($1000)
0	−100	—	—	—	—	—	1	−100
1		−5	25	0.8	50	0.2	0.926	23
2		−5	30	0.8	60	0.2	0.857	27
3		−5	30	0.7	60	0.3	0.794	27
4		−5	30	0.6	60	0.4	0.7354	27
5		−5	30	0.8	60	0.2	0.681	21
Expected Net Present Value:								25

[a] Present value calculations are based on a discount rate of 8%. Probabilities sum to 1.0 in a given year.

3.4.1.4 Expected Value and Risk Attitude

Expected values are useful in explaining risk attitude. Risk attitude may be thought of as a decision maker's preference between taking a chance on an uncertain money payout of known probability versus accepting a sure money amount. Suppose, for example, a person were given a choice between accepting the outcome of a fair coin toss where heads means winning \$10,000, and tails means losing \$5,000 and accepting a certain cash amount of \$2,000. EV analysis can be used to evaluate and compare the choices. In this case, the EV of the coin toss is \$2,500, which is \$500 more than the certain money amount. The "risk-neutral" decision maker will prefer the coin toss because of its higher EV. The decision maker who prefers the \$2,000 certain amount is demonstrating a "risk-averse" attitude. On the other hand, if the certain amount were raised to \$3,000 and the first decision maker still preferred the coin toss, he or she would be demonstrating a "risk-taking" attitude. Such tradeoffs can be used to derive a "utility function" that represents a decision maker's risk attitude.

The risk attitude of a given decision maker is typically a function of the amount at risk. Many people who are risk averse when faced with the possibility of significant loss become risk neutral—or even risk taking—when potential losses are small. Because decision makers vary substantially in their risk attitudes, there is a need to assess not only risk exposure (i.e., the degree of risk inherent in the decision) but also the risk attitude of the decision maker.

3.4.2 Mean-Variance Criterion and Coefficient of Variation

These techniques can be useful in choosing among risky alternatives, if the mean outcomes and standard deviations (variation from the mean) can be calculated.

Consider a choice between two projects—one with higher mean NB and a lower standard deviation than the other. This situation is illustrated in Figure 3.4. In this case, the project whose probability distribution is labeled B can be said to have stochastic dominance over the project labeled A. Project B is preferable to project A, both on grounds that its output is likely to be higher and that it entails less risk of loss. But what if project A, the alternative with higher risk, has the higher mean NB, as illustrated in Figure 3.5? If this were the case, the mean-variance criterion (MVC) would provide inconclusive results.

When there is not stochastic dominance of one project over the other(s), it is helpful to compute the coefficient of variation (CV) to determine the relative risk of the alternative projects. The CV indicates which alternative has the lower risk per unit of project output. Risk-averse decision makers will prefer the alternative with the lower CV, other things being equal. The CV is calculated as follows:

$$CV = \sigma/\mu, \tag{3.8}$$

where CV is the coefficient of variation, σ is the standard deviation, and μ is the mean.

The principal advantage of these techniques is that they provide quick, easy-to-calculate indications of the returns and risk exposure of one project relative to another. The principal disadvantage is that the

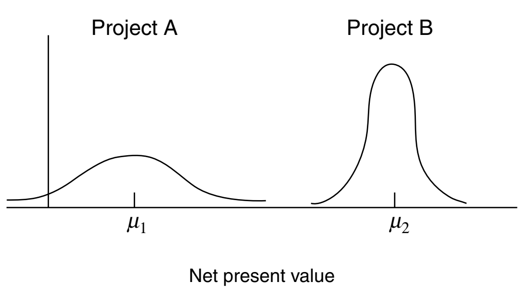


FIGURE 3.4 Stochastic dominance as demonstrated by mean-variance criterion.

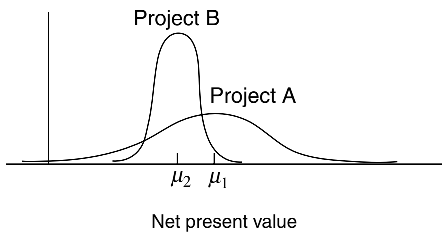


FIGURE 3.5 Inconclusive results from mean-variance criterion.

MVC does not provide a clear indication of preference when the alternative with the higher mean output has the higher risk, or vice versa.

3.4.3 Risk-Adjusted Discount Rate Technique

The risk-adjusted discount rate (RADR) technique takes account of risk through the discount rate. If a project's benefit stream is riskier than that of the average project in the decision maker's portfolio, a higher than normal discount rate is used; if the benefit stream is less risky, a lower than normal discount rate is used. If costs are the source of the higher-than-average uncertainty, a lower than normal discount rate is used and vice versa. The greater the variability in benefits or costs, the greater the adjustment in the discount rate.

The RADR is calculated as follows:

$$\text{RADR} = \text{RFR} + \text{NRA} + \text{XRA}, \tag{3.9}$$

where RADR is the risk-adjusted discounted rate, RFR is the risk-free discount rate, generally set equal to the treasury bill rate, NRA is the "normal" risk adjustment to account for the average level of risk encountered in the decision maker's operations, and XRA is the extra risk adjustment to account for risk greater or less than normal risk.

An example of using the RADR technique is the following: A company is considering an investment in a new type of alternative energy system with high payoff potential and high risk on the benefits side. The projected cost and revenue streams and the discounted present values are shown in Table 3.2. The treasury bill rate, taken as the risk-free rate, is 8%. The company uses a normal risk adjustment of 4% to account for the average level of risk encountered in its operations. The revenues associated with this investment are judged to be more than twice as risky as the company's average investment, so an additional risk adjustment of 6% is added to the RADR. Hence, the RADR is 18%. With this RADR, the NPV of the investment is estimated to be a loss of $28 million. On the basis of this uncertainty analysis, the company would be advised to not accept the project.

Advantages of the RADR technique are that it provides a way to account for both risk exposure and risk attitude. Moreover, RADR does not require any additional steps for calculating NPV once a value of the RADR is established. The disadvantage is that it provides only an approximate adjustment. The value of the RADR is typically a rough estimate based on sorting investments into risk categories and adding a

TABLE 3.2 Risk-Adjusted Discount Rate Example

Year	Costs ($M)	Revenue ($M)	PV Costs[a] ($M)	PV Revenue[b] ($M)	NPV ($M)
0	80	—	80	—	−80
1	5	20	4	17	13
2	5	20	4	14	10
3	5	20	4	12	8
4	5	20	3	10	7
5	5	20	3	9	6
6	5	20	3	7	4
7	5	20	2	6	4
Total NPV					−28

[a] Costs are discounted with a discount rate of 12%.
[b] Revenues are discounted with the RADR discount rate of 18%.

"fudge factor" to account for the decision maker's risk attitude. It generally is not a fine-tuned measure of the inherent risk associated with variation in cash flows. Further, it typically is biased toward investments with short payoffs because it applies a constant RADR over the entire analysis period, even though risk may vary over time.

3.4.4 Certainty Equivalent Technique

The certainty equivalent (CE) technique adjusts investment cash flows by a factor that will convert the measure of economic worth to a "certainty equivalent" amount—the amount a decision maker will find equally acceptable to a given investment with an uncertain outcome. Central to the technique is the derivation of the certainty equivalent factor (CEF) that is used to adjust net cash flows for uncertainty.

Risk exposure can be built into the CEF by establishing categories of risky investments for the decision maker's organization and linking the CEF to the CV of the returns—greater variation translating into smaller CEF values. The procedure is as follows:

1. Divide the organization's portfolio of projects into risk categories. Examples of investment risk categories for a private utility company might be the following: low-risk investments: expansion of existing energy systems and equipment replacement; moderate-risk investments: adoption of new, conventional energy systems; and high-risk investments: investment in new alternative energy systems.
2. Estimate the coefficients of variation (see the section on the CV technique) for each investment-risk category (e.g., on the basis of historical risk-return data).
3. Assign CEFs by year, according to the coefficients of variation, with the highest-risk projects being given the lowest CEFs. If the objectives are to reflect only risk exposure, set the CEFs such that a risk-neutral decision maker will be indifferent between receiving the estimated certain amount and the uncertain investment. If the objective is to reflect risk attitude as well as risk exposure, set the CEFs such that the decision maker with his or her own risk preference will be indifferent.

To apply the technique, proceed with the following steps:

4. Select the measure of economic performance to be used, such as the measure of NPV (i.e., NB).
5. Estimate the net cash flows and decide in which investment-risk category the project in question fits.
6. Multiply the yearly net cash flow amounts by the appropriate CEFs.
7. Discount the adjusted yearly net cash flow amounts with a risk-free discount rate (a risk-free discount rate is used because the risk adjustment is accomplished by the CEFs).
8. Proceed with the remainder of the analysis in the conventional way.

In summary, the certainty equivalent NPV is calculated as follows:

$$\mathrm{NPV_{CE}} = \sum_{t=0}^{N} [(\mathrm{CEF}_t(B_t - C_t))/(1 + \mathrm{RFD})^t], \tag{3.10}$$

where $\mathrm{NPV_{CE}}$ is the NPV adjusted for uncertainty by the CE technique, B_t is the estimated benefits in time period t, C_t is the estimated costs in time period t, and RFD is the risk-free discount rate.

Table 3.3 illustrates the use of this technique for adjusting net present-value calculations for an investment in a new, high-risk alternative energy system. The CEF is set at 0.76 and is assumed to be constant with respect to time.

A principal advantage of the CE Technique is that it can be used to account for both risk exposure and risk attitude. Another is that it separates the adjustment of risk from discounting and makes it possible to make more precise risk adjustments over time. A major disadvantage is that the estimation of CEF is only approximate.

TABLE 3.3 Certainty Equivalent (CE)

Yearly Net Cash Flow ($M)		CV	CEF	RFD Discount Factors[a]	NPV ($M)
1	−100	0.22	0.76	0.94	−71
2	−100	0.22	0.76	0.89	−68
3	20	0.22	0.76	0.84	13
4	30	0.22	0.76	0.79	18
5	45	0.22	0.76	0.75	26
6	65	0.22	0.76	0.7	35
7	65	0.22	0.76	0.67	33
8	65	0.22	0.76	0.63	31
9	50	0.22	0.76	0.59	22
10	50	0.22	0.76	0.56	21
Total NPV					

[a] The RFD is assumed equal to 6%.

3.4.5 Monte Carlo Simulation

A Monte Carlo simulation entails the iterative calculation of the measure of economic worth from probability functions of the input variables. The results are expressed as a probability density function and as a cumulative distribution function. The technique thereby enables explicit measures of risk exposure to be calculated. One of the economic-evaluation methods treated earlier is used to calculate economic worth; a computer is employed to sample repeatedly—hundreds of times—from the probability distributions and make the calculations. A Monte Carlo simulation can be performed by the following steps:

1. Express variable inputs as probability functions. Where there are interdependencies among input values, multiple probability density functions, tied to one another, may be needed.
2. For each input for which there is a probability function, draw randomly an input value; for each input for which there is only a single value; take that value for calculations.
3. Use the input values to calculate the economic measure of worth and record the results.
4. If inputs are interdependent, such that input X is a function of input Y, first draw the value of Y, then draw randomly from the X values that correspond to the value of Y.
5. Repeat the process many times until the number of results is sufficient to construct a probability density function and a cumulative distribution function.
6. Construct the probability density function and cumulative distribution function for the economic measure of worth, and perform statistical analysis of the variability.

The strong advantage of the technique is that it expresses the results in probabilistic terms, thereby providing explicit assessment of risk exposure. A disadvantage is that it does not explicitly treat risk attitude; however, by providing a clear measure of risk exposure, it facilitates the implicit incorporation of risk attitude in the decision. The necessity of expressing inputs in probabilistic terms and the extensive calculations are also often considered disadvantages.

3.4.6 Decision Analysis

Decision analysis is a versatile technique that enables both risk exposure and risk attitude to be taken into account in the economic assessment. It diagrams possible choices, costs, benefits, and probabilities for a given decision problem in "decision trees," which are useful in understanding the possible choices and outcomes.

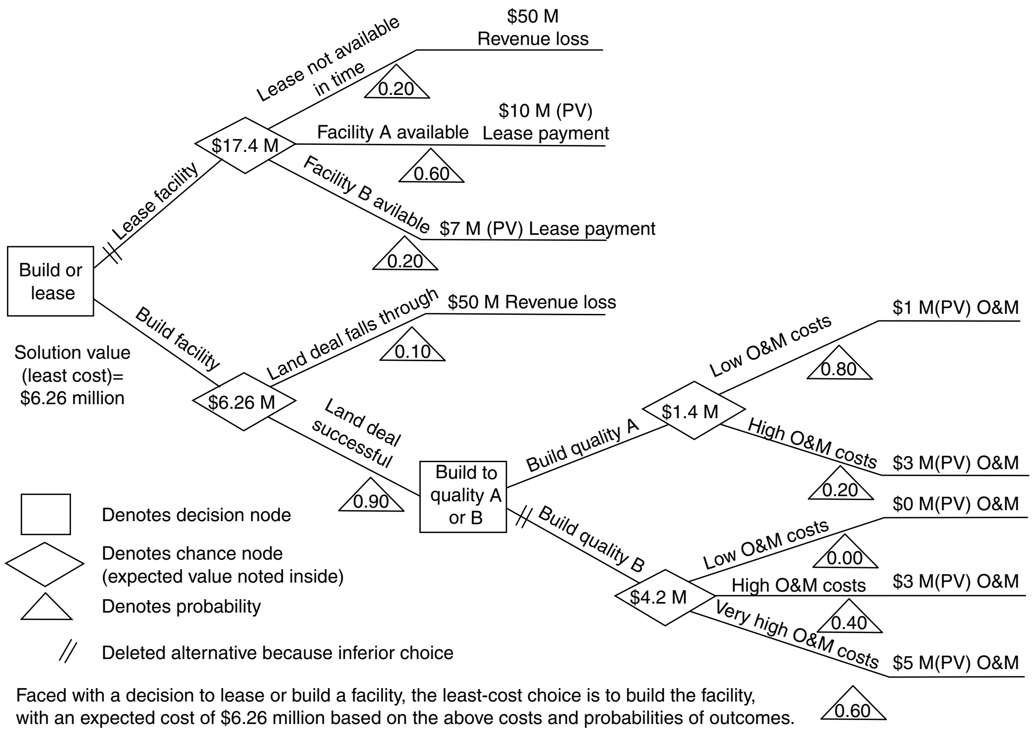


FIGURE 3.6 Decision tree: build versus lease.

Although it is not possible to capture the richness of this technique in a brief overview, a simple decision tree, shown in Figure 3.6, is discussed to give a sense of how the technique is used. The decision problem is whether to lease or build a facility. The decision must be made now, based on uncertain data. The decision tree helps to structure and analyze the problem. The tree is constructed left to right and analyzed right to left. The tree starts with a box representing a decision juncture or node—in this case, whether to lease or build a facility. The line segments branching from the box represent the two alternative paths: the upper one the lease decision and the lower one the build decision. Each has a cost associated with it that is based on the expected cost to be incurred along the path. In this example, the minimum expected cost of $6.26 M is associated with the option to build a facility.

An advantage of this technique is that it helps to understand the problem and to compare alternative solutions. Another advantage is that, in addition to treating risk exposure, it can also accommodate risk attitude by converting benefits and costs to utility values (not addressed here). A disadvantage is that the technique, as typically applied, does not provide an explicit measure of the variability of the outcome.

3.4.7 Real Options Analysis

Real options analysis (ROA) is an adaptation of financial options valuation techniques[4] to real asset investment decisions. ROA is a method used to analyze decisions in which the decision maker has one or more options regarding the timing or sequencing of an investment. It explicitly assumes that the investment is partially or completely irreversible, that there exists leeway or flexibility about the timing of

[4]Financial options valuation is credited to Fisher Black and Myron Scholes who demonstrated mathematically that the value of a European call option—an option, but not the obligation, to purchase a financial asset for a given price (i.e., the exercise or strike price) on a particular date (i.e., the expiry date) in the future—depends on the current price of the stock, the volatility of the stock's price, the expiry date, the exercise price, and the risk-free interest rate (see Black and Scholes 1973).

the investment, and that it is subject to uncertainty over future payoffs. Real options can involve options (and combinations) to defer, sequence, contract, shut down temporarily, switch uses, abandon, or expand the investment. This is in contrast to the NPV method that implies the decision is a "now or never" choice.

The value of an investment with an option is said to equal the value of the investment using the traditional NPV method (that implicitly assumes no flexibility or option) plus the value of the option. The analysis begins by construction of a decision tree with the option decision embedded in it. There are two basic methods to solve for the option value: the risk-adjusted replicating portfolio (RARP) approach and the risk-neutral probability (RNP) approach. The RARP discounts the expected project cash flows at a risk-adjusted discounted rate, whereas the RNP approach discounts certainty-equivalent cash flows at a risk-free rate. In other words, the RARP approach takes the cash flows essentially as-is and adjusts the discount rate per time period to reflect that fact that the risk changes as one moves through the decision tree (e.g., risk declines with time as more information becomes available). In the RNP approach, the cash flows themselves essentially are adjusted for risk and discounted at a risk-free rate.

Copeland and Antikarov provide an overall four-step approach for ROA:[5]

1. Step 1: Compute a base-case traditional NPV (e.g., without flexibility).
2. Step 2: Model the uncertainty using (binominal) event trees (still without flexibility; e.g., without options)—although uncertainty is incorporated, the "expected" value of Step 2 should equal that calculated in Step 1.
3. Step 3: Create a decision tree incorporating decision nodes for options, as well as other (nondecision and nonoption decisions) nodes.
4. Step 4: Conduct a ROA by valuing the payoffs, working backward in time, node by node, using the RARP or RNP approaches to calculate the ROA value of the investment.

3.4.8 Sensitivity Analysis

Sensitivity analysis is a technique for taking into account uncertainty that does not require estimates of probabilities. It tests the sensitivity of economic performance to alternative values of key factors about which there is uncertainty. Although sensitivity analysis does not provide a single answer in economic terms, it does show decision makers how the economic viability of a renewable energy or efficiency project changes as fuel prices, discount rates, time horizons, and other critical factors vary.

Figure 3.7 illustrates the sensitivity of fuel savings realized by a solar energy-heating system to three critical factors: time horizons (zero to 25 years), discount rates (D equals 0%, 5%, 10%, and 15%), and energy escalation rates (E equals 0%, 5%, 10%, and 15%). The present value of savings is based on yearly fuel savings valued initially at $1,000.

Note that, other things being equal, the present value of savings increase with time—but less with higher discount rates and more with higher escalation rates. The huge impact of fuel price escalation is most apparent when comparing the top line of the graph ($D=0.10$, $E=0.15$) with the line next to the bottom ($D=0.10$, $E=0$). The present value of savings at the end of 25 years is approximately $50,000 with a fuel escalation rate of 15%, and only about $8,000 with no escalation, other things equal. Whereas the quantity of energy saved is the same, the dollar value varies widely, depending on the escalation rate.

This example graphically illustrates a situation frequently encountered in the economic justification of energy efficiency and renewable energy projects: the major savings in energy costs, and thus the bulk of the benefits, accrue in the later years of the project and are highly sensitive to both the assumed rate of fuel-cost escalation and the discount rate. If the two rates are set equal, they will be offsetting as shown by the straight line labeled $D=0$, $E=0$ and $D=0.10$, $E=0.10$.

[5]See Dixit Avinash and Pindyck (1994), which is considered the "bible" of real options, and Copeland Antikarov (2001), which offers more practical spreadsheet methods.

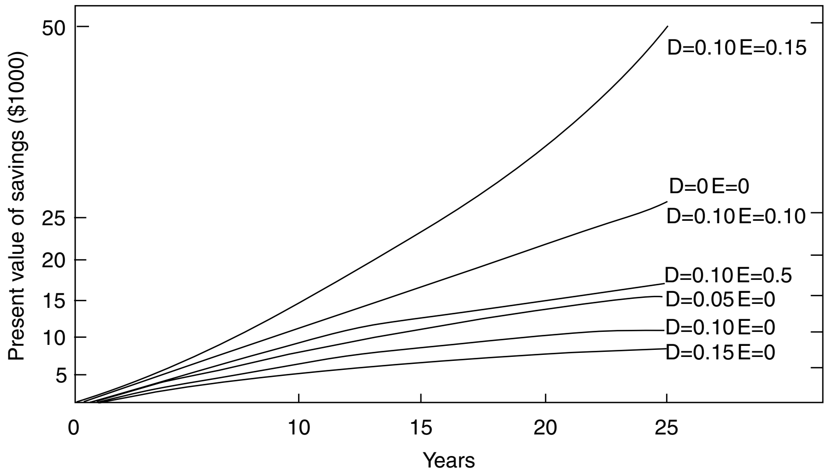


FIGURE 3.7 Sensitivity of present value energy savings to time horizons, discount rates, and energy price escalation rates.

3.5 Building Blocks of Evaluation

Beyond the formula for the basic evaluation methods and risk assessment techniques, the practitioner needs to know some of the "nuts-and-bolts" of carrying out an economic analysis. He or she needs to know how to structure the evaluation process; how to choose a method of evaluation; how to estimate dollar costs and benefits; how to perform discounting operations; how to select an analysis period; how to choose a discount rate; how to adjust for inflation; how to take into account taxes and financing; how to treat residual values; and how to reflect assumptions and constraints, among other things. This section provides brief guidelines for these topics.

3.5.1 Structuring the Evaluation Process and Selecting a Method of Evaluation

A good starting point for the evaluation process is to define the problem and the objective. Identify any constraints to the solution and possible alternatives. Consider if the best solution is obvious, or if economic analysis and risk assessment are needed to help make the decision. Select an appropriate method of evaluation and a risk assessment technique. Compile the necessary data and determine what assumptions are to be made. Apply the appropriate formula(s) to compute a measure of economic performance under risk. Compare alternatives and make the decision, taking into account any incommensurable effects that are not included in the dollar benefits and costs. Take into account the risk attitude of the decision maker, if it is relevant.

Although the six evaluation methods given earlier are similar, they are also sufficiently different in that they are not always equally suitable for evaluating all types of energy investment decisions. For some types of decisions, the choice of method is more critical than for others. Figure 3.8 categorizes different investment types and the most suitable evaluation methods for each. If only a single investment is being considered, the "accept–reject" decision can often be made by any one of several techniques, provided the correct criterion is used.

The accept/reject criteria are as follows:

- LCC technique: LCC must be lower as a result of the energy efficiency or renewable energy investment than without it.
- NPV (NB) technique: NPV must be positive as a result of the investment.

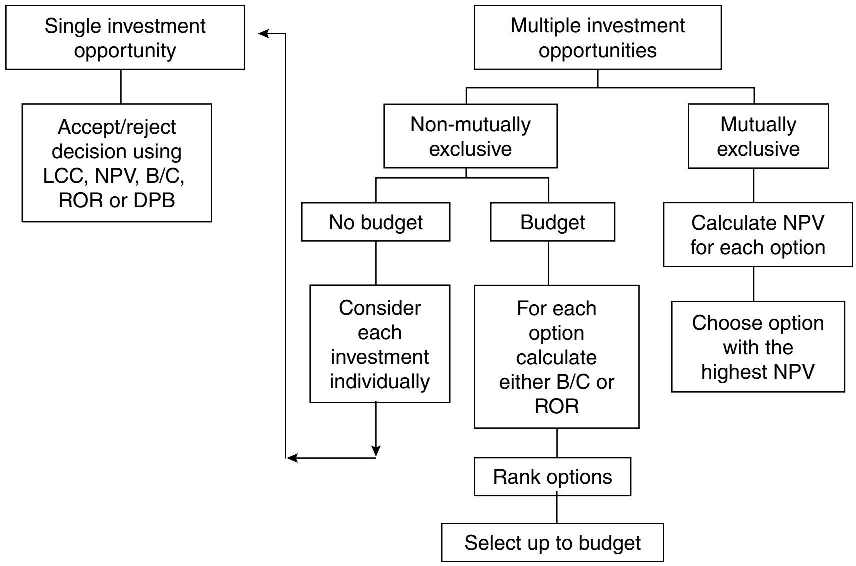


FIGURE 3.8 Investment decisions and evaluation methods.

- B/C (SIR) technique: B/C (SIR) must be greater than one.
- IRR technique: the IRR must be greater than the investor's minimum acceptable rate of return.
- DPB technique: the number of years to achieve DPB must be less than the project life or the investor's time horizon, and there are no cash flows after payback is achieved that would reverse payback.

If multiple investment opportunities are available, but only one investment can be made (i.e., they are mutually exclusive), any of the methods (except DPB) will usually work, provided they are used correctly. However, the NPV method is usually recommended for this purpose, because it is less likely to be misapplied. The NPV of each investment is calculated and the investment with the highest present value is the most economic. This is true even if the investments require significantly different initial investments, have significantly different times at which the returns occur, or have different useful lifetimes. Examples of mutually exclusive investments include different system sizes (e.g., three different photovoltaic array sizes are being considered for a single rooftop), different system configurations (e.g., different turbines are being considered for the same wind farm), and so forth.

If the investments are not mutually exclusive, then (as shown in Figure 3.8), one must consider whether there is an overall budget limitation that would restrict the number of economic investments that might be undertaken. If there is no budget (i.e., no limitation on the investment funds available), than there is really no comparison to be performed and the investor simply makes an accept–reject decision for each investment individually as described above.

If funds are not available to undertake all of the investments (i.e., there is a budget), then the easiest approach is to rank the alternatives, with the best having the highest BCR or rate of return. (The investment with the highest NPV will not necessarily be the one with the highest rank, because present value does not show return per unit investment). Once ranked, those investments at the top of the priority list are selected until the budget is exhausted.

In the case where a fast turnaround on investment funds is required, DPB is recommended. The other methods, although more comprehensive and accurate for measuring an investment's lifetime profitability, do not indicate the time required for recouping the investment funds.

3.5.2 Discounting

Some or all investment costs in energy efficiency or renewable energy systems are incurred near the beginning of the project and are treated as "first costs." The benefits, on the other hand, typically accrue over the life span of the project in the form of yearly energy saved or produced. To compare benefits and costs that accrue at different points in time, it is necessary to put all cash flows on a time-equivalent basis. The method for converting cash flows to a time-equivalent basis is often called *discounting*.

The value of money is time-dependent for two reasons: First, inflation or deflation can change the buying power of the dollar. Second, money can be invested over time to yield a return over and above inflation. For these two reasons, a given dollar amount today will be worth more than that same dollar amount a year later. For example, suppose a person were able to earn a maximum of 10% interest per annum risk-free. He or she would require $1.10 a year from now to be willing to forego having $1 today. If the person were indifferent between $1 today and $1.10 a year from now, then the 10% rate of interest would indicate that person's time preference for money. The higher the time preference, the higher the rate of interest required to make future cash flows equal to a given value today. The rate of interest for which an investor feels adequately compensated for trading money now for money in the future is the appropriate rate to use for converting present sums to future equivalent sums and future sums to present equivalent sums (i.e., the rate for discounting cash flows for that particular investor). This rate is often called the *discount rate*.

To evaluate correctly the economic efficiency of an energy efficiency or renewable energy investment, it is necessary to convert the various expenditures and savings that accrue over time to a lump-sum, time-equivalent value in some base year (usually the present), or to annual values. The remainder of this section illustrates how to discount various types of cash flows.

Discounting is illustrated by Figure 3.9 in a problem of installing, maintaining, and operating a heat pump, as compared to an alternative heating/cooling system. The LCC calculations are shown for two reference times. The first is the present, and it is therefore called a *present value*. The second is based on a yearly time scale and is called an *annual value*. These two reference points are the most common in economic-evaluations of investments. When the evaluation methods are derived properly, each time basis will give the same relative ranking of investment priorities.

The assumptions for the heat pump problem—which are given only for the sake of illustration and not to suggest actual prices—are as follows:

1. The residential heat pump (not including the ducting) costs $1,500 to purchase and install.
2. The heat pump has a useful life of 15 years.
3. The system has annual maintenance costs of $50 every year during its useful life, fixed by contractual agreement.
4. A compressor replacement is required in the eighth year at a cost of $400.
5. The yearly electricity cost for heating and cooling is $425, evaluated at the outset, and increased at a rate of 7% per annum due to rising electricity prices.
6. The discount rate (a nominal rate that includes an inflation adjustment) is 10%.
7. No salvage value is expected at the end of 15 years.

The LCCs in the sample problem are derived only for the heat pump and not for alternative heating/cooling systems. Hence, no attempt is made to compare alternative systems in this discounting example. To do so would require similar calculations of LCCs for other types of heating/cooling systems. Total costs of a heat pump system include costs of purchase and installation, maintenance, replacements, and electricity for operation. Using the present as the base-time reference point, one needs to convert each of these costs to the present before summing them. Assuming that the purchase and installation costs occur at the base reference point (the present), the $1,500 is already in present value terms.

Figure 3.9 illustrates how to convert the other cash flows to present values. The first task is to convert the stream of annual maintenance costs to present value. The maintenance costs, as shown in the cash flow diagram of Figure 3.9, are $50 per year, measured in current dollars (i.e., dollars of the years in which

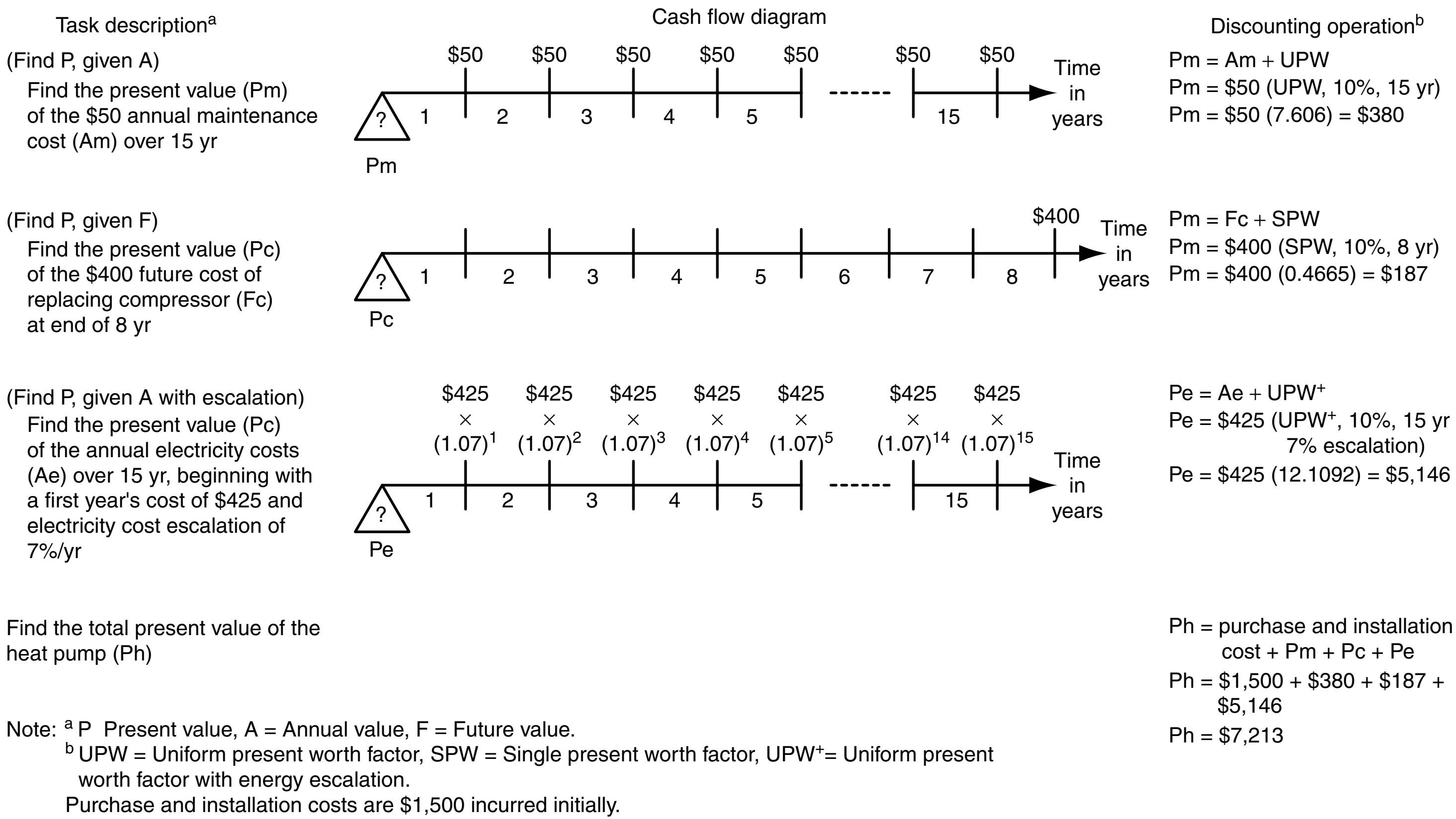


FIGURE 3.9 Discounting for present value: a heat pump example. (From Marshall, H. E. and Ruegg, R. T., in *Simplified Energy Design Economics*, F. Wilson, ed., National Bureau of Standards, Washington, DC, 1980.)

TABLE 3.4 Discount Formulas

Standard Nomenclature	Use When	Standard Notation	Algebraic Form
Single compound amount	Given P; to find F	(SCA, d%, N)	$F=P(1+d)^N$
Single present worth	Given F; to find P	(SPW, d%, N)	$P=F\frac{1}{(1+d)^N}$
Uniform compound amount	Given A; to find F	(UCA, d%, N)	$F=A\frac{(1+d)^N-1}{d}$
Uniform sinking fund	Given F; to find A	(USF, d %, N)	$A=F\frac{d}{(1+d)^N-1}$
Uniform capital recovery	Given P; to find A	(UCR, d%, N)	$A=P\frac{d(1+d)^N}{(1+d)^N-1}$
Uniform present worth	Given A; to find P	(UPW, d%, N)	$P=A\frac{(1+d)^N-1}{d(1+d)^N}$
Uniform present worth modified	Given A escalating at a rate e; to find P	(UPW*, d%, e, N)	$P=A\frac{(1+e)\left[1-\left(\frac{1+e}{1+d}\right)^N\right]}{(d-e)}$

P, a present sum of money; F, a future sum of money, equivalent to P at the end of N periods of time at a discount rate of d; N, number of interest periods; A, an end-of-period payment (or receipt) in a uniform series of payments (or receipts) over N periods at *discount rate d*, usually annually; e, a rate of escalation in A in each of N periods.

they occur). The triangle indicates the value to be found. The practice of compounding interest at the end of each year is followed here. The present refers to the beginning of year one.

The discounting operation for calculating the present value of maintenance costs (last column of Figure 3.9) is to multiply the annual maintenance costs times the uniform present worth (UPW) factor. The UPW is a multiplicative factor computed from the formula given in Table 3.4, or taken from a look-up table of factors that have been published in many economics textbooks. UPW factors make it easy to calculate the present values of a uniform series of annual values. For a discount rate of 10% and a time period of 15 years, the UPW factor is 7.606. Multiplying this factor by $50 gives a present value maintenance cost equal to $380. Note that the $380 present value of $50 per year incurred in each of 15 years is much less than simply adding $50 for 15 years (i.e., $750). Discounting is required to achieve correct statements of costs and benefits over time.

The second step is to convert the one-time future cost of compressor replacement, $400, to its present value. The operation for calculating the present value of compressor replacement is to multiply the future value of the compressor replacement times the single-payment present worth factor (SPW) that can be calculated from the formula in Table 3.4, or taken from a discount factor look-up table. For a discount rate of 10% and a time period of 15 years, the SPW factor is 0.4665. Multiplying this factor by $400 gives a present-value cost of the compressor replacement of $187, as shown in the last column of Figure 3.9. Again, note that discounting makes a significant difference in the measure of costs. Failing to discount the $400 would result in an overestimate of cost in this case of $213.

The third step is to convert the annual electricity costs for heating and cooling to present value. A year's electricity costs, evaluated at the time of installation of the heat pump, are assumed to be $425. Electricity prices, for purposes of illustration, are assumed to increase at a rate of 7% per annum. This is reflected in Table 3.4 by multiplying $425 times $(1.07)^t$ where $t=1,2,\ldots,15$. The electricity cost at the end of the fourth year, for example, is $\$425(1.07)^4=\$557$.

The discounting operation for finding the present value of all electricity costs (shown in Figure 3.9) is to multiply the initial, yearly electricity costs times the appropriate UPW* factor. (An asterisk following UPW denotes that a term for price escalation is included.) The UPW or UPW* discount formulas in Table 3.4 can also be used to obtain present values from annual costs or multiplicative discount factors from look-up tables can be used. For a period of 15 years, a discount rate of l0%, and an escalation rate of 7%, the UPW* factor is 12.1092. Multiplying the factor by $425 gives a present value of electricity costs of $5,146. Note, once again, that failing to discount (i.e., simply adding annual electricity expenses in current prices) would overestimate costs by $1,229 ($6,376 − $5,146). Discounting with a UPW factor that does not incorporate energy price escalation would underestimate costs by $1,913 ($5,146 − $3,233).

The final operation described in Figure 3.9 is to sum purchase and installation cost and the present values of maintenance, compressor replacement, and electricity costs. Total LCCs of the heat pump in present value terms are $7,213. This is one of the amounts that a designer would need for comparing the cost-effectiveness of heat pumps to alternative heating/cooling systems.

Only one discounting operation is required for converting the present- value costs of the heat pump to annual value terms. The total present-value amount is converted to the total annual value simply by multiplying it by the uniform capital recovery factor (UCR)—in this case the UCR for 10% and 15 years. The UCR factor, calculated with the UCR formula found in Table 3.4, is 0.13147. Multiplying this factor by the total present value of $7,213 gives the cost of the heat pump as $948 in annual value terms. The two figures—$7,213 and $948 per year—are time-equivalent values, made consistent through the discounting.

Figure 3.9 provides a model for the designer who must calculate present values from all kinds of benefit or cost streams. Most distributions of values occurring in future years can be handled with the SPW, the UPW, or the UPW* factors.

3.5.3 Discount Rate

Of the various factors affecting the NB of energy efficiency and renewable energy investments, the discount rate is one of the most dramatic. A project that appears economic at one discount rate will often appear uneconomic at another rate. For example, a project that yields net savings at a 6% discount rate might yield net losses if evaluated with a 7% rate.

As the discount rate is increased, the present value of any future stream of costs or benefits is going to become smaller. High discount rates tend to favor projects with quick payoffs over projects with benefits deferred further in the future.

The discount rate should be set equal to the rate of return available on the next-best investment opportunity of similar risk to the project in question, i.e., it should indicate the opportunity cost of the investor.

The discount rate may be formulated as a "real rate" exclusive of general price inflation or as a "nominal rate" inclusive of inflation. The former should be used to discount cash flows that are stated in constant dollars. The latter should be used to discount cash flows stated in current dollars.

3.5.4 Inflation

Inflation is a rise in the general price level. Because future price changes are unknown, it is frequently assumed that prices will increase at the rate of inflation. Under this assumption, it is generally easier to conduct all economic-evaluations in constant dollars and to discount those values using "real" discount rates. For example, converting the constant dollar annual maintenance costs in Figure 3.9 to a present value can be easily done by multiplying by a uniform present-worth factor because the maintenance costs do not change over time. However some cash flows are more easily expressed in current dollars, e.g., equal loan payments, tax depreciation. These can be converted to present values using a nominal discount rate.

3.5.5 Analysis Period

The analysis period is the length of time over which costs and benefits are considered in an economic-evaluation. The analysis period need not be the same as either the "useful life" or the "economic life," two common concepts of investment life. The useful life is the period over which the investment has some value; i.e., the investment continues to conserve or provide energy during this period. Economic life is the period during which the investment in question is the least-cost way of meeting the requirement. Often economic life is shorter than useful life.

The selection of an analysis period will depend on the objectives and perspective of the decision maker. A speculative investor who plans to develop a project for immediate sale, for example, may view the relevant time horizon as that short period of ownership from planning and acquisition of property to the first sale of the project. Although the useful life of a solar domestic hot water heating system, for example, might be 20 years, a speculative home builder might operate on the basis of a two-year time horizon, if the property is expected to change hands within that period. Only if the speculator expects to gain the benefit of those energy savings through a higher selling price for the building will the higher first cost of the solar energy investment likely be economic.

If an analyst is performing an economic analysis for a particular client, that client's time horizon should serve as the analysis period. If an analyst is performing an analysis in support of public investment or a policy decision, the life of the system or building is typically the appropriate analysis period.

When considering multiple investment options, it is best with some evaluation methods (such as LCC, IRR, and ORR) to use the same analysis period. With others like NPV and BCR, different analysis periods can be used. If an investment's useful life is shorter than the analysis period, it may be necessary to consider reinvesting in that option at the end its useful life. If an investment's useful life is longer than the analysis period, a salvage value may need to be estimated.

3.5.6 Taxes and Subsidies

Taxes and subsidies should be taken into account in economic-evaluations because they may affect the economic viability of an investment, the return to the investor, and the optimal size of the investment. Taxes, which may have positive and negative effects, include—but are not limited to—income taxes, sales taxes, property taxes, excise takes, capital gain taxes, depreciation recapture taxes, tax deductions, and tax credits.

Subsidies are inducements for a particular type of behavior or action. They include grants—cash subsidies of specified amounts; government cost sharing; loan-interest reductions, and tax-related subsidies. Income tax credits for efficiency or renewable energy expenditures provide a subsidy by allowing specific deductions from the investor's tax liability. Property tax exemptions eliminate the property taxes that would otherwise add to annual costs. Income tax deductions for energy efficiency or renewable energy expenses reduce annual tax costs. The imposition of higher taxes on nonrenewable energy sources raises their prices and encourages efficiency and renewable energy investments.

It is important to distinguish between a before-tax cash flow and an after-tax cash flow. For example, fuel costs are a before-tax cash flow (they can be expensed), whereas a production tax credit for electricity from wind is an after-tax cash flow.

3.5.7 Financing

Financing of an energy investment can alter the economic viability of that investment. This is especially true for energy efficiency and renewable energy investments that generally have large initial investment costs with returns spread out over time. Ignoring financing costs when comparing these investments against conventional sources of energy can bias the evaluation against the energy efficiency and renewable energy investments.

Financing is generally described in terms of the amount financed, the loan period, and the interest rate. Unless specified otherwise, a uniform payment schedule is usually assumed. Generally, financing improves the economic effectiveness of an investment if the after-tax nominal interest rate is less than the investor's nominal discount rate.

Financing essentially reduces the initial outlay in favor of additional future outlays over time—usually equal payments for a fixed number of years. These cash flows can be treated like any other: The equity portion of the capital cost occurs at the start of the first year, and the loan payments occur monthly or annually. The only other major consideration is the tax deductibility of the interest portion of the loan payments.

3.5.8 Residual Values

Residual values may arise from salvage (net of disposal costs) at the end of the life of systems and components, from reuse values when the purpose is changed, and from remaining value when assets are sold prior to the end of their lives. The present value of residuals can generally be expected to decrease, other things equal, as (1) the discount rate rises, (2) the equipment or building deteriorates, and (3) the time horizon lengthens.

To estimate the residual value of energy efficiency or renewable energy systems and components, it is helpful to consider the amount that can be added to the selling price of a project or building because of those systems. It might be assumed that a building buyer will be willing to pay an additional amount equal to the capitalized value of energy savings over the remaining life of the efficiency or renewable investment. If the analysis period is the same as the useful life, there will be no residual value.

3.6 Summary

There are multiple methods of economic performance and multiple techniques of risk analysis that can be selected and combined to improve decisions in energy efficiency and renewable energy investments. Economic performance can be stated in a variety of ways, depending on the problem and preferences of the decision maker: as NPV, as LCCs, as the cost of energy, as a rate of return, as years to payback, or as a ratio. To reflect the reality that most decisions are made under conditions of uncertainty, risk assessment techniques can be used to reflect the risk exposure of the project and the risk attitude of the decision maker. Rather than expressing results in single, deterministic terms, they can be expressed in probabilistic terms, thereby revealing the likelihood that the outcome will differ from the best-guess answer. These methods and techniques can be used to decide whether or not to invest in a given energy efficiency or renewable energy system; to determine which system design or size is economically efficient; to find the combination of components and systems that are expected to be cost-effective; to estimate how long before a project will break even; and to decide which energy-related investments are likely to provide the highest rate of return to the investor. The methods support the goal of achieving economic efficiency—which may differ from technical efficiency.

Glossary

Analysis period: Length of time over which costs and benefits are considered in an economic-evaluation.

Benefit/cost (B/C) or saving-to-investment (SIR) ratio: A method of measuring the economic performance of alternatives by dividing present value benefits (savings) by present value costs.

Constant dollars: Values expressed in terms of the general purchasing power of the dollar in a base year. Constant dollars do not reflect price inflation or deflation.

Cost-effective investment: The least-cost alternative for achieving a given level of performance.

Current dollars: Values expressed in terms of actual prices of each year (i.e., current dollars reflect price inflation or deflation).

Discount rate: Based on the opportunity cost of capital, this minimum acceptable rate of return is used to convert benefits and costs occurring at different times to their equivalent values at a common time.

Discounted payback period: The time required for the discounted annual NB derived from an investment to pay back the initial investment.

Discounting: A technique for converting cash flows that occur over time to equivalent amounts at a common point in time using the opportunity cost for capital.

Economic efficiency optimization: Maximizing NB or minimizing costs for a given level of benefits (i.e., "getting the most for your money").

Economic life: That period of time over which an investment is considered to be the least-cost alternative for meeting a particular objective.

Future value (worth): The value of a dollar amount at some point in the future, taking into account the opportunity cost of capital.

Internal rate of return: The discount rate that equates total discounted benefits with total discounted costs.

Investment costs: The sum of the planning, design, and construction costs necessary to obtain or develop an asset.

Levelized cost of energy: The before-tax revenue required per unit of energy to cover all costs plus a profit/return on investment equal to the discount rate used to levelize the costs.

Life-cycle cost: The present-value total of all relevant costs associated with an asset or project over the analysis period.

Net benefits: Benefits minus costs.

Present value (worth): Past, present, or future cash flows all expressed as a lump sum amount as of the present time, taking into account the time value of money.

Real options analysis: Method used to analyze investment decisions in which the decision maker has one or more options regarding the timing or sequencing of investment.

Risk assessment: As applied to economic decisions, the body of theory and practice that helps decision makers assess their risk exposures and risk attitudes to increase the probability that they will make economic choices that are best for them.

Risk attitude: The willingness of decision makers to take chances on investments with uncertain outcomes. Risk attitudes may be classified as *risk averse*, *risk neutral*, and *risk taking*.

Risk exposure: The probability that a project's economic outcome will be less favorable than what is considered economically desirable.

Sensitivity analysis: A non-probability-based technique for reflecting uncertainty that entails testing the outcome of an investment by altering one or more system parameters from the initially assumed values.

Time value of money: The amount that people are willing to pay for having money today rather than some time in the future.

Uncertainty: As used in the context of this chapter, a lack of knowledge about the values of inputs required for an economic analysis.

References

Black, F. and Scholes, M. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81, 637–659.

Copeland, T. and Antikarov, V. 2001. *Real Options: A Practitioners Guide*, TEXERE, New York.

Dixit, A. K. and Pindyck, R. S. 1994. *Investment Under Uncertainty*. Princeton University Press, Princeton, NJ.

Marshall, H. E. and Ruegg, R. T. 1980. Principles of economics applied to investments in energy conservation and solar energy systems, In *Economics of Solar Energy and Conservation Systems*, F. Kreith and R. West, eds. pp. 123–173. CRC Press, Boca Raton, FL.

Marshall, H. E. and Ruegg, R. T. 1980. In *Simplified Energy Design Economics*, F. Wilson, ed., National Bureau of Standards, Washington, DC.

Ruegg, R. T. and Marshall, H. E. 1990. *Building Economics: Theory and Practice*. Chapman and Hall, New York.

Ruegg, R. T. and Short, W. 1988. Economic methods. In *Economic Analysis of Solar Thermal Energy Systems*, R. West and F. Kreith, eds. pp. 18–83. MIT Press, Cambridge, MA.

Stermole, F. J. and Stermole, J. M. 2000. *Economic Evaluation and Investment Decision Methods,* Investment Evaluations Corp., Lakewood, CO.

4

Environmental Impacts and Costs of Energy

Ari Rabl and
Joseph V. Spadaro
Centre d'Energetique

4.1 Introduction

Air pollution causes considerable damage to human health, flora and fauna, and materials. The damage costs are externalities to the extent that these costs are not reflected in the prices of goods. When making decisions that affect the emission of pollutants, the damage costs, also called external costs, should be considered. The external costs can be internalized via taxes, tradable permits, or other environmental regulations..

In recent years, there has been much progress in the analysis of environmental damage costs due mostly to several major projects evaluating the external costs of energy in Europe (ExternE 1995, 1998, 2000, 2004) and also in the United States (Lee 1994; Rowe et al. 1995). Of these projects the ExternE (External Costs of Energy) Project of the European Commission has the widest scope and is the most up to date. This paper, authored by participants of all ExternE Projects since 1992, presents an overview of the methodology and the results.

The quantification of damage costs has many important applications:

- Guidance for environmental regulations (e.g., determining the optimal level of the limit for the emission of a pollutant)

- Finding the socially optimal level of a pollution tax
- Identifying technologies with the lowest social cost (e.g., coal, natural gas, or nuclear power for the production of electricity)
- Evaluating the benefits of improving the pollution abatement of an existing installation such as a waste incinerator
- Optimizing the dispatching of power plants
- "Green accounting," i.e., including corrections for environmental damage in the traditional accounts of GNP

4.2 Methodology

4.2.1 Impact Pathway Analysis

To calculate the damage costs, one needs to carry out an impact pathway analysis (IPA), tracing the passage of a pollutant from where it is emitted to the affected receptors (population, crops, forests, buildings, etc.). The principal steps of an IPA can be grouped as follows, as shown in Figure 4.1:

1. Emission: specification of the relevant technologies and pollutants, e.g., kg of NO_x per GWh_e emitted by power plant

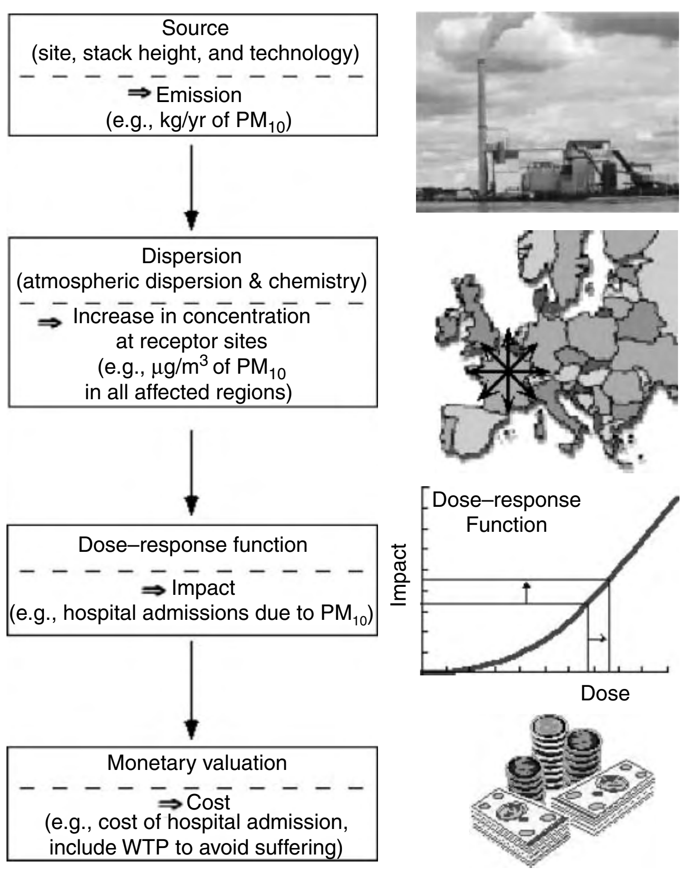


FIGURE 4.1 Impact pathway analysis.

2. Dispersion: calculation of increased pollutant concentrations in all affected regions, e.g., incremental concentration of ozone, using models of atmospheric dispersion and chemistry for ozone formation due to NO_x (this step is also called environmental fate analysis, especially when it involves more complex pathways that pass through the food chain)
3. Impact: calculation of the dose from the increased concentration and calculation of impacts (damage in physical units) from this dose, using a dose–response function (DRF), e.g., cases of asthma due to this increase in ozone
4. Cost: economic valuation of these impacts, e.g., multiplication by the cost of a case of asthma

The impacts and costs are summed over all receptors of concern. The work involves a multidisciplinary system analysis, with inputs from engineers, dispersion modelers, epidemiologists, ecologists, and economists. The result of an IPA is the damage cost per kg of emitted pollutant, as shown in Figure 4.5 of Section 4.3. The steps of the IPA are described in the following sections.

The reader may wonder about the relationship between an IPA and an environmental impact study (EIS) that is required before the approval of a proposed installation (factory, power plant, incinerator, etc.). The purpose of an EIS is to ensure that nobody is exposed to an unacceptable risk or burden. Because the highest exposures are imposed in the local zone, it is sufficient for an EIS to focus on local analysis, up to 10 km depending on the case. Thus, an EIS provides the possibility of a veto if a proposed installation is considered unacceptable. By contrast, the calculation of total damage costs requires an IPA where the damages are summed over all affected receptors. Regarding most air pollutants emitted in Europe the affected receptors are the entire continent, and in the case of greenhouse gases (GHGs) it is the entire globe. Damage costs are needed primarily by decision makers at the national or international level, or generally by anyone concerned with total impacts.

For many environmental choices, one needs to look not only at a particular source of pollutants but also must take into account an entire process chain by means of a life cycle assessment (LCA). For example, a comparison of power generation technologies involves an analysis of the fuel chain sketched in Figure 4.2. Whether an IPA of a single source or an LCA of an entire cycle is required depends on the policy decision in question. When finding the optimal limit for the emission of NO_x from an incinerator, an IPA is sufficient, but the choice between incineration and landfill of waste involves an LCA.

In principle, a site-specific IPA would evaluate the damages and costs for each pollution source in the life cycle. In practice, however, most LCAs have taken the shortcut of first summing the emissions over all stages and then multiplying the result by site-independent impact indices. Additionally, most practitioners of LCA reject the concept of monetary valuation, preferring instead to use approximately ten nonmonetary indicators of “potential impact” that are based on expert judgment.

4.2.2 Dispersion of Pollutants

The principal GHGs, CO_2, CH_4, and N_2O, stay in the atmosphere long enough to mix uniformly over the entire globe. No specific dispersion calculation is needed. However, the calculation of impacts is extraordinarily complex and the reader is referred merely to the main authority, the Intergovernmental Panel on Climate Change (IPCC, http://www.ipcc.ch). For most other air pollutants, in particular PM_{10} (particulate matter with diameter less than 10 μm), NO_x and SO_2, atmospheric dispersion is significant over hundreds to thousands of km, so both local and regional effects are important. ExternE uses a combination of local and regional dispersion models to account for all significant damages. The main model for the local range ($<$ 50 km from the source) has been the Gaussian plume model ISC (Brode and Wang 1992).

At the regional scale one needs to take into account the chemical reactions that lead to the transformation of primary pollutants (i.e., the pollutants as they are emitted) to secondary pollutants. For example when studying the creation of sulfates from SO_2, ExternE uses the Windrose trajectory model (WTM) (Trukenmüller and Friedrich 1995) to estimate the concentration and deposition of acid species. WTM is a user-configurable Lagrangian trajectory model, derived from the Harwell trajectory model (Derwent and Nodop 1986). The modeling of ozone is based on the EMEP MSC-W oxidant model

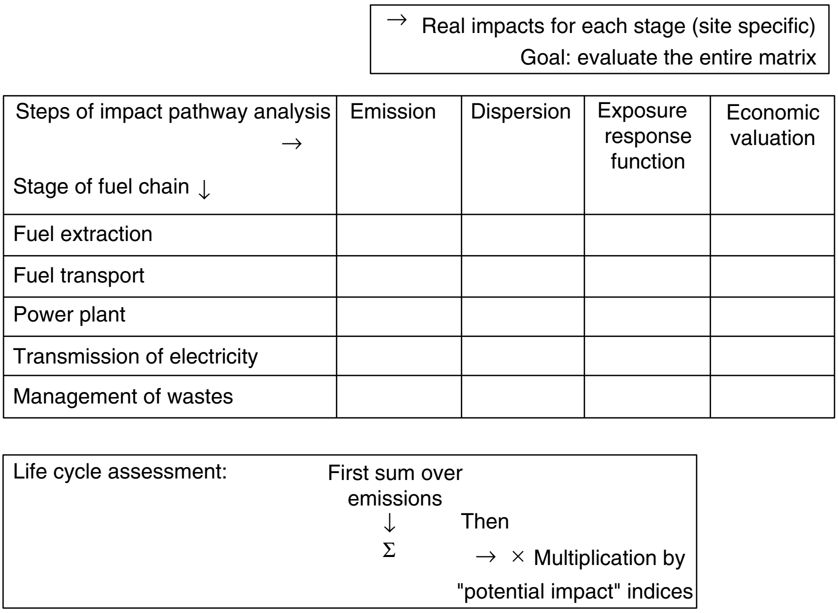


FIGURE 4.2 Relation between impact pathway analysis and current practice of most LCA, illustrated for the example of electricity production. (From Spadaro, J. V. and Rabl, A. 1999. *International Journal of Life Cycle Assessment*, 4(4), 229–243. With permission.)

(Simpson 1992; Simpson and Eliassen 1997). EMEP is the official model used for policy decisions about transboundary air pollution in Europe.

The calculation of damage costs is carried out by using the EcoSense software package (Krewitt et al. 1995), an integrated impact assessment model that combines these atmospheric models with databases for receptors (population, land use, agricultural production, buildings and materials, etc.), DRFs and monetary values. J. V. Spadaro also has developed a simplified analysis tool called RiskPoll (actually a package of several models with different input requirements) that is freely available from www.arirabl.org or www.externe.info. This tool is based on the interpolation of dispersion calculations by EcoSense. Its simplest version yields results that are typically within a factor of two to three of detailed EcoSense calculations for stack heights above 50 m. RiskPoll includes a module for the multimedia pathways of Figure 4.3.

Several tests have been done to confirm the accuracy of the results. The tests have checked the consistency between ISC and ROADPOL, and the concentrations predicted by WTM were compared with measured data and with calculations of the EMEP program, the official program for the modeling of acid rain in Europe.

Whereas only the inhalation dose matters for PM_{10}, NO_x, SO_2 and O_3, toxic metals and persistent organic pollutants affect humans also through food and drink. These pollutants require a much more complex IPA to calculate ingestion doses. Spadaro and Rabl (2004) have developed a model for the assessment of external costs due to the emission of the most toxic metals (As, Cd, Cr, Hg, Ni, and Pb), as well as certain organic pollutants, in particular dioxins. The model takes into account the pathways in Figure 4.3. The output of this model is the damage per kg of pollutant, as a function of the site and conditions (for emissions to air: stack height, exhaust temperature, and velocity) of the source. The model is based mostly on transfer factors published by EPA (1998), with some supplemental data of IAEA (1994, 2001). These transfer factors account in a simple manner for the transport of a pollutant between different environmental compartments, like the uptake by agricultural crops of a pollutant from the soil. The uncertainties are large, but at least approximate values for the pollutants of concern are available.

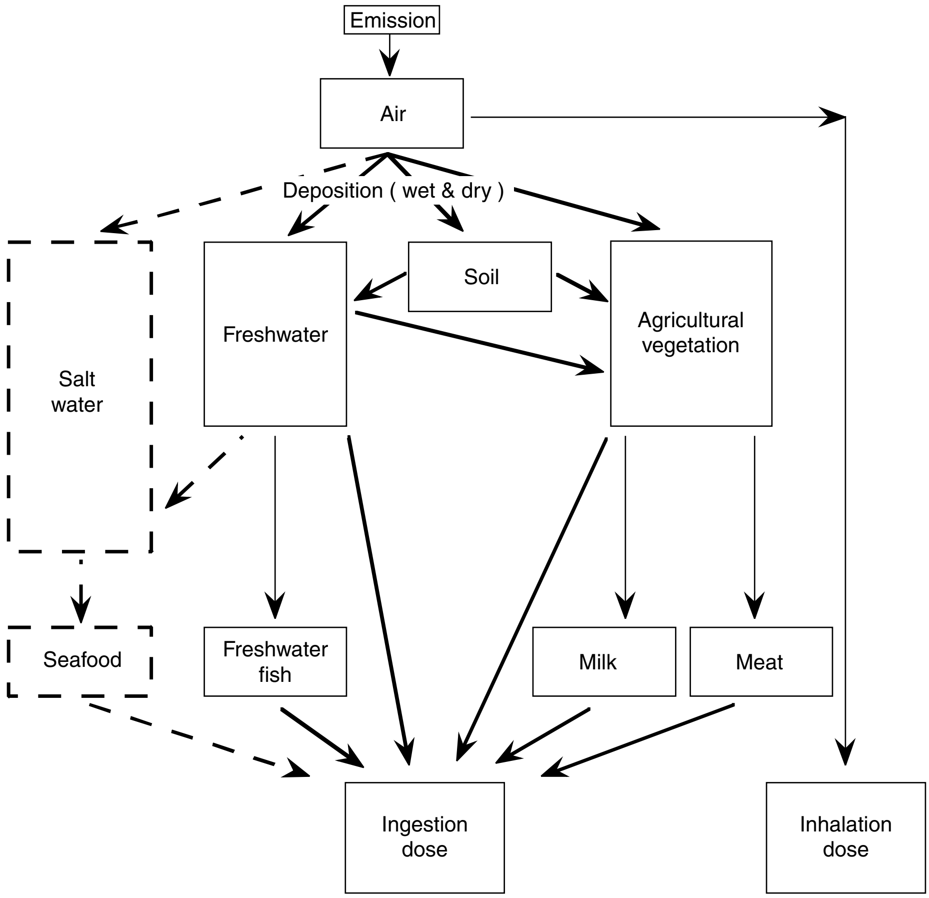


FIGURE 4.3 Pathways taken into account for health impacts of air pollutants. Direct emissions to soil or water are a special case where the analysis begins at the respective "soil" and "water" boxes. In the present version of the model seafood is not yet included.

It is not yet possible to have all of the elements for calculating the dose due to ingestion of seafood. The dose may be potentially large because of bioconcentration and the fact that most fish are oceanic rather than freshwater. Even if the concentration increment in the sea is very small, the collective dose from seafood could be significant if the removal processes (sedimentation) are slow, and the analysis has no cutoff in time.

A general result of this analysis is that when these pollutants are emitted into the air, the ingestion dose can be approximately two orders of magnitude larger than the dose by inhalation. Because nowadays most food is transported over very large distances, the total dose varies little with the site where these pollutants are emitted into the air. As far as damages are concerned, one has to note that the same dose can have a very different effect on the body depending on whether it is inhaled or ingested. Cd, Cr^{VI}, and Ni, for instance, are carcinogenic only through inhalation according to current knowledge.

4.2.3 Dose–Response Functions: General Considerations

The DRF relates the quantity of a pollutant that affects a receptor (e.g. population) to the physical impact on this receptor (e.g., incremental number of hospitalizations). In the narrowest sense of the term, it should be based on the dose actually absorbed by a receptor. However, the term DRF is often used in a wider sense where it is formulated directly in terms of the concentration of a pollutant in the ambient air, accounting implicitly for the absorption of the pollutant from the air into the body. The functions for the classical air pollutants (NO_x, SO_2, O_3, and particulates) are typically of this kind, and the terms exposure–response function or concentration–response function (CRF) are often used.

The DRF is a central ingredient in the IPA, thus meriting special attention. Damage can be quantified only if the corresponding DRF is known. Such functions are available for the impacts on human health, building materials, and crops caused by a range of pollutants such as primary and secondary (i.e., nitrates, sulfates) particles, ozone, CO, SO_2, NO_x, benzene, dioxins, As, Cd, Cr, Ni, and Pb. The most comprehensive reference for health impacts is the IRIS database of EPA (http://www.epa.gov/iriswebp/iris/index.html). For the application in an IPA, the information often has to be expressed in somewhat different form, due to accounting for additional factors such as the incidence rate (ExternE 1998; Spadaro and Rabl 2004). Unfortunately, the DRFs are very uncertain or not known for many pollutants and many impacts. For most substances and noncancer impacts, the only available information covers thresholds, typically the NOAEL (no observed adverse effect level) or LOAEL (lowest observed adverse effect level). Knowing thresholds is not sufficient for quantifying impacts; it only provides an answer to the question whether or not there is a risk. The principal exceptions are carcinogens and the classical air pollutants, for which explicit DRFs are known (often on the assumption of linearity and no threshold).

By definition, a DRF starts at the origin, and in most cases it increases monotonically with dose, as sketched schematically in Figure 4.4. At very high doses, the function may level off in S-shaped fashion due to saturation, but that case is not of interest here. DRFs for health are determined from epidemiological studies or laboratory studies. Because the latter are mostly limited to animals, the extrapolation to humans introduces large uncertainties.

A major difficulty lies in the fact that unless the sample is very large, relatively high doses are needed to obtain observable nonzero responses; such doses are usually far in excess of typical ambient concentrations in the EU or North America. Thus there is a serious problem of how to extrapolate from the observed data towards low doses. Figure 4.4 indicates several possibilities for the case where the point P corresponds to the lowest dose at which a response has been measured. The simplest is the linear model, i.e., a straight line from the origin through the observed data point(s). The available evidence suggests that a DRF is unlikely to go above this straight line in the low dose limit. However, the straight-line model does appear to be appropriate in many cases, especially for many cancers. In fact, most estimates of cancers due to chemicals or radiation assume this linear behavior.

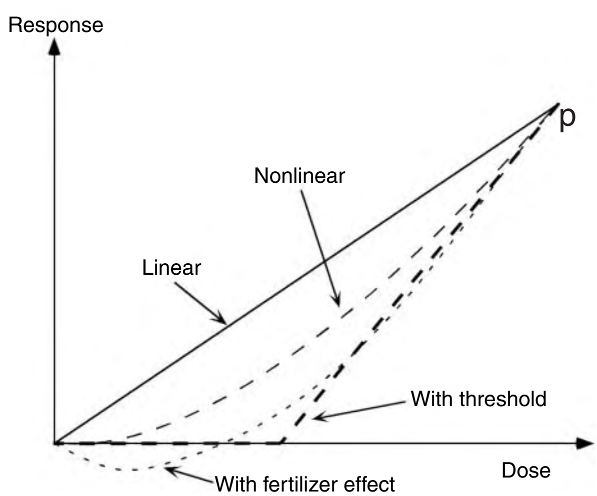


FIGURE 4.4 Possible behavior of dose–response functions at low doses. If *P* is the lowest dose where a nonzero impact has been observed, the extrapolation to lower doses is uncertain but values higher than linear are unlikely.

Another possibility is the “hockey stick”: a straight line down to some threshold, and zero effect below that threshold. Thresholds occur when an organism has a natural repair mechanism that can prevent or counteract damage up to a certain limit.

There is even the possibility of a “fertilizer effect” at low doses, as indicated by the dashed line in Figure 4.4. This effect can be observed in the DRFs for the impact of NO_x and SO_2 on crops: a low dose of these pollutants can increase the crop yield. In other words, the damage is negative. Generally a fertilizer effect can occur with pollutants that provide trace elements needed by an organism.

In practice, most DRFs used by ExternE, particularly all the ones for health, are assumed to be linear (without threshold). However, for the calculation of incremental damage costs, there is no difference between the linear and a hockey stick function with the same slope, if the background concentration is everywhere above this threshold; only the slope matters. For the particles, NO_x, SO_2, O_3, and CO the background in most industrialized countries is above the level where effects are known to occur. Thus the precise form of the DRF at extremely low doses is irrelevant for these pollutants; if there is a no-effect threshold, it is below the background concentrations of interest.

4.2.4 Health Impacts

In terms of costs, health impacts contribute the largest part of the damage estimates of ExternE. A consensus has been emerging among public health experts that air pollution, even at current ambient levels, aggravates morbidity (especially respiratory and cardiovascular diseases) and leads to premature mortality (e.g., Wilson and Spengler 1996; ERPURS 1997; the reports of the World Health Organization, available at http://www.who.int, see Table 4.1). There is less certainty about specific causes, but most recent studies have identified fine particles as a prime culprit; ozone has also been directly implicated. The most important cost comes from chronic mortality due to particles, calculated on the basis of Pope et al. (2002). This term, chosen by analogy with acute and chronic morbidity impacts, indicates that the total or long-term effects of pollution on mortality have been included, by contrast to acute mortality impacts that are observed within a few days of exposure to pollution. Another important contribution

TABLE 4.1 Air Pollutants and Their Effects on Health

Primary Pollutants	Secondary Pollutants	Impacts
Particles (PM_{10}, $PM_{2.5}$, black smoke)		Mortality Cardiopulmonary morbidity (cerebrovascular hospital admissions, congestive heart failure, chronic bronchitis, chronic cough in children, lower respiratory symptoms, cough in asthmatics)
SO_2		Mortality Cardiopulmonary morbidity (hospitalization, consultation of doctor, asthma, sick leave, restricted activity)
SO_2	Sulfates	Like particles?
NO_x		Morbidity?
NO_x	Nitrates	Like particles?
NO_x+VOC	Ozone	Mortality Morbidity (respiratory hospital admissions, restricted activity days, asthma attacks, symptom days)
CO		Mortality (congestive heart failure) Morbidity (cardiovascular)
PAH (diesel soot, benzene, 1,3-butadiene, dioxins)		Cancers
As, Cd, Cr^{VI}, Ni		Cancers Other morbidity
Hg, Pb		Morbidity (neurotoxic)

comes from chronic bronchitis due to particles (Abbey et al. 1995). In addition, there may be significant direct health impacts of SO_2, but for direct impacts of NO_x, the evidence is less convincing.

In ExternE, the working hypothesis has been to use the DRFs for particles and for O_3 as basis. Effects of NO_x and SO_2 are assumed to arise indirectly from the particulate nature of nitrate and sulfate aerosols, and the effects calculated by applying the particle DRFs to these aerosol concentrations. But the uncertainties are large because there is insufficient evidence for the health impacts of the individual components or characteristics (acidity, solubility, etc.) of particulate air pollution. In particular there is a lack of epidemiological studies of nitrate aerosols because until recently air pollution monitoring stations did not monitor this pollutant. In view of the lack of evidence for thresholds at current ambient concentrations, all DRFs for health impacts have been assumed linear at the population level. By contrast to the homogeneous populations of cloned animals studied by toxicologists, the absence of a no-effect threshold is plausible for real populations because they always contain individuals with widely differing sensitivities (for example, at any moment about 1% is within the last nine months of life and thus extremely frail).

4.2.5 Monetary Valuation

The goal of the monetary valuation of damages is to account for all costs, market and nonmarket. The valuation of an asthma attack should include not only the cost of the medical treatment but also the willingness to pay (WTP) to avoid the residual suffering. If the WTP for a nonmarket good has been determined correctly, it is like a price, consistent with prices paid for market goods. Economists have developed several tools for determining nonmarket costs; of these tools contingent valuation (CV) has enjoyed increasing popularity in recent years (Mitchell and Carson 1989). The basic idea of a CV is to ask people how much they would be willing to pay for a certain good if they could buy it. The results of well-conducted CV studies are considered sufficiently reliable.

It turns out that nonmarket goods dominate damage costs of air pollution, especially the valuation of mortality. The single most important parameter is the so-called "value of statistical life" (VSL). This term often evokes hostile reactions from people who think that economists try to measure the value of life. The value of life is limitless. To save an individual in danger no means are spared. In reality VSL is the "willingness to pay for avoiding a small risk of an anonymous premature death". In ExternE (2000), a European-wide value of 3.4 M€ was chosen for VSL, consistent with similar studies in the USA; this value was chosen as average of the VSL studies that had been carried out in Europe.

A crucial question for air pollution mortality is whether one should simply multiply the number of premature deaths by VSL, or whether one should take into account the years of life lost (YOLL) per death. The difference is very important because premature deaths from air pollution tend to involve far fewer YOLL per death than accidents (on which VSL is based). In fact, for air pollution the appropriate measure is the value of a YOLL due to air pollution, called VOLY (value of a life year), whereas the true number of premature deaths due to pollution cannot even be determined, as shown by Rabl (2003).

There is considerable uncertainty because until recently there have been no studies to determine VOLY. ExternE (1998, 2000) calculated VOLY on theoretical grounds by considering VSL as the net present value of a series of discounted annual values. The ratio of VSL and the value of a YOLL thus obtained depends on the discount rate; it is typically in the range of 20–30. More recently ExternE carried out a CV study for VOLY and is now using a value of 50,000 € for a year of life lost due to air pollution. As for cancers, ExternE assumes 0.45 M€ for nonfatal cancers, and 1.5–2.5 M€ for fatal cancers (depending on the YOLL for each cancer type).

4.2.6 Global Warming

The valuation of global warming damages is extremely complex (see, for example, Tol et al. (2001)). Not only is the task difficult because of the large number of different impacts in all countries of the world that should be taken into account, but also as these impacts will occur in future decades and centuries one needs to estimate how these costs will evolve into the distant future. On top of the resulting uncertainties there are controversial ethical issues related to the valuation of mortality in developing countries (where most of the impacts will occur) and the choice of the discount rate for intergenerational costs.

Several major studies have been published with estimates of the cost per tonne of CO_{2eq} (the subscript eq indicates that the result can also be used for other GHGs if their masses are multiplied by their global warming potential (GWP)). Most of the results are in the range of 1–50 €/t_{CO2eq}, the range being so wide because of the large uncertainties. The ExternE team carried out two valuation efforts: the first, in 1998, yielded a range of values with a geometric mean of 29 €/t_{CO2eq}, the second, in 2000, obtained a much lower value of 2.4 €/t_{CO2eq} because of more optimistic assumptions and a better accounting for benefits such as increased agricultural production in cold countries. The current phase of ExternE uses the value of 19 €/t_{CO2eq} because that is the abatement cost in the EU implied by the commitment to the Kyoto Protocol. It represents an implicit valuation by decision makers of the EU. It is also in effect the cost imposed on the EU by incremental emissions of CO_2 in the EU. The choice of this value appears reasonable in view of the estimates published in the literature.

4.2.7 Other Impacts

Air pollution damage to materials and agricultural crops has been found to make a relatively small contribution to the damage costs, only a few percent of the total. Estimations of ecosystem impacts, other than agricultural losses, have remained extremely uncertain, if they have been attempted at all. Some estimates have been made of the costs of forest decline due to acid rain, but more recently doubts have been raised about their validity. In general there is a lack of information on ecosystem impacts and their economic valuation.

However, air pollution at typical ambient concentrations in the EU does not seem to have significant direct (i.e., not acid rain) impacts on ecosystems. At first glance this claim may appear surprising because human health impacts are significant and one might indeed expect similar impacts on animals as on humans. The explanation lies in what is valued: society values ecosystem impacts at the level of a population, human impacts at the level of the individual. Concentrations of air pollutants are generally so small that the incremental mortality is at most a small percentage of the natural rate. Furthermore, most of the deaths from air pollution occur among individuals well beyond reproductive age. If a small percentage of animals die prematurely after having produced and raised offspring, the effect on the ecosystem is negligible. But if any human dies prematurely, society cares a great deal.

The situation is different for certain aquatic impacts. Acidification of rivers and lakes has been shown to be detrimental to aquatic life. A river can collect much of the air pollution from a large region, leading to relatively high concentrations in the water, quite apart from direct emission of pollutants to water.

4.2.8 UWM: A Simple Model for Damage Cost Estimation

A simple and convenient tool for the development of typical values is the "uniform world model" (UWM), first presented by Curtiss and Rabl (1996) and further developed, with detailed validation studies by Spadaro (1999), and Spadaro and Rabl (2002). More recently Spadaro and Rabl (2004) extended it to toxic metals and their pathways through the food chain. The UWM is a product of a few factors; it is simple and transparent, showing at a glance the role of the most important parameters of the IPA. It is exact for tall stacks in the limit where the distribution of either the sources or the receptors is uniform and the key atmospheric parameters do not vary with location. In practice the agreement with detailed models is usually within a factor of two for stack heights above 50 m. For policy applications one needs typical values, and the UWM is more relevant than a detailed analysis for a specific site.

The UWM for the damage cost D_{uni} in €/kg of a particular impact due to the inhalation of a primary pollutant is shown in Equation 4.1:

$$D_{\text{uni}} = \frac{p s_{\text{CR}} \rho}{v_{\text{dep}}}, \tag{4.1}$$

where p is the cost per case ("price") [€/case], s_{CR} is the CRF slope [(cases/yr)/(pers·(μg/m^3))], ρ is the average population density [pers/km^2] within 1000 km of source, and v_{dep} is the deposition velocity of pollutant (dry+wet) [m/s].

For secondary pollutants the equation has the same form, but with an effective deposition velocity that includes the transformation rate of the primary into the secondary pollutant. With this model it is easy to transfer to the results from one region to another (assuming that CRF and deposition velocity are the same): simply rescale the result in proportion to the receptor density and the cost per case.

4.3 Results for Cost per Kilogram of Pollutant

The impacts quantified by ExternE so far are global warming, health, damage to buildings and materials, and loss of agricultural production. Apart from global warming due to CO_2, CH_4, and N_2O, more than 95% of the costs is due to health impacts, especially mortality. Morbidity (especially chronic bronchitis but also asthma, hospital admissions, etc.) account for almost 30% of the damage cost of PM, NO_x, and SO_2. The impacts evaluated and the key assumptions are listed in Table 4.2. The resulting damage costs in €/kg of pollutant are shown in Figure 4.5 for typical sources with stack heights above 50 m in Central Europe. As for variation with site and stack height, the following rules can be recommended for modifying factors:

TABLE 4.2 Impacts Evaluated and Key Assumptions

Atmospheric Dispersion Models	
Local range:	Gaussian plume model ISC (point sources) or ROADPOL (emissions from transport)
Regional range (Europe):	Harwell trajectory model as implemented in EcoSense software of ExternE Ozone impacts based on EMEP model
Physical Impacts	
Impacts on Health	
Form of dose–response functions (DRF)	Linearity without threshold, with slope s_{CR}
Chronic mortality	S_{CR}=4.1E-4 YOLL (years of life lost) per person per year per μg/m^3 derived from increase in age-specific mortality due to $PM_{2.5}$ [Pope et al 2002], by integrating over age distribution
Acute mortality	For SO_2 and ozone, assume 0.75 YOLL per premature death
Nitrate and sulfate aerosols	Dose–response functions for nitrates: 50%×PM_{10} (PM_{10} slope=60%×$PM_{2.5}$ functions) Dose–response functions for sulfates same as for PM_{10}
Micropollutants	Cancers due to As, Cd, Cr, Ni, and dioxins DRF for dioxins according to EPA [2000] IQ decrement due to Hg and Pb
Impacts on plants	Loss of crops due to SO_2 and ozone
Impacts on buildings and materials	Corrosion and erosion due to SO_2 and soiling due to particles
Impacts not quantified but potentially significant	Reduced visibility due to air pollution Eutrophication and acidification Disposal of residues from fossil fuels or incineration
Monetary Valuation	
Valuation of premature death	Proportional to reduction of life expectancy, with value of a of life year (VOLY) =50,000 €
Valuation of cancers	2 M€ per cancer
Valuation of neurotoxicity	10,000 € per IQ point lost
Global warming damage cost	0.019 €/kg$_{CO2eq}$

Source: From ExternE. 2004. Project NewExt "New elements for the assessment of external costs from energy technologies." European commission DG Research, Contract No. ENG1-CT2000-00129. Coordinated by R. Friedrich, IER, University of Stuttgart. Final report. http://www.externe.info. With permission.

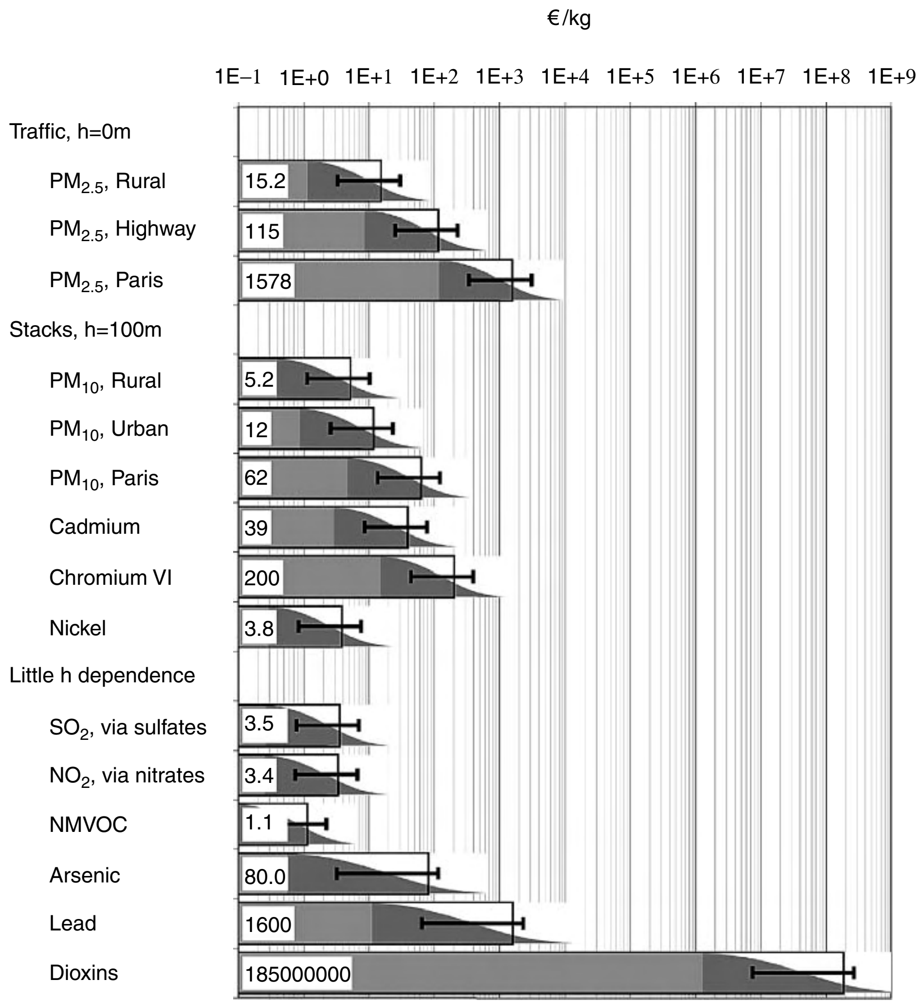


FIGURE 4.5 Results for damage costs of the most important air pollutants (typical values, for LCA applications in the EU15). The error bars indicate the 68% confidence interval; on the logarithmic scale, they are symmetric around the median (equal to the geometric mean of lognormal distribution). The broad hollow bars and the numbers are the mean that is larger than the median. The gray S-shaped curve indicates the probability that the true cost is above a specified value, h= stack height. (From ExternE, 2004. New results of ExternE, after the NewExt project, http://www.externe.info.)

- No variation for globally dispersing pollutants such as CO_2
- Weak variation for As, Pb, and dioxins because noninhalation pathways dominate: approximately 0.7–1.5
- Weak variation for secondary pollutants: approximately 0.5–2.0
- Strong variation for primary pollutants: approximately 0.5–5 for site, and approximately 0.6–3 for stack conditions (up to 15 for ground-level emissions in a large city)

Of course, such rules can only yield rough estimates; site-specific calculations should be carried out when more precise results are needed.

An analysis of the uncertainties is crucial for the credibility of the results. Uncertainties can be grouped into different categories, even though there may be some overlap:

- Data uncertainty (e.g., slope of a DRF, cost of a day of restricted activity, and deposition velocity of a pollutant)
- Model uncertainty (e.g., assumptions about causal links between a pollutant and a health impact), assumptions about form of a DRF (e.g., with or without threshold), and choice of models for atmospheric dispersion and chemistry
- Uncertainty about policy and ethical choices (e.g., discount rate for intergenerational costs, and "VSL")
- Uncertainty about the future (e.g., the potential for reducing crop losses by the development of more resistant species)
- Idiosyncrasies of the analyst (e.g., interpretation of ambiguous or incomplete information)

The first two categories (data and model uncertainties) are scientific in nature. They are amenable to analysis by statistical methods, combining the component uncertainties over the steps of the impact pathway, to obtain formal confidence intervals around a central estimate. For this an approach based on lognormal distributions and multiplicative confidence intervals was followed. The error bars in Figure 4.5 show the results of this analysis; they are one-geometric standard deviation intervals around the median estimate. The largest sources of uncertainty lie in the DRFs for health impacts and in the value of a life year. Details can be found in Rabl and Spadaro (1999).

Quantifying the sources of uncertainty in this field is problematic because of a general lack of information. Usually one has to fall back on subjective judgment, preferably by the experts of the respective disciplines. For ExternE a survey of experts has been done in an informal manner. Also all the relevant information that could be found in the literature was used.

4.4 Results for Energy Production

Once the cost per kg has been determined, multiplication by the emitted quantities of the pollutants yields the cost per activity, for instance per kWh of electricity produced by a power plant. A complete accounting of the damage costs should involve a LCA, i.e., a complete inventory of emissions over the entire chain of processes involved in the activity. The total damage cost per kWh of electricity should include impacts upstream and downstream from the power plant, such as air pollution from the ships, trucks, or trains that transport the fuel to the power plant. That has been done for the fuel chain results of ExternE. A few results for France are shown in Figure 4.6.

For the fossil fuel chains, the lion's share of the external costs comes from air pollutants emitted by the power plant, the main impact categories being global warming and public health. Air pollutants from upstream and downstream activities make a relatively small contribution (roughly 10% of the GHGs). Apart from CO_2, the damage cost is mostly due to health impacts, especially mortality.

The emission of toxic metals is highly uncertain and variable from one source of fuel to another; the numbers shown are not necessarily typical. However, they are in any case so small that their contribution to the total damage cost is negligible.

Recently updated and more complete results for power production in Europe have been calculated during the ExternE-Pol phase of the ExternE series (Rabl et al. 2004); they are shown in Figure 4.7 and Figure 4.8. The numbers are not completely comparable with those in Figure 4.6 because the methodology has been evolving. In particular, the numbers in the following are based on detailed LCA inventories provided by the ecoinvent database. Also the methodology for nuclear damage costs in Figure 4.7 is different from the one in Figure 4.6.

Results obtained for new and current technologies on the basis of ecoinvent are discussed in the remainder of this section. Figure 4.7 presents the results for current and advanced electricity systems, with the external costs per kWh in part a) and the contributions of the individual pollutants in part b). Likewise Figure 4.8 summarizes the results for the different heating systems.

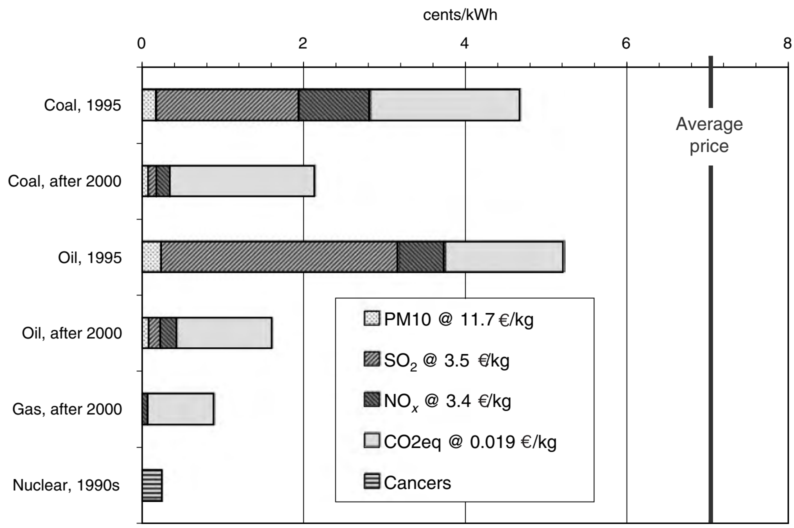


FIGURE 4.6 Typical damage costs for electric power plants in France, for plants operating during the mid-1990s and for plants that respect the new EU regulations issued in 2000. Costs for nuclear are upper bound (0% "effective discount rate"); 90% of cancers from nuclear occur only after the first 100 yr. Retail price of electricity is about 7 cents/kWh; production cost of base load electricity about 3 cents/kWh.

Current fossil systems exhibit the highest external costs, in the range of 1.6–5.8 c€/kWh (Figure 4.7). Introduction of advanced technology (CC and PFBC) substantially reduces the external costs of fossil systems, but they still remain in the range 1–2 c€/kWh. This also applies to cogeneration, for which gas technology generates external costs one-third lower than diesel technology. Greenhouse gas (GHG) contribution to total costs prevails over other species for advanced fossil technologies, making about 80% of total external costs for CC and PFBC. Current averages of coal and oil plants show still high contributions from SO_2, depending on the extension of installation of scrubbers in UCTE.

Nuclear external costs are below 0.19 c€/kWh of which 70% is radioactivity dependent. However, with discounting this contribution would strongly decrease, because most of the calculated damages from radiation are either related to very long-term emissions (e.g., radon from uranium mill tailings) or to very long-lived isotopes giving very small doses. On the other hand, the present estimation of external costs from ionizing radiation is based on a preliminary calculation using the disability-adjusted life years (DALY) concept, a rough attribution of cost/DALY, and an incomplete but important subset of isotope releases from the ecoinvent database. It is recommended to rework the estimation of damage factors from radioactive emissions in future projects of the ExternE series. The nuclear power plant itself contributes 5% or less to external costs from the nuclear chain.

Wind onshore with nearly 0.09 c€/kWh performs slightly better than wind offshore with 0.12 c€/kWh. Monocrystalline silicon photovoltaic (PV) panels of European fabrication, installed in Southern Europe cause nearly 0.28 c€/kWh, which would mean 0.41 c€/kWh for the average yield of 800 kWh/kW_{peak} yr in Central Europe. Assuming improvements in manufacturing technology of crystalline silicon, improved cell efficiency, and an expanded PV market, 0.21 c€/kWh has been estimated for future (2010) systems. External costs associated with imported panels may differ due to different manufacturing technology and electricity supply. Due to the relatively high material intensity of PV and wind, the contribution from heavy metals is about 15% and nearly 25%, respectively. Hydropower exhibits the lowest external costs of all systems, below 0.05 c€/kWh, but costs may increase on sites where higher direct emission of GHG from the surface of reservoir occur.

In general, gas boilers have lower external costs than boilers burning light oil: approximately 0.6 c€/kWh_{th} vs. 0.94 c€/kWh_{th} (Figure 4.8). The upstream chain of gas and light oil contributes

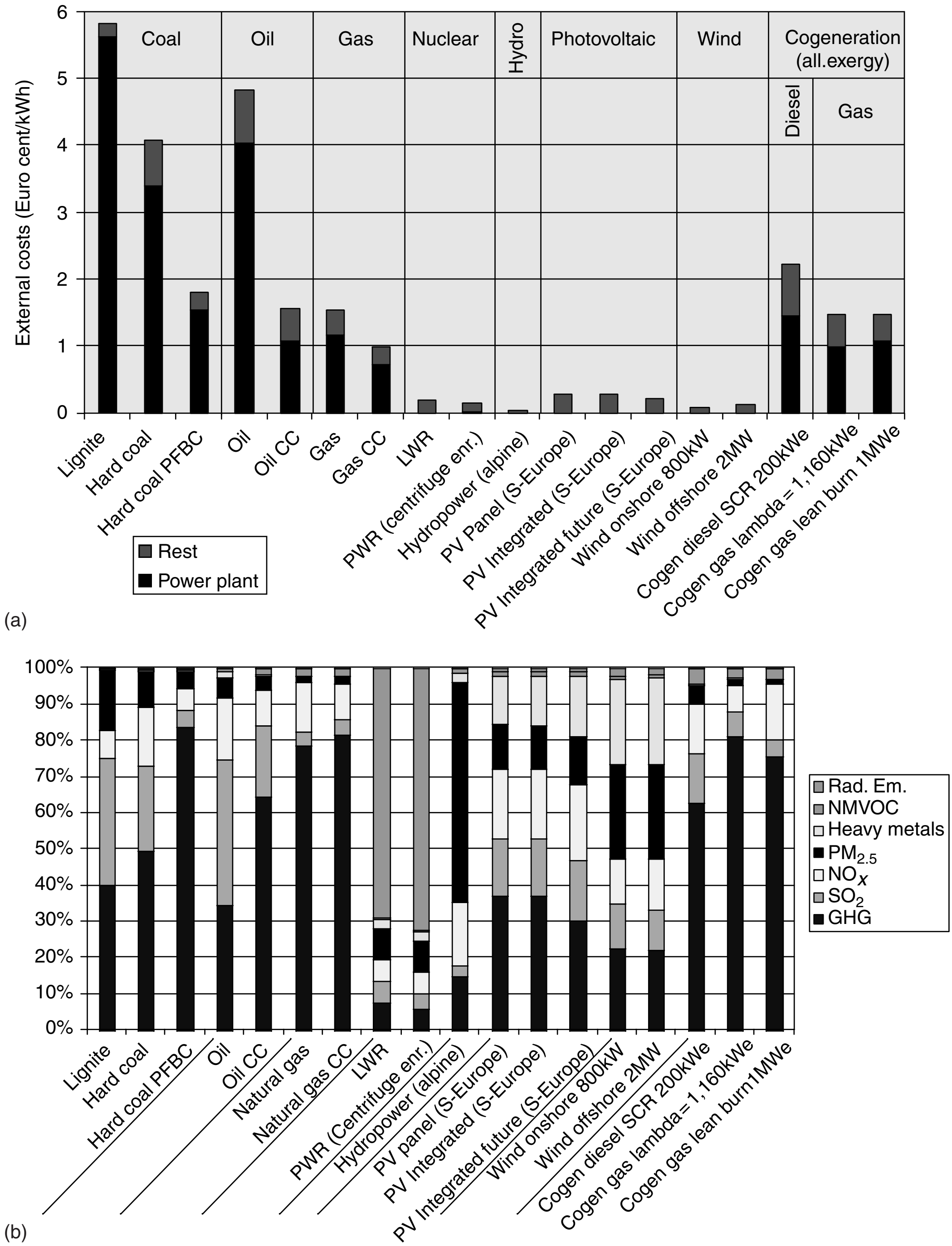


FIGURE 4.7 External costs of current and advanced electricity systems, associated with emissions from the operation of power plant and with the rest of energy chain. (a) The costs in €cent/kWh. (b) The contribution of the individual pollutants.

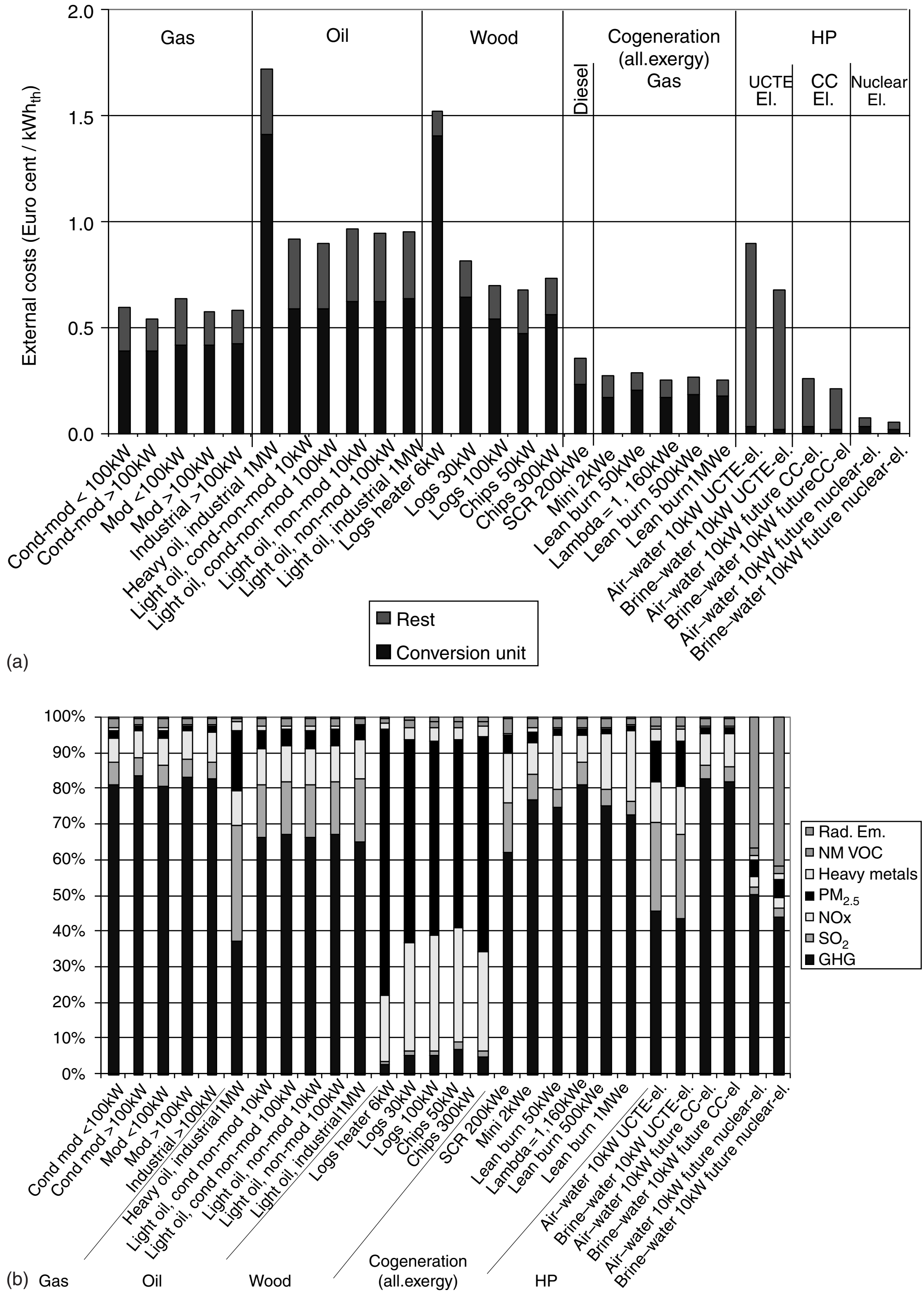


FIGURE 4.8 External costs of heating systems, associated with emissions from the operation of boiler/cogeneration unit and with the rest of energy chain. (a) The costs in €cent/kWh. (b) The contribution of the individual pollutants.

roughly one-third to the total external costs. GHG contribute two-thirds of the total external costs for oil, over 80% for gas boilers. Burning heavy oil gives the highest damages with over 1.7 c€/kWh$_{th}$, where SO_2 makes about 33% and GHG 38% of the damages. A range of about 0.7–0.8 c€/kWh$_{th}$ has been calculated for wood boilers, where the upstream chain contributes 20%–30% to total damages. Particles and nitrogen oxides emissions contribute most, i.e., nearly 60% and about 30%, respectively, to total damages. The modern fireplace gives more than 1.5 c€/kWh$_{th}$, mostly due to the high particle release. GHG contribute 7% or less to total external costs for modern wood systems, because the CO_2 from wood combustion is compensated by tree sequestration.

Cogeneration plants perform well when allocation is based on exergy: 0.36 c€/kWh$_{th}$ for diesel and an average of 0.27 c€/kWh$_{th}$ calculated for the gas units. The magnitude of external costs of heat pumps (HP) is controlled basically by two factors: the seasonal performance factor (SPF) and the energy supply source. For current systems and average UCTE electricity mix the external costs are nearly 0.7 c€/kWh$_{th}$ and 0.9 c€/kWh$_{th}$ for the air–water HP and brine–water HP, respectively. Due to the fact that about 26% of the UCTE electricity mix is from coal systems, damages from SO_2 contribute nearly one quarter to the total external costs. For future HP technologies and electricity delivered by gas CC or nuclear, these costs go down to 0.26 c€/kWh$_{th}$ and 0.21 c€/kWh$_{th}$, or nearly 0.08 c€/kWh$_{th}$ and 0.06 c€/kWh$_{th}$, respectively, for the two heat pump systems and the two electricity supply cases (Figure 4.8).

4.5 Comparison Landfill ↔ Incineration

Because waste can be considered a source of renewable energy, the paper also examines the role of energy recovery from landfill and incineration, the main approaches to waste treatment. The results reported are based on Rabl, Spadaro and McGavran (1998), Rabl and Spadaro (2002), and Zoughaib and Rabl (2004).

4.5.1 Assumptions

Because the comparison necessitates an LCA, the work begins by choosing the boundaries of the analysis. The most appropriate choice is to start at the point where the waste has been collected and sorted. From here the waste must be transported to the landfill or incinerator; for the purpose of illustration, the emissions due to possible differences in transport distance by showing a hypothetical distance of 100 km are included. In addition to the emission of pollutants from the landfill or incinerator, the emissions avoided by recovery of energy and materials are also taken into account, based on the LCA data of ADEME (2000). The assumptions of the analysis are summarized in Table 4.3.

The principal emissions from landfill are CH_4 and CO_2. Figure 4.9 shows the total GHG emissions of a municipal solid waste landfill versus time. CH_4 is expressed as equivalent CO_2, using a GWP of 20. Note that a modern landfill is divided into a large number of individual compartments; they are filled one after another and sealed when full. The time in Figure 4.9 is measured from the date that a compartment is sealed.

TABLE 4.3 Assumptions of the Analysis of Incineration and Landfill of Municipal Waste

Stages taken into account	Transport of waste; Emissions from landfill or incinerator; Avoided emissions due to energy recovery; Avoided emissions due to materials recovery
Emissions from incinerator	Equal to limit values of Directive EC (2000), in reality the average emissions are usually lower
Impact pathway analysis	Assumptions and results of ExternE (2004)
Impacts taken into account	Human health; crops; materials and buildings; global warming
Impacts not taken into account	Effects of air pollutants on ecosystems; Soil and water pollution due to leachates (probably negligible); Impacts from residues of incineration; Amenity impacts (visual intrusion, noise, odor)

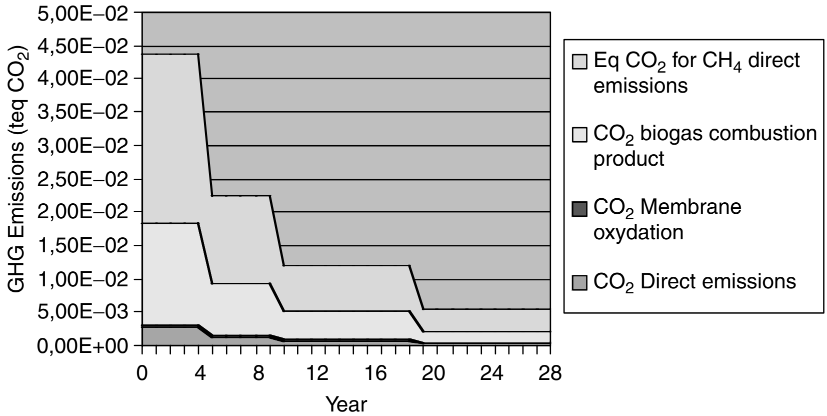


FIGURE 4.9 Total greenhouse gas emissions from a municipal solid waste landfill versus time, t_{CO2eq}/t_{waste} if 70% of the CH_4 is captured. (This figure was prepared from a table with averages over four or more years; in reality, the evolution is smooth rather than in steps).

There may also be emissions to soil and to water. Emissions to soil can occur from slag, leaking liners under the landfill, and the storage site of incinerator fly ash. Emissions to water arise from certain types of flue gas treatment and from the extraction of leachates under a landfill. Emissions to soil are difficult to estimate because they depend on integrity of the liners in the future. If the landfill is operated according to regulations and if there are no mishaps, there are no such impacts during the foreseeable future because the operator has the obligation to maintain and safeguard the facility for 30 years after closure. In any case their impacts would remain limited to the immediate vicinity of the landfill, with the possible exception of sites with sufficient ground water movement. In the present version impacts of emissions of leachates to soil or water have not been considered.

The external costs and the comparison between landfill and incineration turn out to be extremely sensitive to assumptions about energy recovery. For that reason, a fairly large number of options (indicated in the figures by labels such as "$E=c+o$") is condered:

For incineration:

- Recovery of heat and electricity, for typical installations in France, according to ADEME (2000) ($E=\ldots$, $H=\ldots$)
- Recovery of electricity ($E=\ldots$)
- Recovery of heat ($H=\ldots$)

For landfill:

- No energy recovery
- Recovery of electricity, by motor (reciprocating engine) ($E=\ldots$)
- Recovery of electricity, by turbine ($E=\ldots$)
- Recovery of heat ($H=\ldots$)

For recovery of electricity, a year-round demand is assumed such that all the electricity is used. Likewise, for recovery of heat, a year-round demand is assumed (industrial process heat loads or certain district heating systems with year-round demand, e.g., Paris and Vienna), such that none of the heat is wasted. Year-round demand is essential for good recovery rates because the supply of waste tends to be fairly constant.

For each of these options, several suboptions (indicated by the labels in the captions) are considered:

- The recovered electricity displaces coal- and oil-fired power plants, 50% each ($E=c+o$)

- The recovered electricity displaces nuclear power plants ($E=n$)
- The recovered heat displaces gas- and oil-fired heating systems, 50% each ($H=g+o$)
- The recovered heat displaces only oil-fired heating systems ($H=o$)

For example, ($E=c+o$, $H=c+o$) designates a system where heat and electricity are produced, each displacing a fuel mixture of coal and oil. Note that for the purpose of this analysis the benefit of recovered electricity is essentially zero if it displaces nuclear because the damage costs of nuclear are very small compared to those of oil or coal; thus this option is essentially equivalent to no electricity production at all as far as external costs are concerned.

4.5.2 Results

A summary of the total damage cost for all the options is shown in Figure 4.10. More detailed results for some of the options can be found in Figure 4.11, showing the contribution of each stage and of the major pollutants (dioxins and toxic metals are shown as "other"). The benefits of materials recovery make a small or negligible contribution to the total damage cost. The damage costs of waste transport, illustrated with an arbitrary choice of 100 km round-trip by a 16-ton truck, is also negligible. The only significant contributions come from direct emissions (of the landfill or incinerator) and energy recovery.

For landfill the cost is dominated by GHG emissions because only about 70% the CH_4 can be captured. Energy recovery from a landfill is not very significant. By contrast, energy recovery is crucial for the damage cost of incineration. Under favorable conditions (all heat produced by incinerator displaces coal and oil) the total external cost can even be negative, i.e., a net benefit. In contrast to most other countries, in France

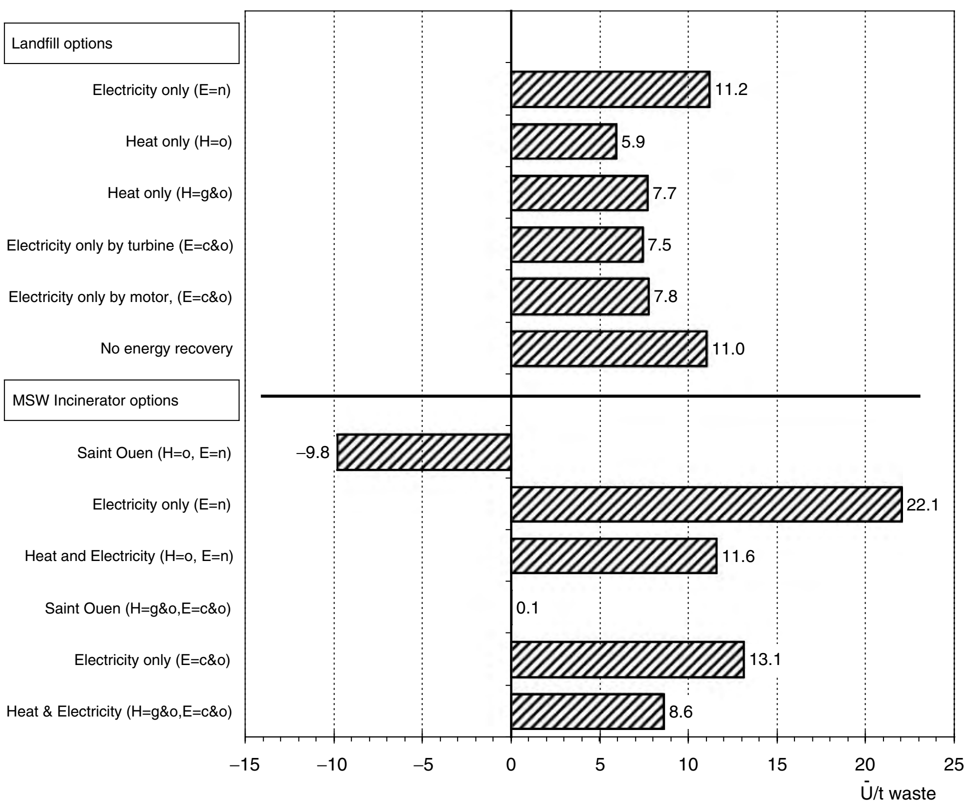


FIGURE 4.10 Results of total damage cost for all options. If electricity displaces nuclear, damage costs are essentially the same as for the case without energy recovery.

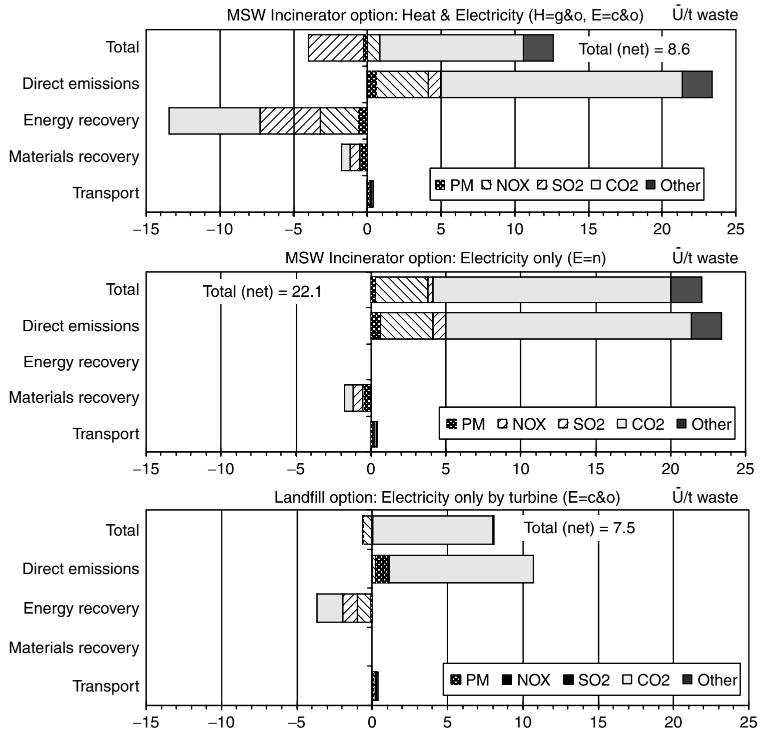


FIGURE 4.11 Some detailed results, by stage and pollutant. "Other" refers to dioxins and toxic metals.

recovery of electricity does not bring significant benefits because it is base load power, and all the base load power is produced by nuclear; the options where it displaces coal or oil are not realistic in France (except near the border where the power can be exported) because these fuels are used only during the heating season.

The uncertainties are very large (see Figure 4.5) and they have different effects on different policy choices. Some comparisons, in particular those between landfill and incineration, are especially sensitive to this uncertainty because GHGs play such a large role for landfills.

4.6 Conclusions

The most relevant damages caused by air pollution can be quantified and monetized using the methodology of ExternE. However, damage cost estimates still show large uncertainties, and the reader may wonder whether it is meaningful to use them as basis for decisions. The first reply is that even a threefold uncertainty is better than infinite uncertainty. Second, in many cases the benefits are either so much larger or so much smaller than the costs that the implication for a decision is clear even in the face of uncertainty. Third, if in the other cases the decisions are made without a significant bias in favor of either costs or benefits, some of the resulting decisions will err on the side of costs, others on the side of benefits. Rabl, Spadaro, and van der Zwaan (2005) have examined the consequences of such unbiased errors and found a very reassuring result. The extra social cost incurred because of uncertain damage costs is remarkably small, less than 10%–20% in most cases even if the damage costs are in error by a factor three. But without any knowledge of the damage costs, the extra social cost (compared to the minimal social cost that one would incur with perfect knowledge) could be very large.

A key argument for the necessity of quantifying damage costs has emerged from recent epidemiology. Whereas in the past there was a general belief that it would be sufficient to reduce the concentration of air pollutants below their no-effect thresholds, epidemiological studies have not been able to find evidence for such no-effect thresholds at the population level, and linear DRFs appear increasingly plausible for low concentrations of most air pollutants. If there are no safe levels of air pollution, then policy makers have no natural criterion for deciding how far to reduce the emission of pollutants. Because the cost per kg of avoided pollutant increases sharply as emissions are reduced, there is a serious risk of spending too much on the fight against air pollution. This could in fact lead to a deterioration of public health if the money thus spent is not available for more cost-effective measures. A comparison of costs and benefits is needed for rational policy making.

Gradually the results of ExternE are diffusing into the world of decision makers. For example, ExternE is recognized as the reference for comparative risk assessment by agencies such as the International Atomic Energy Agency. In the EU, ExternE is increasingly used as input to environmental decisions, e.g., via cost–benefit analyses that justify tighter regulations for the emission of pollutants from incinerators or from power plants.

Glossary and Nomenclature

As Arsenic
CBA Cost–benefit analysis
CO Carbon monoxide
CO_{2eq} Quantity of a greenhouse gas expressed as equivalent quantity of CO_2, using the GWP of the gas
Cr^{VI} Chromium in oxidation state 6
CRF Concentration–response function
CV Contingent valuation
Discount rate Rate that allows comparison of monetary values incurred at different times, defined such that an amount P_n in year n has the same utility as an amount $P_0 = P_n(1+r)^{-n}$ in year zero
DRF Dose–response function
EC European Commission
EPA Environmental Protection Agency of USA
External costs Costs that arise when the social or economic activities of one group of people have an impact on another for which the first group does not fully account, e.g., when a polluter does not compensate others for the damage imposed on them
GWP Global warming potential
Hg Mercury
IPA Impact pathway analysis
ISC Industrial Source Complex Gaussian plume dispersion model
LCA Life cycle assessment
N Nitrogen
Ni Nickel
NO_x Unspecified mixture of NO and NO_2
O_3 Ozone
PAH Polycyclic aromatic hydrocarbons
Pb Lead
PM_d Particulate matter with aerodynamic diameter smaller than d μm
S Sulfur
S_{CR} Slope of concentration–response function [cases/(person yr μg/m^3)]
UCTE Union for the Coordination of Transmission of Electricity
UWM Uniform world model (for simplified estimation of damage costs)
v_{dep} Deposition velocity [m/s]
VOC Volatile organic compounds
VOLY Value of a life year
VSL Value of statistical life

WTP Willingness to pay
YOLL Years of life lost (reduction of life expectancy)

Acknowledgments

This work has been supported in part by the ExternE project series of the EC DG Research.

References

Abbey, D. E., Lebowitz, M. D., Mills, P. K., Petersen, F. F., Lawrence Beeson, W., and Burchette, R. J. 1995. Long-term ambient concentrations of particulates and oxidants and development of chronic disease in a cohort of nonsmoking California residents. *Inhalation Toxicology*, 7, 19–34.

ADEME. 2000. Analyse environnementale de systèmes de gestion de déchets ménagers, Phase 1: Analyse des parameters determinants pour les impacts environnementaux des différents modules. Etude réalisée par BIO Intelligence et Ecobilan pour ADEME et Eco-Emballages.

Brode, R. W. and Wang, J. F. 1992. *User's Guide for the Industrial Source Complex (ISC2) Dispersion Models Volumes I–III, EPA-450/4-92-008a, EPA-450/4-92-008b, EPA-450/4-92-008c*. US Environmental Protection Agency, Research Triangle Park, NC.

Curtiss, P. S. and Rabl, A. 1996. Impacts of air pollution: General relationships and site dependence. *Atmospheric Environment*, 30, 3331–3347.

Derwent, R. G. and Nodop, K. 1986. Long-range transport and deposition of acidic nitrogen species in northwest Europe. *Nature*, 324, 356–358.

European Commission (EC). 2000. Directive 2000/76/EC of the European Parliament and of the Council of 4 December 2000 on the incineration of waste.

ERPURS. 1997. Analyse des liens à court terme entre pollution atmosphérique et santé (Analysis of short term correlations between air pollution and health). Evaluation des Risques de la Pollution Urbaine sur la Santé. Observatoire Régional de Santé d'Ile-de-France, Paris.

ExternE. 1995. *ExternE: Externalities of Energy*. Vol. 1: Summary EUR 16520. Vol. 2: Methodology EUR 16521; Vol. 3: Coal and Lignite. EUR 16522; Vol. 4: Oil and Gas. EUR 16523; Vol. 5: Nuclear. EUR 16524; Vol. 6: Wind and Hydro Fuel Cycles. EUR 16525. Luxembourg: European Commission, Directorate-General XII, Science Research and Development. Office for Official Publications of the European Communities.

ExternE. 1998. *ExternE: Externalities of Energy*. Vol. 7: Methodology 1998 Update. EUR 19083; Vol. 8: Global Warming. EUR 18836; Vol. 9: Fuel Cycles for Emerging and End-Use Technologies, Transport and Waste. EUR 18887; Vol. 10: National Implementation. EUR 18528. Luxembourg: European Commission, Directorate-General XII, Science Research and Development. Office for Official Publications of the European Communities.

ExternE. 2000. External costs of energy conversion—Improvement of the Externe methodology and assessment of enery-related transport externalities. Final Report for contract J0S3-CT97-0015, Published as Environmental External Costs of Transport, R. Friedrich and P. Bickel, eds., Springer, Heidelberg.

ExternE. 2004. Project NewExt New elements for the assessment of external costs from energy technologies. European Commission of Research, Contract No. ENG1-CT2000-00129. Coordinated by R. Friedrich, IER, University of Stuttgart. Final report. http://www.externe.info. (accessed on September 2006).

International Atomic Energy Agency (IAEA). 1994. *Handbook of Parameter Values for the Prediction of Radionuclide Transfer in Temperate Environments,* Technical Report Series No. 364, Produced in collaboration with the International Union of Union of Radioecologists, International Atomic Energy Agency, Vienna, Austria.

International Atomic Energy Agency (IAEA). 2001. *Generic Models for Use in Assessing the Impact of Discharges of Radioactive Substances to the Environment,* Safety Reports Series No. 19, International Atomic Energy Agency, Vienna, Austria.

Krewitt, W., Trukenmueller, A., Mayerhofer, P., and Friedrich, R. 1995. ECOSENSE—An integrated tool for environmental impact analysis, In *Space and Time in Environmental Information Systems*, H. Kremers and W. Pillmann, eds., Umwelt-Informatik aktuell, Band 7, Metropolis-Verlag, Marburg.

Lee, R. ed. 1994. *External Costs and Benefits of Fuel Cycles*, Oak Ridge National Laboratory and Resources for the Future, Oak Ridge, TN.

Mitchell, R. C. and Carson, R. T. 1989. *Using Surveys to Value Public Goods: The Contingent Valuation Method,* Resources for the Future, Washington, DC.

Pope, C. A., Burnett, R. T., Thun, M. J., Calle, E. E., Krewski, D., Ito, K., and Thurston, G. D. 2002. Lung cancer, cardiopulmonary mortality, and long term exposure to fine particulate air pollution. *Journal of the American Medical Association*, 287, 9, 1132–1141.

Rabl, A. 2003. Interpretation of air pollution mortality: Number of deaths or years of life lost? *Journal of the Air and Waste Management*, 53, 1, 41–50.

Rabl, A. and Spadaro, J. V. 1999. Environmental damages and costs: An analysis of uncertainties. *Environment International*, 25, 29–46.

Rabl, A. and Spadaro, J. V. 2002. Health impacts of waste incineration. In *Issues in environmental science and technology. Vol.18 of Environmental Impact of Solid Waste Management Activities*, pp. 171–193. Royal Society of Chemistry, UK.

Rabl, A., Spadaro, J. V., and McGavran, P. D. 1998. Health risks of air pollution from incinerators: A perspective. *Waste Management & Research*, 16, 4, 365–388.

Rabl, A., Spadaro, J. V., and van der Zwaan, B. 2005. Uncertainty of pollution damage cost estimates: To what extent does it matter? *Environmental Science & Technology*, 39, 2, 399–408.

Rowe, R. D., Lang, C. M., Chestnut, L. G., Latimer, D., Rae, D., Bernow, S. M., and White, D. 1995. *The New York Electricity Externality Study,* Oceana Publications, Dobbs Ferry, NY.

Simpson, D. 1992. Long period modelling of photochemical oxidants in Europe. Calculations for July 1985. *Atmospheric Environment*, 26A, 1609–1634.

Simpson, D. and Eliassen, A. 1997. *Control Strategies for Ozone and Acid Deposition—An Iterative Approach. EMEP MSC-W Note 5/97*, Norwegian Meteorological Institute, Oslo.

Spadaro, J. V. 1999. Quantifying the Effects of Airborne Pollution: Impact Models, Sensitivity Analyses and Applications. Doctoral thesis, Ecoledes Mines, 60 boul. St-Michel, F-75272 Paris, France.

Spadaro, J. V. and Rabl, A. 2002. Air pollution damage estimates: The cost per kilogram of pollutant. *International Journal of Risk Assessment and Management*, 3, 1, 75–98.

Spadaro, J. V. and Rabl, A. 2004. Pathway analysis for population-total health impacts of toxic metal emissions. *Risk Analysis*, 24, 5, 1121–1141.

Tol, R. S. J., Downing, T. E., Fankhauser, S., Richels, R. G., and Smith, J. B. 2001. Progress in estimating the marginal costs of greenhouse gas emissions. Special issue. *Pollution Atmosphérique*, December, 155–179.

Trukenmüller, A. and Friedrich, R. 1995. Die abbildung der großräumigen verteilung, chemischen umwandlung und deposition von luftschadstoffen mit dem trajektorienmodell WTM, In *Jahresbericht ALS 1995*, pp. 93–108. Stuttgart.

US Environmental Protection Agency (EPA). 1998. Region 6. Multimedia Planning and Permitting Division Office of Solid Waste, Center for Combustion Science and Engineering. *Human Health Risk Assessment Protocol for Hazardous Waste Combustion Facilities, Support Materials*, http://www.epa.gov/epaoswer/hazwaste/combust/riskvol.htm#volume1 (accessed on September 2004).

Wilson, R. and Spengler, J. D. eds. 1996. *Particles in Our Air: Concentrations and Health Effects*, Harvard University Press, Cambridge, MA.

A. Zoughaib and Rabl, A. 2004. Impacts et coûts de la pollution due au traitement des déchets. *Final Report for ADEME*, Centre Energétique et Procédés, ARMINES/Ecole des Mines, Paris.

5

Distributed Generation and Demand-Side Management

Aníbal T. de Almeida and Pedro S. Moura
Universidade de Coimbra

Clark W. Gellings
Electric Power Research Institute

Kelly E. Parmenter
Global Energy Partners, LLC

5.1 Distributed Generation Technologies

Aníbal T. de Almeida and Pedro S. Moura

5.1.1 Introduction

Distributed generation (DG) can be defined as a source of electric power connected to a distribution network or a customer site, representing an innovative and efficient way to both generate and deliver electricity, since it generates electricity right where it is going to be used. Technological improvements now allow power generation systems to be built in smaller sizes with high efficiency, low cost, and minimal environmental impact.

Distributed generation can serve as a supplement to electricity generated by huge power plants and delivered through the electric grid. Located at a customer's site, DG can be used to manage energy service needs or help meet increasingly rigorous requirements for power quality (PQ) and reliability.

Distributed generation has the potential to provide site-specific reliability improvement, as well as transmission and distribution (T&D) benefits including: shorter and less extensive outages, lower reserve margin requirements, improved PQ, reduced lines losses, reactive power control, mitigation of transmission and distribution congestion, and increased system capacity with reduced T&D investment. Distributed generation also provides economic benefits because DG technologies are modular and provide location flexibility and redundancy as well as short lead times. Economic benefits can also be gained by using DG technologies for peak-shaving purposes, for combined heat and power (CHP) (cogeneration), and for standby power applications. In addition, many DG technologies provide environmental benefits including reduced land requirements, lower or no environmental emissions, and lower environmental compliance costs.

Distributed generation technologies can be divided into two different categories according to availability: firm and intermittent power. The firm power technologies are those that enable the power control of DG units that can be managed as a function of the load requirements. Firm DG plants can be utilized as backup, working only in situations of grid unavailability, in periods of high consumption (when the electricity is more expensive), working continuously, or dispatched to meet the variable load in an optimal manner.

The intermittent power technologies do not allow the management of the produced energy by themselves having a random generation character. Examples of this kind of technology are wind power or solar power that only produces energy when the wind or the sun is available. These technologies can be installed aggregated with energy storage that, by filtering the energy generation fluctuation, enables the management of the delivered energy by the combined system.

In DG applications, traditional technologies can be used, such as internal combustion engines, gas turbines, and, in large installations, steam turbines and combined-cycle turbines. Other kinds of technologies such as microturbines, Stirling engines, fuel cells, or renewable energies, including solar power, geothermal power, or wind power can also be utilized (Figure 5.1).

5.1.2 Gas Turbines

Gas turbines are an often-used electricity production technology. The first studies of gas utilization to actuate turbines started at the end of the nineteenth century; however, the first efficient gas turbines started to operate in 1930. A gas turbine consists of a compressor, a combustion chamber, and a turbine

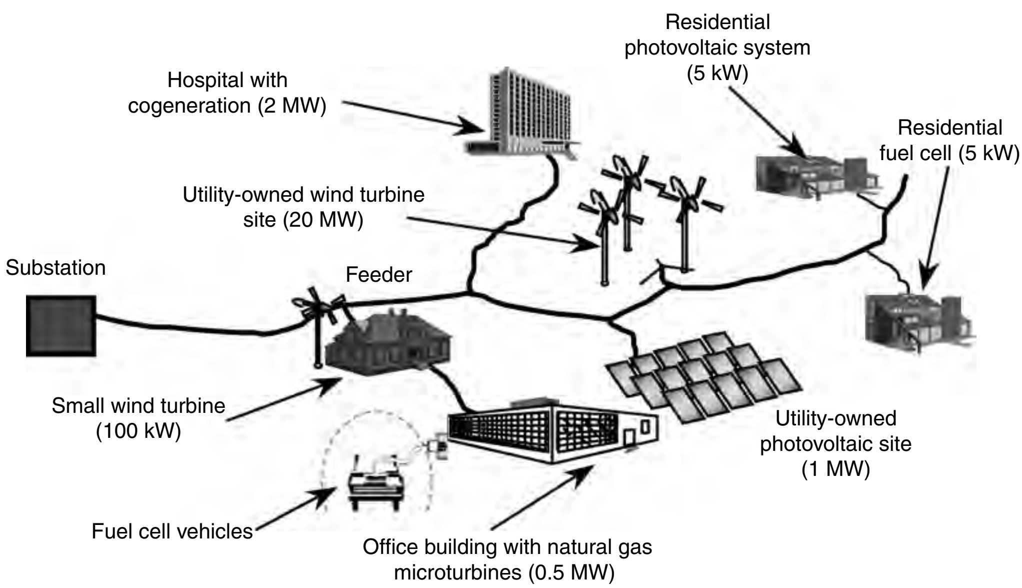


FIGURE 5.1 Power system with multiple energy sources. 1: air intake section; 2: compression section; 3: combustion section; 4: turbine section; 5: exhaust section; 6: exhaust diffuser.

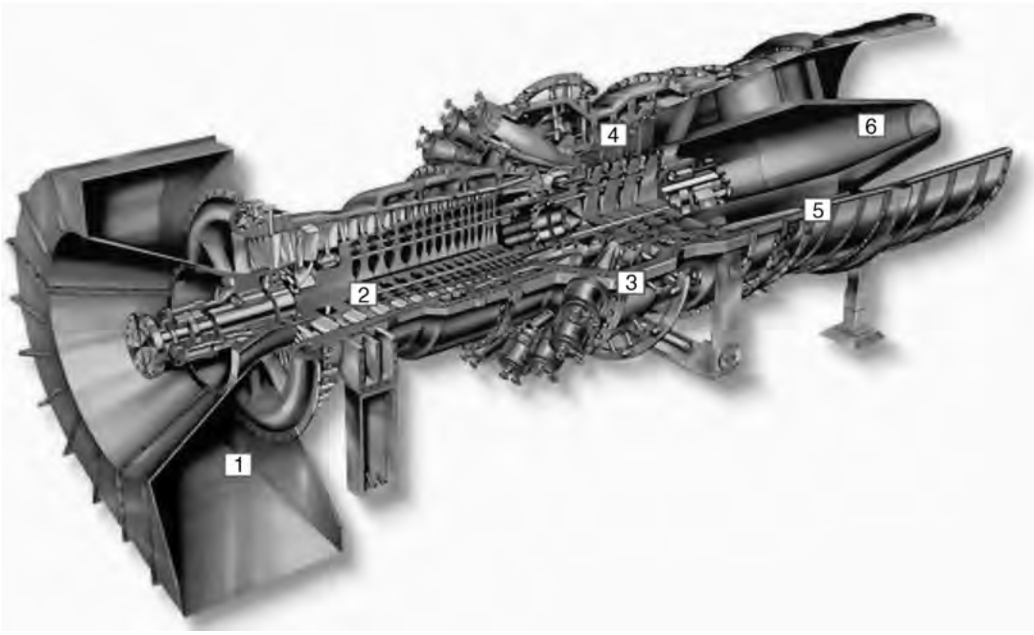


FIGURE 5.2 Combustion turbine. (From Siemens Corp., http://www.siemens.com. With permission.)

coupled to the generator. The turbines with only one axis have all the pieces associated with a continuous axis, all rotating at the same speed. This kind of architecture is used when variations in the turbine speed are not foreseeable. The rotor that drives the generator can be mechanically separated from the rotor driven by the combustion of gases, with more flexibility in the operation speed.

In contrast to internal combustion engines, the gas turbines work in continuous process and not in a repetition of a sequence of different operations. However, the operation can be viewed as a set of four stages similar to the four strokes of the internal combustion engines (Figure 5.2).

1. A compressor drives a rotor that directs the work fluid (air) to the combustion chamber, where the air is compressed, increasing the pressure up to 10 bar and the temperature to 300°C.
2. The compressed air is mixed with the burning fuel, achieving temperatures of 1250°C. This combustion occurs with controlled conditions to maximize the fuel efficiency and minimize the emissions.
3. The air, at high pressure, is passed through the turbine that converts the air energy into mechanical energy. Part of this energy is transmitted to the compressor and the remnant is used for electricity generation through a generator.
4. The exhaust gases are released to the atmosphere, or may be used for generation of process heat or to increase the electricity generated as described below.

Because gas turbines produce a large volume of exhaust gases at high temperatures, the energy of these gases can be utilized for steam production for industrial processes (cogeneration mode) or for electricity production through combined cycle.

In a combined cycle, the gas turbine is used as the first cycle, where the exhaust gases of the gas turbine are used to produce steam in a heat-recovery steam generator. This steam is then used to drive a steam turbine, increasing the electrical global efficiency of the system to values up to 60%.

In the Cheng cycle, the steam is injected in the expansion chamber of the gas turbine (superheated steam injection). In the expansion chamber, the steam is mixed with the gases of the combustion that expand and produce additional work, thereby increasing the electrical efficiency.

TABLE 5.1 General Characteristic of Gas Turbines

Commercial availability	High
Size range	0.5–250 MW
Fuel	Natural gas, biogas, oil derivatives
Efficiency	25%–45%
Environmental emissions	Very low when controls are used, high noise

In cogeneration appliances, the exhaust gases can be used to heat water for residential buildings or for steam production for industrial processes, etc. In some industrial applications of cogeneration, a global efficiency of 60% is reached.

The conversion of mechanical energy to electricity is made almost always through synchronous generators. In DG applications, the rotation speed of the turbines can be higher than the generator synchronous speed which requires a gearbox, reducing the conversion efficiency by about 3%. The generator also works as an auxiliary engine to start the turbine.

Natural gas is the fuel that enables the best efficiency in gas turbines. However, gas turbines can work with other fuels, like fuel oil, diesel, propane, J-5 (used in aeronautics), kerosene, methane, and biogas. The heavy oil utilization decreases the efficiency and the power of the turbine by 5%–8%. Because heavy oil is less expensive, it can decrease the electricity production costs, but the emissions are higher than with other fuels.

The gas turbine generators are available in a wide power range, corresponding at three types of generators:

- Microturbines (20–500 kW)
- Medium turbines (500–10,000 kW)
- Large turbines (more than 10 MW).

These generators use gas turbines with the same working principle, but with different configurations and operation characteristics. The large turbines are not normally considered DG. Table 5.1 shows the general characteristics of gas turbines.

5.1.3 Microturbines

Microturbines are small combustion turbines that produce between 25 and 500 kW of power. Microturbines were derived from turbocharger technologies found in large trucks or in the turbines found in aircraft auxiliary power units. Only the largest class of gas turbine generators, those made for central station utility application, are designed specifically for electric power production. In recent years, due to the new market requirements, microturbines have undergone significant innovations, enabling the energy production with high quality and reliability, with low greenhouse gases emissions, and with moderate costs, thus becoming a competitive technology.

Microturbines have the same working principle as the gas turbines, with various modifications in the system configuration. One of the innovations consists in the adoption of a unique shaft, on which the compressor, the turbine, and the generator are assembled (Figure 5.3). These systems eliminate the gearbox, reducing the cost, and increasing the reliability, but with a reduced overall efficiency.

The rotor rotates at a very high speed (up to 100,000 rpm). Another innovation is utilization of air bearings, avoiding the need of a fluid for refrigeration and lubrication, because the unique utilized element is the air. The air can be continuously renovated and will never be contaminated by the materials wastage and by the combustion products.

One of the key characteristics of the microturbines is heat recovery, which utilizes the thermal energy of the exhaust gases (at high temperatures) for preheating the air supply to the compressor. The mechanical energy is converted to electrical energy by a permanent magnet AC generator, which includes a low inertia rotor rotating at the turbine speed.

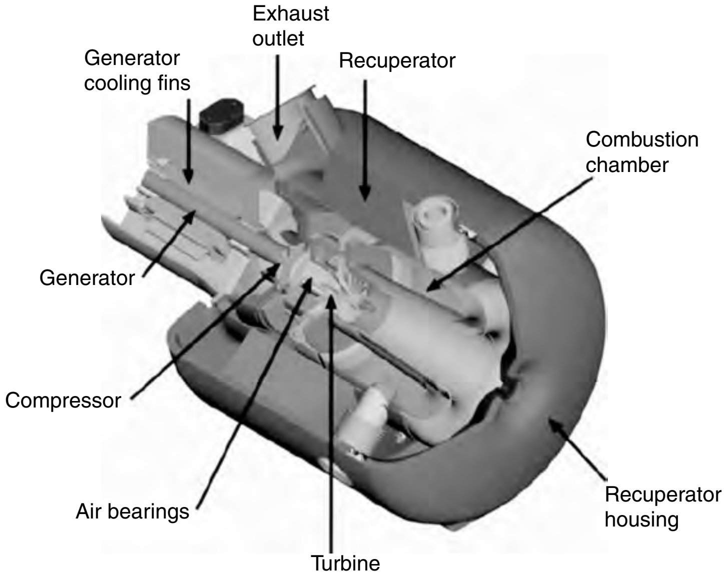


FIGURE 5.3 Microturbine. (From http://www.capstone.com. With permission.)

Because of the high rotor speed, the AC output has a frequency of approximately 2 kHz. To connect the microturbine generator with a 50 or 60 Hz network in normal applications, the microturbine voltage output must be connected to an AC–DC–AC converter. In this converter, microturbine voltage output is rectified, filtered, and converted to an AC voltage through an inverter system synchronized with the 50 or 60 Hz supply.

Microturbines can also operate with a wide variety of fuels, like natural gas (at high or low pressure), propane, diesel, gasoline, biogas (methane), or kerosene. The electrical efficiency of microturbines is between 20 and 30%, with heat-recovery utilization. If the heat recovery does not exist, this value can decrease to 15%. In cogeneration systems, the global efficiency can reach 85% (Table 5.2).

Microturbine generators can be divided into two general classes:

- Heat-recovery microturbines recover heat from the exhaust gas to boost the temperature of the air stream supplied to the combustion and increase the efficiency. Further exhaust heat recovery can be used in a cogeneration configuration.
- Microturbines, without heat recovery (or simple cycle) have lower efficiencies, but also have lower capital costs.

Other applications of microturbine technology include:

- Core power conversion element of vehicles, such as buses, trucks, helicopters, and so on. Automotive companies are interested in microturbines to provide a lightweight and efficient fuel-based energy source for hybrid electric vehicles.

TABLE 5.2 Microturbines Overview

	Yes (only a few manufacturers)
Size range	25–500 kW
Fuel	Natural gas, hydrogen, propane, diesel
Efficiency	20%–30% (recuperated)
Environmental emissions	Low ($<$9–50 ppm) NO_x
Other features	Cogeneration (50°C–80°C water)
Commercial status	Medium volume production

- Standby power, PQ, peak-shaving, and cogeneration applications. Some types of microturbines are well suited for small commercial building establishments such as: restaurants, hotels, small offices, retail stores, and many others.
- Utilization of by-products of processes in oil-processing, gas-transferring, petroleum production, industrial waste utilization for the purpose of optimizing the use of natural gas, associated gas, biogas, landfill gas, etc.

The improvement of microturbine design, resulting in lower costs and higher performance, makes microturbines a competitive DG product.

Development is ongoing in a variety of areas:

- Heat recovery/cogeneration
- Use of waste heat for absorption cooling
- Increase of the efficiency
- Fuel flexibility
- Vehicles
- Hybrid systems (e.g., fuel cell/microturbine, flywheel/microturbine)

5.1.4 Internal Combustion Engines

Internal combustion (IC) engines were one of the first technologies that used fossil fuels for electricity generation. Developed more than a century ago, IC engines are the most common of all DG technologies. They are available from sizes of a few kilowatts for residential backup generation to generators on the order of 10 MW.

An IC engine uses the thermal energy of fuel combustion to move a piston inside a cylinder, converting the linear motion of the piston to rotary motion of a crankshaft and uses that rotation to turn an AC electric generator. Internal combustion engines are also called *reciprocating engines* because of the reciprocating linear motion of the pistons.

Internal combustion engines can be fueled with gasoline, natural gas, diesel fuel, heavy oil, biodiesel, or biogas. The two primary types of IC engines used for DG applications are:

- Four-cycle spark-ignited engines (Otto cycle) that use an electrical spark introduced into the cylinder. This explosion engine, or ignition by spark, that uses the Otto cycle was invented by Nikolaus August Otto in 1867. Fast-burning fuels, like gasoline and natural gas, are commonly used in these engines. Biofuels, such as alcohols and biogas, may also be used.
- In 1892, Rudolph Diesel, developed the diesel engine, in which the combustion is initiated by compression. The compression-ignited (diesel cycle) engines, in which compression of the fuel–air mixture inside the piston cylinder rises it to a temperature where it spontaneously ignites, work best with slow-burning fuels such as diesel. Biofuels, such as biodiesel, vegetable oils, etc., may also be used.

Distributed generation engines have efficiencies that range from 25 to 45% (Table 5.3). In general, diesel engines are more efficient than Otto engines because they operate at higher compression ratios. In the future, engine manufacturers are targeting lower fuel consumption and shaft efficiencies up to 50%–55% in large engines (greater than 1 MW) by 2010. Efficiencies of Otto engines using natural gas are expected to improve and approach those of diesel engines.

Internal combustion engine generators for distributed power applications, commonly called *gensets*, are found universally in sizes from less than 5 kW to over 10 MW. Gensets are frequently used as a backup power supply in residential, commercial, and industrial applications. When used in combination with a 1–5 min uninterruptible power supply (UPS), the system is able to supply seamless power during a utility outage. In addition, large IC engine generators may be used as base-load generation, grid support,

TABLE 5.3 Internal Combustion (IC) Engines Overview

Commercially Available	Yes
Size range	0.005–10 MW
Fuel	Natural gas, diesel, heavy fuel, biogas
Efficiency	25%–45%
Environmental	Emission controls required for NO_x and CO
Other features	Cogeneration (some models)
Commercial status	Widely available

or peak-shaving devices. Internal combustion engine generators have start-up times ranging between 0.5 and 15 min, and a high tolerance for frequent starts and stops. The smaller engines, available in sizes as small as a few kilowatts, are intended for dispersed applications, such as individual homes and small businesses coping with power outages.

5.1.5 Stirling Engines

Stirling engines (Figure 5.4) are a class of reciprocating piston engines and are classed as external combustion engines, invented in 1816 by Robert Stirling. They constitute an efficient thermodynamic machine for the direct conversion of heat into mechanical work with a theoretical efficiency of 40%. Stirling engines were commonly used prior to beginning of the twentieth century. As steam engines improved and the competing compact Otto cycle engine was invented, Stirling engines lost favor. Recent developments in DG and solar thermal power have revived interest in Stirling engines. As a result, research and development efforts in this area have increased in recent years. The principles of Stirling engines are described in another chapter and will not be covered here.

The operation is reversible, i.e., by supplying thermal energy, mechanical energy is produced, and by supplying mechanical energy, thermal energy is produced. Stirling engines can be fueled by any source of

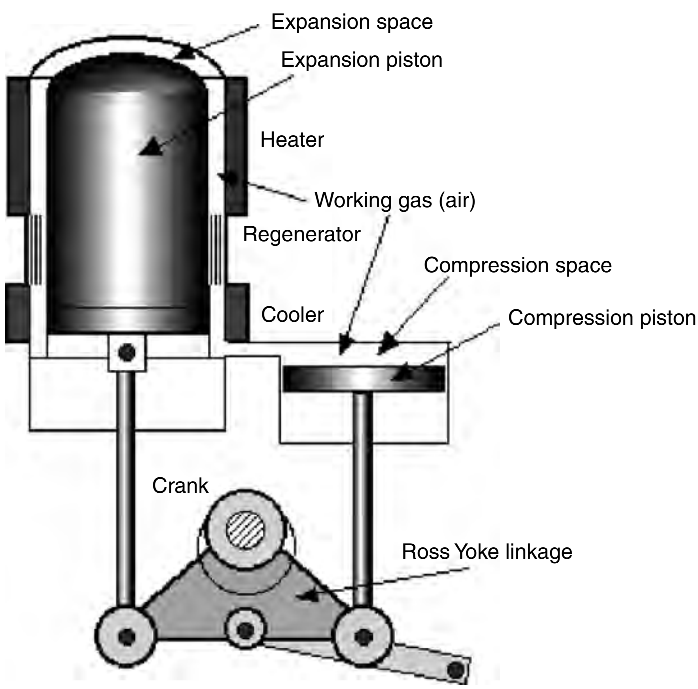


FIGURE 5.4 Stirling engine. (From http://www.ent.ohiou.edu/~urieli/stirling/engines/engines.html)

TABLE 5.4 Stirling Engine Overview

Commercially available	On a limited scale
Size range	<1–25 kW
Fuel	Fuel flexibility—fossil or renewable heat is possible
Efficiency	12%–30%
Environmental	Potential for low emissions
Other features	Cogeneration (some models)
Commercial status	Availability for specialized applications

heat (fossil fuel or renewable) and some models have possibilities to perform cogeneration. In a Stirling engine, the continuous combustion process, which is easier to optimize and control, results in lower emissions compared to the intermittent explosions of fuel air mixtures in IC Otto and diesel engines.

Stirling engines are commercially available for some marine applications and have been available for several years. Trials in domestic CHP are occurring in several countries, on the scale of hundreds of units. Usually Stirling engines are found in sizes from 1 to 25 kW and are currently being produced in small quantities for specialized applications (Table 5.4).

Large 25-kW Stirling motors have an electrical efficiency of approximately 30%, although the goal is to increase this efficiency to greater than 34% with more development. Stirling engines are ideally suited for solar thermal power. Using a gas as an operating fluid, there is no practical limit placed on the solar unit's upper temperature due to its operating fluid. Maximum temperature would be limited only by the materials used in its construction.

Recent Stirling engine developments have been directed at a wide range of applications, including:

- Small scale: residential or portable power generation.
- Solar dish applications: heat reflected from concentrating dish reflectors is used to drive the Stirling engine; several research programs are aimed at enhancing this application.
- Vehicles: auto manufacturers have investigated utilizing Stirling engines in vehicles to improve the fuel economy.
- Refrigeration: Stirling engines are being developed to provide cooling for applications, such as microprocessors and superconductors.
- Aircraft: Stirling engines could provide a quieter-operating engine for small aircraft.
- Space: Power generation units aboard space ships and vehicles.

The primary challenges faced by Stirling engines over the last two decades have been their long-term durability/reliability and their relatively high cost.

5.1.6 Fuel Cells

Fuel cells are a technology for power generation that is quiet and highly efficient with no moving parts. Fuel cells generate electricity through an electrochemical process in which the energy stored in a fuel is converted directly into DC electricity and thermal energy. The chemical energy normally comes from hydrogen contained in various types of fuels (hydrocarbon fuels, such as natural gas, methanol, ethanol, biogas, etc.), including pure hydrogen. Fuel cells are described in detail in another chapter.

Several hundred phosphoric acid fuel cell (PAFC) demonstration and test plants have been built in the mid-1990s to early twenty-first century, mostly with 200-kW capacity appropriate for DG applications, in many commercial buildings to provide premium PQ for demanding loads. The operating temperature is about 200°C, which is suitable for cogeneration applications in buildings and in small industrial plants. They do not offer opportunity of self-reforming and they require platinum for their catalyst. PAFCs' efficiency and peak output capability deteriorate by about 2% per year.

One of the most promising developments is the design of hybrid systems. Solid oxide fuel cells (SOFCs) promise top efficiency, particularly in a combined cycle operation mode, in which they can surpass conventional combined-cycle gas turbine plants. High efficiencies under part-load operation also result in high overall efficiency.

For applications in industrial heat-power cogeneration and public electricity supply, high-temperature fuel cells (SOFCs and molten carbonate fuel cells [MCFCs]) are most suitable. Both systems are still in an early stage of development, but permit the use of a wide range of fuels. Such fuel cell systems compete in the lower rating range with gas turbines and motor cogeneration plants, and in the upper rating range with combined gas and steam turbine power plants. Conventional plants have a clear advantage in terms of practical experience and in terms of comparatively low capital costs compared with fuel cell plants. High-temperature fuel cells are likely to gain market penetration due to a decrease in specific need for primary fuels and also due to a sharp decrease of specific pollutant emissions in comparison to conventional generation. This last factor is a key advantage in DG applications in urban areas.

5.1.7 Solar Photovoltaic

Photovoltaic (PV) cells, or solar cells, convert sunlight directly into electricity. Photovoltaic technology has several applications, including:

- Off-grid/remote
- Grid attached residential and commercial buildings
- Remote communication systems
- Central power plants (above 1 MW)

Traditionally, PV cells have been used to power structures such as individual homes in locations where it is expensive or impossible to send electricity through power lines. Solar power has traditionally been used in remote areas where the grid is not available; such systems store electricity in batteries for use when the sun is not shining and are called *stand-alone power systems*. Currently, PV-generated power is less expensive than conventional power where the load is small and the area is too difficult to serve by electric utilities. However, solar power is now appearing more in urban areas due to innovative policy mechanisms (rebates, feed-in tariffs) to promote PV generation. Here, the surplus solar electricity is injected into the grid. These are called *grid-connected solar systems* because the owner has the security of the grid available.

In the decade 1995–2004, average annual growth has been 20% with global sales reaching over 1000 MW per year at the end of that period. Photovoltaic is the most modular and operationally simple of the clean, distributed power technologies, with benefits that include the ability to provide power during summer peak periods, distribution congestion benefits, environmental benefits, reduced fuel price risk, and local economic development. As a result of private and government research, PV systems are becoming more efficient and affordable.

Distributed PV systems that provide electricity at the point of use are reaching widespread commercialization. Chief among these distributed applications are PV power systems for individual buildings. Interest in the building integration of photovoltaics, where the PV elements actually become an integral part of the building, often serving as the facade or exterior weather skin, is growing worldwide. Photovoltaic specialists and innovative designers in Europe, Japan, and in the U.S. are now exploring creative ways of incorporating solar electricity into their work. A building integrated photovoltaics (BIPV) system consists of integrating photovoltaics modules into the building envelope, such as the roof or the façade (Figure 5.5). By simultaneously serving as building envelope material and power generator, BIPV systems can provide savings in materials and in electricity costs.

Photovoltaics may be integrated into many different assemblies within a building envelope:

- Incorporated into the façade of a building, complementing or replacing traditional glass.
- Incorporated in the external layers of the wall of a building façade.

FIGURE 5.5 Building integrated photovoltaics in a façade and in a roof.

- Use in roofing systems providing a direct replacement for different types of roofing material.
- Incorporated in skylight systems in which part of the solar light is transmitted to the inside of the building and the other part is converted into electricity.

5.1.8 Wind Power

Wind energy became a significant research area in the 1970s during the energy crisis and the resulting search for potential renewable energy sources. Modern wind turbine technology has made significant advances over the last 25 years. Today, wind-power technology is available as a mature, environmentally sound, and convenient alternative. Generally, individual wind turbines are grouped into wind farms containing several turbines. Many wind farms are megawatt scale, ranging from 1 MW to tens of megawatts. Wind turbines may be connected directly to utility distribution systems. The larger wind farms are often connected to transmission lines.

The land can still be used for animal grazing and some agriculture operations. The small-scale wind farms and individual units are typically defined as DG. Residential systems (1–15 kW) are available (Figure 5.6). However, they are generally not suitable for urban or small-lot suburban homes due to large space requirements.

FIGURE 5.6 Building-integrated wind power. (From http://www.oksolar.com/wind/. With permission.)

Utility-scale turbines range in size from 50 to 5000 kW. Single, small turbines, below 50 kW, are used for homes, telecommunications stations, or for water pumping. A detailed description of wind-power technology is given in a different chapter.

5.1.9 Cogeneration

Combined heat and power, also known as *cogeneration*, is an efficient, clean, and reliable approach to producing both electricity and usable thermal energy (heating and/or cooling) at high efficiency and near the point of use, from a single fuel source. Because CHP is highly efficient, it reduces traditional air pollutants and carbon dioxide emissions, the leading greenhouse gas associated with climate change.

Combined heat and power can use a variety of technologies to meet an energy users needs. The range of technologies available allows the design of cogeneration facilities to meet specific onsite heat and electrical requirements. Combined heat and power systems consist of a number of individual components—prime mover, generator, heat recovery, and electrical interconnection—configured into an integrated system. The type of equipment that drives the overall system (i.e., the prime mover) typically identifies the CHP system.

Typical CHP prime movers include:

- Combustion turbines
- Reciprocating engines
- Boilers with steam turbines
- Microturbines
- Fuel cells

Combined heat and power may be used in a variety of applications ranging from small 10-kW systems to very large utility-scale applications approaching 1000 MW. The first step in assessing which CHP application is right for a particular facility is to identify whether there is coincident demand of electrical and thermal energy at the host site. The CHP project will be most economically viable when the system

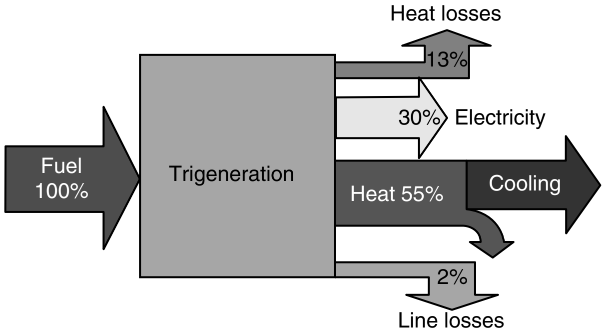


FIGURE 5.7 Trigeneration technology.

provides the maximum amount of energy that can be used. Therefore, CHP project development begins with an analysis of site electrical and thermal load profiles. Based on these profiles, the type of CHP technology which most closely matches the facility's power and demand will be chosen.

In developed countries, about 10% of all electricity is generated in CHP plants, leading to huge primary energy savings and reduction of emissions. Combined heat and power in some industries, such as the pulp and paper industry, uses biomass by-product as fuel input.

Trigeneration can provide even greater efficiency than cogeneration. Trigeneration is the conversion of a single fuel source into three useful energy products: electricity, steam or hot water, and chilled water (Figure 5.7). Trigeneration converts and distributes up to 90% of the energy contained in the fuel burned in a turbine or engine into usable energy. Introducing an absorption chiller into a cogeneration system means that the site is able to increase the operational hours of the plant with an increased utilization of heat, particularly in summer periods. Trigeneration has been applied with very positive results in buildings such as hotels and hospitals, which feature a large space conditioning load during most of the year.

5.1.10 Vehicle-to-Grid

Electric-drive vehicles (EDVs) include battery electric vehicles, hybrid vehicles, and fuel cell vehicles running on gasoline, natural gas, or hydrogen. These vehicles have gained attention in the past few years due to growing public concerns about urban air pollution and other environmental and resource problems. All these vehicles have within them power electronics that generate AC power at power levels from 10 to 100 kW; this power, with suitable electronics, can be fed into the electric grid (Figure 5.8). It allows such vehicles to use their installed power to help balance load in localized grid segments during peak load periods. This concept of bidirectional grid interface is known as *vehicle-to-grid power* or V2G. Sharing power assets between transportation and power generation functions can accelerate commercialization of battery electric vehicles, hybrid vehicles, and fuel cell vehicles.

These vehicles can be recharged during off-peak hours at cheaper rates while helping to absorb excess nighttime generation. There is a potential to supply extra power during peak demand if electric-drive vehicles are grid-connected to allow discharge from their batteries, or run their onboard generators.

To work in vehicle-to-grid power systems, each vehicle must have three required elements:

- A connection to the grid for electrical energy flow.
- Control or logical connection necessary for communication with the grid operator.
- Controls and metering onboard the vehicle.

For fueled vehicles (fuel cell and hybrid), a fourth element, a connection for gaseous fuel (natural gas or hydrogen), could be added so that onboard fuel is not depleted.

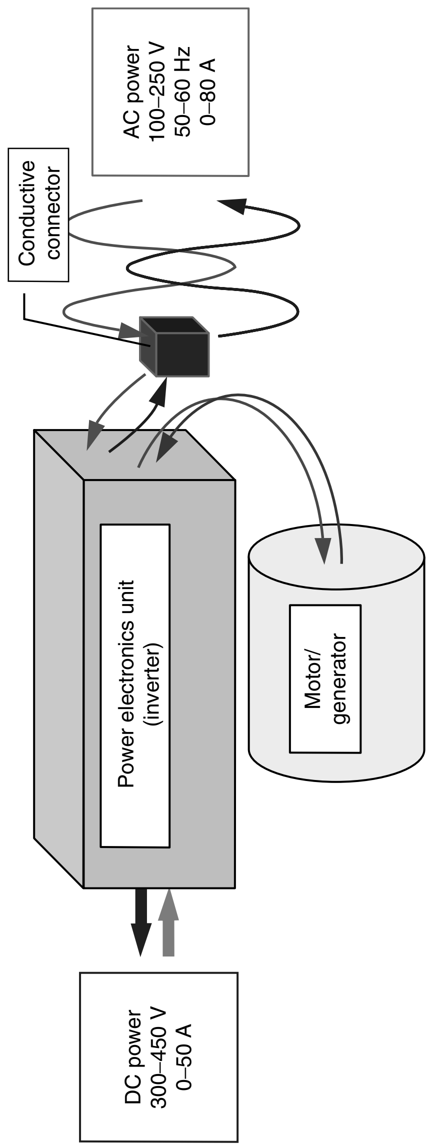


FIGURE 5.8 Vehicle-to-grid enabling technology. (From AC Propulsion. Vehicle-to-grid demonstration project: Grid regulation ancillary service with a battery electric vehicle (December 3–10, 2002, http://www.acpropulsion.com/reports/V2G%20Final%20Report%20R5.pdf. With permission.)

Vehicles with significant energy stored in batteries could perform as uninterruptible power systems for whole houses and support the grid exceptionally well by providing any of a number of functions known collectively as *ancillary services*. These services are vital to the smooth and efficient operation of the power grid. These vehicles could provide:

- Extra power during demand peaks
- Spinning reserve
- Grid regulation (automatic generation control (AGC))
- Uninterruptible power source for businesses and homes
- Active stability control of transmission lines

Hybrid vehicles with IC engines show the potential for power generation at specific emissions levels in some cases better than the best new large power plants. A continuous source of fuel for hybrid or fuel cell vehicles can be provided with a connection to low-pressure natural gas or biogas at compatible parking locations. Over just a decade or two, V2G could revolutionize the ancillary services market, improve grid stability and reliability, and support increased generation from intermittent renewables.

One conceptual barrier is an initial belief that their power would be unpredictable or unavailable because the vehicles would be on the road. Although availability of any one vehicle is unpredictable, the availability of an average number of vehicles is highly predictable and can be estimated from traffic and road-use data.

From the societal point of view, the large-scale application of V2G can lead to a much more cost-effective allocation of resources to provide highly reliable electricity to a wide range of resources.

5.1.11 Conclusions

The utilization of DG technologies enables the creation of a power system with multiple energy sources, allowing the integration conventional central power plants with dispersed DG fossil fuel-based generation, as well as with dispersed DG renewable generation. It is anticipated that DG growth and its large-scale application may lead to an improvement in reliability, to improved security of supply, to the decrease of power costs, and to the minimization of environmental impacts. In Figure 5.9 and Table 5.5, a characterization of the most important DG technologies is presented showing typical parameters associated with each technology.

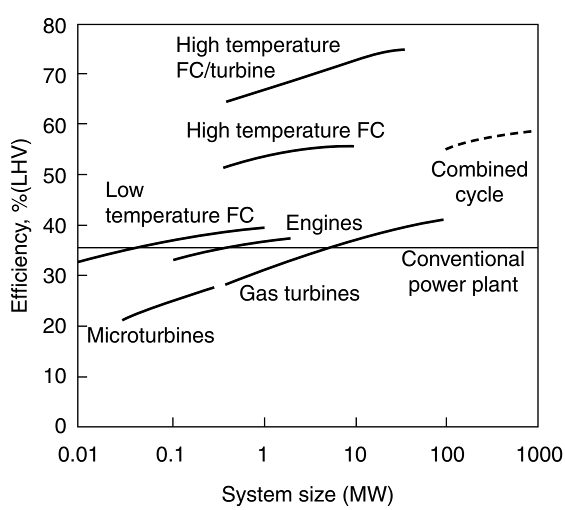


FIGURE 5.9 Distributed generation technologies comparative efficiency range.

TABLE 5.5 Distributed Generation (DG) Technologies Characterization

	Internal Combustion Engines	Gas Turbines	Microturbines	Fuel Cell	Wind Power	Photovoltaic (PV)	Stirling Engine
Power range (kW)	5–50,000	500–250,000	20–500	1–10,000	0.3–5000	0.07–1000	Up to 25
Electric efficiency (%) (LHV)	25–45	25–45	20–30	30–70	25–40	5–15	12–30
Efficiency with partial load	Reasonable until 35%–40% of the rated load	Reasonable until 40% of the rated load	Bad below 40% of the rated load	Good/reasonable until 35%–40% of the rated load	[a]	[a]	[a]
Load following capacity	Very good	Good[b]	Reasonable/low[c]	Very good	[a]	[a]	[a]
Start time	10 s to 15 min	2 min to 1 h	60 s	[a]	[a]	NA	[a]
Availability (%)	90–98	90–98	90–98	90–95	10–40[d]	5–25[d]	High[e]
Interval between the maintenance stops (×1000 h)	0.5–2	30	5–8	10–40	4	NA	[a]
Useful life time (years)	15–20	20–25	10	20[e]	20	20–30	Long [e]
Fuel flexibility	Good	Good	Good	Good	NA	Good	Excellent
Noise	High	Moderate to high	Moderate	Low	Moderate	NA	Low
Acquisition costs (€/kW)	300–900	300–1000	700–1100	> 4000	700–1300	4000	2000–50,000
O&M costs (€/kWh)	0.005–0.015	0.004–0.010	0.005–0.016[e]	0.005–0.030	[a]	[a]	[a]
Energy costs (€/kWh)	0.07–0.15	0.05–0.15	0.05–0.15	> 0.15	0.03–0.20	> 0.20	[a]
Commercial availability	High	High	Moderate	Low	High	High	Very low

[a] Dependent upon conditions.
[b] Installation costs of the system boiler/turbine.
[c] Despite present a high potential to load folowing, the actual models do not present a good fulfillment at this level.
[d] Depending on the climate condition at the location.
[e] Estimated value.
NA, not applicable.

5.2 Integration of Distributed Generation into Grid

Aníbal T. de Almeida and Pedro S. Moura

5.2.1 Introduction

Connecting a distributed power system to the electricity grid has potential impacts on the safety and reliability of the grid, which is one of the most significant barriers to the installation of DG technologies. Electric utilities have understandably always placed a high priority on the safety and reliability of their electrical systems. Faced with the interconnection of potentially large number of distributed generators, utilities have perceived DG as a threat. This has led some utilities to place overly conservative restrictions on interconnected systems, causing added costs that may make an installation economically unfeasible. Several techniques may reduce adverse network impacts allowing DG connection, but those techniques can be project specific and may be expensive, and adversely affect project economics.

Connection of DG fundamentally affects the operation of distribution networks with changes and impacts like:

- Voltage fluctuations
- Increased fault levels
- Degraded protection
- Bidirectional power flow
- Altered transient stability

To reduce the impact in the power grid several requirements are needed. Typical requirements include equipment that prevents power from being fed to the grid when the grid is de-energized, manual disconnects and PQ requirements such as limits on the interconnected system's effects on "flicker," harmonic distortion, and other types of waveform disturbance. Systems may also be required to automatically shut down in the event of electrical failures, to provide an isolation transformer for the system, as well as to provide liability insurance.

Up to recently the lack of a well-defined interconnect standard and failure to adhere to a standard can add considerably to engineering and equipment costs, making process planning difficult. Many interconnection requirements were drafted and adopted without understanding of the protection capabilities of modern DG equipment. As a result, these requirements often unnecessarily burden projects with redundant studies and hardware.

Interconnection requirements for large DG installations ($\sim$10 MW) are well understood because they are very similar to the interconnections required for central power stations. Interconnection requirements for smaller installations are more difficult because the utility must balance the desire for a safe interconnection with the plant owner's desire to have a "quick and easy" interconnection design to get the DG up and running. Interconnection complexity generally increases with project size and is technology dependent.

Grid interconnection is important for three reasons:

- The number of small generators seeking interconnection to the grid will increase in the future.
- Distributed generation advocates contend that the current interconnection requirements and processes are effectively increasing costs unfairly and pricing DG out of the market.
- Distribution companies are concerned that DG will negatively impact the safety and reliability of the grid and unfairly increase the distribution companies' costs.

5.2.2 Power Distribution

Electric grid is broadly divided into two systems: the transmission system that transfers bulk power at high voltages from power plants to utility-owned substations and a few very large customers, and the

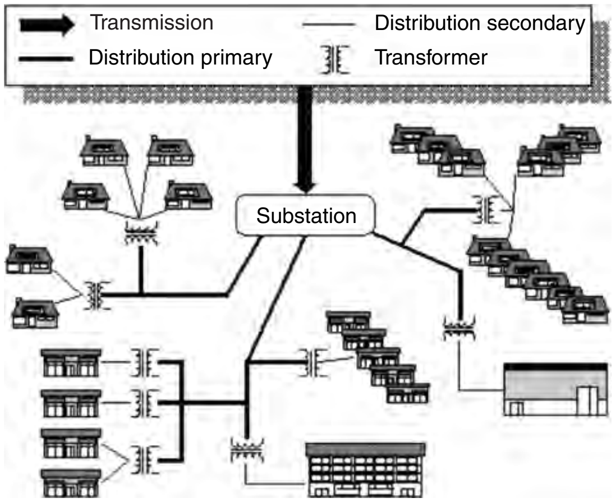


FIGURE 5.10 Radial distribution system. (From California Energy Commission, California interconnection guidebook: a guide to interconnecting customer-owned electric generation equipment to the electric utility distribution system using California's electric rule 21, California Energy Commission, Sacramento, CA, 2003. http://www.energy.ca.gov/reports/2003-11-13_500-03-083F.PDF. With permission.)

distribution system that delivers power at medium and low voltages from the substation to the majority of customers.

The utility distribution systems can be categorized as either radial or networked. Power system design is a tradeoff between complexity and cost to maximize economy and reliability. As a result, the general structure of the power delivery system has a networked nature at the transmission level and a more radial nature at distribution level.

Radial distribution refers to a system where the power lines extend from a common substation to the customer loads coming off at single nodes along the line (Figure 5.10). In these distribution systems, power can only flow in one direction: from the substation to a load. Although there can be many radial distribution lines emanating from a substation, each load is typically served by only one line. A disruption to that feed or substation will typically affect all customer loads on that line. A radial system generally offers a less reliable power source than a networked system because it lacks redundancy. However, the radial system and its protection equipment are less complex and less expensive than the networked system.

The introduction of an energy source such as DG within the radial distribution system will affect the load distribution in the system, and may even cause reverse power flow if it is large relative to the load. Introduction of a sufficiently large power source within the radial distribution normally requires some modification to the protection system.

Networked distribution refers to a system where numerous separate lines form a grid so that customer loads can tap-off of multiple independent feeds, which are then tied to a common bus on the secondary side of the transformers (Figure 5.11). These can be separate lines from a common substation or they can be from independent substations.

The networked system offers reliability advantages over the radial system because it provides multiple power sources for loads. This multipath design is sometimes referred to as a *looped* system. These systems use network protectors that quickly isolate faults to protect the grid and shift customer loads onto the remaining feeds. Understandably, utilities are reluctant to allow the interconnection of anything that they feel will endanger the integrity or safety of this system.

System protection in a networked distribution system is more complex and expensive than in the radial distribution system due to the extra intelligence needed for reliable, effective protection.

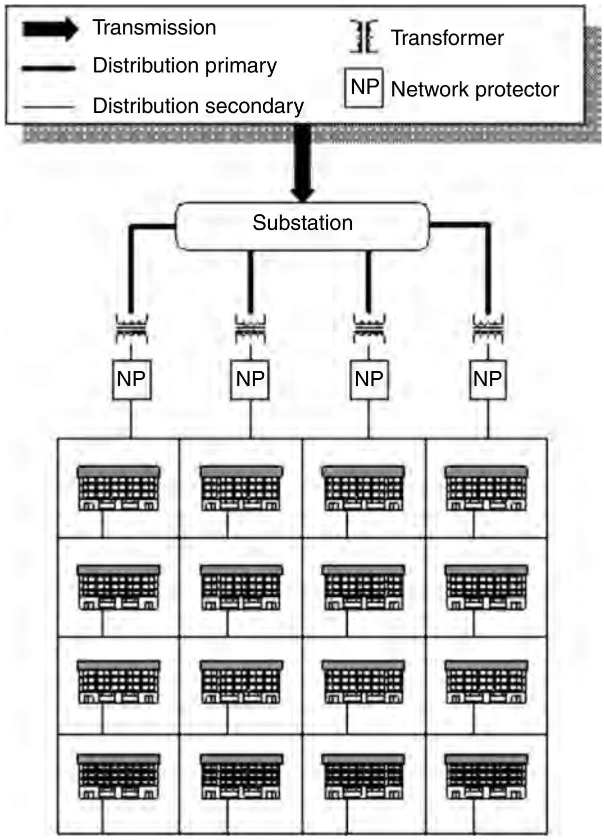


FIGURE 5.11 Networked distribution system. (From California Energy Commission, California interconnection guidebook: a guide to interconnecting customer-owned electric generation equipment to the electric utility distribution system using California's electric rule 21, California Energy Commission, Sacramento, CA, 2003. http://www.energy.ca.gov/reports/2003-11-13_500-03-083F.PDF. With permission.)

Because the networked system is specifically designed to deliver energy from multiple transformers to loads, it is capable of dealing with reverse power flow.

5.2.3 Types of Grid Connections

The electric power system interface is the means by which the DG unit electrically connects to the power system outside the facility in which the unit is installed. Depending on the application and operation of the DG unit, the interface configuration can range from a complex parallel interconnection, to being nonexistent if the DG unit is operated in isolation (Figure 5.12).

In remote applications, due to the high costs of the power grid expansion to the consumption site, the option of interconnecting with the local grid may be impractical. In these cases, the DG units become the unique means of energy supply at low cost. In this configuration, the DG unit provides power for all loads completely isolated from grid, providing the utility no backup or supplemental power.

In near-to-grid applications, the DG unit owner can opt for interconnection, by several types of connections. Depending on the application and the operation mode of the DG unit, the connection system with the grid can represent a complex parallel interconnection or can be inexistent if the DG is operating isolated from the grid. The complexity of the interconnection systems increases with the required interaction level between the DG unit and the distribution grid.

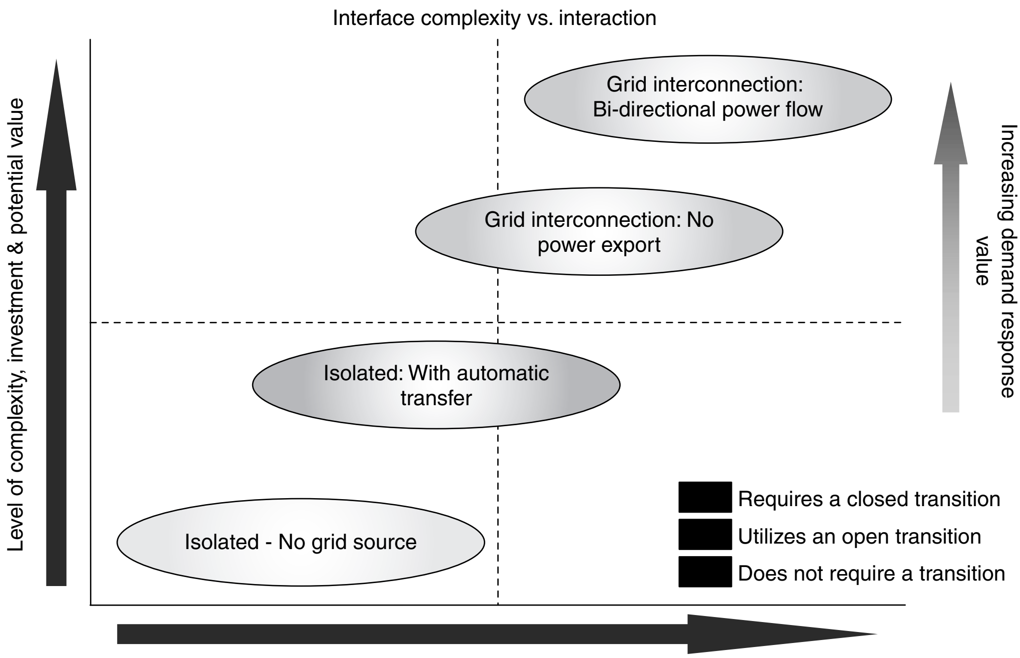


FIGURE 5.12 Complexity vs. interaction. (From Arthur D. Little, *Distributed Generation: System Interfaces*, An Arthur D. Little white paper, ADL Publishing, Boston, MA, 1999. http://www.encorp.comdwnld/pdf/whitepaper/ADLittleWhitePaperDGSystemInterfaces.pdf. With permission.)

For most customers, DG systems are most cost-effective and efficient when they are interconnected with the utility grid. In simple terms, "interconnected with the grid" means that both the DG system and the grid supply power to the facility at the same time. Paralleled systems offer added reliability, because when the DG system is down for maintenance, the grid meets the full electrical load, and vice versa.

Distributed generation systems can be designed to keep a facility up and running without an interruption if the grid experiences an outage. Also, grid-interconnected systems can be sized smaller to meet the customer's base load as opposed to its peak load. Not only is the smaller base-load system cheaper, it also runs closer to its rated capacity and, therefore, is more fuel efficient and cost-effective.

Two different types of grid interconnection are possible: parallel or roll-over. With the parallel operation, the DG system and the grid are interconnected and both are connected to the load. In the roll-over operation, the two sources are interconnected, but only one is connected with the load.

A typical interconnection system includes three kinds of equipment:

- Control equipment for regulating the output of the DG
- A switch and circuit breaker (including a "visible open") to isolate the DG unit
- Protective relaying mechanisms to monitor system conditions and to prevent dangerous operating conditions

5.2.3.1 Isolated Operation

In remote applications, the DG units become the unique means of energy supply at low cost. In this configuration, the DG unit provides power for all loads completely isolated from grid, providing the utility no backup or supplemental power (Figure 5.13). Isolated operation is also possible in sites that are normally connected to the grid but in which continuous supply is required in the event of an outage. Some generating facilities, such as a hospital emergency generator, power the customer's partial or entire load isolated from the utility.

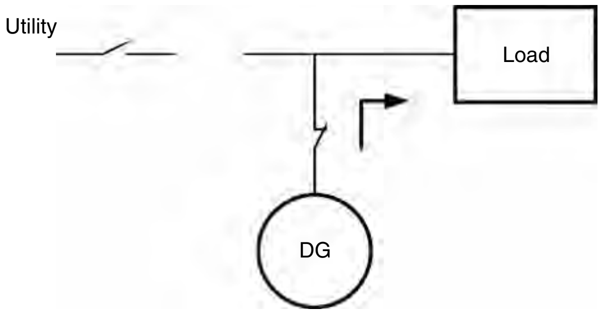


FIGURE 5.13 Isolated operation.

Isolated operation involves no interaction with the utility's distribution system because the generator does not operate in parallel with the utility. In some isolated systems, the generator is sized for a specific load that is always powered from the generator and never from the utility. There are two ways of transferring load to isolated operation:

- A break-before-make transfer switch (also known as *open transition switching*), disconnects the load from the utility prior to making the new connection with the onsite electric generating facility.
- A momentary-parallel (or closed transition) switch, a control system starts the customer's generator and parallels it with the utility's distribution system, quickly ramps the generator output power to meet the customer's load demand and then disconnects the load from the utility.

5.2.3.2 Roll-Over Operation

When a roll-over connection exists (Figure 5.14), the load can be connected only to one of the two sources (grid or DG system) at any given moment. Both the sources are connected to a load control center with a load transfer switch. When the source that is feeding the load fails, this device makes the commutation, ensuring that the other source feeds the load. This commutation can be automatic or manual, to allow the interchange between the two sources due to technical or economical reasons, even when a failure does not exist. In this configuration, the DG unit provides power to load 2 for peaking, base-load, or backup power, and the utility provides power to load 1 and occasionally to load 2.

The automatic commutation devices achieve the fast transition between sources, but there is always a time period in which the load is not fed. This kind of connection does not allow the decrease of

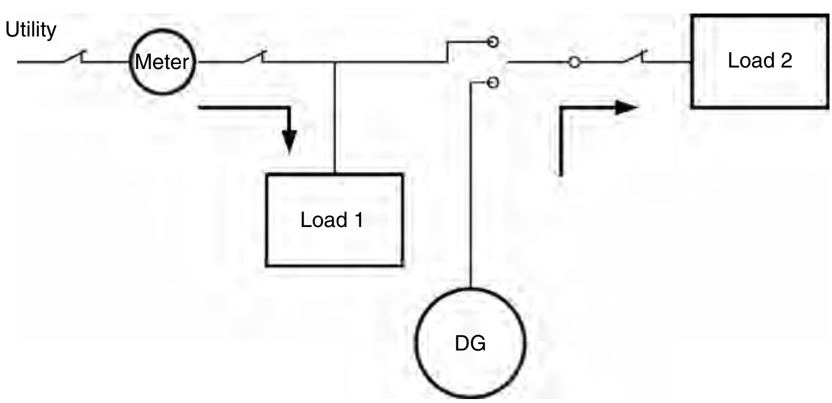


FIGURE 5.14 Roll-over operation.

the frequency of interruption, but reduces considerably their duration. To eliminate the feeding interruption, the installation of energy storage devices between the switch and the load would be necessary. In this configuration, the transfer time of the load switch is typically 0.1–0.15 s.

The roll-over operation is cheaper and simpler because it does not need a high number of control, protection, and coordination equipment. This type of operation can easily ensure the impossibility of the DG system injecting energy into the grid when the grid is out of service. This phenomenon, named *backfeed* or *islanding*, can be harmful to the grid operation and can put people and goods at risk. For example, if a line is disconnected due to technical reasons, a worker can be electrocuted by assuming that the line does not have voltage.

5.2.3.3 Parallel Operation

When a parallel operation exists, the sources are interconnected and both are connected with the load. If one of the sources fails, the load passes to be instantaneously fed exclusively by the other, without any interruption in the load supply.

The fact of the two sources operating in parallel implies that the DG unit will be in operation and in synchronism with the grid, aggregating the necessary conditions to feed the load at the moment of the grid failure. This kind of operation is more expensive because besides the necessary additional protection and control equipment, there are additional fuel and equipment wear out costs that occur in generating equipment, even without electricity generation.

The parallel operation requires a large quantity of monitoring, control, synchronization, and protection devices. Both the sources must be protected against the failures of the other, including the backfeed phenomenon. This kind of connection is necessary in the cases in which the DG unit owner wants to sell energy to the grid.

Several types of parallel connections are available (Figure 5.15), depending on the DG unit localization and the possibility of selling energy to the grid. In the first configuration, the DG unit operates in parallel with the grid, supplying energy to all the loads or to some loads, particularly providing the utility supplemental or backup power. In this configuration, it is impossible to supply energy to the grid.

In the second configuration, the DG unit operates in parallel with the grid, supplying energy to all the loads. With this configuration, it is possible to supply energy to the grid. The DG unit provides peaking or base-load power to load and exports power to the grid, providing the utility supplemental and backup power. In the third configuration, the DG unit operates in parallel with the grid, supplying energy to the grid and to the consumer. In this configuration, the DG unit does not usually belong to the consumer.

5.2.4 Ancillary Services

Ancillary services is the designation given to a number of functions that are necessary to support the reliable and efficient operations of the power system network. Besides energy (kWh) and capacity (kW), DG can provide other additional benefits, including spinning reserve capacity, peaking, load following, reactive power, and voltage support and other ancillary services.

Generally, a customer can use DG in conjunction with the traditional utility service or as a separate service. There are two ways of DG utilization: to supply power and energy during peak periods or during the entire demand period. Distributed generation equipment can also be used as backup or standby power. Some ancillary services provided by the conventional generators can also be provided by DG, thus minimizing the cost of supplying ancillary services.

In addition to generating energy, DG operation can provide the following benefits:

- Eliminate the need to upgrade the size of feeders
- Improve voltage levels at the feeder ends
- Eliminate the need for capacitor banks
- Provide reactive power compensation
- Eliminate the need for voltage regulators

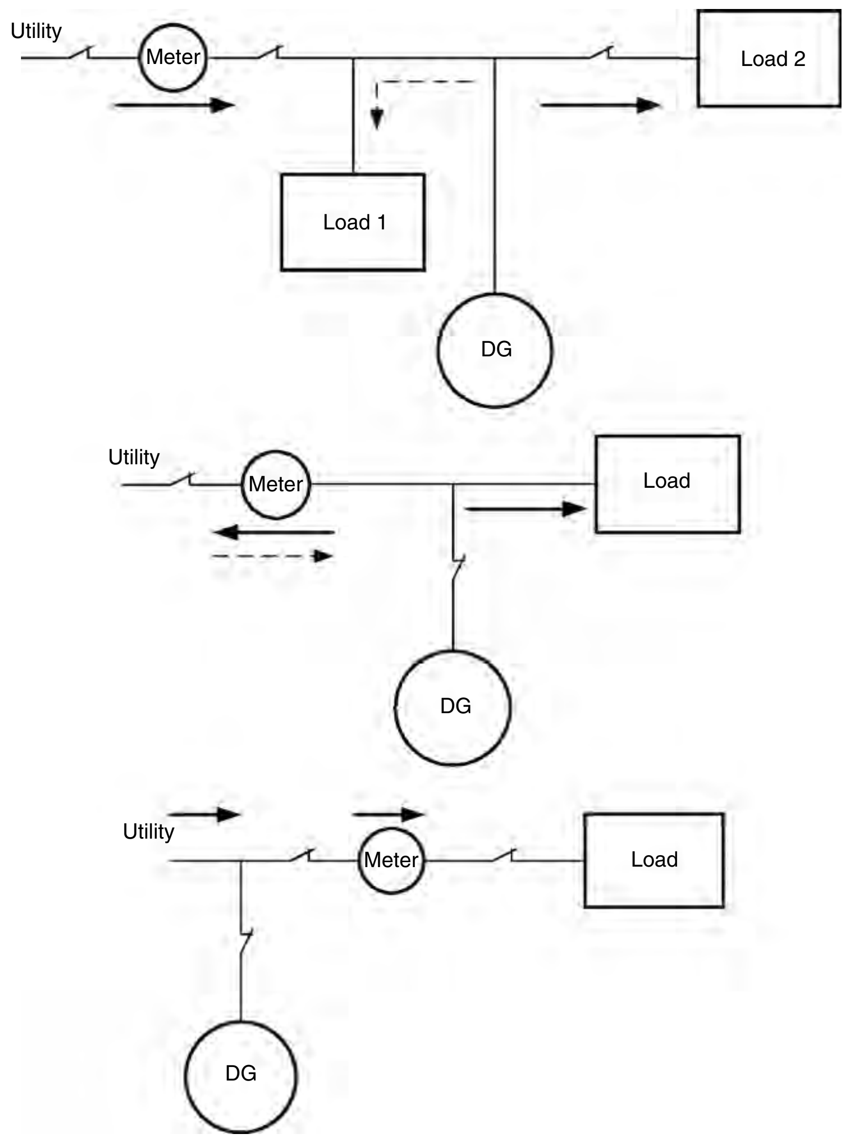


FIGURE 5.15 Parallel operation.

- Reduce feeder loading and delay replacement
- Reduce line losses and transmission system load

The main types of ancillary services, which can be provided by generators include:

Regulation Service and Frequency Response. They provide generation capacity that is available and running, and that can be used to maintain real-time balance in the transmission system. As system loads fluctuate minute-by-minute, generators must be available to match instantly the fluctuations due to the increase or decrease of the loads. An AGC reacts to perceived system fluctuations by adjusting its output to oppose or dampen the fluctuation, whether it is caused by load changes or changes in the output of other bulk system generators as they ramp up or down. Generation units equipped with automated generation control can follow load variations on the time scale of seconds. The load balance is a critical service to the stability of day-to-day grid operation.

The load-following capability service is associated with the "Regulation service and spinning reserves." Load following is the use of available generation capacity to meet the variations in system load. A more detailed description is given below.

Spinning Reserves. This refers to supplemental generation capacity that is ready to quickly ramp up at short notice. Spinning reserves is the incremental generating supply that an active unit can ramp up to within 10 min and then sustain, typically for 30–120 min.

An amount of generating capacity must be kept fully warmed up and ready to take over within seconds in the event of a generator or transmission line failure. The term *spinning* refers to the fact that the generator is on, spinning at rated speed (in the case of turbine generators), and synchronized to the grid. It only needs to adjust its power output to the prescribed level.

Large DG and aggregated small DG alike can provide spinning reserve service. Implicit in the definition, however, is the availability of the capacity to be called upon at any time. Therefore, for example, a DG unit cannot use its full capacity for peak shaving a local load, and at the same time qualify that capacity for spinning reserve.

This limitation is true for nonspinning and replacement reserve services as well. A generator designed to run at 80% of its normal capacity for local purposes can qualify the remaining 20% capacity for spinning reserve, as long as it is synchronized to the grid for the defined reserve period.

Some quick-release hydro units allow a change from zero to full power in 1 min. Alternatively, quick-response loads (using demand response controls) can also contribute to achieve a fast balance between supply and demand.

Supplementary (Nonspinning) Reserves. This refers to generation that is available but not running. Generation kept on standby so that it can be started rapidly in the event that generators or lines suddenly fail, but not as rapidly as spinning reserves above. The incremental generation that can be achieved by units with slower responses, and those requiring start-up, is considered nonspinning reserve. Jurisdictions differ as to the ramping time allowed, varying from 10 to 30 min.

Generators that can start, synchronize, and ramp to full power in short time periods can therefore participate in the quick-response reserve market without running at all times. Fast-start combustion turbines can serve this function. In addition, customers in the form of medium fast-response load may provide nonspinning reserve services.

Nonspinning reserve, in most cases, will be a more appropriate choice over spinning reserve for unused DG capacity. Most distribution level DG technologies do not require 10 min to start-up, and therefore would not gain from remaining synchronized to the grid when not needed. Nonspinning reserves further provide ample opportunity for generators installed as emergency backup systems to participate in the reserve market, where they would not under spinning reserve. These generators are designed to remain off under normal circumstances and serve the customer's load only if the utility experiences an outage; therefore, their capacity during normal utility operation is always available.

Replacement or Operating Reserves. Replacement reserve is the incremental generation that can be obtained in the next hour to replace spinning and nonspinning reserves used in the current hour. Replacement reserve is very similar to nonspinning reserve with the exception that the generator has 60 min to start- and ramp-up instead of only 10 min. Each resource providing replacement reserve must be capable of supplying any level of output up to and including its full reserved capacity within 60 min after issue of dispatch instructions by the independent system operator (ISO). Each resource providing replacement reserve must be capable of sustaining the required output for at least 2 h.

Replacement reserve may be supplied from resources already providing another ancillary service, such as spinning reserve. However, the sum of the ancillary service capacity plus the replacement reserve cannot exceed the capacity of said resource.

Replacement reserve can be provided by large and aggregated small DG that requires more than 10 min (and less than 1 h) to start and ramp to full power. This would be appropriate in cases where the generator technology itself has ramping limitations, or where the generator starting functions are not automated in response to a signal from the ISO, and therefore require delayed manual intervention.

Voltage Support. These services are required to maintain transmission voltage level margins within the criteria in force. Dispatchers at the control center alter the settings on transformers, transmission lines, and other downstream grid-connected equipment, as well as provide sufficient reactive power in areas where needed.

Reactive Power Support. Reactive power support is the injection or absorption of reactive power from generators to maintain transmission-system voltage within required ranges. Generators and loads may be dispatched and operated within a prescribed power factor range to boost the voltage during heavy load periods, or reduce the voltage during light load periods. The service can be provided by generators, loads, and utility distribution companies alike, as long as they have the proper power factor adjustment capabilities.

Black-Start Generation Capability. Black-start generation capability is the ability of a generating unit to go from a shutdown condition to an operating condition without assistance from the electrical grid and to then energize the grid to help other units start after a blackout occurs.

Generators are started in a sequence so that each subsequent generator has an energized bus with which to synchronize. Strategically located black-start generators are a key factor for ensuring timely restoration after a major outage. Each black-start generating unit must be able to start-up with a dead primary and station service bus within 10 min of issue of a dispatch instruction by the ISO requiring a black start.

Each slack-start generating unit must provide sufficient reactive capability to keep the energized transmission bus voltages within emergency voltage limits over the range of no-load to full load.

Each black-start generating unit must be capable of sustaining its output for a minimum period of 12 h from the time when it first starts delivering energy.

The other characteristics that may influence the adoption of DG technologies for ancillary service applications will vary according to the service performed and the ultimate shape of the ancillary service market. Start-up time for all electrical generators is an extremely important parameter to determine if the particular unit can be used as the reserve or can operate in the load following mode.

The part-load capabilities of DG technologies and the start-up time periods of each are presented in Table 5.6.

In addition to their high fuel efficiency, fuel cells appear to offer technical capabilities. Their flexible size enables them to be located close to the load, which can reduce energy losses and transmission and distribution costs. One of the most significant characteristics of the fuel cell is its ability to operate efficiently at part-load, i.e., to respond to sudden increases or decreases in power demands. In addition to meeting changes in power demand, the fuel cell's spinning reverse and load following capabilities enable it to complement effectively the variable output from other renewable power sources, such as solar energy and wind farms.

5.2.5 Advantages of the Grid Interconnection

5.2.5.1 Economical Advantages

The cost of the electricity provided by the grid can be smaller than the cost of the local production in some time periods. Thus, during the periods in which the marginal production cost in the local unit is superior to the grid electricity cost, it makes sense to use the grid energy. During a peak period, the cost of the grid electricity is higher when local production becomes advantageous.

Sometimes is not advantageous to size the DG unit to meet all the required power by the loads. In this situation, when the requested power is higher, the additional required energy is provided by the grid. When the unit is working below the full capacity and the marginal production cost is lesser than the grid price, it is possible to increase the local production thus increasing the profits.

5.2.5.2 Voltage Regulation

The electric power grid is projected to approach an ideal voltage source, with lower internal impedance. In this kind of source, the voltage is the same to all the connected loads, ensuring that the start of a large

TABLE 5.6 Summary Table of Some Performance Characteristics by Distributed Generation Technology Type

Technology	Steam Turbine	Diesel Engine	Natural Gas Engine	Gas Turbine	Microturbine	PAFC	MCFC	SOFC Tubular	SOFC Planar	PEMFC
Part-load	Satisfactory	Good	Satisfactory	Poor	Satisfactory	Satisfactory	Poor			Satisfactory
Start-up time	1 h–1 d	10 s	10 s	10 min–1 h	60 s	1–4 h	More than 10 h	5–10 h	Not available	<0.1 h

Source: From US Environmental Protection Agency, Introduction to CHP Technologies, California Energy Commission, DER Equipment.

load does not disturb the feeding voltage to the other loads. In a nonideal voltage source, the start of a large load (e.g., a large induction motor) causes a momentary voltage sag, affecting the other loads negatively.

In general, the grid achieves a good voltage regulation, because as the power of the load varies, the grid adjusts the power flows automatically, with very small variations in the voltage. Only at some points of the grid, especially toward the end of long lines, this variation is relatively high.

In general, DG units do not have as good a voltage regulation as the grid. As the load varies, the unit controller monitors the output voltage, that tends to decrease with the load increase and increase with the load reduction. When the voltage varies, the unit controller automatically responds, but it is almost impossible to equal the nearly instantaneous response of the grid.

The interconnection, with the DG unit working in parallel with the grid, solves the voltage regulation problem, even in the cases in which all the consumed energy is provided by the DG unit. When the load varies, the grid ensures the transitory instantaneous response, allowing the unit to make a relatively slow change.

5.2.5.3 Reliability

When properly managed, two energy sources work better together than isolated. Either with DG reserve units or with normal operating units, the DG system owner can view the grid as a reserve energy source. In fact, in almost all of the sites, the grid reliability is higher than the reliability of any isolated DG unit. In the DG projects, it is common to use values of 92%–93% for the unit's availability. Considering an availability of 99%, which is extremely difficult to obtain, the corresponding unavailability will be higher than 80 h per year, which is unacceptable to most of the appliances.

Even in areas with poor performance of the electric grid, the grid availability is normally higher than the DG unit, being many times higher than the availability of a DG system, even with several units. The grid utilization like reserve source makes sense in most of the cases, if the charged costs by the grid operator are reasonable.

5.2.6 Disadvantages of the Grid Interconnection

5.2.6.1 Costs with the Grid Operator

In any market, the grid operator will charge a considerable value for interconnection with the grid. The DG unit owner, making the interconnection, will normally pay not only for the energy supplied by the grid, but also a charge dependent upon the maximum power delivered by the grid.

5.2.6.2 Additional Equipment and Maintenance

Distributed generation system operation with interconnection with the grid is more complex than an isolated operation. Besides all the necessary equipment for the system operation, it is necessary to install additional control, metering, and protection devices to isolate the DG system from the grid. Usually the additional equipment, besides increasing the cost, increases the system complexity, thereby increasing the maintenance needs.

5.2.6.3 Increasing Maintenance Needs to Ensure High Reliability Levels

The potential reliability improvement achieved by the interconnection does not appear automatically. It is necessary to have a constrained management of the DG/grid combination to obtain the expected reliability levels. There are three essential points:

- The DG system operating in parallel with the grid needs an exhaustive monitoring of the grid operation conditions of the DG system, the load and the interconnection. The existence of two sources and a load cause a complex control problem that is easily resolved with the modern electronic devices, but is a potential weakness. Problems can occur either because the system is not correctly programmed or due to a failure in a key element of the system that can deactivate the entire system. In critical DG units interconnected with the grid, redundant control equipment is usually installed with auto-monitoring.

- Problems at grid points distant from the DG installation can cause perturbations in the DG unit operation. The most common problems are the atmospheric discharges. Reaching a line, a lightning discharge causes a current impulse that, without the appropriated protection, can reach the DG system. To mitigate this problem, additional protection equipment is needed, increasing the total cost of the system.
- Undesirable events in the grid can disturb the DG unit operation, especially if the control and protection equipment is very sensitive. Momentary failure of a line can cause an abrupt fall in the voltage that can be interpreted by the control system as a risk situation to the DG unit, resulting in its removal from service, or it may be interpreted as an abrupt increase in the load. In the last situation, the control device responds by increasing the power of the DG unit. After the automatic reclosure, the voltage level increases too fast and the DG unit cannot react in a timely manner, while it continues to try follow what it seems a load increase. An overvoltage occurs that is detected by the control system which removes the unit from service, and the installation is then supplied by the grid.

5.2.7 IEEE Standard for Interconnecting Distributed Resources with Electric Power Systems

The IEEE Standards Board approved the IEEE 1547 Standard for Interconnecting Distributed Resources (DR) with Electric Power Systems (EPS) in June, 2003. It was then approved as an American National Standard in October 2003. Many of the technical concerns that the companies of distributed electricity usually raise for the DR interconnection with the grid are related to reliability, security, and quality of service. The IEEE P1547/D07 defines interconnection technical specifications and requirements that are universally needed for interconnection of DR. This standard constitutes an important step to overcome the barriers and increase the development of DG installations.

This American standard establishes criteria and requirements for interconnection of DR with EPS. It provides requirements relevant to the performance, operation, testing, safety considerations, and maintenance of the interconnection. The standard applies to all DR technologies, with aggregate capacity of 10 MVA or smaller at the point of common coupling, and to all EPS at typical primary and/or secondary distribution voltages.

The standard defines interconnection technical specifications and requirements that all interconnection systems and DRs shall meet. General requirements are related to voltage regulation, integration with area electric power system grounding, synchronization, DR on secondary grid and spot networks, inadvertent energization, and reconnection to area EPS, monitoring, and isolation device.

Abnormal conditions can arise on the area EPS that require a response from the connected DR. This response contributes to the safety of utility maintenance personnel and the general public, as well as the avoidance of damage to the connected equipment, including the DR. The abnormal conditions of concern are voltage and frequency excursions above or below the values stated in the IEEE P1547 (Table 5.7), and the isolation of a portion of the area EPS with some DR, presenting the potential for an unintended island.

TABLE 5.7 Interconnection System Response to Abnormal Voltages

Voltage Range (% of Base Voltage)	Clearing Time (s)[a]
$V < 50$	0.16
$50 \leq V < 88$	2
$110 < V < 120$	1
$V \geq 120$	0.16

[a] Clearing time: Time between the start of the abnormal condition and the DR ceasing to energize the area EPS.

TABLE 5.8 Maximum Harmonic Current Distortion in Percent of Current (I)

Individual Harmonic Order (Odd Harmonics)	<11	$11 \leq h < 17$	$17 \leq h < 23$	$23 \leq h < 35$	$35 \leq h$	Total Demand Distortion (TDD)
Percent (%)	4.0	2.0	1.5	0.6	0.3	5.0

I is the greater of the local EPS maximum load current integrated demand (15–30 min) without the DR unit, or the DR unit rated current capacity (transformed to the PCC when a transformer exists between the DR unit and the PCC); even harmonics are limited to 25% of the odd harmonic limits above.

Power quality issues are also addressed in the IEEE P1547, namely, limitation of DC injection, voltage flicker induced by the DR, harmonic current injection (Table 5.8), immunity protection and surge capabilities, as well as the islanding considerations.

The standard also provides test requirements for an interconnection system to demonstrate that it meets all the requirements. The following tests are required for all interconnection systems:

- Interconnection test
- Production tests
- Interconnection installation evaluation
- Commissioning tests
- Periodic interconnection tests

IEEE 1547 is the first in a family of IEEE interconnection standards for DR (Figure 5.16). Other standards in the family currently underway are:

- IEEE P1547.1 Draft Standard for Conformance Tests Procedures for Equipment Interconnecting DR with EPS. This standard specifies the type, production, and commissioning tests that shall be performed to demonstrate that the interconnection functions and equipment of a distributed resource conform to IEEE Standard P1547.
- IEEE P1547.2 Draft Application Guide for IEEE 1547 Standard for Interconnecting DR with EPS. This guide provides technical background and application details to support the understanding of IEEE 1547 Standard for Interconnecting DR with EPS.
- IEEE P1547.3 Draft Guide For Monitoring, Information Exchange, and Control of DR Interconnected with EPS. This document provides guidelines for monitoring, information exchange, and control for DR interconnected with EPS.
- IEEE P1547.4 Draft Guide for Design, Operation, and Integration of Distributed Resource Island Systems with EPS. This document provides alternative approaches and good practices for the design, operation, and integration of distributed resource island systems with EPS.
- IEEE P1547.5 Draft Technical Guidelines for Interconnection of Electric Power Sources Greater than 10 MVA to the Power Transmission Grid. This document provides guidelines regarding the technical requirements, including design, construction, commissioning acceptance testing, and maintenance/performance requirements, for interconnecting dispatchable electric power sources with a capacity of more than 10 MVA to a bulk power transmission grid.
- IEEE P1561 Draft Guide for Sizing Hybrid Stand-Alone Energy Systems. This guide provides the rationale and guidance for operating lead-acid batteries in remote hybrid systems considering the system's load, and the capacities of its renewable-energy generator(s), dispatchable generator(s), and battery(s).

5.2.8 Power Quality Applications

Power quality-related issues are currently of great concern. The widespread use of electronic equipment, such as information technology equipment, power electronics such as adjustable speed drives (ASDs), programmable logic controllers (PLCs), and energy-efficient lighting led to a complete change of electric

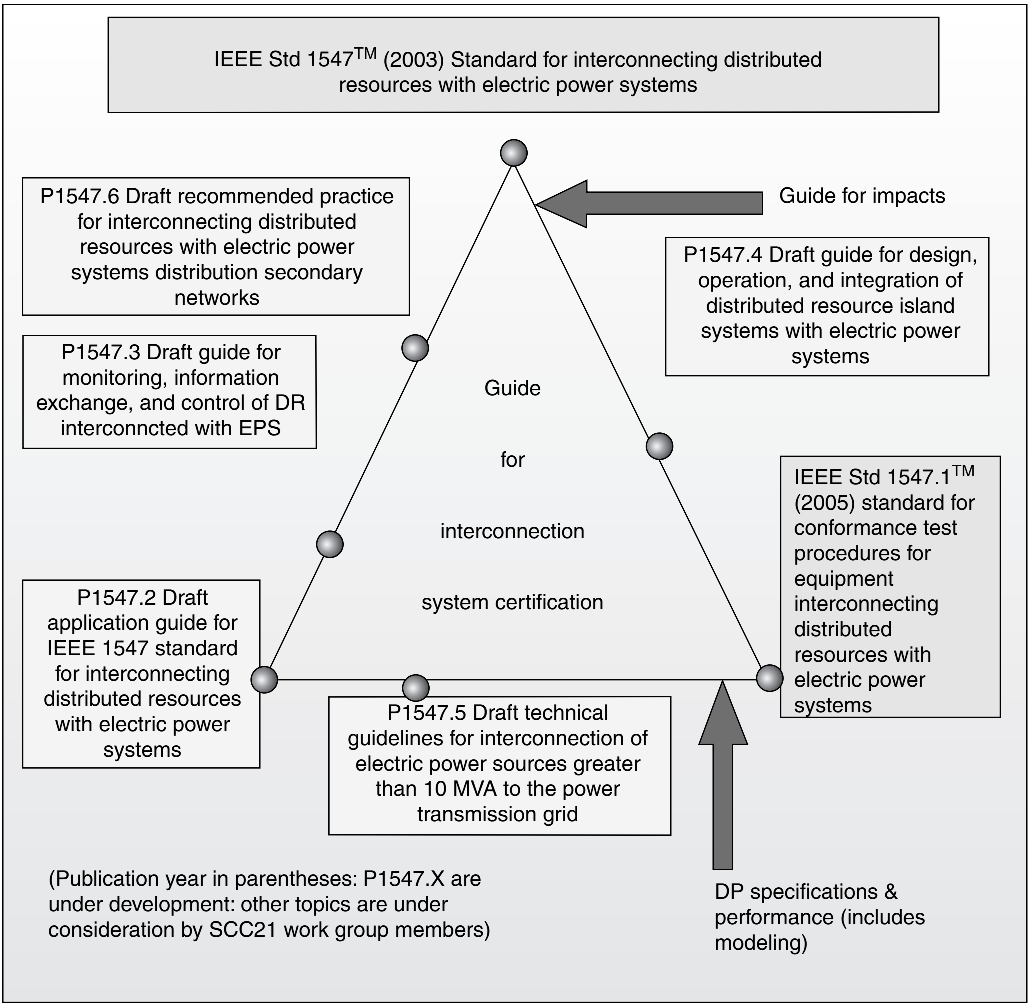


FIGURE 5.16 IEEE SCC21 1547 Series of Interconnection Standards. (From Institute of Electric and Electronics Engineers (IEEE), IEEE P1547/D07. *Standard for Interconnecting Distributed Resources with Electric Power Systems*. IEEE Standards Coordinating Committee 21 (IEEE SCC21) on Fuel Cells, Photovoltaics, Dispersed Generation, and Energy Storage of the IEEE Standards Association, New York, 2001. With permission.)

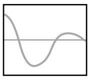
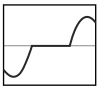
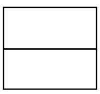
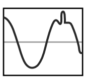
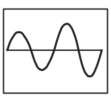
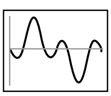
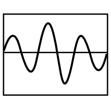
loads nature. These loads are simultaneously the major causers and the major victims of PQ problems. Due to their nonlinearity, all these loads cause disturbances in the voltage waveform.

Along with technology advance, the organization of the worldwide economy has evolved towards globalization and the profit margins of many activities tend to decrease. The increased sensitivity of the vast majority of processes (industrial, services, and even residential) to PQ problems turns the availability of electric power with quality a crucial factor for competitiveness in every activity sector. The most critical areas are the continuous process industry and the information technology services. When a disturbance occurs, huge financial losses may occur, with the consequent loss of productivity and competitiveness.

Although many efforts have been taken by utilities, some consumers require a level of PQ higher than the level provided by modern electric networks. This implies that some measures must be taken in order to achieve higher levels of PQ.

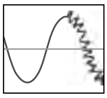
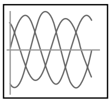
The most common types of PQ problems are presented in Table 5.9.

TABLE 5.9 Most Common Power Quality Problems

1. Voltage sag (or dip) 	*Description*: A decrease of the normal voltage level between 10 and 90% of the nominal rms voltage at the power frequency, for durations of 0,5 cycle to 1 min *Causes*: Faults in the transmission or distribution network (most of the times on parallel feeders). Faults in consumer's installation. Connection of heavy loads and start-up of large motors *Consequences*: Malfunction of information technology equipment, namely microprocessor-based control systems (PCs, programmable logic controllers (PLCs), adjustable speed drives (ASDs), etc.) that may lead to a process stoppage. Tripping of contactors and electromechanical relays. Disconnection and loss of efficiency in electric rotating machines
2. Very short interruptions 	*Description*: Total interruption of electrical supply for duration from few milliseconds to one or two seconds *Causes*: Mainly due to the opening and automatic reclosure of protection devices to decommission a faulty section of the network. The main fault causes are insulation failure, lightning, and insulator flashover *Consequences*: Tripping of protection devices, loss of information, and malfunction of data processing equipment. Stoppage of sensitive equipment, such as ASDs, PCs, PLCs, if they are not prepared to deal with this situation
3. Long interruptions 	*Description*: Total interruption of electrical supply for duration greater than 1–2 s *Causes*: Equipment failure in the power system network, storms, and objects (trees, cars, etc.) striking lines or poles, fire, human error, bad coordination or failure of protection devices *Consequences*: Stoppage of all equipment
4. Voltage spike 	*Description*: Very fast variation of the voltage value for durations from several microseconds to few milliseconds. These variations may reach thousands of volts, even in low voltage *Causes*: Lightning, switching of lines or power factor correction capacitors, disconnection of heavy loads *Consequences*: Destruction of components (particularly electronic components) and of insulation materials, data processing errors or data loss, electromagnetic interference
5. Voltage swell 	*Description*: Momentary increase of the voltage, at the power frequency, outside the normal tolerances, with duration of more than one cycle and typically less than a few seconds *Causes*: Start/stop of heavy loads, badly dimensioned power sources, badly regulated transformers (mainly during off-peak hours) *Consequences*: Data loss, flickering of lighting and screens, stoppage or damage of sensitive equipment, if the voltage values are too high
6. Harmonic distortion 	*Description*: Voltage or current waveforms assume nonsinusoidal shape. The waveform corresponds to the sum of different sine-waves with different magnitudes and phases, having frequencies that are multiples of power-system frequency *Causes*: *Classic sources*: electric machines working above the knee of the magnetization curve (magnetic saturation), arc furnaces, welding machines, rectifiers, and DC brush motors. *Modern sources*: all nonlinear loads, such as power electronics equipment including ASDs, switched mode power supplies, data processing equipment, high efficiency lighting *Consequences*: Increased probability of occurrence of resonance, neutral overload in three-phase systems, overheating of all cables and equipment, loss of efficiency in electric machines, electromagnetic interference with communication systems, errors in measures when using average reading meters, nuisance tripping of thermal protections
7. Voltage fluctuation 	*Description*: Oscillation of voltage value, amplitude modulated by a signal with frequency of 0–30 Hz *Causes*: Arc furnaces, frequent start/stop of electric motors (for instance elevators), oscillating loads *Consequences*: Most consequences are common to undervoltages. The most perceptible consequence is the flickering of lighting and screens, giving the impression of unsteadiness of visual perception

(continued)

TABLE 5.9 **(*Continued*)**

8. Noise 	*Description*: Superimposition of high frequency signals on the waveform of the power-system frequency *Causes*: Electromagnetic interferences provoked by Hertzian waves, such as microwaves, television diffusion, and radiation due to welding machines, arc furnaces, and electronic equipment. Improper grounding may also be a cause *Consequences*: Disturbances on sensitive electronic equipment, usually not destructive. May cause data loss and data processing errors
9. Voltage unbalance 	*Description*: A voltage variation in a three-phase system in which the three voltage magnitudes or the phase-angle differences between them are not equal *Causes*: Large single-phase loads (induction furnaces, traction loads), incorrect distribution of all single-phase loads by the three phases of the system (this may be also due to a fault) *Consequences*: Unbalanced systems imply the existence of a negative sequence that is harmful to all three-phase loads. The most affected loads are three-phase induction machines

Even the most advanced transmission and distribution systems are not able to provide electrical energy with the desired level of reliability for the proper functioning of the loads in the modern society. Modern T&D (transmission and distribution) systems are projected for 99.9%–99.99% availability. This value is highly dependant of redundancy level of the network, which is different according to the geographical location and the voltage level (availability is higher at the HV network). In some remote sites, availability of T&D systems may be as low as 99%. Even with a 99.99% level, there is an equivalent interruption time of 52 min per year. The most demanding processes in the modern digital economy need electrical energy with 99.9999999% availability (9-nines reliability) to function properly.

The mitigation of PQ problems may take place at different levels: transmission, distribution, and the end use equipment. As seen in Figure 5.17, several measures can be taken at these levels. Many PQ problems have origin in the transmission or distribution grid. Thus, a proper transmission and distribution grid, with adequate planning and maintenance, is essential to minimize the occurrence of PQ problems.

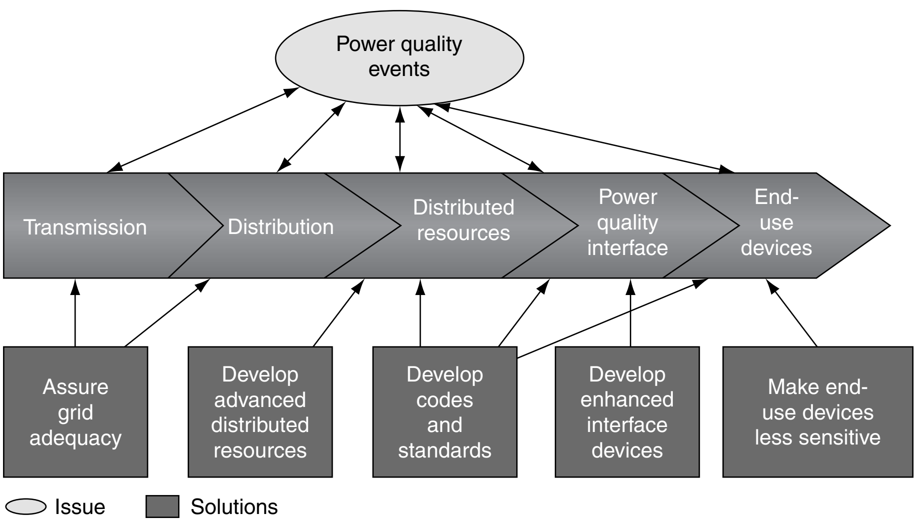


FIGURE 5.17 Solutions for digital power.

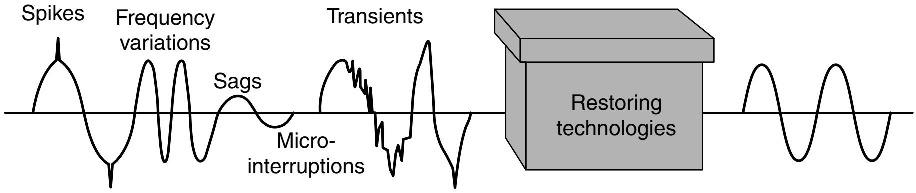


FIGURE 5.18 Restoring technologies principle.

Interest in the use of distributed energy resources has increased substantially over the last few years because of their potential to provide increased reliability. These resources include DG and energy storage systems. Energy storage systems, also known as restoring technologies, are used to provide the electric loads with ride-through capability in poor PQ environment. Recent technological advances in power electronics and storage technologies are turning the restoring technologies as one of the premium solutions to mitigate PQ problems (Figure 5.18).

Distributed generation units can be used to provide clean power to critical loads, isolating them from disturbances with origin in the grid. Distributed generation units can also be used as backup generators to assure energy supply to critical loads during sustained outages. Additionally, DG units can be used for load management purposed to decrease the peak demand.

At present, the reciprocating engine is the prevalent technology in DG market, but with technology advancements, other technologies are becoming more attractive, such as photovoltaics, microturbines, or fuel cells.

If DG units are to be used as backup generation, a storage unit must be used to provide energy to the loads during the period between the origin of the disturbance and the start-up of the emergency generator.

The most common solution is the combination of electrochemical batteries UPS and a diesel genset. At present, the integration of a flywheel and a diesel genset in a single unit (Figure 5.19 and Figure 5.20) is also becoming a popular solution, offered by many manufacturers.

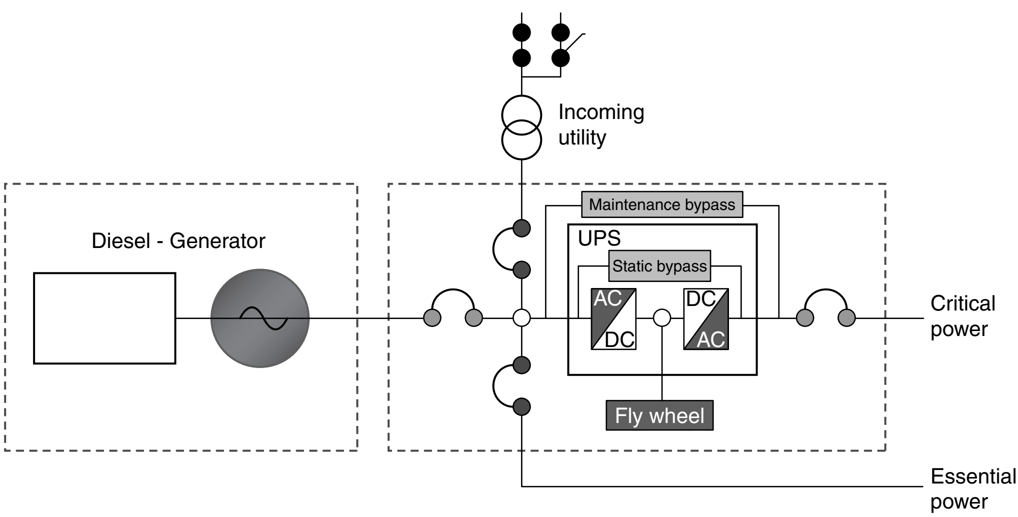


FIGURE 5.19 Scheme of a continuous power system, using a flywheel and a diesel genset. (From http://www.geindustrial.com. With permission.)

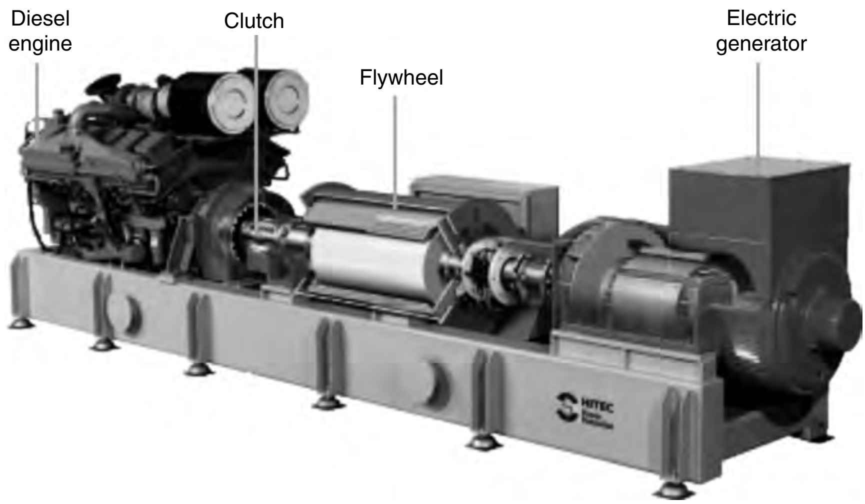


FIGURE 5.20 Dynamic UPS, by Hitec Power Protection. (From http://www.hitecups.com/?RubriekID=1991. With permission.)

5.3 Demand-Side Management

Clark W. Gellings and Kelly E. Parmenter

5.3.1 Introduction

Since the mid-1980s, demand-side management has been an important element of the electric utility planning approach referred to as "integrated resource planning." At that time, annual demand-side management expenditures in the U.S. were measured in billions of dollars, energy savings were measured in billions of kilowatts hours, and peak load reductions were stated in thousands of megawatts. Although activities nationally have slowed since then, there are a number of instances where demand-side management continued to influence the demand for electricity. This article defines demand-side management, describes the role demand-side management plays in integrated resource planning, and discusses the main elements of demand-side management programs. It then presents case studies of four successful demand-side management programs that were offered between 2001 and 2003.

5.3.2 What is Demand-Side Management?

The term *demand-side management* is the result of a logical evolution of planning processes used by utilities in the late 1980s. One of the first terms, *demand-side load management* was introduced by the author, Clark W. Gellings, in an article for IEEE's *Spectrum* in 1981. Shortly after the publication of this article, at a meeting of The Edison Electric Institute (EEI) Customer Service and Marketing Executives in 1982, Mr. Gellings altered the term to *demand-side planning*. This change was made to reflect the broader objectives of the planning process. Mr. Gellings coined the term *demand-side management* and continued to popularize the term throughout a series of more than 100 articles since that time, including the five-volume set *Demand-Side Management* that is widely recognized as a definitive and practical source of information on the demand-side management process.

Perhaps the most widely accepted definition of demand-side management is the following: "Demand-side management is the planning, implementation, and monitoring of those utility activities designed to influence customer use of electricity in ways that will produce desired changes in the utility's load shape, i.e., changes in the time pattern and magnitude of a utility's load. Utility programs falling under the umbrella of demand-side management include: load management, new uses, strategic conservation, electrification, customer generation, and adjustments in market share" (Gellings 1984–1988). However, demand-side management is even more encompassing than this definition implies because it includes the management of all forms of energy at the demand-side, not just electricity. In addition, groups other than just electric utilities (including natural gas suppliers, government organizations, nonprofit groups, and private parties) implement demand-side management programs.

In general, demand-side management embraces the following critical components of energy planning:

1. Demand-side management will influence customer use. Any program intended to influence the customer's use of energy is considered demand-side management.
2. Demand-side management must achieve selected objectives. To constitute a "desired load shape change," the program must further the achievement of selected objectives, i.e., it must result in reductions in average rates, improvements in customer satisfaction, achievement of reliability targets, etc.
3. Demand-side management will be evaluated against non-demand-side management alternatives. The concept also requires that selected demand-side management programs further these objectives to at least as great an extent as non-demand-side management alternatives, such as generating units, purchased power or supply-side storage devices. In other words, it requires that demand-side management alternatives be compared to supply-side alternatives. It is at this stage of evaluation that demand-side management becomes part of the integrated resource planning process.
4. Demand-side management identifies how customers will respond. Demand-side management is pragmatically oriented. Normative programs ("we ought to do this") do not bring about the desired result; positive efforts ("if we do this, that will happen") are required. Thus, demand-side management encompasses a process that identifies how customers will respond not how they should respond.
5. Demand-side management value is influenced by load shape. Finally, this definition of demand-side management focuses upon the load shape. This implies an evaluation process that examines the value of programs according to how they influence costs and benefits throughout the day, week, month, and year.

Subsets of these activities have been referred to in the past as "load management," "strategic conservation," and "marketing."

5.3.3 Demand-Side Management and Integrated Resource Planning

A very important part of the demand-side management process involves the consistent evaluation of demand-side to supply-side alternatives and vice versa. This approach is referred to as "integrated resource planning." Figure 5.21 illustrates how demand-side management fits into the integrated resource planning process. For demand-side management to be a viable resource option, it has to compete with traditional supply-side options.

5.3.4 Demand-Side Management Programs

A variety of programs have been implemented since the introduction of demand-side management in the early 1980s. Mr. Gellings and EPRI have been instrumental in defining a framework for utilities and other implementers to follow when planning demand-side management programs. This section describes the main elements of the demand-side management planning framework. It then discusses the types of end use sectors, buildings, and end use technologies targeted during program development. It also lists the various entities typically responsible for implementing programs, along with several program implementation methods. Lastly, this section summarizes several representative demand-side management programs offered in the US.

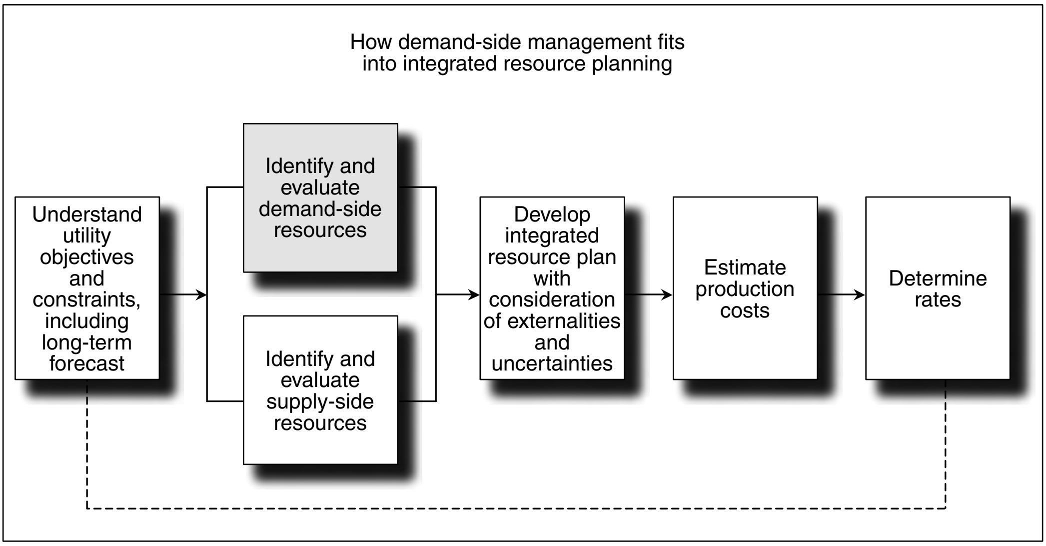


FIGURE 5.21 How demand-side management fits into integrated resources planning.

5.3.4.1 Elements of the Demand-Side Management Planning Framework

Figure 5.22 illustrates the five main elements of the demand-side management planning framework. These five elements are summarized as follows:

1. Set objectives. The first step in demand-side management planning is to establish overall organizational objectives. These strategic objectives are quite broad and generally include examples, such as reducing energy needs, reducing dependence on foreign imports, improving cash flow, increasing earnings, or improving customer and employee relations. The second level of the formal planning process is to operationalize broad objectives to guide policymakers to specific actions. It is at this operational level or tactical level that demand-side management alternatives should be examined and evaluated. For example, an examination of capital investment requirements may show periods of high investment needs. Postponing the need for new construction through a demand-side management program may reduce investment needs and stabilize the financial future of an energy company, or a utility and its state or country. Specific operational objectives are established on the basis of the conditions of the existing energy system—its system configuration, cash reserves, operating environment, and competition. Once designated, operational objectives are translated into desired demand-pattern changes or load-shape changes that can be used to characterize the potential impact of alternative demand-side management programs. Although there is an infinite combination of load-shape-changing possibilities, six have been illustrated in Figure 5.23 to show the range of possibilities, namely peak clipping, valley filling, load shifting, strategic conservation, strategic load growth, and flexible load shape. These six are not mutually exclusive, and may frequently be employed in combinations.
2. Identify alternatives. The second step is to identify alternatives. The first dimension of this step involves identifying the appropriate end uses whose peak load and energy consumption characteristics generally match the requirements of the load-shape objectives established in the previous step. In general, each end use (e.g., residential space heating, commercial lighting) exhibits typical and predictable demand or load patterns. The extent to which load pattern modification can be accommodated by a given end use is one factor used to select an end use for demand-side management. The second dimension of demand-side management alternatives involves choosing appropriate technology alternatives for each target end use. This process

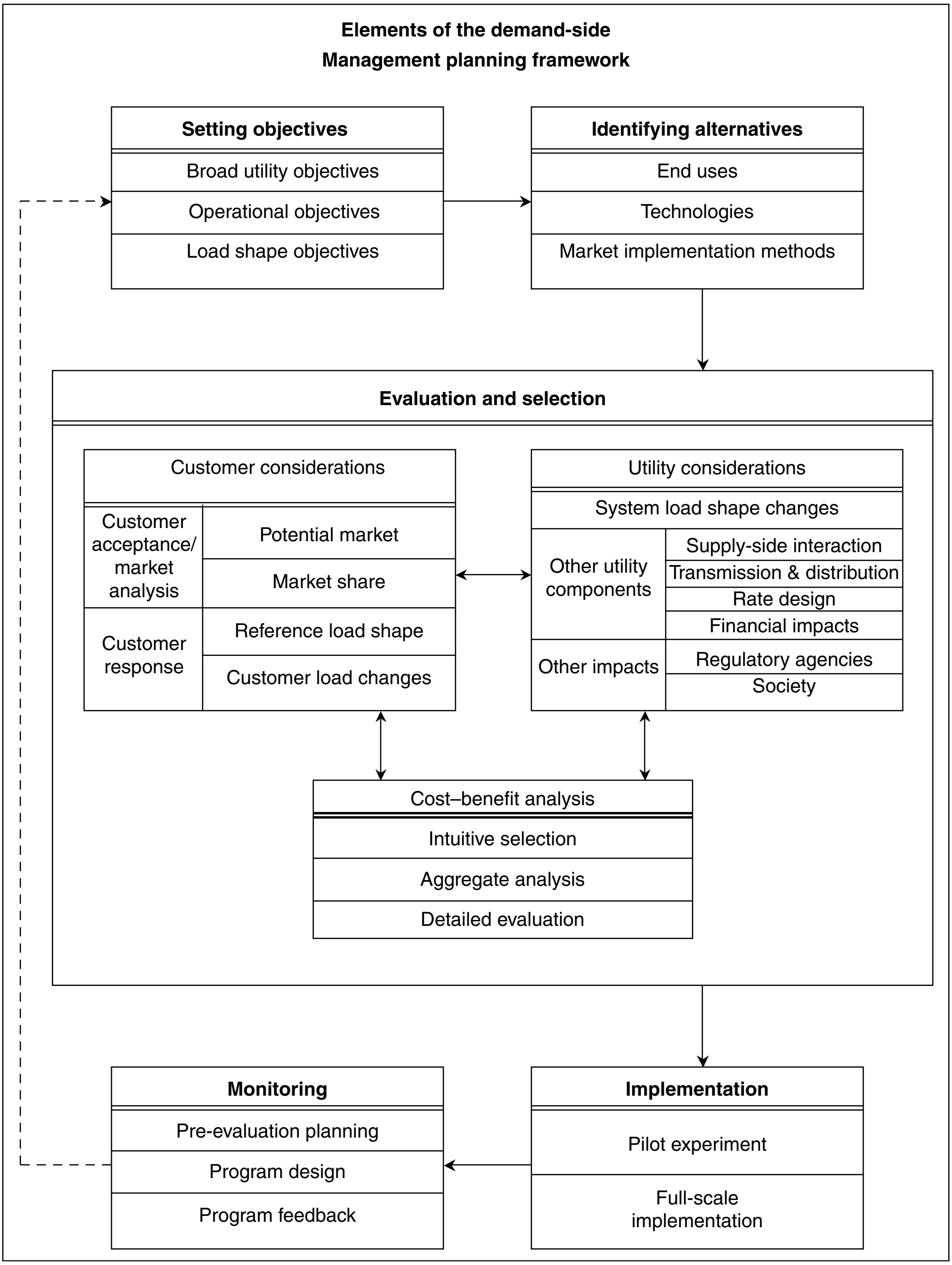


FIGURE 5.22 Elements of the demand-side management planning framework.

should consider the suitability of the technology for satisfying the load-shape objective. Even though a technology is suitable for a given end use, it may not produce the desired results. For example, although water-heater wraps are appropriate for reducing domestic water-heating energy consumption, they are not appropriate for load shifting. In this case, an option such as electric

Peak Clipping - Or the reduction of the system peak loads, embodies one of the classic forms of load management. Peak clipping is generally considered as the reduction of peak load by using direct load control. Direct load control is most commonly practiced by direct utility control of either service to customer facilities or of customers' appliances. While many utilities consider this as means to reduce peaking capacity or capacity purchases and consider control only during the most probable days of system peak, direct load control can be used to reduce operating cost and dependence on critical fuels by economic dispatch.

Valley Filling - Is the second classic form of load management and applies to both gas and electric systems. Valley filling encompasses building off-peak loads. This may be particularly desirable where the long-run incremental cost is less than the average price of energy. Adding properly priced off-peak load under those circumstances decreases the average price. Valley filling can be accomplished in several ways, one of the most popular of which is new thermal energy storage (water heating and/or space heating) that displaces loads served by fossil fuels.

Load Shifting - Is the last classic form of load management and also applies to both gas and electric systems. This involves shifting load from on-peak to off-peak periods. Popular applications include use of storage water heating, storage space heating, coolness storage, and customer load shifts. The load shift from storage devices involves displacing what would have been conventional appliances.

Strategic Conservation - Is the load shape change that results from programs directed at end use consumption. Not normally considered load management, the change reflects a modification of the load shape involving a reduction in consumption as well as a change in the pattern of use. In employing energy conservation, the planner must consider what conservation actions would occur naturally and then evaluate the cost-effectiveness of possible intended programs to accelerate or stimulate those actions. Examples include weatherization and appliance efficiency improvement.

Strategic Load Growth - Is the load shape change that refers to a general increase in sales beyond the valley filling described previously. Load growth may involve increased market share of loads that are or can be, served by competing fuels, as well as economic development. Load growth may include electrification. Electrification is the term being employed to describe the new emerging electric technologies surrounding electric vehicles, industrial process heating, and automation. These have a potential for increasing the electric energy intensity of the industrial sector. This rise in intensity may be motivated by reduction in the use of fossil fuels and raw materials resulting in improved overall productivity.

Flexible Load Shape - Is a concept related to electric system reliability, a planning constraint. Once the anticipated load shape, including demand-side activities, is forecast over some horizon, the power supply planner studies the final optimum supply-side options. Among the many criteria he uses is reliability. Load shape can be flexible – if customers are presented with options as to the variations in quality of service that they are willing to allow in exchange for various incentives. The program involved can be variations of interruptible or curtailable load; concepts of pooled, integrated energy management systems; or individual customer load control devices offering service constraints.

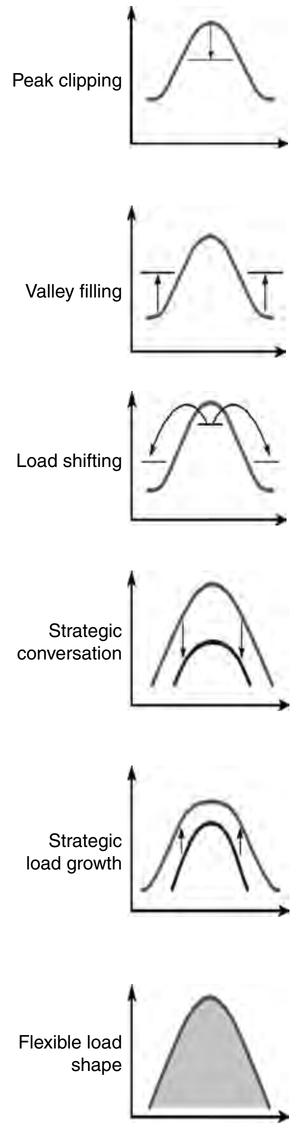


FIGURE 5.23 Six generic load shape objectives that can be considered during demand-side management planning.

water-heating direct load control via receiver/switches would be a better choice. The third dimension involves investigating market implementation methods (see Section 5.3.4.5 for a description of potential implementation methods).

3. Evaluate and select program(s). The third step balances customer considerations, supplier considerations, and cost–benefit analyses to identify the most viable demand-side management alternative(s) to pursue. Although customers and suppliers act independently to alter the pattern of demand, the concept of demand-side management implies a supplier/customer relationship that produces mutually beneficial results. To achieve that mutual benefit, suppliers must carefully

consider such factors as the manner in which the activity will affect the patterns and amount of demand (load shape), the methods available for obtaining customer participation, and the likely magnitudes of costs and benefits to both supplier and customer prior to attempting implementation.

4. Implement program(s). The fourth step, which takes place in several stages, is to implement the program(s). As a first step, a high level, demand-side management project team should be created with representation from the various departments and organizations, and with the overall control and responsibility for the implementation process. It is important for implementers to establish clear directives for the project team, including a written scope of responsibility, project team goals and time frame. When limited information is available on prior demand-side management program experiences, a pilot experiment may precede the program. Pilot experiments can be a useful interim step toward making a decision to undertake a major program. Pilot experiments may be limited either to a subregion or to a sample of consumers throughout an area. If the pilot experiment proves cost-effective, then the implementers may consider initiating the full-scale program.
5. Monitor program(s). The fifth step is to monitor the program(s). The ultimate goal of the monitoring process is to identify deviations from expected performance and to improve both existing and planned demand-side management programs. Monitoring and evaluation processes can also serve as a primary source of information on customer behavior and system impacts, foster advanced planning and organization within a demand-side management program, and provide management with the means of examining demand-side management programs as they develop.

5.3.4.2 Targeted End Use Sectors/Building Types

The three broad categories of end use sectors targeted for demand-side management programs are residential, commercial, and industrial. Each of these broad categories includes several subsectors. In some cases, the program will be designed for one or more broad sectors; in other cases, it may be designed for a specific subsector. For example, the residential sector can be divided into several subsectors including single family homes, multi-family homes, mobile homes, low income homes, etc. In addition, the commercial sector can be split into subsets, such as offices, restaurants, healthcare facilities, educational facilities, retail stores, grocery stores, hotels/motels, etc. There are also numerous specific industrial end users that may be potentially targeted for a demand-side management program. Moreover, the program designer may want to target a specific type or size of building within the chosen sector. The program could focus on new construction, old construction, renovations and retrofits, large customers, small customers, or a combination. Crosscutting programs target multiple end use sectors and/or multiple building types. Figure 5.24 illustrates the broad types of end use sectors and building types and how they relate to other aspects of demand-side management program planning.

5.3.4.3 Targeted End Use Technologies/Program Types

There are several end use technologies or program types targeted in demand-side management programs. (See Figure 5.24 for representative end use technologies and program types.) Some programs are comprehensive, and crossover between end use technologies (e.g., see case study 1 below). Other technologies target-specific end use equipment such as lighting, air conditioners, dishwashers, etc. Still others target load control measures, such as those that shift loads to off-peak hours (e.g., thermal energy storage). Figure 5.24 shows representative end use technologies or program types and how they relate to other aspects of demand-side management program planning.

5.3.4.4 Program Implementers

Implementers of demand-side management programs are often utilities. However, other possible implementers include government organizations, nonprofit groups, private parties, or a collaboration of several entities (see Figure 5.24). Utilities and governments, in particular, have a special interest in influencing customers' demand—treating it not as fate but as choice—to provide better service at lower

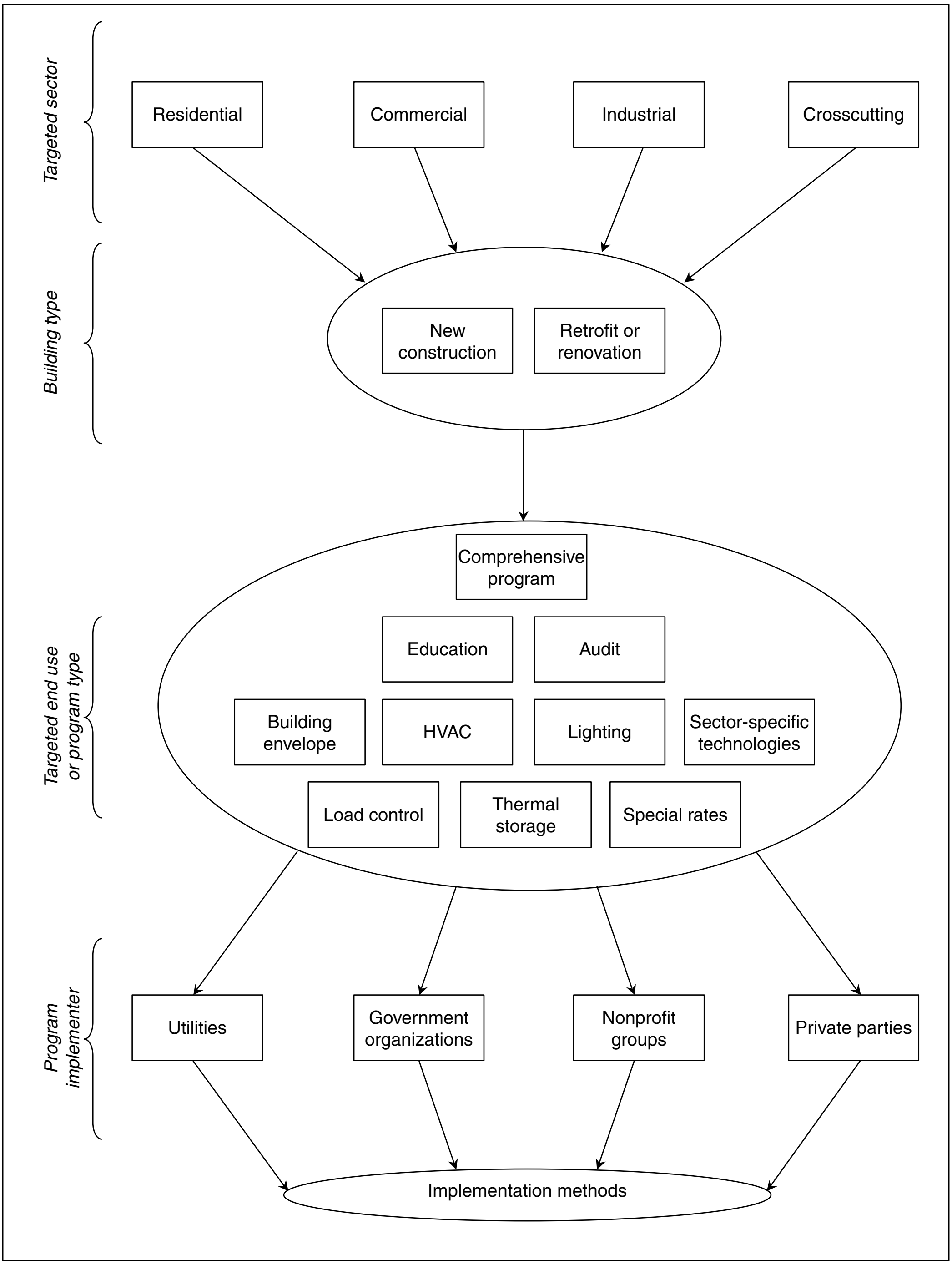


FIGURE 5.24 Relationship between end use sectors, building types, end use programs, and program implementers.

cost while increasing their own profits and reducing their business risks. Energy planners can choose from a wide range of market push and pull methods designed to influence consumer adoption and reduce barriers, as discussed in the next paragraph.

5.3.4.5 Implementation Methods

Among the most important dimension in the characterization of demand-side alternatives is the selection of the appropriate market implementation methods. Planners and policy makers can select from a variety of methods for influencing customer adoption and acceptance of demand-side management programs. The methods can be broadly classified into six categories. Table 5.10 lists examples for each category of market implementation method. The categories include:

1. Customer education. Many energy suppliers and governments have relied on some form of customer education to promote general customer awareness of programs. Brochures, bill inserts,

TABLE 5.10 Examples of Market Implementation Methods

Market Implementation Method	Illustrative Objective	Examples
Customer education	Increase perceived value of energy services Increase customer awareness of programs	Bill inserts Brochures Information packets Displays Clearinghouses Direct mailings
Direct customer contact	Through face-to-face communication, encourage greater customer acceptance, and response to programs	Energy audits Direct installation Store fronts Workshops/energy clinics Exhibits/displays Inspection services
Trade ally cooperation (i.e., architects, engineers, appliance dealers, heating/cooling contractors)	Increase capability in marketing and implementing programs Obtain support and technical advice on customer adoption of demand-side technologies	Cooperative advertising and marketing Training Certification Selected product sales/service
Advertising and promotion	Increase public awareness of new programs Influence customer response	Mass media (radio, TV, and newspaper) Point-of-purchase advertising
Alternative pricing	Provide customers with pricing signals that reflect real economic costs and encourage the desired market response	Demand rates Time-of-use rates Off-peak rates Seasonal rates Inverted rates Variable levels of service Promotional rates Conservation rates
Direct incentives	Reduce up-front purchase price and risk of demand-side technologies to the customer Increase short-term market penetration Provide incentives to employees to promote demand-side management programs	Low- or no-interest loan Cash grants Subsidized installation/modification Rebates Buyback programs Rewards to employees for successful marketing of demand-side management programs

information packets, clearinghouses, educational curricula, and direct mailings are widely used. Customer education is the most basic of the market implementation methods available and should be used in conjunction with one or more other market implementation method for maximum effectiveness.

2. Direct customer contact. Direct customer contact techniques refer to face-to-face communication between the customer and an energy supplier or government representative to encourage greater customer acceptance of programs. Energy suppliers have for some time employed marketing and customer service representatives to provide advice on appliance choice and operation, sizing of heating/cooling systems, lighting design, and even home economics. Direct customer contact can be accomplished through energy audits, specific program services (e.g., equipment servicing), store fronts where information and devices are displayed, workshops, exhibits, onsite inspection, etc. A major advantage of these methods is that they allow the implementer to obtain feedback from the consumer, thus providing an opportunity to identify and respond to major customer concerns. They also enable more personalized marketing, and can be useful in communicating interest in and concern for controlling energy costs.
3. Trade ally cooperation. Trade ally cooperation and support can contribute significantly to the success of many demand-side management programs. A trade ally is defined as any organization that can influence the transactions between the supplier and its customers or between implementers and consumers. Key trade ally groups include home builders and contractors, local chapters of professional societies, technology/product trade groups, trade associations, and associations representing wholesalers and retailers of appliances and energy consuming devices. Depending on the type of trade ally organization, a wide range of services are performed, including development of standards and procedures, technology transfer, training, certification, marketing/sales, installation, maintenance, and repair. Generally, if trade ally groups believe that demand-side management programs will help them (or at least not hinder their business), they will likely support the program.
4. Advertising and promotion. Energy suppliers and government energy entities have used a variety of advertising and promotional techniques. Advertising uses various media to communicate a message to customers in order to inform or persuade them. Advertising media applicable to demand-side management programs include radio, television, magazines, newspapers, outdoor advertising, and point-of-purchase advertising. Promotion usually includes activities to support advertising, such as press releases, personal selling, displays, demonstrations, coupons, and contest/awards. Some prefer the use of newspapers based on consumer research that found this medium to be the major source of customer awareness of demand-side management programs. Others have found television advertising to be more effective.
5. Alternative pricing. Pricing as a market-influencing factor generally performs three functions: (1) transfers to producers and consumers information regarding the cost or value of products and services being provided, (2) provides incentives to use the most efficient production and consumption methods, and (3) determines who can afford how much of a product. These three functions are closely interrelated. Alternative pricing, through innovative schemes can be an important implementation technique for utilities promoting demand-side options. For example, rate incentives for encouraging specific patterns of utilization of electricity can often be combined with other strategies (e.g., direct incentives) to achieve electric utility demand-side management goals. Pricing structures include time-of-use rates, inverted rates, seasonal rates, variable service levels, promotional rates, off-peak rates, etc., A major advantage of alternative pricing programs over some other types of implementation techniques is that the supplier has little or no cash outlay. The customer receives a financial incentive, but over a period of years, so that the implementer can provide the incentives as it receives the benefits.
6. Direct incentives. Direct incentives are used to increase short-term market penetration of a cost control/customer option by reducing the net cash outlay required for equipment purchase or by reducing the payback period (i.e., increasing the rate of return) to make the investment more

attractive. Incentives also reduce customer resistance to options without proven performance histories or options that involve extensive modifications to the building or the customer's lifestyle. Direct incentives include cash grants, rebates, buyback programs, billing credits, and low-interest or no-interest loans. One additional type of direct incentive is the offer of free, or very heavily, subsidized, equipment installation or maintenance in exchange for participation. Such arrangements may cost the supplier more than the direct benefits from the energy or demand impact, but can expedite customer recruitment and allow the collection of valuable empirical performance data.

Energy suppliers, utilities, and government entities have successfully used many of these marketing strategies. Typically, multiple marketing methods are used to promote demand-side management programs. The selection of the individual market implementation method or mix of methods depends on a number of factors, including:

- Prior experience with similar programs
- Existing market penetration
- The receptivity of policy makers and regulatory authorities
- The estimated program benefits and costs to suppliers and customers
- Stage of buyer readiness
- Barriers to implementation

Some of the most innovative demand-side marketing programs started as pilot programs to gauge consumer acceptance and evaluate program design prior to large-scale implementation.

The objective of the market implementation methods is to influence the marketplace and to change customer behavior. The key question for planners and policy makers is the selection of the market implementation method(s) to obtain the desired customer acceptance and response. Customer acceptance refers to customer willingness to participate in a market implementation program, customer decisions to adopt the desired fuel/appliance choice and efficiency, and behavior change as encouraged by the supplier, or state. Customer response is the actual load shape change that results from customer action, combined with the characteristics of the devices and systems being used.

Customer acceptance and responses are influenced by the demographic characteristics of the customer, income, knowledge, and awareness of the technologies and programs available, and decision criteria such as cash flow and perceived benefits and costs, as well as attitudes and motivations. Customer acceptance and response are also influenced by other external factors, such as economic conditions, energy prices, technology characteristics, regulation, and tax credits.

5.3.4.6 Representative Programs in the U.S.

Numerous demand-side management programs are implemented in the U.S. yearly by various organizations. In recent years, the California Best Practices Project Advisory Committee and their contractor, Quantum Consulting, Inc., have reviewed and compared many demand-side management programs that focus on energy conservation and efficiency as part of a National Energy Efficiency Best Practices Study. The results of the study are included in a series of reports. Table 5.11 provides an overview of more than 60 programs evaluated in the National Energy Efficiency Best Practices Study that were implemented between 1999 and 2004. The type of program, program name, implementer(s), achieved energy and demand savings, program cost, and review period are listed for each program. Where values were not available, the abbreviation "NA" is used. The table shows the wide variety of programs offered spanning the residential and nonresidential sectors. Some programs provide general information and training, others target specific end uses such as lighting, heating, ventilation and air conditioning (HVAC), and new construction, and still others are comprehensive in nature. Yearly costs for the programs in Table 5.11 ranged from $150,000 for a nonresidential HVAC program offered to customers in a single service territory to $25.9 million for a statewide comprehensive program.

TABLE 5.11 Examples of Recent Demand-Side Management Programs in the U.S.

Program Type	Program Name	Implementer(s)	Time Period	Cost	Energy Savings	Demand Savings
Residential lighting[a]	2002 California Crosscutting Statewide Residential Lighting Program	Pacific Gas & Electric Co. (PG&E); Southern California Edison (SCE); San Diego Gas & Electric Co. (SDG&E)	2002	\$9.4 million	162,888 MWh	21,365 kW
	2002 Efficient Products Program—Lighting Component	Efficiency Vermont (EVT)	2002	\$1.6 million	11,039 MWh	1740 kW winter 1074 kW summer
	2002 Massachusetts Electric—Residential Lighting Program	Massachusetts Electric	2002	\$3.3 million	18,037 MWh	5084 kW
	2002 Midwest Change a Light, Change the World Campaign	Midwest Energy Efficiency Alliance (MEEA)	Fall 2002	\$630,000	10,198 MWh	NA
	2001 ENERGY STAR® Residential Lighting Program	Northwest Energy Efficiency Alliance (NW Alliance)	2001	\$2.6 million	271,560 MWh	NA
	2000–2001 Retail Lighting Program	United Illuminating	2000–2001	\$3.0 million	7808 MWh	NA
Residential AC[b]	2002 Keep Cool Air Conditioner Bounty Program	New York State Energy Research and Development Authority (NYSERDA)	2002	NA	27,208 MWh	44,813 kW
	2002 California Statewide Single-Family Rebate Program AC Component	PG&E; SCE; SDG&E	2002	NA (included in overall Single-Family Rebate Program budget)	8399 MWh	NA
	2002 New Jersey Clean Energy™ Collaborative Residential AC Component	Conectiv Power Delivery; Jersey Central Power & Light Co. (JCP&L); Public Service Electric & Gas Co. (PSE&G); Rockland Electric Company (RECO)	2002	\$24.2 million	NA	NA

(continued)

TABLE 5.11 (*Continued*)

Program Type	Program Name	Implementer(s)	Time Period	Cost	Energy Savings	Demand Savings
	2003 Air Conditioning Distributor Market Transformation Program	Oncor	2003	\$5.9 million	13,478 MWh	10,800 kW
	2002 Residential Air Conditioning Program	Florida Power and Light (FPL)	2002	\$18.0 million	78,957 MWh	37,360 kW
Single-Family Comprehensive[c]	2001–2002 Central Valley Hard-to-Reach Mobile Home Energy Savings Program	American Synergy Corp.	Oct. 2002–Oct. 2003	\$1.4 million	3,447 MWh	1,329 kW
	2002 California Statewide Single-Family Energy Efficiency Rebate Program	PG&E; SCE; SDG&E	2002	\$25.9 million	36,028 MWh	31,869 kW
	1999–2000 Residential High-Use Program	NSTAR	Aug. 1999–Aug. 2000	\$3.5 million	3,179 MWh	1,164 kW winter 831 kW summer
	2001 Energy Wise Program	National Grid U.S.A.	2001	\$1.2 million	3,461 MWh	743 kW
	2002 Efficiency Equipment Load Program	Sacramento Municipal Utility District (SMUD)	2002	\$2.4 million	1,254 MWh	700 kW
	2002 Residential Weatherization Program	Tacoma Power	2002	\$938,000	2,031 MWh	NA
Multi-Family Comprehensive[d]	2002 Multi-Family Incentive Program	Austin Energy	2002	\$581,300	3,121 MWh	2,080 kW
	2002 California Statewide Multi-Family Program	PG&E; SCE; SDG&E	2002	\$8.3 million	9,050 MWh gross 7,621 MWh net	1,853 kW
	2003 Home Energy Savings Program—Multi-Family Component	The City of Portland/Energy Trust of Oregon, Inc.	Jan.–Dec. 2003	\$1.0 million	7,000 MWh gross 2,578 MWh net	NA
	2002–2003 Apartment & Condo Efficiency Services	Focus on Energy™/Wisconsin Energy Conservation Corp. (WECC)	Sep. 2002–Aug. 2003	\$5.1 million	12,963 MWh net	2,391 kW net
	2002 Energy Wise—Multi-Family Component	National Grid	2002	\$2.3 million	3,487 MWh gross 2,706 MWh net	400 kW winter 600 kW summer
	2000 Multi-Family Conservation Program	Seattle City Light (SCL)	2000	\$1.2 million	2,769 MWh	NA

Audits & Information[e]	2002 Home Performance with ENERGY STAR Program	NYSERDA	2002	$4.0 million	741 MWh	80 kW
	2000 Time-of-Sale Home Inspection Program	SCE; GeoPraxis, Inc.	2000	$282,000	1,974 MWh	NA
	2002 Residential Conservation Services Audit Program	National Grid	2002	$2.8 million	2,677 MWh	406 kW
	2002 E+ Energy Audit for Your Home Program	Northwestern Energy	2002	$1.3 million	4,713 MWh	884 kW
	2002 Residential Energy Advisory Services Program	SMUD	2002	$1.1 million	400 MWh	70 kW
	2002 California Statewide Home Energy Efficiency Program	PG&E; SCE; SDG&E	2002	$2.0 million	8,700 MWh	4,190 kW
Residential New Construction[f]	2001–2002 Austin Green Building Program	Austin Energy	FY 2000–2001	$605,000	7,666 MWh	3,630 kW
	2002 California Energy Star New Homes Program	PG&E; SCE; SDG&E	2002	$15.2 million	10,655 MWh	22,262 kW
	2002 New Jersey ENERGY STAR Homes	Clean Energy for New Jersey	2002	$10.9 million	3,262 MWh	3,415 kW
	2002 Texas ENERGY STAR Homes Program	Oncor	2002	$5.2 million	24,700 MWh	7,410 kW
	2002 Tucson Guarantee Home Program	Tucson Electric Power	2002	$3.0 million	3,023 MWh	4,094 kW
	2001 Vermont ENERGY STAR Homes	EVT	2001	$920,000	841 MWh	278 kW
	2001–2002 Wisconsin ENERGY STAR Program	WECC	2002–2003	$2.9 million	1,049 MWh	247 kW
Nonresidential Lighting[g]	2003 Lighting Efficiency Program	Xcel Energy	2003	$2.3 million for all commercial & industrial $1.1 million for businesses <500kW	41,780 MWh for all commercial & industrial 19,433 MWh for businesses <500kW	7896 kW for all commercial & industrial 3928 kW for businesses <500kW
	2002–2003 Business Energy Services Team Program	KEMA-XENERGY	2002–2003	$941,000	2,704 MWh	559 kW
	2002 EZ Turnkey Program	SDG&E	2002	$1.3 million	3,121 MWh	570 kW

(continued)

TABLE 5.11 (*Continued*)

Program Type	Program Name	Implementer(s)	Time Period	Cost	Energy Savings	Demand Savings
	2003 Small Commercial Prescriptive Lighting Initiative	SMUD	2003	$2.7 million	19,865 MWh	3,920 kW
	2002 Small Business Energy Advantage Program	Connecticut Light & Power (CL&P)	2003	$4.6 million	16,167 MWh	3,570 kW
	2002 California Statewide Express Efficiency Program	PG&E; SCE; SDG&E	2002	$21.7 million	244,346 MWh	43,000 kW
Nonresidential HVAC[h]	New England Efficiency Partnership (NEEP) Cool Choice Program	CL&P; United Illuminating; Cape Light Compact; Massachusetts Electric Co.; Nantucket Electric Co.; NSTAR Electric; Western Massachusetts Electric Co.; Conectiv Power Delivery; JCP&L; PSE&G; Narragansett Electric Co.; Burlington Electric; EVT	2002	$2.3 million	3929 MWh	3518 kW
	Avista Rooftop HVAC Maintenance Program	Avista Utilities	2001	$1.8 million	13,000 MWh	NA
	California Express Efficiency HVAC Component	PG&E; SCE; SDG&E	2002	NA (included in overall Express Efficiency program budget)	2,901 MWh	NA
	Los Angeles Department of Water and Power (DWP) Chiller Efficiency	Los Angeles DWP	2003–2004	$786,430	7,174 MWh	5,666 kW
	FP&L Commercial/Industrial HVAC Program	FPL	2002	$5.4 million	NA	NA
	Glendale Water and Power Check Me!	Glendale Water and Power	2001	$150,000	25,128 MWh	358 kW
Nonresidential Large Comprehensive Incentive[i]	Non-Residential Standard Performance Contract	PG&E; SCE; SDG&E	2002	$23.0 million	167,300 MWh	28,441 kW
	Energy \$mart™ C/I Performance	NYSERDA	2001–2002	$34.2 million	204,500 MWh	53,886 kW
	Energy Opportunities	United Illuminating	2002	$1.3 million	10,772 MWh	2,627 kW

	Power Smart	BC Hydro	2004	$7.8 million industrial $17.6 million commercial and government	54,000 MWh industrial 74,000 MWh commercial and government	NA
	Custom Efficiency	Xcel Energy (Colorado)	2002–2005	$12.2 million	76,167 MWh	40,077 kW
	Custom Services	CL&P	2003	$8.6 million	24,853 MWh	NA
	Energy Initiative	National Grid	2002	$9.7 million	30,862 MWh	6,089 kW
	Energy Shared Savings	WP&L (Alliant) Wisconsin	2001	$21.9 million	104,325 MWh	16,000 kW
	Business Energy Services	EVT	2002	$1.1 million	4,955 MWh	NA
	Commercial & Industrial Custom Retrofit	SMUD	2002	$7.3 million	NA	NA
New Construction Information & Services[j]	Energy Conscious Construction	Northeast Utilities	2002	$7.4 million	33,365 MWh	NA
	Energy Design Assistance	Xcel Energy	2002	$3.4 million	63,093 MWh	19,100 kW
	Design 2000 Plus	National Grid	2002	$13.9 million	31,804 MWh	6,429 kW
	Savings by Design	PG&E; SCE; SDG&E	2002	$22.6 million	82,697 MWh	18,600 kW
	Construction Solutions	NSTAR	2001	$7.9 million	14,230 MWh	1,710 kW
	Commercial & Industrial New Construction Program	Hawaiian Electric Co. (HECO)	1999	$935,000	5584 MWh	821 kW

[a] Quantum Consulting, Inc., 2004. *Residential Lighting Best Practices Report*, *Vol. 1*, *National Energy Efficiency Best Practices Study*. Quantum Consulting, Inc., Berkeley, CA.

[b] Quantum Consulting, Inc., 2004. *Residential Air Conditioning Best Practices Report*, *Vol. R2*, *National Energy Efficiency Best Practices Study*. Quantum Consulting Inc., Berkeley, CA.

[c] Quantum Consulting, Inc., 2004. *Residential Single-Family Comprehensive Weatherization Best Practices Report*, *Vol. R4*, *National Energy Efficiency Best Practices Study*. Quantum Consulting Inc., Berkeley, CA.

[d] Quantum Consulting, Inc., 2004. *Residential Multi-Family Comprehensive Best Practices Report*, *Vol. R5*, *National Energy Efficiency Best Practices Study*. Quantum Consulting Inc., Berkeley, CA.

[e] Quantum Consulting, Inc., 2004. *Residential Audit Programs Best Practices Report*, *Vol. R7*, *National Energy Efficiency Best Practices Study*. Quantum Consulting Inc., Berkeley, CA.

[f] Quantum Consulting, Inc., 2004. *Residential New Construction Best Practices Report*, *Vol. R8*, *National Energy Efficiency Best Practices Study*. Quantum Consulting Inc., Berkeley, CA.

[g] Quantum Consulting, Inc., 2004. *Non-Residential Lighting Best Practices Report*, *Vol. NR1*, *National Energy Efficiency Best Practices Study*. Quantum Consulting Inc., Berkeley, CA.

[h] Quantum Consulting, Inc., 2004. *Non-Residential HVAC Best Practices Report*, *Vol. NR2*, *National Energy Efficiency Best Practices Study*. Quantum Consulting Inc., Berkeley, CA.

[i] Quantum Consulting, Inc., 2004. *Non-Residential Large Comprehensive Incentive Programs Best Practices Report*, *Vol. NR5*, *National Energy Efficiency Best Practices Study*. Quantum Consulting Inc., Berkeley, CA.

[j] Quantum Consulting, Inc., 2004. *Non-Residential New Construction Best Practices Report*, *Vol. NR8*, *National Energy Efficiency Best Practices Study*. Quantum Consulting Inc., Berkeley, CA.

Reported energy savings ranged from 400 MWh for a residential audit and information program offered in a single service territory to 271,560 MWh for a northwest regional ENERGY STAR® residential lighting program. The following section examines four of the programs from Table 5.11 in more detail.

5.3.5 Case Studies

5.3.5.1 Case Study 1: 2001 California 20/20 Rebate Program[1]

In response to a crisis of constrained supply, skyrocketing electricity prices, and fear of summer blackouts, California launched an enormous effort in 2001 to conserve energy and reduce electricity demand. This effort was primarily embodied in emergency legislation that provided additional funding and led to the rapid development and deployment of hundreds of energy efficiency programs administered and implemented by a variety of entities. One of the most successful programs was the 2001 California 20/20 Rebate Program (see Table 5.12 for program summary). This program provided rebates to residential and small commercial/industrial customers of the state's investor-owned utilities for reducing monthly electricity usage from June through September, 2001. Customers were offered a 20% rebate off of the electricity commodity portion of their energy bill for lowering their total monthly electricity use by at least 20% compared to the same month of the previous year. In addition, large commercial/industrial customers with time-of-use meters received a 20% rebate off of their summer on-peak demand and energy charges for reducing on-peak electricity use by at least 20%. The program was executed with cooperation and funding by several state agencies and organizations. Because the program overlapped with other California programs such as the "Flex Your Power" marketing campaign, it was difficult to accurately credit energy savings specifically to the 20/20 Rebate Program. In a study for the California Measurement Advisory Council (CALMAC), which involved evaluating the success of California's 2001 programs, Global Energy Partners (Global) attempted to adjust reported savings from the 20/20 Rebate Program to discount the effect of double counting. The reported accomplishments of the program were 5.3 million MWh in energy savings and 2616 MW in demand savings. With Global's adjustments for double counting among programs, the energy savings were estimated to be 3.1 MWh. The program budget for 2001 was $350 million. Although this program was hugely successful during the crisis of 2001, it is difficult to sustain a program of this type in the absence of an immediate energy crisis. As a result, it was discontinued after 2002.

5.3.5.2 Case Study 2: 2002 California Statewide Residential Lighting Program[2]

The 2002 California Statewide Residential Lighting Program was designed in response to the 2001 energy crisis experienced in California (see Table 5.13 for program summary). Its purpose was to encourage greater penetration of energy efficient lamps and fixtures into the residential sector. The products covered included compact fluorescent lamps, torchieres, ceiling fans, and complete fixtures. This was accomplished by providing rebates to manufacturers (manufacturer upstream buydown) to lower wholesale costs as well as by providing instant rebates to consumers at the point of sale. The program was implemented by three large investor-owned utilities in the state: Pacific Gas & Electric (PG&E), Southern California Edison (SCE), and San Diego Gas & Electric (SDG&E). Each utility had in-house management responsibilities. The program leveraged on relationships with manufacturers and retailers established in previous lighting programs. Progress was tracked by using data on the number of products delivered by manufacturers and retailer sales information. In all, 5,502,518 lamps, 24,932 fixtures, 6736 torchieres, and 50 ceiling fans with bulbs were rebated during 2002. The estimated program accomplishments were 162,888 MWh in energy savings, and 21.4 MW in demand savings. The total program cost was $9.4 million.

[1]Data from Global Energy Partners, 2003. *California Summary Study of 2001 Energy Efficiency Programs*. Publication 02-1099. Global Energy Partners, LLC, Lafayette, CA.

[2]Data from Quantum Consulting, Inc., 2004. *Residential Lighting Best Practices Report, Vol. 1, National Energy Efficiency Best Practices Study*. Quantum Consulting, Inc., Berkeley, CA.

TABLE 5.12 Case Study 1: 2001 California 20/20 Rebate Program

Description

The 20/20 Rebate program was a statewide program designed to address California's energy crisis in 2001. It provided rebates to residential and small commercial/industrial customers of the state's investor-owned utilities for reducing monthly electricity usage from June through September, 2001. Customers were offered a 20% rebate off of the electricity commodity portion of their energy bill for lowering their total monthly electricity use by at least 20% compared to the same month of the previous year. In addition, large commercial/industrial customers with time-of-use meters received a 20% rebate off of their summer on-peak demand and energy charges for reducing on-peak electricity use by at least 20%

Targeted Sector/Building Type

All residential, commercial, and industrial customers of California's investor-owned utilities

Targeted End Use Technology/Program Type

Rebate for reducing electricity use by 20% relative to same month in previous year

Program Implementer

Executive order of Governor Gray Davis; included cooperation and funding by several state agencies and organizations

Budget for Year

$350 million for 2001

Program Results

Reported energy savings = 5,258,000 MWh[a]
Adjusted energy savings = 3,053,000 MWh[b]
Demand savings = 2616 MW[c]

[a] Goldman and Barbose. 2002. *California Customer Load Reductions During the Electricity Crisis: Did they Help to Keep the Lights On?* Lawrence Berkeley National Laboratories, Berkeley, CA.

[b] During an evaluation of California's 2001 energy efficiency programs for the California Measurement Advisory Council (CALMAC), Global Energy Partners adjusted the 20/20 Rebate program's energy savings to correct for double counting. *Data Source*: Global Energy Partners. 2003. *California Summary Study of 2001 Energy Efficiency Programs*. Publication 02-1099. Global Energy Partners, LLC, Lafayette, CA.

[c] This value represents "residual" peak demand reduction for September 2001 attributable to the combined effects of the 20/20 Rebate program, the *Flex Your Power* public awareness campaign, electricity rates, and voluntary demand-side management. *Data Source*: California Energy Commission. 2001. *California Market Report*. Sacramento, CA, October.

TABLE 5.13 Case Study 2: 2002 California Statewide Residential Lighting Program

Description

The 2002 California Residential Lighting program was designed to build upon the success of earlier lighting efficiency programs, while at the same time address the more immediate energy needs brought about by the 2001 energy crisis in California. The program was unique in that it was implemented identically across the State's investor-owned utility service territories. It offered point-of-sale rebates and manufacturer buy-down to reduce lighting costs to residential customers

Targeted Sector/Building type

All residential

Targeted End Use Technology/Program Type

Multiple lighting measures

Program Implementer

California investor-owned utilities; PG&E, SCE, and SDG&E

Budget for Year

$9.4 million for 2002; $7.3 million total incentives paid

Program Results

Reported energy savings = 162,888 MWh
Demand savings = 21.4 MW

Source: From Quantum Consulting, Inc., *Residential Lighting Best Practices Report*, *Vol. 1*, *National Energy Efficiency Best Practices Study*. Quantum Consulting Inc., Berkeley, CA, 2004.

TABLE 5.14 Case Study 3: 2003 Xcel Energy Lighting Efficiency Program

Description
The 2003 Xcel lighting efficiency program offered low-cost energy assessments, low-interest financing, and rebates for replacing lighting in existing buildings and for adding energy efficient lighting during new construction. The rebates were both prescriptive and custom in nature

Targeted Sector/Building Type
All commercial and industrial customers
Small business customers < 500 kW

Targeted End Use Technology/Program Type
Multiple lighting measures

Program Implementer
Xcel energy of minnesota

Budget for Year
All commercial and industrial customers: $2.29 million for 2003; $1.51 million total incentives paid
Small business customers < 500 kW: $1.09 million for 2003; $0.66 million total incentives paid

Program Results
Reported energy savings = 41,780 MWh (net) for all commercial and industrial customers
= 19,433 MWh for small business customers < 500 kW
Demand savings = 7.9 MW (summer) for all commercial and industrial customers
= 3.9 MW for small business customers < 500 kW
Almost 900 prescriptive rebates were offered during 2003

Source: From Quantum Consulting, Inc., *Non-Residential Lighting Best Practices Report*, *Vol. NR1*, *National Energy Efficiency Best Practices Study*. Quantum Consulting, Inc., Berkeley, CA, 2004.

5.3.5.3 Case Study 3: 2003 Xcel Energy Lighting Efficiency Program[3]

The Xcel Energy Lighting Efficiency Program was designed to encourage energy efficient lighting design for commercial and industrial customers in Minnesota (see Table 5.14 for program summary). The implementer was Xcel Energy, a utility company. The program targeted all commercial and industrial customers, as well as small (less than 500 kW) businesses. It offered prescriptive and custom rebates, energy assessments, and low-interest financing for projects. The rebates were applicable both to lighting in new construction and to replacement of lighting in existing buildings. The customers or vendors were responsible for installing the lighting equipment. Xcel Energy managed the program with in-house personnel. The program managers were responsible for approving projects, monitoring applications, and verifying and tracking installations. During 2003, almost 900 prescriptive lighting projects were undertaken. The estimated program accomplishments were energy savings of 41,780 MWh for all commercial and industrial customers and 19,433 MWh for small businesses. The estimated demand savings were 7.9 and 3.9 MW, respectively, for commercial and industrial customers and for small businesses. The program costs for the year were $2.3 million for commercial and industrial customers, and $1.1 million for small businesses.

5.3.5.4 Case Study 4: 2002 Northeast Energy Efficiency Partnership Cool Choice Program[4]

The 2002 Northeast Energy Efficiency Partnership (NEEP) Cool Choice program was designed to increase penetration of high efficiency cooling systems in commercial and industrial buildings in the Northeast states of Connecticut, Rhode Island, Vermont, Massachusetts, and New Jersey (see Table 5.15 for program summary). NEEP, which is a collaboration of over a dozen utilities in the Northeast region,

[3]Data from Quantum Consulting, Inc. 2004. *Non-Residential Lighting Best Practices Report*, *Vol. NR1*, *National Energy Efficiency Best Practices Study*. Quantum Consulting Inc., Berkeley, CA.

[4]Data from Quantum Consulting, Inc. 2004. *Non-Residential HVAC Best Practices Report*, *Vol. NR2*, *National Energy Efficiency Best Practices Study*. Quantum Consulting Inc., Berkeley, CA.

TABLE 5.15 Case Study 4: 2002 Northeast Energy Efficiency Partnership (NEEP) Cool Choice Program

Description
The 2002 Cool Choice program was an incentive-based program that provided rebates to commercial and industrial customers who purchased high efficiency air conditioning systems. The rebates were intended to offset the higher costs associated with energy efficient units

Targeted Sector/Building Type
All commercial and industrial customers

Targeted End Use Technology/Program Type:
High efficiency direct expansion air conditioners and heat pumps; economizers

Program Implementer
NEEP

Budget for Year
$2.3 million for 2002

Program Results
Reported energy savings = 3929 MWh
Demand savings = 3.5 MW

Source: From Quantum Consulting, Inc., *Non-Residential HVAC Best Practices Report*, *Vol. NR2*, *National Energy Efficiency Best Practices Study*, Quantum Consulting Inc., Berkeley, CA, 2004.

administered the program. The 2002 program was a continuation of an on-going program established in 1998 to educate HVAC contractors in the correct installation of HVAC systems and to encourage them to up-sell high efficiency units to consumers. The program offers an incentive to customers to help offset the greater costs associated with high-efficiency air-conditioning equipment. The incentive is based on the incremental improvement in efficiency provided by the energy efficient alternative, and covered 80% of the incremental costs for air conditioning systems of 30 tn. or less. An outside implementer managed the program and outreached to HVAC contractors, who then outreached to customers. The estimated program accomplishments for 2002 were 3929 MWh in energy savings and 3.5 MW in demand savings. The total program cost for the year was $2.3 million.

5.3.6 Conclusions

Since the early 1970s, economic, political, social, technological, and resource supply factors have combined to change the energy industry's operating environment and its outlook for the future. Many are faced with staggering capital requirements for new plants, significant fluctuations in demand and energy growth rates, declining financial performance and political or regulatory and consumer concern about rising prices. Although demand-side management is not a cure-all for these difficulties, it does provide for great many additional alternatives. These demand-side alternatives are equally appropriate for consideration by utilities, energy suppliers, energy-service suppliers, and government entities. Implementation of demand-side measures not only benefits the implementing organization by influencing load characteristics, delaying the need for new energy resources, and in general improving resource value, but it also provides benefits to customers such as reduced energy bills and/or improved performance from new technological options. In addition, society as a whole receives economic, environmental, and national security benefits. For example, because demand-side management programs can postpone the need for new power plants, the costs and emissions associated with fossil-fueled electricity generation are avoided. Demand-side management programs also tend generate more jobs and expenditures within the regions where the programs are implemented, boosting local economies. Moreover, demand-side management programs can help reduce a country's dependence on foreign oil imports, improving national security. Demand-side management alternatives, particularly those focused on energy conservation and efficiency, will continue to hold an important role in resources planning in the US and abroad, and will be a critical element in the pursuit of a sustainable energy future.

References

AC Propulsion. 2002. Vehicle-to-grid demonstration project: Grid regulation ancillary service with a battery electric vehicle (December, 3–10). 2002. Available at: http://www.acpropulsion.com/reports/V2G%20Final%20Report%20R5.pdf

Arthur, D. Little. 1999. *Distributed Generation: System Interfaces.* An Arthur D. Little white paper, ADL publishing, Boston, MA.

Gellings, C. W. 1984–1988. *Demand-Side Management*, *Vols. 1–5*. EPRI, Palo Alto, CA.

Global Energy Partners. 2003. *California Summary Study of 2001 Energy Efficiency Programs*. Publication 02-1099. Global Energy Partners, LLC, Lafayette, CA.

Goldman, E. and Barbose, G. L. 2002. *California Customer Load Reductions During the Electricity Crisis: Did They Help to Keep the Lights On?.* Lawrence Berkeley National Laboratories, Berkeley, CA.

Institute of Electric and Electronics Engineers (IEEE). 2001. IEEE P1547/D07. *Standard for Interconnecting Distributed Resources with Electric Power Systems*. IEEE Standards Coordinating Committee 21 (IEEE SCC21) on Fuel Cells, Photovoltaics, Dispersed Generation, and Energy Storage of the IEEE Standards Association, New York.

Further Reading

Bathie, W. W. 1995. *Fundamentals of Gas Turbines. 2nd Ed.* Wiley, New York.

Borbely, A.-M. and Kreider, J. F. 2001. *Distributed Generation, The Power Paradigm of the New Millennium.* CRC Press, Boca Raton.

California Energy Commission 2000. *The Role of Energy Efficiency and Distributed Generation in Grid Planning.* California Energy Commission, Sacromento, CA.

California Energy Commission 2001. *California Market Report.* California Energy Commission, Sacramento, CA.

California Energy Commission. 2003. California Interconnection Guidebook: A Guide to Interconnecting Customer-Owned Electric Generation Equipment to the Electric Utility Distribution System using California's Electric Rule 21, California Energy Commission, Sacramento, CA.

Davis, M. W. 2002. Mini gas turbines and high speed generators a preferred choice for serving large commercial customers and microgrids. I. Generating system. In *IEEE Power Engineering Society Summer Meeting, 2002*, *Vol. 2*, pp. 21–25.

Davis, M. W., Gifford, A. H., and Krupa, T. J. 1999. Microturbines-an economic and reliability evaluation for commercial, residential and remote load applications, *IEEE Transactions on Power Systems*, *Vol. 14*, (4), 1556–1562.

Del Monaco, J. L. 2001. The role of distributed generation in the critical electric power infrastructure. *IEEE Power Engineering Society Winter Meeting*, *Vol. 1*, p. 28.

El-Khattam, W. 2004. Distributed generation technologies, definitions and benefits, *Electric Power Systems Research*, 71 (2), 119–128.

EPRI 1993. *Principles and Practice of Demand-Side Management.* EPRI, Palo Alto, CA.

Gellings, C. W. 2002. Using demand-side management to select energy efficient technologies and programs, in Efficient Use and Conservation of Energy, edited by Clark W. Gellings, In *Encyclopedia of Life Support Systems (EOLSS)*, Developed under the Auspices of the UNESCO, EOLSS Publishers, Oxford, UK (http://www.eolss.net).

Gellings, C. W. and Chamberlin, J. H. 1993. *Demand-Side Management: Concepts and Methods. 2nd Ed.* The Fairmont Press, Lilburn, GA.

Gutierrez-Vera, J. 2001. Mini cogeneration schemes in Mexico. *IEEE Power Engineering Review*, 21 (8), 6–7.

International Energy Agency 2002. *Distributed Generation in Liberalised Electricity Markets.* International Energy Agency, Paris.

Kolanowski, B. F. 2004. *Guide to Microturbines.* Marcel Dekker, New York.

Lents, J. and Allison., J. E. 2000. Can we have our cake and eat it, too? Creating distributed generation technology to improve air quality. Final Report. The Energy Foundation, San Francisco, CA.

Li, S., Tomsovic, K., Hiyama, T. 2000. Load following function using distributed energy resources. In *IEEE Power Engineering Society Summer Meeting, 2000.Vol. 3*, pp. 16–20.

Makhkamov, K. and Ingham, D. B. 2000. Theoretical investigations on the Stirling engine working process. *Energy Conversion Engineering Conference and Exhibit, 2000. (IECEC) 35th Intersociety, Vol. 1*, pp. 24–28.

Makhkamov, K., Trukhov, V., Orunov, B., Korobkov, A., Lejebokov, A., Tursunbaev, I., Orda, E., et al. 2000. Development of solar and micro cogeneration power installations on the basis of Stirling engines. *Energy Conversion Engineering Conference and Exhibit, 2000. (IECEC) 35th Intersociety, Vol. 2*, pp. 24–28.

Masters, G. M. 2004. *Renewable and Efficient Electric Power Systems.* Wiley-IEEE Press, New York.

National Renewable Energy Laboratory (NREL). *Technologies for Distributed Energy Resources*, US Department of Energy, Federal Energy Management.

Pepermans, G. 2005. Distributed generation: definition, benefits and issues. *Energy Policy*, 33 (6), 787–798.

Puttgen, H. B., MacGregor, P. R., and Lambert, F. C. 2003. Distributed generation: semantic hype or the dawn of a new era? *IEEE Power and Energy Magazine*, 1 (1), 22–29.

Quantum Consulting. 2004. *National Energy Efficiency Best Practices Study*, Vols. R1, R2, R4, R5, R7, R8, NR1, NR2, NR5, NR8. Quantum Consulting, Inc., Berkeley, CA.

Raggi, L., Katsuta, M., Isshiki, N., and Isshiki, S. 1997. Theoretical and experimental study on regenerative rotary displacer Stirling engine. *Energy Conversion Engineering Conference, 1997; IECEC-97. Proceedings of the 32nd Intersociety.*

Resource Dynamics Corporation. 2001. *Assessment of Distributed Generation Technology Applications*, Maine Public Utilities Commission, Vienna, VA.

Sweet, W. 2001. Networking assets [distributed generation]. *IEEE Spectrum*, 38 (1), 84–86 (see also p.88).

World Alliance for Decentralized Energy (WADE) 2003. *Guide to Decentralized Energy Technologies.* WADE, Edinburgh, Scotland.

World Alliance for Decentralized Energy (WADE) 2005. *World Survey of Decentralized Energy 2005.* WADE, Edinburgh, Scotland.

Willis, H. L. and Scott, W. G. 2000. *Distributed Power Generation, Planning and Evaluation.* Marcel Dekker, New York.

Smeers, Y. and Yatchew, A. Distributed resources: toward a new paradigm of the electricity business. *The Energy Journal*, (special issue).

Zhu, Y. and Tomsovic, K. 2002. Development of models for analyzing the load-following performance of microturbines and fuel cells. *Electric Power Systems Research.*

6

Generation Technologies through the Year 2025

Anthony F. Armor
Electric Power Research Institute

Edwin A. Harvego
Idaho National Laboratory

Kenneth D. Kok
WSMS Mid-America

6.1 Fossil Fuels

Anthony F. Armor

6.1.1 Introduction

The generation of electric power from fossil fuels has seen continuing, and in some cases dramatic technical advances over the last 20–30 years. Technology improvements in fossil fuel combustion have been driven largely by the need to reduce emissions, by the need to conserve fossil fuel resources, and by the economics of the competitive marketplace. The importance of fossil fuel-fired electric generation to the world is undeniable—more than 70% of all power in the U.S. is fossil fuel-based and worldwide the percentage is higher, and growing. Today most large power plants worldwide burn coal though many generating companies are adding natural gas plants, particularly where the cost of gas-fired generation, and the long-term supply of gas, appear favorable. This chapter reviews the current status and likely future deployment of competing generation technologies based on fossil fuels.

It is likely, particularly in the developed world, that gas turbine-based plants will continue to be added to the new generation market in the immediate future. The most advanced combustion turbines now achieve more than 40% lower heating value (LHV) efficiency in simple cycle mode and greater than 50% efficiency in combined cycle mode. In addition, combustion turbine/combined cycle (CT/CC) plants offer siting flexibility, swift construction schedules, and capital costs between \$400/kW and \$800/kW. These advantages, coupled with adequate natural gas supplies (though new wells and pipelines will be needed in the United States) and the assurance, in the longer term, of coal gasification backup, have made this technology one important choice for green field and repowered plants in the U.S. and in Europe.

However, fossil steam pulverized coal (PC) plants are dominant in the expanding nations of the East such as China and India. In fact for the developing world there is good reason why the coal-fired power plant may still be the primary choice for many generation companies. Fuel is plentiful and inexpensive, and sulfur dioxide scrubbers have proved to be more reliable and effective than early plants indicated. In fact up to 99% SO_2 removal efficiency is now possible. Removal of nitrogen oxides is also well advanced with over 95% removal possible using selective catalytic reduction (SCR). Ways to remove mercury are currently under study, and the issue of carbon dioxide control and sequestration from fossil plants is receiving renewed attention as ways to control global warming are pursued. Combustion of coal currently occurs in three basic forms, direct combustion of PC, combustion of coal in a suspended bed of coal and inert matter, and coal gasification.

The *pulverized coal plant*, the most common form of coal combustion, has the capability for much improved efficiency even with full flue gas desulfurization (FGD), ferritic materials technology now having advanced to the point where higher steam pressures and temperatures are possible. Advanced supercritical PC plants are moving ahead commercially, particularly in Japan and Europe. Even higher steam conditions for PC plants, perhaps using nickel-based superalloys, are under study.

Worldwide the application of *atmospheric fluidized bed combustion* (AFBC) plants has increased, and such plants offer reductions in both SO_2 and NO_x while permitting the efficient combustion of vast deposits of low-rank fuels such as lignites. Since the early 1990s, AFBC boiler technology has become established worldwide as a mature, reliable technology for the generation of steam and electric power—with its advantage of in-furnace SO_2 capture with limestone. In fact, the major impetus in the widespread deployment of this relatively new boiler technology since the mid-1980s has been its resemblance to a conventional boiler with the added capability for in-situ SO_2 capture, which could eliminate or reduce the need for flue-gas desulfurization.

Coal gasification power plants are operating at the 250–300 MW level. Much of the impetus came from the U.S. DOE clean coal program where two gasification projects are in successful commercial service. Large gasification plants for power are also in operation in Europe. Gasification with combined cycle operation not only leads to minimum atmospheric (SO_2 and NO_x) and solid emissions, but also provides an opportunity to take advantage of new gas turbine advances. With the rapid advances now being introduced in combustion turbine technology, the coal gasification option is seen as a leading candidate for new plant construction within the first half of the 21st century. Several new gasification plants are planned for the U.S. within the 2010–2020 time frame.

6.1.2 Fuels for Electric Power Generation in the United States

The Energy Information Administration lists more than 500 GW of fossil steam units in the U.S. Coal-fired units dominate with 1400 units capable of generating over 300 GW. All told, fossil-steam plants generate more than 70% of all electric energy in the country, Figure 6.2, and these aging units, on average more than 35 years old, will remain the foundation of the power industry for the immediate future, and certainly through the next 10 years.

The U.S. electric power industry burns about \$30 billion worth of fossil fuels each year, accounting for 70%–80% of the operating costs of fossil-fired plants. Coal dominates and recent changes to the fuel mixes include:

- A mix of eastern high-sulfur coal with low-sulfur, low-cost western coals, often from Powder River Basin (PRB) deposits in Montana and Wyoming. Compared with eastern bituminous coals, PRB coals have LHV, sulfur and ash, but higher moisture content and finer size.

- A mix of 10%–20% gas with coal in a boiler designed for coal firing.
- Orimulsion, a bitumen-in-water emulsion produced only from the Orinoco Basin in Venezuela. This fuel is relatively high in sulfur and vanadium. Power plants that use this fuel need to add scrubbers.
- A mix of coal with petroleum coke, a by-product of refining, whose cost is currently low but whose sulfur content is high.

6.1.2.1 Coal as a Fuel for Electric Power

Coal is the altered remains of prehistoric vegetation that originally accumulated as plant material in swamps and peat bogs. The accumulation of silt and other sediments, together with movements in the earth's crust (tectonic movements) buried these swamps and peat bogs, often to great depth.

With burial, the plant material was subjected to elevated temperatures and pressures, which caused physical and chemical changes in the vegetation, transforming it into coal. Initially the peat, the precursor of coal, was converted into lignite or brown coal—coal-types with low organic maturity. Over time, the continuing effects of temperature and pressure produced additional changes in the lignite, progressively increasing its maturity and transforming it into what is known as subbituminous coals. As this process continued, further chemical and physical changes occurred until these coals became harder and more mature, at which point they are classified as bituminous coals. Under the right conditions, the progressive increase in the organic maturity continued, ultimately to form anthracite.

The degree of metamorphism or coalification undergone by a coal, as it matures from peat to anthracite, has an important bearing on its physical and chemical properties, and is referred to as the "rank" of the coal. Low-rank coals, such as lignite and subbituminous coals, are typically softer, friable materials with a dull, earthy appearance; they are characterized by high-moisture levels and a low-carbon content, and hence a low-energy content. Higher rank coals are typically harder and stronger and often have a black vitreous luster. Increasing rank is accompanied by a rise in the carbon and energy contents and a decrease in the moisture content of the coal. Anthracite is at the top of the rank scale and has a correspondingly higher carbon and energy content and a lower level of moisture.

Large coal deposits only started to be formed after the evolution of land plants in the Devonian period, some 400 million years ago. Significant accumulations of coal occurred during the Carboniferous period (350–280 million years ago) in the Northern Hemisphere, the Carboniferous/Permian period (350–225 million years ago) in the Southern Hemisphere and, more recently, the late Cretaceous period to early Tertiary era (approximately 100–15 million years ago) in areas as diverse as the United States, South America, Indonesia and New Zealand. Of all the fossil fuels, coal is the most plentiful in the world. It is geographically dispersed, being spread over 100 countries and all continents. Coal reserves have been identified that confirm over 200 years of resource availability. The percent of the world reserves categorized by type and use is shown in Figure 6.1 below. Almost one half (48%) of the world's coal reserves is made up of lignite and subbituminous coals, and these coals are used primarily for power generation.

6.1.3 Clean Coal Technology Development

At an increasing rate in the last few years, innovations have been developed and tested aimed at reducing emissions through improved combustion and environmental control in the near term, and in the longer term by fundamental changes in the way coal is preprocessed before converting its chemical energy to electricity. Such technologies are referred to as "Clean Coal Technologies" described by a family of precombustion, combustion/conversion, and post-combustion technologies. They are designed to provide the coal user with added technical capabilities and flexibility, and the world with an opportunity to exploit our most abundant fossil source. They can be categorized as:

- Precombustion, where sulfur and other impurities are removed from the fuel before it is burned.
- Combustion, where techniques to prevent pollutant emissions are applied in the boiler while the coal burns.

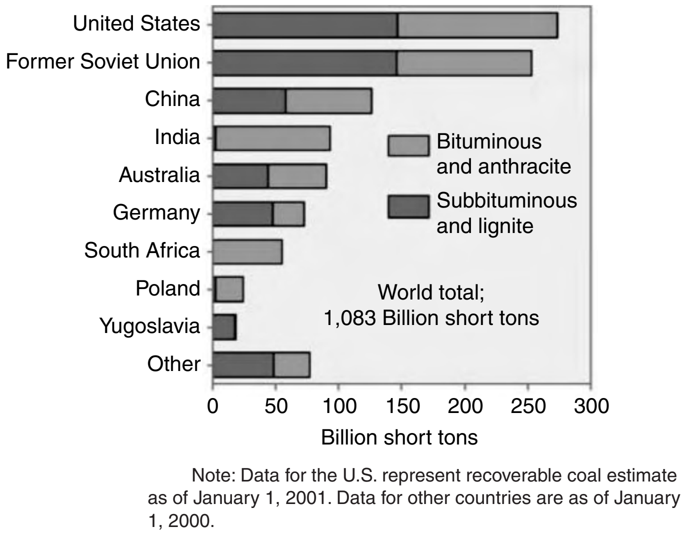


FIGURE 6.1 Worldwide coal reserves. (From Energy Information Administration, International Engergy Annual 2001, DOE/EIA-0219(2001), (Washington, DC, February 2003.) Table 8.2. www.eia.doe.gov/iea/.)

- Post-combustion, where the flue gas released from the boiler is treated to reduce its content of pollutants.
- Conversion, where coal, rather than being burned, is changed into a gas or liquid that can be cleaned and used as a fuel.

6.1.3.1 Coal Cleaning

Cleaning of coal to remove sulfur and ash is well established in the U.S. with more than 400 operating plants, mostly at the mine. Coal cleaning removes primarily pyritic sulfur (up to 70% SO_2 reduction is possible) and in the process increases the heating value of the coal, typically about 10% but occasionally 30% or higher. The removal of organic sulfur, chemically part of the coal matrix, is more difficult, but may be possible using microorganisms or through chemical methods and research is underway [Couch, 1991]. Heavy metal trace elements can be removed also, conventional cleaning removing (typically) 30%–80% of arsenic, mercury, lead, nickel, antimony, selenium and chromium.

6.1.4 New Generation Needs

The exploding global demand for electricity, particularly in the developing world, implies practical electric generating options, and significant increases in efficiency throughout the entire energy chain. A strong portfolio of advanced power generation options would include fossil, renewable and nuclear, all essential to meet these growth requirements, both domestically and globally.

Over the next 20 years, the developed world will meet the need for additional capacity largely with coal and gas-fired plants, while the developing world will continue to rely on indigenous resources, particularly coal in the case of China and India. In the U.S. the likely balance between coal- and gas-fired new generation is a topic of significant debate as discussed above. While gas-fired additions currently predominate, the relative fuel prices (coal to gas), and the timing of additional emissions requirements will ultimately determine the preferred choice between these two fossil fuel options. In the longer term noncarbon sources will begin to supplant both.

6.1.4.1 Central Station Options for New Generation

Coal and gas fuels are expected to continue to dominate U.S. central stations in the next decade, with gas-fired combined cycles supplanting several older fossil steam stations. Timing is still a question though. At the end of 2006, more than 150 new coal-fired plants were being planned in 42 states. Based on the expected mix of coal, oil, gas, nuclear, and renewables through the year 2015 the U.S. central station generation options for fossil fuels may be described as follows:

1. Coal, oil, and gas-fired plants of conventional design mostly Rankine cycles.
   These represent the majority of plants currently in operation. On average they are 35-years old, many (more than 100,000 MW) equipped with SO_2 scrubbers, with NO_x control additions, and with other environmental upgrades. Yet they provide the bulk of our electricity needs, are extremely reliable, and are increasingly in demand as evidenced by an average capacity factor above 70% in the early years of this century.
2. Repowered plants, based on gas-firing and combined cycle operation.
   Many of the gas-fired steam plants that have changed hands are targeted for repowering. That is combustion turbines will be added to provide exhaust heat for producing steam for the existing steam turbines. This combination of gas and steam turbine cycles adds MW, reduces emissions, and improves efficiencies 5% or more.
3. New combined cycles based on gas-firing, and on coal-firing with gasification, utilizing advanced gas and steam turbine technology.
   Gas-fired combined cycles have been the new central plants of choice moving into the new century. Though relatively few are in operation today a considerable number had been planned. However massive deployment of these plants in the future raises questions of gas and gas pipeline availability, and gas prices. Any potential retreat from coal though could have serious future energy consequences.
4. Coal-fired Rankine cycles with advanced steam conditions.
   Advancing steam temperatures and pressures in pulverized steam plants greatly improves overall efficiency. Such ultrasupercritical cycles are already in operation outside the U.S. Advancing steam temperatures to 700°C from current levels of about 590°C enhances efficiency and reduces emissions. When used in coal combined cycles, and with temperatures increased to 750°C or beyond, a coal plant beyond 55% efficiency can be attained. Significant challenges still exist in materials technology.
5. Integrated coal gasification fuel cells perhaps combined with gas turbines with efficiencies of 60% or more.
   The fuel cell is an exciting advance that will change the energy picture in the long term. In the shorter term, and in small sizes, advances are being made in both mobile and stationary applications. If the fuel cell can be used as a combuster for a gas turbine, efficiencies can be raised above 60%. Clearly this is a power source of great promise for the second half of this century.

6.1.5 Pulverized Coal Power Plants

Today more than 50% of total U.S. electricity generation is supplied by about 1000 large coal-fired units, with a concentration in the 500–600 MW unit sizes. These plants are typically aging, 35-plus years old on average, and compete with other regional coal-fired plants at the cost margin. Fuel cost is frequently the determining factor in assuring cost-competitiveness. For this reason—and also to reduce the price of emissions control—many eastern coal-fired units now burn a mixture of eastern coal and coal from the low-sulfur deposits of the Powder River Basin of Wyoming and Montana. Average capacity factor for coal-fired units is about 70%, up from 60% about 10 years ago. Equivalent availabilities average about 85%, close to a 15-year high, and are achieved despite an aging fleet and reduced staffing levels. By the early 2000s sulfur dioxide and nitrogen oxides were down more than 40 and 20%, respectively, since the early 1980s,

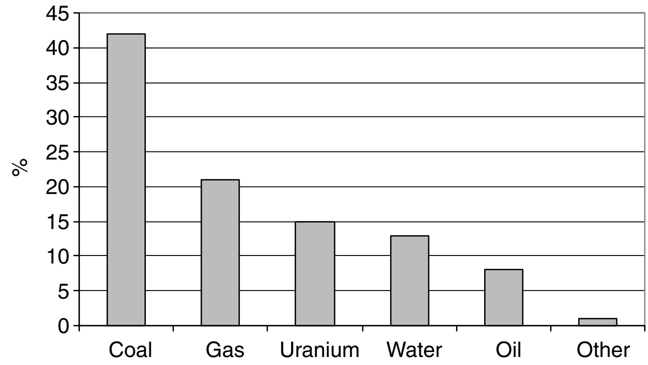


FIGURE 6.2 US installed capacity by fuel percentage.

even though electricity production climbed about 40% in that period. These performance parameters reflect the progressive advances in our understanding of the combustion of coal (Figure 6.2 through Figure 6.4).

6.1.5.1 Cycle Selection

The selection of a supercritical versus a subcritical cycle for a fossil-steam unit is dependent on many site specific factors including fuel cost, emissions regulations, capital cost, load factor, duty, local labor rates, and perceived reliability and availability. In fact the use of subcritical cycles for the limited number of fossil-steam plants that have been build in the U.S. in the last 20 years has been mainly due to relatively low fuel costs which eliminated the cost justification for the somewhat higher capital costs of the higher efficiency cycles. However, in some international markets where fuel cost is a higher fraction of the total cost, the higher efficiency cycles offer a favorable cost-of-electricity comparison and provide lower emissions compared to a subcritical plant. This is true in both Europe (see Figure 6.5) and Japan. Supercritical

FIGURE 6.3 US coal basins.

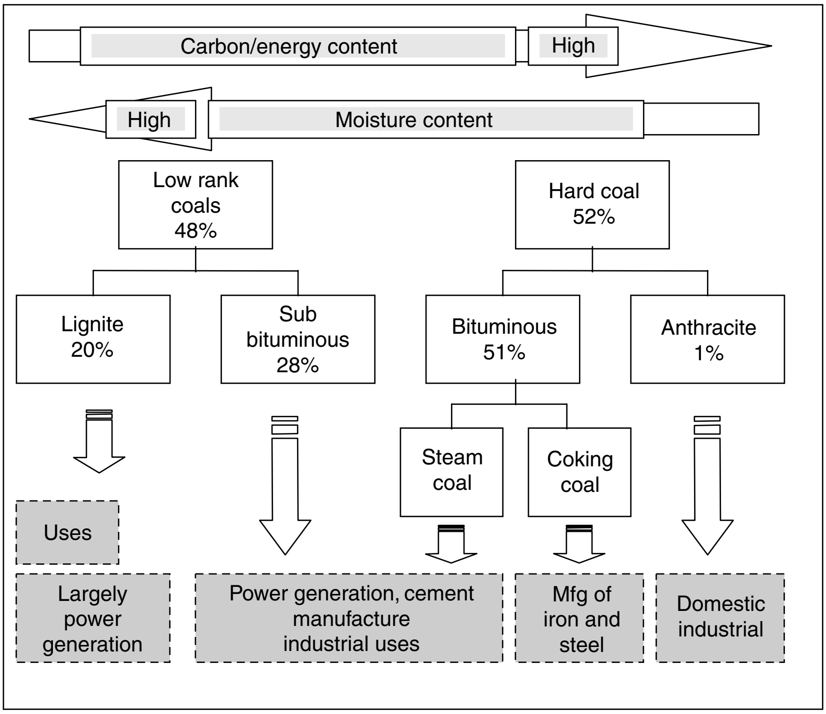


FIGURE 6.4 World use of coal.

cycles have recently been selected for several new fossil plants in the U.S. and the reduction of CO_2 emissions for the supercritical cycle could be a deciding factor as ways are sought to reduce global warming concerns.

6.1.5.2 Environmental Controls for Fossil-Steam Plants

Of all the hurdles facing owners of generating plants, perhaps none is greater than preparing units for meeting environmental limits at minimum cost. In the U.S. by the year 2000, more than 200 SO_2 scrubbers had been installed on more than 100,000 MW of fossil-steam capacity, valuable additions that will permit plants to operate in compliance for many more years. Typically a 450 MW coal-fired plant will emit 75 tn. of SO_2 per day without a scrubber and perhaps 8 tn. per day with a 90% FGD system in place, a difference that can be measured in terms of the market for SO_2 credits. And for NO_x, where most current control activities are focused the same plant might emit 10–35 tn. per day. NO_x control options range from burner optimization to the use of SCR. As for carbon dioxide, the above plant emits about 9000 tn./day at a plant efficiency of 38% which translates to 2452 tn. of carbon. Such emissions are of increasing concern and potential future carbon taxes must be considered. A combined cycle gas plant, for comparison, emits about half of this amount, per MWh, due to the higher plant efficiency and lower carbon content of natural gas. The removal of mercury from the products of fossil fuel combustion, will be an increasingly important task for fossil plant operators. Current plant testing for mercury control is exploring several promising options.

An overall perspective of emissions control technologies from a modern pulverized coal-fired plant is shown in Figure 6.6.

6.1.5.3 NO_x Control

The broad span of options for control may be categorized into combustion and post-combustion methods (Figure 6.7). Low NO_x burners, especially when combined with staging of over-fired air, are currently deployed most often. Combustion optimization techniques may offer a low cost alternative to hardware

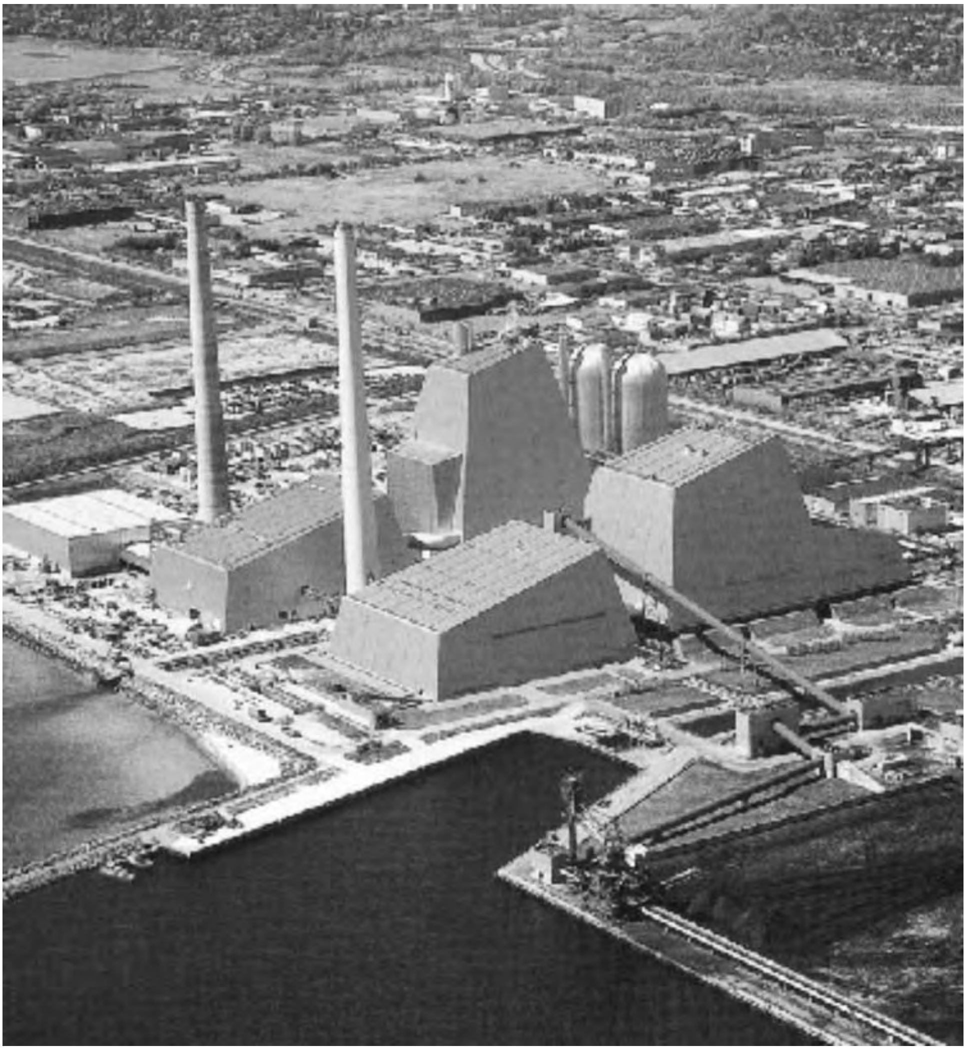


FIGURE 6.5 Modern supercritical coal-fired power plant.

options, particularly where modest reductions up to 30% are needed, and many units are now operating with some form of optimization of air and fuel flows, perhaps utilizing advanced flame diagnostics or software based on neural networks to do the optimizing.

6.1.5.4 Post-Combustion Options for NO_x Control

Selective catalytic reduction is used widely in Europe (especially Germany where it is installed on more than 30,000 MW of coal-fired boilers) and in Japan, and increasingly in the U.S. In an SCR, ammonia is injected into the boiler exhaust gases ahead of the catalyst bank (at about 550°F–750°F). NO_x and NH_3 then react to produce nitrogen and water, the chemical reactions being:

$$4NO + 4NH_3 + O_2 \rightarrow 4N_2 + 6H_2O$$

$$6NO_2 + 8NH_3 \rightarrow 7N_2 + 12H_2O$$

$$2NO_2 + 4NH_3 + O_2 \rightarrow 3N_2 + 6H_2O$$

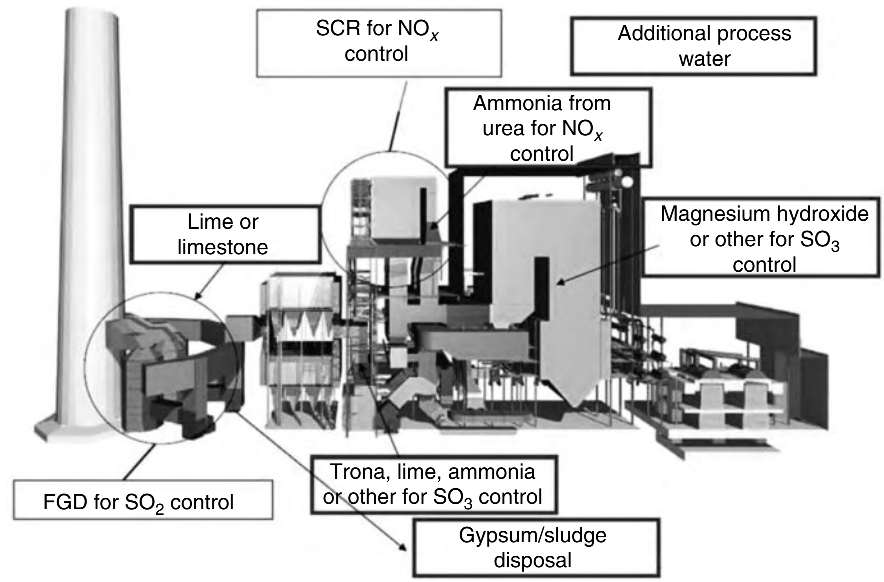


FIGURE 6.6 Emission controls from a modern pulverized coal-fired power plant.

The reaction can result in a potential NO_x removal capability of more than 90%. Retrofit installation of an SCR system can require considerable space, although the reactor can be placed inside the original ductwork if NO_x reduction levels are modest (Figure 6.8).

Selective non-catalytic reduction (SNCR) is a promising lower capital cost alternative to SCR ($10/kW versus more than $50/kW), but with lower performance (20%–35% reduction compared with 50 to as high as 80% for SCR). In SNCR, the injection of a reagent like urea or ammonia into the upper part of the furnace converts NO_x from combustion into nitrogen, this conversion being a direct function of furnace temperature and reagent injection rate.

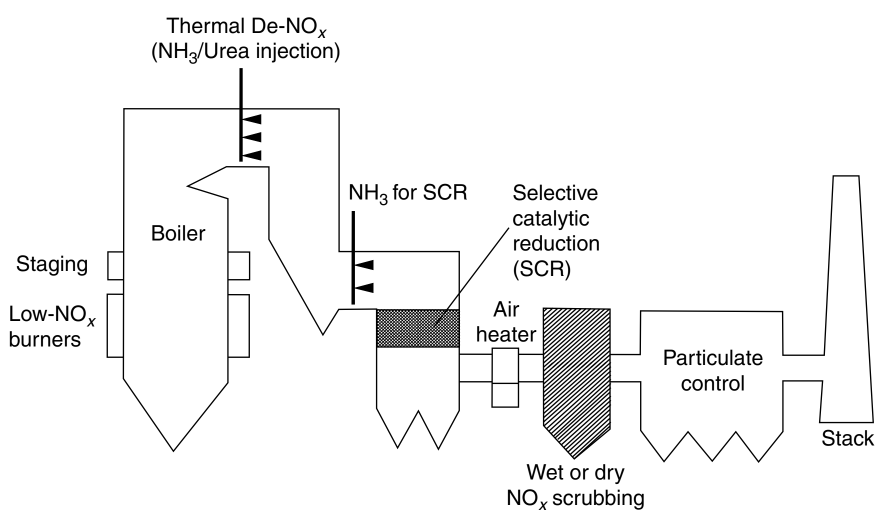


FIGURE 6.7 NO_x control options for fossil boilers.

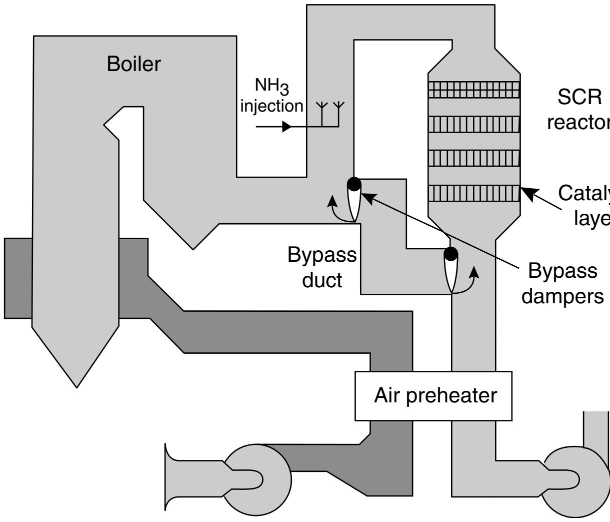


FIGURE 6.8 Selective catalytic reduction system in the boiler flue gas.

6.1.5.5 Sulfur Dioxide Control

The need for removal of sulfur dioxide from flue gases has led to the installation of FGD units in much of the coal fired capacity in the U.S. (Figure 6.9).

More than 200 coal-fired units in the United States use wet or dry scrubbing to remove sulfur. Of these, the majority use wet lime or limestone scrubbers with perhaps between 20 and 30 using dry scrubbing where the sulfur content of the coal is generally lower. Lime may be used alone or in combination with magnesium, carbides, or with alkaline fly ash if the boiler burns subbituminous coals or lignites.

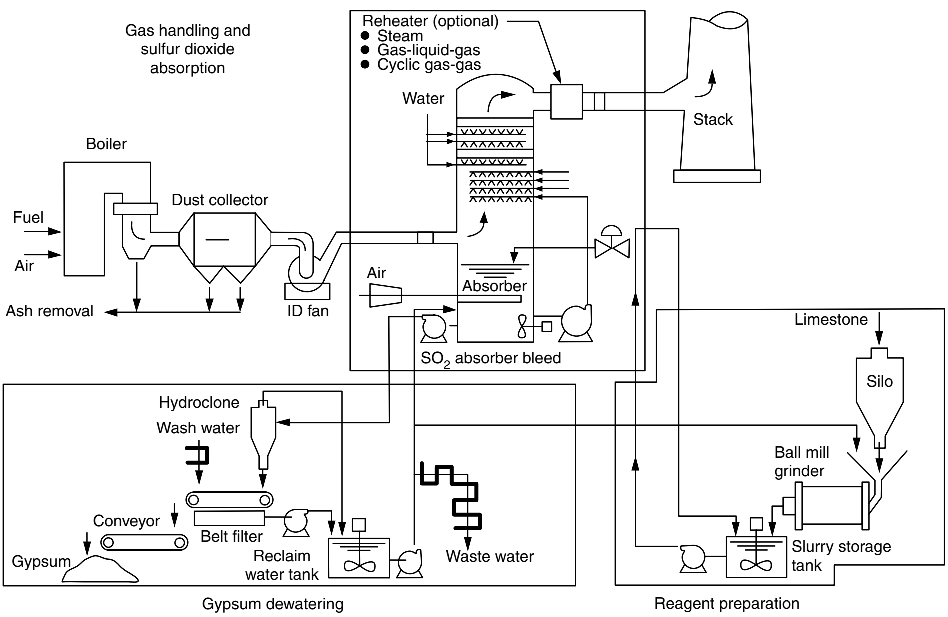


FIGURE 6.9 Wet limestone flue gas desulfurization system.

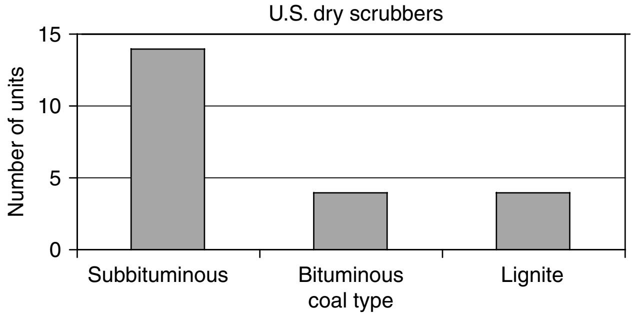


FIGURE 6.10 Use of spray dryers is mainly, though not exclusively, confined to western subbituminous, low sulfur coal-burning units.

The most common wet systems bring lime or limestone slurries into contact with flue gases in a spray tower. SO_2 in the flue gas is absorbed in the slurry and collected in a reaction tank, where it precipitates to produce calcium sulfite or calcium sulfate (gypsum) crystals. A portion of the slurry is then pumped to a thickener where the crystals settle out before going to a filter for final dewatering. Calcium sulfite and/or sulfate are typically disposed of in a landfill. Flue gas desulfurization gypsum can be sold for use in wallboard, cement or agricultural products.

Dry FGD systems employ calcium or sodium reagents that are either injected as dry powders or in spray drying systems as slurries that dry on contact with flue gas. Dry injection systems are most economical for space-constrained sites or applications that require only moderate emissions reductions. Spray drying systems, which can achieve higher SO_2 removal efficiencies, have mainly been applied at units burning low-to-medium-sulfur coals (Figure 6.10).

6.1.5.6 Mercury Control

Various approaches to mercury removal from power plant flue gases are under development though much mercury can be removed through existing air pollution controls for particulate and SO_2 (Figure 6.11).

The removal of mercury from coal-fired units can be accomplished in several ways. Coal cleaning before combustion can remove some mercury and other heavy metals. After combustion, the injection of a sorbent, such as a activated carbon can be very effective. Existing ESPs and SO_2 scubbers can capture from 20 to 60% of mercury. Catalysts and certain chemicals can be injected that oxidize elemental mercury to

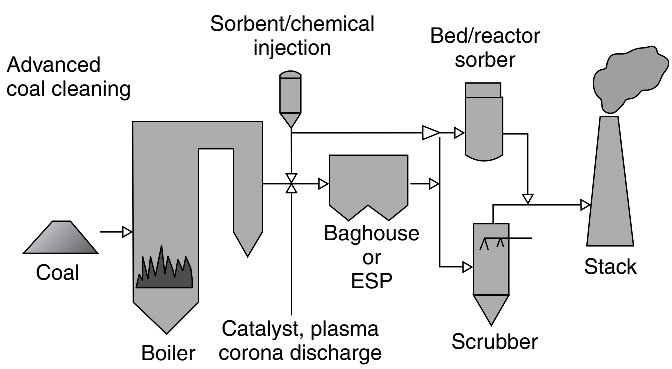


FIGURE 6.11 Options for the removal of mercury.

enhance scrubber capture. Fixed beds, coated with materials such as gold, can form amalgams with mercury.

6.1.6 Fluidized Bed Power Plants

Introduced nearly 30 years ago the atmospheric fluidized bed combustion (AFBC) boiler (Figure 6.12) has found growing application for power generation. From the first FBC boiler, generating 5000 lb/h of steam in 1967, the technology has matured to the 350 MW size units available today.

In the bubbling bed version of the AFBC, the fuel and inert matter, together with limestone or dolomite for SO_2 capture, is suspended through the action of fluidizing air, which flows at a velocity of 3–8 ft./s in essentially a one-pass system. Circulating fluid beds (CFB) differ from bubbling beds in that much of the bed material passes through a cyclone separator before being circulated back to the boiler. In-bed tubes are generally not used for CFB units permitting a much higher fluidizing velocity of 16–26 ft./s. Since the early AFBC designs, attention has been directed towards increasing unit efficiency, and reheat designs are now usual in large units. When SO_2 capture is important a key parameter is the ratio of calcium in the limestone to sulfur in coal. Typical calcium to sulfur ratios for 90% SO_2 reduction are in the range of 3.0–3.5 for bubbling beds and 2.0–2.5 for circulating beds. NO_x levels in AFBCs are inherently low and nominally less than 0.2 lb/MMBtu. It is important to note that for CFBs, boiler efficiencies can be as high as a PC unit. In fact designs now exist for AFBCs with supercritical steam conditions, with prospects for cycles up to 4500 psia, 1100°F with double reheat.

In North America more than 160 units now generate in excess of 9000 MW. Burning coal in a suspended bed with limestone or dolomite permits effective capture of sulfur and fuel flexibility allows a broad range of opportunity fuels. These fuels might include coal wastes (culm from anthracite, gob from bituminous coal), peat, petroleum coke, and a wide range of coals from bituminous to lignite. A low (1500°F) combustion temperature leads to low NO_x formation. Several large size FBC plants are now under construction, including one in Europe with supercritical steam conditions.

Examples of large size generating FBC plants in the Americas are shown in Table 6.1.

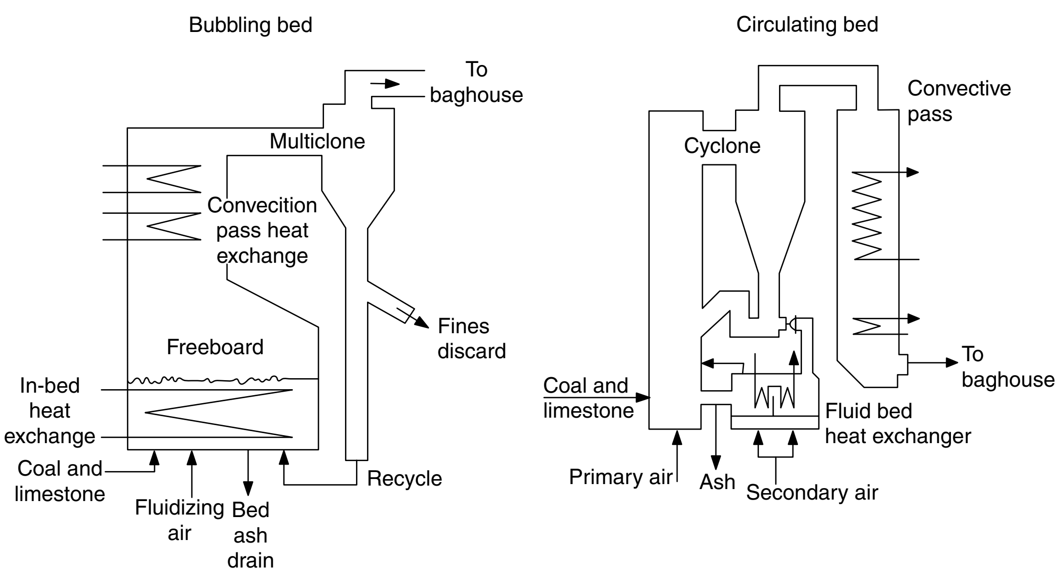


FIGURE 6.12 The primary types of atmospheric fluidized beds are bubbling beds and circulating beds.

TABLE 6.1 U.S., Canadian, and Latin American CFB Units Larger than 75 MW

Plant/Location (Vendor)	Start Up	Capacity, MW (Net)	Fuels
Tri-State Generation and Transmission/Colorado	1987	1×100	Bit. coal
AES Shady Point/Oklahoma	1989	4×75	Bit. coal
AES Thames/Connecticut	1989	2×90	Bit. coal
Schuylkill Energy/Pennsylvania	1989	1×80	Culm
ACE Cogeneration/California	1990	1×97	Low-S bit. Coal
Texas-New Mexico Power/Texas	1990	2×150	Lignite
AES Barbers Point/Hawaii	1992	2×90	Bit. coal
Nelson Industrial Steam Co. (NISCO)/Louisiana	1992	2×110	Coke
Cedar Bay Generating Co./Florida	1993	3×90	Bit. coal
Nova Scotia Power/Nova Scotia	1993	1×165	30% bit. coal and 70% coke
Colver Power/Pennsylvania	1995	1×105	Gob
Northampton Generating Co./Pennsylvania	1995	1×112	Culm
ADM/Illinois	1996/2000	2×132	Bit. coal and up to 5% TDF
ADM/Iowa	2000	1×132	Bit. coal
AES Warrior Run/Maryland	1999	1×180	Bit. coal
Choctaw Generation—the Red Hills project/Mississippi	2001	2×220	Lignite
Bay Shore Power—First Energy/Ohio	2001	1×180	Coke
AES Puerto Rico/Puerto Rico	2002	2×227	Bit. coal
JEA/Florida	2002	2×265	Bit. coal and coke
Southern Illinois Power Cooperative/Illinois	2002	1×113	Waste bit. coal
Termoelectrica del Golfo/Mexico	2002	2×115	Coke
Termoelectrica de Penoles/Mexico	2003	2×115	Coke
Reliant Energy Seward Station/Pennsylvania (ALSTOM)	2004	2×260	Gob and bit. coal
East Kentucky Power Cooperative/Kentucky	2004	1×268	Unwashed high-sulfur bit. coals
Figueira/Brazil	2004	1×128	Bit. coal

6.1.7 Coal Gasification Power Plants

There are currently two coal-based integrated gasification combined cycle (IGCC) commercial-sized demonstration plants operating in the U.S. and two in Europe (Table 6.2). The U.S. projects were supported under the U.S. Department of Energy's (DOE) Clean Coal Technology (CCT) demonstration program.

The 262 MW Wabash River IGCC repowering project in Indiana started up in October 1995 and uses the E-GAS™ (formerly Destec) gasification technology. The 250 MW Tampa Electric Company (TEC) IGCC project in Florida started up in September 1996 and is based on the Texaco (now GE) gasification

TABLE 6.2 Coal-Based, Commercial-Size IGCC Plants

	Gasification Technology	Plant Size, MW	Startup Date
Wabash River, Indiana, U.S.A.	Destec	262	10/95
Tampa Electric, Florida, U.S.A.	Texaco	250	9/96
SEP/Demkolec, Buggenum, The Netherlands	Shell	253	Early 1994
ELCOGAS, Puertollano, Spain	Krupp-Uhde Prenflo	310	12/97 on coal

FIGURE 6.13 The 250-MW coal gasification plant at the polk plant of Tampa Electric.

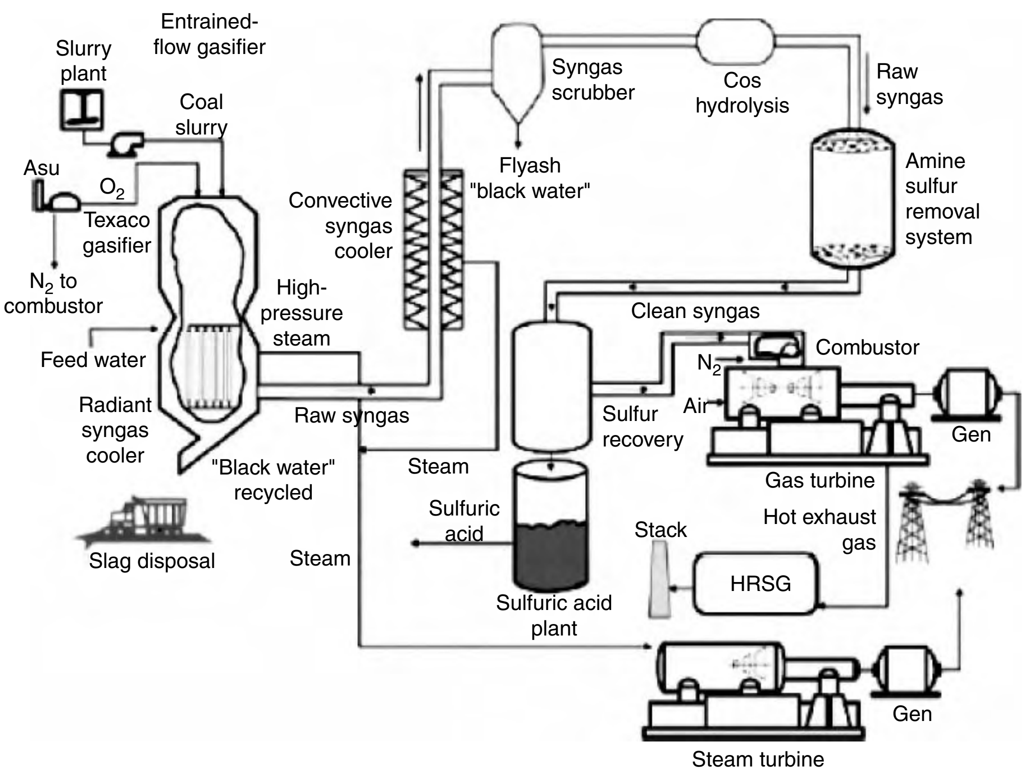


FIGURE 6.14 Tampa electric polk gasification plant.

technology (Figure 6.13). The first of the European IGCC plants the SEP/Demkolec project at Buggenum, the Netherlands, uses the Shell gasification technology and started operations in early 1994. The second European project, the ELCOGAS project in Puertollano, Spain, which uses the Prenflo gasification technology, started coal-based operations in December 1997. New commercial IGCC plants are now in planning in the U.S. and are scheduled for operation within the next 10 years.

6.1.7.1 Tampa Electric Integrated Gasification Combined Cycle Plant

In the DOE Clean Coal project at Tampa Electric coal/water slurry and oxygen are reacted at high temperature and pressure to produce approximately 245 Btu/SCF syngas (LHV) in a Texaco gasifier (Figure 6.14). Molten ash flows out of the bottom of the gasifier into a water-filled sump where it forms a solid slag. The syngas moves from the gasifier to a radiant syngas cooler and a convective syngas cooler (CSC), which cool the syngas while generating high-pressure steam. The cooled gases flow to a water-wash syngas scrubber for particulate removal. Next, a hydrolysis reactor converts carbonyl sulfide (COS) in the raw syngas to hydrogen sulfide (H_2S) that is more easily removed. The raw syngas is then further cooled before entering a conventional amine sulfur removal system and sulfuric acid plant (SAP). The cleaned gases are then reheated and routed to a combined cycle system for power generation. A GE MS 7001FA gas turbine generates 192 MWe. Thermal NO_x is controlled to 0.7 lb/MWh by injecting nitrogen. A steam turbine uses steam produced by cooling the syngas and superheated with the gas turbine exhaust gases in the HRSG to produce an additional 123 MWe. The air separation unit consumes 55 MW and auxiliaries require 10 MW, resulting in 250 MWe net power to the grid. The plant heat rate is 9650 Btu/kWh (HHV).

6.1.7.2 Wabash River Integrated Gasification Combined Cycle Plant

A second U.S. coal gasification plant is at the Cinergy Wabash River plant using a different gasification process from the Tampa Electric approach (Figure 6.15). The Destec, now E-Gas Technology™, process

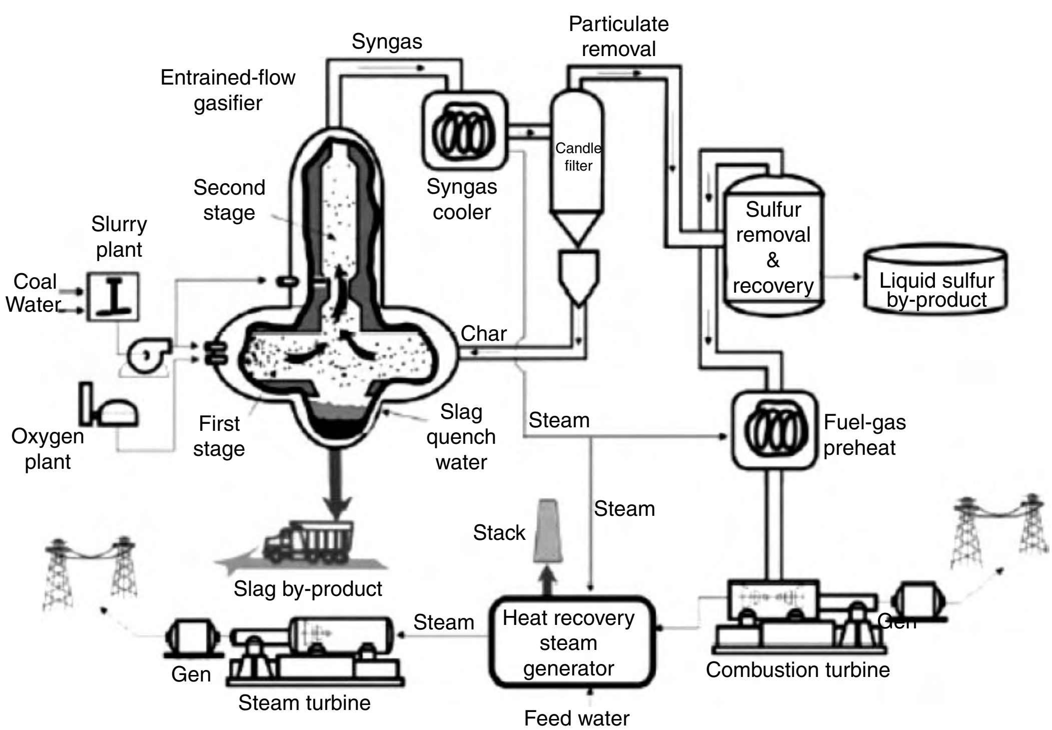


FIGURE 6.15 Sketch of a typical PWR power plant. (From World Nuclear Association, http://www.world-nuclear.org.)

features an oxygen-blown, continuous-slagging, two-stage, entrained flow gasifier. Coal is slurried, combined with 95% pure oxygen, and injected into the first stage of the gasifier, which operates at 2600°F/400 psig. In the first stage, the coal slurry undergoes a partial oxidation reaction at temperatures high enough to bring the coal's ash above its melting point. The fluid ash falls through a tap hole at the bottom of the first stage into a water quench, forming an inert vitreous slag. The syngas flows to the second stage, where additional coal slurry is injected. This coal is pyrolyzed in an endothermic reaction with the hot syngas to enhance syngas heating value and improve efficiency. The syngas then flows to the syngas cooler, essentially a fire tube steam generator, to produce high-pressure saturated steam. After cooling in the syngas cooler, particulates are removed in a hot/dry filter and recycled to the gasifier. The syngas is further cooled in a series of heat exchangers. The syngas is water-scrubbed to remove chlorides and passed through a catalyst that hydrolyzes carbonyl sulfide into hydrogen sulfide. Hydrogen sulfide is removed in the acid gas removal system using absorber/stripper columns. A Claus unit is used to produce elemental sulfur as a salable by-product. The "sweet" gas is then moisturized, preheated, and piped to the power block. The power block consists of a single 192-MWe General Electric MS 7001FA (Frame 7 FA) gas turbine, a Foster Wheeler single-drum heat recovery steam generator with reheat, and a 1952 vintage Westinghouse reheat steam turbine.

TABLE 6.3 Performance Comparisons of Four Clean Coal Technologies

Technology	Units	PC-Sub	PC-SC	AFBC	IGCC
Potential capacity	MW/unit	250–700	350–800	30–300	300–600
Fuel range/diversity		All grades lignite to anthracite	All grades lignite to anthracite	All grades biomass and wastes	All grades bituminous preferred
Build time	Years	3	3	3	3
Manning levels					
Operators	Number/MW	0.16	0.16	0.18	0.18
Total staff	Number/MW	0.31	0.31	0.46	0.33
Expected availability					
Planned outage	%	11.1	11.1	5.7	4.7
Forced outage	%	3.7	3.9	4.1	10.1
Equiv. availability	%	85.7	85.4	90.4	85.7
Expected efficiencies	% (HHV)	34.4–35.7	36.4–37.7	34.6–35.6	39.3–41.1
Expected heat rate	kJ/kWh (HHV)	9570–9930	9050–9380	9600–9870	8310–8680
Emission ranges					
SO_2	Kg/MW-h	0.66–0.68	0.62–0.65	0.66–0.68	0.04–0.22
NO_x	Kg/MW-h	0.66–0.68	0.62–0.65	0.66–0.68	0.23–0.24
CO_2	Kg/MW-h	831–862	786–815	834–857	723–754
Particulates	Kg/MW-h	0.10	0.10	0.10	0.01
Solid waste (Total)	Kg/MW-h	62–113	59–107	55–141	42–44
Ash	Kg/MW-h	45–48	45–48	47–44	42–44
Spent sorbent	Kg/MW-h	12–66	11–62	11–94	0
Cooling water requirements	Cu.M/h/MW	244	236	249	185
Flexibility load range	%	30–100	30–100	30–100 per unit	50–100 per unit 25–100 for 2 trains
Total plant cost	$/kW	1040–1080	1060–1090	1030–1060	1150–1190
O&M cost					
Fixed	$/Kw-yr.	23.9–24.3	23.9–24.4	22.3–22.7	25.6–26.3
Variable	$/MW-h	2–3.1	2–3	2.1–4.4	1.7–1.5
Cost of electricity	$/MW-h	36.2–38.4	36–38	36–39.2	37.1–38.2

6.1.8 Comparison of Clean Coal Options for the Next 10 Years

An overall comparison of the performance and cost of four clean coal technologies discussed above: PC subcritical and supercritical, fluidized bed, and integrated gasification combined cycle, is shown in Table 6.3. It should be recognized that comparisons of this type do not capture the variations in available fuel, site conditions, and operational and business needs that often determine the ultimate choice of technology.

6.1.9 Carbon Dioxide Control

Plant emissions of CO_2 shown in Table 6.3 are mainly a function of the plant heat rate and efficiency. For the combustion technologies (PC, AFBC) there are some minor additional contributions associated with the use of limestone.

It has been shown that if CO_2 removal was ever required that it is much less expensive to remove CO_2 syngas under pressure prior to combustion rather than from the boiler exit ducts of a PC plant where the CO_2 is in more dilute concentrations. Generally, when CO_2 control is deployed, the cost of electricity grows significantly. This is particularly true for PC plants where the CO_2 is at high volume and at atmospheric pressure. In contrast the CO_2 is compressed in gasification plants to less than 1% of the flue gas volume of a PC plant (as are all the gas emissions) and this requires less energy for removal. See Figure 6.16 for rough examples of electricity cost differences. Figure 6.16 assumes plant sizes of about 460 MW. Capacity factor (CF) is 80% except for the second natural gas plant which is 40% CF. Of course these costs also depend on design factors such as coal quality, steam conditions, and coal and gas prices.

Another important question when considering CO_2 control is the impact the additional control equipment will have on the power output and efficiency of the power plant. This is particularly significant when considering the retrofit of control equipment to an existing unit. There are more than 1500 large coal-fired plants in the U.S. that fall into this category. Estimates of the impact of 100% CO_2 control have been made for supercritical PC cycles, natural gas combined cycles, and integrated gasification combined cycles. In this respect the IGCC plant has very significant advantages as shown in Table 6.4. As of the end of 2006, these estimates are still under review as new technology for CO_2 removal suggests that significant reductions in these penalties may be possible.

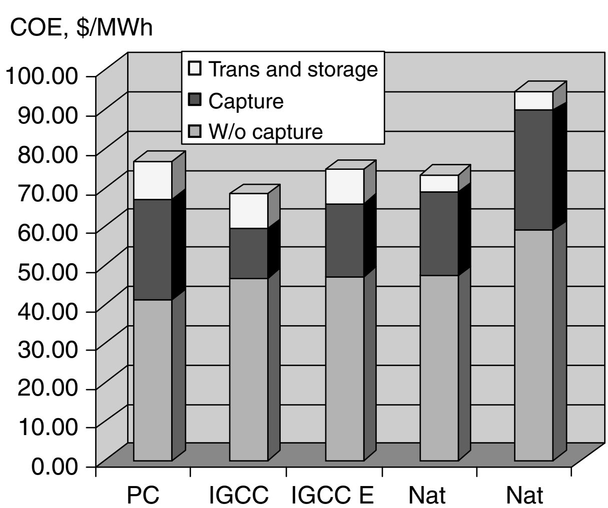


FIGURE 6.16 Cost of electricity with CO_2 removal and storage from fossil-fuel power plants for a PC unit, two gasification units, and two natural gas units. Coal cost is assumed at \$1.5/MBtu and natural gas at \$5/MBtu.

TABLE 6.4 Estimated Costs for CO_2 Removal

	Output Penalty (%)	Efficiency Penalty	$/$CO_2$ tonne
NGCC	20	−10% pts	60
IGCC	5	−6% pts	18
PC	28	−12% pts	43

6.1.10 Conclusions

Essentially all suggested energy scenarios covering the next 20 years include the continued reliance on large central generating plants, with coal, natural gas, and nuclear fuel as the predominant energy sources. Nearly 800 GW of electric power for the U.S. in 2020 is not going to be possible in this time frame without a substantial number of large unit sizes in the 500 MW+ range. Further, in the light of fuel availability and cost forecasts, most industry observers foresee large coal and nuclear plants essential if the U.S. is to maintain its electrical energy supply and independence in the near term. It is important to note also the question of carbon dioxide control for fossil plants, that use of the nuclear option mitigates.

It is clear though that coal remains a crucial fuel resource for both U.S. and worldwide growth. In Europe and North America, natural gas-fired combustion turbines have been recently favored for new and repowered units, but it appears unlikely that the natural gas supply can meet all the new capacity demand and also be the major replacement fuel for the existing fleet. Substantial recent increases in natural gas prices have thus renewed interest in coal-based generation. Coal's sustained viability beyond the next 20 years will clearly depend on new technologies to improve efficiency, cost, and plant emissions.

It is also clear that, in small sizes, solar, wind, biomass, and other renewable technologies offer very attractive opportunities for cleaner energy sources to reduce mankind's environmental impact. State renewable energy programs are valuable in moving renewables into the market, while increased R&D can be a potent force in driving up the efficiency of these new technologies, simultaneously driving down costs and making them competitive in world markets.

It is clear that regulatory policies, siting-related issues, fuel availability, and risk assessment by financiers and owners will determine which technologies will succeed, and on what timetable.

In a broader perspective, the defining challenge of the coming century will be to balance the "trilemma" of interlocking sustainability issues—the economic aspirations of rapidly expanding populations in the developing world, environmental quality, and natural resource availability. Technology innovation is the best prospect for resolving conflicts between population, prosperity, and pollution. With much advanced technology now mature and ready for application in other parts of the globe, the developing world can leapfrog the slow technology development process and enjoy a vastly improved quality of life while further advances are pursued.

References

1. Generation technologies through the year 2005, CRC Handbook of Energy Efficiency, CRC Press Inc., 1997.
2. Coal Power for Progress, 2000. *4th Ed.* World Coal Institute.
3. The Electricity Technology Roadmap: Powering Progress, 1999 Summary and Synthesis, EPRI Report 1006315, C1-112677-V1.
4. Armor, A. F. and Wolk, R. H. 2002. *Productivity Improvement Handbook for Fossil Steam Plants, 3rd Ed.* EPRI Report 1006315.
5. Armor, A. F. and Wolk, R. H. 2005. *Productivity Improvement for Fossil steam Power Plants 2005: One Hundred Case Studies*. EPRI Report 1012098.
6. EPRI. 2000. *An Assessment of Mercury Emissions from U.S. Coal-Fired Power Plants*. EPRI Report 1000608.
7. Dalton, S. M., Viswanathan, R., Gehl, S. M., Armor, A. F., and Purgert, R. 2001. *Ultrasupercritical Materials*. DOE Clean Coal Conference, Washington, DC.

8. Armor, A. F., Viswanathan, R., and Dalton, S. M. 2003. *Ultrasupercritical Steam Turbines: Design and Materials Issues for the Next Generation, Coal Utilization and Fuel Systems*. DOE Annual Conference, Clearwater, FL.
9. DOE Fossil Energy-Tomorrow's Turbines, 2001. http://fossil.energy.gov/coal_power/turbines/index.shtml, April 30.
10. EPRI. 2002. *Atmospheric Fluidized—Bed Combustion Handbook*. EPRI Report 1004493.
11. Skowyra, R. S. et al. 1995. Design of a supercritical sliding pressure circulating fluidized bed boiler with vertical waterwalls. In *Proceedings of 13th International Conference on Fluidized Bed Combustion*. ASME, New York.
12. Technical Status. 2002. *Operating Experience, and Risk Assessment of Clean Coal Technologies-2002*. EPRI Report 1004480.
13. Courtright, H. A., Armor, A. F., Holt, N. H., and Dalton, S. M. 2003. *Clean Coal Technologies and their Commercial Development*, POWER-GEN International, Conference Proceedings, Las Vegas, NV.
14. EPRI. 2004. *Decommissioning Handbook for Coal-Fired Power Plants*. EPRI Report 1011220.
15. U.S. Department of Energy. 2002. *Clean Coal Technology Demonstration Program*, DOE/FE-0444.

For Further Information

Annual Energy Outlook, 2003, Energy Information Administration: www.eia.doe.gov
National Engineering Technology Laboratory: www.netl.doe.gov
EPRI: www.epri.com
Steam, Its Generation and Use. Babcock and Wilcox, New York.
Combustion: Fossil Power Systems. Combustion Engineering, Inc., Windsor, CT.
Tapping global expertise in coal technology. *EPRI Journal*, Jan/February, 1986.
IGCC: new fuels, new players. *EPRI Journal*, July/August 1994.
A brighter future for PFBC. *EPRI Journal*, December 1993.
Fuel cells for urban power. *EPRI Journal*, September 1991.
Distributed generation. *ERPI Journal*, April/May 1993.
Plant repowering. *EPRI Journal*, September/October 1995.
Smart materials. *EPRI Journal*, July/August 1998.
Merchant plants. *EPRI Journal*, Summer 1999.
Energy and air emissions. *EPRI Journal*, Summer 2000.
Global coal initiative. *EPRI Journal*, Summer 2001.

6.2 Nuclear Power Technologies

Edwin A. Harvego and Kenneth D. Kok

6.2.1 Introduction

Nuclear power is derived from the fission of heavy element nuclei or the fusion of light element nuclei. This chapter will discuss nuclear power derived from the fission process because fusion as a practical power source will not reach the stage of commercial development in the next 20–25 years. In a nuclear reactor, the energy available from the fission process is captured as heat that is transferred to working fluids that are used to generate electricity. Uranium-235 (^{235}U) is the primary fissile fuel currently used in nuclear power plants. It is an isotope of uranium that occurs naturally at about 0.72% of all natural uranium deposits. When ^{235}U is "burned" (fissioned) in a reactor, it provides about one megawatt day of energy for each gram of ^{235}U fissioned (3.71×10^{10} Btu/lb).

Nuclear power technology includes not only the nuclear power plants that produce electric power but also the entire nuclear fuel cycle. Nuclear power begins with the mining of uranium. The ore is processed and converted to a form that can be enriched in the ^{235}U isotope so that it can be used efficiently in today's light-water-moderated reactors. The reactor fuel is then fabricated into appropriate fuel forms for use in nuclear power plants. Spent fuel can then be either reprocessed or stored for future disposition. Radioactive waste materials are generated in all of these operations and must be disposed of. The transportation of these materials is also a critical part of the nuclear fuel cycle.

In this chapter, the development, current use, and future of nuclear power will be discussed. The second section of this chapter is a brief review of the development of nuclear energy as a source for production of electric power, and looks at nuclear power as it is deployed today both in the United States and worldwide. The third section examines the next generation of nuclear power plants that will be built. The fourth section reviews concepts being proposed for a new generation of nuclear power plants. The fifth section describes the nuclear fuel cycle, beginning with the availability of fuel materials and ending with a discussion of fuel reprocessing technologies. The sixth section discusses nuclear waste and the options for its management. The seventh section addresses nuclear power economics. Conclusions are presented in Section 6.2.8.

6.2.2 Development of Current Power-Reactor Technologies

The development of nuclear reactors for power production began following World War II when engineers and scientists involved in the development of the atomic bomb recognized that controlled nuclear chain reactions could provide an excellent source of heat for the production of electricity. Early research on a variety of reactor concepts culminated in President Eisenhower's 1953 address to the United Nations in which he gave his famous "Atoms for Peace" speech, in which he pledged the United States "to find the way by which the miraculous inventiveness of man shall not be dedicated to his death, but consecrated to his life." In 1954, President Eisenhower signed the 1954 Atomic Energy Act that fostered the cooperative development of nuclear energy by the Atomic Energy Commission (AEC) and private industry. This marked the beginning of the commercial nuclear power program in the United States.

The world's first large-scale nuclear power plant was the Shippingport Atomic Power Station in Pennsylvania, which began operation in 1957. This reactor was a pressurized-water reactor (PWR) nuclear power plant designed and built by the Westinghouse Electric Company and operated by the Duquesne Light Company. The plant produced 68 MWe and 231 MWt.

The first commercial-size boiling-water reactor (BWR) was the Dresden Nuclear Power Plant that began operation in 1960. This 200 MWe plant was owned by the Commonwealth Edison Company and was built by the General Electric Company at Dresden, Illinois, about 50 miles southwest of Chicago.

Although other reactor concepts, including heavy-water-moderated, gas-cooled and liquid-metal-cooled reactors, have been successfully operated, the PWR and BWR reactor designs have dominated the commercial nuclear power market, particularly in the U.S. These commercial power plants rapidly increased in size from the tens of MWe generating capacity to over 1000 MWe. Today, nuclear power plants are operating in 33 countries. The following section presents the current status of nuclear power plants operating or under construction around the world.

6.2.2.1 Current Nuclear Power Plants Worldwide

At the end of 2004 there were 439 individual nuclear power reactors operating throughout the world. More than half of these nuclear reactors are PWRs. The distribution of current reactors by type is listed in Table 6.5. As shown in Table 6.5, there are six types of reactors currently used for electricity generation throughout the world. The following sections provide a more detailed description of the different reactor types shown in the table.

6.2.2.2 Pressurized-Water Reactors

Pressurized-water reactors represent the largest number of reactors used to generate electricity throughout the world. They range in size from about 400–1500 MWe. The PWR shown in Figure 6.17 consists of

TABLE 6.5 Nuclear Power Units by Reactor Type, Worldwide

Reactor Type	Main Countries	# Units Operational	GWe	Fuel
Pressurized light-water reactors (PWR)	U.S., France, Japan, Russia	263	237	Enriched UO_2
Boiling light-water reactors (BWR and AWBR)	U.S., Japan, Sweden	92	81	Enriched UO_2
Pressurized heavy-water reactors—CANDU (PHWR)	Canada	38	19	Natural UO_2
Gas-cooled reactors (Magnox & AGR)	U.K.	26	11	Natural U (metal), enriched UO_2
Graphite-moderated light-water reactors (RBMK)	Russia	17	13	Enriched UO_2
Liquid-metal-cooled fast-breeder reactors (LMFBR)	Japan, France, Russia	3	1	PuO_2 and UO_2
		439	362	

Source: Information taken from World Nuclear Association Information Paper "Nuclear Power Reactors".

a reactor core that is contained within a pressure vessel and is cooled by water under high pressure. The nuclear fuel in the core consists of uranium dioxide fuel pellets enclosed in zircaloy rods that are held together in fuel assemblies. There are 200–300 rods in an assembly and 100–200 fuel assemblies in the reactor core. The rods are arranged vertically and contain 80–100 tons of enriched uranium.

The pressurized water at 315°C is circulated to the steam generators. The steam generator is a tube-and shell-type of heat exchanger with the heated high-pressure water circulating through the tubes. The steam generator isolates the radioactive reactor cooling water from the steam that turns the turbine generator. Water enters the steam generator shell side and is boiled to produce steam that is used to turn

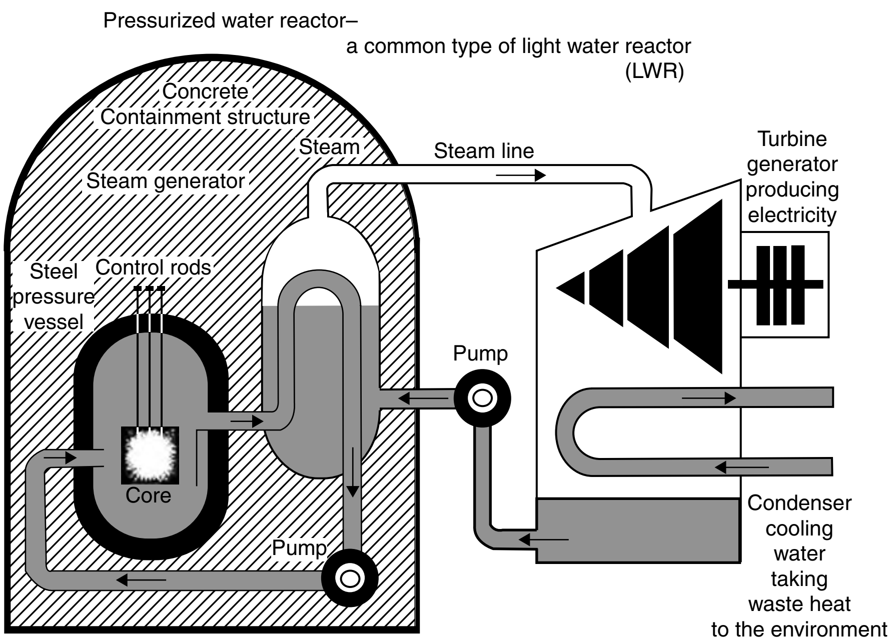


FIGURE 6.17 Sketch of a typical PWR power plant. (From World Nuclear Association, http://www.world-nuclear.org.)

the turbine generator producing electricity. The pressure vessel containing the reactor core and the steam generators are located in the reactor containment structure. The steam leaving the turbine is condensed in a condenser and returned to the steam generator. The condenser cooling water is circulated to cooling towers where it is cooled by evaporation. The cooling towers are often pictured as an identifying feature of a nuclear power plant.

6.2.2.3 Boiling-Water Reactors

The BWR power plants represent the second-largest number of reactors used for generating electricity. The BWRs range in size from 400 to 1200 MWe. The BWR, shown in Figure 6.18, consists of a reactor core located in a reactor vessel that is cooled by circulating water. The cooling water is heated to 285°C in the reactor vessel and the resulting steam is sent directly to the turbine generators. There is no secondary loop as there is in the PWR. The reactor vessel is contained in the reactor building. The steam leaving the turbine is condensed in a condenser and returned to the reactor vessel. The condenser cooling water is circulated to the cooling towers where it is cooled by evaporation.

6.2.2.4 Pressurized Heavy-Water Reactor

The so-called CANDU reactor was developed in Canada beginning in the 1950s. It consists of a large tank called a calandria containing the heavy-water moderator. The tank is penetrated horizontally by pressure tubes that contain the reactor fuel assemblies. Pressurized heavy water is passed over the fuel and heated to 290°C. As in the PWR, this pressurized water is circulated to a steam generator where light water is boiled, thereby forming the steam used to drive the turbine generators.

The pressure-tube design allows the CANDU reactor to be refueled while it is in operation. A single pressure tube can be isolated and the fuel can be removed and replaced while the reactor continues to operate. The heavy water in the calandria is also circulated and heat is recovered from it. The CANDU reactor is shown in Figure 6.19.

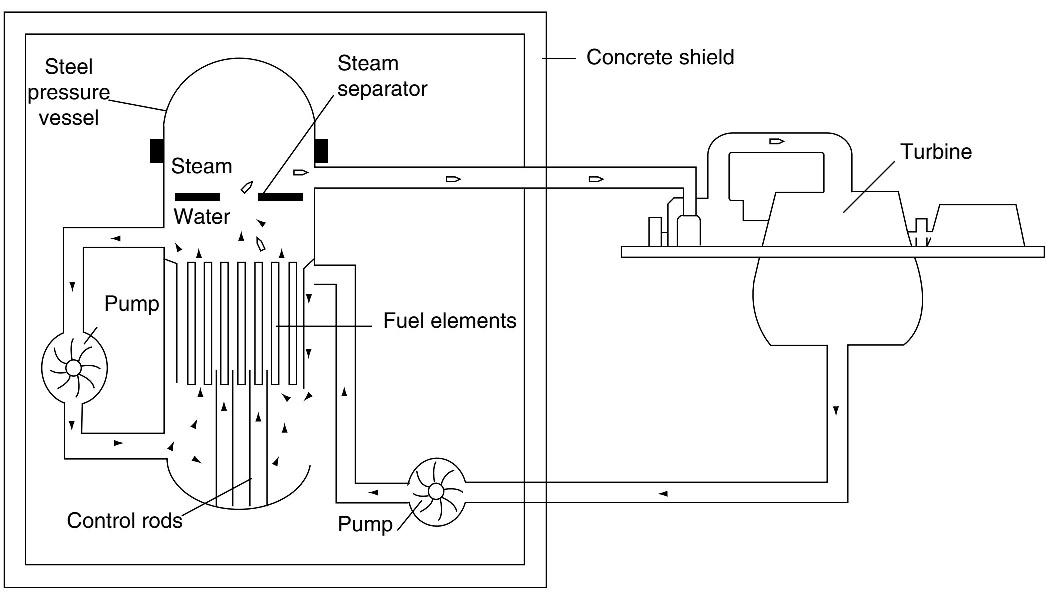


FIGURE 6.18 Sketch of a typical BWR power plant. (From World Nuclear Association, http://www.world-nuclear.org.)

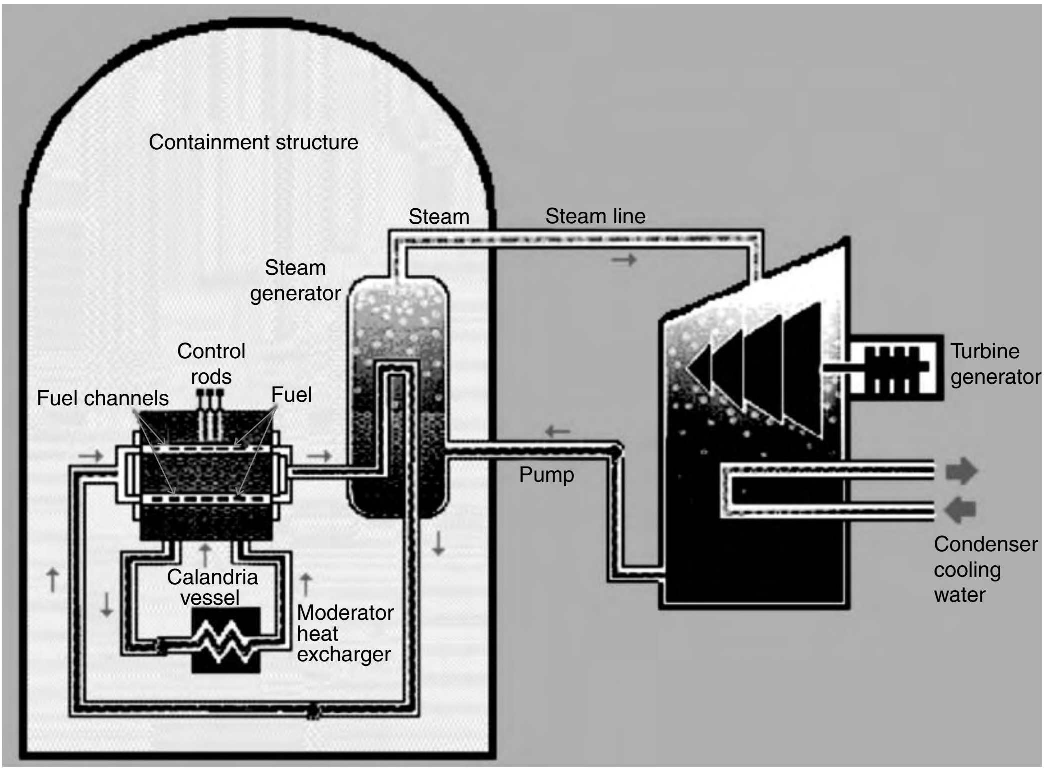


FIGURE 6.19 Sketch of a typical CANDU reactor power station. (From World Nuclear Association, http://www.world-nuclear.org.)

6.2.2.5 Gas-Cooled Reactors

Gas-cooled reactors were developed and implemented in the U.K. The first generation of these reactors was called Magnox, followed by the advanced gas-cooled reactor (AGR). These reactors are graphite moderated and cooled by CO_2. The Magnox reactors are fueled with uranium metal fuel, whereas the AGRs use enriched UO_2 as the fuel material. The CO_2 coolant is circulated through the reactor core and then to a steam generator. The reactor and the steam generators are located in a concrete pressure vessel. As with the other reactor designs, the steam is used to turn the turbine generator to produce electricity. Figure 6.20 shows the configuration for a typical gas-cooled reactor design.

6.2.2.6 Other Power Reactors

The remaining reactors listed in Table 6.5 are the light-water graphite-moderated reactors used in Russia, and the liquid-metal-cooled fast-breeder reactors (LMFBRs) in Japan, France, and Russia. In the light-water graphite-moderated reactors, the fuel is contained in vertical pressure tubes where the cooling water is allowed to boil at 290°C and the resulting steam is circulated to the turbine generator system as it is in a BWR. In the case of the LMFBR, sodium is used as the coolant and a secondary sodium cooling loop is used to provide heat to the steam generator.

6.2.2.7 Growth of Nuclear Power

The growth of nuclear power generation is being influenced by three primary factors. These factors are: (1) current plants are being modified to increase their generating capacity, (2) the life of old plants is being

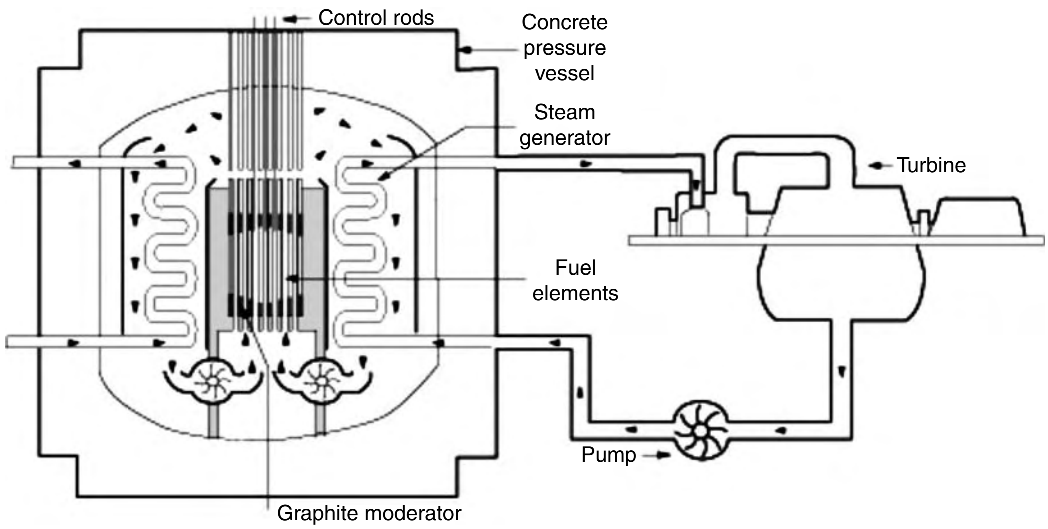


FIGURE 6.20 Sketch of a typical gasd-cooled reactor power station. (From World Nuclear Association, http://www.world-nuclear.org.)

lengthened by life-extension practices that include relicensing, and (3) new construction is adding to the number of plants operating worldwide. According to the IAEA, in May 2005, there were 440 nuclear power plants in operation with a total net installed capacity of 367 GWe. They now anticipate that 60 new plants will be constructed in the next 15 years, increasing the installed capacity to 430 GWe by 2020.

6.2.2.7.1 *Increased Capacity*

Operating nuclear plants are being modified to increase their generating capacity. Reactors in the U.S., Belgium, Sweden, Germany, Switzerland, Spain, and Finland are being uprated. In the U.S., 96 reactors have been uprated since 1977, with some of them having capacity increased up to 20%. The number of operating reactors in the U.S. peaked in 1991 with a gross electrical generation of over 70,000 MW-years; however, in 2003, the net electrical generation approached 90,000 MW-years from six fewer reactors. The generating capacity increase was due to both power uprating and improvements in operation and maintenance practices to produce higher plant availability. Switzerland increased the capacity of its plants by over 12%, whereas in Spain, uprating has added 11% to that country's nuclear capacity. The uprating process has proven to be a very cost effective way to increase overall power production capacity while avoiding the high capital cost of new construction.

6.2.2.7.2 *Plant-Life Extension*

Life extension is the process by which the life of operating reactors is increased beyond the original planned and licensed life. Most reactors were originally designed and licensed for an operational life of 40 years. Without life extension, many of the reactors that were built in the 1970 s and 1980 s would reach the end of their operational lives during the years 2010–2030. If they were not replaced with new plant construction, there would be a significant decrease in nuclear-based electricity generation as these plants reached the end of their useful life.

Engineering assessments of current nuclear plants have shown that they are able to operate for longer than their original planned and licensed lifetime. Fifteen plants in the U.S. have been granted 20-year extensions to their operating licenses by the U.S. Nuclear Regulatory Commission (NRC). The operators of most of the remaining plants are also expected to apply for license extensions. This will give the plants an operating life of 60 years. In Japan, operating lifetimes of 70 years are envisaged.

The oldest nuclear power stations in the world were operated in Great Britain. Chalder Hall and Chaplecross were built in the 1950 s and were expected to operate for 20–25 years. They were authorized to operate for 50 years, but were shut down in 2003 and 2004 for economic reasons. In 2000, the Russian government extended the lives of their 12 oldest reactors by 15 years for a total of 45 years.

Although life extension has become the norm throughout the world, many reactors have been shut down due to economic, regulatory, and political reasons. Many of these reactors were built early in the development of nuclear power. They tended to be smaller in size and were originally built for demonstration purposes. However, the political and regulatory process in some countries has led to the termination of nuclear power programs and the shutdown of viable reactor plants.

6.2.2.7.3 New Nuclear Plant Construction

New nuclear power plants are currently being constructed in several countries. The majority of the new construction is in Asia. Plants currently under construction are listed in Table 6.6.

6.2.3 Next-Generation Technologies

The next generation, generation-III nuclear power reactors, are being developed to meet power production needs throughout the world. These reactors incorporate the lessons that have been learned by operation of

TABLE 6.6 Power Reactors under Construction

Start Operation	Country, Organization	Reactor	Type	MWe (net)
2005	Japan, Tohoku	Higashidori 1	BWR	1067
2005	India, NPCIL	Tarapur 4	PHWR	490
2005	China, CNNC	Tianwan 1	PWR	950
2005	Ukraine, Energoatom	Khmelnitski	PWR	950
2005	Russia, Rosenergoatom	Kalinin 3	PWR	950
2006	Iran, AEOI	Bushehr 1	PWR	950
2006	Japan, Hokuriku	Shika 2	ABWR	1315
2006	India, NPCIL	Tarapur 3	PHWR	490
2006	China, CNNC	Tianwan 2	PWR	950
2006	China, Taipower	Lungmen 1	ABWR	1300
2007	India, NPCIL	Rawatbhata 5	PHWR	202
2007	Romania, SNN	Cernavoda 2	PHWR	650
2007	India, NPCIL	Kudankulam 1	PWR	950
2007	India, NPCIL	Kaiga 3	PHWR	202
2007	India, NPCIL	Kaiga 4	PHWR	202
2007	USA, TVA	Browns Ferry 1	BWR	1065
2007	China, Taipower	Lungmen 2	ABWR	1300
2008	India, NPCIL	Kudankulam 2	PWR	950
2008	India, NPCIL	Rawatbhata 6	PHWR	202
2008	Russia, Rosenergoatom	Volgodonsk-2	PWR	950
2008	Korea, KHNP	Shin Kori 1	PWR	950
2009	Finland, TVO	Oikiluoto 3	PWR	1600
2009	Japan, Hokkaido	Tomari 3	PWR	912
2009	Korea, KHNP	Shin Kori 2	PWR	950
2009	Korea, KHNP	Shin Wolsong 1	PWR	950
2010	Russia, Rosenergoatom	Balakovo 5	PWR	950
2010	Russia, Rosenergoatom	Kalinin 4	PWR	950
2010	India, NPCIL	Kalpakkam	FBR	440
2010	Pakistan, PAEC	Chashma 2	PWR	300
2010	Korea, KHNP	Shin Wolsong 2	PWR	950
2010	North Korea, KEDO	Sinpo 1	PWR(KSNP)	950
2010	China, Guangdong	Lingao 3	PWR	950
2010	Russia, Rosenergoatom	Beloyarsk 4	FBR	750
2011	China, Guangdong	Lingao 4	PWR	950
2011	China, CNNC	Sanmen 1 and 2	PWR	?
2011	China, CNNC	Yangiang 1 and 2	PWR	?

Source: From World Nuclear Association, Plans for New Reactors Worldwide, http://www.world-nuclear.org, 2005.

nuclear power systems since the 1950 s. The reactors are designed to be safer, more economical, and more fuel efficient. The first of these reactors were built in Japan and began operation in 1996.

The biggest change in the generation-III reactors is the addition of passive safety systems. Earlier reactors relied heavily on operator actions to deal with a variety of operational upset conditions or abnormal events. The advanced reactors incorporate passive or inherent safety systems that do not require operator intervention in the case of a malfunction. These systems rely on such things as gravity, natural convection, or resistance to high temperatures.

Generation-III reactors also have:

- Standardized designs with many modules of the reactor being factory constructed and delivered to the construction site leading to expedited licensing, reduction of capital cost and reduced construction time
- Simpler designs with fewer components that are more rugged, easier to operate, and less vulnerable to operational upsets
- Longer operating lives of 60 years and designed for higher availability
- Reduced probability of accidents leading to core damage
- Higher fuel burnup reducing refueling outages and increasing fuel utilization with less waste produced

The following sections describe the different types of generation-III reactors being developed worldwide.

6.2.3.1 Light-Water Reactors

Generation-III advanced light-water reactors are being developed in several countries. These will be described below on a country by country basis.

6.2.3.1.1 United States

Even though no new reactors are being built in the U.S., U.S. companies have continued to design advanced systems in anticipation of sales both in the U.S. and other parts of the world. In the U.S., the commercial nuclear industry in conjunction with the U.S. Department of Energy (DOE) has developed four advanced light-water reactor designs.

Two of these are based on experience obtained from operating reactors in the U.S., Japan, and western Europe. These reactors will operate in the 1300-MW range. One of the designs is the advanced boiling-water reactor (ABWR). This reactor was designed in the U.S. and is already being constructed and operated in Asia. The NRC gave final design certification to the ABWR in 1997. It was noted that the design exceeded NRC "safety goals by several orders of magnitude." The other type, designated System 80+, is an advanced PWR. This reactor system was ready for commercialization, but the sale of this design is not being pursued.

The AP-600 (AP = advanced passive), designed by Westinghouse, was the second reactor system to receive NRC certification. The certification came in 1999. The reactor is designed with passive safety features that result in projected core damage frequencies nearly 1000 times less than current NRC licensing requirements.

The Westinghouse AP-1000 (a scaled up version of the AP-600) received final design approval from the NRC and is scheduled for full design certification in 2005. The passive safety systems in this reactor design lead to a large reduction in components including 50% fewer valves, 35% fewer pumps, 80% less pipe, 45% less seismic building volume, and 70% less cable.

Another aspect of the AP-1000 is the construction process. After the plant is ordered, the plant will be constructed in a modular fashion, with modules being fabricated in a factory setting and then transported to the reactor site. The anticipated design construction time for the plant is 36 months. The construction cost of an AP-1000 is expected to be $1200/kW and the generating costs are postulated to be less then 3.5 cents/kWh. The plant is designed to have a 60-year operating life. China, Europe, and the U.S. are considering purchases of the AP-1000.

General Electric has created a modification of the ABWR for the European market. The European simplified BWR is a 1300 MWe reactor with passive safety systems. It is now called the economic and simplified boiling-water reactor (ESBWR). General Electric has a 1500-MWe version of this reactor in the preapplication stage for design certification by the NRC.

An international project being led by Westinghouse is designing a modular 335-MWe reactor known as the international reactor innovative & secure (IRIS). This PWR is being designed with integral steam generators and a primary cooling system that are all contained in the reactor pressure vessel. The goal of this system is to reach an eight-year refueling cycle using 10% enriched fuel with an 80,000-MWd/t burn-up. U.S. Nuclear Regulatory Commission design certification of this plant is anticipated by 2010.

6.2.3.1.2 Japan

Japan has three operating ABWRs. The first two, Kashiwazaki Kariwa-6 and Kashiwazaki Kariwa-7, began operation in 1996, and the third, Hamaoka-5, started up in 2004. These plants are expected to have a 60-year life and produce power at about $0.07/kWh. Several of these plants are under construction in Japan and Taiwan.

Hitachi has completed systems design of three additional ABWRs. These are rated at 600, 900, and 1700 MWe and are based on the design of the 1350-MWe plants. The smaller versions are designed with standardized components that will allow construction times on the order of 34 months.

Westinghouse and Mitsubishi, in conjunction with four utilities, are developing a large, 1500-MWe advanced PWR. This design will have both active and passive cooling systems and will have a higher fuel burn-up of 55 GWd/t of fuel. Mitsubishi is also participating with Westinghouse in the design of the AP-1000.

6.2.3.1.3 South Korea

The South Koreans have the APR-1400 system that evolved from the U.S. System 80+ and is known as the Korean next generation reactor. The first of these will be Shin-Kori-3 and Shin-Kori-4. Capital cost for the first systems is estimated to be $1400/kW with future plants coming in at $1200/kW with a 48-month construction time.

6.2.3.1.4 Europe

Four designs are being developed in Europe to meet the European utility requirements that were derived from French and German requirements. These systems have stringent safety requirements.

Framatome ANP has designed a large (1600–1750 MWe) European pressurized-water reactor (EPR). This reactor is the new standard design in France and it received design approval in 2004. The first of these units is scheduled to be built at Olkiluoto in Finland and the second at Flamanville in France. It is capable of operating in a load following manner and will have a fuel burn-up of 65 GWd/t. It has the highest thermal efficiency of any light-water reactor at 36%.

Framatome ANP, in conjunction with German utilities and safety authorities, is developing the supercritical-water-cooled reactor (SWR), a 1000–1290-MWe BWR. This design was completed in 1999 and is ready for commercial deployment. Framatome ANP is seeking U.S. design certification for this system.

General Electric and Westinghouse are also developing designs for the European market. The General Electric system, known as the ESBWR, is 1390 MWe and is based on the ABWR. They are in the preapplication stage for a 1500-MWe version of this reactor for design certification by the U.S. NRC. Westinghouse is working with European and Scandinavian authorities on the 90+ PWR to be built in Sweden. These reactors all have passive safety systems.

6.2.3.1.5 Russia

Russia has also developed several advanced PWR designs with passive safety systems. The Gidropress 1000 MWe V-392 is being built in India with another planned for Novovoronezh. They are also building two VVER-91 reactors in China at Jiangsu Tainwan. The VVER-91 is designed with western control systems.

OKBM is developing the VVER-1500 for replacement of two plants each in Leningrad and Kursk. The design is planned to be complete in 2007 and the first units will be commissioned in 2012–2013.

Gidropress is developing a 640-MWe PWR with Siemans control systems which will be designated the VVER-640. OKBM is designing the VVER-600 with integral steam generators. Both of these designs will have enhanced safety systems.

6.2.3.2 Heavy-Water Reactors

Heavy-water reactors continue to be developed in Canada by AECL. They have two designs under development. The first, designated CANDU-9, is a 925–1300-MWe extension of the current CANDU-6. The CANDU-9 completed a two-year license review in 1997. The interesting design feature of this system is the flexible fuel requirements. Fuel materials include natural uranium, slightly enriched uranium, uranium recovered from the reprocessing of PWR fuel, mixed oxide (MOX) fuels, direct use of spent PWR fuel, and also thorium. The second design is the advanced CANDU Reactor (ACR). It uses pressurized light water as a coolant and maintains the heavy water in the calandria. The reactor is run at higher temperature and pressure, which gives it a higher thermal efficiency than earlier CANDU reactors.

The ACR-700 is smaller, simpler, cheaper, and more efficient than the CANDU-6. It is designed to be assembled from prefabricated modules that will cut the construction time to a projected 36 months. Heavy-water reactors have been plagued with a positive-void reactivity coefficient, which led some to question their safety. The ACR-700 will have a negative-void reactivity coefficient that enhances the safety of the system, as do the built-in passive safety features. AECL is seeking certification of this design in Canada, China, the U.S., and the U.K.

A follow-up to the ACR-700 is the ACR-1000, which will contain additional modules and operate in the range of 1100–1200 MWe. Each module of this design contains a single fuel channel and is expected to produce 2.5 MWe. The first of these systems is planned for operation in Ontario by 2014.

The long-range plan of AECL is to develop the CANDU-X, which will operate at a much higher temperature and pressure, yielding a projected thermal efficiency of 40%. The plan is to commercialize this plant after 2020 with a range of sizes from 350 to 1150 MWe.

India is also developing an advanced heavy-water reactor (AHWR). This reactor is part of the Indian program to utilize thorium as a fuel material. The AHWR is a 300-MWe heavy-water-moderated reactor. The fuel channels are arranged vertically in the calandria and are cooled by boiling light water. The fuel cycle will breed ^{233}U from ^{232}Th.

6.2.3.3 High-Temperature Gas-Cooled Reactors

The third generation of HTGRs is being designed to directly drive a gas turbine generating system using the circulating helium that cools the reactor core. The fuel material is a uranium oxycarbide in the form of small particles coated with multiple layers of carbon and silicon carbide. The coatings will contain the fission products and are stable up to 1600°C. The coated particles can be arranged in fixed graphite fuel elements or contained in "pebbles" for use in a pebble-bed-type reactor.

In South Africa, a consortium lead by the utility Eskom is developing the pebble-bed-modular reactor (PBMR). This reactor will have modules with power outputs of 165 MWe. It will utilize the direct gas turbine technology and is projected to have a thermal efficiency of 42%. The goal is to obtain a fuel burn-up of 90 GWd/t at the outset and eventually reach 200 GWd/t. The intent is to build a demonstration plant for operation in 2006 and obtain commercial operation in 2010.

In the U.S., a larger system is being design by General Atomics in conjunction with Minatom of Russia and Fuji of Japan. This reactor, designated the gas turbine-modular helium reactor (GT–MHR), utilizes hexagonal fuel elements of the kind that were used in the Fort St. Vrain reactor. The initial use of this reactor is expected to be to burn the weapons-grade plutonium at Tomsk in Russia.

6.2.3.4 Fast-Neutron Reactors

Several nations are working on developing improved fast-breeder reactors (FBRs). Fast-breeder reactors are fast-neutron reactors and about 20 of these reactors have operated since the 1950 s. They are able to use both ^{238}U and ^{235}U as reactor fuel, thus making use of all the uranium. These reactors use liquid metal as a coolant. In Europe, research work on the 1450-MWe European FBR has been halted.

In India, at the Indira Gandhi Centre for Atomic Research, a 40-MWt fast-breeder reactor has been operating since 1985. This reactor is used to research the use of thorium as reactor fuel by breeding ^{233}U. India has used this experience and began the construction of a 500-MWe prototype fast-breeder reactor in 2004. This unit at Kalpakkam is expected to be operating in 2010.

In Japan, the Joyo experimental reactor has been operating since 1977 and its power is now being raised to 140 MWt.

In Russia, the BN-600 FBR has been supplying electricity since 1981. It is considered to be the best operating reactor in Russia. The BN-350 FBR operated in Kazakhstan for 27 years and was used for water desalinization as well as electricity production. The BN-600 is being reconfigured to burn plutonium from the military stockpiles.

Russia has also begun construction of the BN-800 (880 MWe), which has enhanced safety features and improved fuel economy. This reactor will also be used to burn stockpiled plutonium. Russia has also experimented with lead-cooled reactor designs. A new Russian design is the BREST fast-neutron reactor. It will operate at 300 MWe or more and is an inherently safe reactor design. A pilot unit is being built at Beloyarsk. The reactor is fueled with plutonium nitride fuel and it has no blanket so no new plutonium is produced.

In the U.S., General Electric is involved in the design of a 150-MWe modular liquid-metal-cooled inherently safe reactor called PRISM. This design, along with a larger 1400-MWe design being developed jointly by GE and Argonne, has been withdrawn from NRC review.

6.2.3.5 Summary of Generation-III Reactors

As can be seen from the discussion above, there are many reactor systems of many types under development. The key feature of all of these reactors is the enhancement of safety systems. Some of these reactors have already been built and are in operation, whereas others are under construction. This activity indicates that there will be a growth of nuclear-reactor-generated electricity during the next 20 years. Table 6.7, taken from World Nuclear Association information on advanced nuclear power reactors, shows the advanced thermal reactors that are being marketed around the world.

6.2.4 Generation-IV Technologies

As discussed earlier, the development of nuclear power occurred in three general phases. The initial development of prototype reactor designs occurred in the 1950s and 1960s, development and deployment of large commercial plants occurred in the 1970s and 1980s, and development of advanced light-water reactors occurred in the 1990s.

Although the earlier generations of reactors have effectively demonstrated the viability of nuclear power, the nuclear industry still faces a number of challenges that need to be overcome for nuclear power to achieve its full potential. Among these challenges are (1) public concern about the safety of nuclear power in the wake of the Three Mile Island accident in 1979 and the Chernobyl accident in 1986, (2) high capital costs and licensing uncertainties associated with the construction of new nuclear power plants, (3) public concern over potential vulnerabilities of nuclear power plants to terrorist attacks, and (4) issues associated with the accumulation of nuclear waste and the potential for nuclear material proliferation in an environment of expanding nuclear power production.

To address these concerns and to fully realize the potential contributions of nuclear power to future energy needs in the United States and worldwide, the development of a new generation of reactors, termed generation IV, was initiated in 2001. The intent or objective of this effort is to develop multiple generation-IV nuclear power systems that would be available for international deployment before the year 2030. The development of the generation-IV reactor systems is an international effort, initiated by the U.S. DOE with participation from 10 countries. These countries established a formal organization referred to as the Generation IV International Forum (GIF). The GIF countries included Argentina, Brazil, Canada, France,

TABLE 6.7 Advanced Thermal Reactors Being Marketed

Country and Developer	Reactor	Size (MWe)	Design Progress	Main Features
U.S.-Japan (GE-Hitachi-Toshiba)	ABWR	1300	Commercial operation in Japan since 1996–1997, in U.S.: NRC certified 1997, first-of-a-kind engineering	Evolutionary design More efficient, less waste Simplified construction (48 months) and operation
South Korea (derived from Westinghouse)	APR-1400 (PWR)	1400	NRC certified 1997, Further developed for new S. Korean Shin Kori 3 and 4, expected to be operating in 2010	Evolutionary design Increased reliability Simplified construction and operation
U.S.A (Westinghouse)	AP-600	600	AP-600: NRC certified 1999, FOAKE	Passive safety features Simplified construction and operation 3 years to build 60-year plant life
	AP-1000 (PWR)	1100	AP-1000 NRC design approval 2004	
Japan (Utilities, Westinghouse, Mitsubishi)	APWR	1500	Basic design in progress, planned at Tsuruga	Hybrid safety features Simpliified construction and operation
France–Germany (Framatome ANP)	EPR (PWR)	1600	Confirmed as future French standard, French design approval, to be built in Finland	Evolutionary design Improved safety features High fuel efficiency Low-cost electricity
U.S.A (GE)	ESBWR	1390	Developed from the ABWR, precertification in U.S.A	Evolutionary design Short construction time Enhanced safety features
Germany (Framatome ANP)	SWR-1000 (BWR)	1200	Under development, precertification in U.S.A	Innovative design High-fuel efficiency Passive safety features
Russia (OKBM)	V-448 (PWR)	1500	Replacement for Leningrad and Kursk plants	High-fuel efficiency Enhanced safety

Russia (Gidropress)	V-392 (PWR)	950	Two being build in India, likely bid for China	Evolutionary design 60-year plant life Enhanced safety features
Canada (AECL)	CANDU-9	925-1300	Licensing approval 1997	Evolutionary design Single stand-alone unit Flexible fuel requirements Passive safety features
Canada (AECL)	ACR	700	ACR-700: precertification in U.S.A	Evolutionary design Light-water cooling Low-enriched fuel Passive safe features
		1000	ACR-1000 proposed for U.K.	
South Africa (Eskom, BNFL)	PBMR	165 (module)	Prototype due to start building, precertification in U.S.A	Modular plant, low cost Direct cycle gas turbine High-fuel efficiency Passive safety features
U.S.A-Russia et al. (General Atomics, Minatom)	GT-MHR	285 (module)	Under development in the U.S.A. and Russia by multinational joint venture	Modular plant, low cost Direct-cycle gas turbine High-fuel efficiency Passive safety features

Source: From World Nuclear Association, Plans for New Reactors Worldwide, http://www.world-nuclear.org, 2005.

Japan, the Republic of Korea, the Republic of South Africa, Switzerland, the United Kingdom, and the United States. The intent of the GIF is "…to develop future-generation nuclear energy systems that can be licensed, constructed, and operated in a manner that will provide competitively priced and reliable energy products while satisfactorily addressing nuclear safety, waste, proliferation, and public perception concerns."

The process used by the GIF to identify the most promising reactor concepts for development (referred to as the Generation IV Technology Roadmap) consisted of three steps. These steps were (1) to develop a set of goals for new reactor systems, (2) solicit proposals from the worldwide nuclear community for new reactor systems to meet these goals, and (3) using experts from around the world, evaluate the different concepts to select the most promising candidates for further development.

The eight goals developed by the GIF for generation-IV nuclear systems were:

- Sustainability 1: Generation-IV nuclear energy systems will provide sustainable energy generation that meets clean air objective and promotes long-term availability of systems and effective fuel utilization for worldwide energy production.
- Sustainability 2: Generation-IV nuclear energy systems will minimize and manage their nuclear waste and notably reduce the long-term stewardship burden in the future, thereby improving protection for the public health and the environment.
- Economics 1: Generation-IV nuclear energy systems will have a clear life-cycle cost advantage over other energy sources.
- Economics 2: Generation-IV nuclear energy systems will have a level of financial risk comparable to other energy projects.
- Safety and reliability 1: Generation-IV nuclear energy systems operations will excel in safety and reliability.
- Safety and reliability 2: Generation IV nuclear energy systems will have a very low likelihood and degree of reactor core damage.
- Safety and reliability 3: Generation-IV nuclear energy systems will eliminate the need for offsite emergency response.
- Proliferation resistance and physical protection: Generation-IV nuclear energy systems will increase the assurance that they are a very unattractive and the least desirable route for diversion or theft of weapons-usable materials, and provide increased physical protection against acts of terrorism.

Over 100 generation-IV candidates were evaluated by experts from the GIF countries and six reactor systems were selected for further evaluation and potential development. The six reactor systems selected were:

6.2.4.1 Gas-Cooled Fast-Reactor System

The gas-cooled fast-reactor system (GFR) is a fast-neutron spectrum reactor that uses helium as the primary coolant. It is designed to operate at relatively high helium outlet temperatures, making it a good candidate for the high-efficiency production of electricity or hydrogen. As shown in Figure 6.21 below, a direct Brayton cycle is used for the production of electricity with the helium gas delivered from the reactor outlet to a high-temperature gas turbine connected to a generator that produces electricity. In alternative designs, the high-temperature helium can also be used to produce hydrogen using either a thermochemical process or high-temperature electrolysis, or for other high-temperature process heat applications.

The reference plant is designed to produce 288 MWe using the direct Brayton cycle with a reactor outlet temperature of 850°C. The fuel forms being considered for high-temperature operation include composite ceramic fuel, advanced fuel particles, or ceramic clad elements of actinide compounds. Alternative core configurations include prismatic blocks and pin- or plate-based assemblies. The GFR's fast-neutron spectrum also makes it possible to efficiently use available fissile and fertile materials in a once-through fuel cycle.

6.2.4.2 Very-High-Temperature Reactor

The very-high-temperature reactor (VHTR) is a helium-cooled reactor designed to provide heat at very high temperatures, in the range of 1000°C for high-temperature process heat applications. In particular, the

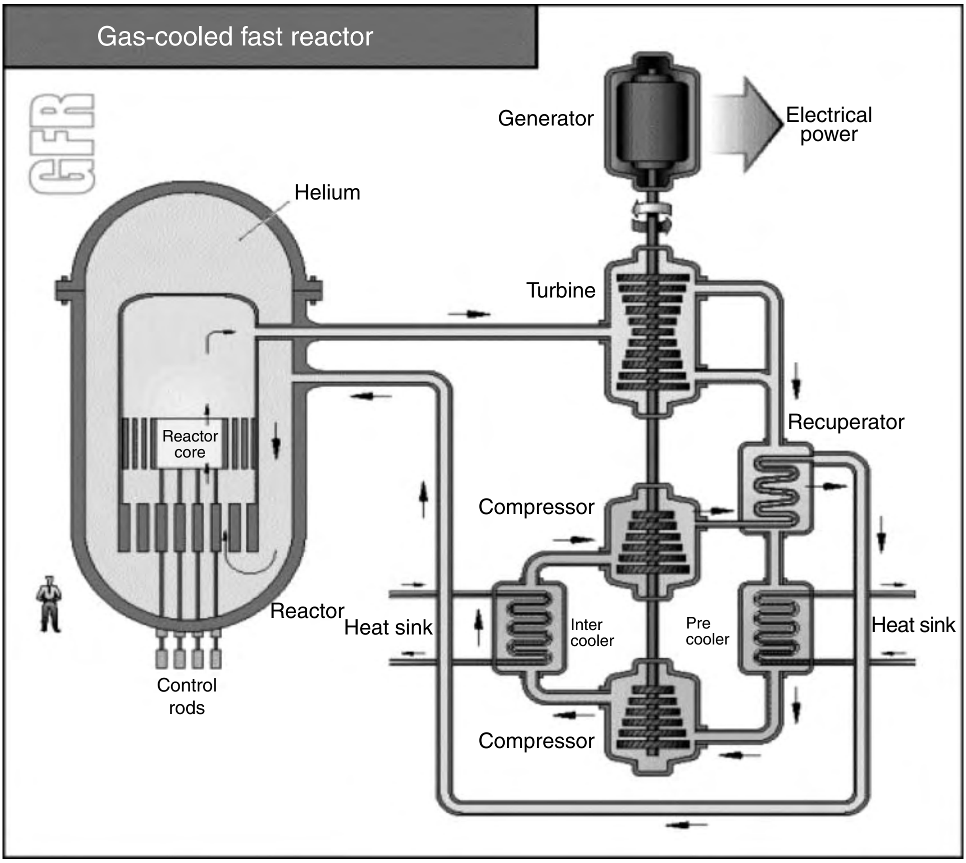


FIGURE 6.21 Gas-cooled fast reactor. (From US DOE Nuclear Energy Research Advisory Committee and the Generation IV International Forum. 2002. *A Technology Roadmap for the Generation IV Nuclear Energy Systems*.)

1000°C reactor outlet temperature makes it a good candidate for the production of hydrogen using either thermochemical or high-temperature electrolysis processes. As shown in Figure 6.22 below, heat for the production of hydrogen is delivered through an intermediate heat exchanger that serves to isolate the reactor system from the hydrogen production process.

The reference design for the VHTR is a 600-MWt reactor with an outlet temperature of 1000°C. The reactor core uses graphite as a moderator to produce the thermal neutrons for the fission process. The core configuration can be either graphite blocks or pebbles about the size of billiard balls in which fuel particles are dispersed. For electricity production, either a direct Brayton cycle gas turbine using the primary helium coolant as the working fluid, or an indirect Rankine cycle using a secondary working fluid can be used. The high-temperature characteristics of this reactor concept also make it an ideal candidate for cogeneration applications to meet both electricity and hydrogen production or other high-temperature process heat needs.

6.2.4.3 Supercritical-Water-Cooled Reactor

The supercritical-water-cooled reactor (SWR) is a relatively high-temperature, high-pressure reactor designed to operate above the thermodynamic critical point of water, which is 374°C and 22.1 MPa. Because there is no phase change in the supercritical coolant water, the balance of plant design, shown in Figure 6.23, utilizes a relatively simple direct-cycle power-conversion system. The reference design for this concept is a 1700-MWe reactor operating at a pressure of 25 MPa with a reactor outlet temperature ranging

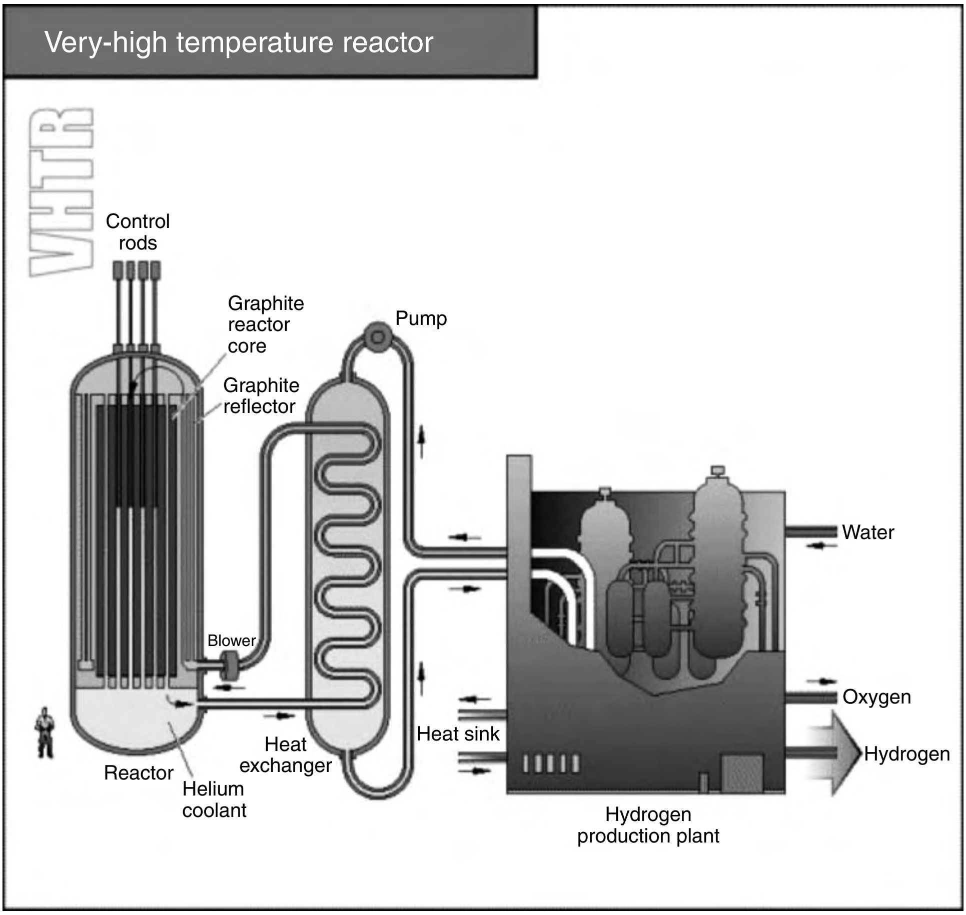


FIGURE 6.22 Very-high temperature reactor. (From US DOE Nuclear Energy Research Advisory committee and the Generation IV International Forum, *A Technology Roadmap for the Generation IV Nuclear Energy Systems*.)

between 510 and 550°C. This reactor can be designed as either a fast-neutron-spectrum or thermal-neutron-spectrum reactor. The relatively simple design also allows for the incorporation of passive safety features similar to those of the simplified boiling-water reactor discussed earlier. However, unlike the previously discussed concepts, the lower reactor outlet temperature is not well suited for the efficient production of hydrogen, which requires minimum temperatures in the range of 850°C–900°C. Therefore, this reactor concept is primarily intended for the efficient, low-cost production of electricity.

6.2.4.4 Sodium-Cooled Fast Reactor

The sodium-cooled fast reactor (SFR), shown in Figure 6.24, is a sodium-cooled fast-neutron-spectrum reactor designed primarily for the efficient management of actinides and conversion of fertile uranium in a closed fuel cycle. Two reference designs to support different fuel reprocessing options have been defined for this concept. The first is a medium-sized sodium-cooled reactor with a power output between 150 and 500 MWe that utilized uranium-plutonium-minor-actinide-zirconium metal alloy fuel. This reactor concept is supported by a fuel cycle based on pyrometallurgical processing in which the processing facilities are an integral part of the reactor plant design.

The second reactor reference design is a large sodium-cooled reactor with a power output capability between 500 and 1500 MWe that utilizes uranium-plutonium oxide fuel. This reactor design is supported

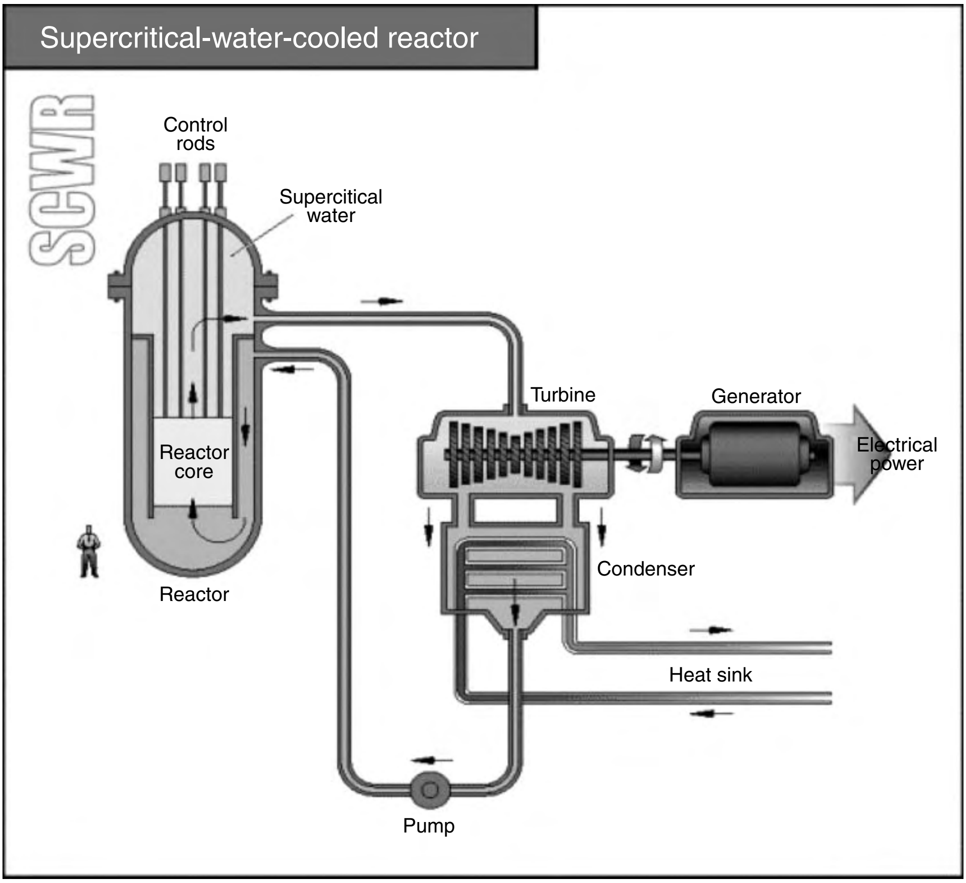


FIGURE 6.23 Supercritical-water-cooled reactor. (From US DOE Nuclear Energy Research Advisory Committee and the Generation IV International Forum, 2002. *A Technology Roadmap for the Generation IV Nuclear Energy Systems.*)

by a fuel cycle based on an advanced aqueous process that would include a centrally located processing facility supporting a number of reactors.

Both versions of this reactor concept would operate at coolant outlet temperatures in the range of 550°C, and are intended primarily for the management of high-level waste and the production of electricity. In addition to design innovations to reduce capital costs, these reactors incorporate a number of enhanced safety features that include:

- Long thermal response time
- Large margin to coolant boiling
- Primary system that operates near atmospheric pressure
- Intermediate sodium system between the radioactive sodium in the primary system and the water and steam in the power plant.

6.2.4.5 Lead-Cooled Fast Reactor

The lead-cooled fast reactor (LFR) is a fast-neutron-spectrum reactor cooled by either molten lead or a lead-bismuth eutectic liquid metal. It is designed for the efficient conversion of fertile uranium and the management of actinides in a closed fuel cycle. The reactor core for this design, shown in Figure 6.25,

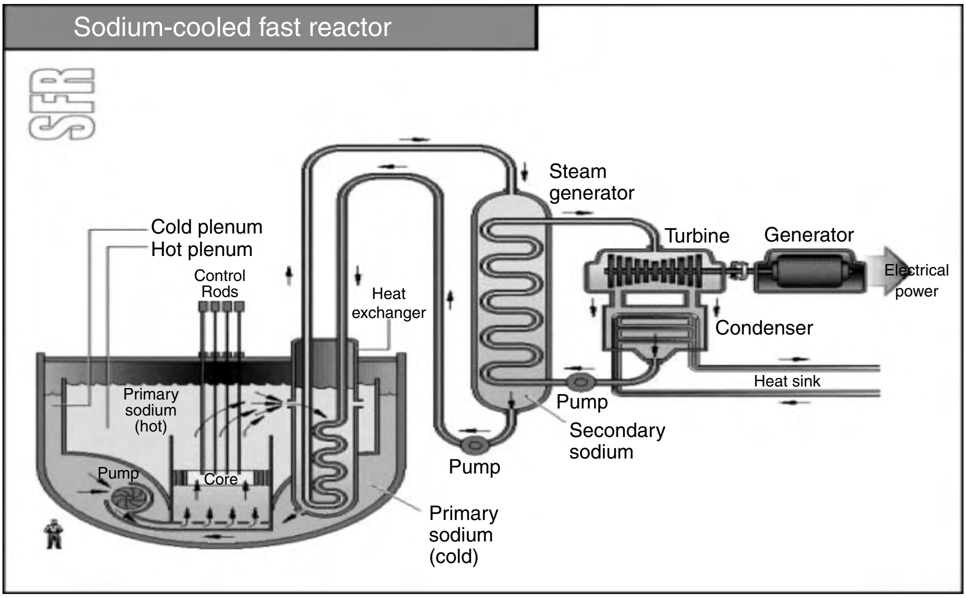


FIGURE 6.24 Sodium-cooled fast reactor. (From US DOE Nuclear Energy Research Advisory Committee and the Generation IV International Forum, 2002. *A Technology Roadmap for the Generation IV Nuclear Energy Systems.*)

utilizes a metal or nitride-based fuel containing fertile uranium and transuranics. As shown in Figure 6.25, the LFR relies on natural convection to cool the reactor core. The outlet temperature for the current reactor concept is about 550°C, but with advanced materials, reactor outlet temperatures of 800°C may be possible. An indirect-gas Brayton cycle is used to produce electrical power.

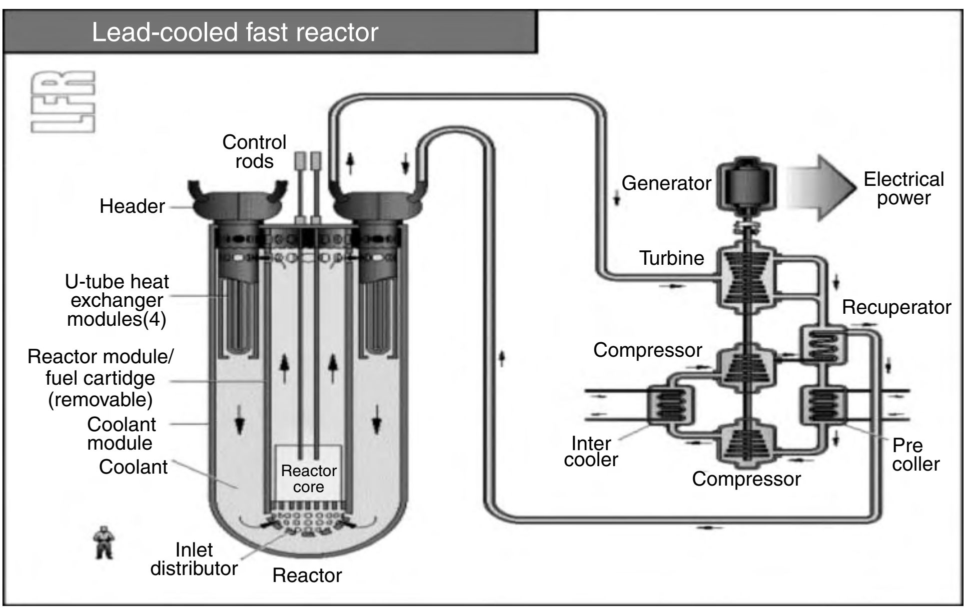


FIGURE 6.25 Lead-cooled fast reactor. (From US DOE Nuclear Energy Research Advisory Committee and the Generation IV International Forum, 2002. *A Technology Roadmap for the Generation IV Nuclear Energy Systems.*)

There are currently three versions of the reference design for this concept. The smallest design, rated at 50–150 MWe is intended for distributed power applications or electricity production on small grids. This reactor design, referred to as a battery, features modular design with a factory fabrication "cassette" core. The reactor is designed for very long refueling intervals (15–20 years), with refueling accomplished by replacement of the cassette core or reactor module.

The other two versions of this design are a modular system rated at 300–400 MWe, and a large plant rated at 1200 MWe. The different power options for this design are intended to fill different needs or opportunities in the power market, and be economically competitive with comparable alternative power sources.

6.2.4.6 Molten-Salt Reactor

The molten-salt reactor (MSR), shown in Figure 6.26, produces power by circulating a molten salt and fuel mixture through graphite-core flow channels. The slowing down of neutrons by the graphite moderator in the core region provides the epithermal neutrons necessary to produce the fission power for sustained operation of the reactor. The heat from the reactor core is then transferred to a secondary system through an intermediate heat exchanger and then through a tertiary heat exchanger to the power conversion system that produces the electric power. The circulating coolant flow for this design is a mixture of sodium, uranium, and zirconium fluorides. In a closed fuel cycle, actinides such as plutonium can be efficiently burned by adding these constituents to the liquid fuel without the need for special fuel fabrication.

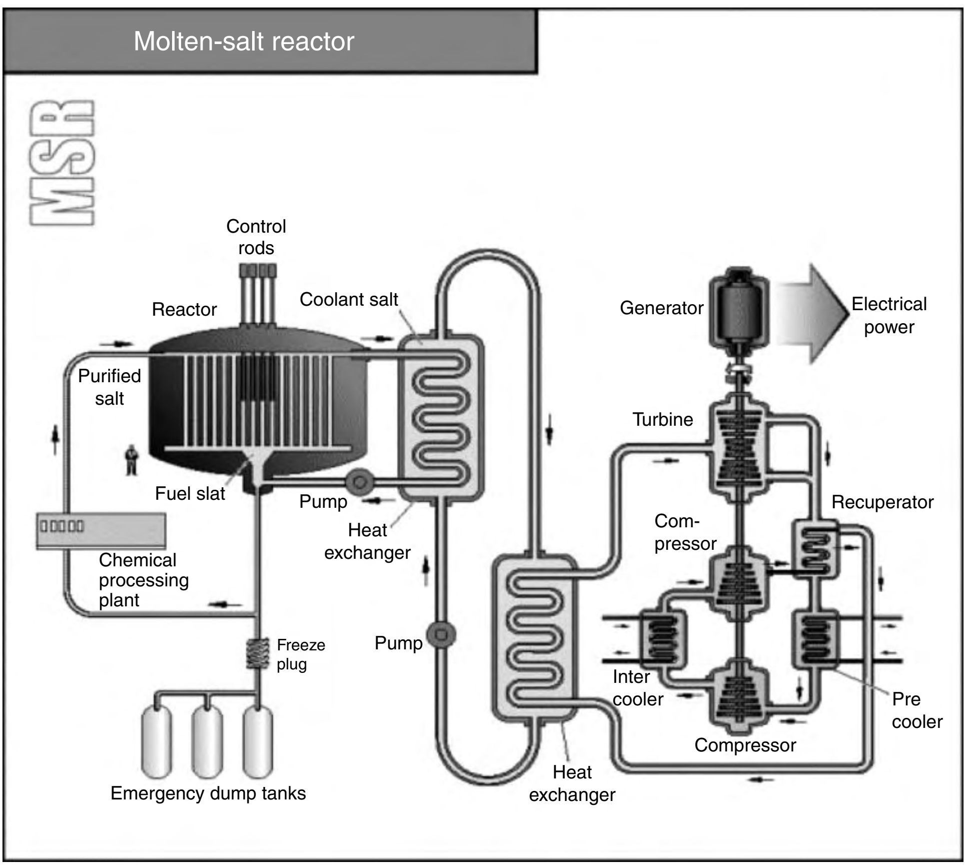


FIGURE 6.26 Molten-salt reactor. (From US DOE Nuclear Energy Research Advisory Committee and the Generation IV International Forum, 2002. *A Technology Roadmap for the Generation IV Nuclear Energy Systems.*)

The reference design for this concept is a 1000 MWe power plant with a coolant outlet temperature of 700°C. To achieve higher thermal efficiencies for this concept, coolant outlet temperatures as high as 800°C may also be possible.

6.2.5 Fuel Cycle

The process of following the fuel material from the uranium or thorium mine through processing and reactor operation until it becomes waste is called the fuel cycle for nuclear systems. After a discussion of the fuel cycle in general, the fuel cycle will be examined by looking at uranium and thorium resources, mining and milling, enrichment, reactor fuel use, spent fuel storage, nuclear materials transportation, and reprocessing. Nuclear waste will be addressed in a separate section.

General discussion of the fuel cycle will often include the terms “open” or “closed.” The open fuel cycle is also called the *once-through* cycle. In the once-through fuel cycle, the uranium fuel is fabricated and run through the reactor once and then disposed of as waste. There is no reprocessing of the fuel. In the closed cycle, the fuel is reprocessed after leaving the reactor so that it can be reused to improve overall fuel utilization.

In the open cycle, the fuel is introduced into the reactor for one to two years. It is then removed and placed into long-term storage for eventual disposal. The impact of this cycle is the waste of about 95% of the energy contained in the fuel. The U.S. adopted the open cycle in 1977 when President Carter issued an executive order to stop reprocessing as a part of the fuel cycle. Canada has also adopted the open cycle.

The closed cycle was envisioned when the development of nuclear power began. The uranium and plutonium removed from reactors would be reprocessed and returned to reactors as fuel. Breeder reactors would be used to breed additional plutonium for use in thermal reactors. Thorium could also be used as a breeding material to generate ^{233}U as a reactor fuel. The intent of the closed fuel cycle was to maximize the use of available reactor fuel resources while minimizing waste generated by operating reactors.

Currently, reprocessing is used in Europe and Japan, but the benefits of the closed cycle have not been fully realized because there has only been limited use of the separated plutonium. As discussed above, the U.S. and Canada, for reasons described later, have not pursued closed cycle reprocessing of spent fuel. As a result, only a small fraction of the available fuel resources are utilized, and disposal of large quantities of potentially usable spent fuels has become a major issue for the U.S. nuclear industry.

6.2.5.1 Uranium and Thorium Resources

Uranium is a common material in the earth’s crust. It is also present in sea water. Thorium is about three times more plentiful then uranium. Typical concentrations of uranium measured in parts per million (ppm) are shown in Table 6.8.

The amount of recoverable uranium is dependent upon the price. As the price increases, more material is economically recoverable. Also, more exploration will occur and it is likely that additional orebodies will be discovered. An orebody is defined as an occurrence of mineralization from which the metal, in this case

TABLE 6.8 Typical Concentrations of Uranium

Source	Uranium Concentration (ppm)
High-grade ore: 2% U	20,000
Low-grade ore: 0.1% U	1000
Granite	4
Sedimentary rock	2
Earth’s continental crust (avg)	2.8
Seawater	0.003

Source: From World Nuclear Association, Supply of Uranium, http://www.world-nuclear.org, 2004.

TABLE 6.9 Known Recoverable Resources of Uranium

Country	Tons of Uranium	Percentage of Total
Australia	989,000	28
Kazakhstan	622,000	18
Canada	439,000	12
South Africa	298,000	8
Namibia	213,000	6
Brazil	143,000	4
Russian Federation	158,000	4
U.S.A	102,000	3
Uzbekistan	93,000	3
World total	3,537,000	—

Source: From World Nuclear Association, Supply of Uranium, http://www.world-nuclear.org, 2004.

uranium, can be recovered economically. Because of the uncertainties of price and its impact on exploration, any statement of recoverable amounts of uranium is simply a picture at an instant in time and is likely to change many times in the future. There is also a store of highly enriched uranium that is being recovered as nuclear weapons are dismantled. In addition, there are millions of tons of ^{238}U that are the results of previous enrichment activities around the world. The ^{238}U can be blended with highly enriched uranium or plutonium to make fuel for nuclear power plants. The ^{238}U can also be used to breed plutonium in FBR fuel cycles.

Table 6.9 presents a list of recoverable resources of uranium. The table is taken from information gathered by the World Nuclear Association from other sources and was generated in 2004.

The 3.5 Mt is enough to fuel the world's current reactors for 50 years assuming the same fuel cycles currently in use. IAEA estimates the world supply at over 14 Mt, which provides a supply exceeding 200 years at the current rate of use. This estimate does not include the uranium in phosphate deposits estimated at 22 Mt or the uranium available in seawater estimated at 1400 Mt. In addition, the ability of nuclear reactors to achieve higher burn-ups (utilize more of the uranium in the fuel) has also increased. This increases the efficiency of uranium use. Because thorium is not included in these fuel supply numbers, and as noted above is about three times as plentiful as uranium, there does not appear to be a fuel supply limitation for nuclear power in the foreseeable future.

6.2.5.2 Mining and Milling

Uranium is being mined using traditional underground and open-pit excavation technologies, and also using in situ leaching or solution-mining techniques.

Underground mining is used when the orebody is deep underground, usually greater than 120 m deep. In underground mines, only the orebody material is extracted. Underground mining is hazardous and made more so by high concentrations of radon from the radioactive decay of the uranium. Once mined, the extracted ore is sent to a mill where the uranium in the ore is concentrated.

Open-pit technology is used when the orebody is near the surface. This leads to the excavation of large amounts of material that does not contain the ore itself. The ore that is recovered is also sent to a mill for further processing.

Solution mining involves the introduction of an aqueous solution into the orebody. The solution, oxygenated ground water, is pumped into the porous orebody and the uranium is dissolved. The uranium-rich solution is then extracted and sent to the mill for further processing.

The milling process for the solid ore material involves crushing the ore and then subjecting it to a highly acidic or alkaline solution to dissolve the uranium. Mills are normally located close to the mining activity and a single mill will often support several mines. The solution containing the uranium goes through a precipitation process that yields a material called *yellow cake*. The yellow cake contains about

80% uranium oxide. The yellow cake is packaged and sent to a conversion and enrichment facility for further processing.

6.2.5.3 Conversion and Enrichment

Prior to entering the enrichment process, the impure U_3O_8 is converted through a series of chemical processing steps to UF_6. During these processes, the uranium is purified. Conversion facilities are operating commercially in the U.S., Canada, France, the U.K., and Russia. UF_6 is a solid at room temperature but converts to its gaseous form at moderate temperature levels, making the compound suitable for use in the enrichment process. UF_6 is very corrosive and reacts readily with water. It is transported in large cylinders in the solid state.

Conversion of the U_3O_8 to UO_2 is also done at conversion facilities. The natural UO_2 is used in reactors such as the CANDU that do not require enriched uranium as fuel.

The first enrichment facilities were operated during the 1940 s. The electromagnetic isotope-separation process was used to separate the ^{235}U used in the first atomic bomb. The process used a magnetic field to separate the ^{235}U from the ^{238}U. As the ions were accelerated and turned, they moved differently because of the difference in their masses. Multiple stages were required and the process was very difficult to run efficiently; it was therefore soon abandoned.

Today, only two processes—gaseous diffusion and gas centrifugation—are used commercially. The capacity of enrichment plants is measured in separative work units (SWU). The SWU is a complex term that is dependent on the amount of uranium that is processed, and the concentration of ^{235}U in the product and in the tails. It is a measure of the amount of energy used in the process.

The first commercial enrichment was carried out in large gaseous diffusion plants in the U.S. It has also been used in Russia, the U.K., France, China, and Argentina. Today, operating plants remain in the U.S., France, and China, with a total nominal capacity of 30 million SWU.

In the gaseous diffusion process, UF_6 is pumped through a series of porous walls or membranes that allow more of the light ^{235}U to pass through. Because the lighter ^{235}U particles travel faster then the heavier ^{238}U particles, more of them penetrate the membrane. This process continues through a series of membranes with the concentration of ^{235}U increasing each time. For commercial reactor fuel, the process continues until the ^{235}U concentration is 3%–5%. The slower ^{238}U particles are left behind and collect as a product referred to as *tails*. The tails have a reduced concentration of ^{235}U and are commonly referred to as *depleted uranium*. This process uses a very large amount of energy and thus is very expensive to operate.

In the centrifuge enrichment process, the gaseous UF_6 is placed in a high-speed centrifuge. The spinning action forces the heavier ^{238}U particles to the outside while the lighter ^{235}U particles remain closer to the center. To obtain the enrichment required for power reactor fuel, many stages of separation are required. The arrangement is know as a *cascade*. Again, the process is continued until the ^{235}U concentration is 3%–5%. The centrifuge process uses only about 2% of the energy required by gaseous diffusion.

Table 6.10 shows the location and size of enrichment facilities around the world.

6.2.5.4 Fuel Fabrication and Use

Following enrichment, the UF_6 is shipped to a fuel fabrication facility. Here, the UF_6 is converted to UO_2 and pressed into cylindrical ceramic pellets. The pellets are sintered, heated to high temperature, and inserted in the fuel cladding tubes. The tubular material is zircaloy, an alloy of zirconium. The tubes are sealed forming fuel rods that are assembled into fuel assemblies and shipped to a reactor for use. All of the dimensions of the pellets and fuel rods are very carefully controlled to assure uniformity throughout the fuel assemblies.

The primary hazard in the fabrication facility is the potential for an accidental criticality because they are working with enriched uranium. Therefore, all of the processing quantities and the dimensions of the processing vessels must be controlled. This must be done even with low-enriched uranium.

TABLE 6.10 Location, Size, and Type of Enrichment Facilities Around the World

Country	Owner/Controller	Plant Name/Location	Capacity (1000 SWU)
		Gaseous Diffusion Plants	
China	CNNC	Lanzhou	900
France	EURODIF	Tricastin	10,800
		Paducah, KY	11,300
United States	U.S. Enrichment Corporation	Portsmouth, OH (Closed since May, 2001)	7400
Subtotal			30,400
		Centrifuge Plants	
		Hanzhong	500
China	CNNC	Lanzhou	500
Germany	Urenco	Gronau	1462.5
	JNC	Ningyo Toge	200
Japan	Japan Nuclear Fuel Limited (JNFL)	Rokkasho-mura	1050
The Netherlands	Urenco	Almelo	1950
Pakistan	Pakistan Atomic Energy Commission (PAEC)	Kahuta	5
		Ural Electrochemical Integrated Enterprise (UEIE), Novouralsk	7000
		Siberian Chemical Combine (SKhK), Seversk	4000
Russia	Minatom	Electrochemical Plant (ECP), Zelenogorsk	3000
		Angarsk Electrolytic Chemical Combine (AEKhK), Angarsk	1000
United Kingdom	Urenco	Capenhurst	2437.5
Subtotal			23,105
Total			53,505

Source: From WISE Uranium Project, World Nuclear Fuel Facilities, http://www.wise-uranium.org, 2005.

A typical 1000-MWe reactor will use about 27 tons of UO_2 each year. Typical burn-up in current reactors is 33 GWd/t of uranium fed to the reactor. The energy available from the fission of uranium is 1 MW/g of uranium or 1000 GW/t. Using these numbers, the actual amount of uranium burned is only 3%–5%. This means that the unused energy available from the spent fuel, if it could be completely burned, is over 95%. During the operation of the reactor, some of the ^{238}U is converted to plutonium, which also contributes to the thermal energy of the reactor.

Advanced fuel use in reactors is estimated to be up to 200 GWd/t. In this case, about 80% of the energy available from the uranium remains in the spent fuel. These facts are the driving force behind the questions regarding reprocessing. In the once-through fuel cycle, the spent fuel will be disposed of as waste. In the closed cycle, the spent fuel is reprocessed and the remaining uranium and also the plutonium are recovered.

6.2.5.5 Reprocessing

In the 1940 s, reactors were operated solely for the production of plutonium for use in weapons. The fuels from the production reactors were reprocessed to recover the plutonium. The chemical processes were developed to separate the fission products and the uranium from the plutonium. The most common process was the PUREX process. This is the process that is used today by countries that reprocess power reactor fuels.

The purpose of reprocessing is to recover the uranium and plutonium in the spent fuel. As discussed above, these materials contain a large amount of potential energy if they are reused as reactor fuel. Plutonium separated in the PUREX process can be mixed with uranium to form a MOX fuel. Plutonium from the dismantlement of weapons can be used in the same way.

The potential availability of separated plutonium is seen by some as a potential mechanism for the proliferation of nuclear weapons. This was the basis of the U.S. decision to halt reprocessing. In the 1970 s, research began into methods for modifying the chemical process so that the plutonium and uranium would remain together at the end of the process. In this method, called coprocessing, the short-lived fission products would be separated and the remaining uranium, plutonium, and other actinide elements would remain together. This remaining mixture would be highly radioactive, but could be remotely processed into new reactor fuel. A blend of fast neutron and thermal reactors could be used to maximize the use of this material.

The current worldwide reprocessing capability is shown in Table 6.11. These facilities all use the PUREX technology. More then 80,000 tons of commercial fuel have been reprocessed in these facilities.

Three processes are considered to be mature options for reprocessing fuel: PUREX, UREX+, and pyroprocessing. Each of these processes has certain advantages and disadvantages.

TABLE 6.11 World Commercial Reprocessing Capacity

Type of Fuel	Location	Tons/year
LWR fuel	France, La Hague	1700
	U.K., Sellafield (THORP)	900
	Russia, Ozersk (Mayak)	400
	Japan	14
	Subtotal	3000
Other nuclear fuels	U.K., Sellafield	1500
	India	275
	Subtotal	1750
Civilian capacity	Total	4750

Source: From Uranium Information Centre, Nuclear Issues Briefing Paper 72, Processing of Used Nuclear Fuel, http://www.uic.com.au, 2005.

6.2.5.5.1 PUREX

The PUREX process is the oldest and most common reprocessing option. It uses liquid–liquid extraction to process light-water reactor spent fuel. The spent fuel is dissolved in nitric acid, and then the acid solution is mixed with an organic solvent consisting of tributyl phosphate in kerosene. The uranium and plutonium are extracted in the organic phase and the fission products remain in the aqueous phase. Further processing allows the separation of the uranium and plutonium. The advantage of this process is the long-term experience with the process. The disadvantage is that it cannot separate fission products such as technetium, cesium, and strontium, nor can it separate actinides such as neptunium, americium, and curium.

6.2.5.5.2 UREX+

The UREX+ process is a liquid–liquid extraction process like PUREX. It can be used for light-water reactor fuels and it includes additional extraction steps that allow separation of neptunium/plutonium, technetium, uranium, cesium/strontium, americium, and curium. The advantage of this process is that it meets the requirements for continuous recycle in light-water reactors and it builds on current technology. The disadvantage is that it cannot be used to process short-cooled fuels and it cannot be used for some specialty fuels being developed for advanced reactors.

6.2.5.5.3 Pyroprocessing

This process was developed and tested at Experimental Breeder Reactor-2 (EBR-2) by Argonne National Laboratory in the U.S. It is an electrochemical process rather then a liquid–liquid extraction process. Oxide fuels are first converted to metals to be processed. The metallic fuel is then treated to separate uranium and the transuranic elements from the fission products. The advantage of this process is the ability to process short-cooled and specialty fuels designed for advanced reactors. The disadvantage is that it does not meet the requirements for continuous recycle from thermal reactors; however, it is ideal for fuel from fast-neutron reactors.

6.2.5.6 Spent-Fuel Storage

Spent fuel is routinely discharged from operating reactors. As it is discharged, it is moved to the spent-fuel storage pool that is an integral part of the reactor facility. Reactors are built with storage pools that will hold fuel from many years of operation. The pools are actively cooled by circulating cooling water. The fuel stored at many of the older reactors is reaching the capacity of the on-site storage pools. At this point, the fuel is being transferred to dry storage. Dry storage takes place in large metal or concrete storage facilities. These dry facilities are passively cooled by the air circulating around them.

6.2.5.7 Spent-Fuel Transportation

Spent fuel is transported in large engineered containers designated as type-B containers (casks). The casks provide shielding for the highly radioactive fuel so that they can be safely handled. They are constructed of cast iron or steel. Many of them use lead as the shielding material. They are also designed to protect the environment by maintaining their integrity in the case of an accident. They are designed to withstand severe accidents, including fires, impacts, immersion, pressure, heat and cold, and are tested as part of the design certification process.

Casks have been used to transport radioactive materials for over 50 years. The IAEA has published advisory regulations for safe transportation of radioactive materials since 1961. Casks are built to standards designed to meet the IAEA advisory regulations specified by licensing authorities such as the NRC in the U.S.

Spent fuel is shipped from reactor sites by road, rail, or water. The large casks can weigh up to 110 tons and hold about 6 tons of spent fuel. Since 1971, about 7000 shipments of spent fuel (over 35,000 tons) have been transported over 30 million km with no property damage or personal injury, no breach of containment, and a very low dose rate to the personnel involved.

6.2.6 Nuclear Waste

Radioactive wastes are produced throughout the reactor fuel cycle. The costs of managing these wastes are included in the costs of the nuclear fuel cycle and thus are part of the electricity cost. Because these materials are radioactive, they decay with time. Each radioactive isotope has a half life, which is the time it takes for half of the material to decay away. Eventually, these materials decay to a stable nonradioactive form.

The process of managing radioactive waste involves the protection of people from the effects of radiation. The longer lived materials tend to emit alpha and beta particles. It is relatively easy to shield people from this radiation but if these materials are ingested the alpha and beta radiation can be harmful. The shorter lived materials usually emit gamma rays. These materials require greater amounts of shielding.

6.2.6.1 Types of Radioactive Wastes

The strict definitions of types of radioactive waste may vary from country to country. In the following discussion, the more generally accepted terminology will be used.

6.2.6.1.1 Mine Tailings

Mining and milling of uranium produces a sandy type of waste that contains the naturally occurring radioactive elements that are present in uranium ore. The decay of these materials produces radon gas that must be contained. This is often accomplished by covering the tailings piles with clay to contain the radon gas. Technically, tailings are not classified as radioactive waste.

6.2.6.1.2 Low-Level Wastes

Low-level wastes (LLW) is generated from medical and industrial uses of radioactive materials as well as from the nuclear fuel cycle. In general, these wastes include materials such as paper, clothing, rags, tools, filters, soils, etc., that contain small amounts of radioactivity. The radioactivity tends to be short-lived. These materials generally do not have to be shielded during transport and they are suitable for shallow land burial. The volume of these materials may be reduced by compacting or incinerating prior to disposal. They make up about 90% of the volume of radioactive waste but contain only about 1% of the radioactivity of all the radioactive waste.

6.2.6.1.3 Intermediate-Level Wastes

Intermediate-level wastes (ILW) are generated during the operation of nuclear reactors, in the reprocessing of spent fuel, and from the decommissioning of nuclear facilities. These materials contain higher amounts of radioactivity and generally require some shielding during storage and transportation. Intermediate-level wastes is generally made up of resins, chemical sludges, fuel cladding, and contaminated materials from decommissioned nuclear facilities. Some of these materials are processed before disposal by solidifying them in concrete or bitumen. They make up about 7% of the volume and have about 4% of the radioactivity of all the radioactive waste.

6.2.6.1.4 High-Level Wastes

High-level wastes (HLW) is generated in the operation of a nuclear reactor. This waste consists of fission products and transuranic elements generated during the fission process. This material is highly radioactive and it is also thermally hot so that it must be both shielded and cooled. It accounts for 95% of the radioactivity produced by nuclear power reactors.

6.2.6.1.5 Managing HLW from Spent Fuel

The form of HLW from spent fuel is either the spent fuel itself or the waste products from reprocessing. The level of radioactivity from spent fuel falls to about one thousandth of the level it was when removed from the reactor in 40–50 years. This means the heat generated is also greatly reduced.

Currently, 270,000 tons of spent fuel are in storage at reactor sites around the world. An additional 12,000 tons are generated each year and about 3,000 tons of this are sent for reprocessing.

When spent fuel reprocessing is used, the uranium and plutonium are first removed during reprocessing, and then the much smaller volume of remaining HLW is solidified using a vitrification process. In this process, the fission products are mixed in a glass material, vitrified in stainless steel canisters and stored in shielded facilities for later disposal.

High-level waste will eventually be disposed of in deep geologic facilities. Several countries have selected sites for these facilities and they are expected to be commissioned for use after 2010.

6.2.6.1.6 Managing Other Radioactive Wastes

Generally, ILW and LLW are disposed of by burial. Intermediate-level wastes generated from fuel reprocessing will be disposed of in deep geological facilities. Some low-level liquid wastes from reprocessing plants are discharged to the sea. These liquids include some distinctive materials such as ^{99}Tc that can be discerned hundreds of kilometers away. Such discharges are tightly controlled and regulated so that the maximum dose any individual receives is a small fraction of natural background radiation.

Nuclear power stations and reprocessing facilities release small quantities of radioactive gases to the atmosphere. Gases such as ^{85}Kr and ^{133}Xe are chemically inert, and gases such as ^{131}I have short half-lives. The net effect of these gases is too small to warrant further consideration.

Table 6.12 provides a summary of waste management adopted by countries throughout the world.

6.2.7 Nuclear Power Economics

Any discussion of the economics of nuclear power involves a comparison with other competitive electric generation technologies. The competing technologies are usually coal and natural gas.

Nuclear power costs include capital costs, fuel cycle costs, waste management costs and the cost of decommissioning after operation. The costs vary widely depending on the location of the generating plant. In countries such as China, Australia and the U.S. coal remains economically attractive because of large accessible coal resources. This advantage could be changed if a charge is made on carbon emissions. In other areas nuclear energy is competitive with fossil fuels even though nuclear costs include the cost of all waste disposal and decommissioning.

As previously stated, nuclear power costs include spent fuel management, plant decommissioning, and final waste disposal. These costs are not generally included in the costs of other power generation technologies.

Decommissioning costs are estimated to be 9%–15% of the initial cost of a nuclear plant. Because these costs are discounted over the life of the plant, they contribute only a few percent to the investment cost of the plant and have an even lower impact on the electricity generation cost. This impact in the U.S. is about 0.1–0.2 cent/kWh or about 5% of the cost of electricity produced.

Spent-fuel interim storage and ultimate disposal in a waste repository contribute another 10% to the cost of electricity produced. This cost is reduced if the spent fuel is disposed of directly. This does not account for the energy that could be extracted from the fuel if it was reprocessed.

Costs for nuclear-based electricity generation have been dropping over the last decade. This reduction in the cost of nuclear-generated electricity is a result of reductions in nuclear plant fuel, operating costs, and maintenance costs. However, the capital construction costs for nuclear plants are significantly higher than coal- and gas-fired plants. Because the capital cost of nuclear plants contribute more to the cost of electricity than coal- or gas-fired generation, the impact of changes in fuel, operation costs, and maintenance costs on the cost of electricity generation is less than those for coal- or gas-fired generation.

One of the primary contributors to the capital cost of nuclear plants has been the cost of money used to finance nuclear plant construction. The financing costs increase when the time required to license and construct a plant increases. Two factors are leading to the reduction in this portion of the cost. First, especially in the U.S., the licensing process is changing so that a plant receives both the construction permit and the operating license prior to the start of construction. Under this process, there is no large investment in plant hardware prior to completion of a significant portion of the licensing process, leading to a reduction in time required for the plant to begin producing revenue. Second, the new generation of nuclear

TABLE 6.12 Waste Management Policies for Spent Fuel for Countries Throughout the World

Country	Policy	Facilities and Progress Toward Final Disposition
Belgium	Reprocessing	Central waste storage and underground laboratory established Construction of repository to begin about 2035
Canada	Direct disposal	Underground repository laboratory established Repository planned for use 2025
China	Reprocessing	Central spent fuel storage in LanZhou
Finland	Direct disposal	Spent fuel storages in operation Low and intermediate-level repositories in operation since 1992 Site near Olkiluoto selected for deep repository for spent fuel, from 2020
France	Reprocessing	Two facilities for storage of short-lived wastes Site selection studies underway for deep repository for commissioning 2020
Germany	Reprocessing but moving to direct disposal	Low-level waste sites in use since 1975 Intermediate-level wastes stored at Ahaus Spent fuel storage at Ahaus and Gorleben High-level repository to be operational after 2010
India	Reprocessing	Research on deep geological disposal for HLW
Japan	Reprocessing	Low-level waste repository in operation High-level waste storage facility at Rokkasho-mura since 1995 Investigations for deep geological repository begun, operation from 2035
Russia	Reprocessing	Sites for final disposal under investigation Central repository for low and intermediate-level wastes planned from 2008 Central interim HLW store planned for 2016
South Korea	Direct disposal	Central low- and ILW repository planned from 2008 Investigating deep HLW repository sites Low and intermediate-level waste repository in operation
Spain	Direct disposal	Final HLW repository site selection program for commissioning 2020 Central spent fuel storage facility in operation since 1985
Sweden	Direct disposal	Final repository for low to intermediate waste in operation since 1988 Underground research laboratory for HLW repository Site selection for repository in two volunteered locations Central interim storage for high-level wastes at Zwilag since 2001 Central low and intermediate-level storages operating since 1993
Switzerland	Reprocessing	Underground research laboratory for high-level waste repository with deep repository to be finished by 2020 Low-level waste repository in operation since 1959
United Kingdom	Reprocessing	High-level waste is vitrified and stored at Sellafield Underground HLW repository planned
U.S.A	Direct disposal	Three low-level waste sites in operation 2002 decision to proceed with geological repository at Yucca Mountain

Source: From World Nuclear Association, Waste Management in the Nuclear Fuel Cycle, http://www.world-nuclear.org, 2004.

plants will be highly standardized and modularized. This will allow a significant reduction in the time required to construct a new plant. It is estimated that the time from the start of construction to the start of operation will be reduced from nearly 10 years to 4–5 years. This will have a significant impact on capital costs.

The reduced capital costs associated with the licensing and construction of new nuclear power plants, and the fact that nuclear power is inherently less susceptible to large fluctuations in fuel costs, have made nuclear power an attractive energy option for many countries seeking to diversify their energy mix in the face of rising fossil fuel costs.

6.2.8 Conclusions

The development of nuclear power began after World War II and continues today. The first power-generating plants were constructed in the late 1950 s. During the 1960 s and 1970 s, there was a large commitment to nuclear power until the accidents occurred at Three Mile Island in 1979 and then at Chernobyl in 1986. The new safety requirements and delays caused by these accidents drove up the costs and at the same time caused a loss of public acceptance. In the U.S., many plant orders were canceled; in other countries, entire nuclear programs were canceled.

The ability of nuclear reactors to produce electricity economically and safely without the generation of greenhouse gasses has revitalized the interest in nuclear power as an alternative energy source. Many lessons have been learned from the operation of current power plants that have allowed the safety of newly designed plants to be improved. This, coupled with the desire of many nations to develop secure energy sources and a diversity of energy options, have resulted in the continuing development of a whole new generation of nuclear plants to meet future energy needs.

Nuclear power is also not as susceptible to fluctuation in fuel costs as petroleum and natural gas. As shown, the supply of uranium is very large, and if it is supplemented with thorium, the fuel supply is seemingly unlimited. This drives many other aspects of the fuel cycle, such as the choice between closed and open fuel cycles discussed earlier. For example, because of the large uranium resource and the fears of nuclear proliferation, the once-through (open) fuel cycle is favored by many. This will require large deep geologic waste repositories for the disposal of large quantities of spent fuel. However, when reprocessing is included in the closed fuel cycle, the amount of needed repository space is greatly reduced, but the expense of operation is increased. Finally, it may be possible to essentially eliminate the need for repositories by utilizing advanced fuel cycles that utilize almost all of the energy available in the uranium and the other transuranic products of reactor operation.

The need for energy and the use of electricity as the primary energy source for the end user will drive the increase in electricity generation around the world. The drive to reduce the production of greenhouse gases will contribute to a wider use of nuclear power for electricity generation. The recognition that nuclear power can safely provide large base-load generating capacity at a reasonable cost using known technologies will also be a major factor in its future development.

References

1. CRC Press. 1997. Generation technologies through the year 2005. In *CRC handbook of energy efficiency*. CRC Press.
2. World Coal Institute. 2000. *Coal Power for Progress. 4th Ed*. World Coal Institute.
3. EPRI. 1999. *The Electricity Technology Roadmap: Powering Progress*. 1999 Summary and Synthesis. EPRI Report C1-112677-V1.
4. Armor, A. F. and Wolk, R. H. 2002. *Productivity Improvement Handbook for Fossil Steam Plants*. 3rd Ed. EPRI Report 1006315.
5. Armor, A. F. and Wolk, R. H. 2005. *Productivity Improvement for Fossil Steam Power Plants 2005: One Hundred Case Studies*. EPRI Report 1012098.

6. EPRI. 2000. *An Assessment of Mercury Emissions from U.S. Coal-Fired Power Plants*. EPRI Report 1000608.
7. Dalton, S. M., Viswanathan, R., Gehl, S. M., Armor, A. F., Purgert, R. 2001. Ultrasupercritical materials. In *DOE Clean Coal Conference*.
8. Armor, A. F., Viswanathan, R., and Dalton, S. M. 2003. Ultrasupercritical steam turbines: Design and materials issues for the next generation, coal utilization and fuel systems. In *DOE Annual Conference*.
9. US Department of Energy. 2001. DOE fossil energy-Tomorrow's turbines. http://fossil.energy.gov/coal_power/turbines/index.shtml (accessed on).
10. EPRI. 2002. *Atmospheric Fluidized—Bed Combustion Handbook*. EPRI Report 1004493.
11. Skowyra, R. S. et al. 1995. Design of a supercritical sliding pressure circulating fluidized bed boiler with vertical waterwalls. In *Proceedings of 13th International Conference on Fluidized Bed Combustion*
12. EPRI. 2002. *Technical Status, Operating Experience, and Risk Assessment of Clean Coal Technologies-2002*. EPRI Report 1004480.
13. Courtright, H. A., Armor, A. F., Holt, N. H., and Dalton, S. M. 2003. Clean coal technologies and their commercial development. In *POWER-GEN International, Conference Proceedings*.
14. EPRI. 2004. Decommissioning Handbook for Coal-Fired Power Plants. EPRI Report 1011220.
15. US Department of Energy. 2002. Clean Coal Technology Demonstration Program. DOE/FE-0444, US DOE.
16. El-Wakil, M. M. 1962. Nuclear Power Engineering. McGraw-Hill, New York.
17. Nuclear News. 2005.
18. American Nuclear Society. 2005. Nuclear News.
19. Deutch, J. et al. 2003. The Future of Nuclear Power. Massachusetts Institute of Technology, Cambridge, MA.
20. Sutherland, J. K. 2003. Nuclear power comparisons and perspectives, http://www.energypulse.net (accessed on April 26, 2004).
21. World Nuclear Association. 2005. Plans for new reactors worldwide. World Nuclear Association information paper, http://www.world-nuclear.org, (accessed on June 16, 2005).
22. World Nuclear Association. 2005. Advanced nuclear power reactors. World Nuclear Association information paper, http://www.world-nuclear.org, (accessed on June 3, 2005).
23. World Nuclear Association. 2001. The nuclear fuel cycle. World Nuclear Association information paper. http://www.world-nuclear.org, (accessed on June 22, 2005).
24. WISE Uranium Project. 2005. World nuclear fuel facilities. http://www.wise-uranium.org, (accessed on June 22, 2005).
25. World Nuclear Association. 2004. Waste management in the nuclear fuel cycle. World Nuclear Association information paper, http://www.world-nuclear.org, (accessed on June 29, 2005).
26. World Nuclear Association. 2005. Nuclear waste disposal concepts. World Nuclear Association information paper, http://www.world-nuclear.org (accessed on June 29, 2005).
27. Uranium Information Centre. Processing of used nuclear fuel. UIC Nuclear Issues Briefing Paper 72, http://www.uic.com.au (accessed on June 29, 2005).
28. Sutherland, J. K. 2003. Nuclear cycles and nuclear resources, http://www.energypulse.net, (accessed on April 26, 2004).
29. World Nuclear Association. 2003. Uranium enrichment. World Nuclear Association information paper, http://www.world-nuclear.org (accessed on June 22, 2005).
30. World Nuclear Association. 2003. Transport of nuclear materials. World Nuclear Association information paper, http://www.world-nuclear.org (accessed on June 22, 2005).
31. Finck, P. J. 2005. Congressional testimony on nuclear fuel reprocessing. http://www.anl.gov/Media_-Center/News/2005/testimony050616.html (accessed on June 29, 2005).
32. World Nuclear Association. 2004. Supply of uranium. World Nuclear Association information paper, http://www.world-nuclear.org (accessed on June 22, 2005).

33. World Nuclear Association. 2005. The economics of nuclear power. World Nuclear Association information paper, http://www.world-nuclear.org (accessed on June 3, 2005).
34. World Nuclear Association. 2004. Energy analysis of power systems. World Nuclear Association information paper, http://www.world-nuclear.org (accessed on June 16, 2005).
35. Bruschi, H. J. The Westinghouse AP1000—Final design approved. Nuclear News, (November 2004).
36. US DOE Nuclear Energy Research Advisory Committee and the Generation IV International Forum. 2002. A technology roadmap for the generation IV nuclear energy systems. GIF-002-00.

For Further Information

Energy Information Administration. Annual energy outlook. 2003. http://www.eia.doe.gov.
National Engineering Technology Laboratory. http://www.netl.doe.gov.
EPRI. http://www.epri.com.
Babcock and Wilcox Co. *Steam, its generation and use*. Babcock and Wilcox, New York.
Combustion Engineering, Inc. *Combustion: Fossil power systems*. Combustion Engineering, Inc, Windsor, CT.
EPRI Journal. 1986. Tapping global expertise in coal technology. *EPRI Journal*. (Jan/Feb).
IGCC. 1994. New fuels, new players. *EPRI Journal*. (July/August).
EPRI Journal. 1993. A brighter future for PFBC. *EPRI Journal*. (December).
EPRI Journal. 1991. Fuel cells for urban power. *EPRI Journal*. (September).
EPRI Journal. 1993. Distributed generation. *ERPI Journal*. (April/May).
EPRI Journal. 1995. Plant repowering. *EPRI Journal*. (September/October).
EPRI Journal. 1998. Smart materials. *EPRI Journal*. (July/August).
EPRI Journal. 1999. Merchant plants. *EPRI Journal*. (Summer).
EPRI Journal. 2000. Energy and air emissions. *EPRI Journal*. (Summer).
EPRI Journal. 2001. Global coal initiative. *EPRI Journal*. (Summer).

7

Outlook for U.S. Energy Consumption and Prices in the Midterm

Andy S. Kydes

Energy Information Administration

7.1 Introduction

All projections of energy-economic systems, particularly those that incorporate consumer and producer behavior, are inherently uncertain. However, consumer and producer behavior are not the only sources of uncertainty. Some of the other critical uncertainties on which projections depend include:

- The rate of technological progress for end use and supply technologies
- Changes in energy market regulations and efficiency standards
- The quantity, location, and depth of energy resources in the ground (e.g., coal, crude oil, natural gas, nuclear material)
- The costs to explore, locate, and ultimately produce energy resources
- The willingness of financial markets to make energy investments

Consequently, the projections described in this chapter are not statements of what will happen, but rather statements of what might happen given the assumptions and methodologies used.

The outlook of U.S. energy markets through 2025, as presented in this chapter in Section 7.1 through Section 7.9, was developed using the Energy Information Administration's (EIA) National Energy Modeling System (NEMS) [1]. With the exception of higher world oil prices, the October oil futures case projection uses the assumptions of the *Annual Energy Outlook 2005 (AEO2005)* reference case [2]. The reference case projection for *AEO2006* is available on the EIA Internet website.* Central to this scenario is the assumption that current technological and demographic trends and current laws and regulations will continue as usual into the future. The October oil futures case of *AEO2005* assumes that crude oil (using the imported refiner's acquisition cost (IRAC)), in real 2003 dollars, is almost $5 per barrel higher than the *AEO2005* reference case in 2025.†

The EIA does not propose, advocate, or speculate on future legislative or regulatory changes. Thus, the *AEO2005* projections, including the October oil futures case described in this chapter, provide a policy-neutral basis that can be used as an adjunct to analyze policy initiatives. For scenarios based on alternative macroeconomic growth rates, world oil prices (both higher and lower than those used in this outlook) or alternative rates of technological progress, the reader is encouraged to review the alternative scenarios developed in the *AEO2005* [3].

7.2 Key Energy Issues to 2025

World crude oil prices—defined by the U.S. average imported refiner's acquisition cost of crude oil (IRAC)—reached a recent low of $10.29 per barrel (in 2003 dollars) in December 1998. For the next three years, crude oil prices ranged between just under $20 and just over $30 per barrel. Since December 2001, however, prices have steadily increased to about $46 per barrel in October 2004 and into the mid-$60 dollar per barrel range in the second half of 2005.

Strong growth in the demand for oil worldwide, particularly in China and other developing countries, is generally cited as the driving force behind the sharp price increases seen over the 2002–2005 period. Other factors contributing to the upward world oil price trend include a tight supply situation that has shown only limited production response by countries outside of the Organization of Petroleum Exporting Countries (OPEC) to higher prices; changing views on the economics of oil (and gas) production; and concerns about economic and political situations in the Middle East, Venezuela, Nigeria, China, and the former Soviet Union. However, the rate of technological progress, new international emission control protocols such as the Kyoto Protocol and significant global concerns about the security

*http://www.eia.doe.gov/oiaf/aeo/index.html. See http://www.eia.doe.gov/oiaf/aeo/assumption/pdf/0554(2005).pdf for a detailed description of the assumptions of the *AEO2005* reference case.

†The outlook and commentary described herein represent solely the views of the author and not necessarily the views of EIA, the U.S. Department of Energy, the administration, or any agency of the U.S. government.

of the petroleum supply represent additional factors that may trigger actions that reduce the upward pressure on crude oil prices, i.e., by increasing production from previously unexpected or uneconomic sources or by reducing petroleum demand. For example, technological progress may lead to cost reductions for gas-to-liquids and coal-to-liquids technologies, the development of new methods and technologies to economically develop domestic oil shale resources that are currently plentiful but uneconomic, and the discovery of significant new 'finds' of crude oil in other relatively unexplored regions of the world. The future path of world oil prices is a major uncertainty facing world oil markets.

The October oil futures case was based on the October 2004 prices from the New York Mercantile Exchange (NYMEX) futures market (corrected for the difference between futures prices and the IRAC). The NYMEX crude oil outlook implies that the annual average price in 2005 will exceed the 2004 average price level, and that prices will then decline only slowly in constant 2003 dollars over the next five years, resulting in a 2010 price of about $31 per barrel.* The IRAC is then projected to rise to about $35 per barrel by 2025.

From 1986 to 2000, when U.S. natural gas consumption grew from 16.2 trillion $ft.^3$ to a high of 23.3 trillion $ft.^3$, 40% of the increased demand was met by imports, predominantly from Canada. Based on EIA's assessment of recent data and recent projections by Canada's National Energy Board, EIA has revised expectations about Canadian natural gas production, particularly coalbed methane and conventional production in Alberta. It is unlikely that future production from Canada will be able to support a continued increase in U.S. imports.

In the October oil futures case, U.S. natural gas consumption is projected to grow from 22 trillion $ft.^3$ in 2003 to over 30.5 trillion $ft.^3$ in 2025. Most of the additional supply is expected to come from Alaska and imports of liquefied natural gas (LNG). A key issue for U.S. energy markets is whether the investments and regulatory approvals needed to make those natural gas supplies available will occur. The following sections summarize the key trends in the projection. A summary of the projection is provided in Table 7.1.

7.3 Economic Growth

The output of the Nation's economy, measured by gross domestic product (GDP), is projected to grow by 3.1% per year between 2003 and 2025 (with GDP based on 2000 chain-weighted dollars). Figure 7.1 shows the trend in the annual real growth for GDP, including the projection. The labor force is projected to increase by 0.9% per year between 2003 and 2025. Labor productivity growth in the nonfarm business sector is projected to be over 2.0% per year.

Compared with the second half of the 1990s, the rates of growth in GDP and nonfarm employment were lower from 2000 through 2002. Economic growth has been more robust since 2003. Real GDP growth was 3.0% in 2003 and 4.4% in 2004. Total population growth (including armed forces overseas) is expected to remain fairly constant after 2003, growing by 0.8% per year on average. Labor force growth is expected to slow as a result of demographic changes, but more people over 65 are expected to remain in the work force. Nonfarm business productivity growth has been strong recently, averaging 3.8% per year from 2000 to 2003. Productivity growth from 2003 to 2025 is expected to average about 2% per year, supported by investment growth of about 5.0% per year.

*There are large oil projects world-wide that are scheduled to come on line between 2008 and 2009 which should increase supply; the high world oil prices will dampen world petroleum demand. Between the lower demand resulting form higher prices and the increased supply which is in response to the higher prices, we should see a dip from current prices, unless you expect to see continued or increasing war fears and global instability. There is no question that the current situation with the potential for war-generated supply disruptions could send prices into the $100–$200 per barrel range for a while. I don't consider that situation "normal." If such instability is sustained another 2–3 years, the situation could become part of the "normal" view and then I would be inclined to agree on the higher prices. So if one believes that the current level of fear and instability that the world will face in the next 25 years is likely, high sustained prices over the next 10–20 years are likely. Financial incentives to find and develop alternative technologies for petroleum substitutes and higher-efficiency end use appliances will be high and will eventually put downward price pressures latter as they significantly penetrate the market.

TABLE 7.1 October Oil Futures Case (AEO2005): Total Energy Supply and Disposition Summary (Quadrillion Btu per Year, Unless Otherwise Noted)

	2002	2003	2005	2010	2015	2020	2025	2003–2025 (%)
Primary energy production (quadrillion Btu)								
Petroleum	14.71	14.38	14.67	15.74	14.98	14.22	13.38	0.3
Dry natural gas	19.48	19.58	19.60	21.34	21.89	23.09	22.72	0.7
Coal	22.70	22.66	24.12	25.15	25.52	27.07	29.91	1.3
Nuclear power	8.14	7.97	8.31	8.49	8.62	8.67	8.67	0.4
Renewable energy	5.79	5.89	6.48	6.85	7.15	7.60	8.26	1.5
Other	1.12	0.93	1.02	1.16	0.94	0.90	0.97	0.2
Total production	71.94	71.42	74.20	78.64	79.09	81.54	83.91	0.7
Net imports (quadrillion Btu)								
Petroleum	22.64	24.10	25.64	27.15	31.23	35.09	38.96	2.2
Natural gas	3.59	3.32	3.43	4.86	6.97	7.98	8.83	4.5
Coal/other (indicates export)	0.47	0.43	0.48	0.14	0.19	0.25	0.58	N/A
Total net imports	25.75	26.99	28.55	31.87	38.39	43.32	48.37	2.7
Consumption (quadrillion Btu)								
Petroleum products	38.41	39.09	40.28	43.98	47.11	50.22	53.21	1.4
Natural gas	23.59	22.54	22.892	26.29	29.01	31.21	31.71	1.6
Coal	21.98	22.71	23.27	24.91	25.67	27.31	30.50	1.3
Nuclear power	8.14	7.97	8.31	8.49	8.62	8.67	8.67	0.4
Renewable energy	5.79	5.89	6.48	6.85	7.15	7.61	8.27	1.6
Other	0.07	0.02	0.01	0.03	0.07	0.05	0.04	4.1
Total consumption	97.99	98.22	101.2	110.6	117.6	125.10	132.4	1.3
Petroleum (million barrels per day)								
Domestic crude production	5.74	5.68	5.76	6.16	5.78	5.36	4.98	0.5
Other domestic production	3.60	3.72	3.77	3.95	3.77	4.19	4.29	1.0
Net imports	10.54	11.24	11.6	12.61	14.44	16.21	18.06	2.1
Consumption	19.71	20.00	20.65	22.54	24.1	25.76	27.30	1.4
Natural gas (trillion ft.^{3})								
Production	19.03	19.13	19.16	20.86	21.39	22.56	22.20	0.6
Net imports	3.50	3.24	3.34	4.73	6.79	7.77	8.60	4.6
Consumption	22.98	21.95	22.29	25.61	28.27	30.42	30.91	1.5

Coal (million short tons)								
Production	1105	1083	1178	1235	1268	1344	1484	1.5
Net imports	−23	−18	−23	−9	3	7	20	N/A
Consumption	1066	1095	1136	1227	1271	1351	1504	1.5
Prices (2003 dollars)								
World oil price (dollars per barrel)	24.10	27.73	43.63	30.99	32.33	33.67	35.00	1.0
Gas Wellhead price (dollars per thousand cubic ft.)	3.06	4.98	5.56	3.63	4.11	4.45	4.83	0.2
Coal Minemouth price (dollars per ton)	18.23	17.93	18.78	17.45	16.99	17.54	18.52	0.1
Average electricity (cents per kilowatthour)	7.4	7.4	7.4	6.6	7.0	7.2	7.3	0.1
Economic indicators								
Real gross domestic product (billion 2000 dollars)	10,075	10,381	11,191	13,063	15,216	17,641	20,293	3.1
GDP Chaintype price index (index, 2000 = 1.000)	1.041	1.060	1.106	1.219	1.371	1.559	1.811	2.5
Real disposable personal income (billion 2000 dollars)	7560	7734	8213	9540	11,152	12,745	14,945	3.0
Value of industrial shipments (billion 1996 dollars)	5067	5105	5464	6167	6872	7669	8506	2.3
Energy intensity (thousand Btu per 2000 dollar of GDP)	9.73	9.47	9.05	8.47	7.74	7.09	6.53	1.7
Carbon dioxide emissions (million metric tons)	5751	5789	5982	6561	6988	7461	7981	1.5

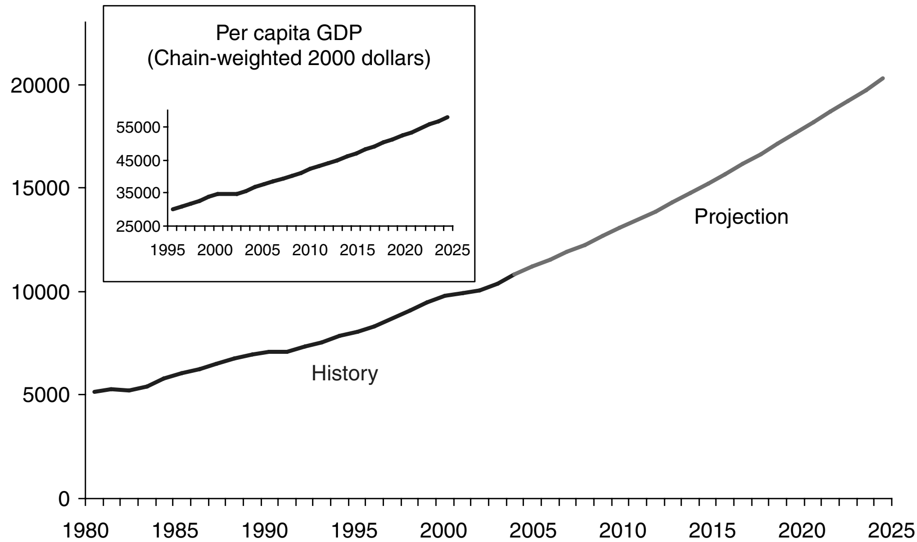


FIGURE 7.1 October oil futures case, U.S. gross domestic product 1982–2025 (billions of chain weighted year 2000 dollars).

From 2003 through 2025, personal disposable income is projected to grow by about 3% per year and disposable income per capita by 2.2% per year. Nonfarm employment is projected to grow by 1.2% per year, and employment in manufacturing is projected to shrink by about 0.6% per year. From 2003 to 2025, industrial output in real value terms is projected to grow by 2.3% per year, compared with 3.2% average annual growth in the services sector.

7.4 Energy Prices

The annual average imported refiners' acquisition cost (IRAC) of crude oil in this projection increases from $27.73 per barrel (2003 dollars) in 2003 to $46.00 per barrel in 2004 and then declines to $31 per barrel in 2010 as new supplies enter the market. It then rises to $35 per barrel in 2025 (Figure 7.2). Note that fuel prices in the graph have all been converted to dollars per million Btu to allow for comparison. In 2025, the average world oil price is expected to be about $62 per barrel in 2025 dollars. The NEMS projection provides the price of a basket of crude oil with types ranging from light 'sweet' (low specific gravity and low sulfur) to heavy 'sour' (high specific gravity and high sulfur). Many readers are accustomed to seeing the price of west Texas intermediate (WTI) crude oil which has typically commanded a premium of from $1.00 to $8.00 dollars per barrel above the IRAC (in 2004 dollars). In 2004 dollars, the average margin between WTI and IRAC for the period 1984 to 2004 is $2.70 per barrel. Like WTI, imported light sweet crude oil has also been sold at a price premium relative to IRAC on the world oil market, with a premium ranging between about $0.50 to $5.00 dollars per barrel above the IRAC price. Because the United States has a large refinery capacity to process heavy sour crude oils, use of the IRAC price rather than the WTI price or the price of imported light sweet crude is more representative of crude oil prices in U.S. energy markets. For countries with refinery capacity that can mostly handle only light sweet crude oils, the reader may assume a price which is about $1 to $5 per barrel higher than the IRAC price assumed for the United States in this projection. The specific value will depend on the extent to which foreign refiners also increase their ability to handle heavy sour crude oils over the forecast horizon.

The projected world oil price forecast is characterized by decreasing prices through 2010 and moderately increasing prices thereafter. This is consistent with a forecast that projects increases in world petroleum demand from about 80 million barrels per day in 2003 to about 116 million barrels per

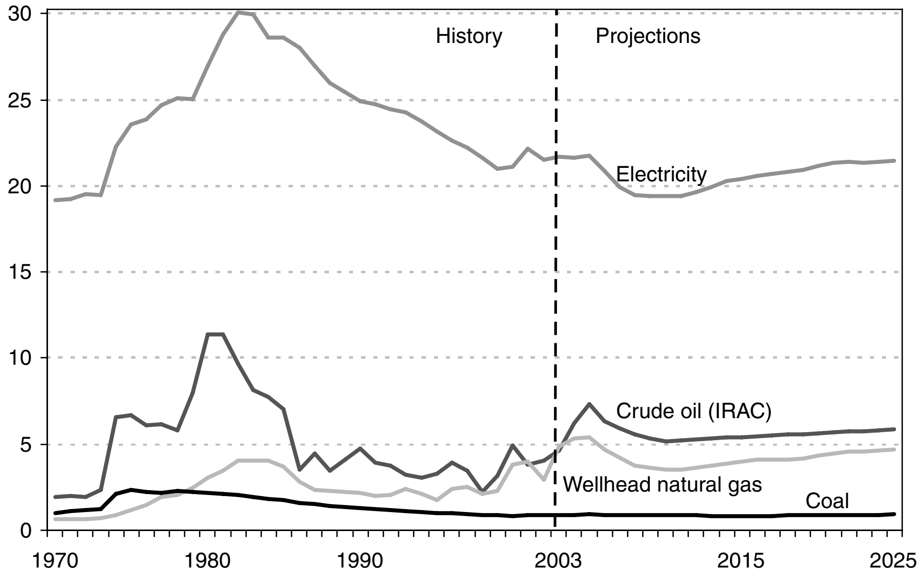


FIGURE 7.2 October oil futures case, energy prices, 1970–2025 (2003 dollars per million Btu).

day in 2025. The projected demand is met by increased oil production both from the OPEC and non-OPEC nations. OPEC oil production is projected to rise to 48.5 million barrels per day in 2025, 57% higher than the 30.6 million barrels per day produced in 2003. The forecast assumes that OPEC will pursue policies intended to increase production, that sufficient resources exist, and that access and capital will be available to expand production. Non-OPEC oil production is expected to increase from 49 to 67.8 million barrels per day between 2003 and 2025.

Average wellhead prices in constant 2003 dollars for natural gas in the United States are projected generally to decrease, from \$5.47 per thousand ft.^{3} (\$5.32 per million Btu) in 2004 to \$3.63 per thousand ft.^{3} (\$3.53 per million Btu) in 2010 as the availability of new natural gas import sources and increased drilling expands available supply. After 2010, wellhead prices are projected to increase gradually, to \$4.83 per thousand ft.^{3} (\$4.70 per million Btu) in 2025 (equivalent to about \$8.50 per thousand ft.^{3} in nominal (current-year) dollars) as demand increases require production from higher-cost domestic unconventional natural gas sources. Growth in LNG imports, Alaska production, and lower-48 production from non-conventional sources are not expected to increase sufficiently to offset the impacts of resource depletion and increased demand.

In the projection, the combination of more moderate increases in coal production, expected improvements in mine productivity, and a continuing shift to low-cost coal from the Powder River Basin in Wyoming leads to a gradual decline in the average minemouth price to approximately \$17.00 per ton (\$0.84 per million Btu) in 2003 dollars in 2015. The price is projected to remain nearly constant between 2015 and 2020, increasing after 2020 as rising natural gas prices and the need for baseload generating capacity lead to the construction of many new coal-fired generating plants. By 2025, the average minemouth price (in 2003 dollars) is projected to be \$18.52 per ton (\$0.92 per million Btu). The projected minemouth coal price in nominal (current-year) terms is expected to be \$31.25 per in 2025.

Average delivered electricity prices in 2003 dollars are projected to decline from 7.4 cents per kWh (\$21.74 per million Btu) in 2003 to a low of 6.6 cents per kWh (\$19.38 per million Btu) in 2010 as a result of an increasingly competitive generation market and a decline in natural gas prices. After 2010, average real electricity prices are projected to increase, reaching 7.3 cents per kWh (\$21.47 per million Btu) in 2025 (equivalent to 12.9 cents per kWh in nominal (current-year) dollars).

7.5 Energy Consumption

Total primary energy consumption in the projection is expected to increase from 98.2 quadrillion Btu in 2003 to 132.4 quadrillion Btu in 2025, an average annual increase of 1.3%, which is less than half the annual rate of growth for projected for GDP.

7.5.1 Residential Energy Consumption

Consistent with population growth rates and household formation, delivered residential energy consumption is projected to grow from 11.6 quadrillion Btu in 2003 to 14.1 quadrillion Btu in 2025 (Figure 7.3), at an average rate of 0.9% per year between 2003 and 2025 (1.0% per year between 2003 and 2010, slowing to 0.8% per year between 2010 and 2025). The most rapid growth in residential energy demand in the projection is expected to be for electricity used to power computers, electronic equipment, and appliances. Natural gas use in the residential sector is projected to grow at an annual rate of 1.1% from 2003 to 2010 and 0.6% from 2010 to 2025.

The projection includes changes in the residential sector that have offsetting influences on the forecast of energy consumption, including more rapid growth in the total number of U.S. households, higher delivered prices for natural gas, electricity, and distillate fuel, and a better accounting of additions to existing homes and the height of ceilings in new homes.

7.5.2 Commercial Energy Consumption

The forecast for commercial energy consumption is largely driven by an expected annual rate of growth in commercial floor space that is projected to average 1.7% per year between 2003 and 2025. Consistent with the projected increase in commercial floor space, delivered commercial energy consumption is projected to grow at an average annual rate of 1.9% between 2003 and 2025, reaching 12.4 quadrillion Btu in 2025. The most rapid increase in commercial energy demand is projected for electricity used for computers, office equipment, telecommunications, and miscellaneous small appliances.

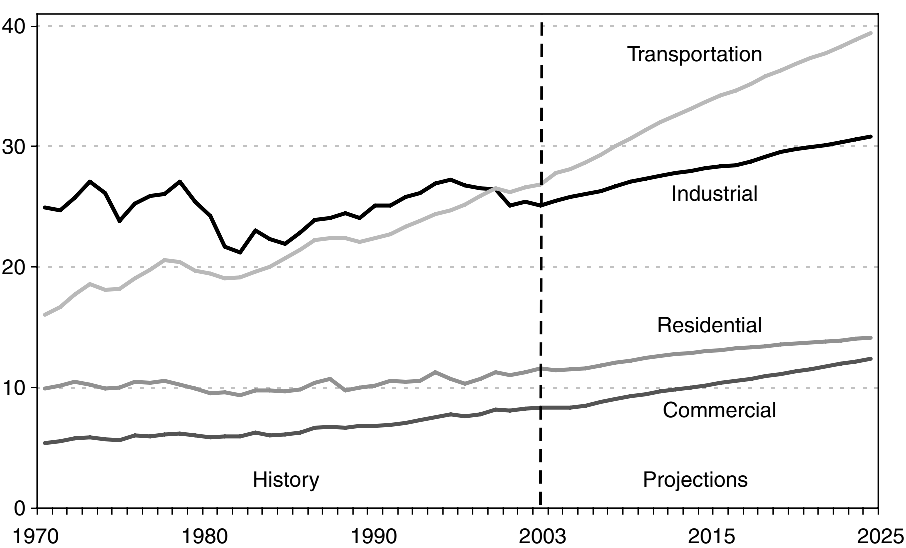


FIGURE 7.3 October oil futures case, delivered energy consumption by sector, 1970–2025 (quadrillion Btu).

7.5.3 Industrial Energy Consumption

Industrial energy consumption in the October oil futures case is projected to increase at an average rate of 0.9% per year between 2003 and 2025, reaching 30.8 quadrillion Btu in 2025. Key to the slower growth rate of industrial energy consumption as compared to the annual U.S. energy consumption rate of 1.3% is a continued shift of the U.S. economy toward services and away from energy-intensive industries. The value of shipments, a measure of industrial economic activity, is projected to increase at an annual rate of 2.3% as compared to the annual growth in the economy as a whole of 3.1%.

7.5.4 Transportation Energy Consumption

Energy consumption in the transportation sector is projected to grow at an average annual rate of 1.7% between 2003 and 2025 in the projection, reaching 39.4 quadrillion Btu in 2025. The growth in transportation energy demand is largely driven by the increasing personal disposable income, projected to grow annually at about 3%, consumer preferences for driving larger cars with more horsepower, and an increase in the share of light trucks and sports utility vehicles that make up light-duty vehicles. Total vehicle miles traveled by light-duty vehicles is projected to increase at an annual rate of 2% between 2003 and 2025 because of the increase in personal disposable income and other demographic factors.

7.5.5 Electricity Sector

Total electricity consumption, including both purchases from electric power producers and on-site generation, is projected to grow from 3657 billion kWh in 2003 to 5470 billion kWh in 2025, increasing at an average rate of about 1.8% per year. Rapid growth in electricity use for computers, office equipment, and a variety of electrical appliances in the end use sectors is partially offset in the forecast by improved efficiency in these and other, more traditional electrical applications and by slower growth in electricity demand in the industrial sector.

7.5.6 Demand for Natural Gas

Total demand for natural gas is projected to increase at an average annual rate of 1.5% from 2003 to 2025 (Figure 7.4), primarily as a result of increasing use for electricity generation and industrial applications that together account for about 75% of the projected growth in natural gas demand from 2003 to 2025. Total projected consumption of natural gas in 2025 is 30.9 trillion $ft.^3$, or 31.8 quadrillion Btu. The growth in demand for natural gas slows in the later years of the forecast (0.9% per year from 2015 to 2025, compared with 2.2% per year from 2003 to 2010), as rising natural gas prices lead to the construction of more coal-fired capacity for electricity generation.

7.5.7 Demand for Coal

Total coal consumption is projected to increase from 1095 million short tons (about 22.7 quadrillion Btu) in 2003 to 1505 million short tons (30.5 quadrillion Btu) in 2025, an annual growth rate of 1.5% per year. The increase in coal consumption results almost entirely from increased coal demand for electricity generation because of the expected price increases of natural gas. Total coal consumption for electricity generation is projected to increase by an average of 1.6% per year, from 1004 million short tons (20.5 quadrillion Btu) in 2003 to 1421 million short tons (28.6 quadrillion Btu) in 2025.

7.5.8 Demand for Petroleum

Total petroleum demand is projected to grow at an average annual rate of 1.4% in the projection, from 20.0 million barrels per day (39.1 quadrillion Btu) in 2003 to 27.3 million barrels per day

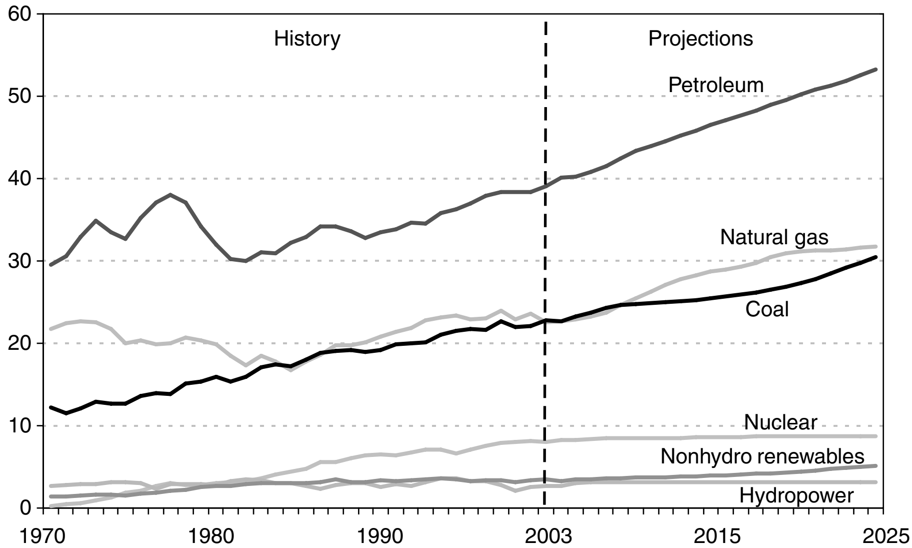


FIGURE 7.4 October oil futures case, energy consumption by fuel, 1970–2025 (quadrillion Btu).

(53.2 quadrillion Btu) in 2025. Almost all of the growth is attributable to the growth in transportation sector demand for petroleum products.

Total marketed renewable fuel consumption (including ethanol for gasoline blending, of which 0.2 quadrillion Btu is included with "petroleum products" consumption in Table 7.1), is projected to grow by 1.6% per year in the projection, from 6.1 quadrillion Btu in 2003 to 8.8 quadrillion Btu in 2025, as a result of state mandates for renewable electricity generation, higher natural gas prices, and the effect of production tax credits. About 60% of the projected demand for renewables in 2025 is for grid-related electricity generation (including combined heat and power), and the rest is for dispersed heating and cooling, industrial uses, and fuel blending. Renewable generating technologies are usually not as competitive in the projection as natural-gas-fired or advanced coal-fired technologies, because the costs for natural gas and advanced coal-fired technologies are usually lower, on a kWh basis, than those for renewable generation.

7.6 Energy Intensity

Energy intensity, as measured by energy use per 2000 dollar of GDP, is projected to decline at an average annual rate of 1.6% in the projection, with efficiency gains and structural shifts in the economy offsetting growth in demand for energy services (Figure 7.5). The projected rate of decline in the October oil futures case falls between the historical averages of 2.3% per year from 1970 to 1986, when energy prices increased in real terms, and 0.7% per year from 1986 to 1992, when energy prices were generally falling.

Since 1992, energy intensity has declined on average by 1.9% per year. During this period, the role of energy-intensive industries in the U.S. economy has fallen sharply. The share of industrial output from the energy-intensive industries declined on average by 1.3% per year from 1992 to 2003. In the projection, the energy-intensive industries' share of total industrial output is projected to continue declining but at a slower rate of 0.8% per year, which leads to the projected slower annual rate of reduction in energy intensity.

Historically, energy use per person has varied over time with the level of economic growth, weather conditions, and energy prices, among many other factors. During the late 1970s and early 1980s, energy

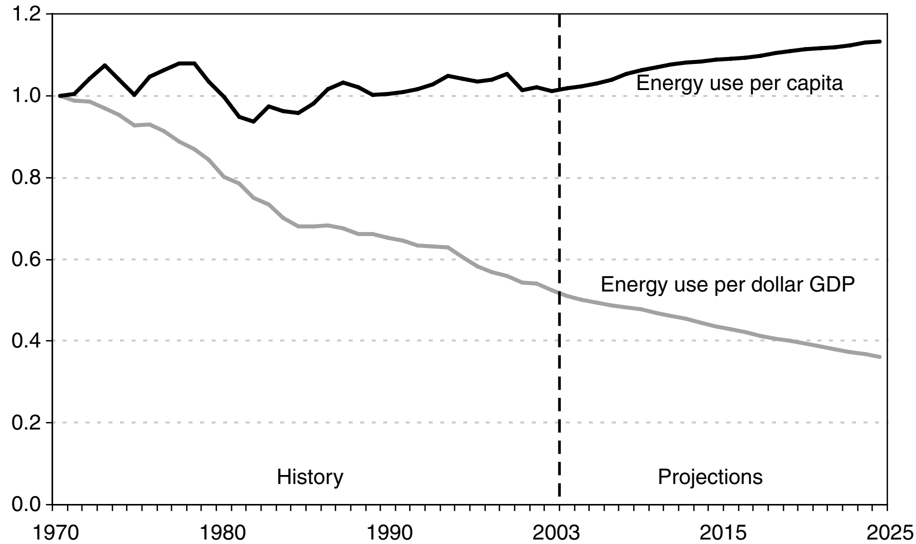


FIGURE 7.5 October oil futures case, energy use per capita and per dollar of gross domestic product, 1970–2025 (index, 1970 = 1).

consumption per capita fell in response to high energy prices and weak economic growth. Starting in the late 1980s and lasting through the mid-1990s, energy consumption per capita increased with declining energy prices and strong economic growth. Per capita energy use is expected to increase in this projection, with growth in demand for energy services only partially offset by efficiency gains. Per capita energy use increases by an average of almost 0.5% per year between 2003 and 2025 in the projection.

The potential for more energy conservation has received increased attention recently as energy prices have risen. Although energy conservation is projected to be induced through energy price increases, the projection does not assume policy-induced conservation measures beyond those in existing legislation and regulation, nor does it assume behavioral changes beyond those experienced in the past.

7.7 Electricity Generation

The natural gas share of electricity generation (including generation in the end use sectors) is projected to increase from 16% in 2003 to 24% in 2025. The share from coal is projected to decrease from 51% in 2003 to 50% in 2025. The projection estimates that 89 GW of new coal-fired generating capacity will be constructed between 2004 and 2025.

In the electric power sector, natural gas consumption increases from 5.0 trillion $ft.^3$ in 2003 to 9.4 trillion $ft.^3$ in 2025, accounting for about 31% of total demand for natural gas in 2025 as compared with 23% in 2003. The increase in natural gas consumption for electricity generation results from both the construction of new gas-fired generating plants and higher capacity utilization at existing plants. Most new electricity generation capacity is expected to be fueled by natural gas, because natural-gas-fired generators are projected to have advantages over coal-fired generators that include lower capital costs, higher fuel efficiency, shorter construction lead times, and lower emissions. Toward the end of the forecast, however, when natural gas prices rise substantially, coal-fired power plants are expected to be competitive for new capacity additions.

Nuclear generating capacity in the October oil futures case is projected to increase from 99.2 GW in 2003 to 102.7 GW in 2025 as a result of upgrades of existing plants between 2003 and 2025. All existing

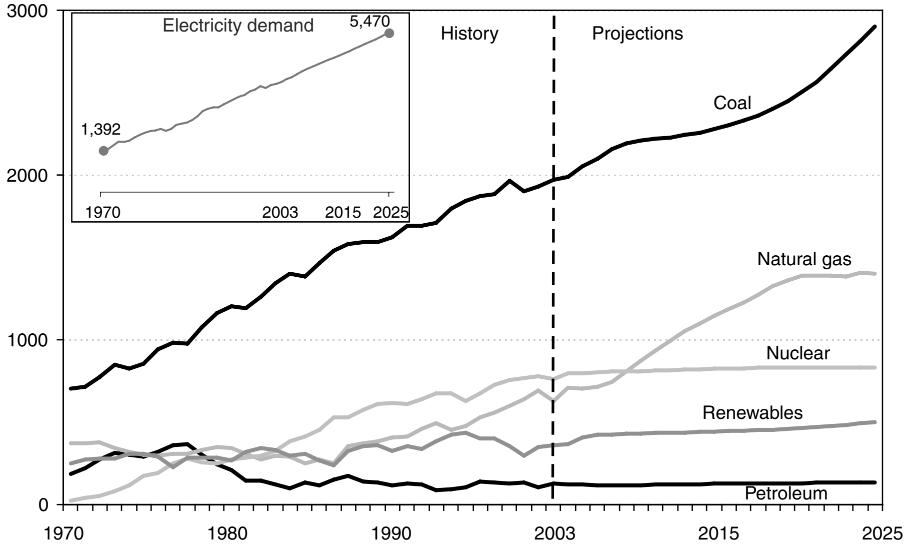


FIGURE 7.6 October oil futures case, electricity generation by fuel, 1970–2025 (billion kilowatthours).

nuclear plants are projected to continue to operate, but new plants are not expected to be economical. Total nuclear generation is projected to grow from 764 billion kWh in 2003 to 830 billion kWh in 2025 in the projection (Figure 7.6).

The use of renewable technologies for electricity generation is projected to grow slowly, both because of the relatively low costs of fossil-fired (primarily natural gas and coal) generation and because competitive electricity markets favor less capital-intensive technologies. Where enacted, state renewable portfolio standards, which specify a minimum share of generation or sales from renewable sources, are included in the forecast. The projection also includes the extension of the production tax credit for wind and biomass through December 31, 2005, as enacted in H.R. 1308, the Working Families Tax Relief Act of 2004. Current Congressional energy bills pending before Congress, including H.R. 6 (or the Senate energy bill) have not been included in this projection because they had not been enacted at the time this projection was developed.

Total renewable generation, including combined heat and power generation, is projected to grow from 359 billion kWh in 2003 to 497 billion kWh in 2025, increasing by 1.5% per year.

7.8 Energy Production and Imports

Total energy consumption is expected to increase more rapidly than domestic energy supply through 2025. As a result, net imports of energy are projected to meet a growing share of energy demand (Figure 7.7). Net imports are expected to constitute 38% of total U.S. energy consumption in 2025, up from 27% in 2003.

7.8.1 Petroleum Supply and Imports

Projected U.S. crude oil production is projected to increases from 5.7 million barrels per day in 2003 to a peak of 6.3 million barrels per day in 2009 as a result of expected higher prices than the 1990s, increased production offshore, predominantly from the deep waters of the Gulf of Mexico. Beginning in 2010, U.S.

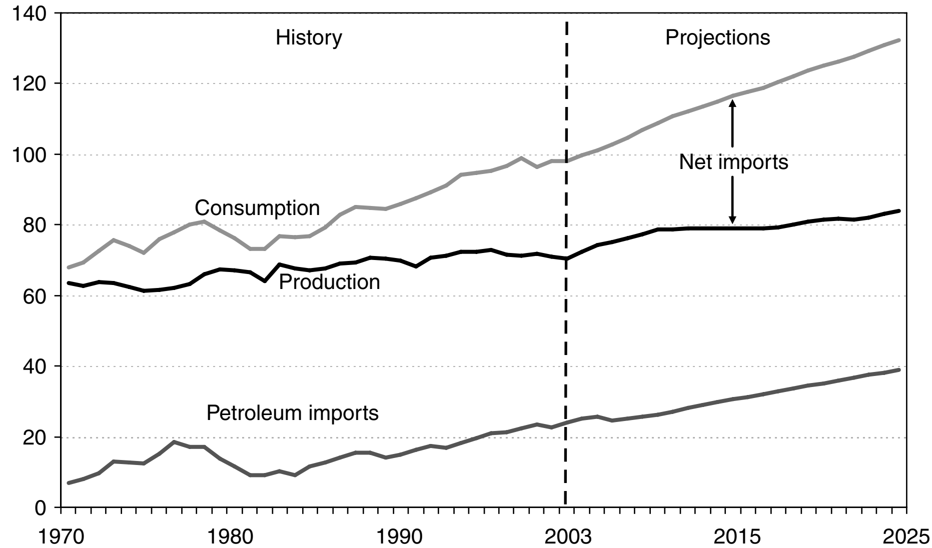


FIGURE 7.7 October oil futures case, total energy production and consumption, 1970–2025 (quadrillion Btu).

crude oil production begins to decline, falling to 5.0 million barrels per day in 2025 (Table 7.1) as depletion effects dominate technological advances and gradual price increases.

Total domestic petroleum supply (crude oil, natural gas plant liquids, refinery processing gains, and other refinery inputs) follows the same pattern as crude oil production in the forecast, increasing from 9.1 million barrels per day in 2003 to a peak of 10.0 million barrels per day in 2009, then declining to 9.3 million barrels per day in 2025.

In 2025, net petroleum imports, including both crude oil and refined products, are expected to account for 66% of demand (on the basis of barrels per day), up from 56% in 2003. Net imports of refined petroleum products account for between 12% and 14% of total net imports for the forecast horizon.

7.8.2 Natural Gas Supply and Imports

Growth in U.S. natural gas supplies will depend on unconventional domestic production, Canadian imports, natural gas from Alaska, and imports of LNG. Domestic natural gas production is projected to increase from 19.1 trillion $ft.^3$ in 2003 to 22.2 quadrillion Btu in 2025 in the projection (Table 7.1). Lower-48 onshore natural gas production is projected to increase from 13.9 trillion $ft.^3$ in 2003 to a peak of 15.7 trillion $ft.^3$ in 2012 before falling to 14.7 trillion $ft.^3$ in 2025. Lower-48 offshore production, which was 4.7 trillion $ft.^3$ in 2003, is projected to increase in the near term (to 5.6 trillion $ft.^3$ by 2015) because of the expected development of some large deepwater fields, including Mad Dog, Entrada, and Thunder Horse. After 2015, offshore production is projected to decline to 5.3 trillion $ft.^3$ in 2025.

Unconventional production is expected to become the largest source of U.S. natural gas supply. As a result of technological improvements and rising natural gas prices, natural gas production from relatively abundant unconventional sources (tight sands, shale, and coalbed methane) is projected to increase more rapidly than conventional production. Lower-48 unconventional gas production grows from 6.6 trillion $ft.^3$ in 2003 to 8.6 trillion $ft.^3$ in 2025 and from 35% of total lower-48 production in 2003 to 43% in 2025.

Production of lower-48 non-associated (NA) conventional natural gas declines from 9.5 trillion ft^3 in 2003 to 8.9 trillion $ft.^3$ in 2025, as resource depletion causes exploration and development costs to increase. Offshore NA natural gas production is projected to rise slowly to a peak of 3.9 trillion $ft.^3$ in 2008, then decline to 3.8 trillion $ft.^3$ in 2025.

Production of associated-dissolved (AD) natural gas from lower-48 crude oil reserves is projected to increase from 2.5 trillion $ft.^3$ in 2003 to 3.2 trillion $ft.^3$ in 2010 due to a projected increase in offshore AD gas production. After 2010, both onshore and offshore AD gas production are projected to decline, and total lower-48 AD gas production falls to 2.5 trillion $ft.^3$ in 2025.

Decreases in natural gas imports from Canada are expected to be offset by substantial increases in LNG imports and the development of the North Slope Alaska natural gas pipeline. Canadian imports are projected to decline from 2003 levels of 3.1 trillion $ft.^3$ to about 2.5 trillion $ft.^3$ in 2010, followed by an increase after 2010 to 3.1 trillion $ft.^3$ in 2016 as a result of rising natural gas prices, the introduction of gas from the Mackenzie Delta, and increased production of coalbed methane. After 2016, because of reserve depletion effects and growing domestic demand in Canada, net U.S. imports of Canadian natural gas are projected to decline to 2.7 trillion $ft.^3$ in 2025. That is, pipeline imports from Canada decline at the end of the forecast, because Canada's gas consumption increases more rapidly than its production. The forecasted supply of natural gas from Canada reflects revised Energy Information Administration (EIA) expectations about Canadian natural gas production, particularly coalbed methane and conventional production in Alberta, based in part on data and projections from Canada's National Energy Board and other sources.

With the exception of the facility at Everett, Massachusetts, three of the four existing U.S. LNG terminals (Cove Point, Maryland; Elba Island, Georgia; and Lake Charles, Louisiana) are expected to expand by 2007; and additional facilities are expected to be built in New England and elsewhere in the lower-48 States, serving the Gulf, mid-Atlantic, and south Atlantic states, including a new facility in the Bahamas serving Florida via a pipeline. Another facility is projected to be built in Baja California, Mexico, serving a portion of the California market. Total net LNG imports to the United States and the Bahamas are projected to increase from 0.4 trillion $ft.^3$ in 2003 to 6.2 trillion $ft.^3$ in 2025.

The North Slope Alaska natural gas pipeline is projected to begin transporting Alaskan natural gas to the lower-48 states in 2017 as a result of favorable investment economics and increasing natural gas prices. In 2025, total Alaskan natural gas production is projected to be 2.2 trillion $ft.^3$ in the October oil futures case, compared with 0.4 trillion $ft.^3$ in 2003.

7.8.3 Coal Supply

As domestic coal demand grows in the forecast, U.S. coal production is projected to increase at an average rate of 1.4% per year, from 1083 million short tons in 2003 to 1484 million short tons in 2025. Production from mines west of the Mississippi River is expected to provide the largest share of the incremental coal production. In 2025, nearly two-thirds of coal production is projected to originate from the western states.

7.9 Carbon Dioxide Emissions

Carbon dioxide emissions from energy use are projected to increase from 5789 million metric tons in 2003 to 7981 million metric tons in 2025, an average annual increase of 1.5% (Figure 7.8). The carbon dioxide emissions intensity of the U.S. economy is projected to fall from 558 metric tons per million dollars of GDP in 2003 to 393 metric tons per million dollars in 2025—an average decline of 1.6% per year.

By sector, including the emissions associated with the electricity consumed, carbon dioxide emissions for 2003 to 2025 are projected to grow at 1.1% per year in the residential sector, 2.1% per year for the commercial sector, 1.0% per year for the industrial sector, and 1.7% per year for transportation. Power sector carbon dioxide emissions are expected to grow at 1.7% per year but these are already included in the sectoral emissions listed above.

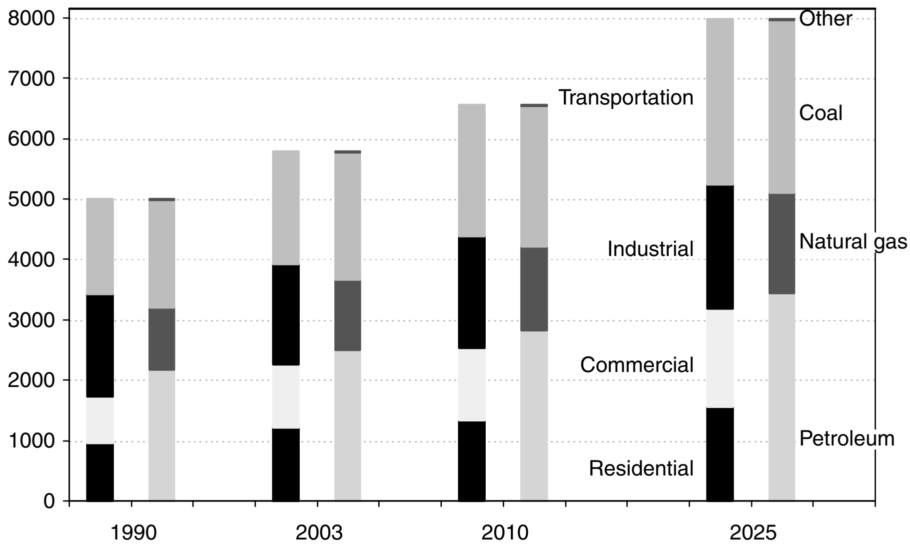


FIGURE 7.8 October oil futures case, U.S. carbon dioxide emissions by sector and fuel, 1990–2025 (million metric tons)-in progress.

7.10 Summary of the AEO2006 Reference Case Projection

7.10.1 Major Changes Reflected in the AEO2006 Reference Case

AEO2006 is the first edition of the *Annual Energy Outlook* (*AEO)* to provide projections through 2030. The *AEO2006* reference case projection incorporates new regulatory changes and laws that were issued after the release of *AEO2005* and before completion of *AEO2006*. Major regulatory changes that affect air emissions from electricity power plants include two rules by the Environmental Protection Agency (EPA) that were not represented in the October oil futures case: the Clean Air Interstate Rule (CAIR), issued on March 10, 2005 and the Clean Air Mercury Rule (CAMR), issued on March 15, 2005.* The new rules are expected to significantly reduce sulfur dioxide, nitrogen oxides, and mercury emissions from electric power plants over the next two decades.

The Energy Policy Act of 2005 (*EPACT2005*, Public Law 109–58), a statute that was signed into to law on August 8, 2005, is also new for *AEO2006*. The act provides tax incentives and loan guarantees for energy production of various types including new nuclear generation (up to 6 GW), integrated gasification combined cycle (IGCC) coal power generation, renewable generation, ethanol use for transportation, and a variety of efficiency programs intended to reduce energy consumption, primarily in the residential and commercial sectors. The *AEO2006* reference case only includes those sections of the *EPACT2005* that establish specific tax credits, incentives, or standards, comprising about 30 of the roughly 500 sections in the legislation.

The world crude oil prices in the *AEO2006* reference case are also significantly higher than those in the October oil futures case. The world crude oil price is now expressed in terms of the average price of imported *low-sulfur* crude oil to U.S. refiners to provide a crude oil price that is more consistent with those reported in the media. Light crude oil import prices are projected to increase from $40.49 per barrel (2004 dollars) in 2004 to $54.08 per barrel in 2025 and to $56.97 per barrel in 2030.

*see http://www.epa.gov/CAIR/ and http://www.epa.gov/oar/mercuryrule/

The higher prices for imported light, low-sulfur crude oil result from major changes in expectations for key drivers of the crude oil market. First, major oil companies lack access to resources in some key oil-rich countries and that situation is not expected to change over the projection period. Second, in attractive oil resource regions where foreign investment may be welcome, political instability and economic risk limit foreign investments. Third, revenues from national oil companies in key oil-rich countries flow to governments which are facing increasing demands placed on them for social, welfare, and physical infrastructure projects to service rapidly growing restive populations. In some cases, such demands have resulted in unilateral changes to agreements that have increased government revenues and increased investment risks to potential investors. Unilateral changes to agreements to contracts are likely to reduce the rate of new investments to develop oil reserves. Finally, since the world economies have continued to grow robustly over the past 2–3 years in spite of crude oil prices that have exceeded $40 per barrel, and petroleum demand has seen only small changes, there appears to be little motivation for key OPEC suppliers to aggressively expand production, substantially lower crude oil prices, and reduce their revenues.

Most of the trends and the reasons for the trends in the October oil futures case are similar to the trends in the *AEO2006* reference case. Consequently, only the major differences in the trends and their explanations will be highlighted in this section.

None of the projections or scenarios provided in *AEO2005* or *AEO2006* reflect the onset of "peak oil." Historically, crude oil prices have shown significant variations with prices in constant 2004 dollars; for example, crude oil prices exceeded $100 per barrel in the late 1970s and early 1980s. The new high world oil price path in *AEO2006* reflects a new investment and access constrained outlook for crude oil in which the world economies appear able to assimilate the new world oil price regime and crude oil exporters are happy to reap the revenues.

The challenge to oil-rich producers is balance the desire to increase prices and revenues with the potential to cause worldwide economic damage stimulate technological breakthroughs that could undermine their revenues in the longer term. Although specific innovations and technological successes cannot be anticipated or guaranteed, innovation and technological and process advances have historically preceded historical price declines from lofty highs (e.g., horizontal drilling, 3-D and 4-D simulations, the breakthroughs for inexpensive computing, and many others). Breakthroughs that allow relatively inexpensive exploitation of oil shale, for example, could be one such future breakthrough for crude oil production if prices remain too high; breakthroughs in coal-to-liquids technologies, and exploitation of ultraheavy crude oils might be others. For this reason, it is appropriate for investors and planners to consider a plausible range of world oil price scenarios as provided in this chapter or in the *AEO2006* alternative oil price scenarios.

The challenge and potentially serious pitfall for forecasters and investors is to avoid extrapolating short-term trends into long-term projections without supporting analysis that focuses on the changing fundamentals of the industry.

7.10.2 Implications of Higher World Oil Prices

The higher world oil prices in the *AEO2006* reference case have important implications for the projected evolution of energy markets. The most significant impact is in the outlook for petroleum imports. Net imports of petroleum are projected to meet a growing share of total petroleum demand. However, the higher world oil prices in the *AEO2006* reference case lead to greater domestic crude oil production and lower demand, which reduces the need for petroleum imports to 60% in 2025 and 62% of petroleum demand (on the basis of barrels per day) in 2030, up from 58% in 2004. Table 7.2 provides a tabular summary of the *AEO2006* reference case.

The higher world oil prices also impact fuel choice and vehicle efficiency decisions in the transportation sector. Higher oil prices increase the demand for unconventional sources of transportation fuel, such as ethanol and biodiesel, and stimulate coal-to-liquids (CTL) production for the first time. The move to alternative liquids production is highly sensitive to the level of oil prices.

TABLE 7.2 Total Energy Supply and Disposition in the *AEO2006* Reference Case: Summary, 2003–2030

Energy and Economic Factors	2003	2004	2010	2015	2020	2025	2030	2004–2030(%)
Primary energy production (quadrillion Btu)								
Petroleum	14.40	13.93	14.83	14.94	14.41	13.17	12.25	−0.5
Dry natural gas	19.63	19.02	19.13	20.97	22.09	21.80	21.45	0.5
Coal	22.12	22.86	25.78	25.73	27.30	30.61	34.10	1.6
Nuclear power	7.96	8.23	8.44	8.66	9.09	9.09	9.09	0.4
Renewable energy	5.69	5.74	7.08	7.43	8.00	8.61	9.02	1.8
Other	0.72	0.64	2.16	2.85	3.16	3.32	3.44	6.7
Total	70.52	70.42	77.42	80.58	84.05	86.59	89.36	0.9
Net imports (quadrillion Btu)								
Petroleum	24.19	25.88	26.22	28.02	30.39	33.11	36.49	1.3
Natural gas	3.39	3.49	4.45	5.23	5.15	5.50	5.72	1.9
Coal/other (− indicates export)	−0.45	−0.42	−0.58	0.20	0.90	1.54	2.02	NA
Total	27.13	28.95	30.09	33.44	36.44	40.15	44.23	1.6
Consumption (quadrillion Btu)								
Petroleum products	38.96	40.08	43.14	45.69	48.14	50.57	53.58	1.1
Natural gas	23.04	23.07	24.04	26.67	27.70	27.78	27.66	0.7
Coal	22.38	22.53	25.09	25.66	27.65	30.89	34.49	1.7
Nuclear power	7.96	8.23	8.44	8.66	9.09	9.09	9.09	0.4
Renewable energy	5.70	5.74	7.08	7.43	8.00	8.61	9.02	1.8
Other	0.02	0.04	0.07	0.08	0.05	0.05	0.05	0.9
Total	98.05	99.68	107.87	114.18	120.63	126.99	133.88	1.1
Petroleum (million barrels per day)								
Domestic crude production	5.69	5.42	5.88	5.84	5.55	4.99	4.57	−0.7
Other domestic production	3.10	3.21	3.99	4.50	4.90	5.45	5.84	2.3
Net imports	11.25	12.11	12.33	13.23	14.42	15.68	17.24	1.4
Consumption	20.05	20.76	22.17	23.53	24.81	26.05	27.57	1.1
Natural gas (trillion ft.^{3})								
Production	19.11	18.52	18.65	20.44	21.52	21.24	20.90	0.5
Net imports	3.29	3.40	4.35	5.10	5.02	5.37	5.57	1.9
Consumption	22.34	22.41	23.35	25.91	26.92	26.99	26.86	0.7

(continued)

TABLE 7.2 (***Continued***)

Energy and Economic Factors	2003	2004	2010	2015	2020	2025	2030	2004–2030(%)
Coal (million short tons)								
Production	1083	1125	1261	1272	1355	1530	1703	1.6
Net imports	−18	−21	−26	5	36	63	83	NA
Consumption	1095	1104	1233	1276	1390	1592	1784	1.9
Prices (2004 dollars)								
Imported low-sulfur light crude oil (dollars per barrel)	31.72	40.49	47.29	47.79	50.70	54.08	56.97	1.3
Imported crude oil (dollars per barrel)	28.46	35.99	43.99	43.00	44.99	47.99	49.99	1.3
Domestic natural gas at wellhead (dollars per thousand cubic ft.)	5.08	5.49	5.03	4.52	4.90	5.43	5.92	0.3
Domestic coal at minemouth (dollars per short ton)	18.40	20.07	22.23	20.39	20.20	20.63	21.73	0.3
Average electricity price (cents per kilowatthour)	7.6	7.6	7.3	7.1	7.2	7.4	7.5	0.0
Economic indicators								
Real gross domestic product (billion 2000 dollars)	10,321	10,756	13,043	15,082	17,541	20,123	23,112	3.0
GDP chain-type price index (index, 2000 = 1.000)	1.063	1.091	1.235	1.398	1.597	1.818	2.048	2.5
Real disposable personal income (billion 2000 dollars)	7742	8004	9622	11,058	13,057	15,182	17,562	3.1
Value of manufacturing shipments (billion 2000 dollars)	5378	5643	6355	7036	7778	8589	9578	2.1
Energy intensity (thousand Btu per 2000 dollar of GDP)	9.51	9.27	8.28	7.58	6.88	6.32	5.80	−1.8
Carbon dioxide emissions (million metric tons)	5815	5919	6365	6718	7119	7587	8115	1.2

Quantities are derived from historical volumes and assumed thermal conversion factors. Other production includes liquid hydrogen, methanol, supplemental natural gas, and some inputs to refineries. Net imports of petroleum include crude oil, petroleum products, unfinished oils, alcohols, ethers, and blending components. Other net imports include coal coke and electricity. Some refinery inputs appear as petroleum consumption. Other consumption includes net electricity imports, liquid hydrogen, and methanol.

Source: From *AEO2006* National Energy Modeling System, run AEO2006.D111905A.

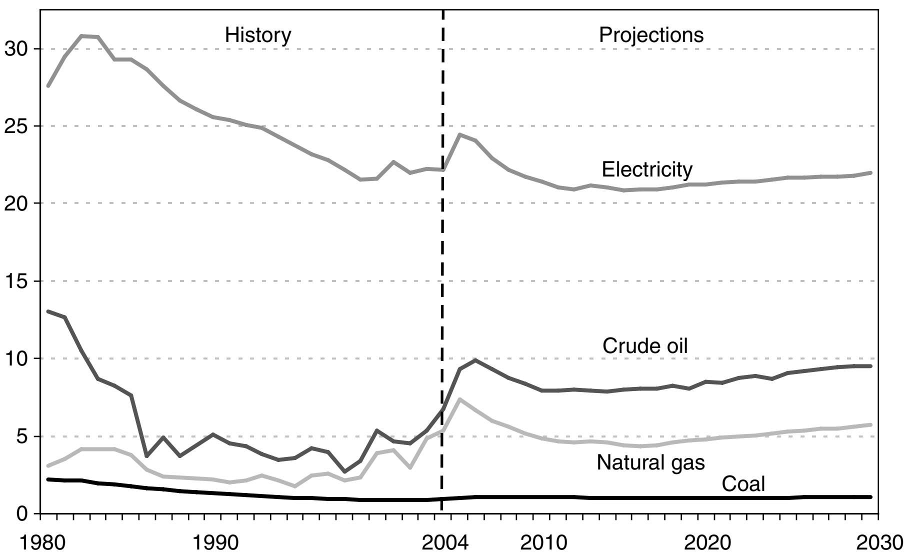


FIGURE 7.9 *AEO2006* reference case, energy prices, 1980–2030 (2004 dollars per million Btu).

7.10.2.1 Economic Growth

Key interest rates, the federal funds rate, the nominal yield on the 10-year treasury note, and the AA utility bond rate are slightly lower compared to the October oil futures case. Also, the projected value of industrial shipments has been revised downward, in part in response to the higher projected energy prices in the *AEO2006* reference case. Despite the higher energy prices in the *AEO2006* reference case relative to the October oil futures case, GDP is projected to grow at an average annual rate of 3.0% from 2004 to 2030 in *AEO2006*, identical to the 3.0% per year growth from 2004 through 2025 projected in the October oil futures case.

7.10.2.2 Energy Prices

In the *AEO2006* reference case, the average world crude oil price increases from $40.49 per barrel (2004 dollars) in 2004 to $59.10 per barrel in 2006 and then declines to $46.90 per barrel in 2014 as new supplies enter the market. It then rises slowly to $54.08 per barrel in 2025 (Figure 7.9).

Average U.S. wellhead natural gas prices are projected to gradually decline from their present level as increased drilling brings on new supplies and new import sources become available, falling to $4.46 per thousand ft.^{3} (2004 dollars) in 2016. After 2016, wellhead prices are projected to increase gradually, to over $5.40 per thousand ft.^{3} in 2025 and over $5.90 per thousand ft.^{3} in 2030. The projected wellhead natural gas prices in the *AEO2006* reference case from 2016 to 2025 are consistently higher than the Octobers oil futures case, ranging roughly from 25 to 60 cents per thousand ft.^{3} higher, primarily due to higher exploration and development costs.

Average delivered electricity prices are projected to decline from 7.6 cents per kWh (2004 dollars) in 2004 to a low of 7.1 cents per kWh in 2018 as a result of falling natural gas prices and, to a lesser extent, coal prices. After 2018, average real electricity prices are projected to increase, reaching 7.4 cents per kWh in 2025 and 7.5 cents per kWh in 2030.

7.10.2.3 Energy Consumption

Total primary energy consumption in the *AEO2006* reference case is projected to increase from 99.7 quadrillion Btu in 2004 to 127.0 quadrillion Btu in 2025 (an average annual increase of 1.2%), 5.4 quadrillion Btu less than in October futures case. In 2025, coal, nuclear, and renewable energy consumption are higher, while petroleum and natural gas consumption are lower in the *AEO2006*

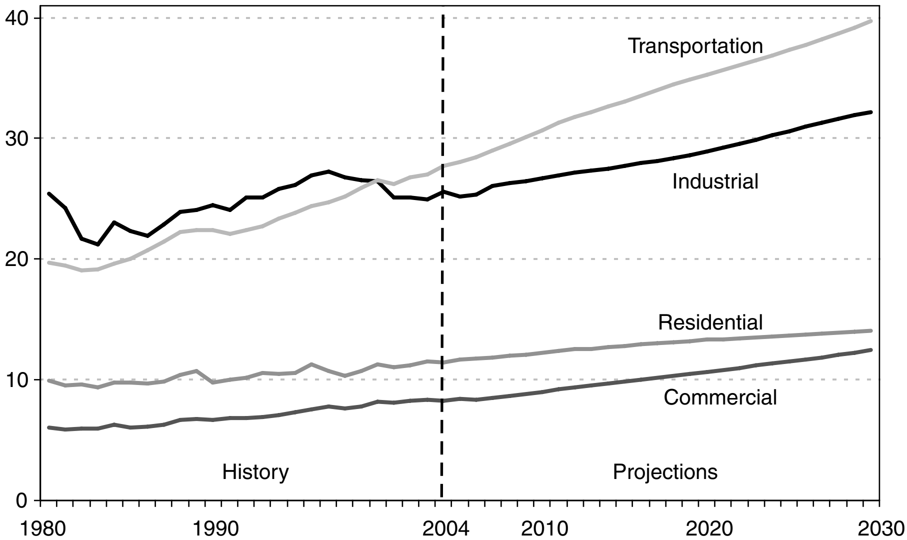


FIGURE 7.10 *AEO2006* reference case, delivered energy consumption by sector, 1980–2030 (quadrillion Btu).

reference case. Among the most important factors accounting for the differences are higher energy prices, particularly petroleum and natural gas prices; lower projected growth rates in the manufacturing portion of the industrial sector, which traditionally includes the most energy-intensive industries; greater penetration by hybrids and diesel vehicles in the transportation sector as consumers focus more on efficiency; and the impact of the recently passed *EPACT2005*, which reduces energy consumption in the residential and commercial sectors and lowers projected growth in electricity demand.

Total petroleum consumption is projected to grow in the *AEO2006* reference case, from 20.8 million barrels per day in 2004 to 26.1 million barrels per day in 2025, 1.2 million barrels per day lower in 2025 than in the October futures case. Petroleum demand growth in the *AEO2006* reference case is lower in all sectors due largely to the impact of the much higher oil prices in *AEO2006*, with almost two-thirds of the decline taking place in the transportation sector.

Total consumption of natural gas in the *AEO2006* reference case is projected to increase from 22.4 tcf in 2004 to 27.0 tcf by 2025 (Figure 7.10), 3.9 tcf lower then projected in the October oil futures case, due mostly to the impact of higher natural gas prices. After peaking at 27.0 tcf, natural gas consumption is projected to fall slightly by 2030 as natural gas loses market share to coal for electricity generation in the later years of the projection due to the higher natural gas prices.

In the *AEO2006* reference case, total coal consumption is projected to increase from 1104 million short tons in 2004 to 1592 million short tons in 2025—84 million short tons more than the 1508 million tons in the October oil futures case. Coal consumption is projected to grow at a faster rate in *AEO2006* toward the end of the projection, particularly after 2020, as coal captures market share from natural gas in power generation due to increasing natural gas prices and as coal use for CTL production grows. Coal was not projected to be used for CTL production in the October oil futures case. In the *AEO2006* reference case, total coal consumption for electricity generation is projected to increase from 1235 million short tons in 2020 to 1502 million short tons by 2030, an average rate of 2.0% per year, and coal use for CTL production increases from 62 million short tons in 2020 to 190 million short tons in 2030.

7.10.2.4 Electricity Generation

In the *AEO2006* reference case, the projected average prices of natural gas and coal delivered to electricity generators in 2025 are higher than the comparable prices in the October oil futures case. While coal consumption in 2025 is similar in both projections, the increase in natural gas prices together with slower

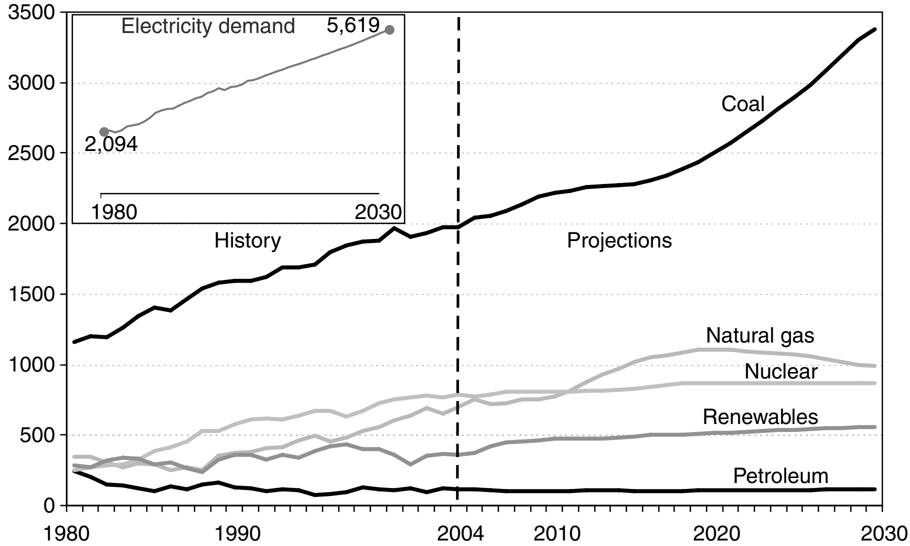


FIGURE 7.11 *AEO2006* reference case, electricity generation by fuel, 1980–2030 (billion kilowatthours).

projected growth in electricity demand and incentives for renewable, nuclear and advanced coal generation technologies leads to significantly lower levels of natural gas consumption for electricity generation. As a result, projected cumulative capacity additions and generation from natural-gas-fired power plants are lower in the *AEO2006* reference case and capacity additions and generation from coal-fired power plants are similar to the October oil futures case through 2025. In fact, in the later years of the projection, natural gas generation is expected to decline as it is displaced with generation from new coal plants (Figure 7.11).

The natural gas share of electricity generation (including generation in the end use sectors) is projected to increase from 18% in 2004 to 22% around 2020 before falling to 17% 2030. The share from coal is projected to decrease from 50% in 2004 to 49% in 2020 and then increase to 57% by 2030. In the *AEO2006* reference case, 87 GW of new coal-fired generating capacity are projected to be constructed between 2004 and 2025 (a comparable amount to the October oil futures case). Over the entire projection (2004 to 2030), 154 GW of new coal-fired generating capacity is projected to be added in the *AEO2006* reference case, including 14 GW at CTL plants.

Nuclear generating capacity in the *AEO2006* reference case is projected to increase from 99.6 GW in 2004 to 108.8 GW by 2020 (10% of total generating capacity) and remain at this level through 2030. The 9 GW increase in nuclear capacity between 2004 and 2030 consists of 3 GW of uprates of existing plants and 6 GW of new plants stimulated by provisions in *EPACT2005*. The nuclear plants that are projected to be added in 2014 and beyond will be the first new plants ordered in the United States in over 30 years. Total nuclear generation is projected to grow from 789 billion kWh in 2004 to 871 billion kWh in 2030 in the *AEO2006* reference case, but nuclear capacity in 2030 is projected to account for only about 15% of total generation.

The use of renewable technologies for electricity generation is projected to grow, stimulated both by higher fossil fuel prices and extended tax credits in the *EPACT2005* and state renewable programs. The expected impacts of state renewable portfolio standards, which specify a minimum share of generation or sales from renewable sources, are included in the projection. The *AEO2006* reference case also includes the extension and expansion of the production tax credit for wind and biomass through December 31, 2007, as enacted in the *EPACT2005*. Total renewable generation in the *AEO2006* reference case, including combined heat and power generation, is projected to grow by 1.7% per year, from 358 billion kWh in 2004 to 559 billion kWh in 2030.

In combination, the Clean Air Interstate Rule (CAIR) and the Clean Air Mercury Rule (CAMR) will lead to large reductions in pollutant emissions from power plants. In the *AEO2006* reference case, sulfur dioxide

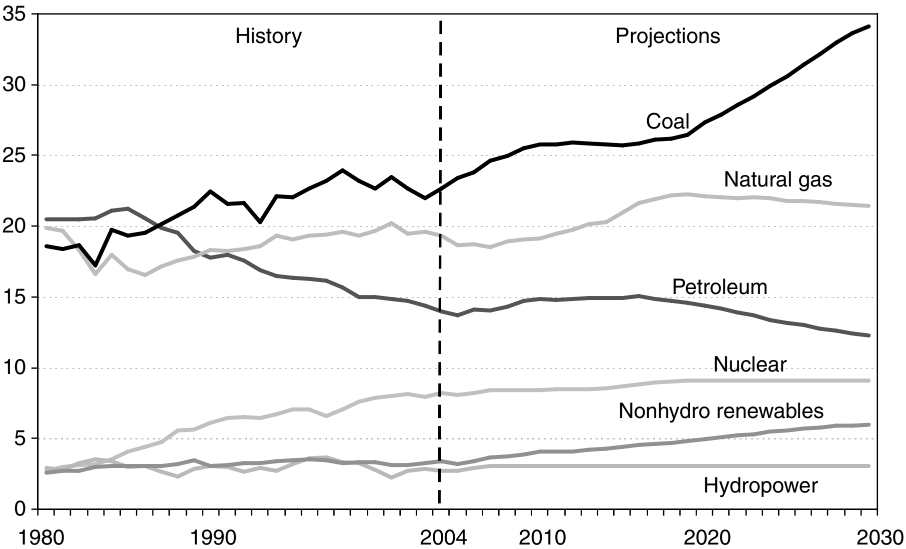


FIGURE 7.12 *AEO2006* reference case, energy production by fuel, 1980–2030 (quadrillion Btu).

emissions are 59% lower, nitrogen oxide emissions are 47% lower, and mercury emissions are 70% lower in 2025 when compared to the October oil futures case because of the inclusion of these regulations.

7.10.2.5 Energy Intensity

Energy intensity, as measured by energy use per 2000 dollar of GDP, is projected to decline at an average annual rate of 1.8% in the *AEO2006* reference case between 2004 and 2030, with efficiency gains and structural shifts in the economy dampening growth in demand for energy services. The rate of decline in energy intensity is faster than the projected 1.6% per year rate in October oil futures case between 2004 and 2025, largely because the higher energy prices in the *AEO2006* reference are projected to result in generally lower levels of energy consumption.

7.10.2.6 Energy Production and Imports

Net imports of energy on a Btu basis are projected to meet a growing share of total energy demand (Figure 7.12). In the *AEO2006* reference case, net imports are expected to constitute 32% and 33% of total U.S. energy consumption in 2025 and 2030, respectively, up from 29% in 2004, much lower then import share projected in the October oil futures case (36.5%) of *AEO2005* for 2025. The higher crude oil and natural gas prices in *AEO2006* lead to greater domestic energy production and lower demand, reducing the projected growth in imports.

In 2025, net petroleum imports, including both crude oil and refined products, are expected to account for 60% of demand (on the basis of barrels per day) in the *AEO2006* reference case compared to 66% in the October oil futures case, up from 58% in 2004. The market share of net petroleum imports grows to 60% in 2025 and 62% of demand in the *AEO2006* reference case by 2030.

Total domestic natural gas production increases from 18.5 tcf in 2004 to 21.2 tcf in 2025, before declining to 20.8 tcf in 2030 in the *AEO2006* reference case. Growth in LNG imports is projected to meet much of the increased demand for natural gas in the *AEO2006* reference case, but the increase is less than projected in the October oil futures case. The growth of LNG imports in the *AEO2006* reference case is moderated by higher domestic and imported LNG gas prices that reduce domestic natural gas demand. Because of the higher petroleum product prices, the international natural gas demand is expected to rise and raise LNG prices with it. Higher international LNG prices are projected to limit the demand for LNG demanded in the United States.

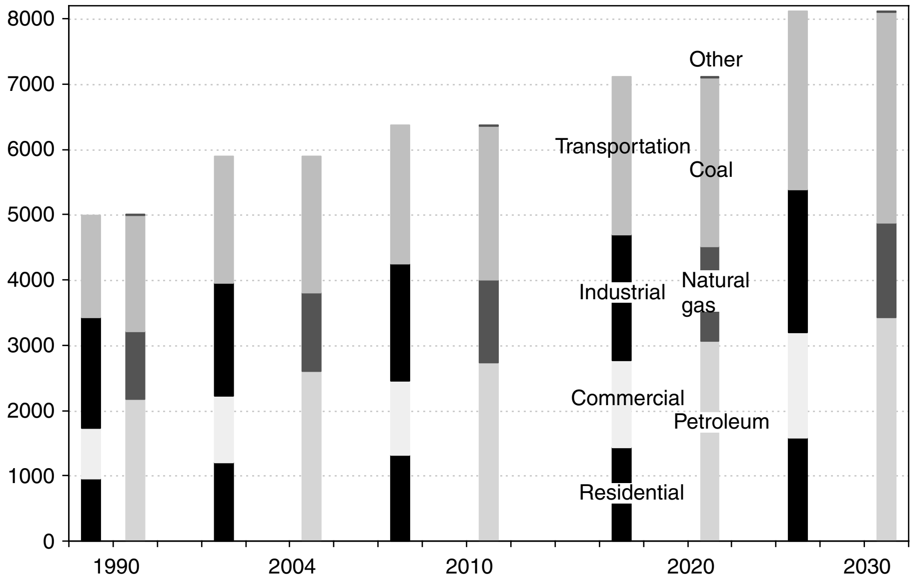


FIGURE 7.13 *AEO2006* reference case, U.S. carbon dioxide emissions by sector and fuel, 1990–2030 (million metric tons).

7.10.2.7 Carbon Dioxide Emissions

Carbon dioxide emissions from energy use are projected to increase from 5919 million metric tons in 2004 to 7587 million metric tons in 2025 and 8115 million metric tons in 2030 in the *AEO2006* reference case, an average annual increase of 1.2% per year (Figure 7.13). The carbon dioxide emissions intensity of the U.S. economy is projected to fall from 550 metric tons per million dollars of GDP in 2004 to 377 metric tons per million dollars of GDP in 2025—an average decline of 1.8% per year—and 351 metric tons per million dollars of GDP in 2030.

7.11 Overview Impacts of the AEO2006 High-Price Case*

Although short-term weather-related events (e.g., hurricanes) and geopolitical instabilities (e.g., terrorist actions and attempts by some producers to withhold supply) could raise crude oil prices well above $100 per barrel in 2004 dollars for a few months or even years, the author believes that sustained crude oil prices which exceed $75 per barrel for the next 25 years are unlikely for technical reasons.† Nevertheless, it is prudent to consider such a case for contingency planning purposes. This section evaluates the impacts

*This case is provided at the request of the editor. Energy markets in the past 3 years have demonstrated the volatility of energy prices to slight supply–demand imbalances. For example, the Henry Hub spot prices have fallen from a high of over $14 per million Btu in December 2005 to a low of about $4 per million Btu in September 2006.

†Numerous technology and supply options for producing petroleum product substitutes are currently economically viable at prices well under $75 per barrel including: oil sands from Canada, coal-to-liquids which the Germans and South Africans used heavily (Fischer–Tropsch technology), and oil shale which was economic at about $75 per barrel using the 1980's underground mining and surface retorting process. New processes using a true in situ process are targeted to be economic at under $45 per barrel. Stranded natural gas, that is, natural gas without a currently accessible market like gas in portions of Alaska and Nigeria, could be used to economically produce high-quality petroleum products at well under $45 per barrel if the stranded natural gas price is less than $1 per thousand ft.^{3}. Biofuels for ethanol and bio-diesel are also likely to grow if high world oil prices are sustained. Finally, ultra-heavy oil, a product currently shunned by most refineries, could be economic at prices well under $75 per barrel. Given sustained high crude oil prices from conventional sources, investments in unconventional liquids production are likely to rise rapidly and eventually place downward pressure on high crude oil prices.

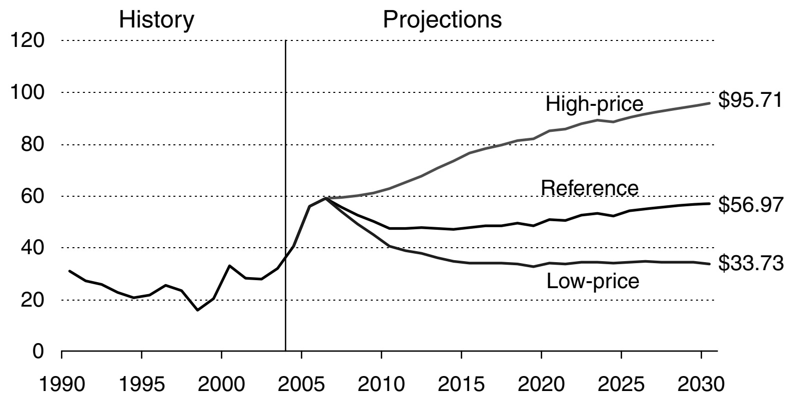


FIGURE 7.14 World oil prices (Foreign low-sulfur light crude oil prices) in three AEO2006 cases, 1990–2030 (2004 dollars per barrel). (From Energy Information Administration. *Annual outlook 2006, with projections to 2030*. DOE/EIA-0383 2006, Energy Information Administration, Washington, D.C., 2006, http://www.eia.doc.gov/oiaf/aeo/index.html)

of high oil and natural gas prices, as defined by the *Annual Energy Outlook 2006* (*AEO2006)* high-price case, on the transportation and power generation sectors of the U.S. energy economy. The *AEO2006* low-price and reference price cases are also briefly described to provide balance and context.

The *AEO2006* high- and low-price cases reflect, respectively, lower or higher domestic and international unproven/undiscovered oil and natural gas resources than the reference case. The high-price path reaches $96 per barrel in 2030 for foreign low-sulfur light crude oil (FLL) based on the assumptions that unproven/undiscovered international and domestic oil resources are 15% lower than the reference case and that the OPEC cartel adjusts its output to maintain the higher prices. In this case, the OPEC share of world oil production is projected to decline to 31% compared to over 38% in 2005. The low-price case projects that crude oil prices (FLL) will fall to about $34 per barrel in constant 2004 dollars because undiscovered oil and natural gas resources are assumed to be 15% higher than the reference case and that OPEC chooses to approximately retain its 2005 market share of about 38% through 2030* (see Figure 7.14).

All other things being equal, higher oil prices will stimulate more exploration and production, particularly in non-OPEC countries, and increase crude oil and alternative liquids production from oil sands, coal-to-liquids (CTL), gas-to-liquids (GTL), shale oil, ethanol, and other biofuel liquids. In addition, higher prices will reduce consumption through price-induced conservation (e.g., lower vehicle miles traveled or lower thermostat settings for space heating) and through increased efficiency uptake (e.g., more efficient cars). Table 7.3 summarizes the major outcomes of the AEO2006 price cases.

7.11.1 Domestic Oil Production

U.S. oil production is marginally sensitive to world oil prices. Higher (or lower) oil prices induce more (or less) exploration activity and the development of more (or less) oil resources. In all cases, a significant portion of total domestic oil production is projected to be produced from large, existing oil fields, such as the Prudhoe Bay Field.

*OPEC's projected market share in 2030 is around 40 for the low and reference price cases. In nominal or "current-year" dollars terms, prices in 2030 will be about 70 higher than the quoted constant 2004 dollar values shown in this section. For example, a $96 dollar per barrel crude oil price (light, low sulfur) in 2004 dollars in 2030 would be over $160 per barrel in year 2030 dollars.

TABLE 7.3 AEO2006 Price Case Comparisons

		2010			2015			2020			2025			2030		
	2004	Low	Ref	High	Low	Ref	High	Low	Ref	High	Low	Ref	High	Low	Ref	High
Domestic Production																
Crude oil and lease condensate	11.47	12.71	12.45	12.24	12.69	12.37	12.20	11.77	11.75	11.91	10.56	10.56	11.22	9.51	9.68	10.50
Natural gas plant liquids	2.46	2.43	2.39	2.35	2.59	2.57	2.53	2.61	2.67	2.65	2.61	2.62	2.66	2.65	2.57	2.62
Dry natural gas	19.02	19.58	19.13	18.67	21.21	20.97	20.59	21.66	22.09	21.90	21.64	21.80	22.15	22.09	21.45	21.83
Coal	22.86	25.38	25.78	25.91	24.76	25.73	26.47	24.81	27.30	29.53	26.02	30.61	34.08	27.86	34.10	39.52
Nuclear power	8.23	8.44	8.44	8.44	8.60	8.66	8.77	9.03	9.09	9.09	9.03	9.09	9.09	9.03	9.09	9.09
Renewable energy	5.73	7.00	7.08	7.25	7.12	7.43	7.71	7.64	8.00	8.15	8.22	8.61	8.79	8.73	9.02	9.12
Other	0.64	2.13	2.16	2.20	2.90	2.85	3.01	3.04	3.16	3.37	3.10	3.32	3.74	3.15	3.44	3.91
Total	70.42	77.67	77.42	77.05	79.87	80.58	81.27	80.57	84.05	86.61	81.18	86.59	91.73	83.00	89.36	96.57
Imports																
Crude oil	22.02	22.25	22.01	21.23	24.20	22.91	21.13	27.19	24.63	22.09	30.80	26.96	23.05	33.90	29.54	24.59
Petroleum products	5.93	6.55	6.36	5.90	7.91	7.29	6.22	9.08	8.01	6.31	10.17	8.41	6.25	11.68	9.27	5.88
Natural gas	4.36	5.19	5.01	4.74	6.89	5.81	4.21	8.32	5.83	3.33	9.78	6.37	3.37	10.75	6.72	3.63
Other imports	0.83	0.45	0.45	0.46	0.73	0.74	0.98	1.49	1.36	1.61	1.97	2.02	2.17	2.09	2.42	2.69
Total	33.14	34.43	33.83	32.33	39.73	36.75	32.53	46.08	39.83	33.35	52.73	43.76	34.84	58.43	47.95	36.79
Exports																
Petroleum	2.07	2.17	2.15	2.11	2.36	2.18	2.12	2.68	2.24	2.16	2.94	2.26	2.20	2.86	2.31	2.24
Natural gas	0.86	0.57	0.55	0.53	0.66	0.58	0.48	0.83	0.68	0.50	1.10	0.86	0.58	1.35	1.01	0.57
Coal	1.25	1.03	1.03	1.03	0.54	0.54	0.54	0.46	0.46	0.46	0.48	0.48	0.45	0.39	0.40	0.40
Total	4.18	3.78	3.74	3.67	3.56	3.30	3.14	3.97	3.39	3.11	4.53	3.61	3.22	4.61	3.72	3.21
Discrepancy	−0.31	−0.32	−0.36	−0.49	−0.07	−0.16	−0.39	−0.10	−0.15	−0.24	−0.13	−0.25	−0.04	0.00	−0.30	0.07
Consumption																
Petroleum products	40.08	43.78	43.14	41.88	47.55	45.69	43.11	50.67	48.14	44.72	53.96	50.57	46.62	57.55	53.58	48.87
Natural gas	23.07	24.65	24.04	23.34	27.91	26.67	24.78	29.62	27.70	25.05	30.80	27.78	24.75	31.97	27.66	24.71
Coal	22.53	24.68	25.09	25.22	24.84	25.66	26.59	25.77	27.65	30.01	27.44	30.89	34.08	29.49	34.49	38.25
Nuclear power	8.23	8.44	8.44	8.44	8.60	8.66	8.77	9.03	9.09	9.09	9.03	9.09	9.09	9.03	9.09	9.09
Renewable energy	5.73	7.00	7.08	7.25	7.12	7.43	7.71	7.64	8.00	8.16	8.22	8.61	8.80	8.73	9.02	9.12
Other	0.04	0.07	0.07	0.08	0.08	0.08	0.09	0.04	0.05	0.06	0.05	0.05	0.05	0.05	0.05	0.05
Total	99.68	108.64	107.87	106.21	116.11	114.18	111.05	122.78	120.63	117.09	129.50	126.99	123.38	136.82	133.88	130.09

(continued)

TABLE 7.3 (***Continued***)

		2010			2015			2020			2025			2030		
	2004	Low	Ref	High	Low	Ref	High	Low	Ref	High	Low	Ref	High	Low	Ref	High
Net Imports—Petroleum	25.88	26.62	26.22	25.02	29.74	28.02	25.22	33.59	30.39	26.25	38.03	33.11	27.10	42.72	36.49	28.23
Prices (2004 dollars per unit)																
World oil price ($ per bbl)[a]	35.99	37.00	43.99	58.99	28.99	43.00	71.98	27.99	44.99	79.98	27.99	47.99	84.98	27.99	49.99	89.98
Gas wellhead price ($/mcf)	5.49	4.55	5.03	5.95	3.62	4.52	5.65	4.09	4.90	5.94	4.42	5.43	6.55	4.97	5.92	7.71
Coal minemouth price ($/ton)	20.07	21.74	22.23	22.53	19.78	20.39	21.13	19.19	20.20	21.54	19.50	20.63	21.90	20.66	21.73	22.66
Electricity (cents/kWh)	7.6	7.1	7.3	7.6	6.8	7.1	7.6	7.0	7.2	7.6	7.1	7.4	7.5	7.3	7.5	7.9
Transportation Petroleum cons																
Distillate fuel	5.91	6.90	6.82	6.71	7.61	7.48	7.25	8.25	8.13	7.90	9.08	8.95	8.78	10.10	9.98	9.93
Jet fuel	3.35	3.93	3.89	3.83	4.32	4.27	4.21	4.58	4.53	4.44	4.66	4.61	4.38	4.82	4.79	4.33
Motor gasoline[b]	16.93	18.62	18.33	17.72	20.37	19.54	18.03	21.98	20.73	18.60	23.63	21.81	19.28	25.33	22.99	20.00
Residual fuel	0.61	0.62	0.62	0.62	0.63	0.63	0.63	0.64	0.64	0.64	0.65	0.65	0.65	0.66	0.65	0.65
Liquefied petroleum gas	0.03	0.06	0.06	0.06	0.08	0.07	0.07	0.09	0.09	0.07	0.10	0.10	0.08	0.12	0.11	0.09
Other Petroleum	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.19	0.19	0.19	0.19	0.19	0.19
Petroleum subtotal	27.02	30.30	29.91	29.11	33.19	32.18	30.37	35.73	34.30	31.83	38.31	36.30	33.35	41.21	38.71	35.18
New car mpg	29.3	31.3	31.4	31.8	31.9	32.2	33.3	32.1	32.7	34.6	32.4	33.5	35.7	32.6	33.8	36.1
New light truck mpg	21.5	23.1	23.2	23.5	23.7	24.0	24.9	24.3	24.9	26.5	24.8	25.8	27.6	25.2	26.4	28.2

All energy is in quadrillion Btu. All prices are in 2004 dollars.
[a] IRAC price. Add between $4 and $7 per barrel for Imported light, low-sulfur foreign crude oil price.
[b] Ethanol as gasohol is in gasoline.

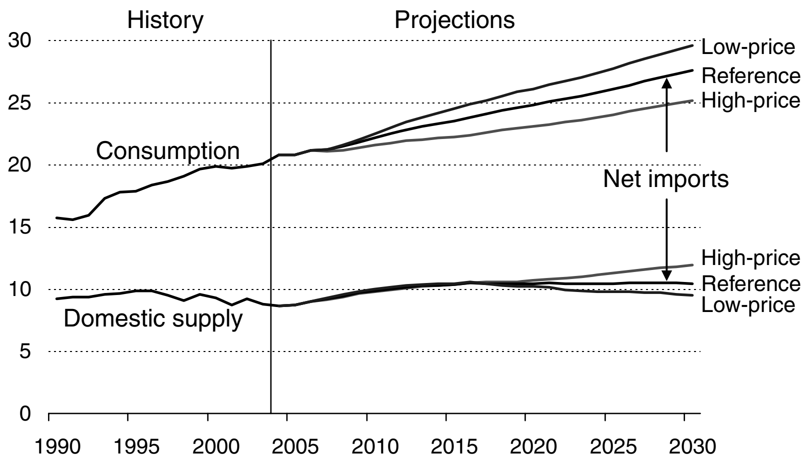


FIGURE 7.15 Petroleum supply, consumption, and imports in AEO2006 1990–2030 (million barrels per day).

In the high-price case, FLL crude oil prices in 2030 are 68% higher than the reference case and yield only 9% higher domestic production because of the smaller resource base. Domestic liquids consumption is 9% lower because of the higher prices.

7.11.2 Impacts of High World Oil Prices on Oil Imports

Oil import dependence is sensitive to world oil prices. In the low-price case, imported oil is projected to grow from a 58% share of total oil supply in 2004 to 68% in 2030; the reference and high-price cases project imported oil shares to be 62% and 53%, respectively. In the reference case, net oil imports are expected to rise from a little over 12 million barrels per day in 2004 to over 17 million barrels per day in 2030. In the low-price case, oil imports in 2030 are expected to rise above 20 million barrels per day. The high-price case is expected to moderate demand so that oil imports are expected to remain near 2005's level, about 13 million barrels per day (see Figure 7.15).

As in the United States, projected oil prices directly affect world oil consumption: higher projected oil prices result in lower projected world oil consumption. In the reference case, world oil consumption is projected to rise from about 82.5 million barrels per day in 2004 to about 118 million barrels per day in 2030, whereas world oil consumption is expected to be 128 and 102 million barrels per day in the low- and high-price cases, respectively.

7.11.3 Impacts of High World Oil Prices on Domestic Unconventional Liquids Supply

High world oil prices provide economic incentives for the development of unconventional liquids technologies. Gas-to-liquids (GTL) plants are projected to be built on the Alaskan North Slope in the high-price case to convert stranded gas to zero-sulfur distillates to be transported via the Trans-Alaska pipeline to Valdez and shipped to the lower-48 states for use on the west coast. GTL production is expected to be economic and reach 200,000 barrels per day in 2030 only in the high-priced case because access to inexpensive Alaskan natural gas supplies are expected to be committed to the Alaska natural gas pipeline that is expected to be completed by 2015.

Coal-to-liquids (CTL) plants are projected to be built in the lower-48 states for the reference and high-price cases. Full-scale CTL production is projected to start in 2011 in both the reference and high-price cases, reaching just under 800,000 barrels per day in 2030 in the reference case and about 1.7 million

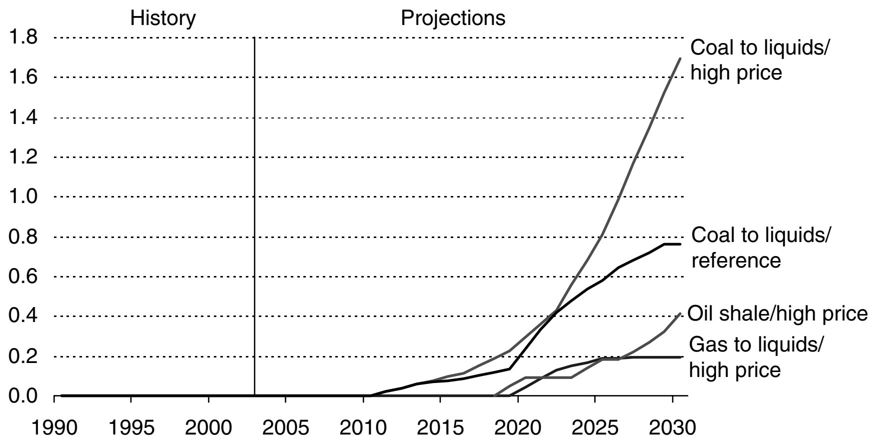


FIGURE 7.16 Gas-to-liquids, coal-to-liquids, and oil shale production in the price cases, 1990–2030 (million barrels per day).

barrels per day in the high-price case (see Figure 7.16). Major environmental and water uncertainties may affect the potential growth of CTL plant investments.

Production costs for oil shale syncrude are even more uncertain than CTL production costs.* Development of this domestic resource came to a halt in the mid-1980s, during a period of low oil prices. The cost assumptions used in developing the *AEO2006* projections represent an oil shale industry based on underground mining and surface retorting. However, the development of a true in situ retorting technology could substantially reduce the cost of producing oil shale syncrude and possibly make its production economic in the reference case. Oil shale production is not economic in the low-price and reference cases of *AEO2006* but is expected to become economic around 2020 in the high-price case; oil shale production is projected to steadily rise to over 400,000 barrels per day by 2030 (see Figure 7.16).

7.11.4 Impacts of High World Oil Prices on Ethanol Production

The Energy Policy Act of 2005 (EPAct 2005) requires that ethanol be used for gasoline-based transportation and rise to at least 7.5 billion gallons by 2012. Consumption is required to rise proportionately with gasoline consumption thereafter. World oil prices are high enough in all of the *AEO2006* price cases to exceed the minimum levels required by the law. Ethanol consumption in 2012 is projected to be between 9.6 and 9.9 billion gallons in the three price cases and rise to consumption levels between 11 and 15 billion gallons in 2030 (see Figure 7.17). As expected, the highest gasoline prices encourage the highest ethanol consumption. Almost all of the ethanol is projected to be produced from corn, with only small amounts of cellulosic ethanol expected to penetrate the market (250 million gallons per year) because of EPAct 2005 incentives. Capital costs reductions and breakthroughs in the cost of manufacturing the enzyme needed for cellulosic ethanol conversion could alter the competitive outlook for cellulosic ethanol.

*The United States has more than half of the world's undiscovered recoverable oil shale resources. The U.S. government has not funded any significant new research in oil shale since the early 1980s when the oil shale technologies were estimated to become economic at prices in the $70 to $80 per barrel range. New leases on prime oil shale lands and current pilot project plans and lease bids on oil shale using new approaches by the Shell Oil Corporation, among others, will provide better information and will hopefully identify lower-cost applications of technologies and processes for producing syncrude from oil shale.

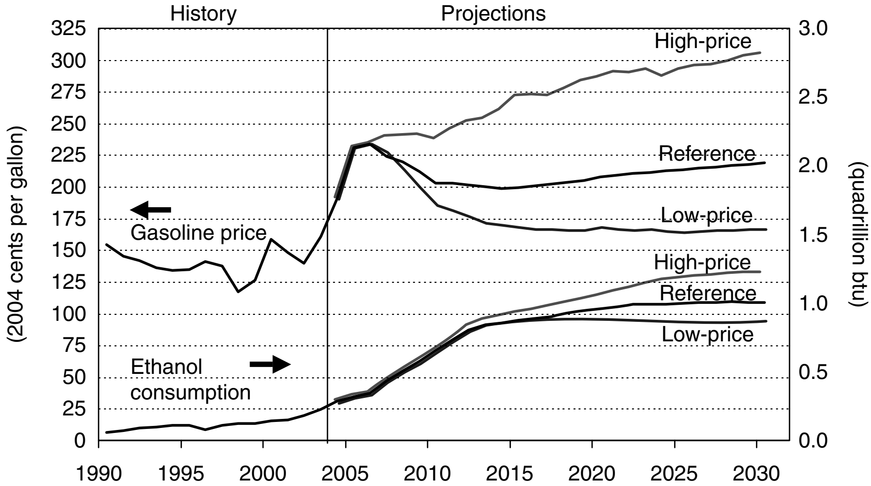


FIGURE 7.17 *AEO2006* ethanol consumption and gasoline price in transportation sector, 1990–2030.

7.11.5 LNG Imports are the Source of Natural Gas Supply Most Affected in the Price Cases

Higher world oil prices are expected to result in a shift away from petroleum consumption and toward natural gas and coal consumption in all sectors of the international energy market. LNG prices are expected to roughly follow the pattern of world crude oil prices because many LNG contract prices are tied directly to crude oil prices and higher oil prices are expected to promote increased GTL production. Both of these factors are expected to put upward price pressure on world natural gas supplies. Because of the higher LNG prices in the high-price case, it is expected that U.S. LNG imports, new LNG receiving capacity, and the utilization rates for LNG terminals will be lower than the reference case.

Net imports of LNG in the reference case are 4.4 trillion $ft.^3$ in 2030. In the low-price case, net LNG imports increase to 7.4 trillion $ft.^3$, and in the high-price case they fall to 1.9 trillion $ft.^3$ (see Figure 7.18).

7.11.6 Natural Gas Wellhead Prices

In all three *AEO2006* price cases, projected wellhead natural gas prices are projected to decline from current levels, reach a low point around 2015, and subsequently rise (see Figure 7.19).

The relatively high projected natural gas prices, particularly in the reference case, both expand the development of new gas supplies, particularly of LNG and Alaska gas, while also constraining future gas consumption growth. Collectively, the interaction of greater gas supply and slower gas demand growth leads to the projected decline in gas prices through roughly 2015 in all three cases. Natural gas from Alaska is projected to be economic in all *AEO2006* price cases by about 2015.

After 2015, natural gas prices are projected to rise again as the increasing marginal costs of developing the remaining U.S. natural gas resources increase. As domestic natural gas resources are produced, the remaining gas resources are more costly and riskier to develop and produce.

In the reference case, the projected wellhead natural gas price is projected to increase from about \$4.45 per thousand $ft.^3$ in 2016 to about \$5.90 per thousand $ft.^3$ (2004 dollars) in 2030. Because the low- and high-price cases respectively increased or reduced domestic unproven gas resources by 15%, natural gas prices in 2030 are projected to be about \$5 per thousand $ft.^3$ in the low-price case and \$7.70 per thousand $ft.^3$ in the high-price case.

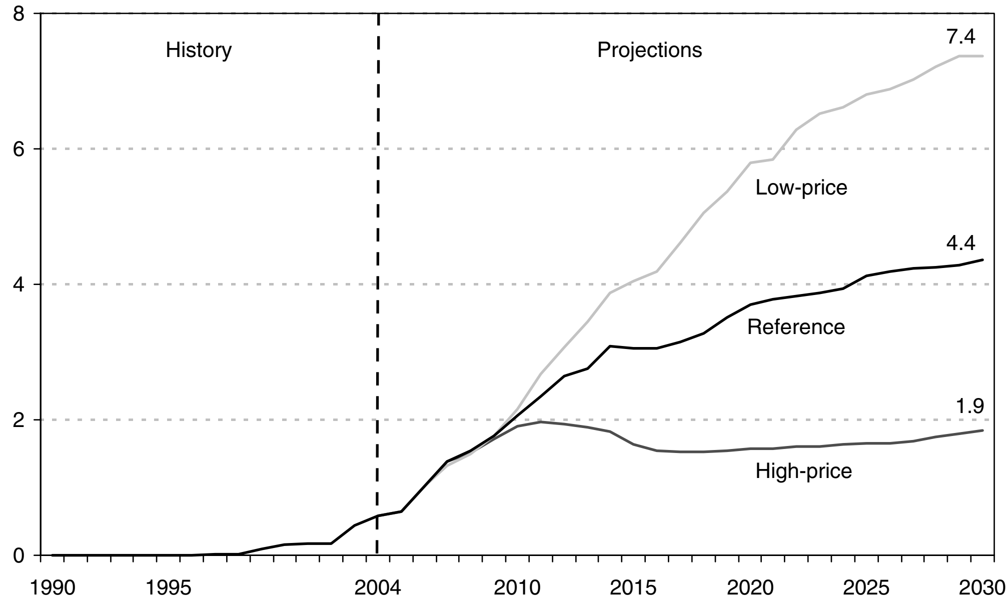


FIGURE 7.18 Net imports of liquefied natural gas in three cases, 1990–2030 (trillion ft.^{3}).

7.11.7 Petroleum Demand Is Significantly Lower in the High-Price Case

Petroleum demand in 2030 in the high-price case is projected to be about 2.4 million barrels per day lower than the reference case. Petroleum imports in 2030 are expected to be roughly comparable to 2005 import levels. Most of the petroleum reduction in the high-price case occurs in gasoline consumption by light-duty vehicles (cars and light trucks). Gasoline consumption in the high-price case is projected to be about 1.5 million barrels per day lower than the reference case in 2030 (see Figure 7.20).

Motor gasoline consumption is driven by vehicle miles traveled and efficiency uptake. High oil prices reduce personal disposable income, which lowers vehicle miles traveled and increases the adoption of more efficient transportation technologies.

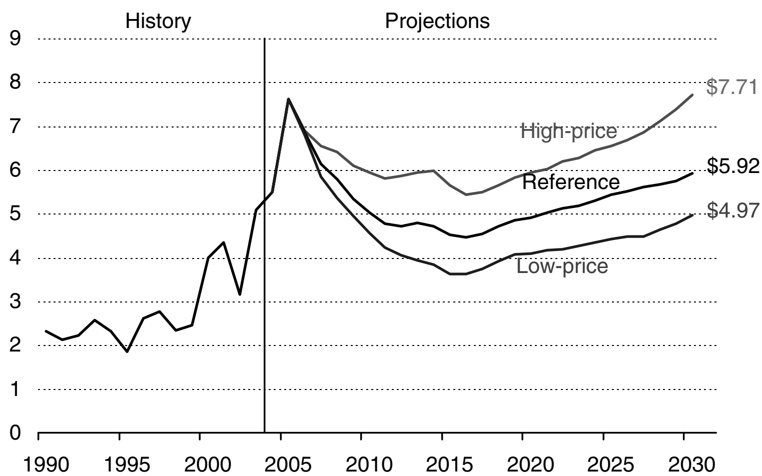


FIGURE 7.19 Natural gas wellhead prices, 1990–2030 (2004 dollars per thousand ft.^{3}).

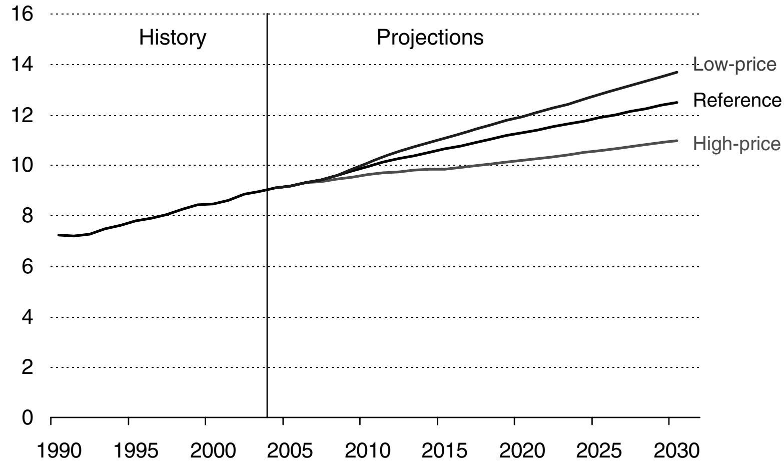


FIGURE 7.20 Motor gasoline consumption, 1990–2030 (million barrels per day).

- Gasoline consumption is expected to rise from 9.1 million barrels per day in 2004 to 13.7, 12.5, and 11.0 million barrels per day in the low-price, reference, and high-price cases, respectively, in 2030.
- Vehicle miles traveled by light-duty vehicles is projected to rise from 2.6 trillion miles in 2005 to 4.4, 4.1, and 3.9 trillion miles in the low-price, reference, and high-price cases, respectively, in 2030.
- New car vehicle efficiency is projected to rise from 29.3 in 2004 to 32.6, 33.8, and 36.1 miles per gallon in the reference case in 2030 (see Figure 7.21).
- Three of the most promising transportation technologies are advanced drag reduction, variable valve timing, and extension of four valves per cylinder technology to six valves per cylinder engines. Each of these would provide more than an 8% boost to fuel economy.

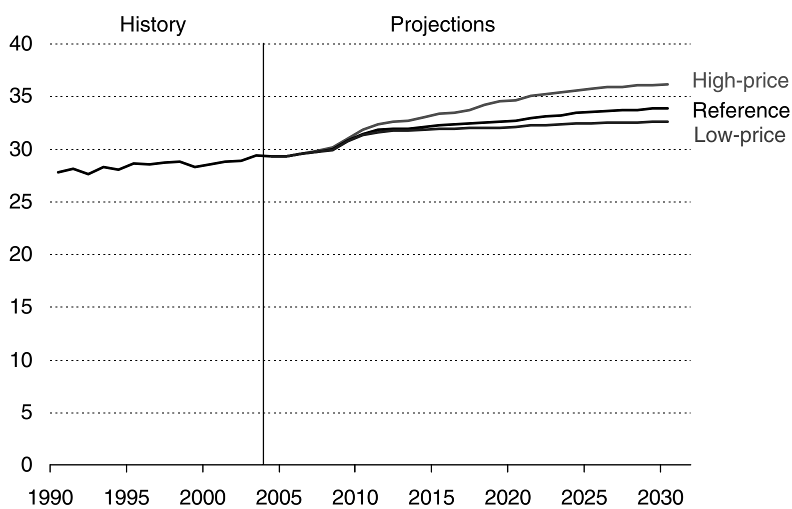


FIGURE 7.21 New car miles per gallon, 1990–2030.

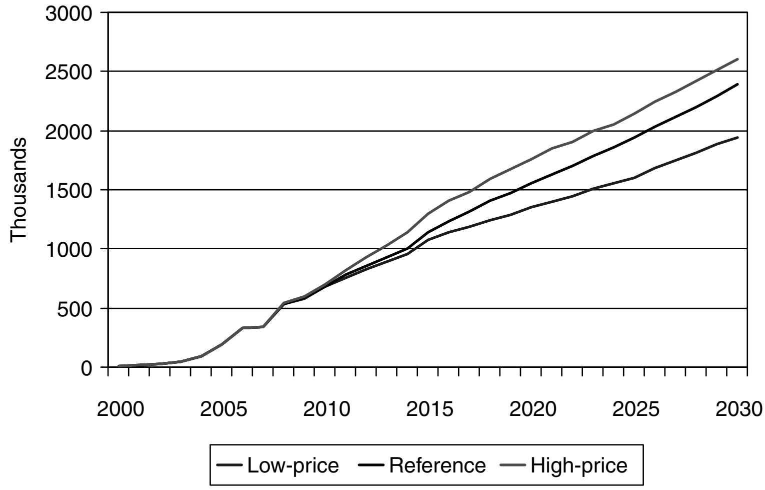


FIGURE 7.22 Hybrid vehicle annual fuel savings.

7.11.8 EPAct2005 Accelerates the Early Adoption of Hybrid Vehicles

EPAct2005 provides financial incentives for the early adoption of hybrid vehicles prior to 2010. Those incentives are projected to accelerate the adoption of hybrid fuel vehicles and help reduce the incremental vehicle costs for midsize full hybrid vehicles from about $2500 in 2005 to about $1500 by 2015. However, because the incentives are limited to the first 60,000 sales by each manufacturer and based on the relative efficiency improvements, the relative impacts are also limited. After 2010, hybrid adoption loses its EPAact 2005 incentives. In the high-price case, fuel savings for a typical midsize car are approximately $300 per year while the savings are reduced to about $225 per year in the reference case and to under $200 per year in the low-price case (see Figure 7.22).

The projected energy savings between the price cases are not sufficiently large, however, to drastically change the hybrid vehicle adoptions by 2030. Hybrid vehicle sales are projected to increase from about 92

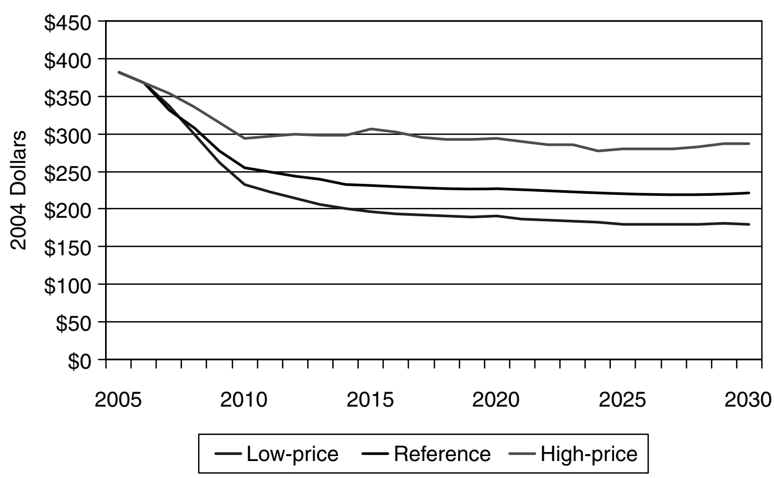


FIGURE 7.23 Hybrid vehicle sales.

thousand in 2004 to about 2.4 million in 2030 in the reference case, 2.6 million in the high-price case, and 1.9 million in the low-price case. With faster hybrid penetration, vehicle costs decline due to manufacturer learning, which leads to greater adoption (see Figure 7.23). Hybrids like the Prius can be viewed as full hybrids and the less aggressive types with the integrated starter generator (ISG), no power assist, like those made by GM are called *mild hybrids*. Full hybrids are expected to represent about 80% of the hybrid sales by 2030.

7.11.9 High Oil and Gas Prices Increase Coal-Based Generation

The U.S. power generation industry is projected to increasingly rely on coal-based generation as natural gas prices increase. Of the four major gas consumption sectors (residential, commercial, industrial, and electric power), the electric power sector's consumption of natural gas is the most affected by natural gas prices over time.

Gas-fired generation facilities can operate in the base, intermediate, and peaking portions of the electricity load duration curve. As gas prices become increasingly more expensive, the operating costs become a more important factor in making the capacity addition and dispatch decisions. With higher natural gas prices, an increasing proportion of new electricity generators are projected to be coal-fired facilities, which are placed in preference to gas-fired facilities in the load duration curve, thereby displacing gas-fired generators from operating in the base and later in the intermediate-load portions of the load duration curve, limiting their operation mostly to the peaking portion of the curve.

In the reference case, electricity sector gas consumption is projected to grow from 5.3 trillion ft.^{3} in 2004 to 6.4 trillion ft.^{3} in 2030. In comparison, electric power sector is projected to consume 9.9 and 4.1 trillion ft.^{3} in the low- and high-price cases, respectively, in 2030.

Figure 7.24 illustrates that steam coal facility construction increases at higher natural gas prices, thereby reducing the level of gas-fired capacity by 2030 relative to the *AEO2006* reference case. In the reference case, combined cycle gas-fired facilities increase from 159 GW in 2004 to 231 GW in 2030. In the high-price case, gas-fired combined cycle capacity only grows to 191 GW in 2030, while the low-price case projects 281 GW of gas-fired combined cycle capacity.

Most of the reduction in gas-fired electricity generator construction in the high gas price case is compensated by an increase in coal-fired electricity generation facilities. In the reference case, coal-fired steam capacity grows from 310 GW in 2004 to 457 GW in 2030. In the low- and high-price cases, coal-fired steam capacity is projected to be 380 and 509 GW, respectively.

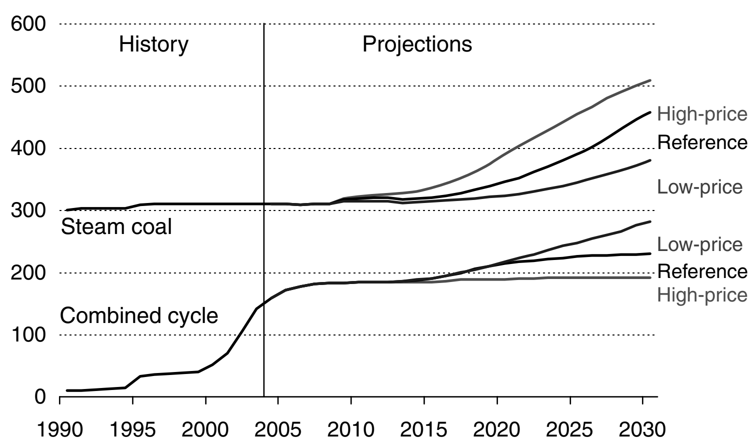


FIGURE 7.24 Generation capacity from steam coal and combined cycle, 1990-2030 (GW).

Acknowledgments

I am indebted to the forecasting staff of the Office of Integrated Analysis and Forecasting in the Energy Information Administration (EIA) for their integral part in helping to develop this projection. Credits for this effort are shared with the staff and management who produced the *Annual Energy Outlook 2005* (*AEO2005*) and the updated reference case forecast in the *Annual Outlook 2006* (*AEO2006*) which is summarized in the last section of this chapter. The *AEO2006* reference case was released to the public on December 12, 2005 on the EIA website and the complete publication in February, 2006. Like all projections, *these projections represent what might occur under laws, regulations, policies, and other key assumptions of the scenarios at the time of their development, not what will occur*. All errors remain the sole responsibility of this author.

References

1. Energy Information Administration. 2003. *National Energy Modeling System, An overview 2003*. DOE/EIA-058. Energy Information Administration, Washington, DC. http://www.eia.doe.gov/oiaf/aeo/overview/index.html (accessed on March 4, 2003).
2. Energy Information Administration. 2005. *Assumptions to the Annual Outlook 2005, with Projections to 2025*. DOE/EIA-0383 (2005). Energy Information Administration, Washington, DC, http://www.eia.doe.gov/oiaf/aeo/assumption/pdf/0554(2005).pdf.
3. Energy Information Administration. 2005. *Annual Outlook 2005, with Projections to 2025*. DOE/EIA-0383(2005). Energy Information Administration, Washington, D.C.
4. Energy Information Administration. *Annual Outlook 2006, with Projections to 2030*. 2006. DOE/EIA-0383(2006). Energy Information Administration, Washington, D.C. http://www.eia.doe.gov/oiaf/aeo/index.html (accessed on February 2006).

8

Transportation Systems

Beth Isler
Resource Systems Group, Inc.

8.1 Introduction

This chapter presents trends in land use, freight, ground-transportation modes for people and freight, transportation fuel supply, and the opportunities for conservation that exist within each area. The chapter starts with a discussion of the transportation–land use relationship for a better understanding of the framework within which the transportation system functions and the design theories that aim to influence mode choice and trip generation. Next is a description of mass transit, with particular emphasis on how its energy use compares to the energy use of the automobile. The movement of freight, its modes, and energy consumption relative to the rest of the transportation system follows. Then, emerging future technologies are described; the focus of this section is on vehicle efficiencies to conserve energy resources. Finally, the well-to-wheel energy analysis combining fuel production and vehicle performance is presented, focusing on what feedstocks are available and how they can be refined efficiently into a fuel.

8.2 Land Use

8.2.1 Land Use and Its Relationship to Transportation

There is a fundamental relationship between transportation and land use, because the distance between one's origin and destination will determine the feasibility, route, mode, cost, and time necessary to travel from one place to another. Likewise, transportation influences land use as it impacts people's decisions about where to live and work, considering factors such as commute time and cost, the distance to a

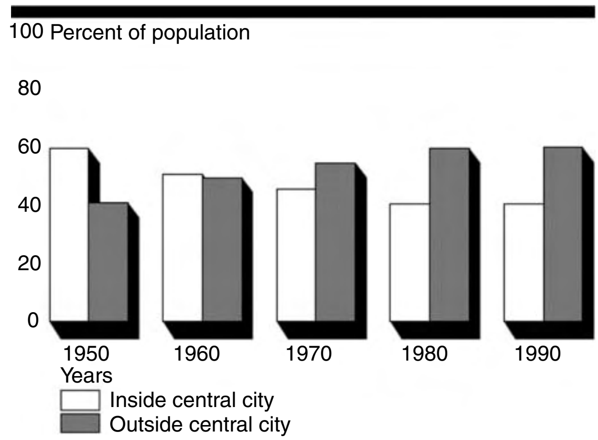


FIGURE 8.1 Location of population relative to central city. (From U.S. General Accounting Office (GAO), *Community Development: Extent of Federal Influence on "Urban Sprawl" is Unclear*, GAO-RCED-99-87, Washington, DC, 1999.)

quality school for a family's children, the safety and convenience of the routes to school, work, activities, and access to goods and services.

The best opportunity for conservation in transportation begins with the transportation–land use relationship. An energy-efficient transportation system exploits and integrates all modes rather than just the highway. However, current land use regulations, codes, and development trends are designed exclusively for the single-occupant vehicle (SOV) and do not efficiently support other travel options. A more balanced system that incorporates mass transit, walking, bicycling, and other alternatives would be more energy-efficient. These modes are less energy intensive and would reduce traffic congestion, vehicle idling, and inefficient stop-and-go traffic. However, land use must be designed for multimodal movement for such a balanced system to be realized.

Land use and the population in the U.S. have become more decentralized over time (see Figure 8.1). The distribution of land uses into residential, commercial, and business areas increases the distances between the many daily necessities of life so that walking and bicycling are either infeasible or unsafe; it also makes mass transit inefficient because stops would be required to serve each individual's needs. Therefore, personal vehicles are the most convenient and most widely chosen mode of transportation for daily travel needs given the type of development most commonly used in the U.S. A more systems-oriented approach, integrating pedestrian, bicycle, automobile, and mass-transit networks within a higher-density developmental structure would be more energy-efficient, but this situation is not the norm in the U.S. today.

8.2.2 Smart Growth

The terms "urban/suburban sprawl" and "smart growth" first appeared in the 1990s and often together, with the first assumed to be a problem and the latter assumed to be the solution. Without making vague generalizations, it should first be recognized that at the heart of the smart growth debate are "the rights of the individual versus the goal of the community" (Miller and Hoel 2002); but there is no reason that both individuals' rights and community goals cannot be achieved at the same time. In fact, the realization of community goals can enhance the realization of individuals' rights by increasing choices and improving livability. The point at which one or the other becomes threatened is when regulations begin to restrict individual rights or a few individuals prohibit the community from realizing its goals (Miller and Hoel 2002).

TABLE 8.1 Fundamental Principles of Smart Growth

Downs (2001)	Cervero (2001)
"Preserving large amounts of open spaces and protecting the quality of the environment"	Embracing "urban planning by anticipating and creating a vision of the future"
"Redeveloping inner-core areas and developing infill sites"	"Balanc[ing] the twin and often competing aims of urban design-form versus function"
"Removing barriers to urban design innovation in both cities and new suburban areas"	"Infrastructure investments are cleverly used to shape and leverage development"
"Creating a greater sense of community…and a greater recognition of regional interdependence and solidarity"	Regional governance to "deal with spillover and cross-boundary problems"

Source: From Downs, A., *Planning*, April, 20–25, 2001; Cervero, R., *Australian Planner*, 38(1), 29–37, 2001.

The term "urban/suburban sprawl" is generally defined as the growth of "low-density, automobile-dependent development on the fringe of cities" with both positive (increased home ownership, lower prices for business real estate) and negative (high infrastructure costs due to low-density development, increased traffic congestion, and consumption of green space) results (U.S. General Accounting Office 1999; Miller and Hoel 2002). In contrast, "smart growth" is considered the antidote to urban sprawl, by planning land use and transportation simultaneously for an integrated result. Specifically, the goals and strategies listed in Table 8.1 have been offered as the fundamental principles of smart growth.

Usually, these principles translate to high-density, mixed-use growth that incorporates transportation alternatives into design, such as transit-oriented development, pedestrian pockets, and bicycle networks. Shortening distances and providing transportation alternatives can conserve energy by decreasing the use of the automobile and increasing the use of less energy-intensive modes. The key is to maintain mobility and accessibility while curtailing the need for a motor vehicle for each and every trip.

Like most aspects of transportation and land use, the strategies to implement smart growth overlap and are interrelated, as shown in Figure 8.2 (Miller and Hoel 2002). As with any project, it is important to take each situation on a case-by-case basis, remembering that there is no silver bullet to design and execute the ideal transportation–land use plan.

8.2.3 Designing for Smart Growth

In an effort to implement smart growth, architectural movements, such as New Urbanism, neo-traditional or traditional neighborhood development (TND), and transit-oriented development (TOD), have appeared with the intention of countering sprawl and improving livability. These design strategies stress the importance of pedestrian accessibility through high-density development to reduce distances that would otherwise require an automobile or other motorized mode to travel. Typically, the standard design distance for pedestrians is one-quarter mile. Ten factors to improve the walkability of an area are shown in Figure 8.3. It is often assumed that people can walk that distance in about 5 min and that for any distance greater than that, they will choose to drive rather than walk. The designs mix land uses and housing affordability within an area so that residents can easily access shops, services, schools, employment centers, and other facilities from their homes. An emphasis on interconnectedness encourages streets laid out in a grid pattern as opposed to winding roads that terminate in cul de sacs. The designs often incorporate traffic calming features to reduce automobile speeds for safe and enjoyable walking and bicycling, as well as strict parking management to conserve the amount of land traditionally devoted to vehicle storage. TODs offer the added benefit of focusing on transit access for greater regional accessibility. Studies suggest that these high-density, mixed-use designs reduce automobile dependence (Cervero and Gorham 1995; Ewing 1995) and, therefore, the congestion and pollution associated with it. Table 8.2 indicates the impact of the design elements on vehicle travel.

Among the obstacles to New Urbanism and other smart growth approaches is the perceived risk of investment that financiers associate with the multiuse nature of such designs: developers understandably

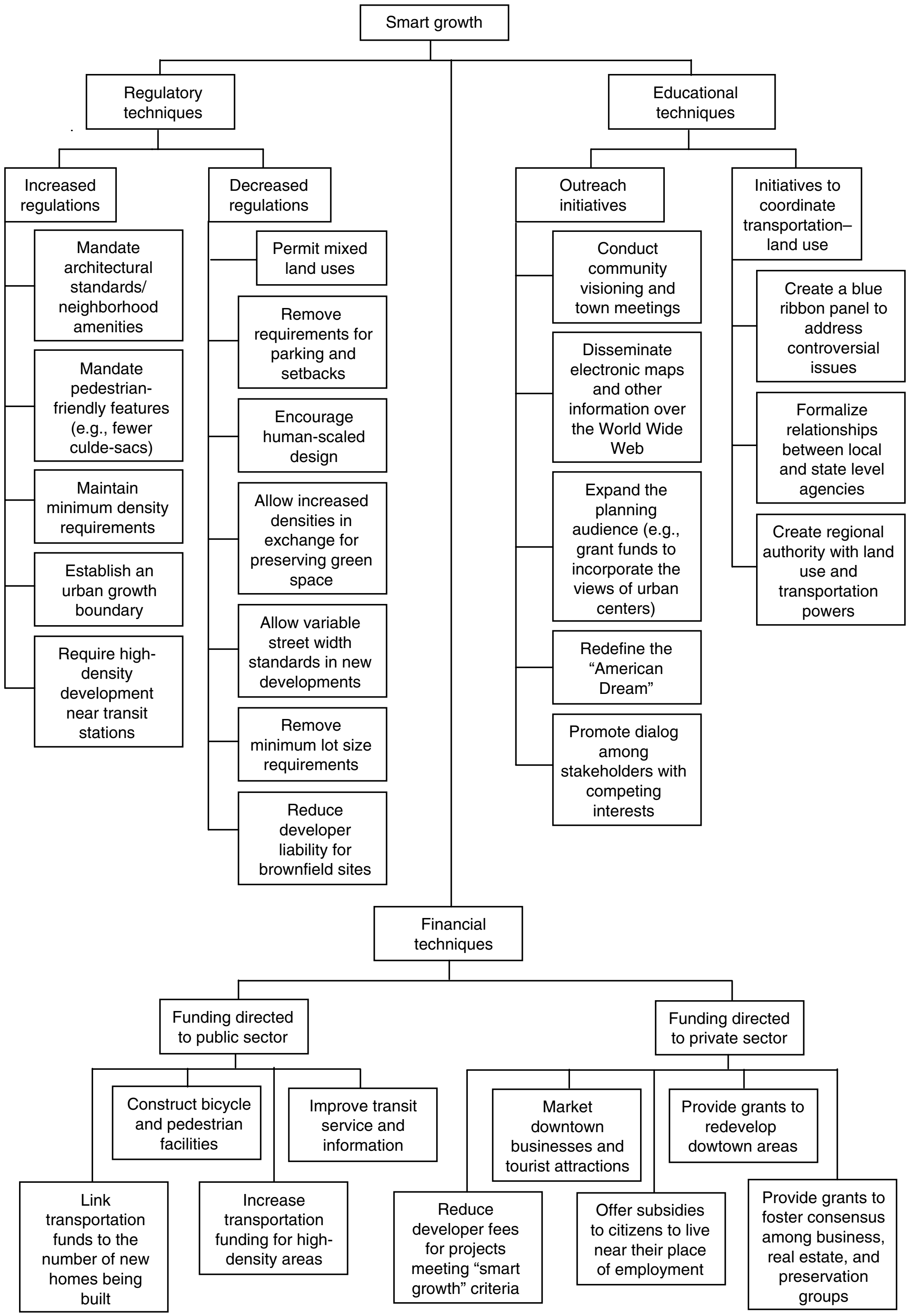


FIGURE 8.2 Emphasis areas for smart growth techniques. (From Miller, J. S., and Hoel, L. A., The "smart growth" debate: Best practices for urban transportation planning, pp. 1–24, 2002.)

10. Narrow streets
9. Traffic volumes
8. Sidewalks
7. Street trees
6. Interconnected streets

5. On-street parking
4. Lower traffic speeds
3. Mixed land use
2. Buildings fronting the street
1. Small block size

FIGURE 8.3 Top 10 walkability factors. (From Hall, R., Walkable thoroughfares through balanced design. Presentation at The Nuts & Bolts of Traditional Neighborhood Development Conference, Richmond, VA, 2005.)

want evidence that a new design concept will produce a high return on their investment. While a public–private partnership would help to dissipate the risk, these are frequently difficult to form and/or work within. Moreover, the cost of untouched land (greenfields) is often lower than land available for redevelopment (brownfields) that is usually located in infill areas between suburbs and central business districts. Brownfield redevelopment frequently carries with it the condition that the developer upgrade the infrastructure serving the area and/or assume responsibility for any known or unknown liabilities on the property, which can ultimately become a very expensive contingency. Finally, outdated zoning ordinances that prohibit the mixing of land uses and/or high-density development are another existing barrier to new urbanist designs (Farris 2001).

8.3 Alternative Transportation: Mass Transit

The efficiency of mass-transit service typically decreases with the density of land uses. However, density is not the single factor determining the success or failure of a transit system. Vuchic (1999) notes the success of the transit networks in spread-out areas of San Francisco, Washington, Montreal, Calgary, and particularly the suburbs of Philadelphia (with a lower population density than that of Los Angeles: 3500 people per square mile). Many planners and architects suggest a "hierarchy" of modes rather than the single mode system that dominates most areas: at the base is a network of bicycle- and pedestrian-friendly streets that support the local bus system, which in turn feeds a regional transit network. As each component relies on the others, their integration is essential for transit's success (Calthorpe and Fulton 2001). Furthermore, "the balance between car and transit use in central cities is strongly influenced by the character of the area (its physical design, organization of space, and types of development) and by the relative convenience and attractiveness of the two systems" (Vuchic 1999).

TABLE 8.2 Travel Impacts of Land use Design Features

Design Feature	Reduced Vehicle Travel (%)
Residential development around transit centers	10
Commercial development around transit centers	15
Residential development along transit corridor	5
Commercial development along transit corridor	7
Residential mixed-use development around transit centers	15
Commercial mixed-use development around transit centers	20
Residential mixed-use development along transit corridors	7
Commercial mixed-use development along transit corridors	10
Residential mixed-use development	5
Commercial mixed-use development	7

Source: From Victoria Transport Policy Institute (VTPI), In *Transportation Demand Management Encyclopedia*, 2005. http://www.vtpi.org/tdm/tdm45.htm

Several different types of transit exist to serve the needs of the public. "Demand response" describes the paratransit mode, by which a passenger calls a dispatcher who sends the transit vehicle (a shuttle bus or taxi) to the passenger's door and delivers her to her destination. Commuter rail denotes regional rail operating between a city and its suburban areas; light rail implies one or two cars using overhead electricity as a power source and operating within a city, often sharing the streets with automobiles; heavy rail operates at high speeds within a separate right-of-way. Bus rapid transit (BRT) is gaining popularity as a system that grants buses their own right-of-way so that they do not get caught in traffic congestion. BRT operates parallel to the street, such as in the median between travel lanes or in an exclusive bus-only lane (see Figure 8.4), and depending on the system, may also get prioritization at traffic signals so that upon approach, the light turns green and the bus will not have to wait at a red light. Table 8.3 summarizes the characteristics of each mode. Table 8.4 illustrates what percentages of the transit fleets use alternative fuels (i.e., fuels other than the conventionally used gasoline).

The factors that determine what mode and what technology are best for a given transit system include:

- The availability of a separate right-of-way
- The distance between/frequency of stops (i.e., will it be regional, express or local service?)
- The density of the surrounding area (to determine at what speeds the vehicle can safely travel)
- Expected passenger volumes
- Size of the city being served

A separate right-of-way is not dependent on the existing conditions of the street network and provides great reliability (since there are no traffice congestion delays), high speed, short trip times, and overall convenience for passengers.

The potential of mass transit to conserve energy is a large, untapped resource. Table 8.5 illustrates how much fuel could be saved by one person switching to mass transit for their daily commute to work. The reason for mass transit's high efficiency is its energy intensity, which is a result of the load factor of each vehicle. Table 8.6 provides passenger travel and energy use data for 2002, while Figure 8.5 provides the transit mode split on a passenger-mile basis (i.e., the distribution of travel on each mode per passenger per mile). Mass transit's efficiency could certainly be much higher compared to automobiles if more passengers used it and increased its load factor (Greene and Schafer 2003).

FIGURE 8.4 BRT photo. (From U.S. General Accounting Office (GAO), *Mass Transit: Bus Rapid Transit Shows Promise*, GAO-01-984, Washington, DC, 2001.)

TABLE 8.3 Summary of Transit Mode Characteristics

Mode	Vehicle	Fuel Options	Right-of-way	Notes
Technologies currently in use				
Bus	30- to 70-passenger bus	Gasoline, diesel, hybrid, battery, alternative fuel (e.g., natural gas, ethanol, etc.)	Existing street network	Can be caught in traffic congestion, potentially making service slow and unreliable Flexible routes can be adjusted as needed because travel medium is the existing street network
Bus rapid transit (BRT)	Conventional or guided buses	Gasoline, diesel, hybrid, battery, alternative fuel (e.g., natural gas, ethanol, etc.)	Separate from street, guideway, or exclusive "bus only"/high occupancy vehicle (HOV) lane	Generally has lower capital costs per mile than LRT (U.S. General Accounting Office 2001) Often combined with intelligent transportation systems for fast fare collection and traffic prioritization
Light rail (LRT)	One to two rail cars	Overhead electricity	Existing street network, elevated railway, subway, or at-grade track system (separate from street)	Also known as trolleys or streetcars Lower construction costs than conventional rail systems High-design flexibility because of many travel medium options Maximum speed: 65 mph
Metro/rapid transit system/heavy rail	Train	Electricity	Subway, elevated railway, or at-grade track system	Has very high passenger-carrying capacities and can operate at high speeds
Commuter/regional rail	Train	Electricity or diesel locomotive	Track system	Has very high passenger-carrying capacities and can operate at high speeds
Advanced technologies				
Monorail	Train	Electricity	Single rail, beam, or tube	Example: Seattle Center Monorail
Magnetic Levitation (MagLev)	Train	Electricity	Magnetic guideway	Uses magnetism to lift and propel train over tracks No wheels or moving parts; therefore, no friction Can operate at 300+ mph Still in research stage in most areas

TABLE 8.4 Alternative Power Vehicles by Mode, 2005

Mode	Percent Using Alternative Power
Bus	16.0
Commuter rail	47.8
Commuter rail locomotive	31.2
Demand response	4.9
Ferryboat	41.5
Heavy rail	100.0
Jitney	0.0
Light rail	100.0
Other rail	74.9
Trolleybus	100.0
Vanpool	0.8

Source: From Danchenko, D., *Public Transportation Fact Book*, American Public Transportation Association, Washington, DC, 2005.

8.4 Freight

The movement of goods in the U.S. is increasing. On an average day in 1993, 37 million tons of goods valued at $20 billion traveled 10 billion ton-miles; in 2002, those numbers rose to 43 million tons of goods valued at $29 billion moving almost 12 billion ton-miles (U.S. Department of Transportation/ Bureau of Transportation Statistics 2005a, 2005b). Freight is typically transported by air, truck, rail, pipeline, water, or any multimodal combination of these; the energy use and intensities of the modes are shown in Table 8.7 and Table 8.8. In 2002, the average value per ton shipped by air was $75,000, followed by truck at $725 and rail at $205. The U.S. DOT notes that, as value per ton rises, shipment sizes are likely to shrink to save costs: "shipments weighing less than 50,000 pounds (average payload of a typical truck) grew twice as fast (28%), measured by weight, than those weighing more than 50,000 pounds (13%) between 1993 and 2002, reflecting growth in smaller sized just-in-time deliveries" (U.S. Department of Transportation/Bureau of Transportation Statistics 2005a). Air is increasingly chosen for its timely deliveries, and air-freight shipments nearly doubled between 1993 and 2002 to $770 billion. As air shipments increase, so do truck shipments in order to fill the intermodal gap and deliver goods from their origin to the airport and from the airport to their destination. Trucking is by far the most widely used freight mode (Figure 8.6), increasing its ton-miles 44.4% between 1993 and 2002. For other shipments such as perishables and time-sensitive goods that need to travel very long distances (Figure 8.7), rail is used to ship low value-per-ton goods like coal, ores, and grains. It therefore has a relatively low share by value compared to its ton-miles,

TABLE 8.5 Examples of Fuel Savings to a Person Commuting to Work on Public Transportation

Length of Trip (miles)	Miles Traveled per Year (Based on 472 Trips per Year)	Annual Fuel Savings (gallons) Based on the Following Personal Vehicle Fuel Efficiencies					
		15 mpg	20 mpg	25 mpg	30 mpg	35 mpg	40 mpg
2	944	62.9	47.2	37.8	31.5	27.0	23.6
5	2,360	157.3	118.0	94.4	78.7	67.4	59.0
10	4,720	314.7	236.0	188.8	157.3	134.9	118.0
20	9,440	629.3	472.0	377.6	314.7	269.7	236.0
30	14,160	944.0	708.0	566.4	472.0	404.6	354.0
40	18,880	1,258.7	944.0	755.2	629.3	539.4	472.0
50	23,600	1,573.3	1,180.0	944.0	786.7	674.3	590.0
60	28,320	1,888.0	1,416.0	1,132.8	944.0	809.1	708.0

Source: From Danchenko, D., *Public Transportation Fact Book*, American Public Transportation Association, Washington, DC, 2005.

TABLE 8.6 Passenger Travel and Energy Use, 2002

	Number of Vehicles (Thousands)	Vehicle-miles (Millions)	Passenger-miles (Millions)	Load Factor (Persons/Vehicle)	Energy Intensities (Btu per Vehicle-mile)	Energy Intensities (Btu per Passenger-mile)	Energy Use (Trillion Btu)
Automobiles	135,920.7	1,658,640	2,604,065	1.57	5,623	3,581	9,325.9
Personal trucks	65,268.2	698,324	1,201,117	1.72	6,978	4,057	4,872.7
Motorcycles	5,004.2	9,553	10,508	1.22	2,502	2,274	23.9
Demand Response	34.7	803	853	1.1	14,449	13,642	11.6
Vanpool	6.0	77	483	6.3	8,568	1,362	0.7
Buses	[a]	[a]	[a]	[a]	[a]	[a]	191.6
Transit	76.8	2,425	22,029	9.1	37,492	4,127	90.0
Intercity[b]	[a]	[a]	[a]	[a]	[a]	[a]	29.2
School[b]	617.1	[a]	[a]	[a]	[a]	[a]	71.5
Air	[a]	[a]	[a]	[a]	[a]	[a]	2,212.9
Certified route[c]	[a]	5,841	559,374	95.8	354,631	3,703	2,071.4
General aviation	211.2	[a]	[a]	[a]	[a]	[a]	141.5
Recreational boats	12,409.7	[a]	[a]	[a]	[a]	[a]	187.2
Rail	18.2	1,345	29,913	22.2	74,944	3,370	100.8
Intercity[d]	0.4	379	5,314	14.0	67,810	4,830	25.7
Transit[e]	12.5	682	15,095	22.1	72,287	3,268	49.3
Commuter	5.3	284	9,504	33.5	90,845	2,714	25.8

[a] Data are not available.
[b] Energy use is estimated.
[c] Includes domestic scheduled services and ½ of international scheduled services. These energy intensities may be inflated because all energy use is attributed to passengers; cargo energy use is not taken into account.
[d] Amtrak only.
[e] Light and heavy rail.
Source: From Davis, S. and Diegel, S., *Transportation Energy Data Book*, Oak Ridge National Laboratory/U.S. Department of Energy, Oak Ridge, TN, 2004.

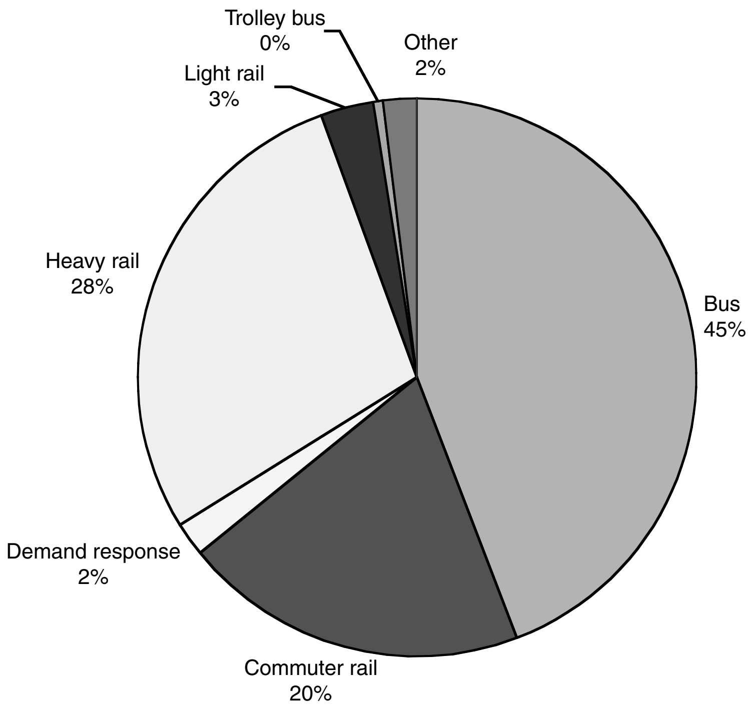


FIGURE 8.5 Transit mode split for 2003. (From Danchenko, D., *Public Transportation Fact Book*, American Public Transportation Association, Washington, DC, 2005. http://www.apta.com/research/stats/factbook/index.cfm)

which increased 33.8% between 1993 and 2002. U.S. pipelines transport crude oil and petroleum products around the nation, thereby playing a role not only in the movement of transportation fuel as freight, but also in supplying fuel for other transportation modes. In 2002, pipelines moved 750 billion ton-miles of crude oil and petroleum products. Water transportation modes are classified by shallow-draft and

TABLE 8.7 2002 Transportation Energy Use by Mode

		Trillion Btu	Thousand Barrels per Day Crude Oil Equivalent
Medium/heavy trucks		5,026.8	2,397.7
Air		2,212.9	1,071.1
	General aviation	141.5	70.2
	Domestic air carriers	1,734.5	838.1
	International air	336.9	162.8
Water		1,184.8	541.1
	Freight	997.6	444.6
	Recreational	187.2	96.5
Pipeline		935.4	12.8
Rail		621.0	259.9
	Freight (class I)	520.3	244.7
	Passenger	100.7	15.2
	Transit	49.3	1.9
	Commuter	25.8	5.4
	Intercity	25.6	7.9

Source: From Davis, S. and Diegel, S., *Transportation Energy Data Book*, Oak Ridge National Laboratory/U.S. Department of Energy, Oak Ridge, TN, 2004.

TABLE 8.8 2002 Energy Intensities of Freight Modes

Heavy single unit and combination trucks	23,432 Btu per vehicle-mile
Class I freight railroad	15,003 Btu per freight car-mile
	345 Btu per ton-mile
Domestic waterborne commerce	471 Btu per ton-mile

Source: From Davis, S. and Diegel, S., *Transportation Energy Data Book*, Oak Ridge National Laboratory/U.S. Department of Energy, Oak Ridge, TN, 2004.

deep-draft vessels: shallow-draft vessels operate on rivers, canals, harbors, the Great Lakes, the Saint Lawrence Seaway, the Intracoastal Waterway, the Inside Passage to Alaska, major bays and inlets, and in the ocean along the shoreline; deep-draft vessels operate in the open ocean. Water's domestic share of freight has declined since 1982, but its international share has increased to almost 80% of all U.S. international freight. The shares of multimodal freight include parcel, postal, and courier services that since 1993 have grown heavily to 11.8% of all U.S. freight shipments and average $39,000 per ton (the overall multimodal

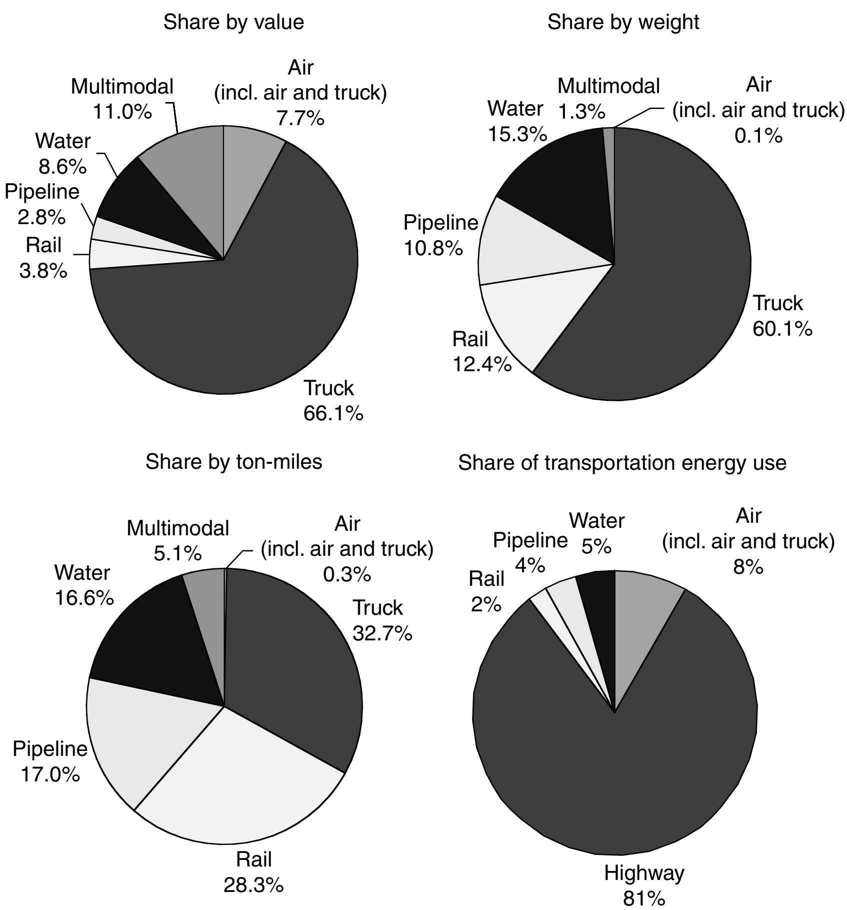


FIGURE 8.6 Shares of commercial U.S. freight activity in 2002. Transportation energy share data reflects all transportation modes, not just freight. (From U.S. Department of Transportation/Bureau of Transportation Statistics, Freight, 2005, http://www.bts.gov/programs/freight_transportation/html/more_freight.html; Davis, S. and Diegel, S., *Transportation Energy Data Book*, Oak Ridge National Laboratory/U.S. Department of Energy, Oak Ridge, TN, 2004.)

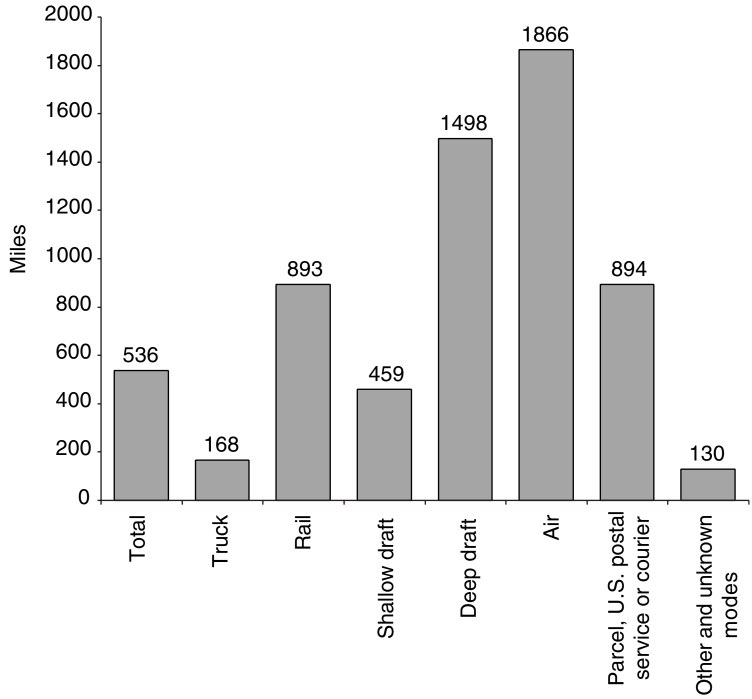


FIGURE 8.7 Average miles per shipment by mode in 2002. (From U.S. Department of Transportation/Bureau of Transportation Statistics, *2002 Commodity Flow Survey, United States*, 2004. http://www.bts.gov/publications/commodity_flow_survey/2002/united_states_final/)

average is $5000 per ton). The truck–rail combination is the most widely used multimodal combination (U.S. Department of Transportation/Bureau of Transportation Statistics 2004, 2005a).

Table 8.9 illustrates the top commodities shipped in 2002. The U.S. Commodity Flow Survey notes that pharmaceutical shipments are among the fastest growing commodities and hold the highest value per ton at almost $19,000, whereas gravel and crushed stone are the lowest at $7 (U.S. Department of Transportation/Bureau of Transportation Statistics 2004).

TABLE 8.9 Top Three Commodities Shipped in 2002 by Value, Weight, and Ton-Mile

Top Three Commodities	Total	Percent of Total
Shipped by value		
Electronic, electrical, and office equipment	$948 billion	11.2
Mixed freight (includes supplies and food for restaurants, grocery and convenience stores; hardware and plumbing supplies, office supplies, and miscellaneous)	$858 billion	10.1
Motorized and other vehicles (including parts)	$736 billion	8.7
Shipped by weight		
Gravel and crushed stone	1775 million tons	15.3
Coal	1255 million tons	10.8
Nonmetallic mineral products	910 million tons	7.9
Shipped by ton-mile		
Coal	562 billion ton-miles	17.6
Cereal grains	264 billion ton-miles	8.2
Basic chemicals	174 billion ton-miles	5.4

Source: From U.S. Department of Transportation/Bureau of Transportation Statistics, 2002 Commodity Flow Survey, United States, 2004. http://www.bts.gov/publications/commodity_flow_survey/2002/united_states_final/

Within freight, long-haul trucks have a vast potential for energy conservation. Trucks used 3653 trillion Btu in 2002 and have an energy intensity of 3476 Btu/ton-mile (Davis and Diegel 2004). Although heavy-duty trucks are equipped with turbo-charged, direct-injection diesel engines that are the most efficient internal-combustion engine (ICE) available with approximately 46% peak thermal efficiency (Greene and Schafer 2003), long-haul trucks consume more than 800 million gallons of fuel per year just idling. It is estimated that, on average, a tractor-trailer spends 6 h per day idling to heat and cool the sleeper cabs, keep the fuel warm, and eliminate the need for cold starts of the engine. While cab heaters, auxiliary power units, and truck-stop electrification could meet these needs with much greater efficiency, economics, regulatory legislation, and driver behavior are barriers. For example, there is concern that a truck's engine will be difficult to start after being shutdown all night, as well as the fact that the hum of the engine masks exterior noises that would otherwise disrupt the driver's sleep (Stodolsky, Gaines, and Vyas 2000; Sharke 2005).

Increases in the efficiency of air travel are expected to come through technological improvements in engine efficiency and aerodynamics (Greene and Schafer 2003). The efficiency of rail and water transport can be improved by advances in diesel engine technology, which powers nearly all locomotives and some ships, barges, and ferries. All modes (with the exception of the pipeline) can reduce consumption through improved aerodynamic designs (e.g., ship hulls), reduced friction (e.g., between moving parts), and weight reductions.

8.5 Motor Vehicles: Tank-to-Wheel Technologies

8.5.1 The Well-to-Wheel Efficiency Analysis

It is essential that any discussion of vehicle efficiency include the well-to-wheel analysis, shown in Figure 8.8. This section focuses on vehicle engines and their operational—or "tank-to-wheel"—efficiency (i.e., how efficiently the vehicle converts the fuel in its tank to energy to rotate its wheels and move the vehicle). Section 8.5 on transportation fuels will consider the other half of the well-to-wheel cycle, specifically, the well-to-tank portion of the cycle that examines the efficiency of obtaining a feedstock, refining it to a fuel fit for engine use, and distributing that fuel to the vehicles.

8.5.2 Background: Conventional Vehicles

In 2003, 228 million light vehicles and trucks were registered in the U.S.; since 1998, there has been an average increase in these registrations of 2% per year. If this trend continues, there will be 353 million light vehicles registered by 2025 (U.S. Department of Transportation/Bureau of Transportation Statistics 2005b). These vehicles accounted for 5,203,140 U.S. passenger-miles traveled, with 90% by highway.

Given these figures, the nation's dependence on oil is not surprising. The 1970s gasoline crisis created by OPEC's restriction of oil production prompted the federal government to regulate fuel consumption

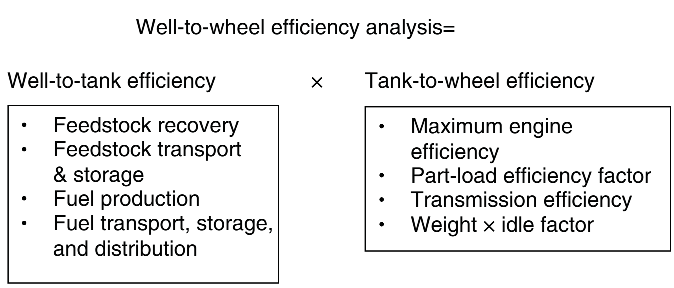


FIGURE 8.8 Well-to-wheel efficiency analysis. (From Kreith, F., West, R. E., and Isler, B., *Transportation Quarterly*, 56(1), 51–73, 2002a.)

TABLE 8.10 Automobile CAFE Standards Versus Sales-Weighted Fuel Economy Estimates, 1978–2004[a] (mpg)

Model Year[b]	CAFE Standards	CAFE Estimates[c]: Autos and Light Trucks Combined
1978	18.0	19.9
1979	19.0	20.1
1980	20.0	23.1
1981	22.0	24.6
1982	24.0	25.1
1983	26.0	24.8
1984	27.0	25.0
1985	27.5	25.4
1986	26.0	25.9
1987	26.0	26.2
1988	26.0	26.0
1989	26.5	25.6
1990	27.5	25.4
1991	27.5	25.6
1992	27.5	25.1
1993	27.5	25.2
1994	27.5	24.7
1995	27.5	24.9
1996	27.5	24.9
1997	27.5	24.6
1998	27.5	24.7
1999	27.5	24.5
2000	27.5	24.8
2001	27.5	24.5
2002	27.5	24.7
2003	27.5	25.0
2004	27.5	24.7

[a] Only vehicles with at least 75% domestic content can be counted in the average domestic fuel economy for a manufacturer.
[b] Model year as determined by the manufacturer on a vehicle by vehicle basis.
[c] All CAFE calculations are sales weighted.
Source: From Davis, S. and Diegel, S., *Transportation Energy Data Book*, Oak Ridge National Laboratory/U.S. Department of Energy, Oak Ridge, TN, 2004.

by lowering the national speed limit to 55 mph and by legislating via the U.S. Energy Policy and Conservation Act of 1975 that automakers increase vehicle mileage by creating corporate average fuel economy (CAFE) standards. It has been suggested that it takes 15 years or more for an increase in fuel economy to be reflected in the global fleet due to market penetration, so any immediate advances take a while to make an impact (Greene and Schafer 2003). Table 8.10 provides the average estimated fuel economy of vehicles between 1978 and 2004. The data show that the CAFE standards have not changed since 1990, reflecting the supply-side approach to transportation management taken by the federal government. CAFE estimates are weighted by sales and therefore illustrate the increase in sales of light trucks (i.e., sport utility vehicles (SUVs), minivans, and pickup trucks) shown in Table 8.11, the share of which has increased from 20% in 1976 to over half of total light vehicle sales today. Reducing vehicle weight, aerodynamic drag, and rolling resistance are basic areas (other than the engine) to target when attempting to improve fuel economy[1]: large SUVs and minivans are designed from the opposite end of the spectrum with their heavy weight, high and wide frontal cross-sections, and large tires. Fuel economy projections to 2025 from the Energy Information Administration predict that:

[1]It is suggested that a 10% reduction in weight on an average production vehicle can result in a 6% increase in fuel economy; a 10% reduction in aerodynamic drag can improve fuel economy by 3%; and a 10% reduction in rolling resistance can result in an increase in fuel economy of 2% (Office of Technology Assessment 1995; Bosch 1996).

TABLE 8.11 Light Vehicle Market Shares by Size Class, Sales Periods[a] 1976–2003

Sales Period[a]	1976	1980	1985	1990	1995	2000	2002	2003
Mini-compact	0.0%	3.8%	0.3%	0.6%	0.3%	0.1%	0.3%	0.5%
Sub-compact	21.7%	30.4%	15.7%	14.8%	10.4%	10.4%	3.7%	2.8%
Compact	23.5%	5.3%	23.2%	23.0%	22.4%	13.9%	18.8%	18.5%
Midsize	15.0%	27.2%	20.5%	18.3%	17.0%	19.4%	17.2%	16.1%
Large	18.2%	11.8%	10.0%	9.3%	9.0%	7.5%	8.1%	8.3%
Two-seater	1.7%	1.9%	2.5%	1.2%	0.4%	0.7%	0.8%	1.0%
Small pickup	1.4%	4.6%	5.7%	8.3%	7.3%	6.2%	4.5%	4.6%
Large pickup	13.1%	9.9%	11.1%	8.1%	10.0%	11.4%	13.0%	12.7%
Small van	0.2%	0.1%	2.9%	7.4%	8.6%	7.4%	6.9%	6.5%
Large van	4.8%	2.9%	3.5%	2.3%	9.1%	2.1%	2.1%	2.0%
Small utility	0.0%	0.5%	2.9%	2.9%	3.5%	4.4%	5.2%	5.2%
Medium utility	0.4%	1.3%	1.2%	3.2%	7.3%	12.5%	14.3%	16.5%
Large utility	0.1%	0.3%	0.5%	0.7%	1.0%	4.1%	5.1%	5.3%
Total light vehicles sold	12,096,613	11,311,043	15,203,880	13,739,090	14,658,736	17,285,055	17,009,538	16,315,470
Cars	80.1%	80.4%	72.1%	67.1%	59.5%	51.9%	49.0%	47.2%
Light trucks	19.9%	19.6%	27.9%	32.9%	40.5%	48.1%	51.0%	52.8%

[a] Sales period is October 1 of the current year through September 30 of the next year.

Source: From Davis, S., and Diegel, S., *Transportation Energy Data Book*, Oak Ridge National Laboratory/U.S. Department of Energy, Oak Ridge, TN, 2004.

"…in addition to increases in market penetration of advanced technologies, sales of hybrid and diesel vehicles will continue to increase. As a result, new car fuel economy in 2025 is projected to average 31.0 mpg, and new light truck fuel economy is projected to average 24.6 mpg—increases of 5.4% for cars and 14.1% for light trucks over the respective model year 2003 CAFE levels. Similar to historic trends, average engine power output is projected to increase to 215 horsepower for new cars sold in 2025 (26.3% higher than model year 2003) and 243 horsepower for new light trucks sold in 2025 (18.0% higher than model year 2003). Light truck sales are projected to account for 58.6% of new light-duty vehicle sales in 2025, and as a result the average fuel economy for all new light-duty vehicles sold is projected to increase by 7.2%, to 26.9 mpg in 2025." (U.S. Department of Energy/Energy Information Administration 2005)

The Energy Efficiency and Renewable Energy sector of the U.S. Department of Energy (DOE) offers suggestions for reducing gasoline consumption to improve vehicle mileage and efficiency. Table 8.12

TABLE 8.12 U.S. Department of Energy Recommended Gasoline Saving Methods

	Method	Estimated Fuel Economy Benefit
1.	Avoid driving aggressively	5%–33%
2.	Observe speed limit	7%–23%
3.	Avoid keeping unnecessary items in vehicle/ remove excess weight	1%–2% per 100 lbs.
4.	Avoid excessive idling	NA
5.	Use cruise control	NA
6.	Use overdrive gears	NA
7.	Keep engine properly tuned	4%–40%
8.	Check and replace air filters regularly	Up to 10%
9.	Keep tires properly inflated	Up to 3%
10.	Use vehicle's recommended grade of motor oil	1%–2%

Source: From U.S. Department of Energy/Energy Efficiency and Renewable Energy, Gas Mileage Tips, 2005e. http://www.fueleconomy.gov/feg/drive.shtml

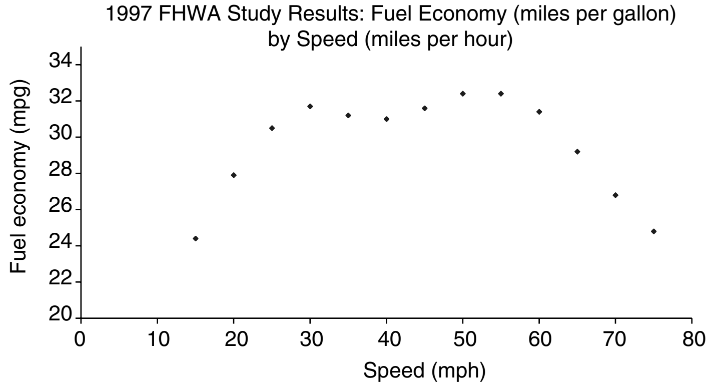


FIGURE 8.9 Relationship of engine speed to fuel economy. (From Davis, S. and Diegel, S., *Transportation Energy Data Book*, Oak Ridge National Laboratory/U.S. Department of Energy, Oak Ridge, TN, 2004.)

summarizes these methods and the estimated fuel savings of each. Method 2 is illustrated by a 1997 FHWA study examining fuel economy by speed in which nine conventional vehicles were tested. The result (Figure 8.9) shows that mileage suffers at speeds over 55 mph, mainly due to pumping and mechanical friction losses, wind resistance, and the fact that complete gas exchange is difficult to achieve at high engine speeds.

8.5.3 Internal Combustion: Spark Ignition and Compression Ignition

The ICEs that power motor vehicles function on a four-stroke cycle as shown in Figure 8.10, consisting of the compression, power/expansion, exhaust, and intake strokes.

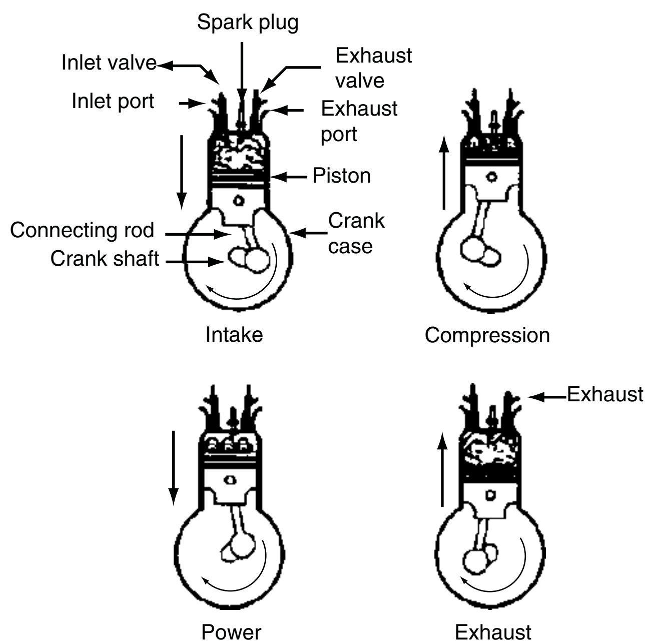


FIGURE 8.10 The four-stroke cycle of an IC engine. (From Klett, D. E. and Afify, E. M., In *The CRC Handbook of Mechanical Engineering*, ed. F. Kreith and D. Y. Goswami, Taylor & Francis, Boca Raton, FL, 2005.)

The difference between spark ignition (typically gasoline) and compression ignition (typically diesel fuel) is that spark ignition (SI) engines operate on the Otto cycle and use an air–fuel mixture in the cylinder at the beginning of the cycle and ignite this mixture via a spark plug. A compression ignition (CI, or diesel) engine operates on the diesel cycle, which begins with air in the cylinder that is compressed to a temperature above that of the autoignition temperature of the fuel; the fuel is then injected into the chamber and ignites on contact with the high-temperature air (Lichty and MacCoull 1967).

The thermal efficiency of diesel engines is higher than that of spark ignition engines: 29.5% tank-to-wheel efficiency for diesel engines compared to 22.0% for SI engines[2] (Kreith, West, and Isler 2001). One estimate suggests that "replacing a gasoline engine with a diesel in a 3000 pound car could result in a 30- to 40-percent improvement in fuel efficiency and such an exchange in a sport utility vehicle could provide a similar improvement in the range of 40- to 50-percent" (DeGaspari 2005). The primary reason for the difference in efficiency is that the diesel engines have higher compression ratios (the ratio of the cylinder's volume when the piston is at bottom dead center to the volume at top dead center) than spark ignition engines. SI engines have compression ratios between 8:1 and 12:1, while CI engines operate with compression ratios between 20:1 and 24:1. However, these higher combustion temperatures cause greater NO_x emissions from diesel engines (Bosch 1996).

Further reasons for the higher efficiency of diesel engines are that the engines are not throttled; they run lean and at lower revolutions per minute, and they burn the fuel more completely than SI engines. Although diesel engines usually have a higher initial cost than gasoline engines, diesel fuel tends to be less expensive than gasoline and it contains approximately 10% more energy per gallon than gasoline. These facts make diesels the choice for large engines such as those used in heavy equipment, large trucks, ships, train locomotives, and emergency power generators. In the past, poor emissions control and noise offset the efficiency of light-vehicle diesels, handicapping their popularity. But improvements in direct-injection, turbo-charging, and electronic controls are causing many consumers to reconsider diesels and to take advantage of their efficiency and torque.

These improvements have been so significant that in the spring of 2005, the Bush administration endorsed clean-diesel vehicles as part of its energy policy. Previously, diesels had only been popular in areas where gasoline is more heavily taxed than diesel fuel. A 2003 JD Power LMC forecast expects the global sales of light-diesel vehicles to increase from 12.5 million in 2003 to 27 million by 2015. The report notes a significant portion of this increase will be in North America, predicting that light diesels will secure 16% of new light-vehicle sales in 2015, compared to 4.5% in 2002 (JD Power LMC 2003). Provided that diesels can reduce their emissions to meet future standards, there is no reason why they should not become more popular in the market given their performance and efficiency.

A simpler version of the four-stroke cycle is the two-stroke that, unlike the four-stroke, does not require separate compression and expansion strokes to exhaust and intake gases; rather, the fresh air–fuel mixture enters the chamber at the end of the expansion stroke, while exhaust gases are forced out at the beginning of the compression stroke. This lighter and simpler SI engine is often used for motorcycles and mopeds; however, its less efficient gas exchange results in higher fuel consumption and greater hydrocarbon emissions. Two-stroke engines are frequently found in developing countries where motorized bicycles and rickshaws are popular modes of transportation.

Four sources of inefficiency in ICEs are: (1) the fact that combustion is not instantaneous, so the ideal cycle cannot be replicated; (2) mechanical friction losses (from the piston, crankshaft, and valves), particularly at high engine speeds; (3) aerodynamic frictional and pressure losses from air flow through the muffler and catalytic converter; (4) pumping losses due to throttling (in SI engines) (Office of Technology Assessment 1995). Therefore, opportunities for improved efficiency start in the actual combustion of the fuel in the chamber. Partial combustion occurs due to poor circulation of the air–fuel mixture in the chamber that is determined by the kinetic energy of the fuel spray, the thermal energy of the space, the combustion chamber shape, air flow, and utilization of partial combustion in a swirl chamber

[2]Tank-to-wheel efficiency used by Kreith, West, and Isler (2001) is the peak brake engine efficiency $\times$ part-load efficiency factor $\times$ transmission efficiency $\times$ weight $\times$ idle factor.

(Bosch 1996). Direct injection uses combinations of these attributes to influence the turbulence within the combustion chamber to improve air and fuel mixing and thereby induce more complete combustion.

Further opportunities exist in gas exchange to ensure that the available chamber volume is filled with a fresh air–fuel mixture instead of old exhaust gases from the previous stroke cycle. Gas exchange in four-stroke engines is entirely dependent on valve timing to open and close the valves at the right instant for exhaust release and fresh air intake, and for combustion. Variable valve timing, made possible by the introduction of electronics into IC engines, has helped to drastically improve the efficiency of conventional SI and CI engines with its ability to precisely control the valves.

Bosch (1996) further suggests that efficiency may be improved through "selective interruption of the fuel supply to individual cylinders to allow the remaining cylinders to operate at higher efficiency levels with improved combustion and gas exchange. Valve deactivation provides further reductions in power loss by allowing the intake and exhaust valves for the deactivated cylinders to remain closed. Cylinder deactivation entails immobilizing the mechanical power transfer components in these resting cylinders for further increases in mechanical efficiency." However, these measures "are not yet ready for general series production."

One version of the SI engine that may become competitive with diesel engines' efficiency is the direct injection stratified charge (DISC) engine. Rather than premixing the air and fuel, the fuel is injected into the chamber and aimed at the spark plug. The DISC engine is almost completely unthrottled (thereby reducing pumping losses), uses variable valve timing, and has a higher compression ratio than conventional engines (estimated at 13) (Office of Technology Assessment 1995). It is estimated that the DISC engine may offer 15%–20% higher fuel economy than conventional gasoline engines (U.S. Department of Energy/Energy Efficiency and Renewable Energy 2005b). However, because the DISC engine has not been able to meet emissions standards, it is not yet available in the U.S.

8.5.4 Electric Vehicles

Electric vehicles (EVs) have three main components: the battery pack, the controller, and the electric motor. The controller links the battery pack and the motor, regulating the amount of energy provided to the motor depending on the demands communicated by the accelerator. A significant feature of EVs is regenerative braking, which recovers the energy from the momentum that would otherwise be wasted when slowing down. EVs have a tank-to-wheel efficiency of 82% (Electric Power Research Institute 2004).

An EV battery must have/be:

- Quick discharge and recharge capability
- A long life cycle
- Low cost
- Recyclable
- Safe
- High specific energy (i.e., energy per unit mass, watt-hours per pound or Wh/kg), so that the battery pack's weight will not restrict vehicle performance
- High energy density (energy per unit volume), so that the battery pack will not take up too much room in the vehicle
- High specific power, so that the EV's performance will be comparable to that of a conventional IC vehicle
- Ability to function in a range of extreme operating temperatures

A summary of viable EV battery types and their attributes is shown in Table 8.13. Zinc and aluminum air batteries, ultracapacitors, and flywheels are also under research as energy storage devices for EVs, although whether they can deliver on their promises of long life cycles and high specific energy ($>$200 Wh/kg) is not fully known (U.S. Department of Energy/Energy Efficiency and Renewable Energy 2005c).

TABLE 8.13 Summary of Viable EV Batteries

Battery Type	Life Cycle (80% DOD)	Cost ($/kWh)	Recyclable	Specific Energy (Wh/kg)	Energy Density (Wh/l)	Operating Temperature
Lead acid	600–900 cycles[a] (85,000 miles[b])	150–200[3]	Yes[b]	45–50[b]	100–130[a]	Ambient[b]
Nickel cadmium (NiCad)	700–1200 cycles[a] (100,000 miles[b])	300–500[3]	Yes[b]	55[b]	100–150[a]	NA
Nickel metal hydride (NiMH)	>1200 cycles[c] (130,000–150,000 miles[c])	300–700[3]	Yes[b]	63[c]	150–200[a]	NA
Lithium ion (Li ion)	400–1200 cycles[a]	150–220[3]	NA	100[c]	100–200[a]	NA
USABC[d]	1000 cycles	100	NA	200	300	−40 to +85°C

[a] Data from Kreith, Potestio, and Kimbell (1999).
[b] Data from U.S. Department of Energy/Energy Efficiency and Renewable Energy (2005c).
[c] Data from Electric Power Research Institute (2004).
[d] U.S. Advanced Battery Consortium long-term goal for advanced batteries for EVs (U.S. Advanced Battery Consortium 1999).

Lawrence Berkeley National Laboratory notes that cost is the biggest barrier to battery technology: 1 kWh costs between $150 and $300 to store in an advanced battery, and an electric vehicle requires at least 30 kWh (Lawrence Berkeley National Laboratory, Environmental Energy Technologies Division 2004). However, the popularity of cell phones, laptop computers, and other portable electronic devices has led to significant advances in battery technology. A recent report by the Electric Power Research Institute (EPRI) found that although past estimates of electric vehicle battery lifetimes were only about 6 years or 75,000 miles, 5-year-old test vehicles had already traveled over 100,000 miles “with no appreciable degradation in battery performance or vehicle range” (Electric Power Research Institute 2004). This finding is extremely significant as it means that the high cost of replacing the battery pack of an electric vehicle is no longer a market barrier if the battery pack lasts the lifetime of the vehicle. Original projections estimated that battery packs would have to be replaced at least once during an EV’s lifetime, but the EPRI report says that “it is highly probable that (nickel metal hydride) batteries can meet 130,000–150,000 lifetime mileage,” thereby making hybrid and electric vehicles cost-competitive with conventional gasoline vehicles over the vehicle’s lifetime (see Table 8.14). The results of the EPRI study show that the long-term life cycle goal (1000 cycles at 80% depth of discharge (DOD)) of the United States Advanced Battery Consortium for advanced batteries for EVs—and perhaps the largest barrier to consumer acceptance—have been surpassed.

8.5.5 Hybrid Technologies

The introduction of the Toyota Prius and Honda Insight into the vehicle market in 2000 brought hybrid technology to the general public. The combination of a smaller ICE with an electric motor and battery

TABLE 8.14 Battery Cost to Break-Even Net Present Value

	Battery Price, $/kWh (at Which Net Present Values Are Equal After 10 Years, 117,000 miles)	
	Gasoline @ $1.75 per Gallon, EPRI	Gasoline @ $2.50 per Gallon, Our Recalculation of EPRI Values
Conventional vehicle	—	—
HEV 0	$385	$1135
PHEV 20	$316	$1648

Source: From Kreith, F. and West, R. E., Personal communication, 2005.

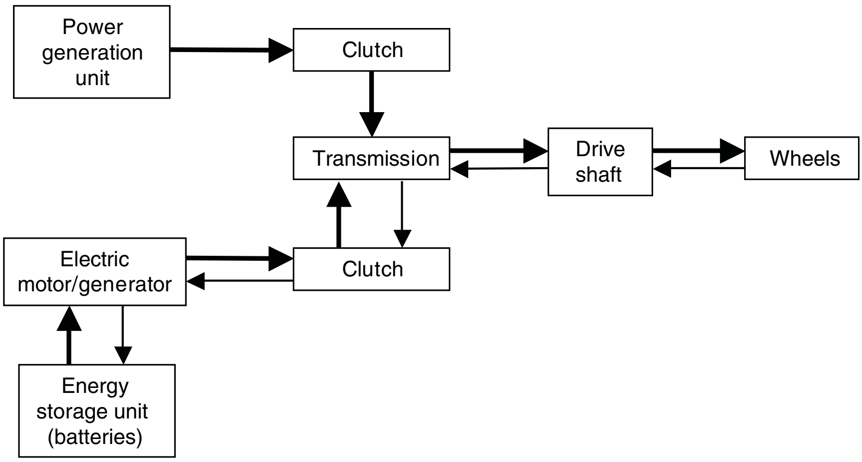


FIGURE 8.11 Parallel hybrid–electric vehicle (HEV) schematic. (From Kreith, F., Potestio, D. S., and Kimbell, C., *Ground Transportation for the 21st Century*, National Conference of State Legislatures, Denver, CO, 1999.)

pack (typically nickel metal hydride) has proven very popular; potential buyers are placed on wait lists because demand exceeds supply for hybrid vehicles. Hybrid vehicles can be configured in series, with the ICE generating power for the electric motor that then powers the wheels, or in parallel, where both the ICE and the electric motor are directly connected to the transmission. Commercially available hybrid vehicles (HEVs) use the parallel configuration (Figure 8.11) because it does not require energy to be converted as many times as in the series configuration. Using a "power split device," a hybrid such as the Toyota Prius optimizes the "blend" of energy provided by the gasoline engine and the electric motor for maximum efficiency, even shutting off the gasoline engine and operating exclusively on the electric motor in situations that do not demand powerful acceleration, such as during idling or low-speed, stop-and-go driving (see Table 8.15). Table 8.16 compares four hybrids with two of the most popular conventional vehicles on the market: the Honda Accord (a sedan) and the Ford Explorer (a sport utility vehicle).

The fundamental reason that hybrid configurations are so efficient is that ICEs are most efficient near full loads, at which most vehicles rarely operate. The part-load efficiency factor is defined as the vehicle's average efficiency (averaged over a specified driving cycle) over its efficiency at full load. Because of the driving schedule on which most vehicles operate, the part-load efficiency factor of conventional vehicles and hybrids is less than 1, as schematically shown in Figure 8.12. Supplementing an ICE with an electric motor provides optimum efficiency by providing additional energy when needed, and gives the hybrid a higher part-load efficiency factor than conventional vehicles because at partial loads such as low speeds and stop-and-go driving, the electric motor can tap the battery for energy to power the vehicle. At high speeds, such as highway driving, the ICE is mainly used. Regenerative braking also helps hybrid efficiency,

TABLE 8.15 Energy Optimization Schedule for Operation of a Toyota Prius

Operation	Power Is Supplied by…	
	Gasoline Engine	Electric Motor
Low speeds		☼
Heavy acceleration	☼	☼
Highway cruising	☼	☼
Deceleration/braking (i.e., regenerative braking)		☼

Source: From Toyota Motor Corporation, Hybrid Synergy Drive: Prius Demo, 2005b. http://www.toyota.com

TABLE 8.16 Comparison of Hybrids and Conventional Gasoline Vehicles

2005 Model	Toyota Prius	Honda Insight	Honda Civic Hybrid	Honda Accord Hybrid	Honda Accord EX (Conventional)	Ford Explorer 4WD (Conventional)
Engine	1.5-L, 4-cyl, CVT[a]	1.0-L, 3-cyl, CVT	1.3-L, 4-cyl, CVT	3.0-L, 6-cyl, 5-speed automatic	3.0-L, 6-cyl, 5-speed automatic	4.0-L, 6-cyl, 5-speed automatic
Valves	16	12	8	24	24	12
Body/max. cargo capacity	4-dr hatchback/16 $ft.^3$	2-dr hatchback/16 $ft.^3$	4-dr sedan/10 $ft.^3$	4-dr sedan/11 $ft.^3$	4-dr sedan/14 $ft.^3$	4-dr SUV/86 $ft.^3$
EPA estimated mileage (mpg)	60 city/51 hwy	57 city/56 hwy	48 city/47 hwy	29 city/37 hwy	21 city/30 hwy	14 city/20 hwy
Gas tank capacity (gallons)	11.9	10.6	13.2	17.1	17.1	22.5
Range (miles per tank)	714 city/607 hwy	604 city/594 hwy	634 city/620 hwy	496 city/633 hwy	359 city/513 hwy	315 city/450 hwy

[a] CVT, continuously variable transmission.

Source: From Toyota Motor Corporation, Vehicle Comparison, 2005a. http://www.toyota.com/toyotacomparator/displayComparator.do?toyotaModelCode=prius

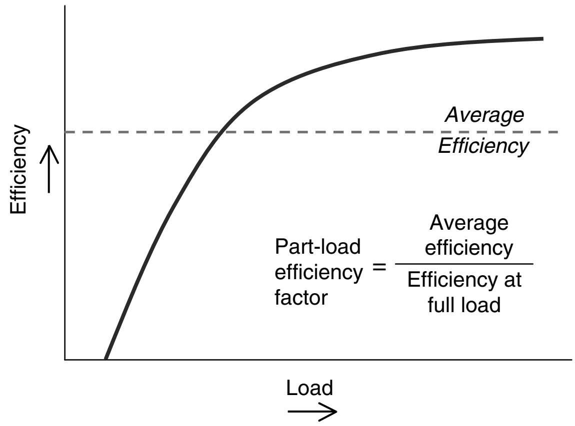


FIGURE 8.12 Schematic of ICE efficiency curve.

as the forward momentum of the vehicle—which is normally wasted when braking in a conventional vehicle—is recovered and the energy is used to recharge the batteries. Different driving situations require different power sources as shown in Table 8.15, so hybrids can utilize a computer to manage energy sources in any given situation to optimize efficiency.

Aside from the hybrid technology currently on the market, there are more varieties of hybrids on the horizon. One such option is plug-in hybrids (PHEVs) that can reportedly achieve 100–180 mpg (based on a converted Toyota Prius model). The only difference in the configuration of a PHEV and an HEV is more batteries and the ability to be plugged into a standard household outlet to recharge the batteries. The excess battery power enables the PHEV to travel 20–60 miles on electricity alone, depending on the size of the battery pack (Carey 2005). A close examination of the well-to-tank portion of the efficiency analysis is necessary to evaluate the PHEV's efficiency because a large portion of the fuel consumed by the vehicle will be from electricity that typically has a low production efficiency (but which will be partially offset by load-leveling of the grid) (Kreith, West, and Isler 2001).

Hybrids were first introduced to the U.S. in 1999 with the Honda Insight, and the 2004 Toyota Prius received over 12,000 purchase requests before it was even available. In 2004, approximately 88,000 hybrids were sold, constituting 0.5% of total vehicle sales. Studies suggest that hybrid sales will plateau at about 3% around 2011 due to the higher price of hybrids and the increasing availability of high-efficiency gasoline and diesel engines (Greene, Duleep, and McManus 2004; Porretto 2005). Hybrid tank-to-wheel efficiency (using natural gas as the fuel for the ICE) is estimated to be 35.7% when the ICE is a spark-ignition engine and between 37.6% and 44.7% using various diesel configurations for the ICE (Kreith, West, and Isler 2001). It has been suggested that the long-term impact of current hybrid technology could raise the fleet fuel economy by about 10%, although this estimate clearly depends on the market penetration of hybrids into the global fleet (Greene, Duleep, and McManus 2004). More optimistic estimates project an increase in fuel economy up to 40% after 30 years (National Research Council 2004).

8.5.6 Fuel Cells

The simplest description of a fuel cell that it is an electrochemical device similar to a battery, except that a battery is recharged with electricity, whereas a fuel cell is refueled with hydrogen. There are several varieties of fuel cells, but the most promising option for transportation applications is the proton exchange membrane (PEM) fuel cell, mainly due to the fact that its operating temperature is lower than that of most other types of fuel cells and it has a high power density. The fuel cell is made up of an anode, a cathode, and an electrolyte that, in the case of the PEM fuel cell, is the proton exchange membrane.

When provided with pure hydrogen and oxygen from the air, the following chemical reactions take place:

$$\text{Anode}: \quad 2H_2 \rightarrow 4e^- + 4H^+$$

$$\text{Cathode}: \quad O_2 + 4H^+ + 4e^- \rightarrow 2H_2O$$

$$\text{Net reaction}: \quad 2H_2 + O_2 \rightarrow 2H_2O$$

The fuel cell provides energy by sending the electrons released from the hydrogen through an external circuit. The only by-product of hydrogen fuel cell operation is water (see Figure 8.13).

A fuel cell stack is several fuel cells bundled together, much like a battery pack in an EV. The part-load efficiency factor plays a very large part in the inherently high efficiency of a fuel cell stack. Whereas conventional and hybrid vehicles are more efficient near full loads, fuel cells are most efficient at part loads, giving them a part-load efficiency factor greater than 1 (see Figure 8.14).

For use in a vehicle, several additional components are necessary, such as an air compressor, cooling system, water management system, and hydrogen fuel supply system. Although onboard reformers were once considered a possible tool for converting a fuel such as methanol or gasoline to hydrogen, their cost,

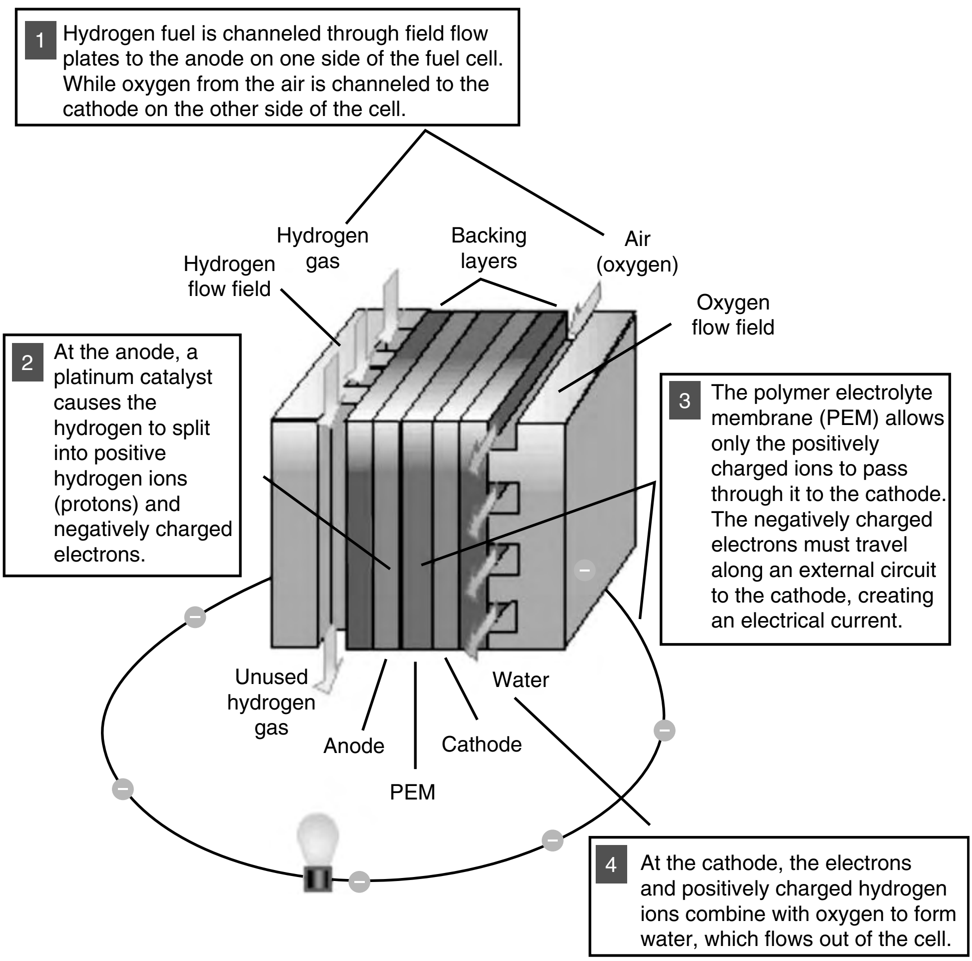


FIGURE 8.13 Proton exchange membrane fuel cell. (From U.S. Department of Energy/Energy Efficiency and Renewable Energy, Fuel Cell Vehicles, 2005d. http://www.fueleconomy.gov/feg/fuelcell.shtml)

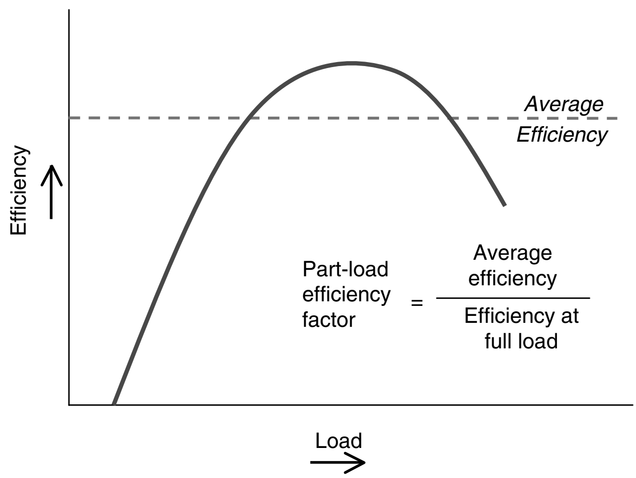


FIGURE 8.14 Schematic of fuel cell efficiency curve.

complexity, increases in emissions, decreases in system efficiency, and safety concerns led all major automobile manufacturers to abandon this initiative by 2003 (National Research Council 2004). It has been estimated that the hydrogen fuel cell stack efficiency is between 40% and 50%, but that with the auxiliary systems needed to support the stack in a transportation application (directly supplied with hydrogen), the result is an overall system efficiency (tank-to-wheel) of between 35% and 40% (Kolke 1999; Kreith, West, and Isler 2001). Fuel cells can also be arranged in a hybrid configuration like conventional hybrids to optimize energy use by storing excess energy in a battery pack.

Several other obstacles lie in the path of mass-produced fuel cell vehicles, such as cost, fuel supply, storage, and safety. Much more research is required before fuel cells can be considered a candidate for a vehicle propulsion system: for transportation purposes, a fuel cell stack will need to be able to last 5000 h (translating to 150,000–200,000 miles) as well as be affordable. Current fuel cell stacks are very fragile, lasting only thousands of hours in the laboratory because they decay under operational pressures (Baard 2003). Moreover, the fuel cell is not the “green” technology that the tank-to-wheel operation makes it appear to be because of the energy required for and the pollution potentially created during hydrogen production (depending on production method). Until the obstacles with hydrogen production (see Section 8.6) are solved to make hydrogen safe and efficient as a fuel, it is not a feasible option as a transportation energy source for any type of propulsion system. Even if these obstacles are overcome, there is still the colossal feat of building a fueling infrastructure to distribute hydrogen to vehicles.

8.6 Transportation Fuels

Generating energy sources for the technologies discussed in Section 8.5 makes up most of the phases of the well-to-wheel efficiency analysis shown in Figure 8.8. Vehicle operations would not be possible without feedstock recovery, feedstock transport and storage, fuel production, and fuel transport, storage, and distribution. It is therefore imperative to consider what options are available for fuel sources before determining what technology is to be used for a mass-produced automobile propulsion system.

8.6.1 Feedstocks

Estimating when world oil supplies will peak is a difficult task that leads to a wide range of timeframes. Due to the many variables involved, such as market forces and technology, it is believed that world production of conventional oil will peak between year 2015 and 2030. Both "conventional" and "unconventional" feedstocks are available to supplement conventional oil production. According to Greene, Hopson, and Li (2004), "conventional oil includes liquid hydrocarbons of light and medium gravity and viscosity, occurring in porous and permeable reservoirs…unconventional oil comprises deposits of greater density than water (e.g., heavy oil), viscosities in excess of 10,000 cP (e.g., oil sands), or occurrences in tight formations (e.g., shale oil)." Although unconventional feedstocks for oil can be expensive and environmentally damaging to recover and refine, it is expected that they will be used as conventional supplies dwindle because the gasoline distribution infrastructure is already in place (Greene and Schafer 2003).

Coal, of which the U.S. has the world's largest reserves (26%) (U.S. Department of Energy/Energy Information Administration 2004), can be used as a feedstock for diesel fuel via the Fischer–Tropsch process. Coal can also be used to create a synthetic diesel fuel called dimethyl ether (DME), as well as methanol. Under ambient conditions, DME is a gas, but only mild pressure is required to liquefy it. Being a domestic fuel keeps coal relatively inexpensive and does not involve the national security concerns or energy dependence associated with foreign oil.

Natural gas (NG) is also abundant and relatively inexpensive, which, along with its versatility, make it an attractive feedstock for many fuels. It can be directly used in vehicles in its liquid form (LNG) or its compressed gaseous form (CNG). Like coal, it can be a feedstock to produce diesel fuel via the Fischer–Tropsch process. Currently, the most widely used feedstock for hydrogen is natural gas because of its low cost and abundance. Table 8.17 provides the well-to-tank efficiencies of several fuels using natural gas as the feedstock. Combining these results with the tank-to-wheel efficiencies described in Section 8.5 provides the total cycle efficiency (Table 8.18). The results indicate that of the fuels considered, CNG and diesel (created using NG via the Fischer–Tropsch process) are the most efficient to produce. The well-to-wheel results indicate that SI and CI hybrid configurations are more efficient than a fuel cell powered by hydrogen produced using NG as the feedstock. Similarly, a well-to-wheel energy and greenhouse gas emissions analysis at MIT found "no current basis for preferring either fuel cell or ICE hybrid power plants for midsize automobiles over the next 20 years or so. That conclusion applied even with optimistic assumptions about the pace of fuel cell development" (Weiss et al. 2003). Hydrogen as a fuel will only fulfill its implied environmental promises if it is produced via electrolysis of water using renewable or nuclear energy. If an efficient hydrogen production method is developed, there is still the difficulty of replacing the existing gasoline distribution infrastructure with a hydrogen one.

TABLE 8.17 Fuel Production Efficiencies Using Natural Gas (NG) as Feedstock

Fuel	NG Feedstock Production[a]	Conversion NG to Fuel	Fuel Storage, Transmission, and Distribution	Additional Compression[b]	Overall Efficiency of Fuel Production
CNG	0.95		0.97	0.89[c]	87.5
Hydrogen (gaseous)	0.95	0.785[c]	0.97[d]	0.86[c]	59
Fischer–Tropsch (F–T) diesel	0.95	0.72	0.97		67
Electricity					48
Methanol	0.95	0.624[e]	0.97		57.5

[a] 95% = 97.0% (recovery) × 97.5% (processing).

[b] Assuming 90% compressor efficiency, 55% conversion efficiency (NG to electricity), and 93% electricity transmission and distribution.

[c] Efficiency of hydrogen from natural gas by steam reforming (see Kreith, West, and Isler (2001) for additional information).

[d] Assuming gaseous hydrogen from centralized plants.

[e] Via conventional steam reforming.

Source: From Kreith, F., West, R. E., and Isler, B., *Journal of Energy Resources Technology*, 124(September), 173–179, 2002b; Kreith, F. and West, R. E., *Journal of Energy Resources Technology*, 126(12), 249–256, 2004.

TABLE 8.18 Comparison of Natural Gas Well-to-Wheel Efficiencies of Technologies Using Natural Gas as Feedstock

Vehicle Drive Technology	Fuel	Well-to-Wheel Efficiency[a] (%)
Battery + electric motor	Electricity from NG combined cycle	39[b]
Hybrid SI	NG	32
Hybrid diesel	NG + F–T diesel[c]	32
Hybrid diesel	F–T diesel[c]	30
Fuel cell + electric motor	Hydrogen[d]	27
Hybrid SI	Hydrogen[d]	22
Conventional diesel	NG + F–T diesel[c]	22
Conventional SI	NG	19
Conventional diesel	F–T diesel[c]	19
Fuel cell + electric motor	Methanol[e]	16
Conventional SI	Hydrogen[d]	14
Fuel cell + electric	Hydrogen[f]	13

[a] Well-to-wheel efficiency is the efficiency of use of the natural gas, starting with gas in well.
[b] Given the advances in battery technology noted in Electric Power Research Institute (2004), the well-to-wheel efficiency for an EV has been reevaluated: "The tank-to-wheel efficiency for a battery all-electric vehicle according to Electric Power Research Institute (2004) is 0.82…Combining this with a well-to-grid efficiency of 48% (for an NG combined cycle plant) and a distribution efficiency of 93%, gives $(0.93 \times 0.48 \times 0.82) = 0.36$" (Kreith and West 2005).
[c] Diesel fuel made by Fischer–Tropsch synthesis from natural gas.
[d] Hydrogen made by steam reforming of natural gas.
[e] Methanol made from natural gas and converted to hydrogen by onboard reactor.
[f] Hydrogen made by electrolysis with electricity from natural gas combined cycle.
Source: From Kreith, F. and West, R. E., *Journal of Energy Resources Technology*, 126(12), 249–256, 2004.

8.6.2 Fuels

One of the often overlooked aspects of transportation fuel is storage onboard the vehicle. In addition to the phases of fuel production and distribution, critical aspects of the fuel include whether it is liquid or gas, its energy density, and the size, type, and weight of the tank needed to store it on the vehicle. Table 8.19 compares the storage needs of various fuels to gasoline. High-pressure or high-volume tanks can be quite large and/or heavy, thereby affecting the vehicle's carrying capacity and potentially degrading fuel economy. Fuels with high energy densities will have fewer requirements for onboard fuel storage systems.

Biodiesel is a combination of oils or fats (usually agricultural residue such as soybeans or animal fats) with an alcohol such as ethanol or methanol. In a vehicle, it can be used alone (neat) or blended with

TABLE 8.19 Onboard Storage Requirements for Various Fuels

Fuel	Tank System Containing 15 GGE of Fuel	
	Fuel Volume (gallons)	Total Mass (lbs)
Diesel	13.6	115
Gasoline	15.0	115
Liquefied petroleum gas	20.7	115
Ethanol	22.7	179
Methanol	31.0	240
Compressed natural gas (3600 psi)	46.3	268
Compressed hydrogen (5000 psi)	175.1	408
Liquid hydrogen	56.9	298

GGE, gallons of gasoline equivalent.
Source: From Greene, D., and Schafer, A., *Reducing Greenhouse Gas Emissions From U.S. Transportation*, Pew Center on Global Climate Change, Arlington, VA, 2003.

petroleum diesel (also known as *number 2 diesel*) using 20% biodiesel and 80% petroleum diesel, which helps to alleviate the fact that it freezes at a higher temperature than conventional diesel. With the exception of NO_x, biodiesel has lower tailpipe emissions than conventional diesel. The energy content is comparable to conventional diesel and it can be used in existing diesel engines without modifications.

Compressed NG is used in light-duty vehicles at 3000–4000 psi, whereas liquefied NG is used in heavy-duty vehicles and stored at about −260°F. Unlike LNG, CNG requires heavy pressurized tanks for onboard storage and has a lower energy content. Both burn more cleanly than gasoline or diesel, but have expensive infrastructure requirements because one is a compressed gas and the other is a cryogenic liquid. As noted, natural gas can be—and currently is—used directly in commercially available vehicles: it is estimated that there are up to 120,000 NG vehicles now on the road. Honda recently announced that it will offer a lease on an at-home refueling device with the purchase of a NG-powered Civic. The device fills the vehicle's tank overnight using the household's NG connection, thereby freeing the vehicle's owner from finding and using a NG filling station (of which there are 1100 in the U.S.) (Woodyard 2005). This obstacle has prevented many private parties from buying an NG vehicle, and most sales have been to government or corporate fleets that have private filling stations.

Electricity has a very low energy density that requires many heavy batteries for onboard energy storage. Most electricity is generated using coal as the feedstock, which benefits the U.S. because it has large reserves of coal, but this process can increase sulfur emissions. However, some of the advantages of electricity are that it can be produced using renewable sources and it creates no tailpipe emissions.

Ethanol is primarily made from corn; therefore, it can be produced domestically with a renewable feedstock. It may, however, have to compete with food uses in the corn market, but research is underway to extract the sugars in cellulose (e.g., plant material such as corn stalks and wheat straw, municipal solid waste) and convert it to ethanol. As an agricultural residue, cellulose would be much more abundant (300 million tons produced in the U.S. each year) than corn kernels (Morris 2003). It is estimated that while corn-derived ethanol generates about 1.4 times the energy required to produce it, cellulosic materials can produce ethanol with a 10–1 return on energy input (Gartner 2005a).

As a fuel, ethanol can be used neat (ethanol/gasoline mixture of 85% or more ethanol, known as E85) or as a blend (ethanol/gasoline mixture of 10% or less ethanol). The blend scenario helps ethanol to penetrate the market because any conventional gasoline vehicle can use it and it can be distributed with the current infrastructure. Neat ethanol requires a unique refueling infrastructure (there are currently 200 E85 stations in the U.S.) as well as small modifications (costing about $160 per vehicle; Morris 2003) to the vehicle so that it can run on ethanol, gasoline, or any combination of the two. There are currently 2.3 million of these "flexible-fuel" vehicles on the road. Ethanol is also being tested in fuel cell vehicles that use reformers to extract the hydrogen from the ethanol; for this purpose, the ethanol can be of a lower grade than that used in conventional ICEs, which translates to more efficient ethanol production. Ethanol is a liquid at ambient temperatures, so onboard storage requirements are not difficult to meet, although if used neat, a larger tank is required to maintain the same range as a conventional gasoline vehicle.

In the U.S., 95% of the hydrogen is made from NG through steam–methane reforming, compared to 50% for global production. Electrolysis of water can be used to produce hydrogen, but this method consumes large quantities of energy. The nuclear industry expects to be a major player in hydrogen production through electrolysis because it does not contribute to greenhouse gas emissions. Coal gasification produces 20% of the world's hydrogen, but creates significant emissions during hydrogen extraction. Solar power could be used to power electrolysis, but so far this method has not been cost-competitive with other alternatives (Morris 2003). Interest in using ethanol as a feedstock for hydrogen is growing, as ethanol uses a renewable feedstock and the process of reforming it into hydrogen is similar to that of reforming NG into hydrogen. The idea that a feedstock could be reformed into hydrogen onboard the vehicle does not yet show any promise, as no onboard reformer technology has been developed that is efficient enough to make this process feasible.

Gaseous hydrogen has a very low density and must be stored under extremely high pressure (5000–10,000 psi), creating another safety hazard in addition to its explosiveness. Liquid hydrogen has a higher energy density but must be stored at −253°C. In terms of energy density, it is

TABLE 8.20 Fuel Comparison Chart

	Gasoline	No. 2 Diesel	Biodiesel	CNG	Electricity	Ethanol (E85)	Hydrogen	LNG	Liquefied Petroleum Gas (LPG)	Methanol (M85)
Main fuel source	Crude oil	Crude oil	Soybean oil, waste cooking oil, animal fats, grapeseed oil	Underground reserves	Coal, natural gas, hydroelectric, renewables	Corn, grains, agricultural waste	Natural gas, methanol, other sources	Underground reserves	A by-product of petroleum refining or natural gas processing	Natural gas, coal, woody biomass
Energy content per gallon	109,000–125,00 Btu	128,000–130,000 Btu	117,000–120,000 Btu	33,000–38,000 Btu @ 3000 psi; 38,000–44,000 Btu @ 3600 psi	N/A	~88,000 Btu	N/A	~73,500 Btu	~84,000 Btu	56,000–66,000 Btu
Energy ratio gasoline to fuel	1.0	1.04–1.17	1.1–1 or 90% (relative to diesel)	3.94–1 or 25% at 3000 psi; 3.0–1 @ 3600 psi	N/A	1.42–1 or 70%	N/A	1.55–1 or 66%	1.36–1 or 74%	1.75–1 or 57%
Physical state	Liquid	Liquid	Liquid	Compressed gas	N/A	Liquid	Compressed gas or liquid	Liquid	Liquid	Liquid
Types of vehicles currently available	All types of vehicle classes	Many types of vehicle classes	Any vehicle that runs on diesel—no modifications are needed for up to 5% blends. Many engines also compatible with up to 20% blends	Many types of vehicle classes	Neighborhood electric vehicles, bicycles, light-duty vehicles, medium- and heavy-duty trucks and buses	Light-duty vehicles, medium- and heavy-duty trucks and buses—these vehicles are flexible-fueled vehicles that can be fueled with ethanol, gasoline, or any combination of the two fuels	No vehicles are available for commercial sale yet, but some vehicles are being leased for demonstration purposes	Medium- and heavy-duty trucks and buses	Light-duty vehicles that can be fueled with propane or gasoline, medium- and heavy-duty trucks and buses that run on propane	Mostly heavy-duty buses
Environmental impacts of burning the fuel	Produces harmful emissions; but gasoline and gasoline vehicles are rapidly improving and emissions are being reduced	Produces harmful emissions; but diesel and diesel vehicles are rapidly improving and emissions are being reduced especially with after-treatment devices	Reduces particulate matter and global warming gas emissions compared to conventional diesel; but NO_x emissions may be increased	CNG vehicles can demonstrate a reduction in ozone-forming emissions compared to some conventional fuels; but hydrocarbon (HC) emissions may be increased	EVs have zero emissions; but some amount of emissions can be contributed to power generation	E85 vehicles can demonstrate a 25% reduction in ozone-forming emissions compared to reformulated gasoline	Zero regulated emissions for fuel cell-powered vehicles; NO_x emissions possible for ICEs operating on hydrogen	LNG vehicles can demonstrate a reduction in ozone-forming emissions compared to some conventional fuels; but HC emissions may be increased	LPG vehicles can demonstrate a 60% reduction in ozone-forming emissions compared to reformulated gasoline	M85 vehicles can demonstrate a 40% reduction in ozone-forming emissions compared to reformulated gasoline

Energy security impacts	Manufactured using mostly imported oil, which is not an energy secure option	Manufactured using mostly imported oil, which is not an energy secure option	Biodiesel is domestically produced and has a fossil energy ratio of 3.3–1, which means that its fossil energy inputs are similar to those of petroleum	CNG is domestically produced. The U.S. has vast natural gas reserves	Electricity is generated mainly through coal-fired power plants. Coal is the U.S.'s most plentiful fossil energy resource and coal is its most economical and price stable fossil fuel	Ethanol is produced domestically and it is renewable	If produced from renewable resources, hydrogen can reduce dependence on foreign oil	LNG is domestically produced and typically costs less than gasoline and diesel fuels	LPG is the most widely available alternative fuel with 3400 refueling sites nationwide. The disadvantage of LPG is that 45% of the fuel in the U.S. is derived from foreign oil	Methanol can be domestically produced from renewable resources
Fuel availability	Available at all fueling stations	Available at select fueling stations	Available in bulk from an increasing number of suppliers. There are 22 states that have some biodiesel stations available to the public	More than 1100 CNG stations can be found across the U.S., with the highest concentration of stations in California. Home fueling is now available	Most homes, government facilities, fleet garages, and businesses have adequate electrical capacity for charging, but special hookups or upgrades may be required. More than 600 electric charging stations are available in California and Arizona	Most of the E85 fueling stations are located in the Midwest, but in all, approximately 150 stations are available in 23 states	There are only a small number of hydrogen stations across the county. Most are available for private use only	Public LNG stations are limited (only 35 nationally). LNG is available through several suppliers of cryogenic liquids	LPG is the most accessible alternative fuel in the U.S. There are more than 3300 stations nationwide	Methanol remains a qualified alternative fuel as defined by EPA, but it is not commonly used
Maintenance issues			Hoses and seals may be affected with higher-percent blends; lubricity is improved over that of conventional diesel fuel		Service requirements are expected to be reduced, since tune-ups, oil changes, timing belts, water pumps, radiators, and fuel injectors are not required	Special lubricants may be required. Practices are very similar, if not identical to those for conventionally fueled operations	N/A	High-pressure tanks required periodic inspection and certification	Some fleets report services lives that are 2–3 years longer, as well as extended intervals between required maintenance	Special lubricants must be used as directed by the supplier and M85-compatible replacement of parts must be used

(continued)

TABLE 8.20 (*Continued*)

	Gasoline	No. 2 Diesel	Biodiesel	CNG	Electricity	Ethanol (E85)	Hydrogen	LNG	Liquefied Petroleum Gas (LPG)	Methanol (M85)
Safety	Gasoline is a relatively safe fuel since people have learned to use it safely. Gasoline is not biodegradable, however, so a spill could pollute soil and water	Diesel is a relatively safe fuel since people have learned to use it safely. Diesel is not biodegradable though, so a spill could pollute soil and water	Less toxic and more biodegradable than conventional fuel; can be transported, delivered, and stored using the same equipment as for diesel fuel	Pressurized tanks have been designed to withstand a severe impact, high external temperatures, and automotive and environmental exposure	Meet all the same vehicle safety standards as conventional vehicles	Ethanol can form an explosive vapor in fuel tanks. In accidents, however, ethanol is less dangerous than gasoline because its low evaporation speed keeps alcohol concentration in the air low and nonexplosive	Hydrogen is extremely explosive and acceptable systems for widespread distribution and storage for mass-produced vehicles are yet to be developed	Cryogenic fuels require special handling procedures and equipment to properly store and dispense	Adequate ventilation is important for fueling LPG vehicles due to increased flammability of LPG. LPG tanks are 20 times more puncture resistant than gasoline tanks and can withstand high impact	Methanol can form an explosive vapor in fuel tanks. In accidents, however, methanol is less dangerous than gasoline because its low evaporation speed keeps alcohol concentration in the air low and nonexplosive

Source: Adapted from U.S. Department of Energy/Energy Efficiency and Renewable Energy, Alternative Fuels Comparison Chart, 2005a. http://www.eere.energy.gov/afdc/pdfs/afv_info.pdf suggested that a hydrogen tank storing gaseous hydrogen at 5000 psi would need to be 175 gallons in volume to contain as much energy as a 15-gallon gasoline tank, as shown in Table 8.19 (Greene 2004). Hydrogen embrittles steel and requires special alloys for the storage system; it also leaks very easily, causing yet another obstacle for researchers studying storage systems.

These characteristics influence the required hydrogen distribution infrastructure that currently consists of 15 fueling stations in the U.S. There is also the question of whether to produce hydrogen at a central location and then distribute it, or to produce it at decentralized distribution areas. The cost of a decentralized hydrogen refueling station is $600,000; the cost of an ethanol refueling station that serves several times that number of vehicles is $50,000 (Morris 2003). Moreover, it is estimated that for hydrogen fuel cell cars to gain 10% market penetration, 80% of the existing conventional refueling stations would need to be retrofitted for hydrogen distribution; providing a 90/10 gasoline/ethanol blend (that would not require any station modifications) would achieve almost the same amount of petroleum displacement. It has therefore been suggested that "if hydrogen and fuel cells do prove to be a cost-effective alternative, expanding the use of alcohols (i.e., ethanol) in our engines could become a stepping-stone to using hydrogen derived from those alcohols" (Miller 2003; Morris 2003).

Liquified petroleum gas (LPG) is a by-product of petroleum refining and NG processing and is the most widely used alternative fuel in the U.S. with about 3400 refueling stations. It has lower carbon monoxide and hydrocarbon tailpipe emissions than gasoline while maintaining about 70% of the energy content. However, almost half of LPG comes from oil; consequently, in terms of feedstocks, it does not offer much of an advantage over conventional gasoline.

Methanol (M85) is not a widely used transportation fuel, but like ethanol it can be blended with gasoline (85% methanol/15% gasoline) for use in heavy-duty buses. Feedstocks for methanol include coal, wood, methane, and NG. It is a liquid at ambient temperatures but has a low energy content compared to gasoline. Its corrosiveness and toxicity require special materials for the onboard storage system and the distribution infrastructure. Although it reduces particulate, hydrocarbon, and benzene tailpipe emissions, it increases formaldehyde emissions (Table 8.20).

Acknowledgments

The author would like to extend her sincere gratitude to Dr. Frank Kreith of the American Society of Mechanical Engineers and Dr. Lester A. Hoel of the University of Virginia Center for Transportation Studies for their support during the writing of this chapter.

References

Baard, M. 2003. Hydrogen is no gas, yet. *Wired*. (June 23). http://www.wired.com/news/technology/0,1282,59322,00.html?tw=wn_story_related (accessed June 25, 2005).

Bosch 1996. *Automotive Handbook* Robert Bosch GmbH, Stuttgart, Germany.

Calthorpe, P. and Fulton, W. 2001. *The Regional City: Planning for the End of Sprawl*. Island Press, Washington, DC.

Carey, J. 2005. Giving hybrids a real jolt. *Business Week*, April 11, 70–72.

Cervero, R. 2001. Transport and land use: Key issues in metropolitan planning and smart growth. *Australian Planner*, 38, 1, 29–37.

Cervero, R. and Gorham, R. 1995. Commuting in transit versus automobile neighborhoods. *Journal of the American Planning Association*, 61, 2, 210–224.

Danchenko, D. 2005. *Public Transportation Fact Book* American Public Transportation Association, Washington, DCwww.apta.com/research/stats/factbook/index.cfm (accessed May 29, 2005).

Davis, S. and Diegel, S. 2004. *Transportation Energy Data Book*. Oak Ridge National Laboratory/U.S. Department of Energy, Oak Ridge, TN.

DeGaspari, J. 2005. A new dawn for diesel. *Mechanical Engineering*, January, 26–31.

Downs, A. 2001. What does 'smart growth' really mean? *Planning*, April, 20–25.

Electric Power Research Institute (EPRI) 2004. *Advanced Batteries for Electric-drive Vehicles: A Technology and Cost-effectiveness Assessment for Battery Electric Behicles, Power Assist Hybrid Vehicles, and Plug-in Hybrid Electric Vehicles* Publication 1009299 Electric Power Research Institute, Palo Alto, CA.

Ewing, R. 1995. Beyond density, mode choice, and single-purpose trips. *Transportation Quarterly*, 49, 4, 15–24.

Farris, T. 2001. The barriers to using urban infill development to achieve smart growth. *Housing Policy Debate*, 12, 1–45.

Fishman, R. 1999. The American metropolis at century's end: Past and future influences. *Housing Facts & Findings*, 1, 4.

Gannett News Service. 2005. How a hybrid car works. (February 6). http://www.clarionledger.com/apps/pbcs.dll/article?AID=/20050206/BIZ/50206007 (accessed April 30, 2005).

Gartner, J. 2005a. Biomass adds to ethanol debate. *Wired*, (June 2). http://www.wired.com/news/planet/0,2782,67691,00.html (accessed June 25, 2005).

Gartner, J. 2005b. Ethanol grows as gas alternative. *Wired*, (May 4). http://www.wired.com/news/planet/0,2782,67416,00.html (accessed June 25, 2005).

Greene, D. L., 2004. Transportation and energy, in *The Geography of Urban Transportation*, S Hanson, ed., pp. 274–293. New York: Guilford Press.

Greene, D. and Schafer, A. 2003. *Reducing Greenhouse Gas Emissions from US Transportation*. Pew Center on Global Climate Change, Arlington, VA.

Greene, D., Duleep, K. G., and McManus, W. 2004. *Future Potential of Hybrid and Diesel Powertrains in the US Light-duty Vehicle Market*. Oak Ridge National Laboratory, United States Department of Energy, Oak Ridge, TN.

Greene, D., Hopson, J., and Li, J. 2004. Running out of and into oil: Analyzing global oil depletion and transition through 2050. *Proceedings of Transportation Research Board 83rd Annual Meeting*, Washington, DC.

Hadder, G. R. 2000. *Ethanol Demand in United States Regional Production of Oxygenate-limited Gasoline*. Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, TN.

Hall, R. 2005. Walkable thoroughfares through balanced design. *Presentation at The Nuts & Bolts of Traditional Neighborhood Development Conference*, Richmond, VA.

JD Power LMC. 2003. News release September 29: JD Power-LMC reports: Annual Global Diesel Light-vehicle Sales to reach 27 million by 2015. http://www.prnewswire.co.uk/cgi/news/release?id=109015 (accessed February 27, 2005).

Klett, D. E. and Afify, E. M., 2005. Reprinted with permission from internal combustion engines, In *The CRC Handbook of Mechanical Engineering*, F. Kreith and D.Y. Goswami, eds., Boca Raton, FL: Taylor & Francis.

Kolke, R. 1999. *Technical Options for Abating Road Traffic Impacts: Comparative Study of Fuel Cell Vehicles and Vehicles with Internal Combustion Engines. Translated by S. Smith*. Umwelt Bundes Amt (Federal Environmental Agency), Berlin, Germany.

Kreith, F. and West, R. E. 2004. Fallacies of a hydrogen economy: A critical analysis of hydrogen production and utilization. *Journal of Energy Resources Technology*, 126, 12, 249–256.

Kreith, F. and West, R. E. 2005. Personal communication.

Kreith, F., Potestio, D. S., and Kimbell, C. 1999. *Ground Transportation for the 21st Century*. National Conference of State Legislatures, Denver, CO.

Kreith, F., West, R. E., and Isler, B. 2001. Efficiency of advanced transportation technologies. *Proceedings of 36th Intersociety Energy Conversion Engineering Conference,* IECEC2001-EI-01, Savannah, GA.

Kreith, F., West, R. E., and Isler, B. 2002a. Legislative and technical perspectives for advanced ground transportation systems. *Transportation Quarterly*, 56, 1, 51–73.

Kreith, F., West, R. E., and Isler, B. 2002b. Efficiency of advanced transportation technologies. *Journal of Energy Resources Technology*, 124, September, 173–179.

Lawrence Berkeley National Laboratory (LBNL), Environmental Energy Technologies Division (EETD). 2004. An update on the EETD research program on batteries for advanced transportation technologies. *EETD*, 5, 2.

Lichty, L. and MacCoull, N., 1967. Internal-combustion engines, in *Marks' Standard Handbook for Mechnaical Engineers*, T Baumeister, ed., pp. 9-103–9-149.

Miller, K. 2003. *Environment and Energy Daily*, (April 7). As cited in Morris 2003.

Miller, J. S. and Hoel, L. A. 2002. The "smart growth" debate: Best practices for urban transportation planning, 1–24. Reprinted from *Socio-Economic Planning Sciences*, 36, with permission from Elsevier.

Morris, D. 2003. *A Better Way to Get from Here to There: A Commentary on the Hydrogen Economy and a Proposal for an Alternative Strategy*. The Institute for Local Self-Reliance, Minneapolis, MN.

National Research Council 2004. *The Hydrogen Economy: Opportunities, Costs, Barriers, and R&D Needs* National Academies Press, Washington, DC.

Office of Technology Assessment (OTA) 1995. *Advanced Automotive Technology: Visions of a Super-efficient Family Car* OTA-ETI-638, Office of Technology Assessment, Congress of the U.S., Washington, DC.

Porretto, J. 2005. *JD Power: Hybrids to Top out at 3 Percent of US Market*. Associated Press, (February 3). http://www.detnews.com/2005/autosinsider/0502/03/01-79067 (accessed February 27, 2005).

Sharke, P. 2005. Idle hour. *Mechanical Engineering*, January, 32–34.

Stodolsky, F., Gaines, L., and Vyas, A. 2000. *Analysis of Technology Options to Reduce the Fuel Consumption of Idling Trucks*. Argonne National Laboratory, Argonne, IL.

Toyota Motor Corporation. 2005a. Vehicle comparison. http://www.toyota.com/toyotacomparator/displayComparator.do?toyotaModelCode=prius (accessed May 25, 2005).

Toyota Motor Corporation. 2005b. Hybrid synergy drive: Prius demo. http://www.toyota.com (accessed December 18, 2005).

U.S. Advanced Battery Consortium (USABC). 1999. USABC goals for advanced batteries for EVs. http://www.uscar.org/consortia&teams/consortiahomepages/con-usabc.htm (accessed June 10, 2005).

U.S. Department of Energy (DOE)/Energy Efficiency and Renewable Energy (EERE). 2005a. Alternative fuels comparison chart. http://www.eere.energy.gov/afdc/pdfs/afv_info.pdf (accessed June 18, 2005).

U.S. Department of Energy (DOE)/Energy Efficiency and Renewable Energy (EERE). 2005b. Direct-injection stratified charge (DISC) engines. http://www.fueleconomy.gov/feg/DISC_tech.shtml (accessed June 12, 2005).

U.S. Department of Energy (DOE)/Energy Efficiency and Renewable Energy (EERE). 2005c. Electric vehicle batteries fact sheet. http://www.eere.energy.gov/consumerinfo/factsheets/fa1.html (accessed April 30, 2005).

U.S. Department of Energy (DOE)/Energy Efficiency and Renewable Energy (EERE). 2005d. Fuel cell vehicles. http://www.fueleconomy.gov/feg/fuelcell.shtml (accessed June 10, 2005).

U.S. Department of Energy (DOE)/Energy Efficiency and Renewable Energy (EERE). 2005e. Gas mileage tips. http://www.fueleconomy.gov/feg/drive.shtml (accessed June 10, 2005).

U.S. Department of Energy (DOE)/Energy Information Administration (EIA). 2004. Coal reserves. http://www.eia.doe.gov/neic/infosheets/coalreserves.htm (accessed June 18, 2005).

U.S. Department of Energy (DOE)/Energy Information Administration (EIA). 2005. Annual energy outlook 2005. Report No. DOE/EIA-0383(2005).

U.S. Department of Transportation (DOT)/Bureau of Transportation Statistics (BTS). 2004. 2002 Commodity flow survey, United States. http://www.bts.gov/publications/commodity_flow_survey/2002/united_states_final/ (accessed June 26, 2005).

U.S. Department of Transportation (DOT)/Bureau of Transportation Statistics (BTS). 2005a. Freight. http://www.bts.gov/programs/freight_transportation/html/more_freight.html (accessed June 25, 2005).

U.S. Department of Transportation (DOT)/Bureau of Transportation Statistics (BTS) 2005b. *National Transportation Statistics 2004* U.S. Government Printing Office, Washington, DC.

U.S. General Accounting Office (GAO). 1999. *Community Development: Extent of Federal Influence on "Urban Sprawl" is Unclear*. GAO-RCED-99-87, Washington, DC.

U.S. General Accounting Office (GAO). 2001. *Mass Transit: Bus Rapid Transit Shows Promise*. GAO-01-984, Washington, DC.

Vaitheeswaran, V. 2005. Interview by Robert Siegel: Bush's energy proposals echo earlier plan. *All Things Considered*. National Public Radio (27 April 2005).

Victoria Transport Policy Institute (VTPI). 2005. Transit oriented development. In *Transportation Demand Management Encyclopedia*. http://www.vtpi.org/tdm/tdm45.htm (accessed May 28, 2005).

Vuchic, V. 1999. *Transportation for Livable Cities*. Center for Urban Policy Research, New Brunswick, NJ.

Weiss, M. A., Heywood, J. B., Schafer, A., and Natarajan, V. K. 2003. *Comparative Assessment of Fuel Cell Cars*. MIT LFEE 2003-001RP. Massachusetts Institute of Technology, Cambridge, MA.

Woodyard, C. 2005. Device offers at-home natural-gas fill-ups. *USA Today*, April 22, 5B.

9

Infrastructure Risk Analysis and Security

Bilal M. Ayyub
University of Maryland

Massoud Amin
University Minnesota

9.1 Infrastructure Risk Analysis and Management

Bilal M. Ayyub

9.1.1 Introduction

Risk is associated with all projects and business ventures taken by individuals and organizations regardless of their sizes, their natures, and their time and place of execution and utilization. Risk is present in various forms and levels even in small domestic projects such as adding a deck to a residential house, and in large multibillion-dollar projects such as developing and producing a space shuttle. These risks could result in significant budget overruns, delivery delays, failures, financial losses, environmental damages, and even injuries and loss of life. Risks are taken even though they could lead to devastating consequences because of potential benefits, rewards, survival, and future return on investment. The chapter defines and discusses risk and its dimensions, risk analysis, risk management and control, and risk communication.

9.1.2 Risk Terminology

Definitions that are needed for presenting risk-based technology methods and analytical tools are presented in this section.

9.1.2.1 Hazards

A hazard is an act or phenomenon posing potential harm to some person(s) or thing(s), i.e., a source of harm, and its potential consequences. For example, uncontrolled fire is a hazard, water can be a hazard, and strong wind is a hazard. For the hazard to cause harm, it needs to interact with person(s) or thing(s) in a harmful manner. The magnitude of the hazard is the amount of harm that might result, including the seriousness and the exposure levels of people and the environment. Hazards need to be identified and considered in projects' life cycle analyses because they could pose threats and could lead to project failures.

The interaction between a person (or a system) and a hazard can be voluntary or involuntary. For example, exposing a marine vessel to a sea environment might lead to its interaction with extreme waves in an uncontrollable manner, i.e., an involuntary manner. Although the decision of a navigator of the vessel to go through a storm system that is developing can be viewed as a voluntary act in nature, and might be needed to meet schedule constraints or other constraints, the potential rewards of delivery of shipment or avoidance of delay charges offer an incentive that warrants such an interaction. Other examples can be constructed where individuals interact with hazards for potential financial rewards, fame, and self-fulfillment and satisfaction ranging from investment undertaking to climbing cliffs.

9.1.2.2 Reliability

Reliability can be defined for a system or a component as its ability to fulfill its design functions under designated operating or environmental conditions for a specified time period. This ability is commonly measured using probabilities. Reliability is, therefore, the occurrence probability of the complementary event to failure as provided in the following expression:

$$\text{Reliability} = 1 - \textit{Failure Probability} \tag{9.1}$$

9.1.2.3 Event Consequences

For an event of failure, *consequences* can be defined as the degree of damage or loss from some failure. Each failure of a system has some consequence(s). A failure could cause economic damage, environmental damage, injury, loss of human life, or other possible events. Consequences need to be quantified in terms of failure consequence severities using relative or absolute measures for various consequence types to facilitate risk analysis.

For an event of success, consequences can be defined as the degree of reward or return or benefits from success. Such an event could cause economic outcomes, environmental effects, or other possible events. Consequences need to be quantified using relative or absolute measures for various consequence types to facilitate risk analysis.

9.1.2.4 Risk

The concept of risk can be linked to uncertainties associated with events. Within the context of projects, risk is commonly associated with an uncertain event or condition that, if it occurs, has a positive or a negative effect on a project's objectives.

Risk originates from the Latin term *risicum*, meaning the challenge presented by a barrier reef to a sailor. The Oxford dictionary defines risk as the chance of hazard, bad consequence, loss, etc. Also, risk is the chance of a negative outcome. To measure risk, one must accordingly assess both of its defining components, the chance, its negativity, and potential rewards or benefits. Estimation of risk is usually based on the expected result of the conditional probability of the event occurring times the consequence of the event given that is has occurred.

A risk results from an event or sequence of events called a *scenario*. The event or scenario can be viewed as a cause and, if it occurs, result in consequences with severities. For example, an event or cause may be shortage of personnel needed to perform a task needed to produce a project. The event in this case of personnel shortage for the task will lead to a consequence on the project cost, schedule, and/or quality. The events can reside in the project environment that may contribute to project success or failure, such as project management practices, or external partners or subcontractors.

Risk has certain characteristics that should be used in the risk assessment process. Risk is a characteristic of an uncertain future, and is neither a characteristic of the present nor the past. After uncertainties are resolved and/or the future is attained, the risk becomes nonexistent. Therefore, risks cannot be described for historical events or for events that are currently being realized. Similarly, risks cannot be directly associated with a success. Although risk management through risk mitigation of selected events could result in project success leading to rewards and benefits, these rewards and benefits cannot be considered as outcomes only of the nonoccurrence of these events associated with the risks. The occurrence of risk events leads to adverse consequences that are clearly associated with their occurrence; however, their nonoccurrences are partial contributors to the project success that lead to rewards and benefits. The credit in the form of rewards and benefits cannot be given solely to the nonoccurrence of these risk events. Some risk assessment literature defines risk to include both potential losses and rewards. They need to be treated separately as (1) risks leading to adverse consequences, and (2) risks that contribute to benefits or rewards in trade-off analyses. An appropriate risk definition is this context is a threat (or opportunity) that could affect adversely (or favorably) achievement of the objectives of a project and its outcomes.

Developing an economic, analytical framework for a decision situation involving risks requires examining the economic and finance environments of a project. This environment could have significant impacts on the occurrence probabilities of events associated with risks. This complexity might be needed for certain projects in order to obtain justifiable and realistic results. The role of such an environment in risk analysis is discussed in subsequent sections.

Formally, risk can be defined as the potential of losses and rewards resulting from an exposure to a hazard or as a result of a risk event. Risk should be based on identified risk events or event scenarios. Risk can be viewed to be a multidimensional quantity that includes event occurrence probability, event occurrence consequences, consequence significance, and the population at risk; however, it is commonly measured as a pair of the probability of occurrence of an event, and the outcomes or consequences associated with the event's occurrence. This pairing can be represented by the following equation:

$$Risk \equiv [(p_1, c_1), (p_2, c_2), \ldots, (p_\mathrm{i}, c_\mathrm{i}), \ldots, (p_\mathrm{n}, c_\mathrm{n})], \tag{9.2}$$

where p_i is the occurrence probability of an outcome or event i out of n possible events, and c_i is the occurrence consequences or outcomes of the event. A generalized definition of risk can be expressed as

$$Risk \equiv [(l_1, o_1, u_1, cs_1, po_1), (l_2, o_2, u_2, cs_2, po_2), \ldots, (l_n, o_n, u_n cs_n po_n)], \tag{9.3}$$

where l is likelihood, o is outcome, u is utility (or significance), cs is causal scenario, po is population affected by the outcome, and n is the number of outcomes. The definition according to Equation 9.3 covers all attributes measured in risk assessment that are described in this chapter, and offers a complete description of risk, from the causing event to the affected population and consequences. The population size effect should be considered in risk studies because society responds differently for risks associated with a large population in comparison to a small population. For example, a fatality rate of 1 in 100,000 per event for an affected population of 10 results in an expected fatality of 10^{-4} per event, whereas the same fatality rate per event for an affected population of 10,000,000 results in an expected fatality of 100 per event. Although the impact of the two scenarios might be the same on the society (same risk value), the total number of fatalities per event/accident is a factor in risk acceptance. Plane travel may be "safer" than for example recreational boating, but 200–300 injuries per accident are less acceptable to society. Therefore, the size of the population at risk and the number of fatalities per event should be considered as factors in setting acceptable risk.

Risk is commonly evaluated as the product of likelihood of occurrence and the impact severity of occurrence of the event:

$$RISK\left(\frac{Consequence}{Time}\right) = LIKELIHOOD\left(\frac{Event}{Time}\right) \times IMPACT\left(\frac{Consequence}{Event}\right) \tag{9.4}$$

In Equation 9.4, the likelihood can also be expressed as a probability. Equation 9.4 presents risk as an expected value of loss or an average loss. A plot of occurrence probabilities and consequences is called a *risk profile* or a *Farmer curve*. An example Farmer curve is given in Figure 9.1 based on a nuclear case

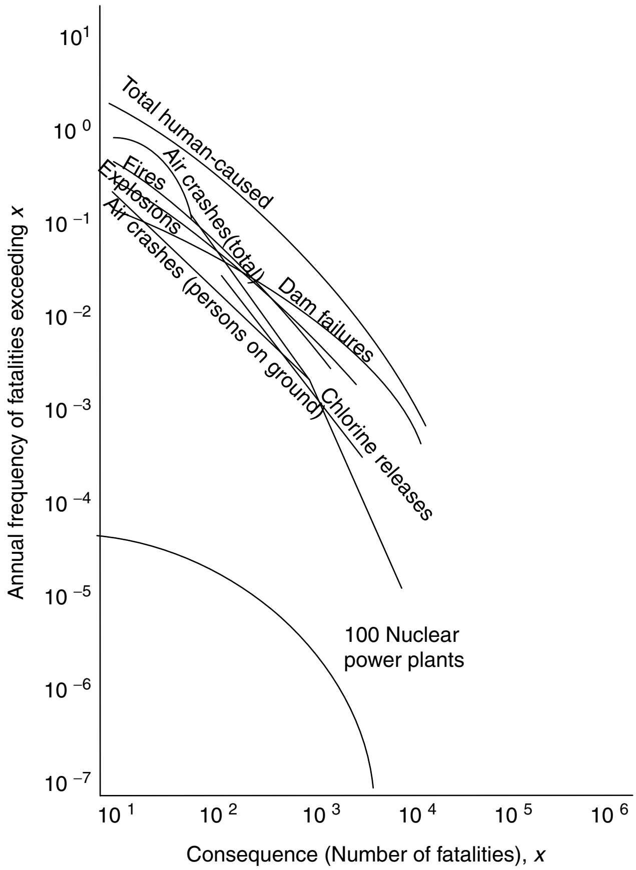


FIGURE 9.1 Example risk profile.

study, provided herein for illustration purposes. It should be noted that the abscissa provides the number of fatalities, and the ordinate provides the annual frequency of exceedence for the corresponding number of fatalities. These curves are sometimes constructed using probabilities instead of frequencies. The curves represent or median average values. Sometimes, bands or ranges are provided to represent uncertainty in these curves. They represent confidence intervals for the average curve or for the risk curve. Figure 9.2 shows examples curves with uncertainty bands. This uncertainty is sometimes called *meta-uncertainty*. A complete treatment of uncertainty analysis is provided by Ayyub and Klir (2006).

The occurrence probability (p) of an outcome (o) can be decomposed into an occurrence probability of an event or threat (t), and the outcome occurrence probability given the occurrence of the event ($o|t$). The occurrence probability of an outcome can be expressed as follows using conditional probability concepts:

$$p(o) = p(t)p(o|t). \tag{9.5}$$

In this context, threat is defined as a hazard or the capability and intention of an adversary to undertake actions that are detrimental to a system or an organization's interest. In this case, threat is a function of only the adversary or competitor, and usually cannot be controlled by the owner or user of the system. However, the adversary's intention to exploit his capability may be encouraged by vulnerability of the system or discouraged by an owner's countermeasures. The probability ($p(o|t)$) can be interpreted as the vulnerability of the system in case of this threat occurrence. Vulnerability is

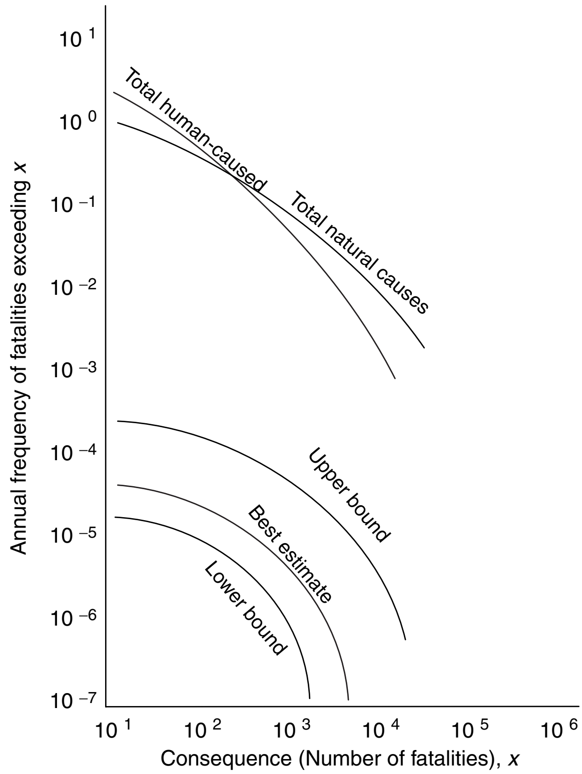


FIGURE 9.2 Uncertain risk profile.

a result of any weakness in the system or countermeasure that can be exploited by an adversary or competitor to cause damage to the system.

9.1.2.5 Performance

The performance of a system or component can be defined as its ability to meet functional requirements. The performance of an item can be described by various elements, including such items as speed, power, reliability, capability, efficiency, and maintainability. The design and operation of the product or system influence performance.

9.1.2.6 Risk-Based Technology

Risk-based technologies (RBT) are methods or tools and processes used to assess and manage the risks of a component or system. RBT methods can be classified into risk management that includes risk assessment/risk analysis and risk control using failure prevention and consequence mitigation, and risk communication as shown in Figure 9.3.

Risk assessment consists of hazard identification, event-probability assessment, and consequence assessment. Risk control requires the definition of acceptable risk and comparative evaluation of options and/or alternatives through monitoring and decision analysis. Risk control also includes failure prevention and consequence mitigation. Risk communication involves perceptions of risk, which depends on the audience targeted, hence, classified into risk communication to the media, the public, and to the engineering community.

9.1.2.7 Safety

Safety can be defined as the judgment of risk acceptability for the system. Safety is a relative term since the decision of risk acceptance may vary depending on the individual making the judgment. Different people

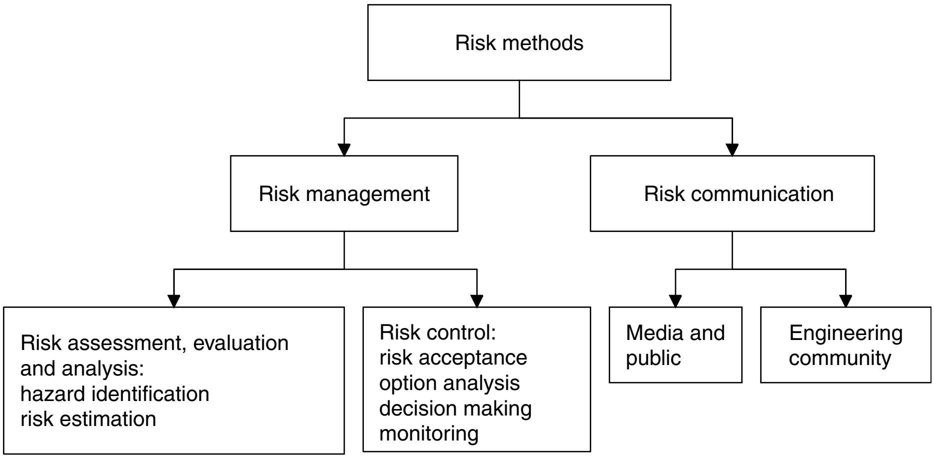


FIGURE 9.3 Risk-based technology methods.

are willing to accept different risks as demonstrated by different factors such as location, method or system type, occupation, and lifestyle. The selection of these different activities demonstrates an individual's safety preference despite a wide range of risk values. Table 9.1 identifies varying annual risks for different activities based on typical exposure times for these activities. Also Figure 9.4 from the Imperial Chemical Industries, Ltd. shows the variation of risk exposure during a typical day that starts by waking up in the morning from sleep and getting ready to go to work, then commuting and working during morning hours, followed by a lunch break, then additional work hours followed by commuting back to having dinner, and round-trip on a motorcycles to a local pub. The ordinate in this figure is the fatal accident frequency rate (FAFR) with a FAFR of 1.0 corresponding to one fatality in 11,415 years, or 87.6 fatalities per one million years. The figure is based on an average number of deaths in 10^8 h of exposure to a particular activity.

Risk perceptions of safety may not reflect the actual level of risk in some activity. Table 9.2 shows the differences in risk perception by three groups of the league of women voters, college students, and experts of 29 risk items. Only the top items are listed in the table. Risk associated with nuclear power was ranked

TABLE 9.1 Relative Risk of Different Activities

Risk of Death	Occupation	Lifestyle	Accidents/Recreation	Environmental Risk
1 in 100	Stunt person			
1 in 1,000	Race car driver	Smoking (one pack/day)	Skydiving rock climbing snowmobile	
	Firefighter miner		Canoeing automobile	
1 in 10,000	Farmer police officer	Heavy drinking	All home accidents frequent air travel	
1 in 100,000	Truck driver engineer banker insurance agent	Using contraceptive pills light drinking	Skiing home fire	Substance in drinking water living downstream of a dam
1 in 1,000,000		Diagnostic x-rays smallpox vaccination (per occasion)	Fishing poisoning occasional air travel (one flight per year)	Natural background radiation living at the boundary of a nuclear power
1 in 10,000,000		Eating charcoal-broiled steak (once a week)		Hurricane tornado lightning animal bite or insect sting

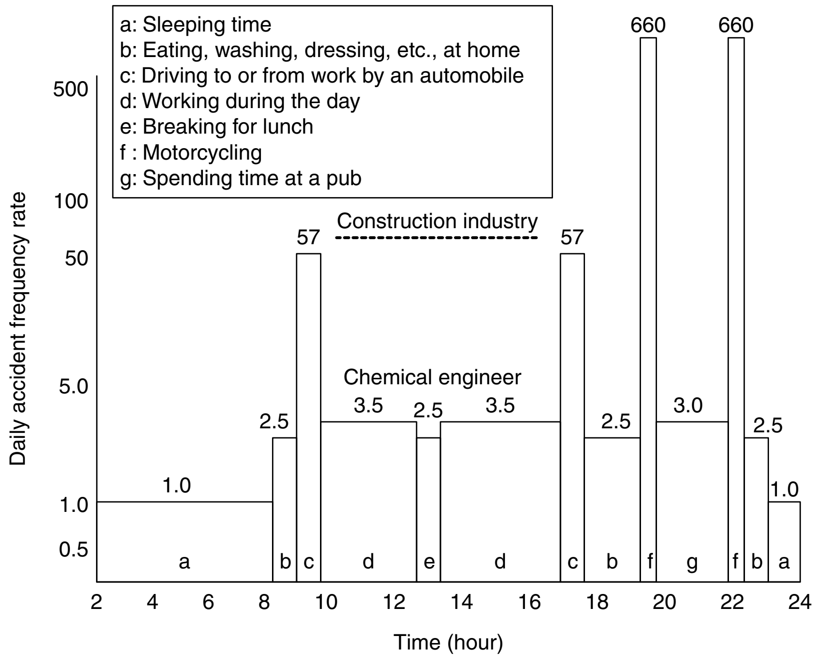


FIGURE 9.4 Daily death risk exposure for a working healthy adult.

as the highest type by women voters and college students, whereas it was placed as the 20th by experts. Experts place motor vehicles as the first risk. Public perception of risk and safety varies by age, gender, education, attitudes, and culture among other factors. Individuals sometimes do not recognize uncertainties associated with risk event or activity causing an unwarranted confidence in an individual's perception of risk or safety. Rare causes of death are often overestimated and common causes of death are often underestimated. Perceived risk is often biased by the familiarity of the hazard. The significance or the impact of safety perceptions stems from that decisions are often made on subjective judgments. If the judgments hold misconceptions about reality, this bias affects the decision. For example, the choice of a transportation mode—train, automobile, motorcycle, bus, bicycle, etc.—results in a decision based on many criteria including such items as cost, speed, convenience, and safety. The weight and evaluation of the decision criteria in selecting a mode of transportation rely on the individual's perception of safety that may deviate sometimes significantly from the actual values of risks. Understanding these differences in risk and safety perceptions is vital to performing risk management decisions and risk communications as provided in subsequent sections on risk management and control.

9.1.3 Risk Assessment

9.1.3.1 Risk Assessment Methodologies

Risk studies require the use of analytical methods at the system level that considers subsystems and components in assessing their failure probabilities and consequences. Systematic, quantitative, qualitative or semiquantitative approaches for assessing the failure probabilities and consequences of engineering systems are used for this purpose. A systematic approach allows an analyst to evaluate expediently and easily complex systems for safety and risk under different operational and extreme conditions. The ability to quantitatively evaluate these systems helps cut the cost of unnecessary and often expensive redesign, repair, strengthening, or replacement of components, subsystems, and systems. The results of risk analysis can also be utilized in decision analysis methods that are based on cost–benefit tradeoffs.

TABLE 9.2 Risk Perception

Activity or Technology	League of Women Voters	College Students	Experts
Nuclear power		1	20
Motor vehicles		5	1
Hand guns		2	4
Smoking		3	2
Motorcycles		6	6
Alcoholic beverages		7	3
General aviation		15	12
Police work		8	17
Pesticides		4	8
Surgery		11	5
Firefighting		10	18
Large construction		14	13
Hunting		18	23
Spray cans		13	25
Mountain climbing		22	28
Bicycles		24	15
Commercial aviation		16	16
Electric (nonnuclear) power		19	9
Swimming		29	10
Contraceptives		9	11
Skiing		25	29
X-rays		17	7
High school or college sports		26	26
Railroads		23	19
Food preservatives		12	14
Food coloring		20	21
Power mowers		28	27
Prescription antibiotics		21	24
Home applications		27	22

Risk assessment is a technical and scientific process by which the risks of a given situation for a system are modeled and quantified. Risk assessment can require and/or provide both qualitative and quantitative data to decision makers for use in risk management.

Risk assessment or risk analysis provides the process for identifying hazards, event-probability assessment, and consequence assessment. The risk assessment process answers three basic questions: (1) What can go wrong? (2) What is the likelihood that it will go wrong? (3) What are the consequences if it does go wrong? Answering these questions requires the utilization of various risk methods as discussed in this chapter.

A risk assessment process should utilize experiences gathered from project personnel including managers, other similar projects and data sources, previous risk assessment models, experiences from other industries and experts, in conjunction with analysis and damage evaluation/prediction tools. A risk assessment process is commonly a part of a risk-based or risk-informed methodology that should be constructed as a synergistic combination of decision models, advanced probabilistic reliability analysis algorithms, failure consequence assessment methods, and conventional performance assessment methodologies that have been employed in related industry for performance evaluation and management. The methodology should realistically account for the various sources and types of uncertainty involved in the decision-making process (Ayyub and McCuen 2003; Ayyub and Klir 2006).

In this section, a typical overall methodology is provided in the form of a workflow or block diagram. The various components of the methodology are described in subsequent sections. Figure 9.5 provides an overall description of a methodology for risk-based management of structural systems for the purpose of demonstration. The methodology consists of the following primary steps:

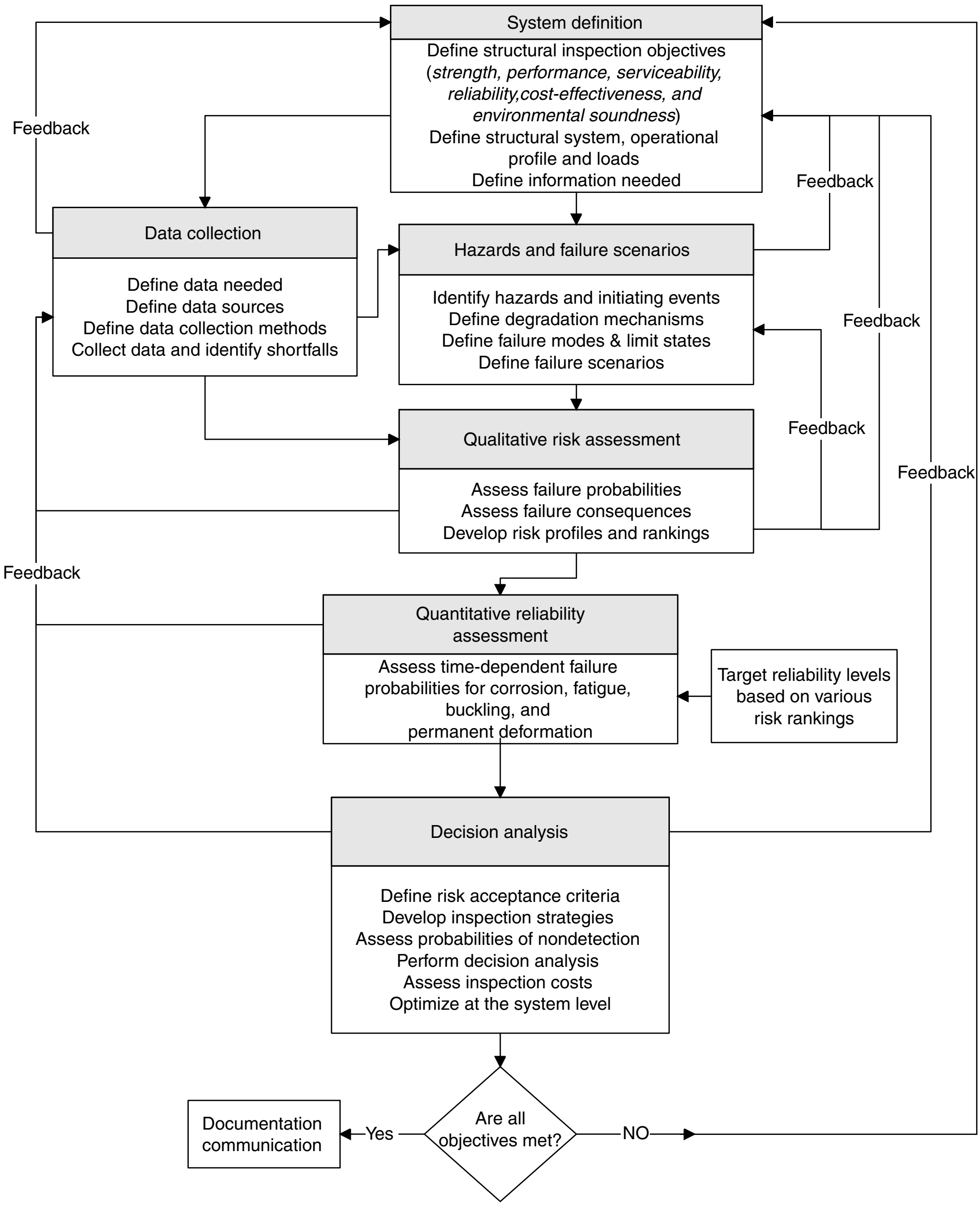


FIGURE 9.5 Methodology for risk-based life cycle management of structural systems.

1. Definition of analysis objectives and systems
2. Hazard analysis, definition of failure scenarios, and hazardous sources and their terms
3. Collection of data in a life cycle framework
4. Qualitative risk assessment
5. Quantitative risk assessment
6. Management of system integrity through failure prevention and consequence mitigation using risk-based decision making.

These steps are briefly described below with additional background materials provided in subsequent sections.

The first step of the methodology is to define the system. This definition should be based on a goal that is broken down into a set of analysis objectives. A system can be defined as an assemblage or combination of elements of various levels and/or details that act together for a specific purpose. Defining the system provides the risk-based methodology with the information it needs to achieve the analysis objectives. The system definition phase of the proposed methodology has four main activities. The activities are to

- Define the goal and objectives of the analysis
- Define the system boundaries
- Define the success criteria in terms of measurable performances
- Collect information for assessing failure likelihood
- Collect information for assessing failure consequences

For example, structural systems require a structural integrity goal that can include objectives stated in terms of strength, performance, serviceability, reliability, cost-effectiveness, and environmental soundness. The objectives can be broken down further to include other structural integrity attributes, such as alignment and water tightness in case of marine vessels. A system can be defined based on a stated set of objectives. The same system can be defined differently depending on these stated objectives. A marine vessel structural system can be considered to contain individual structural elements such as plates, stiffened panels, stiffeners, longitudinals, etc. These elements could be further separated into individual components and/or details. Identifying all of the elements, components and details allows an analysis team to collect the necessary operational, maintenance, and repair information throughout life cycle on each item so that failure rates, repair frequencies, and failure consequences can be estimated. The system definition might need to include nonstructural subsystems and components that would be affected in case of failure. The subsystems and components are needed to assess the consequences.

To understand failure and the consequences of failure, the states of success need to be defined. For the system to be successful, it must be able to perform its designed functions by meeting measurable performance requirements. But the system may be capable of various levels of performance, all of which might not be considered a successful performance. While a marine vessel may be able to get from point A to point B only at a reduced speed due to a fatigue failure that results in excessive vibration at the engine room, its performance would probably not be considered successful. The same concept can be applied to individual elements, components, and details. It is clear from this example that the vessel's success and failure impacts should be based on the overall vessel performance that can easily extend beyond the structural systems.

With the development of the definition of success, one can begin to assess the likelihood of occurrence and causes of failures. Most of the information required to develop an estimate of the likelihood of failure might exist in maintenance and operating histories available on the systems and equipment, and based on judgment and expert opinion. This information might not be readily accessible, and its extraction from its current source might be difficult. Also, assembling it in a manner that is suitable for the risk-based methodology might be a challenge.

Operation, maintenance, engineering, and corporate information on failure history needs to be collected and analyzed for the purpose of assessing the consequences of failures. The consequence information might not be available from the same sources as the information on the failure itself. Typically, there are documentations of repair costs, reinspection or recertification costs, lost person-hours of labor, and possibly even lost opportunity costs due to system failure. Much more difficult to find and assess are costs associated with the effects on other systems, the cost of shifting resources to cover lost production, and things like environmental, safety–loss or public relations costs. These may be attained through carefully organized discussions and interviews with cognizant personnel including the use of expert-opinion elicitation.

9.1.3.2 Risk Events and Scenarios

To adequately assess all risks associated with a project, the process of identification of risk events and scenarios is an important stage in risk assessment. Risk events and scenarios can be categorized as follows:

- Technical, technological, quality, or performance risks, such as unproven or complex technology, unrealistic performance goals, and changes to the technology used or to the industry standards during the project.
- Project management risks, such as poor allocation of time and resources, inadequate quality of the project plan, and poor use of project management disciplines.
- Organizational risks, such as cost, time, and scope objectives that are internally inconsistent, lack of prioritization of projects, inadequacy or interruption of funding, resource conflicts with other projects in the organization, errors by individuals or by an organization, and inadequate expertise and experience by project personnel.
- External risks, such as shifting legal or regulatory environment, labor issues, changing owner priorities, country risk, and weather.
- Natural hazards, such as earthquakes, floods, strong wind, and waves generally require disaster recovery actions in addition to risk management. Within these categories, several risk types can be identified.

9.1.3.3 Identification of Risk Events and Scenarios

The risk assessment process starts with the question “What can go wrong?“ The identification of what can go wrong entails defining hazards, risk events, and risk scenarios. The previous section provided categories of risk events and scenarios. Risk identification involves determining which risks might affect the project and documenting their characteristics. The risk identification generally requires the participation from a project team, risk management team, subject matter experts from other parts of the company, customers, end users, other project managers, stakeholders, and outside experts on as needed basis. Risk identification can be an iterative process. The first iteration may be performed by selected members of the project team, or by the risk management team. The entire project team and primary stakeholders may take a second iteration. To achieve an unbiased analysis, persons who are not involved in the project may perform the final iteration. Risk identification can be a difficult task, because it is often highly subjective, and there are no unerring procedures that may be used to identify risk events and scenarios other than relying heavily on the experience and insight of key project personnel.

The development of the scenarios for risk evaluation can be created deductively (e.g., fault tree) or inductively (e.g., failure mode and effect analysis (FMEA)) as provided in Table 9.3. The table shows methods of multiple uses including likelihood or frequency estimation expressed either deterministically or probabilistically. Also, they can be used to assess varying consequence categories including such items as: economic loss, loss of life, or injuries.

The risk identification process and risk assessment requires the utilization of these formal methods as shown in Table 9.3. These different methods contain similar approaches to answer the basic risk assessment questions; however, some techniques may be more appropriate than others for risk analysis depending on the situation.

9.1.3.4 Risk Breakdown Structure

Risk sources for a project can be organized and structured to provide a standard presentation that would facilitate understanding, communication and management. The previously presented methods can be viewed as simple linear lists of potential sources of risk, providing a set of headings under which risks can

TABLE 9.3 Risk Assessment Methods

Method	Scope
Safety/Review Audit	Identifies equipment conditions or operating procedures that could lead to a casualty or result in property damage or environmental impacts.
Checklist	Ensures that organizations are complying with standard practices.
What-If	Identifies hazards, hazardous situations, or specific accident events that could result in undesirable consequences.
Hazard and Operability Study (HAZOP)	Identifies system deviations and their causes that can lead to undesirable consequences and determine recommended actions to reduce the frequency and/or consequences of the deviations.
Preliminary Hazard Analysis (PrHA)	Identifies and prioritizes hazards leading to undesirable consequences early in the life of a system. It determines recommended actions to reduce the frequency and/or consequences of the prioritized hazards. This is an inductive modeling approach.
Probabilistic Risk Analysis (PRA)	Methodology for quantitative risk assessment developed by the nuclear engineering community for risk assessment. This comprehensive process may use a combination of risk assessment methods.
Failure Modes and Effects Analysis (FMEA)	Identifies the components (equipment) failure modes and the impacts on the surrounding components and the system. This is an inductive modeling approach.
Fault Tree Analysis (FTA)	Identifies combinations of equipment failures and human errors that can result in an accident. This is an deductive modeling approach.
Event Tree Analysis (ETA)	Identifies various sequences of events, both failures and successes that can lead to an accident. This is an inductive modeling approach.
The Delphi Technique	Assists to reach consensus of experts on a subject such as project risk while maintaining anonymity by soliciting ideas about the important project risks that are collected and circulated to the experts for further comment. Consensus on the main project risks may be reached in a few rounds of this process.
Interviewing	Identifies risk events by interviews of experienced project managers or subject matter experts. The interviewees identify risk events based on experience and project information.
Experience-Based Identification	Identifies risk events based on experience including implicit assumptions.
Brainstorming	Identifies risk events using facilitated sessions with stakeholders, project team members, and infrastructure support staff.

be arranged. These lists are sometimes called risk taxonomy. A simple list of risk sources might not provide the richness needed for some decision situations since it only presents a single level of organization. Some applications might require a full hierarchical approach to define the risk sources, with as many levels as are required to provide the necessary understanding of risk exposure. Defining risk sources in such a hierarchical structure is called a risk breakdown structure (RBS). The RBS is defined as a source-oriented grouping of project risks organized to define the total risk exposure of a project of interest. Each descending level represents an increasingly detailed definition of risk sources for the project. The value of the RBS can be in aiding an analyst to understand the risks faced by the project.

An example RBS is provided in Table 9.4. In this example, four risk levels are defined as shown in the table. The project's risks are viewed as level 0. Three types of level 1 risks are provided in the table for the purpose of demonstration. The number of risk sources in each level varies and depends on the application at hand. The subsequent level 2 risks are provided in groups that are detailed further in level 3. The RBS provides a means to systematically and completely identify all relevant risk sources for a project.

The risk breakdown structure should not be treated as a list of independent risk sources since commonly they have interrelations and common risk drivers. Identifying causes behind the risk sources is a key step towards an effective risk management plan including mitigation actions. A process of risk interrelation assessment and root-cause identification can be utilized to potentially lead to identifying credible scenarios that could lead to snowball effects for risk management purposes.

TABLE 9.4 Risk Breakdown Structure for a Project

Level 0	Level 1	Level 2	Level 3
			History, experiences, culture, personnel
		Corporate	Organization structure, stability, communication
			Finances conditions
			Other projects
			M
	Management		History, experiences, culture, personnel
			Contracts and agreements
		Customers &	Requirement definition
		Stakeholders	Finances and credit
			M
			Physical environment
		Natural	Facilities, site, equipment, materials
		environment	Local services
			M
			Political
			Legal, regulatory
Project Risks	External	Cultural	Interest groups
			Society and communities
			M
		Economic	Labor market, conditions, competition
			Financial markets
			M
	Technology	Requirements	Scope and objectives
			Conditions of use, users
			Complexity
			M
		Performance	Technology maturity
			Technology limitations
			New technologies
			New hazards or threats
			M
		Application	Organizational experience
			Personnel skill sets & experience
			Physical resources
			M

9.1.3.5 System Definition for Risk Assessment

Defining the system is an important first step in performing a risk assessment. A system can be defined as a deterministic entity comprising an interacting collection of discrete elements and commonly defined using deterministic models.

The word “deterministic” implies that the system is identifiable and not uncertain in its architecture. The definition of the system is based on analyzing its functional and/or performance requirements. A description of a system may be a combination of functional and physical elements. Usually, functional descriptions are used to identify high information levels on a system. A system may be divided into subsystems that interact. Additional detail leads to a description of the physical elements, components, and various aspects of the system.

The examination of a system needs to be made in a well-organized and repeatable fashion so that risk analysis can be consistently performed, therefore insuring that important elements of a system are defined and extraneous information is omitted. The formation of system boundaries is based upon the objectives of the risk analysis.

The establishment of system boundaries can assist in developing the system definition. The decision on what the system boundary is partially based on what aspects of the system’s performance are of concern.

The selection of items to include within the external boundary region is also reliant on the goal of the analysis. Beyond the established system boundary is the external environment of the system.

Boundaries beyond the physical/functional system can also be established. For example, time may also be a boundary since an overall system model may change, as a product is further along in its life cycle. The life cycle of a system is important because some potential hazards can change throughout the life cycle. For example, material failure due to corrosion or fatigue may not be a problem early in the life of a system; however, this may be an important concern later in the life cycle of the system.

Along with identifying the boundaries, it is also important to establish a resolution limit for the system. The selected resolution is important because it limits the detail of the analysis. Providing too little detail might not provide enough information for the problem. Too much information may make the analysis more difficult and costly due to the added complexity. The depth of the system model needs to be sufficient for the specific problem. Resolution is also limited by the feasibility of determining the required information for the specific problem. For failure analysis, the resolution should be to the components level where failure data are available. Further resolution is not necessary and would only complicate the analysis.

The system breakdown structure is the top-down division of a system into subsystems and components. This architecture provides internal boundaries for the system. Often the systems/subsystems are identified as functional requirements that eventually lead to the component level of detail. The functional level of a system identifies the function(s) that must be performed for the operation of the system. Further decomposition of the system into "discrete elements" leads to the physical level of a system definition identifying the hardware within the system. By organizing a system hierarchy using a top-down approach rather than fragmentation of specific systems, a rational, repeatable, and systematic approach to risk analysis can be achieved.

Further system analysis detail is addressed from modeling the system using some of the risk assessment methods described in Table 9.3. These techniques develop processes that can assist in decision making about the system. The logic of modeling based on the interaction of a system's components can be divided into induction and deduction. This difference in the technique of modeling and decision making is significant. Induction logic provides the reasoning of a general conclusion from individual cases. This logic is used when analyzing the effect of a fault or condition on a systems operation. Inductive analysis answers the question, "What are the system states due to some event?" In reliability and risk studies, this "event" is some fault in the system. Several approaches using the inductive approach include: PrHA, FMEA, and ETA. Deductive approaches provide reasoning for a specific conclusion from general conditions. For system analysis this technique attempts to identify what modes of a system/subsystem/component failure can be used to contribute to the failure of the system. This technique answers the question, "How a system state can occur?" Inductive reasoning provides the techniques for FTA or its complement success tree analysis (STA).

9.1.3.6 Selected Risk Assessment Methods

Qualitative versus Quantitative Risk Assessment. The risk assessment methods can be categorized according to how the risk is determined: by quantitative or qualitative analysis. Qualitative risk analysis uses judgment and sometimes "expert" opinion to evaluate the probability and consequence values. This subjective approach may be sufficient to assess the risk of a system, depending on the available resources.

Quantitative analysis relies on probabilistic and statistical methods, and databases that identify numerical probability values and consequence values for risk assessment. This objective approach examines the system in greater detail to assess risks.

The selection of a quantitative or qualitative method depends upon the availability of data for evaluating the hazard and the level of analysis needed to make a confident decision. Qualitative methods offer analyses without detailed information, but the intuitive and subjective processes may result in differences in outcomes by those who use them. Quantitative analysis generally provides a more uniform understanding among different individuals, but requires quality data for accurate results. A combination of both qualitative and quantitative analyses can be used depending on the situation.

Risk assessment requires estimates of the failure likelihood at some identified levels of decision making. The failure likelihood can be estimated in the form of lifetime failure likelihood, annual failure likelihood, mean time between failures, or failure rate. The estimates can be in numeric or nonnumeric form. An example numeric form for an annual failure probability is 0.00015, and for a mean time between failures is 10 years. An example nonnumeric form for "an annual failure likelihood" is large, and for a "mean time between failures" is medium. In the latter nonnumeric form, guidance needs to be provided regarding the meaning of terms such as large, medium, small, very large, very small, etc. The selection of the form should be based on the availability of information, the ability of the personnel providing the needed information to express it in one form or another, and the importance of having numeric vs. nonnumeric information in formulating the final decisions.

The types of failure consequences that should be considered in a study need to be selected. They can include production loss, property damage, environmental damage, and safety–loss in the form of human injury and death. Approximate estimates of failure consequences at the identified levels of decision making need to be determined. The estimates can be in numeric or nonnumeric form. An example numeric form for production loss is 1000 units. An example nonnumeric form for production loss is large. In the latter nonnumeric form, guidance needs to be provided regarding the meaning of terms such as large, medium, small, very large, very small, etc. The selection of the form should be based on the availability of information, the ability of the personnel providing the needed information to express it in one form or another, and the importance of having numeric vs. nonnumeric information in formulating the final decisions.

Risk estimates can be determined as a pair of the likelihood and consequences, and computed as the arithmetic multiplication of the respective failure likelihood and consequences for the equipment, components and details. Alternatively, for all cases, plots of failure likelihood versus consequences can be developed. Then, approximate ranking of them as groups according to risk estimates, failure likelihood, and/or failure consequences can be developed.

Preliminary Hazard Analysis. Preliminary hazard analysis (PrHA) is a common risk-based technology tool with many applications. The general process is shown in Figure 9.6. This technique requires experts to identify and rank the possible accident scenarios that may occur. It is frequently used as a preliminary method to identify and reduce the risks associated with major hazards of a system.

Failure Mode and Effects Analysis. Failure mode and effects analysis (FMEA) is another popular risk-based technology tool as shown in Figure 9.7. This technique has been introduced both in the national

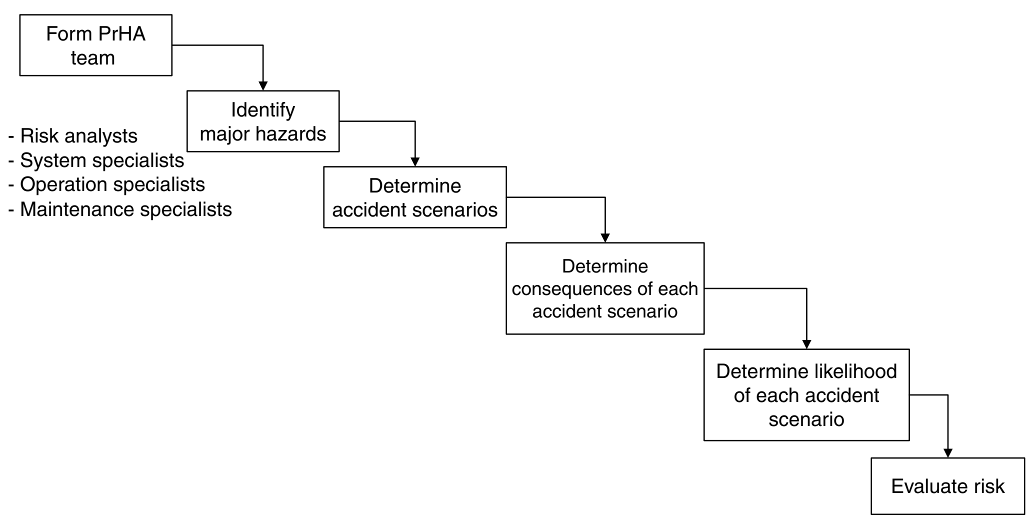


FIGURE 9.6 Preliminary hazard analysis (PrHA) process.

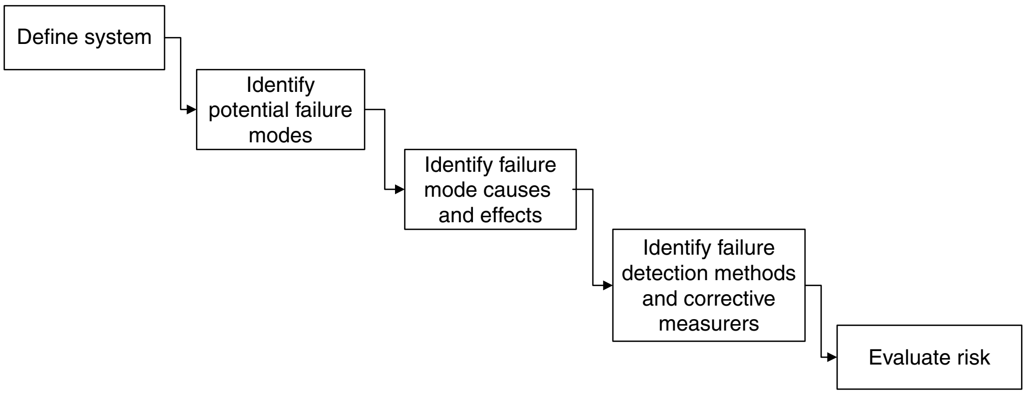


FIGURE 9.7 Failure mode and effects analysis (FMEA) process.

and international regulations for the aerospace (US MIL-STD-1629A), processing plant, and marine industries. The Society of Automotive Engineers, in its recommended practice, introduces two types of FMEA: design and process FMEA. This analysis tool assumes a failure mode occurs in a system/component through some failure mechanism; the effect of this failure on other systems is then evaluated. A risk ranking can be developed for each failure mode for the effect on the overall performance of the system.

The various terms used in FMEA with examples based on the manufacturing of personal flotation devices (PFDs) are provided under subsequent headings to include failure mode, failure effect, severity rating, causes, occurrence rating, controls, detection rating, and risk priority number.

Risk Matrices. Risk can be assessed and presented using matrices for preliminary screening by subjectively estimating probabilities and consequences in a qualitative manner. A risk matrix is a two-dimensional presentation of likelihood and consequences using qualitative metrics for both dimensions. According to this method, risk is characterized by categorizing probability and consequence on the two axes of a matrix. Risk matrices have been used extensively for screening of various risks. They may be used alone or as a first step in a quantitative analysis. Regardless of the approach used, risk analysis should be a dynamic process, i.e., a living process where risk assessments are reexamined and adjusted. Actions or inactions in one area can affect risk in another; therefore continuous updating is necessary.

The likelihood metric can be constructed using the categories shown in Table 9.5, whereas the consequences metric can be constructed using the categories shown in Table 9.6 with an example provided in Table 9.7. The consequence categories of Table 9.6 focus on the health and environmental aspects of consequences. The consequence categories of Table 9.7 focus on the economic impact, and should be adjusted to meet specific needs of industry and/or applications. An example risk matrix is shown in Figure 9.8. In the figure, each boxed area is shaded depending on a subjectively assessed risk level. Three risk levels are used herein for illustration purposes: low (L), medium (M), and high (H). Other risk levels may be added using a scale of five levels instead of three levels if needed. These risk levels

TABLE 9.5 Likelihood Categories for a Risk Matrix

Category	Description	Annual Probability Range
A	Likely	>0.1 (1 in 10)
B	Unlikely	>0.01 (1 in 100) but <0.1
C	Very Unlikely	>0.001 (1 in 1,000) but <0.01
D	Doubtful	>0.0001 (1 in 10,000) but <0.001
E	Highly Unlikely	>0.00001 (1 in 100,000) but <0.0001
F	Extremely Unlikely	<0.00001 (1 in 100,000)

TABLE 9.6 Consequence Categories for a Risk Matrix

Category	Description	Examples
I	Catastrophic	Large number of fatalities and/or major long-term environmental impact.
II	Major	Fatalities and/or major short-term environmental impact.
III	Serious	Serious injuries and/or significant environmental impact.
IV	Significant	Minor injuries and/or short-term environmental impact.
V	Minor	First aid injuries only and/or minimal environmental impact.
VI	None	No significant consequence.

are also called *severity factors*. The high (H) level can be considered as unacceptable risk level, the medium (M) level can be treated as either undesirable or as acceptable with review, and the low (L) level can be treated as acceptable without review.

Event Modeling: Event, Success Trees, and Fault Trees. Event modeling is a systematic—and often most complete—way to identify accident scenarios and quantify risk for risk assessment. This risk-based technology tool provides a framework for identifying scenarios to evaluate the performance of a system or component through system modeling. The combination of event tree analysis (ETA), success tree analysis (STA), and fault tree analysis (FTA) can provide a structured analysis to system safety.

Event tree analysis is often used if the successful operation of a component/system depends on a discrete (chronological) set of events. The initiating event is first followed by other events leading to an overall result (consequence). The ability to address a complete set of scenarios is developed because all combinations of both the success and failure of the main events are included in the analysis. The probability of occurrence of the main events of the event tree can be determined using a fault tree or its complement the success tree. The scope of the analysis for event trees and fault trees depends on the objective of the analysis.

Event tree analysis is appropriate if the operation of some system/component depends on a successive group of events. Event trees identify the various combinations of event successes and failures as a result of an initiating event to determine all possible scenarios. The event tree starts with an initiating event followed by some reactionary event. This reaction can either be a success or failure. If the event succeeds, the most commonly used indication is the upward movement of the path branch. A downward branch of the event tree marks the failure of an event. The remaining events are evaluated to determine the different possible scenarios. The scope of the events can be functions/systems that can provide some reduction to the possible hazards from the initiating event. The final outcome of a sequence of events identifies the overall state resulting from the scenario of events. Each path represents a failure scenario with varying levels of probability and risk. Different event trees can be created for different event initiators. Figure 9.9 shows an example event tree for the basic elements of a sprinkler system that might be critical for maintaining the integrity of a marine vessel.

Based on the occurrence of an initiating event, event tree analysis examines possible system outcomes or consequences. This analysis tool is particularly effective in showing interdependence of system components which is important in identifying events, that at first might appear insignificant, but due to

TABLE 9.7 Example Consequence Categories for a Risk Matrix in 2003 Monetary Amounts (US$)

Category	Description	Cost
I	Catastrophic Loss	> $10,000,000,000
II	Major Loss	> $1,000,000,000 but < $10,000,000,000
III	Serious Loss	> $100,000,000 but < $1,000,000,000
IV	Significant Loss	> $10,000,000 but < $100,000,000
V	Minor Loss	> $1,000,000 but < $10,000,000
VI	Insignificant Loss	< $1,000,000

Severity factors. The high (H) level can be considered as unacceptable risk level, the medium (M) level can be treated as either undesirable or as acceptable with review, and the low (L) level can be treated as acceptable without review.

Table 9-5. Likelihood categories for a risk matrix

Category	Description	Annual probability range
A	Likely	≥ 0.1 (1 in 10)
B	Unlikely	≥ 0.01 (1 in 100) but < 0.1
C	Very unlikely	≥ 0.001 (1 in 1,000) but < 0.01
D	Doubtful	≥ 0.0001 (1 in 10,000) but < 0.001
E	Highly unlikely	≥ 0.00001 (1 in 100,000) but < 0.0001
F	Extremely unlikely	< 0.00001 (1 in 100,000)

Table 9-6. consequence categories for a risk matrix

Category	Description	Examples
I	Catastrophic	Large number of fatalities and/or major long-term environmental impact.
II	Major	Fatalities and/or major short-term environmental impact.
III	Serious	Serious injuries and/or significant environmental impact.
IV	Significant	Minor injuries and/or short-term environmental impact.
V	Minor	First aid injuries only and/or minimal environmental impact.
VI	None	No significant consequence.

Table 9-7. Example consequence categories for a risk matrix in 2003 monetary amounts (US$)

Category	Description	Cost
I	Catastrophic loss	≥ $10,000,000,000
II	Major loss	≥ $1,000,000,000 but < $10,000,000,000
III	Serious loss	≥ $100,000,000 but < $1,000,000,000
IV	Significant loss	≥ $10,000,000 but < $100,000,000
V	Minor loss	≥ $1,000,000 but < $10,000,000
VI	Insignificant loss	< $1,000,000

Probability category	A	L	M	M	H	H	H
	B	L	L	M	M	H	H
	C	L	L	L	M	M	H
	D	L	L	L	L	M	M
	E	L	L	L	L	L	M
	F	L	L	L	L	L	L
		VI	V	IV	III	II	I
	Consequence category						

FIGURE 9.8 Example Risk Matrix

the interdependency result in devastating results. Event tree analysis is similar to fault tree analysis because both methods use probabilistic reliability data of the individual components and events along each path to compute the likelihood of each outcome.

A quantitative evaluation of event tree probability values can be used for each event to evaluate the probability of the overall system state. Probability values for the success or failure of the events can be used to identify the probability for a specific event tree sequence. The probabilities of the events in a sequence can be provided as an input to the model or evaluated using fault trees. These probabilities for various events in a sequence can be viewed as conditional probabilities and therefore can be multiplied to obtain the occurrence probability of the sequence. The probabilities of various sequences can be summed up to determine the overall probability of a certain outcome. The addition of consequence evaluation of a scenario allows for generation of a risk value. For example, the occurrence probability of the top branch, i.e., scenario, in Figure 9.9 is computed as the product of the probabilities of the composing events to this scenario, i.e., $F \cap PO \cap SF \cap SS \cap FE$ or (F)(PO)(SF)(SS)(FE) for short.

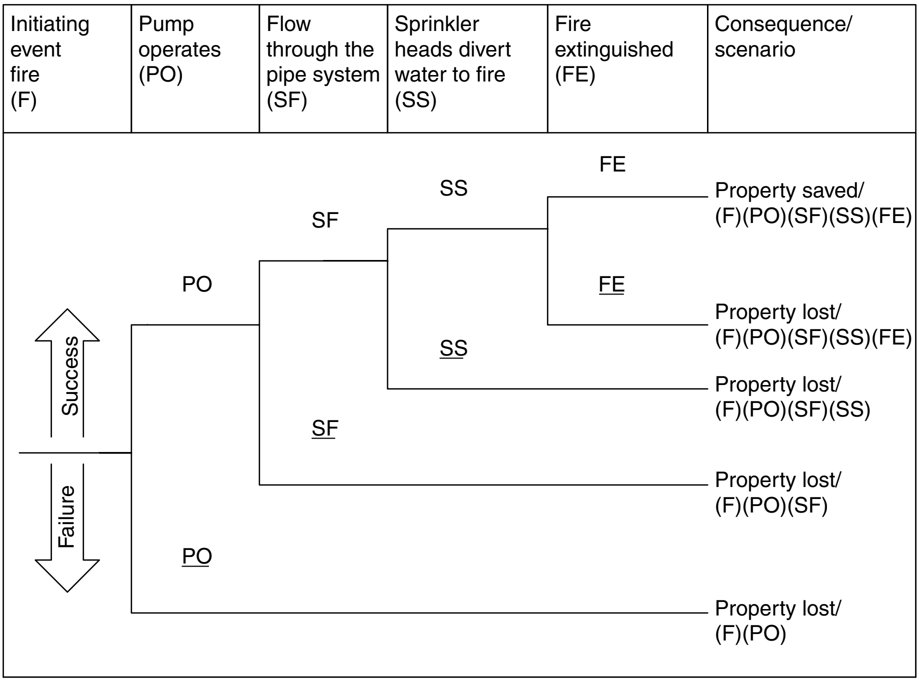


FIGURE 9.9 Event tree example for sprinkler system.

Complex systems are often difficult to visualize and the effect of individual components on the system as a whole is difficult to evaluate without an analytical tool. Two methods of modeling that have greatly improved the ease of assessing system reliability/risk are fault trees (FT) and success trees (ST). A fault tree is a graphical model created by deductive reasoning leading to various combinations of events that lead to the occurrence of some top event failure. A success tree shows the combinations of successful events leading to the success of the top event. A success tree can be produced as the complement (opposite) of the fault tree as illustrated in this section. Fault trees and success trees are used to further analyze the event tree headings (the main events in an event tree) to provide further detail to understand system complexities. In constructing the FT/ST only those failure/success events that are considered significant are modeled. This determination is assisted by defining system boundaries. For example, the event "pump operates (PO)" in Figure 9.9 can be analyzed by developing a top-down logical breakdown of failure or success using fault tress or event trees, respectively.

Fault tree analysis (FTA) starts by defining a top event that is commonly selected as an adverse event. An engineering system can have more than one top event. For example, a ship might have the following top events for the purpose of reliability assessment: power failure, stability failure, mobility failure, or structural failure. Then, each top event needs to be examined using the following logic: in order for the top event to occur, other events must occur. As a result, a set of lower-level events is defined. Also, the form in which these lower-level events are logically connected (i.e., in parallel or in series) needs to be defined. The connectivity of these events is expressed using "AND" or "OR" gates. Lower-level events are classified into the following types:

1. *Basic events*: These events cannot be decomposed further into lower-level events. They are the lowest events that can be obtained. For these events, failure probabilities need be obtained.
2. *Events that can be decomposed further*: These events can be decomposed further to lower levels. Therefore, they should be decomposed until the basic events are obtained.
3. *Undeveloped events*. These events are not basic and can be decomposed further. However, because they are not important, they are not developed further. Usually, the probabilities of these events are

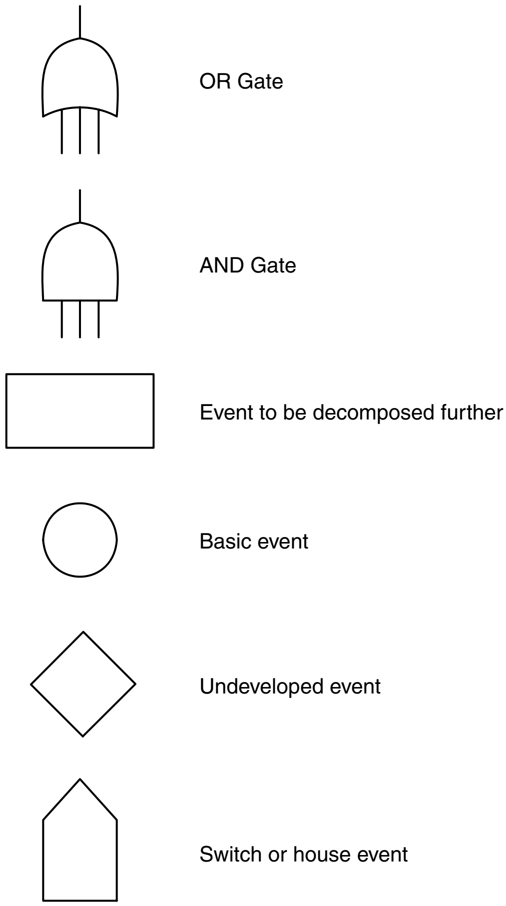


FIGURE 9.10 Symbols used in fault tree analysis.

very small or the effect of their occurrence on the system is negligible, or can be controlled or mediated.

4. *Switch (or house) events*. These events are not random, and can be turned on or off with full control. The symbols shown in Figure 9.10 are used for these events. Also, a continuation symbol is shown that is used to break up a fault tree into several parts for the purpose of fitting it in several pages.

FTA requires the development of a tree-looking diagram for the system that shows failure paths and scenarios that can result in the occurrence of a top event. The construction of the tree should be based on the building blocks and the Boolean logic gates.

The outcome of interest from the fault tree analysis is the occurrence probability of the top event. Because the top event was decomposed into basic events, its occurrence can be stated in the form of "AND" and "OR" of the basic events. The resulting statement can be restated by replacing the "AND" with the intersection of the corresponding basic events, and the "OR" with the union of the corresponding basic events. Then, the occurrence probability of the top event can be computed by evaluating the probabilities of the unions and intersections of the basic events. The dependence between these events also affects the resulting probability of the system.

For large fault trees, the computation of the occurrence probability of the top event can be difficult because of their size. In this case, a more efficient approach is needed for assessing the reliability of a system; such as the minimal cut set approach. According to this approach, each cut set is defined as a set of basic events where the joint occurrence of these basic events results in the occurrence of the top event. A minimal cut set is a cut set with the condition that the nonoccurrence of any one basic event from this set results in the nonoccurrence of the top event. Therefore, a minimal cut set can be viewed as a subsystem in parallel. In general, systems have more than one minimal cut sets. The occurrence of the top event of the system can, therefore, be due to any one of these minimal cut sets. As a result, the system can be viewed as the union of all the minimal cut sets for the system. If probability values are assigned to the cut sets, a probability for the top event can be determined.

A simple example of this type of modeling is shown in Figure 9.11 for a pipe system using a reliability block diagram. If the goal of the system is to maintain water flow from one end of the system to the other, then the individual pipes can be related with a Boolean logic. Both pipe (a) and pipe (d) and pipe (b) or pipe (c) must function for the system to meet its goal as shown in the success tree Figure 9.12a. The compliment of the success tree is the fault tree. The goal of the fault tree model is to construct the logic for system failure as shown in Figure 9.12b. After these tree elements have been defined, possible failure scenarios of a system can be defined.

As previously described, a failure path is often referred to as a *cut set*. One objective of the analysis is to determine the entire minimal cut sets, where a minimal cut set is defined as a failure combination of all

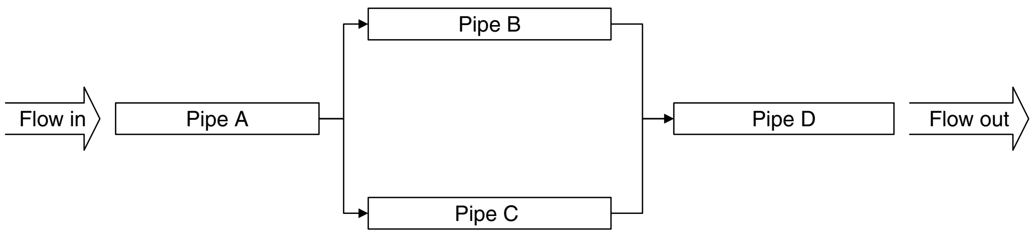


FIGURE 9.11 A reliability block diagram for a piping system.

essential events that can result in the failure top event. A minimal cut set includes in its combination all essential events, i.e., the nonoccurrence of any of these essential events in the combination of a minimal cut set results in the nonoccurrence of the minimal cut set. These failure combinations are used to compute the failure probability of the top event. The concept of the minimal cut sets applies only to the fault trees. A similar concept can be developed in the complementary space of the success trees, and is called the minimal pass set. In this case, a minimal pass set is defined as a survival (or success) combination of all essential success events that can result in success as defined by the top event of the success tree. For the piping example, the minimal cut sets are

$$\text{A} \tag{9.6a}$$

$$\text{D} \tag{9.6b}$$

$$\text{B and C} \tag{9.6c}$$

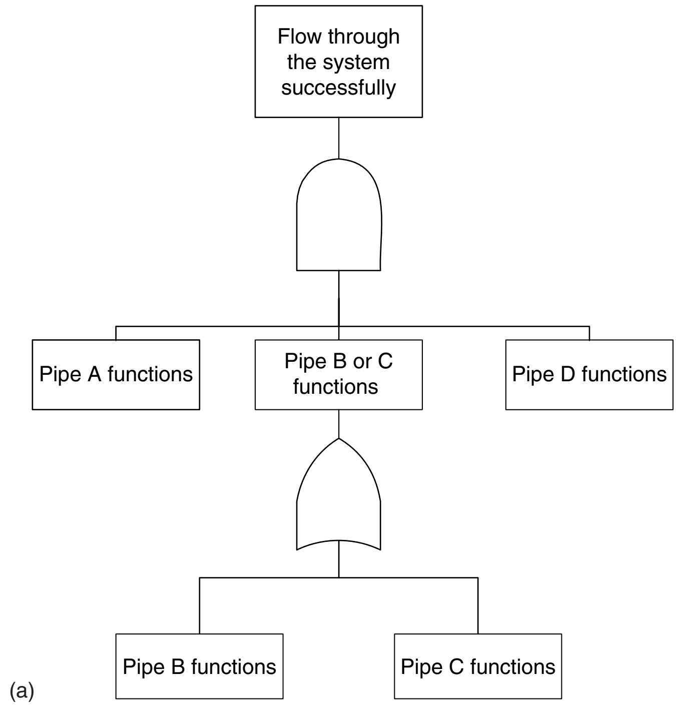
A minimal cut set includes events that are all necessary for the occurrence of the top event. For example, the following cut set is not a minimal cut set:

$$\text{A and B} \tag{9.7}$$

Example 9.1 Trends in Fault Tree Models and Cut Sets

This example demonstrates how the cut sets can be identified and constructed for different arrangements of OR and AND gates logically defining a top event occurrence. Generally, the number of cut sets increases by increasing the number of OR gates in the tree. For example, Figure 9.13 shows this trend by comparing cases a, b, and d. On the other hand, increasing the number of AND gates results in increasing the number of events included in the cut sets as shown in case c of Figure 9.13.

Common cause scenarios are events or conditions that result in the failure of seemingly separate systems or components. Common cause failures complicate the process of conducting risk analysis because a seemingly redundant system can be rendered ineffective by a common cause failure. For example, an emergency diesel generator fed by the same fuel supply as the main diesel engine will fail with the main diesel generator, if the fuel supply is the root source of the failure. The redundant emergency diesel generator is not truly redundant due to a common cause failure. Another example of common cause events is the failure of two separate but similar pieces of machinery due to a common maintenance problem, two identical pieces of equipment failing due to a common manufacturing defect, or two pieces of equipment failing due to a common environmental condition such as the flooding of a compartment or a fire in the vicinity of both pieces of machinery. A method for calculating the reliability of a system while taking into account common cause effects is the beta-factor model. Other methods include multiple Greek letter model, alpha-factor model, and beta-binomial failure-rate model.



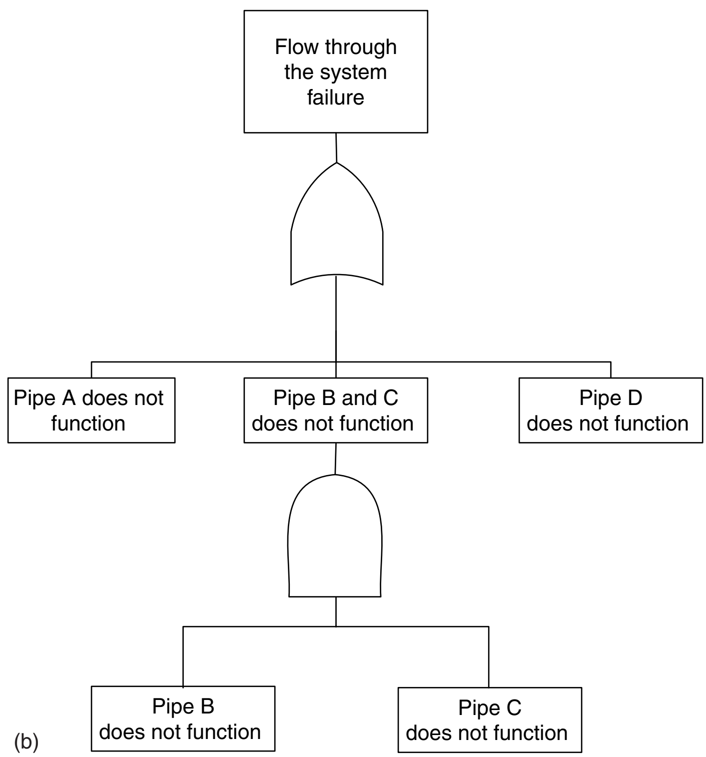


FIGURE 9.12 (a) Success tree for the pipe system example. (b) Fault tree for the pipe system example.

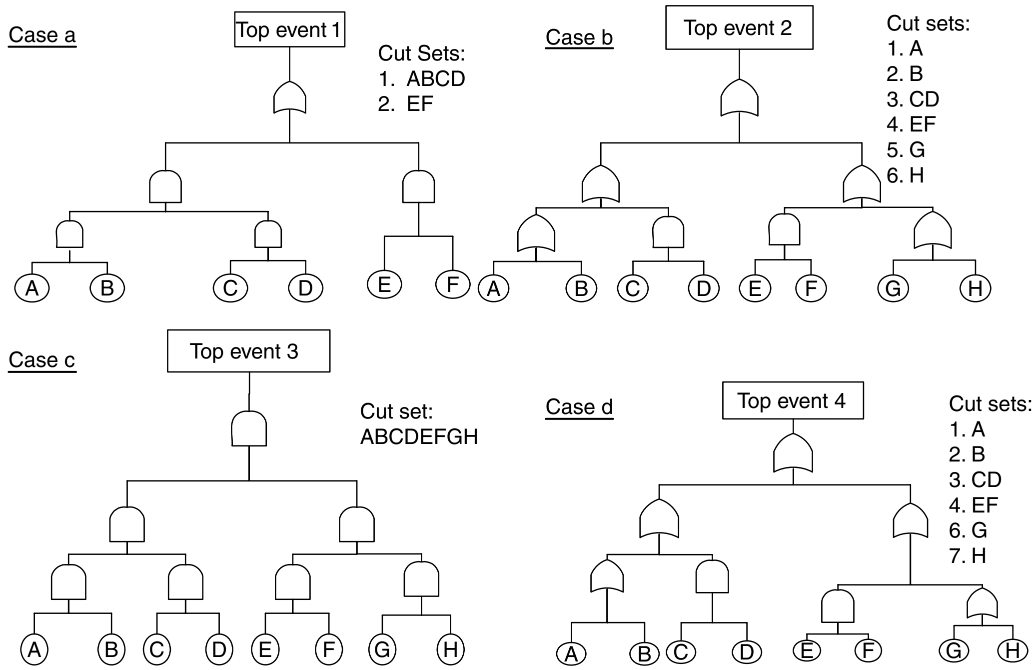


FIGURE 9.13 Trends in fault tree models and cut sets. (From Maryland Emergency Management Agency (MEMA), 2006, State of Maryland Guide for the Protection of Critical Infrastructure and Key Resources for Homeland Security, Volume 1: Critical Asset & Portfolio Risk Assessment (CAPRA) Methodology, Office of Homeland Security, Annapolis, MD.)

Part of risk-based decision analysis is pinpointing the system components that result in high-risk scenarios. Commercial system reliability software provides this type of analysis in the form of system reliability sensitivity factors to changes in the underlying component reliability values. In performing risk analysis, it is desirable to assess the importance of events in the model, or the sensitivity of final results to changes in the input failure probabilities for the events. Several sensitivity or importance factors are available and can be used. The most commonly used factors include (1) Fussell–Vesely factor and (2) Birnbaum factor. Also, a weighted combination of these factors can be used as an overall measure.

9.1.3.7 Human-Related Risks

Risk assessment requires the performance analysis of an entire system composed of a diverse group of components. The system definition readily includes the physical components of the system; however, humans are also part of most systems and provide significant contributions to risk. It has been estimated that nearly 90% of the accidents at sea are contributed to human error. The human contribution to risk can be estimated from an understanding of behavioral sciences. Both the "hardware failure" and human error should be addressed in the risk assessment since they both contribute to risks associated with the system. After the human error probabilities are determined, human error/failures are treated in the same fashion as hardware failures in performing risk assessment quantification.

The determination of the human error contribution to risk is determined by human reliability analysis (HRA) tools. HRA is the discipline that enables the analysis and impact of humans on the reliability and safety of systems. Important results of HRA are determining the likelihood of human error as well as ways in which human errors can be reduced. When combined with system risk analysis, HRA methods provide an assessment of the detrimental effects of humans on the performance of the system. Human reliability analysis is generally considered to be composed of three basic steps: error identification, modeling, and quantification.

Other sources of human-related risks are in the form of deliberate sabotage of a system from within a system or as threat from outside the system, such as a computer hacker or a terrorist. The hazard in this case is not simply random but intelligent. The methods introduced in earlier sections might not be fully applicable for this risk type. The threat scenarios to the system in this case have a dynamic nature that are affected by the defense or risk mitigation and management scenarios that would be implemented by an analyst. The use of game theory methods might be needed in this case in combination with other risk analysis and management methods. Game theory is introduced in the last subsection herein.

Human Error Identification. Human errors are unwanted circumstances caused by humans that result in deviations from expected norms that place systems at risk. It is important to identify the relevant errors to make a complete and accurate risk assessment. Human error identification techniques should provide a comprehensive structure for determining significant human errors within a system. Quality HRA allows for accuracy in both the HRA assessment and overall system risk assessment.

Identification of human errors requires knowledge about the interactions of humans with other humans or machines (the physical world). It is the study of these interfaces that allows for the understanding of human errors. Potential sources of information for identifying human error may be determined from task analysis, expert judgment, laboratory studies, simulation, and reports. Human errors may be considered active or latent depending on the time delay between when the error occurs and when the system fails.

January 09, 2006 9–32. It is important to note the distinction between human errors and human factors. Human errors are generally considered separately from human factors that applies information about human behavior, abilities, limitations, and other characteristics to the design of tools, machines, systems tasks, jobs, and environments for productive, safe, comfortable, and effective human use. Human factors are determined from performing descriptive studies for characterizing populations and experimental research. However, human factors analysis may contribute to the human reliability analysis.

Human Error Modeling. After human errors have been identified, they must be represented in a logical and quantifiable framework along with other components that contribute to the risk of the system. This framework can be determined from development of a risk model. Currently, there is no consensus on how to model human reliably. Many of these models utilize human event trees and fault trees to predict human reliability values. The identifications of human failure events can also be identified using failure mode and effects analysis. The human error rate estimates are often based on simulation tests, models, and expert estimation.

Human Error Quantification. Quantification of human error reliability promotes the inclusion of the human element in risk analysis. This is still a developing science requiring understanding of human performance, cognitive processing, and human perceptions. Because an exact model for human cognition has not been developed, much of the current human reliability data relies on accident databases, simulation, and other empirical approaches. Many of the existing data sources were developed for from specific industry data such as nuclear and aviation industries. The application of these data sources for a specific problem should be thoroughly examined prior to application for a specific model. The result of the quantification of human reliability in terms of probability of occurrence is typically called a *human error probability (HEP)*. There are many techniques that have been developed to help predict the HEP values. The technique for human error rate prediction (THERP) is one of the most widely used methods for HEP. This technique is based on data gathered from the nuclear and chemical processing industries. THERP relies on HRA event tree modeling to identify the events of concern. Quantification is performed from data tables of basic HEP for specific tasks that may be modified based on the circumstances affecting performance.

The degree of human reliability is influenced by many factors often called *performance shaping factors (PSFs)*. PSFs are those factors that affect the ability of people to carry out required tasks. For example, the knowledge people have on how to don/activate a personal flotation device (PFD) will affect the performance of this task. Training (another PSF) in donning PFDs can also assist in the ability to perform this task. Another example is the training that is given to passengers on airplanes before takeoff on using seatbelts, emergency breathing devices, and flotation devices. Often, the quantitative estimates of reliability are generated from a base error rate that is then altered based on the PSFs of the particular circumstances. Internal performance shaping factors are an individual's own attributes (experience, training, skills, abilities, attitudes) that affect the ability of the person to perform certain tasks. External PSFs are the dynamic aspects of situation, tasks, and system that affect the ability to perform certain tasks. Typical external factors include environmental stress factors (such as heat, cold, noise, situational stress, time of day), management, procedures, time limitations, and quality of person–machine interface. With these PSFs, it is easy to see the dynamic nature of HEP evaluation based on the circumstances of the analysis.

Reducing Human Errors. Error reduction is concerned with lowering the likelihood for error in an attempt to reduce risk. The reduction of human errors may be achieved by human factors interventions or by engineering means. Human factors interventions include improving training or improving the human–machine interface (such as alarms, codes, etc.) based on an understanding of the causes of error. Engineering means of error reduction may include automated safety systems or interlocks. The selection of the corrective actions to take can be done through decision analysis considering cost–benefit criteria.

Game Theory for Intelligent Threats. Game theory can be used to model human behavior as a threat to a system. Generally, game theory utilizes mathematics, economics, and the other social and behavioral sciences to model human behavior.

An example of intelligent threats is terrorism and sabotage as an ongoing battle between coordinated opponents representing a two-party game, where each opponent seeks to achieve their own objectives within a system. In the case of terrorism, it is a game of a well-established political system as a government vs. an emerging organization that uses terrorism to achieve partial or complete dominance. Each player in this game seeks a utility, i.e., benefit, that is a function of the desired state of the system. In this case, maintaining system survival is the desired state for the government, whereas the opponent seeks a utility based on the failure state of the system. The government, as an opponent, is engaged in risk mitigation whose actions seek to reduce the threat, reduce the system vulnerability, and/or mitigate the consequences of any successful attacks. The terrorists, as an opponent, can be viewed as the aggressor who strives to alter or damage their opponent's desired system state. This game involves an intelligent threat and is dynamic. The game is ongoing until the probability of a successful disruptive attempt of the aggressor reaches an acceptable level of risk—a stage where risk is considered under control—and the game is brought to an end. Classical game theory can be used in conjunction with probabilistic risk analysis to determine optimal mitigation actions that maximize benefits.

A classical example used to introduce game theory is called the *prisoners' dilemma* and is based on two suspects that are captured near the scene of a crime and are questioned separately by authority such as the police. Each has to choose whether or not to confess and implicate the other. If neither person confesses, then both will serve, for example, one year on a charge of carrying a concealed weapon. If each confesses and implicates the other, both will go to prison for, say, 10 years. However, if one person confesses and implicates the other, and the other person does not confess, the one who has collaborated with the police will go free, while the other person will go to prison for, say, 20 years on the maximum penalty. The strategies in this case are: confess or do not confess. The payoffs, herein *penalties*, are the sentences served. The problem can be expressed compactly in a payoff table of a kind that has become standard in game theory as provided in Table 9.8. The entries of this table mean that each prisoner chooses one of the two strategies, i.e., the first suspect chooses a row and the second suspect chooses a column. The two numbers in each cell of the table provide the outcomes for the two suspects for the corresponding pair of strategies chosen by the suspects as an ordered pair. The number to the left of the comma is the payoff to the person who chooses the rows, i.e., the first suspect, whereas the number to the right of the comma is the payoff to the person who chooses the columns, i.e., the second suspect. Thus reading down the first column, if they both confess, each gets 10 years, but if the second suspect confesses and first suspect does not, the first suspect gets 20 years and second suspect goes free. This example is not a zero-sum game because the payoffs are all losses. However, many problems can be cast with losses (negative numbers) and gains (positive numbers) with a total for each cell in the payoff table. A problem with a payoff table such that the payoffs in each cell add up to zero is called a *zero-sum game*.

TABLE 9.8 Payoff Table in Years for the Prisoners' Dilemma Game

		Second Suspect	
		Confess	Don't Confess
First suspect	Confess	(10, 10)	(0, 20)
	Don't confess	(20, 0)	(1, 1)

The solution to this problem needs to be based on identifying rational strategies that can be based on both persons wanting to minimize the time they spend in jail. One suspect might reason that "two things can happen: either the other suspect confesses or keeps quiet. Suppose the second suspect confesses, I will get 20 years if I don't confess, 10 years if I do; therefore, in this case, it's best to confess. On the other hand, if the other suspect doesn't confess, and I don't either, I get a year; but in that case, if I confess I can go free. Either way, it's best if I confess. Therefore, I'll confess." But the other suspect can and presumably will reason in the same way. In this case, they both confess and go to prison for 10 years each, although if they had acted irrationally, and kept quiet, they each could have gotten off with one year each. The rational strategies of the two suspects have fallen into something called *dominant strategy equilibrium*. The meaning of the term dominant strategy equilibrium requires defining the term dominant strategy that results from an individual player (the suspect, in this case) in a game evaluating separately each of the strategy combinations being faced, and, for each combination, choosing from these strategies the one that gives the greatest payoff. If the same strategy is chosen for each of the different combinations of strategies the player might face, that strategy is called a dominant strategy for that player in that game. The dominant strategy equilibrium occurs if, in a game, each player has a dominant strategy, and each player plays the dominant strategy, then that combination of (dominant) strategies and the corresponding payoffs are said to constitute the dominant strategy equilibrium for that game. In the prisoners' dilemma game, to confess is a dominant strategy, and when both suspects confess, dominant strategy equilibrium is established. The dominant strategy equilibrium is also called *Nash equilibrium*. The definition of Nash equilibrium is a set of strategies with the property that no player can benefit by changing his/her strategy while the other players keep their strategies unchanged, then that set of strategies and the corresponding payoffs constitute the Nash equilibrium.

The prisoners' dilemma game is based on two strategies per suspect that can be viewed as deterministic in nature, i.e., nonrandom. In general, many games, especially ones permitting repeatability in choosing strategies by players, can be constructed with strategies that have associated probabilities. For example, strategies can be constructed based on probabilities of 0.4 and 0.6 that sum to one. Such strategies with probabilities are called *mixed strategies* as opposed to *pure strategies* that do not involve probabilities of the prisoners' dilemma game. A mixed strategy occurs in a game if a player chooses among two or more strategies at random according to specific probabilities.

In general, gaming could involve more than two players. In the prisoners' dilemma game, a third player that could be identified is the authority and its strategies. The solution might change as a result of adding the strategies of this third player. The use of these concepts in risk analysis and mitigation needs further development and exploration.

Risk Methods for Protecting Infrastructure and Key Resources. The protection of critical infrastructure and key resources (CI/KR) for homeland security requires choosing among a large set of protective, response, and recovery actions for reducing risk to an acceptable level. The selection of investment alternatives for improving asset security and increasing infrastructure resilience depends on two factors: their cost to implement and relative cost-effectiveness. To accomplish this task, the Department of Homeland Security (DHS) has identified risk methods as the primary underlying framework for system evaluations, operational assessments, technology assessments, resource and support analyses, and field operations analyses. According to the draft DHS National Infrastructure Protection Plan, cost–benefit analysis is the hallmark of critical infrastructure protection decision making. Figure 9.14 shows a methodology for informing decisions relating to CI/KR protection

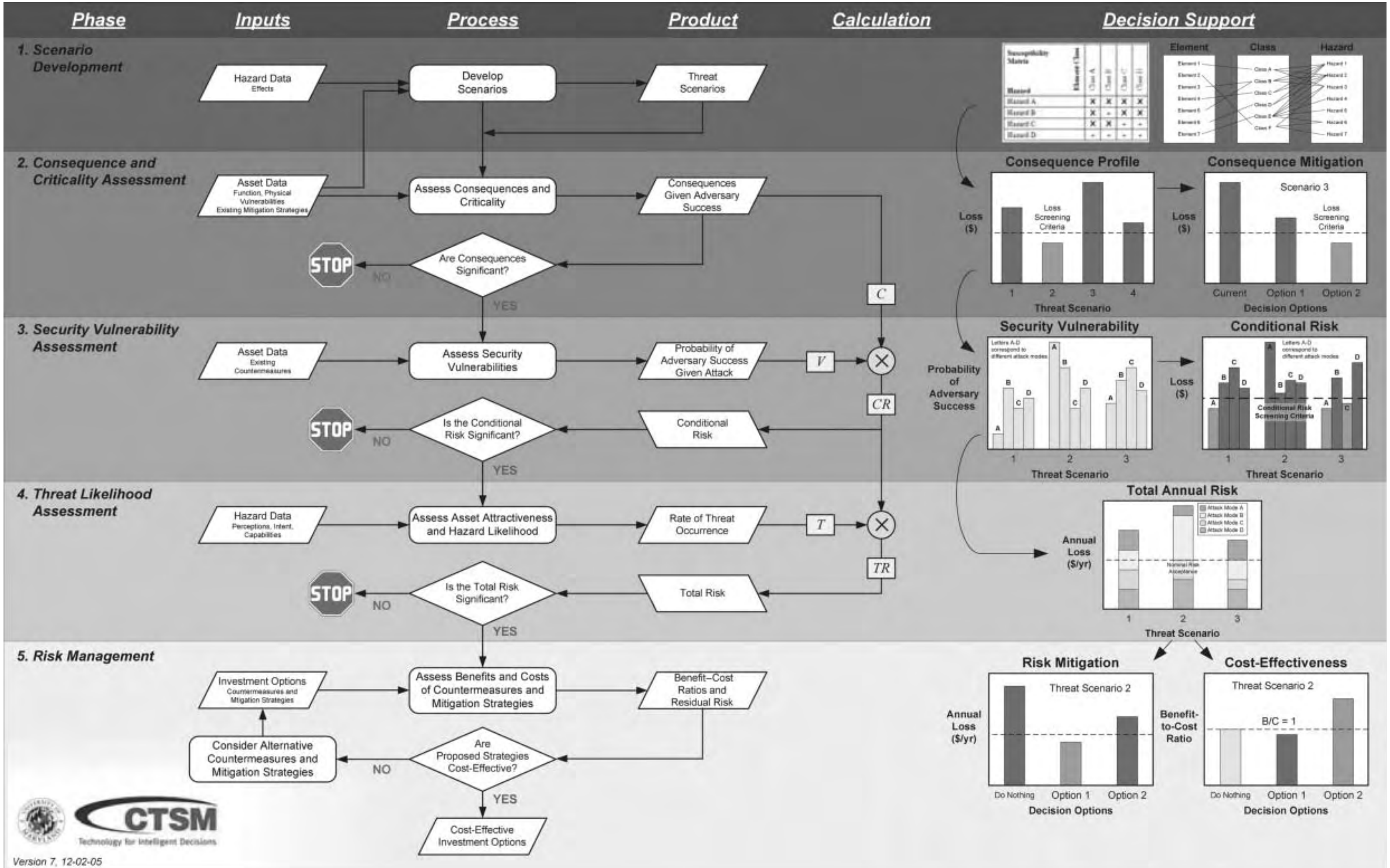


FIGURE 9.14 Critical Asset and Portfolio risk analysis methodology.

(MEMA 2006). It is a quantitative, asset-driven framework for assessing the risk of consequential CI/KR exploitation and disruption due to a variety of plausible threat scenarios and hazard events. This process is divided into five phases as shown in the Figure 9.14. The results from each phase supports decision making of some type, whether it be asset screening and selection based on consequences and their criticality, or countermeasure selection and evaluation based on the results from the security vulnerability assessment phase.

9.1.3.8 Economic and Financial Risks

Economic and financial risks can be grouped into categories that include market risks, credit risks, operation risks, and reputation risks. These four categories are described in subsequent sections.

Market Risks. Governments and corporations operate in economic and financial environments with some levels of uncertainty and instability. A primary contributor to defining this environment is interest rates. Interest rates can have significant impact on the costs of financing a project, and corporate cash flows and asset values. For example, interest rates in the U.S. shot up in 1979 and peaked in 1981, followed by gradual decline with some fluctuations until 2002.

For projects that target global markets, exchange rate instability can be a major risk source. Exchange rates have been volatile ever since the breakdown of the Bretton Woods system to fixed exchange rates in the early 1970s. An example of a bust-up in exchange rates is the fall of the value of the British sterling and Italian lira as a result of the failure of the exchange rate mechanism in September 1992.

Many projects are dependent on availability of venture capital and the stock performance of corporation thereby introducing another risk source related to stock market volatility. Stock prices rose significantly in the inflationary booms of the early 1970s, then fell considerably a little later. They recovered afterward, and fell again in the early 1981. The market rose to a peak until it crashed in 1987, followed by an increase with some swings until reaching a new peak fueled by Internet technologies until its collapse in 2001.

Other contributing factors to economic and finance instability is commodity prices in general and energy prices in particular, primarily crude oil. The hikes in oil prices in the 1973–1974 affected commodity prices greatly and posed series challenges to countries and corporations.

Another contributing source to volatility is derivatives for commodities, foreign currency exchange rates, and stock prices and indices, among others. Derivatives are defined as contracts whose values or payoffs depend on those of other assets, such as the options to buy commodities in the future or options to sell commodities in the future. They offer not only opportunities for hedging positions and managing risks that can be stabilizing, but also speculative opportunities to others that can be destabilizing and a contributor to volatility.

Credit Risks. Credit risks are associated with potential defaults on notes or bonds, as examples, by corporations, including subcontractors. Also, credit risks can be associated with market sentiments that determine a company likelihood of default that could affect its bond rating and ability to purchase money, and maintain projects and operations.

Operational Risks. Operational risks are associated with several sources that include out-of-control operations risk that could occur when a corporate branch undertake significant risk exposure that is not accounted for by a corporate headquarters leading potentially to its collapse, an example being the British Barings Bank that collapsed primarily as a result of its failure to control the market exposure being created within a small overseas branch of the bank.

Another risk source in this category is liquidity risk in which a corporation needing funding more than it can arrange. Also, it could include money transfer risks and agreement breach.

Operational risks include model risks. Model risks are associated with the models and underlying assumptions used to incorrectly value financial instruments and cash flows.

Reputation Risks. The loss of business attributable to decrease in a corporation's reputation can pose another risk source. This risk source can affect its credit rating, ability to maintain clients, workforce, etc. This risk source usually occurs at a slow attrition rate. It can be an outcome of poor management decisions and business practices.

9.1.3.9 Data Needs for Risk Assessment

In risk assessment, the methods of probability theory are used to represent engineering uncertainties. In this context, it refers to event occurrence likelihoods that occur with periodic frequency, such as weather, yet also to conditions that are existent but unknown, such as probability of an extreme wave. It applies to the magnitude of an engineering parameter, yet also to the structure of a model. By contrast, probability is a precise concept. It is a mathematical concept with an explicit definition. The mathematics of probability theory are used to represent uncertainties, despite that those uncertainties are of many forms.

The term *probability* has a precise mathematical definition, but its meaning when applied to the representation of uncertainties is subject to differing interpretations. The frequentist view holds that probability is the propensity of a physical system in a theoretically infinite number of repetitions; i.e., the frequency of occurrence of an outcome in a long series of similar trials (for example, the frequency of a coin landing heads-up in an infinite number of flips is the probability of that event). In contrast, the Bayesian view holds that probability is the rational degree of belief that one holds in the occurrence of an event or the truth of a proposition; probability is manifest in the willingness of an observer to take action upon this belief. This latter view of probability, which has gained wide acceptance in many engineering applications, permits the use of quantified professional judgment in the form of subjective probabilities. Mathematically, such subjective probabilities can be combined or operated on as any other probability.

Data are needed to perform quantitative risk assessment or provide information to support qualitative risk assessment. Information may be available if data have been maintained on a system and components of interest. The relevant information for risk assessment included the possible failures, failure probabilities, failure rates, failure modes, possible causes, and failure consequences. In the case of a new system, data may be used from similar systems if this information is available. Surveys are a common tool used to provide some means of data. Statistical analysis can be used to assess confidence intervals and uncertainties in estimated parameters of interest. Expert judgment may also be used as another source of data (Ayyub 2002). The uncertainty with the quality of the data should be identified to assist in the decision making process.

Data can be classified to including generic and project or plant specific types. Generic data are information from similar systems and components. This information may be the only information available in the initial stages of system design. Therefore, potential differences due to design or uncertainty may result from using generic data on a specific system. Plant-specific data are specific to the system being analyzed. This information is often developed after the operation of a system. Relevant data need to be identified and collected as data collection can be costly. The data collected can then be used to update the risk assessment. Bayesian techniques can be used to combine objective and subjective data.

Data can be classified as failure probability data and failure consequence data. The failure probability data can include failure rates, hazard functions, times between failures, results from reliability studies, and any influencing factors and their effects. Failure consequence data include loss reports, damages, litigation outcomes, repair costs, injuries, and human losses. Also included are influencing factors and effects of failure prevention and consequence mitigation plans. Areas of deficiency in terms of data availability should be identified, and sometimes failure databases need to be constructed. Data deficiency can be used as a basis for data collection and expert-opinion elicitation.

9.1.4 Risk Management and Control

Adding risk control to risk assessment produces risk management. Risk management is the process by which system operators, managers, and owners make safety decisions, regulatory changes, and choose different system configurations based on the data generated in the risk assessment. Risk management involves using information from the previously described risk assessment stage to make educated decisions about system safety. Risk control includes failure prevention and consequence mitigation.

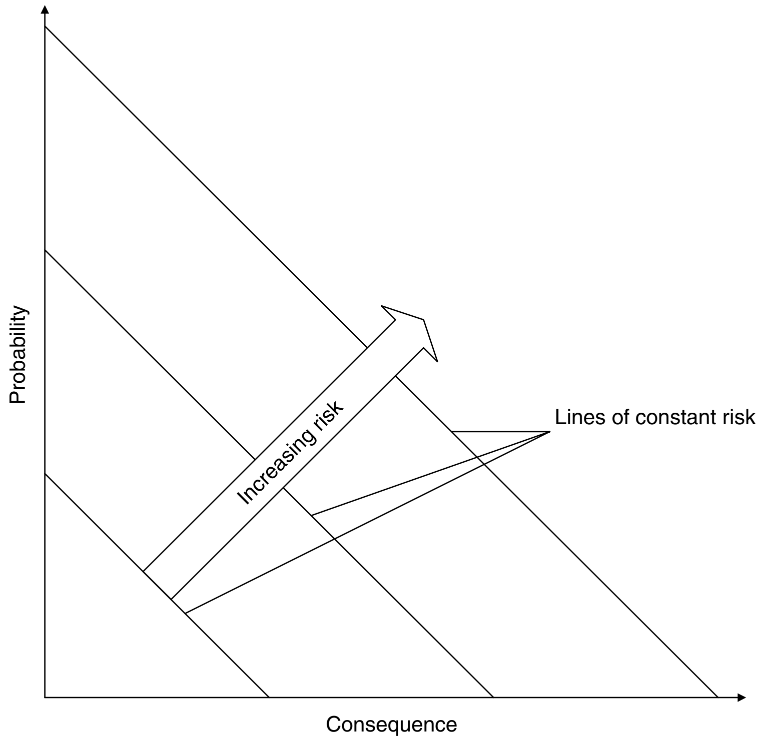


FIGURE 9.15 Risk plot.

Risk management requires the optimal allocation of available resources in support of group goals. Therefore, it requires the definition of acceptable risk, and comparative evaluation of options and/or alternatives for decision making. The goals of risk management are to reduce risk to an acceptable level and/or prioritize resources based on comparative analysis. Risk reduction is accomplished by preventing an unfavorable scenario, reducing the frequency, and/or reducing the consequence. A graph showing the risk relationship is shown in Figure 9.15 as linear contours of constant risk, although due to risk aversion these lines are commonly estimated as nonlinear curves and should be treated as nonlinear curves. Moreover, the vertical axis is termed as probability whereas it is commonly expressed as an annual exceedence probability or frequency as shown in Figure 9.1. In cases involving qualitative assessment, a matrix presentation can be used as shown in Figure 9.8. The figure shows probability categories, severity categories, and risk ratings. A project's base value is commonly assumed as zero. Each risk rating value requires a different mitigation plan.

9.1.4.1 Risk Acceptance

Risk acceptance constitutes a definition of safety as discussed in previous sections. Therefore, risk acceptance is considered a complex and controversial subject that is often subject to debate. The determination of acceptable levels of risk is important to determine the risk performance a system needs to achieve to be considered safe. If a system has a risk value above the risk acceptance level, actions should be taken to address safety concerns and improve the system through risk reduction measures. One difficulty with this process is defining acceptable safety levels for activities, industries, structures, etc. Because the acceptance of risk depends upon society's perceptions, the acceptance criteria do not depend on the risk value alone. This section describes several methods that have been developed to assist in determining acceptable risk values as summarized in Table 9.9.

Risk managers make decisions based on risk assessment and other considerations including economical, political, environmental, legal, reliability, producibility, safety, and other factors. The answer to the question "How safe is safe enough?" is difficult and constantly changing due to different perceptions and understandings of risk. To determine "acceptable risk," managers need to analyze alternatives for the best choice. In some industries, an acceptable risk has been defined by consensus. For example, the U.S. Nuclear Regulatory Commission requires that reactors be designed such that the probability of a large radioactive release to the environment from a reactor incident shall be less than 1×10^{-6} per year. Risk levels for certain carcinogens and pollutants have also been given acceptable concentration levels based on some assessment of acceptable risk. However, risk acceptance for many other activities are not stated.

For example, qualitative implications for risk acceptance are identified in the several existing maritime regulations. The International Maritime Organization High Speed Craft Code and the U.S. Coast Guard Navigation and Vessel Inspection Circular (NVIC) 5–93 for passenger submersible guidance both state that if the end effect is hazardous or catastrophic, a backup system and a corrective operating procedure is required. These references also state that a single failure must not result in a catastrophic event, unless the likelihood is extremely remote.

Often the level of risk acceptance with various activities is implied. Society has reacted to risks through the developed level of balance between risk and potential benefits. Measuring this balance of accepted safety levels for various risks provides a means for assessing society values. These threshold values of acceptable risk depend on a variety of issues including the activity type, industry, and users, and the society as a whole.

Target risk or reliability levels are required for developing procedures and rules for ship structures. For example, the selected reliability levels determine the probability of failure of structural components. The following three methods were used to select target reliability values:

1. Agreeing upon a reasonable value in cases of novel structures without prior history
2. Calibrating reliability levels implied in currently successfully used design codes
3. Choosing target reliability level that minimizes total expected costs over the service life of the structure for dealing with design for which failure results in only economic losses and consequences

The first approach can be based on expert-opinion elicitation. The second approach, called *code calibration*, is the most commonly used approach as it provides the means to build on previous experiences. For example, rules provided by classification and industry societies can be used to determine

TABLE 9.9 Methods for Determining Risk Acceptance

Risk Acceptance Method	Summary
Risk conversion factors	This method addresses the attitudes of the public about risk through comparisons of risk categories. It also provides an estimate for converting risk acceptance values between different risk categories
Farmer's curve	It provides an estimated curve for cumulative probability risk profile for certain consequences (e.g., deaths). It demonstrates graphical regions of risk acceptance/nonacceptance
Revealed preferences	Through comparisons of risk and benefit for different activities, this method categorizes society preferences for voluntary and involuntary exposure to risk.
Evaluation of magnitude of consequences	This technique compares the probability of risks to the consequence magnitude for different industries to determine acceptable risk levels based on consequence
Risk effectiveness	It provides a ratio for the comparison of cost to the magnitude of risk reduction. Using cost–benefit decision criteria, a risk reduction effort should not be pursued if the costs outweigh the benefits. This may not coincide with society values about safety
Risk comparison	The risk acceptance method provides a comparison between various activities, industries, etc., and is best suited to comparing risks of the same type

TABLE 9.10 Risk Conversion Values for Different Risk Factors

Risk Factors	Risk Conversion (RF) Factor	Computed RF Value*
Origin	Natural/human made	20
Severity	Ordinary/catastrophic	30
Volition	Voluntary/involuntary	100
Effect	Delayed/immediate	30
Controllability	Controlled/uncontrolled	5 to 10
Familiarity	Old/new	10
Necessity	Necessary/luxury	1
Costs	Monetary/nonmonetary	NA
Origin	Industrial/ Regulatory	NA
Media	Low profile/ high profile	NA

* NA, not available.

the implied reliability and risk levels in respective rules and codes, then target risk levels can be set in a consistent manner, and new rules and codes can be developed to produce future designs and vessels that are of similar levels that offer reliability and/or risk consistency. The third approach can be based on economic and trade-off analysis. In subsequent sections, the methods of Table 9.9 for determining risk acceptance are discussed.

Risk Conversion Factors. Analysis of risks shows that there are different taxonomies that demonstrate the different risk categories, often called *risk factors*. These categories can be used to analyze risks on a dichotomous scale comparing risks that invoke the same perceptions in society. For example, the severity category may be used to describe both ordinary and catastrophic events. Grouping events that could be classified as ordinary and comparing the distribution of risk to a similar grouping of catastrophic categories yields a ratio describing the degree of risk acceptance of ordinary events as compared to catastrophic events. The comparison of various categories determined the risk conversion values as provided in Table 9.10. These factors are useful in comparing the risk acceptance for different activities, industries, etc. By computing the acceptable risk in one activity, an estimate of acceptable risk in other activities can be calculated based on the risk conversion factors. A comparison of several common risks based on origin and volition is shown in Table 9.11.

Farmer's Curve. The Farmer's curve is graph of the cumulative probability vs. consequence for some activity, industry, or design, as shown in Figure 9.1 and Figure 9.2. This curve introduces a probabilistic approach in determining acceptable safety limits. Probability (or frequency) and consequence values are calculated for each level of risk generating a curve that is unique to hazard of concern. The area to the right (outside) of the curve is generally considered unacceptable because the probability and consequence values are higher than the average value delineated by the curve. The area to the left (inside) of the curve is considered acceptable because probability and consequence values are less than the estimated valve of the curve.

TABLE 9.11 Classification of Common Risks

		Voluntary		Involuntary	
Source	Size	Immediate	Delayed	Immediate	Delayed
Human	Catastrophic	Aviation		Dam failure building fire nuclear accident	Pollution building fire
Made	Ordinary	Sports boating automobiles	Smoking occupation carcinogens	Homicide	
Natural	Catastrophic			Earthquakes hurricanes tornadoes epidemics	
	Ordinary			Lighting animal bites	Disease

Method of Revealed Preferences. The method of revealed preferences provides a comparison of risk versus benefit and categorization for different risk types. The basis for this relationship is that risks are not taken unless there is some form of benefit. Benefit may be monetary or some other item of worth such as pleasure. The different risk types are for the risk category of voluntary versus involuntary actions as shown in Figure 9.16.

Magnitudes of Risk Consequence. Another factor affecting the acceptance of risk is the magnitude of consequence of the event that can result from some failure. In general, the larger the consequence, the less the likelihood that this event may occur. This technique has been used in several industries to demonstrate the location of the industry within societies' risk acceptance levels based on consequence magnitude as shown in Figure 9.17. Further evaluation has resulted in several estimates for the relationship.

January 09, 2006 9–45 between the accepted probability of failure and the magnitude of consequence for failure as provided by Allen in 1981 and called herein the CIRIA (Construction Industry Research and Information Association) equation:

$$P_{\mathrm{f}} = 10^{-4}\frac{KT}{n} \tag{9.8}$$

where T is the life of the structure, K is a factor regarding the redundancy of the structure, and n is the number of people exposed to risk. Another estimate is Allen's equation that is given by:

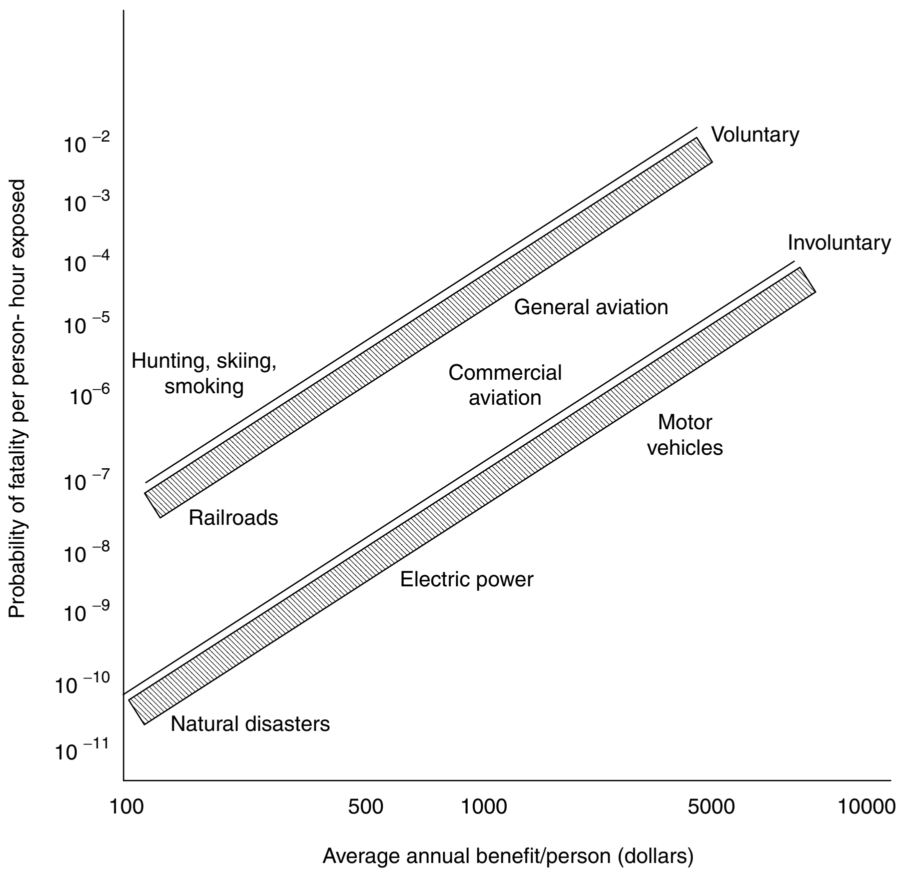


FIGURE 9.16 Accepted rist of voluntary and involuntary activities.

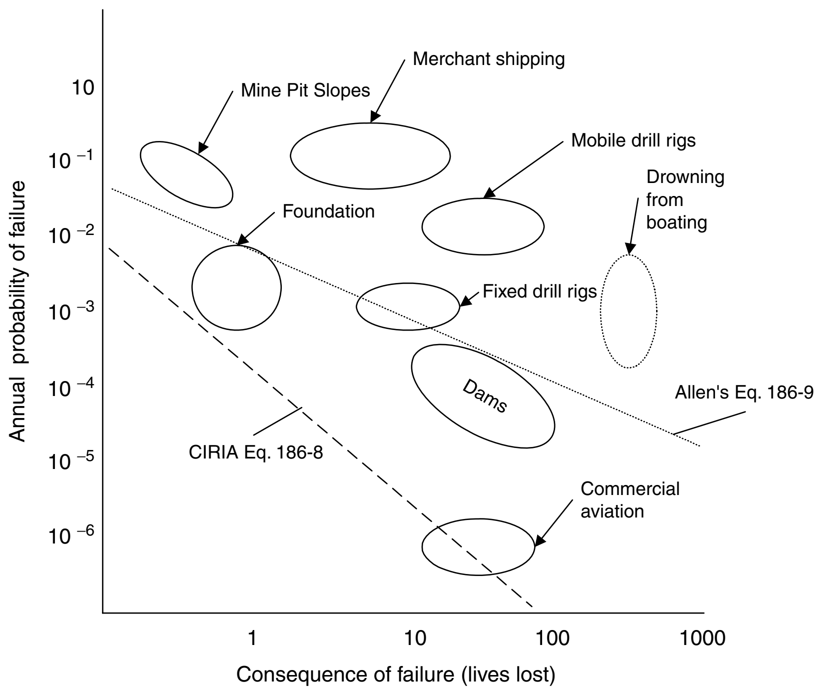


FIGURE 9.17 Comparison of risk and control costs.

$$P_f = 10^{-5} \frac{TA}{W\sqrt{n}} \tag{9.9}$$

where T is the life of the structure, n is the number of persons exposed to risk, and A and W are factors regarding the type and redundancy of the structure. Equation 9.8 offers a lower bound, whereas Equation 9.9 offers a middle line.

Risk Reduction Cost-Effectiveness Ratio. Another measuring tool to assess risk acceptance is the determination of risk reduction effectiveness:

$$\mathit{Risk\ Reduction\ Effectiveness} = \frac{\mathit{Cost}}{\Delta\ \mathit{Risk}} \tag{9.10}$$

where the cost should be attributed to risk reduction, and $\Delta\mathit{Risk}$ is the level of risk reduction as follows:

$$\Delta\ \mathit{Risk} = (\mathit{Risk\ before\ mitigation\ action}) - (\mathit{Risk\ after\ mitigation\ action}) \tag{9.11}$$

The difference in Equation 9.11 is also called the *benefit attributed to a risk reduction action*. Risk effectiveness can be used to compare several risk reduction efforts. The initiative with the smallest risk effectiveness provides the most benefit for the cost. Therefore, this measurement may be used to help determine an acceptable level of risk. The inverse of this relationship may also be expressed as cost-effectiveness.

Risk Comparisons. This technique uses the frequency of severe incidents to directly compare risks between various areas of interest to assist in justifying risk acceptance. Risks can be presented in different ways that can impact how the data are used for decisions. Often values of risk are manipulated in different forms for comparison reasons as demonstrated in Table 9.12. Comparison of risk values should be taken in the context of the values' origin and uncertainties involved.

TABLE 9.12 Ways to Identify Risk of Death

Ways to Identify Risk of Death	Summary
Number of fatalities	This measure shows the impact in terms of the number of fatalities on society. Comparison of these values is cautioned since the number of persons exposed to the particular risk may vary. Also, the time spent performing the activity may vary. Different risk category types should also be considered to compare fatality rates
Annual mortality rate/individual	This measure shows the mortality risk normalized by the exposed population. This measure adds additional information about the number of exposed persons; however, the measure does not include the time spent on the activity.
Annual mortality	This measure provides the most complete risk value since the risk is normalized by the exposed population and the duration of the exposure.
Loss of life exposure (LLE)	This measure converts a risk into a reduction in the expected life of an individual. It provides a good means of communicating risks beyond probability values.
Odds	This measure is a layman format for communicating probability, for example, 1 in 4.

This technique is most effective for comparing risks that invoke the same human perceptions and consequence categories. Comparing risks of different categories is cautioned because the differences between risk and perceived safety may not provide an objective analysis of risk acceptance. The use of risk conversion factors may assist in transforming different risk categories. Conservative guidelines for determining risk acceptance criteria can be established for voluntary risks to the public from the involuntary risk of natural causes.

9.1.4.2 Rankings Based on Risk Results

Another tool for risk management is the development of risk ranking. The elements of a system within the objective of analysis can be analyzed for risk and consequently ranked. This relative ranking may be based on the failure probabilities, failure consequences, risks, or other alternatives with concern towards risk. Generally, risk items ranked highly should be given high levels of priority; however, risk management decisions may consider other factors such as costs, benefits, and effectiveness of risk reduction measures. The risk ranking results may be presented graphically as needed.

9.1.4.3 Decision Analysis

Decision analysis provides a means for systematically dealing with complex problems to arrive at a decision. Information is gathered in a structured manner to provide the best answer to the problem. A decision generally deals with three elements: alternatives, consequences, and preferences. The alternatives are the possible choices for consideration. The consequences are the potential outcomes of a decision. Decision analysis provides methods for quantifying preference tradeoffs for performance along multiple decision attributes while taking into account risk objectives. Decision attributes are the performance scales that measure the degree to which objectives are satisfied. For example, one possible attribute is reducing lives lost for the objective of increasing safety. Additional examples of objectives may include minimize the cost, maximize utility, maximize reliability, and maximize profit. The decision outcomes may be affected by uncertainty; however, the goal is to choose the best alternative with the proper consideration of uncertainty. The analytical depth and rigor for decision analysis depends on the desired detail in making the decision. Cost–benefit analysis, decision trees, influence diagrams and the analytic hierarchy process are some of the tools to assist in decision analysis. Also, decision analysis should consider constraints, such as availability of system for inspection, availability of inspectors, preference of certain inspectors, and availability of inspection equipment.

9.1.4.4 Cost–Benefit Analysis

Risk managers commonly weigh various factors including cost and risk. The analysis of three different alternatives is shown graphically in Figure 9.18 as an example. The graph shows that alternative (C) is the best choice because the level of risk and cost is less than alternatives (A) and (B). However, if the only alternatives were A and B, the decision would be more difficult. Alternative (A) has higher cost and lower risk than alternative (B); alternative (B) has higher risk but lower cost than alternative (A). A risk manager needs to weigh the importance of risk and cost in making this decision and availability of resources, and make use of risk-based decision analysis.

Risk–benefit analysis can also be used for risk management. Economic efficiency is important to determine the most effective means of expending resources. At some point, the costs for risk reduction do not provide adequate benefit. This process compares the costs and risk to determine where the optimal risk value is on a cost basis. This optimal value occurs, as shown in Figure 9.19, when costs to control risk are equal to the risk cost due to the consequence (loss). Investing resources to reduce low risks below this equilibrium point is not providing a financial benefit. This technique may be used when cost values can be attributed to risks. This analysis might be difficult to perform for certain risk such as risk to human health and environmental risks because the monetary values are difficult to estimate for human life and the environment.

The present value of incremental costs and benefits can be assessed and compared among alternatives that are available for risk mitigation or system design. Several methods are available to determine which, if any, option is most worth pursuing. In some cases, no alternative will generate a net benefit relative to the base case. Such a finding would be used to argue for pursuit of the base case scenario. The following are the most widely used present value comparison methods: (1) net present value (NPV), (2) benefit-cost ratio, (3) internal rate of return, and (4) payback period. The net present value (NPV) method requires that each alternative need to meet the following criteria to warrant investment of funds: (1) having a positive NPV; and (2) having the highest NPV of all alternatives considered. The first condition insures that the alternative is worth undertaking relative to the base case, e.g., it contributes more in incremental benefits than it absorbs in incremental costs. The second condition insures that maximum benefits are obtained in a situation of unrestricted access to capital funds. The NPV can be calculated as follows:

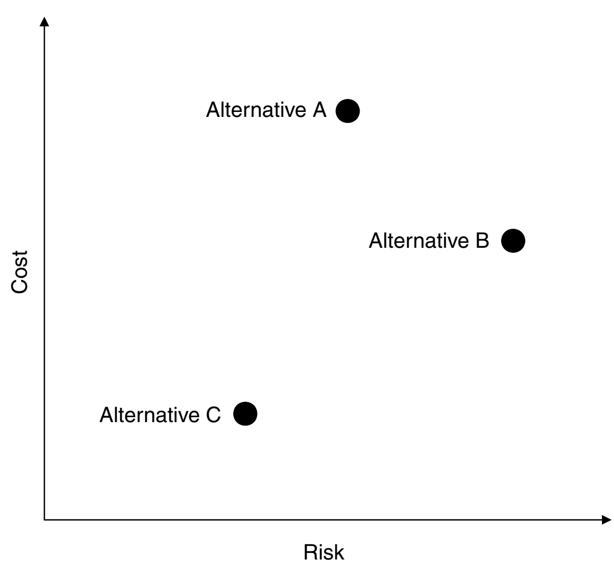


FIGURE 9.18 Risk benefit for three alternatives. (From CERT.)

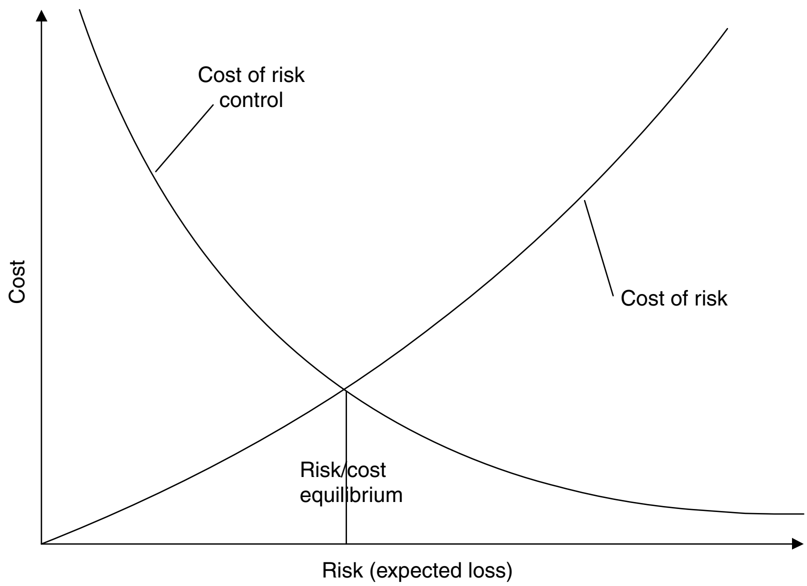


FIGURE 9.19 Comparison of risk and control costs. (From CIAO 2001.)

$$\text{NPV} = \sum_{t=0}^{k} \frac{(B-C)_t}{(1+r)^t} = \sum_{t=0}^{k} \frac{B_t}{(1+r)^t} - \sum_{t=0}^{k} \frac{C_t}{(1+r)^t} \tag{9.12}$$

where B is future annual benefits in constant dollars, C is future annual costs inconstant dollars, r is annual real discount rate, k is number of years from the base year over which the project will be evaluated, and t is an index running from 0 to k representing the year under consideration.

The benefit of a risk mitigation action can be assessed as follows:

$$\text{Benefit} = \text{unmitigated risk} - \text{mitigated risk}. \tag{9.13}$$

The cost in Equation 9.13 is the cost of the mitigation action. The benefit minus the cost of mitigation can be used to justify the allocation of resources. The benefit-to-cost ratio can be computed, and may also be helpful in decision making. The benefit-to-cost ratio (B/C) can be computed as

$$\text{Benefit-to-Cost Ratio } (B/C) = \frac{\text{Benefit}}{\text{Cost}} = \frac{\text{Unmitigated Risk} - \text{Mitigated Risk}}{\text{Cost of Mitigation Action}} \tag{9.14}$$

Cost of mitigation action ratios greater than one are desirable. In general, the larger the ratio, the better the mitigation action.

Accounting for the time value of money would require defining the benefit–cost ratio as the present value of benefits divided by the present value of costs. The benefit–cost ratio can be calculated as follows:

$$B/C = \frac{\sum_{t=0}^{k} \frac{B_t}{(1+r)^t}}{\sum_{t=0}^{k} \frac{C_t}{(1+r)^t}} \tag{9.15}$$

where Bt is future annual benefits in constant dollars, Ct is future annual costs inconstant dollars, r is annual real discount rate, and t is an index running from 0 to k representing the year under consideration. A proposed activity with a B/C ratio of discounted benefits to costs of 1 or more is expected to return at least as much in benefits as it costs to undertake, indicating that the activity is worth undertaking.

The internal rate of return (*IRR*) is defined as the discount rate that makes the present value of the stream of expected benefits in excess of expected costs zero. In other words, it is the highest discount rate at which the project will not have a negative *NPV*. To apply the *IRR* criterion, it is necessary to compute the IRR and then compare it with a base rate of, say, a 7% discount rate. If the real *IRR* is less than 7%, the project would be worth undertaking relative to the base case. The *IRR* method is effective in deciding whether or not a project is superior to the base case; however it is difficult to utilize it for ranking projects and deciding among mutually exclusive alternatives. Project rankings established by the *IRR* method might be inconsistent with those of the *NPV* criterion. Moreover, a project might have more than one *IRR* value, particularly when a project entails major final costs, such as cleanup costs. Solutions to these limitations exist in capital budgeting procedures and practices that are often complicated or difficult to employ in practice and present opportunities for error.

The payback period measures the number of years required for net undiscounted benefits to recover the initial investment in a project. This evaluation method favors projects with near-term and more certain benefits, and fails to consider benefits beyond the payback period. The method does not provide information on whether an investment is worth undertaking in the first place.

The previous models for cost–benefit analysis presented in this section do not account for the full probabilistic characteristics of B and C in their treatment. Concepts from reliability assessment 4 can be used for this purpose. Assuming B and C to normally distributed, a benefit–cost index ($\beta_{B/C}$) can be defined as follows:

$$\beta_{B/C} = \frac{\mu_B - \mu_C}{\sqrt{\sigma_B^2 + \sigma_C^2}} \tag{9.16}$$

where μ and σ are the mean and standard deviation. The failure probability can be computed as

$$P_{\mathrm{f},B/C} = P(C > B) = 1 - \Phi(\beta) \tag{9.17}$$

In the case of lognormally distributed B and C, the benefit–cost index ($\beta_{B/C}$) can be computed as

$$\beta_{B/C} = \frac{\ln\left(\frac{\mu_\mathrm{B}}{\mu_\mathrm{C}}\sqrt{\frac{\delta_C^2+1}{\delta_B^2+1}}\right)}{\sqrt{\ln[(\delta_B^2+1)(\delta_C^2+1)]}} \tag{9.18}$$

where δ is the coefficient of variation. Equation (9.18) also holds for the case of lognormally distributed B and C. In the case of mixed distributions or cases involving basic random variables of B and C, the advanced second moment method or simulation method can be used. In cases where benefit is computed as revenue minus cost, benefit might be correlated with cost requiring the use of other methods.

Example 9.2 Protection of Critical Infrastructure

This example is used to illustrate the cost of cost–benefit analysis using a simplified decision situation. As an illustration, assume that there is a 0.01 probability of an attack on a facility containing hazardous material during the next year. If the attack occurs, the probability of a serious release to the public is 0.01 with a total consequence of \$100B. The total consequence of an unsuccessful attack is negligible. The unmitigated risk can therefore be computed as

$$\textit{Unmitigated Risk} = 0.01(\$0.01)(100\text{B}) = \ \$\ 10\text{M}.$$

If armed guards are deployed at each facility, the probability of attack can be reduced to 0.001 and the probability of serious release if an attack occurs can be reduced to 0.001. The cost of the guards for all plants is assumed to be \$100 M per year. The mitigated risk can therefore be computed as

$$\textit{Mitigated Risk} = 0.001(0.001)(\$100\text{B}) = \$0.10\text{M}.$$

The benefit in this case is

$$\textit{Benefit} = \$10\text{M} - \$0.1\text{M}\ \textit{or}\ \sim\$10\text{M}.$$

The benefit-to-cost ratio is about 0.1. Therefore, the \$100 M cost might be difficult to justify.

9.1.4.5 Risk Mitigation

A risk mitigation strategy can be presented from a financial point of view. Risk mitigation in this context can be defined as an action to either reduce the probability of an adverse event occurring or to reduce the adverse consequences if it does occur. This definition captures the essence of an effective management process of risk. If implemented correctly a successful risk mitigation strategy should reduce any adverse (or downside) variations in the financial returns from a project, which are usually measured by either (1) the net present value (NPV) defined as the difference between the present value of the cash flows generated by a project and its capital cost and calculated as part of the process of assessing and appraising investments, or (2) the internal rate of return (IRR) defined as the return that can be earned on the capital invested in the project, i.e., the discount rate that gives an NPV of zero, in the form of the rate that is equivalent to the yield on the investment.

Risk mitigation involves direct costs like increased capital expenditure or the payment of insurance premiums; hence might reduce the average overall financial returns from a project. This reduction is often a perfectly acceptable outcome, given the risk aversion of many investors and lenders. A risk mitigation strategy is the replacement of an uncertain and volatile future with one where there is less exposure to adverse risks and so less variability in the return, although the expected NPV or IRR may be reduced. These two aspects are not necessarily mutually exclusive. Increasing risk efficiency by simultaneously improving the expected NPV or IRR and simultaneously reducing the adverse volatility is sometimes possible and should be sought. Risk mitigation should cover all phases of a project from inception to closedown or disposal.

Four primary ways are available to deal with risk within the context of a risk management strategy as follows:

- Risk reduction or elimination
- Risk transfer, e.g., to a contractor or an insurance company
- Risk avoidance
- Risk absorbance or pooling

These four methods are described in subsequent sections.

Risk Reduction or Elimination. Risk reduction or elimination is often the most fruitful form for exploration. For example, could a design of a system be amended so as to reduce or eliminate either the probability of occurrence of a particular risk event or the adverse consequences if it occurs? Alternatively, could the risks be reduced or eliminated by retaining the same design but using different materials or a different method of assembly? Other possible risk mitigation options in this category include as examples: a better labor relations policy to minimize the risk of stoppages, training of staff to avoid hazards, better site security to avoid theft and vandalism, a preliminary investigation of possible site pollution, advance ordering of key components, noise abetment measures, good signposting, and liaisons with the local community.

Risk Transfer. A general principle of an effective risk management strategy is that commercial risks in projects and other business ventures should be borne wherever possible by the party that is best able to manage them, and thus mitigate the risks. Contracts and financial agreements are the principal forms to transfer risks. Companies specializing in risk transfer can be consulted that could appropriately meet the needs of a project. Risks can be transferred alternately to an insurance company which, in return for a payment (i.e., premium) linked to the probability of occurrence and severity associated with the risk, is obliged by the contract to offer compensation to the party affected by the risk. Insurance coverage can range from straight insurance for expensive risks with a low probability, such as fire, through performance bonds, which ensure that the project will be completed if the contractor defaults, to sophisticated financial derivatives such as hedge contracts to avoid such risks as unanticipated losses in foreign exchange markets.

Risk Avoidance. A most intuitive way of avoiding a risk is to avoid undertaking the project in a way that involves that risk. For example, if the objective is to generate electricity but a nuclear power source, although cost-efficient, is considered to have a high risk due to potentially catastrophic consequences, even after taking all reasonable precautions, the practical solution is to turn to other forms of fuel to avoid that risk. Another example would be the risk that a particularly small contractor would go bankrupt. In this case, the risk could be avoided by using a well-established contractor for that particular job.

Risk Absorbance and Pooling. Cases where risks cannot, or cannot economically, be eliminated, transferred, or avoided, they must be absorbed if the project is to proceed. Normally, a sufficient margin in the project's finances needs to be created to cover the risk event should it occur. However, it is not always essential for one party alone to bear all these absorbed risks. Risks can be reduced through pooling possibly through participation in a consortium of contractors, when two or more parties are able to exercise partial control over the incidence and impact of risk. Joint ventures and partnerships are other examples of organizational forms for pooling risks.

Uncertainty Characterization. Risk can be mitigated through proper uncertainty characterization. The presence of improperly characterized uncertainty could lead to higher adverse event occurrence likelihood and consequences. Also, it could result in increasing estimated cost margins as a means of compensation. Therefore, risk can be reduced by a proper characterization of uncertainty. The uncertainty characterization can be achieved through data collection and knowledge construction.

9.1.5 Risk Communication

Risk communication can be defined as an interactive process of exchange of information and opinion among stakeholders such as individuals, groups, and institutions. It often involves multiple messages about the nature of risk or expressing concerns, opinions, or reactions to risk managers or to legal and institutional arrangements for risk management. Risk communication greatly affects risk acceptance and defines the acceptance criteria for safety.

Risk communication provides the vital link between the risk assessors, risk managers, and the public to understand risk. However, this does not necessarily mean that risk communication will always lead to agreement among different parties. An accurate perception of risk provides for rational decision making. The Titanic was deemed an unsinkable ship, yet was lost on its maiden voyage. Space shuttle flights were perceived to be safe enough for civilian travel until the Space Shuttle Challenger disaster. These disasters obviously had risks that were not perceived as significant until after the disaster. Risk communication is a dynamic process that must be considered prior to management decisions.

The communication process deals with technical information about controversial issues. Therefore, it needs to be skillfully performed by risk managers and communicators who might be viewed as adversaries to the public. Risk communication between risk assessors and risk managers is necessary to effectively apply risk assessments in decision making. Risk managers must participate in determining the criteria for determining what risk is acceptable and unacceptable. This communication between the risk managers and risk assessors is necessary for a better understanding of risk analysis in making decisions.

Risk communication also provides the means for risk managers to gain acceptance and understanding by the public. Risk managers need to go beyond the risk assessment results and consider other factors in making decisions. One of these concerns is politics, which is largely influenced by the public. Risk managers often fail to convince the public that risks can be kept to acceptable levels. Problems with this are shown by the public's perception of toxic waste disposal and nuclear power plant operation safety. As a result of the public's perceived fear, risk managers may make decisions that are conservative to appease the public.

The value of risk calculated from risk assessment is not the only consideration for risk managers. All risks are not created equal and society has established risk preferences based on public preferences. Decision makers should take these preferences into consideration when making decisions concerning risk.

To establish a means of comparing risks based on the society preferences, risk conversion factors (RCF) may be used. The RCF expresses the relative importance of different attributes concerning risk. An example of possible risk conversion factors is shown in Table 9.10. These values were determined by inferences of public preferences from statistical data with the consequence of death considered.

For example, the voluntary and involuntary classification depends on whether the events leading to the risk are under the control of the persons at risk or not, respectively. Society, in general, accepts a higher level of voluntary risk than involuntary risk by an estimated factor of 100. Therefore, an individual will accept a voluntary risk that is 100 times greater than an involuntary risk.

The process of risk communication can be enhanced and improved in three aspects: (1) the process, (2) the message, and (3) the consumers. The risk assessment and management process needs to have clear goals with openness, balance, and competence. The contents of the message should account for audience orientation and uncertainty, provide risk comparison, and be complete. There is a need for consumer's guides that introduce risks associated with a specific technology, the process of risk assessment and management, acceptable risk, decision making, uncertainty, costs and benefits, and feedback mechanisms. Improving risk literacy of consumers is an essential component of the risk communication process.

The USACE has a 1992 Engineering Pamphlet (EP) on risk communication (EP 1110-2-8). The following are guiding considerations in communicating risk:

- Risk communication must be free of jargon
- Consensus of expert needs to be established
- Materials must be cited, and their sources must be credible
- Materials must be tailored to audience
- The information must be personalized to the extent possible
- Motivation discussion should stress a positive approach and the likelihood of success
- Risk data must be presented in a meaningful manner

9.2 Electricity Infrastructure Security

Massoud Amin

9.2.1 Introduction

The massive power outages in the United States, Canada, U.K. and Italy in 2003 underscored electricity infrastructure's vulnerabilities (President's Commission on Critical Infrastructure Protection 1997; Amin 2000, 2001, 2002a, 2002b, 2003; U.S. Department of Energy 2002; Energy Information Administration 2003; North American Electric Reliability Council; EPRI 2003). The North American power network may be considered to be the largest and most complex machine in the world; its transmission lines connect all the electric generation and distribution on the continent. In that respect, it exemplifies many of the complexities of electric power infrastructure and how technological innovation combined with efficient markets and enabling policies can address them. This network represents an enormous investment, including more than 15,000 generators in 10,000 power plants, and hundreds of thousands of miles of transmission lines and distribution networks, whose estimated worth is over US$800 billion. In 2000, transmission and distribution was valued at US$358 billion (EPRI 1999, 2000, 2001, 2003a, 2003b; Hauer and Dagle 1999; Energy Information Administration 2003; North American Electric Reliability Council).

Through the North American electricity infrastructure, every user, producer, distributor, and broker of electricity buys and sells, competes, and cooperates in an "electric enterprise." Every industry, every business, every store, and every home is a participant, active or passive, in this continent-scale conglomerate. Over the last decade and during the next few years, the electric enterprise will undergo dramatic transformation as its key participants—the traditional electric utilities—respond to deregulation, competition, renewable energy portfolio standards, tightening environmental/land-use restrictions, and other global trends.

However, this network has evolved without formal analysis of the system-wide implications of this evolution, including its diminished transmission and generation shock absorber capacity under the forces of deregulation, the digital economy, and interaction with other infrastructures. Only recently, with the advent of deregulation, unbundling, and competition in the electric power industry, has the possibility of power delivery beyond neighboring areas become a key design and engineering consideration, yet the existing grid is still expected to handle a growing volume and variety of long-distance, bulk power transfers. To meet the needs of a pervasively digital world that relies on microprocessor-based devices in vehicles, homes, offices, and industrial facilities, grid congestion and atypical power flows are increasing, as are customer reliability expectations.

The vulnerability of the energy infrastructure to natural disasters and physical attacks has long been recongnized, but this vulnerability has significantly increased in recent years because (1) relatively little infrastructure (especially transmission lines) has been added to handle increased demand with an adequate safety "cushion" (EPRI 2001, NERC 2001) and (2) terrorists might in fact be planning attacks on the system (EPRI 2001, EEI 2003, DOE and DHS 2006).

Ongoing efforts at NERC, DOE, and EPRI—including EPRI's Enterprise Information Security Program and Infrastructure Security Initiative—have highlighted utility-specific threats and the technologies to counteract them. As an example, the Infrastructure Security Assessment, developed by EPRI in response to the September 11, 2001, terrorist attacks, discusses some of the specific threats and recommends counter measures; as noted in these and other reports (EPRI 2001, 2004, 2005, EEI 2003, NERC, 2001, 2004, 2006, DOE and DHS 2006), the existing power delivery system is vulnerable to natural disasters and intentional attack. Regarding the latter, a successful terrorist attempt to disrupt the power delivery system could have adverse effects on national security, the economy, and the lives of every citizen.

The existing power delivery system is vulnerable to natural disasters and intentional attack. Regarding the latter, a successful terrorist attempt to disrupt the power delivery system could have adverse effects on national security, the economy, and the lives of every citizen.

Both the importance and difficulty of protecting power systems have long been recognized. In 1990, the Office of Technology Assessment (OTA) of the U.S. Congress issued a detailed report, *Physical Vulnerability of the Electric System to Natural Disasters and Sabotage,* concluding: "Terrorists could emulate acts of sabotage in several other countries and destroy critical [power system] components, incapacitating large segments of a transmission network for months. Some of these components are vulnerable to saboteurs with explosives or just high-powered rifles." The report also documented the potential cost of widespread outages, estimating them to be in the range of $1 to $5/kWh of disrupted service, depending on the length of outage, the types of customers affected, and a variety of other factors, such as the unintended release of pollutants/wastewater and the corresponding environmental cost, social unrest, damage from looting and arson as components of the total cost.

The reality of a coordinated attack has raised the issue of security to be considered along with power systems' reliability, which posits more random and independent failures. The system's vulnerability to natural disasters and physical attacks has long been recognized, but this vulnerability has significantly increased in recent years, in part because the system is operating closer to its capacity and in part because terrorist attacks are no longer hypothetical.

The situation has become even more complex because accounting for all critical assets includes thousand of transformer, line reactors, series capacitors, and transmission lines. Protection of ALL the widely diverse and dispersed assets is impractical because there are so many involved:

- Over 230,000 miles of HV lines (230 kV and above)
- Over 6,644 transformers in the Eastern Interconnection
- Over 6,000 HV transformers in the North American Interconnection (of which less than 300 are critical assets)
- Control centers
- Interdependence with gas pipelines
- Compressor stations
- Dams
- Rail lines
- Telecommunication equipment (monitoring and control of the system)

In this section, the security, agility, and robustness/survivability of large-scale power delivery infrastructure that face new threats and unanticipated conditions is presented. In addition, focus is placed, in part, on challenges associated with development of a smart self-healing electric power grid for enhanced system security, reliability, efficiency, and quality.

As an example, trends show that worldwide cyber attacks are on the rise; the number of documented attacks and intrusions has been rising very rapidly in recent years. Due to the increasingly sophisticated nature and speed of some malicious code, intrusions, and denial-of-service attacks, human response may be inadequate. Some trends and documented nonsensitive modes of attack are shown in Figure 9.20 and Figure 9.21.

The vulnerability of the energy infrastructure to natural disasters and physical attacks has long been recognized, but this vulnerability has significantly increased in recent years because relatively little infrastructure (especially transmission lines) has been added to handle increased demand with an adequate safety "cushion."[1]

Secure and reliable operation of these systems is fundamental to national and international economy, security and quality of life. Their very interconnectedness makes them more vulnerable to global disruption, initiated locally by material failure, natural calamities, intentional attack, or human error.

In addition to the security problems associated with electric transmission systems additional concerns arise with nuclear power systems. Nuclear power serves the United States as well as many countries around the world as a fairly reliable source of electricity that will not contribute to global warming producing carbon dioxide. In that sense it is a "clean" source of electricity. But there are ways in which nuclear power can be a hazardous technology.

There are two pressing safety concerns with nuclear power in this age of terrorism. The first has to do with the spent fuel storage areas located outside and near the reactors. These spent fuel "swimming pools" are laden with a far higher inventory of radioactive material than is present in the reactor itself and are far less protected from attack than the reactor. If a "9/11" plane had hit the Indian Point Nuclear Plant spent fuel storage area (about 35 miles from New York City), it is possible that millions of people would have had to be evacuated to escape radiation and that parts of New York would have had to be abandoned. There might of course be solutions to avoid this catastrophic scenario, but they would add to the cost of nuclear power.

The other long-term safety concern is the eventual transportation of thousands of high-level nuclear waste canisters across open roads of the United States (through 43 states) on their way to a final resting place, e.g., Yucca Mountain. These shipments would be enticing targets for terrorists. Technology may be able to ameliorate this type of accident, but it is currently not in place.

[1] NERC 2001, "Reliability Assessment 2001–2010"

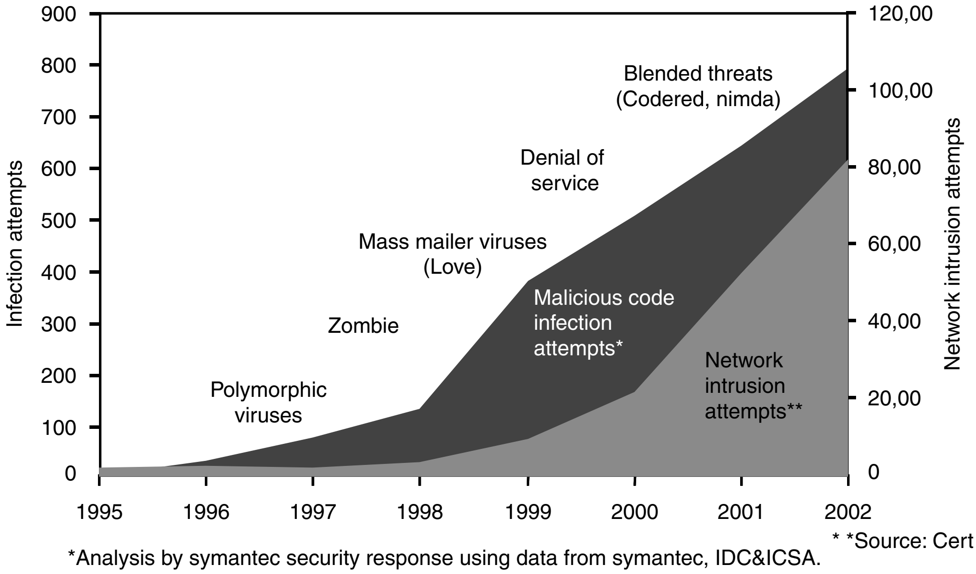


FIGURE 9.20 Attack trends: Documented network intrusion and infection attempts. (From CERT.)

9.2.2 The Electricity Enterprise: Today and Tomorrow

The North American power network's transmission lines connect all generation and distribution on the continent to form a vertically integrated hierarchical network. The question is raised as to whether there is a unifying paradigm for the simulation, analysis, and optimization of time-critical operations (both financial transactions and actual physical control) in these multiscale, multicomponent, and distributed systems. In addition, mathematical models of interactive networks are typically vague (or may not even exist); moreover, existing and classical methods of solution are either unavailable, or are not

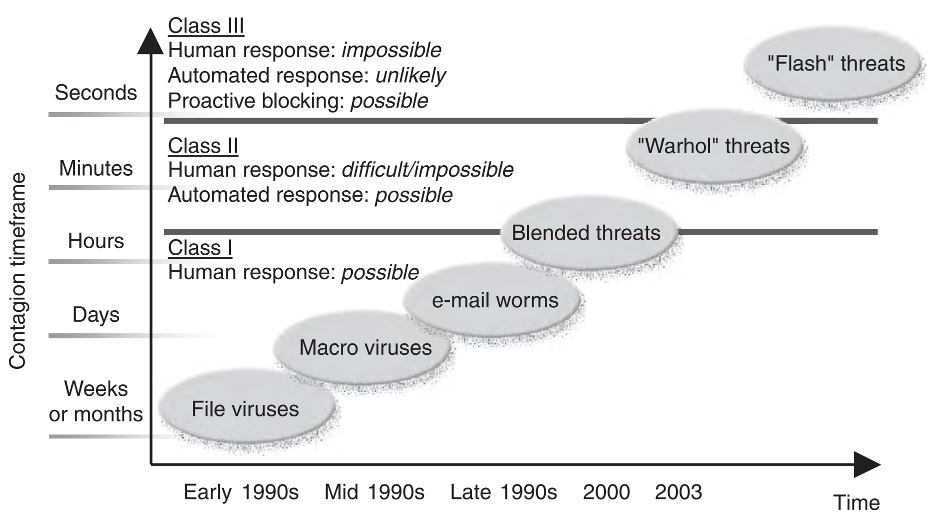


FIGURE 9.21 Malicious codes' threat evolution: Increased sophistication and much faster than a few years ago. (From CIAO 2001.)

sufficiently powerful. For the most part, no present methodologies are suitable for understanding their behavior.

Another important dimension is the effect of deregulation and economic factors on a particular infrastructure. Although other and more populous countries, such as China and India, will have greater potential electricity markets and demands, the United States is presently the largest national market for electric power. Its electric utilities have been mostly privately owned, vertically integrated, and locally regulated. National regulations in areas of safety, pollution, and network reliability also constrain their operations to a degree, but local regulatory bodies, mostly at the state level, have set their prices and their return on investment, and have controlled their investment decisions while protecting them from outside competition. That situation is now rapidly changing; state regulators are moving toward permitting and encouraging a competitive market in electric power.

The electric power grid was historically operated by separate utilities, each independent in its own control area and regulated by local bodies, to deliver bulk power from generation to load areas reliably and economically. As a noncompetitive, regulated monopoly, emphasis was on reliability (and security) at the expense of economy. Competition and deregulation have created multiple energy producers that must share the same regulated energy delivery network. Traditionally, new delivery capacity would be added to handle load increases, but because of the current difficulty in obtaining permits and the uncertainty about achieving an adequate rate of return on investment, total circuit miles added annually are declining while total demand for delivery resources continues to grow. In recent years, the "shock absorbers" have been shrinking; e.g., during the 1990s actual demand in the United States increased some 35%, while capacity increased only 18%, the most visible parts of a larger and growing U.S. energy crisis that is the result of years of inadequate investments in the infrastructure. According to EPRI analyses, since 1995 to the present the amortization/depreciation rate exceeds utility construction expenditures (Figure 9.22).

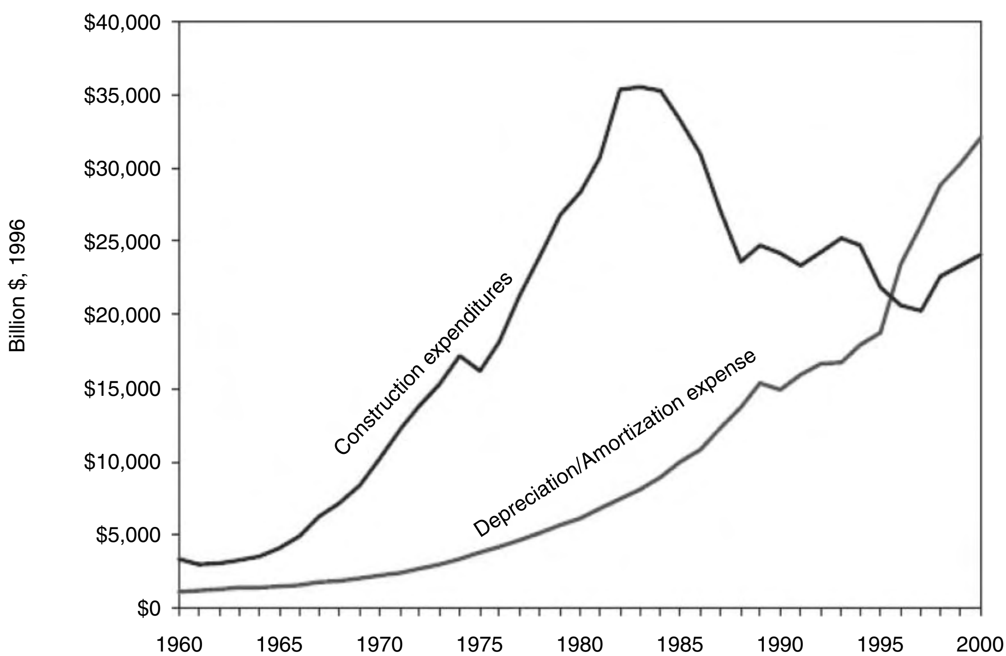


FIGURE 9.22 Since the "crossover" point in about 1995 utility construction expenditures have lagged behind asset depreciation. This has resulted in a mode of operation of the system analogous to "harvesting the farm far more rapidly then planting new seeds." (Data provided by EEI and graph courtesy of EPRI.)

As a result of these "diminished shock absorbers," the network is becoming increasingly stressed, and whether the carrying capacity or safety margin will exist to support anticipated demand is in question. The complex systems used to relieve bottlenecks and clear disturbances during periods of peak demand are at great risk to serious disruption, creating a critical need for technological improvements.

9.2.3 Reliability Issues

Several cascading failures during the past 40 years spotlighted the need to understand the complex phenomena associated with power network systems and the development of emergency controls and restoration. Widespread outages and huge price spikes during the past few years raised public concern about grid reliability at the national level (President's Commission on Critical Infrastructure Protection, 1997; Hauer and Dagle 1999; U.S. Department of Energy 2002; Energy Information Administration, 2003; House Committee on Energy and Commerce, 2003; North American Electric Reliability Council). According to data from the North American Electric Reliability Council (NERC) and analyses from the Electric Power Research Institute (EPRI), average outages from 1984 to the present have affected nearly 700,000 customers per event annually. Smaller outages occur much more frequently and affect tens to hundreds of thousands of customers every few weeks or months, whereas larger outages occur every two to nine years and affect millions. Much larger outages affect seven million or more customers per event each decade. These analyses are based on data collected for the U.S. Department of Energy (DOE), which requires electric utilities to report system emergencies that include electric service interruptions, voltage reductions, acts of sabotage, unusual occurrences that can affect the reliability of bulk power delivery systems, and fuel problems (North American Electric Reliability Council; Hauer and Dagle 1999; Amin 2000; 2001, 2002b, 2003; Energy Information Administration 2003; Amin 2005).

Coupling these analyses with diminished infrastructure investments, and noting that the crossover point for the utility construction investment vs. depreciation occurred in 1995 (Figure 9.20), the number and frequency of major outages along with the number of customers affected during the decade 1991–2000 were analyzed. The data were split into the two time periods: 1991–1995 and 1996–2000 (Figure 19.21). Based on EPRI's analyses (Amin 2003; EPRI 2003) of data in NERC's Disturbance Analysis Working Group (DAWG) database (Amin 2003; Energy Information Administration 2003; North American Electric Reliability Council), 41% more outages affected 50,000 or more consumers in the second half of the 1990s than in the first half (58 outages in 1996–2000 versus 41 outages in 1991–1995). The average outage affected 15% more consumers from 1996 to 2000 than from 1991 to 1995 (average size per event was 409,854 customers affected in the second half of the decade versus 355,204 in the first half of the decade). In addition, there were 76 outages of size 100 MW or more in the second half of the decade, compared to 66 such occurrences in the first half. During the same period, the average lost load caused by an outage increased by 34%, from 798 MW from 1991 to 1995 to 1067 MW from 1996 to 2000 (Figure 9.23) (Amin, 2003; Energy Information Administration, 2003; North American Electric Reliability Council; EPRI 2003).

Electric power utilities typically own and operate at least parts of their own telecommunications systems which often consist of backbone fiber-optic or microwave connecting major substations, with spurs to smaller sites. Increased use of electronic automation raises significant issues regarding the adequacy of operational security. As is true of other critical infrastructures, increased use of automated technologies raises significant security issues, however:

- Reduced personnel at remote sites makes the sites more vulnerable to hostile threats;
- Interconnecting automation and control systems with public data networks makes them accessible to individuals and organizations, from any worldwide location using an inexpensive computer and a modem; and
- Use of networked electronic systems for metering, scheduling, trading, or e-commerce imposes numerous financial risks associated with network failures.

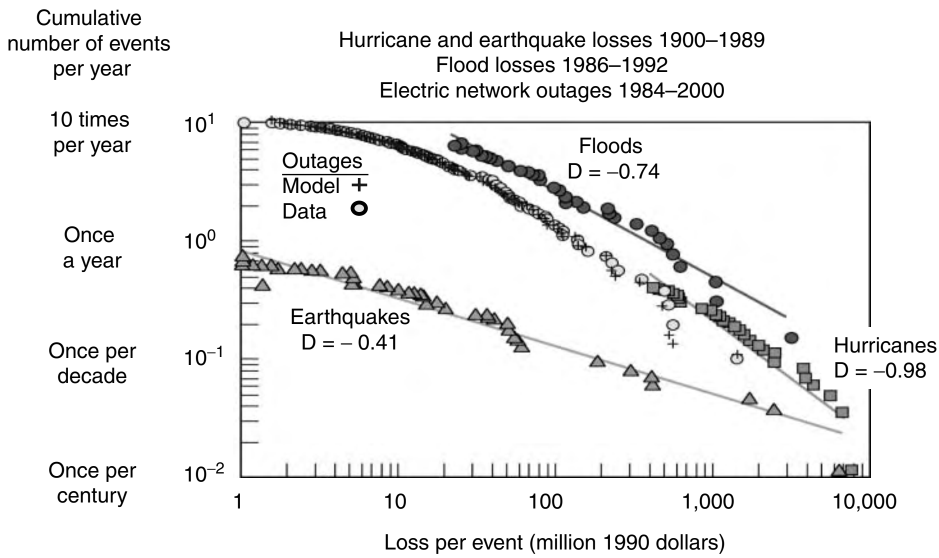


FIGURE 9.23 Understanding complex systems and global dynamics. Economic losses from disasters were found to follow a power law distribution—for hurricanes, floods, earthquakes, and even electrical outages. Fundamental power law distributions also were found for forest fires, Internet congestion, and other systems. CIN/SI results such as these translate in new approaches for optimizing complex systems in terms of productivity and robustness to disaster. Our goal is to move the power outage curve down toward the origin, i.e., to make outages less frequent and with smaller impact on customers. (From The EPRI/DoD Complex Interactive Networks/Systems Initiative [CIN/SI].)

In what follows, a brief overview of some key areas is given and selected security aspects of operational systems are presented, without discussing potentially sensitive material; these aspects include:

- Operational systems rely very heavily on the exchange of information amongst disparate systems
- Utilities rely on very extensive private and leased telecommunication systems
- Networking of these systems is expanding rapidly
- This networking is expanding beyond utility doors, to encompass other utilities, corporations, and customers
- Standard communication protocols and integration techniques are a MUST, despite the increased security risks
- Increased security concerns in the aftermath of tragic events of September 11, 2001
- Deregulation is increasing the incentives for unauthorized access to information

9.2.4 Infrastructures Under Threat

The terrorist attacks of September 11 have exposed critical vulnerabilities in America's essential infrastructures: Never again can the security of these fundamental systems be taken for granted. Electric power systems constitute the fundamental infrastructure of modern society. A successful terrorist attempt to disrupt electricity supplies could have devastating effects on national security, the economy, and the lives of every citizen. Yet power systems have widely dispersed assets that can never be absolutely defended against a determined attack.

Because critical infrastructures touch us all, the growing potential for infrastructure problems stems from multiple sources. These sources include system complexity, deregulation, economic effects, power market impacts, terrorism, and human error. The existing power system is also vulnerable to natural disasters and intentional attacks. Ongoing efforts at NERC, DOE, and EPRI—including EPRI's

Enterprise Information Security Program and Infrastructure Security Initiative—have highlighted utility-specific threats and the technologies to counteract them. Regarding the latter, a November 2001 EPRI assessment developed in response to the September 11, 2001 attacks highlights three different kinds of potential threats to the U.S. electricity infrastructure (Amin, 2001, 2002a, 2003, EPRI 2001):

- Attacks upon the power system. In this case, the electricity infrastructure itself is the primary target, with ripple effects, in terms of outages, extending into the customer base. The point of attack could be a single component, such as a critical substation, or a transmission tower. However, there could also be a simultaneous, multipronged attack intended to bring down the entire grid in a region of the United States. Similarly, the attack could target electricity markets, which, because of their transitional status, are highly vulnerable.
- Attacks by the power system. In this case, the ultimate target is the population, using parts of the electricity infrastructure as a weapon. Power plant cooling towers, for example, could be used to disperse chemical or biological agents.
- Attacks through the power system. In this case, the target is the civil infrastructure. Utility networks include multiple conduits for attack, including lines, pipes, underground cables, tunnels and sewers. An electromagnetic pulse, for example, could be coupled through the grid to with the intention of damaging computer and/or telecommunications infrastructure.

Protection against all three modes of attack thus represents a potentially critical "showstopper" for realizing a digital economy and enabling customer-managed service networks. Indeed, if infrastructure security is not ensured, even maintaining current levels of productivity and service will be jeopardized. Conversely, deploying some of the advanced technologies needed to enhance security will have a positive effect on efforts to improve grid reliability and coordinate power system operations with those of other infrastructures.

Therefore, the imperative for enhancing security in the electric power system has reached a new level that demands industry attention. To address this imperative in a logical and deliberate way, understanding is required of what is involved and how to measure the current and future levels of secure performance.

The technologies that support the operational control of electrical networks range from energy management systems (EMSs) to remote field devices. Critical systems include:

- Energy management system (EMS): The objective of the EMS is to manage the production, purchasing, transmission, distribution, and sale of electrical energy in the power system at a minimal cost with respect to safety and reliability. Management of the real-time operation of an electric power system is a complex task requiring interaction of human operators, computer systems, communications networks, and real-time data-gathering devices in power plants and substations. An EMS consists of computers, display devices, software, communication channels, and remote terminal units (RTUs) that are connected to other remote terminal units, control actuators, and transducers in power plants and substations. The main tasks that an EMS performs have to do with generator control and scheduling, network analysis, and operator training. Control of generation requires that the EMS maintain system frequency and tie-line flows while economically dispatching each generating unit. Management of the transmission network requires that the EMS monitor up to thousands of telemetered values, estimate the electrical state of the network, and inform the operator of the best strategy to handle potential outages that could result in an overload or voltage limit violation. EMSs can have real-time two-way communication links between substations, power plants, independent system operators, and other utility EMSs.
- Supervisory control and data acquisition (SCADA) system: A SCADA system supports operator control of remote (or local) equipment, such as opening or closing a breaker. A SCADA system provides three critical functions in the operation of an electric power system: data acquisition,

supervisory control, and alarm display. It consists of one or more computers with appropriate applications software connected by a communications system to a number of RTUs placed at various locations to collect data, perform intelligent control of electrical system devices, and report results back to an EMS. SCADAs can also be used for similar applications in natural gas pipeline transmission and distribution applications. A SCADA can have real-time communication links with one or more EMSs and hundreds of substations.

- Remote terminal units (RTUs): RTUs are special-purpose microprocessor-based computers that contain analog-to-digital converters (ADCs) and digital-to-analog converters (DACs), digital inputs for status, and digital output for control. There are transmission substation RTUs and distribution automation (DA) RTUs. Transmission substation RTUs are deployed at substation and generation facilities, where a large number of status and control points are required. DA RTUs are used to:
- Control air switches and Var-compensation capacitor banks on utility poles
- Control pad-mounted switches
- Monitor and automate feeders
- Monitor and control underground networks
- Monitor, control, and automate smaller distribution substations

RTUs are also used as indicated above in natural gas transmission and distribution. RTUs can be configured and interrogated using telecommunication technologies. They can have hundreds of real-time communication links with other substations, EMSs, and power plants.

- Programmable logic controllers (PLCs): PLCs have been used extensively in manufacturing and process industries for many years and are now being used to implement relay and control systems in substations. PLCs have extended I/O systems similar to transmission substation RTUs. The control outputs can be controlled by software residing in the PLC and via remote commands from a SCADA system. The PLC user can make changes in the software stored in EEPROM without making any major hardware or software changes. In some applications, PLCs with RTU-reporting capability may have advantages over conventional RTUs. PLCs are also used in many power plant and refinery applications. They were originally designed for use in discrete applications like coal handling. They are now being used in continuous control applications such as feedwater control. PLCs can have many real-time communication links inside and outside substations or plants.
- Protective relays: Protective relays are designed to respond to system faults and short circuits. When faults occur, the relays must signal the appropriate circuit breakers to trip and isolate the faulted equipment. Distribution system relaying must be coordinated with fuses and reclosures for faults while ignoring cold load pickup, capacitor bank switching, and transformer energization. Transmission line relaying must locate and isolate a fault with sufficient speed to preserve stability, reduce fault damage, and minimize the impact on the power system. Certain types of "smart" protective relays can be configured and interrogated using telecommunication technologies.
- Automated metering: Automated metering is designed to upload residential and/or commercial gas and/or electric meter data. These data can then be automatically downloaded to a PC or other device and transmitted to a central collection point. With this technology, real-time communication links exist outside the utility infrastructure.
- Plant-distributed control systems (DCSs): DCSs are plant-wide control systems that can be used for control and/or data acquisition. The I/O count can be higher than 20,000 data points. Often, the DCS is used as the plant data highway for communication to and from intelligent field devices, other control systems (such as PLCs), RTUs, and even the corporate data network for enterprise resource planning (ERP) applications. The DCS traditionally has used a proprietary operating system. Newer versions are moving toward open systems such as Windows NT,

Sun Solaris, and so on. DCS technology has been developed with operating efficiency and user configurability as drivers, rather than system security. Additionally, technologies have been developed that allow remote access, usually via PC, to view and potentially reconfigure the operating parameters.

- Field devices: Examples of field devices are process instrumentation such as pressure and temperature sensors and chemical analyzers. Other standard types of field devices include electric actuators. Intelligent field devices include electronics to enable field configuration, upload of calibration data, and so on. These devices can be configured off-line. They also can have real-time communication links between plant control systems, maintenance management systems, standalone PCs, and other devices inside and outside the facility.

Security of these cyber and communication networks is fundamental to the reliable operation of the grid. As power systems rely more heavily on computerized communications and control, system security has become increasingly dependent on protecting the integrity of the associated information systems. Part of the problem is that existing control systems, which were originally designed for use with proprietary, standalone communication networks, were later connected to the Internet (because of its productivity advantages and lower costs), but without adding the technology needed to make them secure. Communication of critical business information and controlled sharing of that information are essential parts of all business operations and processes.

As the deregulation of the energy industry unfolds, information security will become more important. For the energy-related industries, the need to balance the apparently mutually exclusive goals of operating system flexibility with the need for security will need to be addressed from a business perspective. Key electric energy operational systems depend on real-time communication links (both internal and external to the enterprise). The functional diversity of these organizations has resulted in a need for these key systems to be designed with a focus on open systems that are user configurable to enable integration with other systems (both internal and external to the enterprise). In many cases, these systems can be reconfigured using telecommunication technologies. In nearly all cases, the systems dynamically exchange data in real time. This results in a need for highly reliable, secure control and information management systems.

Power plant DCS systems produce information necessary for dispatch and control. This requires real-time information flow between the power plant and the utility's control center, system dispatch center, regulatory authorities, and so on. A power plant operating as part of a large wholesale power network may have links to an independent system operator, a power pool, and so on. As the generation business moves more and more into market-driven competitive operation, both data integrity and confidentiality will become major concerns for the operating organizations.

Any telecommunication link that is even partially outside the control of the organization that owns and operates power plants, SCADA systems, or EMSs represents a potentially insecure pathway into the business operations of the company as well as a threat to the grid itself. The interdependency analyses done by most companies during Y2K preparations have identified these links and the system's vulnerability to their failures. Thus they provide an excellent reference point for a cyber-vulnerability analysis.

In particular, monitoring and control of the overall grid system is a major challenge. Existing communication and information system architectures lack coordination among various operational components, which usually is the cause for the unchecked development of problems and delayed system restoration. Like any complex dynamic infrastructure system, the electricity grid has many layers and is vulnerable to many different types of disturbances. While strong centralized control is essential to reliable operations, this requires multiple, high-data-rate, two-way communication links, a powerful central computing facility, and an elaborate operation control center, all of which are especially vulnerable when they are needed most—during serious system stresses or power disruptions. For deeper protection, intelligent distributed control is also required, which would enable parts of the network to remain operational and even automatically reconfigure in the event of local failures or threats of failure.

9.2.5 The Dilemma: Security and Quality Needs

The specter of terrorism raises a profound dilemma for the electric power industry: How to make the electricity infrastructure more secure without compromising the productivity advantages inherent in today's complex, highly interconnected electric networks? Resolving this dilemma will require both short-term and long-term technology development and deployment, affecting some of the fundamental characteristics of today's power systems:

- Centralization/decentralization of control. For several years, there has been a trend toward centralizing control of electric power systems. Emergence of regional transmission organizations (RTOs) as agents of wide-area control, for example, offers the promise of greatly increased efficiency and improved customer service. But if terrorists can exploit the weaknesses of centralized control, security would seem to demand that smaller, local systems become the system configuration of choice. In fact, strength and resilience in the face of attack will increasingly rely upon the ability to bridge simultaneous top-down and bottom-up decision making in real time.
- Increasing complexity. The North American electric power system has been called the "most complex machine ever built." System integration helps move power more efficiently over long distances and provides redundancy to ensure reliable service, but it also makes the system more complex and harder to operate. In response, new mathematical approaches are needed to simplify the operation of complex power systems and to make them more robust in the face of natural or manmade interruptions.
- Dependence on Internet communications. Today's power systems could not operate without tightly knit communications capability—ranging from high-speed data transfer among control centers to interpretation of intermittent signals from remote sensors. Because of the vulnerability of Internet communications, however, protection of the electricity supply system requires new technology to enhance the security of power system command, control, and communications, including both hardware and software.
- Accessibility and vulnerability. Because power systems are so widely dispersed and relatively accessible, they are particularly vulnerable to attack. Although "hardening" of some key components, such as power plants and critical substations, is certainly desirable, it is simply not feasible or economic to provide comprehensive physical protection to all components. Probabilistic assessments can offer strategic guidance on where and how to deploy security resources to greatest advantage.

A survey of electric utilities revealed real concerns about grid and communications security. Figure 9.24 ranks the perceived threats to utility control centers. The most likely threats were bypassing controls, integrity violations, and authorization violations. Concern about the potential threats generally increased as the size of the utility (peak load) increased.

The system's equipment and facilities are dispersed throughout the North American continent which complicates protection of the system from a determined terrorist attack. In addition, another complexity—the power delivery systems' physical vulnerabilities and susceptibility to disruptions in computer networks and communication systems—must also be considered. For example, terrorists might exploit the increasingly centralized control of the power delivery system to magnify the effects of a localized attack. Because many consumers have become more dependent on electronic systems that are sensitive to power disturbances, an attack that leads to even a momentary interruption of power can be costly. A 20-minute outage at an integrated circuit fabrication plant, for example, could cost US$30 million.

Despite increasing concerns in these areas as well as the continuing erosion of reserve margins, the infrastructures of today do generally function well. The electricity is on, the phones work, traffic flows nearly all of the time. But more and more, the traditional level of performance is no longer good enough;

more robust infrastructures are needed for the "digital society" envisioned for tomorrow. For example in the electric power area, there is a need for an increase in reliability from today's average of about 99.9% (approximately 8 hours of outage per year) to 99.9999% (about 32 seconds outage per year) or even 99.9999999% (one outage lasting less than a single AC cycle per year). Such near-perfect power is needed today for error-free operation of the microprocessor chips finding their way into just about everything, including billions of embedded applications.

Fortunately, the core technologies needed to strategically enhance system security are the same as those needed to resolve other areas of system vulnerability, as identified in the *Electricity Technology Roadmap*. These result from open access, exponential growth in power transactions, and the reliability needed to serve a digital society.

The North American electric power system needs a comprehensive strategy to prepare for the diverse threats posed by terrorism. Such a strategy should both increase protection of vital industry assets and ensure the public that they are well protected. A number of actions will need to be considered in formulating an overall security strategy:

- The grid must be made secure from cascading damage.
- Pathways for environmental attack must be sealed off.
- Conduits for attack must be monitored, sealed off and "sectionalized" under attack conditions.
- Critical controls and communications must be made secure from penetration by hackers and terrorists.
- Greater intelligence must be built into the grid to provide flexibility and adaptability under attack conditions, including automatic reconfiguration.
- Ongoing security assessments, including the use of game theory to develop potential attack scenarios, will be needed to ensure that the power industry can stay ahead of changing vulnerabilities.

The dispersed nature of the power delivery system's equipment and facilities complicates the protection of the system from a determined attack. Furthermore, both physical vulnerabilities and susceptibility of power delivery systems to disruptions in computer networks and communication systems must be considered. For example, terrorists might exploit the increasingly centralized control of the power delivery system to magnify the effects of a localized attack. Because many consumers have become more dependent on electronic systems that are sensitive to power disturbances, an attack that leads to even a momentary interruption of power can be costly.

9.2.5.1 Human Performance

Because humans interact with these infrastructures as managers, operators, and users, human performance plays an important role in their efficiency and security. In many complex networks, the human participants themselves are both the most susceptible to failure and the most adaptable in the management of recovery. Modeling and simulating these networks, especially their economic and financial aspects, will require modeling the bounded rationality of actual human thinking, unlike that of a hypothetical "expert" human as in most applications of artificial intelligence (AI). Even more directly, most of these networks require some human intervention for their routine control and especially when they are exhibiting anomalous behavior that may suggest actual or incipient failure.

Operators and maintenance personnel are obviously "inside" these networks and can have direct, real-time effects on them. But the users of a telecommunication, transportation, electric power, or pipeline system also affect the behavior of those systems, often without conscious intent. The amounts, and often the nature, of the demands put on the network can be the immediate cause of conflict, diminished performance and even collapse. Reflected harmonics from one user's machinery degrade power quality for all. Long transmissions from a few users create Internet congestion. Simultaneous lawn watering drops the water pressure for everyone. In a very real sense, no one is "outside" the infrastructure.

Given that there is some automatic way to detect actual or immanent local failures, the obvious next step is to warn the operators. Unfortunately, the operators are usually busy with other tasks, sometimes even responding to previous warnings. In the worst case, the detected failure sets off a multitude of almost simultaneous alarms as it begins to cascade through the system, and, before the operators can determine the real source of the problem, the whole network has shut itself down automatically.

Unfortunately, humans have cognitive limitations that can cause them to make serious mistakes when they are interrupted. In recent years, a number of systems have been designed that allow users to delegate tasks to intelligent software assistants ("softbots") that operate in the background, handling routine tasks and informing the operators in accordance with some protocol that establishes the level of their delegated authority to act independently. In this arrangement, the operator becomes a supervisor, who must either cede almost all authority to subordinates or be subject to interruption by them. At present, this is a very limited understanding of how to design user interfaces to accommodate interruption.

9.2.5.2 Broader Technical Issues

In response to the above challenges, several enabling technologies and advances are/will be available that can provide necessary capabilities when combined in an overall system design. Among them are the following:

- Flexible AC transmission system (FACTS) devices, which are high-voltage thyristor-based electronic controllers that increase the power capacity of transmission lines and have already been deployed in several high-value applications. At peak demand, up to 50% more power can be controlled through existing lines.
- Fault current limiters (FCLs), which absorb the shock of short circuits for a few cycles to provide adequate time for a breaker to trip. It is noteworthy that preliminary results of the August 14 outage show that FCLs could have served as large electrical "shock absorbers" to limit the size of blackouts.
- Wide-area measurement systems (WAMS), which integrate advanced sensors with satellite communication and time stamping using global positioning systems (GPS) to detect and report angle swings and other transmission system changes.
- Innovations in materials science and processing, including high-temperature superconducting (HTS) cables, oxide-power-in-tube technology for HTS wire, and advanced silicon devices and wide bandgap semiconductors for power electronics.
- Distributed resources such as small combustion turbines, solid oxide and other fuel cells, photovoltaics, superconducting magnetic energy storage (SMES), transportable battery energy storage systems (TBESS), etc.
- Information systems and online data processing tools such as the open-access same-time information system (OASIS), and transfer capability evaluation (TRACE) software, which determines the total transfer capability for each transmission path posted on the OASIS network, while taking into account the thermal, voltage, and interface limits.
- Monitoring and use of IT: Wide-area measurement/management systems (WAMS), open-access same-time information system (OASIS), supervisory control and data acquisisiton (SCADA) systems, energy management systems (EMS).
- Analysis tools: Several software systems for dynamic security assessment of large/wide-area networks augmented with market/risk assessment.
- Intelligent electronic devices with security provisions bulit in- combining sensors, computers, telecommunication units, and actuators; integrated sensor; two-way communication; "intelligent agent" functions: assessment, decision, learning; actuation, enabled by advances in several areas including semiconductors and resource-constrained encryption.

However, if most of the above technologies are developed, still the overall systems' control will remain a major challenge. This is a rich area for research and development of such tools, as well as to address systems and infrastructure integration issues of their deployment in the overall network, especially now because of increased competition, the demand for advanced technology to gain an advantage, and the challenge of providing the reliability and quality consumers demand.

9.2.5.3 Western States' Power Crises: A Brief Overview of Lessons Learned

An example of "urgent" opportunities is within the now seemingly calm California energy markets; the undercurrents that led to huge price spikes and considerable customer pain in recent years are yet to be fully addressed and alleviated. Such "perfect storms" may appear once again during another cycle of California economic recovery and growth. The California power crisis in 2000 was only the most visible parts of a larger and growing U.S. energy crisis that is the result of years of inadequate investments in the infrastructure.

For example, at the root of the California crisis was declining investment in infrastructure components that led to a fundamental imbalance between growing demand for power and an almost stagnant supply. The imbalance had been brewing for many years and is prevalent throughout the nation (see EPRI's Western States Power Crises White Paper; http://www.epri.com/WesternStatesPowerCrisisSynthesis.pdf).

California is a good downside example of a societal testbed for the ways that seemingly "good" theories can fail in the real world. For example, inefficient markets provide inadequate incentives for infrastructure investment:

- Boom–bust cycle may be taking shape in generation investment
- Transmission investment running at one-half of 1975 levels
- Congestion in transmission network is rising, as indicated by increase of number of transmission loading relief (TLR) during the last three years.

The cost of market failure can be also very high; as indicated by the exercise of market power in California during summer of 2000 which cost consumers $4 billion initially, while the ongoing intermediate loss to businesses may well be considerably higher.

More specifically regarding the electricity under investments and persisting undercurrents, very specific "investments" by the state were made, on the order of $10 billion, paid to subsidize (hold down) electricity prices, and to bail out bankrupt companies through long-term noncompetitive contracts that did not address the undercurrents and shortcomings of the earlier policies.

To address these issues there are both tactical as well as strategic needs; for example, the so-called "low-hanging fruits" to improve transmission networks include:

- Deploy existing technologies to improve use of already in place transmission assets (e.g., FACTS, dynamic thermal circuit rating, and energy storage–peak shaving technologies). For example, through the integration of load management technologies shaving nearly 5,000 MW which amounts to about 10% of total demand, combined with a more precise control enabled by the use of FACTS devices, which enable nearly 50% more transfer capability over existing transmission lines.
- Develop and deploy new technologies to improve transmission reliability and throughput (e.g., low sag composite conductors, high-temperature superconducting cables, extra high-voltage AC and DC transmission systems, hierarchical control systems)
- Improve real-time control of network via monitoring and data analysis of dynamic transmission conditions
- Develop and deploy self-healing grid tools to adaptively respond to overload and emergency conditions
- Digital control of the power delivery network (reliability, security, and power quality).
- Integrated electricity and communications for the user
- Transformation of the meter into a two-way energy/information portal
- Integration of distributed energy resource into the network

- The complex grid can operate successfully if technology is deployed and operated in an integrated manner (there is no "silver bullet")

In addition, longer-term strategic considerations must be addressed; they include:

- Greater fuel diversity including renewable energy technologies—regional and national priorities
- Risk assessment of long-term U.S. reliance—analysis of the value of risk management through fuel diversity
- Introduce time-varying prices and competitive market dynamics for all customers
- Create a planning process and in silico testing of designs, devices, and power markets
- Model market efficiencies, environmental constraints, and renewables
- Develop advanced EM threat detection, shielding, and surge suppression capabilities
- Develop the tools and procedures to ensure a robust and secure marketplace for electricity
- Develop the portfolio of advanced power generation technologies to assure energy security
- Transmission network expansion and RTOs, e.g., would an RTO, compliment a competitive wholesale power market and result in a sustainable and robust system? How large should they be?
- Comprehensive architecture for power supply and delivery infrastructure that anticipates rapidly escalating demands of digital society
- Enable self-healing power delivery infrastructure
- Significant investment in R&D, transmission, generation, and conservation resources are needed
- Incentives for technology innovation and accountability for R&D
- Revitalize the national public/private electricity infrastructure partnership needed to fund the "self-healing grid" deployment
- The "law of unintended consequences" should be considered in crafting any solution

Having discussed the above technology-intensive "push," the fact that adoption of new technologies often creates equally new markets must also be considered. For example, wireless communication creates the market of spectrum, and broadband technologies create the market of bandwidth. Reduced regulation of major industries has required new markets wherever the infrastructure is congested: airlines compete for landing rights, power generators for transmission rights, oil, and gas producers for pipeline capacity.

From a national perspective, a key grand challenge is how to redesign, retrofit, and upgrade the nearly 200,000 miles of electromechanically controlled system into a smart self-healing grid that is driven by a well-designed market approach?

In addressing this challenge, as technology progresses, and the economy becomes increasingly dependent on markets, infrastructures such as electric power, oil/gas/water pipelines, telecommunications, financial, and transportation networks becomes increasingly critical and complex. In particular, since it began in 1882, electric power has grown to become a major industry essential to a modern economy

Over the past two decades, governments around the globe have introduced increasing amounts of competition into network industries. With the advent of restructuring in the electric power industry, the onset of a historical transformation of the energy infrastructure in the context of global trends is underway:

- Increasing electricity demand as a consequence of economic and population growth
- Technological innovations in power generation, delivery, control and communications
- Increasing public acceptance of market mechanisms
- Growing public concerns about environmental quality and depletion of exhaustible resources

Services previously supplied by vertically integrated, regulated monopolies are now provided by multiple firms. The transition to competition has fundamentally altered important aspects of the engineering and economics of production. This presents unique opportunities and challenges. Clearly, this change will have far-reaching implications for the future development of the electricity industry. More fundamentally, as we look beyond the horizon, this change will further power the information revolution and increasing global interdependence. The long-term socioeconomic impacts of such a transformation will be huge, and the tasks are just as daunting, going well beyond the boundary of existing knowledge.

To meet such a challenge, collaborative research between engineers and economists is critical to provide a holistic and robust basis that will support the design and management of complex technological and economic systems in the long term. The electric power industry offers an immediate opportunity for launching such research, as new ways are being sought to improve the efficiency of electricity markets while maintaining the reliability of the network. Complexity of the electric power grid combined with ever more intricate interactions with markets offers a plethora of new and exciting research opportunities.

In what follows we provide our vision and approach to enabling a smart self-healing electric power system that can respond to a broad array of destabilizers.

9.2.5.4 How to Make an Electric Power System Smart?

To add intelligence to an electric power transmission system, independent processors in each component and each substation and power plant are needed. These processors must have a robust operating system and be able to act as independent agents that can communicate with and cooperate with other forming a large distributed computing platform. Each agent must be connected to sensors associated with its own component or its own substation so that it can assess its own operating conditions and report them to its neighboring agents via the communications paths. Thus for example, a processor associated with a circuit breaker would have the ability to communicate with sensors built into the breaker and communicate those sensor values using high-bandwidth fiber communications connected to other such processor agents.

9.2.5.5 Complex System Failure

Beyond the human dimension, there is a strategic need to understand the societal consequences of infrastructure failure risks along with benefits of various tiers of increased reliability. From an infrastructure interdependency perspective, power, telecommunications, banking and finance, transportation and distribution, and other infrastructures are becoming more and more congested, and are increasingly vulnerable to failures cascading through and between them. A key concern is the avoidance of widespread network failure due to cascading and interactive effects. Moreover, interdependence is only one of several characteristics that challenge the control and reliable operation of these networks. Other factors that place increased stress on the power grid include dependencies on adjacent power grids (increasing because of deregulation), telecommunications, markets, and computer networks. Furthermore, reliable electric service is critically dependent on the whole grid's ability to respond to changed conditions instantaneously.

More specifically, secure and reliable operation of critical infrastructures poses significant theoretical and practical challenges in analysis, modeling, simulation, prediction, control, and optimization. To address these challenges, a research initiative—the EPRI/DOD Complex Interactive Networks/Systems Initiative (CIN/SI)—was undertaken during 1998–2001 to enable critical infrastructures to adapt to a broad array of potential disturbances, including terrorist attacks, natural disasters, and equipment failures.

The CIN/SI overcame the longstanding problems of complexity, analysis, and management for large interconnected systems—and systems of systems—by opening up new concepts and techniques. Dynamical systems, statistical physics, information and communication science, and computational complexity were extended to provide practical tools for measuring and modeling the power grid, cell phone networks, Internet, and other complex systems. For the first time, global dynamics for such systems can be understood fundamentally (Figure 9.25).

As an example, related to numerous major outages, narrowly programmed protection devices have contributed to worsening the severity and impact of the outage, typically performing a simple on/off logic which locally acts as preprogrammed while destabilizing a larger regional interconnection.

From a broader perspective, any critical national infrastructure typically has many layers and decision-making units and is vulnerable to various types of disturbances. Effective, intelligent, distributed control is required that would enable parts of the constituent networks to remain operational and even automatically reconfigure in the event of local failures or threats of failure. In any situation subject to rapid changes, completely centralized control requires multiple, high-data-rate, two-way communication links, a powerful central computing facility, and an elaborate operations control center. But all of these are liable to disruption at the very time when they are most needed (i.e., when the system is stressed by natural disasters, purposeful attack, or unusually high demand).

When failures occur at various locations in such a network, the whole system breaks into isolated "islands," each of which must then fend for itself. With the intelligence distributed, and the components acting as independent agents, those in each island have the ability to reorganize themselves and make efficient use of whatever local resources remain to them in ways consonant with the established global goals to minimize adverse impact on the overall network. Local controllers will guide the isolated areas to operate independently while preparing them to rejoin the network, without creating unacceptable local conditions either during or after the transition. A network of local controllers can act as a parallel, distributed computer, communicating via microwaves, optical cables, or the power lines themselves, and intelligently limiting their messages to only that information necessary to achieve global optimization and facilitate recovery after failure.

9.2.6 Conclusions: Toward a Secure and Efficient Infrastructure

How to control a heterogeneous, widely dispersed, yet globally interconnected system is a serious technological problem in any case. It is even more complex and difficult to control it for optimal efficiency and maximum benefit to the ultimate consumers while still allowing all its business components to compete fairly and freely. A similar need exists for other infrastructures, where future advanced systems are predicated on the near-perfect functioning of today's electricity, communications, transportation, and financial services.

From a strategic R&D viewpoint, agility and robustness/survivability of large-scale dynamic networks that face new and unanticipated operating conditions will be presented. A major challenge is posed by the lack of a unified mathematical framework with robust tools for modeling, simulation, control and optimization of time-critical operations in complex multicomponent and multiscaled networks.

Given the state of art in electricity infrastructure security and control as indicated in this chapter, creating a smart grid with self-healing capabilities is no longer a distant dream; we have made considerable progress. But considerable technical challenges as well as several economic and policy issues remain to be addressed

Acknowledgments

Most of the material and findings presented here were developed while the author was at the Electric Power Research Institute (EPRI) in Palo Alto, California. The author is grateful for EPRI's support and feedback from numerous colleagues at EPRI, universities, industry, and government agencies.

References

Amin, M. 2000. Toward self-healing infrastructure systems, *IEEE Computer Magazine*, 33 (8), 44–53.

Amin, M. 2001. Toward self-healing energy infrastructure systems, *IEEE Computer Applications in Power*, 14 (1), 20–28.

Amin, M. 2002a. Security challenges for the electricity infrastructure, *IEEE Computer Magazine*, April 8–10.

Amin, M. 2002b. *Special Issue* of *IEEE Control Systems Magazine on Control of Complex Networks*, December, Vol. 21, No. 6.

Amin, M. 2002c. *Special Issue* of *IEEE Control Systems Magazine on Control of Complex Networks*, February, Vol. 22, No. 1.

Amin, M. 2003. North America's electricity infrastructure: Are we ready for more perfect storms? *IEEE Security and Privacy Magazine*, 1 (5), 19–25.

Amin, M., *Special Issue of the Proceedings of the IEEE on Energy Infrastructure Defense Systems*. in press.

Amin, M., Gerhart, V., and Rodin, E. Y. 1997. System identification via artificial neural networks: Application to on-line aircraft parameter estimation, *Proceedings of AIAA/SAE 1997 World Aviation Congress*, p. 22, Anaheim, CA.

Ang, A. H-S. and Tang, W. H. 1990. *Decision, Risk, and Reliability, Volume 2 of Probability Concepts in Engineering Planning and Design*, Wiley, New York.

Ayyub, B. M. 2002. *Elicitation of Expert Opinions for Uncertainty and Risks*, CRC Press, Boca Raton, FL.

Ayyub, B. M. 2003. *Risk Analysis in Engineering and Economics*, Chapman and Hall/CRC Press, Boca Raton, FL.

Ayyub, B. M. and Klir, G. J. 2006. *Uncertainty Modeling and Analysis for Engineers and Scientists*, Chapman and Hall/CRC Press, Boca Raton, FL.

Ayyub, B. M. and McCuen, R. 2003. *Probability, Statistics and Reliability for Engineers and Scientists*, *2nd Ed.*, Chapman and Hall/CRC Press, Boca Raton, FL.

Dy Liacco, T. E. 1967. The adaptive reliability control system, *IEEE Transactions on Power Apparatus and Systems*, May, 517–561.

Energy Information Administration (EIA). 2006. *Annual Energy Outlook 2006 with projections to 2030*, U.S. Department of Energy, Washington, DC, February 2006, http://www.eia.doe.gov/oiaf/aeo/figure_3.html

EPRI. 1999. Electricity technology roadmap: 1999 summary and synthesis report, Palo Alto, CA, July 1999, http://www.epri.com/corporate/discover_epri/roadmap/CI-112677-V1_all.pdf

EPRI. 2000. Communication security assessment for the United States electric utility infrastructure, EPRI Report 1001174, EPRI, Palo Alto, CA.

EPRI. 2001. *Electricity Infrastructure Security Assessment*, *Vol. I–II*, EPRI, Palo Alto, CA.

EPRI. 2003. Complex interactive networks/systems initiative: Final summary report—Overview and summary final report for joint EPRI and US Department of Defense University Research Initiative, EPRI, Palo Alto, CA.

EPRI. 2003. *Electricity Technology Roadmap Report*, EPRI, Palo Alto, CA, July 2003, http://www.epri.com/corporate/discover_epri/roadmap/index.html

Edison Electric Institute (EEI), 2003. Critical infrastructure protection, http://www.eei.org/industry_issues/energy_infrastructure/critical_infrastructure_protection/#1

EPRI. 2004. *Supervisory Control and Data Acquisition (SCADA) Systems Security Guide*, EPRI, Rep. 1002604, Palo Alto, CA.

EPRI. 2005. Guideline for securing control system & corporate network interfaces, EPRI, Rep. 1010714, Palo Alto, CA.

Fink, L. H. and Carlsen, K. 1978. Operating under stress and strain, *IEEE Spectrum,* March, 48–53.

Gellings, C. W. and Yeager, K. E. 2004. Transforming the electric infrastructure. *Physics Today,* December, Vol. 57, 45–51.

Hauer, F. F. and Dagle, J. E. 1999. *Review of Recent Reliability Issues and System Events*, Consortium for Electric Reliability Technology Solutions, Transmission Reliability Program, Office of Power Technologies, U.S. Department of Energy (DOE), Washington, DC, August 30.

House Committee on Energy and Commerce. 2003. Blackout 2003: How did it happen and why?, 108th Congress House Hearings, the U.S. Government Printing Office Via GPO Access (DOCID: F: 89467. wais), Washington, DC, September 3, 2003, http://energycommerce.house.gov/108/Hearings/09032003 hearing 1061/print.htm

Kumamoto, H. and Henley, E. J. 1996. *Probabilistic Risk Assessment and Management for Engineers and Scientists*, *2nd Ed.*, IEEE Press, New York.

Kundur, P. 1994. *Power System stability and control, EPRI Power System Engineering Series*, McGraw-Hill, New York.

Maryland Emergency Management Agency (MEMA). 2006. State of Maryland guide for the protection of critical infrastructure and key resources for homeland security, Volume 1: Critical asset & portfolio risk assessment (CAPRA) methodology, Office of Homeland Security, Annapolis, MD.

Modarres, M. 1993. *What every engineer should know about Reliability and Analysis*, Marcel Dekker, New York.

Modarres, M., Kaminskiy, M., and Krivstov, V. 1999. *Reliability Engineering and Risk Analysis: A Practical Guide*, Marcel Dekker, New York.

National Science Foundation, Division of Science Resources Statistics. 2003. Research and development in industry: 2000. NSF 03-318, NSF, Arlington, VA, http://www.nsf.gov/sbe/srs/nsf03318/pdf/taba19.pdf

North American Electric Reliability Council (NERC), 2006. Disturbance Analysis Working Group (DAWG) database (1984–2002), Princeton, NJ, Available: http://www.nerc.com/~dawg

North American Electric Reliability Council (NERC), 2001. *Reliability Assessment 2001-2010*, Princeton, NJ, Available: http://www.nerc.com/~filez/rasreports.html

North American Electric Reliability Council (NERC), 2004. *NERC security guidelines for the electricity sector*, Princeton, NJ, Available: http://www.esisac.com/library-guidelines.htm

North American Electric Reliability Council (NERC), 2006. *Top 10 Vulnerabilities of control system and Their Associated mitigations-2006*, Control Systems Security Working Group, U.S. Department of Energy, National SCADA Test Bed Program, Princeton, NJ, March 16.

President's Commission on Critical Infrastructure Protection. 1997. Critical foundations: Protecting America's infrastructures, http://www.ciao.ncr.gov

Samotyj, M., Gellings, C., and Amin, M. 2003. Power system infrastructure for a digital society: creating the new frontiers, *Proceedings of the GIGRE/IEEE-PES Symposium on Quality and Security of Electric Power Delivery*, p. 10, Montreal, October 7–10.

Silberman, S. 2001. The energy web, *Wired*, 9 (7), 116.

Starr, C., and Amin, M. 2003. Global transition dynamics unfolding the full social implications of national decision pathways, http://cdtlnet.cdtl.umn.edu/Amin/GlobalTransition.pdf

US Department of Energy. 2002. National transmission grid study, http://tis.eh.doe.gov/ntgs/gridstudy/main_screen.pdf

US Department of Homeland Security (DHS), 2006. The National Infrastructure Protection Plan (NIPP),Washington, DC, June, Available: http://www.dhs.gov/nipp

US DOE and the US DHS, 2006. Roadmap to Secure Control Systems in the Energy Sector, Office of Energy Delivery and Energy Reliability, 58 pages, Washington, DC, January.

10

Electrical Energy Management in Buildings

Craig B. Smith
Dockside Consultants, Inc.

Kelly E. Parmenter
Global Energy Partners, LLC

10.1 Principal Electricity Uses in Buildings

10.1.1 Introduction: The Importance of Energy Efficiency in Buildings

A typical building is designed for a forty-year economic life. This implies that the existing inventory of buildings—with all their good and bad features—is turned over very slowly. Today, we know it is cost-effective to design a high degree of energy efficiency into new buildings because the savings on operating and maintenance costs will repay the initial investment many times over. Many technological advances have occurred in the last two decades, resulting in striking reductions in the energy usage required to operate buildings safely and comfortably. An added benefit of these developments is the reduction in air pollution, which has occurred as a result of generating less electricity.

There are hundreds of building types, and buildings can be categorized in many ways: by use, type of construction, size, thermal characteristics, etc. For simplicity, two designations will be used here: residential and nonresidential.

The residential category includes features common to single-family dwellings, apartments, and hotels. In 2001, there were 107 million housing units in the U.S. The nonresidential category mainly emphasizes office buildings, but also includes a less detailed discussion of features common to retail stores, hospitals, restaurants, and laundries. There are approximately 5 million commercial buildings, totaling 72 billion ft.^{2}, in the U.S. (2001 data). Most of this space is contained in buildings larger than 10,000 ft.^{2}. Industrial facilities are not included here, but are discussed in Chapter 14. The extension to other types is either obvious, or can be pursued by referring to the literature.

Total energy consumption in the two sectors has evolved as follows since the previous two editions of this book (1 quad = 10^{15} BTU):

	Year		
	1975 (quads)	1992 (quads)	2004 (quads)
Residential	—	16.8	21.2
Commercial	—	12.9	17.5
Total	25.1	29.7	38.7

There has been a remarkable shift in the residential and commercial sectors since 1975. The use of natural gas, which increased rapidly in these sectors prior to 1975, flattened out and has remained essentially constant. The use of petroleum has decreased. The biggest change has been the dramatic increase in electricity use, which doubled from 1975 to 2004.

The approach taken in this chapter is to list two categories of specific strategies that are cost-effective methods for conserving electricity. The first category includes those measures that can be implemented at low capital cost using existing facilities and equipment in an essentially unmodified state. The second category includes technologies that require retrofitting, modification of existing equipment, or new equipment or processes. Generally, moderate to substantial capital investments are also required.

10.1.2 Electricity Use in Residential and Commercial Buildings

Table 10.1 summarizes electricity consumption data by major end use for the residential and commercial sectors. The data are from the Energy Information Administration's (EIA's) most recently available energy consumption surveys that took place in 2001 for the residential sector and 1999 for the commercial sector.

The single most significant residential end use of electricity is space cooling (16.1%), followed by refrigeration (13.7%), space heating (10.2%), water heating (9.1%) and lighting (8.8%). The combination of other uses, such as entertainment systems, personal computers, printers, etc., is also substantial, accounting for one-quarter of residential electricity use. In the commercial sector, space conditioning, i.e., the combination of heating, ventilating, and air conditioning (HVAC), uses of the most electricity, accounting for 37.9%. Space cooling accounts for the majority of space conditioning electricity use; indeed, by itself, space cooling represents one-quarter of all commercial electricity use. The next two largest end users of electricity in the commercial sector are lighting at 23.1% and office equipment at 17.9%.

10.1.2.1 Residential Electricity Use

Space conditioning is the most significant end use of electricity in residential buildings, accounting for approximately one-quarter of total electricity use. Electricity is used in space heating and cooling to drive fans and compressors, to provide a direct source of heat (resistance heating), to provide an indirect source of heat or "cool" (heat pumps),* and for controls.

At 8.8% in 2001, residential lighting electricity use was down from 10% in the 1980s due to the introduction of more efficient lamps, principally compact fluorescents, as well as to a greater share of electricity being consumed by other end uses. The bulk of residential lighting is still incandescent, and offers substantial opportunities for improved efficiency.

The share of residential electricity used by water heating has also decreased during the last decade. At 9.1% in 2001, it is down from 10.7% in 1992. Electricity use for this purpose currently occurs in regions where there is cheap hydroelectricity, or where alternative fuels are not available. Solar water heating (discussed in Chapter 20) is another alternative that is used on a limited basis.

*Heat pumps are discussed in detail elsewhere in this handbook (Chapter 12 and Chapter 25).

TABLE 10.1 Electricity Consumption by End Use in the Residential and Commercial Sectors

End Use	Residential Sector, 2001 (billion kWh)	Percent of Total Residential Electricity Use (%)	Commercial Sector, 1999 (billion kWh)	Percent of Total Commercial Electricity Use (%)
Space Heating	116	10.2	45	5.0
Space Cooling	183	16.1	232	25.6
Ventilation	[a]	[a]	66	7.3
Water Heating	104	9.1	11	1.2
Refrigeration	156	13.7	78	8.6
Cooking	80	7.0	19	2.1
Clothes Washers & Dryers	76	6.7	[a]	[a]
Freezers	39	3.4	[a]	[a]
Lighting	101	8.8	210	23.1
Office Equipment	[a]	[a]	163	17.9
Other Uses	285	25.0	84	9.2
Total	1,140	100	908	100

[a]Included in "Other Uses".

Source: From Energy Information Administration, Residential Energy Consumption Survey, 2001; Energy Information Administration, Commercial Building Energy Consumption Survey, 1999.

Refrigerators are another important energy end use in the residential sector, accounting for 13.7% of residential electricity consumption in 2001. For the last 40 years, virtually every home in the U.S. has had a refrigerator. Therefore, refrigerators have fully penetrated the residential sector for some time. However, significant changes related to energy use have occurred during this period as new standards have been implemented. For one, the average size of refrigerators has more than doubled from less than 10 ft.^{3} in 1947 to over 20 ft.^{3} in recent years. Meanwhile, the efficiency of refrigerators has increased dramatically. In the early 1960s, average electricity consumption of a new refrigerator was around 1000 kWh/year, and they were about 12 ft.^{3} (adjusted volume). Since 2001, new refrigerators consume less than 500 kWh/year and have adjusted volumes around 20 ft.^{3}. The net result is that current refrigerators, although 70% larger than those forty years ago, consume about 50% less electricity.

Cooking, clothes washing and drying, and freezers account for another 17.1%, while "other" uses (including home entertainment systems, personal computers, and miscellaneous items) make up the balance (25%) of electricity used in the residential sector. Computers and other consumer electronic devices, such as VCRs, DVD players, and electronic gaming systems, have proliferated in homes during the last few decades. For example, only 8% of households had personal computers (either desktop or laptop) in 1984, whereas 51% had desktops and 13% had laptops in 2001. The energy used by personal computers is continuing to rise. The EIA estimates that energy use by PCs increased by 9% between 2001 and 2005, and that it will increase by 29% between 2005 and 2010.

10.1.2.2 Nonresidential Electricity Use

For the commercial sector as a whole, HVAC dominates electricity use. In most nonresidential buildings where space conditioning is used, HVAC is the major electricity end use. There are exceptions of course—in energy-intensive facilities such as laundries, the process energy will be most important. Electricity is used in space conditioning to run fans, pumps, chillers, and cooling towers. Other uses include electric resistance heating (for example, in terminal reheat systems) or electric boilers.

Nonresidential lighting is generally next in importance to HVAC in regards to electricity use, except in those nonresidential facilities with energy-intensive processes. In a typical office building, lighting consumes nearly one-quarter of the electricity. Interior lighting in the commercial sector is generally fluorescent, with a growing use of metal halide lamps, and with a small fraction of incandescent lamps. High-efficiency fluorescent lamps, electronic ballasts, compact fluorescent lamps and improved lighting controls are now the norm. Incandescent lamps still see major use in retail for display lighting, as well as in older buildings, or for decorative or esthetic applications. The EIA's *Commercial Building Energy Consumption Surveys* in 1992 and 1999 show that the percentage of commercial building electricity use consumed by lighting dropped from 27.7% in 1992 to 23.1% in 1999. Furthermore, the EIA's *Annual Energy Outlook 2005* predicts that lighting's share of commercial building electricity use will continue to drop, reaching an estimated 21.4% by 2025.

Water heating is another energy use in nonresidential buildings, but here circulating systems (using a heater, storage tank, and pump) are more common. Many possibilities exist for using heat recovery as a source of hot water. The amount of electricity used for water heating in commercial buildings has remained relatively consistent at a little over 1% during the last decade. However, the EIA predicts electricity use for water heating may increase to 2.3% by 2025.

Refrigeration is an important use of energy in supermarkets and several other types of nonresidential facilities. It is now common practice to include heat recovery units on commercial refrigeration systems. As in residential applications, commercial refrigeration electricity use has decreased in the past few decades due to efficiency gains. For example, in 1999 refrigeration accounted for 8.6% of commercial electricity use, down from 10.1% in 1992. EIA predictions estimate refrigeration will only represent 3.9% in 2025.

Commercial electricity use by office equipment has increased substantially since the last edition of this handbook; it grew from 6.7% in 1992 to 17.9% in 1999. Much of this increase is due to a greater use of computers. For larger computing units used in central data processing systems, specially designed rooms with temperature and humidity control are required. As a rule of thumb, the electricity used by the computer must be at least doubled because cooling must be provided to remove the heat from the equipment, lights, and personnel. Now, with the trend toward the widespread use of microcomputers, special space conditioning is not required. But, the sheer numbers of computers increase the air conditioning load and electrical demand.

These, with their peripheral equipment including printers, scanners, and data storage systems, have grown rapidly in the last decade. Electronic mail and Internet conferencing is displacing other communication systems. For the commercial sector as a whole, 431 out of every 1000 employees in 1992 had a computer, compared with 707 out of every 1000 employees in 1999. The ratio of computers to employees is even higher in office buildings. In 1992, there were 6 computers to every 10 employees, while, in 1999, there were 9.5 computers to every 10 employees.

In nonresidential facilities, the balance of the electricity use is for elevators, escalators, and miscellaneous items.

10.2 Strategies for Electricity End-Use Management

10.2.1 Setting up an Energy Management Program

The general procedure for establishing an energy management program in buildings involves five steps:

- Review historical energy use.
- Perform energy audits.
- Identify energy management opportunities.
- Implement changes to save energy.
- Monitor the energy management program, set goals, and review progress.

Each step will be described briefly.

10.2.1.1 Review of Historical Energy Use

Utility records can be compiled to establish electricity use for a recent 12-month period. These should be graphed on a form (see Figure 10.1) so annual variations and trends can be evaluated. By placing several

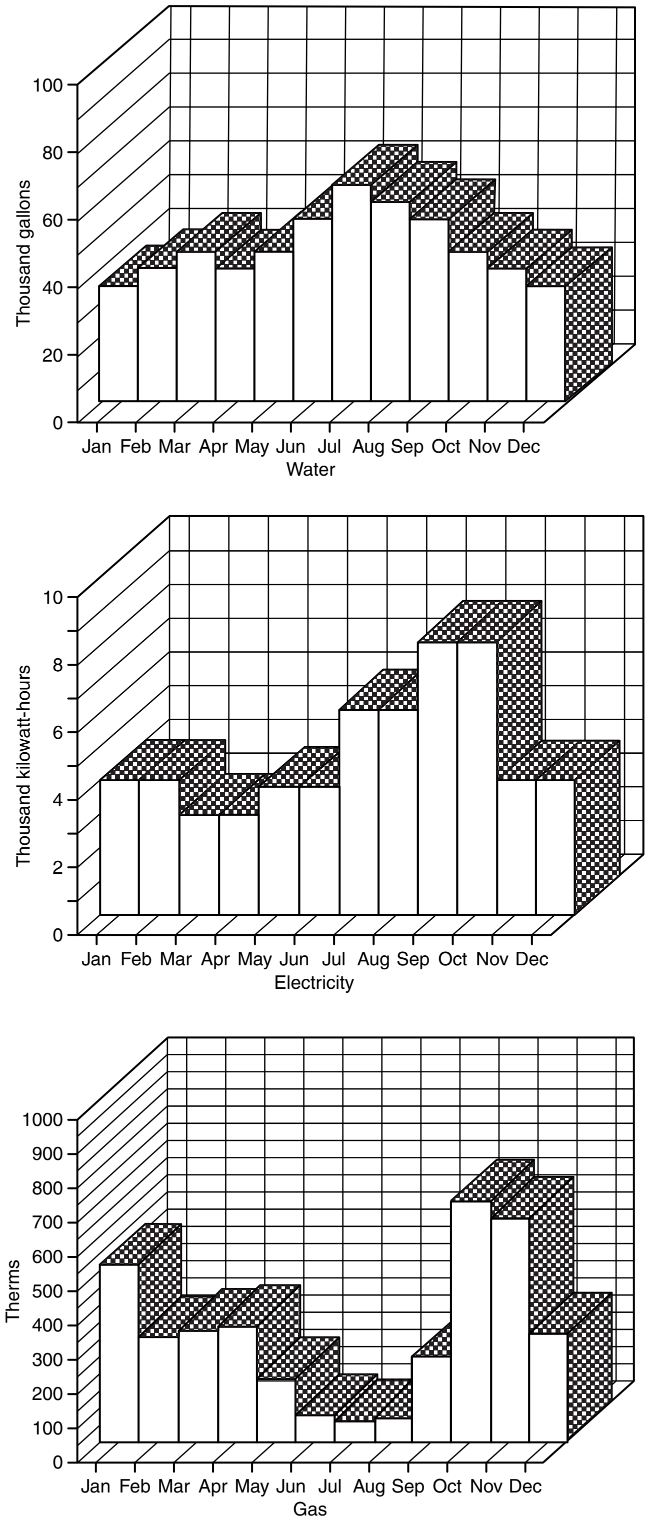


FIGURE 10.1 Sample graph: Historial energy use in an office building.

Building description

- Name:______________ Age: ___years heating degree days ____
- Location: ____
- No. of floors____Gross floor area___ m^2 ($ft.^2$) Net floor area___m^2($ft.^2$)
- Percentage of surface area which is glazed ___% cooling degree days _____
- Type of air conditioning system; heating only ____ evaporative _____
  dual duct ____ other (describe) ____________________________________
- Percentage breakdown of lighting equipment: Incandescent ______%

Building mission

- What is facility used for:
- Full time occupancy (employees) ____ persons
- Transient occupancy (visitors or public) ____ persons
- Hours of operations per year ______
- Unit of production per year ____ Unit is _______

Installed capacity

- Total installed capacity for lighting ____kW
- Total installed capacity of electric drives greater than 7.5 kW (10hp)
  (motors, pumps, fans, elevators, chillers, etc.)____hp × 0.746=____kW
- Total steam requirements _____lbs/day or ____kg/day
- Total gas requirements ___$ft.^3$/day or BTU/hr or ____m^3/day
- Total other fuel requirements _______________________________________

Annual energy end use

Energy form	× conversion			kBTU/yr	metric units		Conversion		MJ/yr
• Electricity	__kWh/yr	× 3.41	=	_____	_____ kWh/yr	×	3.6	=	_______
• Steam	__lb/yr	× 1.00	=	_____	_____ kg/yr	×	2.32	=	_______
• Natural gas	__cf/yr	× 1.03	=	_____	_____ m^3/yr	×	38.4	=	_______
• Oil	__gls/yr	× {#2 139, #6 150}	=	_____	_____ l/yr	×	{#2 38.9, #6 41.8}	=	_______
• Coal	__tons/yr	× 24,000	=	_____	_____ kg/yr	×	28.0	=	_______
• Other	__	×	=	_____	_____	×		=	_______
	Totals			_____					_______

Energy use performance factors (EUPF's) for building

- EUPF 1 = MJ/yr (kBTU/yr) ÷ Net floor area = _______ MJ/m^2yr (kBTU/$ft.^2$yr)
- EUPF 2 = MJ/yr (kBTU/yr) ÷ Average annual occupancy = _______ MJ/person ·yr) (kBTU/person ·yr)
- EUPF 3 = MJ/yr (kBTU/yr) ÷ Annual units of production = _______ MJ/unit ·yr) (kBTU/unit ·yr)

FIGURE 10.2 Building energy survey form.

years (e.g. last year, this year, and next year projected) on the form, past trends can be reviewed and future electricity use can be compared with goals. Alternatively, several energy forms can be compared for energy use vs. determined production (e.g., meals served for a restaurant, or kilograms of laundry washed for a laundry, etc.).

10.2.1.2 Perform Energy Audits

Figure 10.2 and Figure 10.3 are data sheets used in performing an energy audit of a building. The building energy survey form, Figure 10.2, provides a gross indication of how energy is used in the building in meeting the particular purpose for which it was designed. This form would not be applicable to single-family residences, but it could be used with apartments. It is primarily intended for commercial buildings.

Figure 10.3 is a form used to gather information concerning energy used by each piece of equipment in the building. When totaled, the audit results can be compared with the historical energy use records plotted on Figure 10.1. The energy audit results show a detailed breakdown and permit identification of major energy-using items.

Facility name ________________ Date ________________ By __________ Sheet ___ of ___

Location ____________________ Period of survey: 1 day 1wk 1mo 1yr

Symbol: K = 10^3 M = 10^3 Notes ____

Conversion factors

Multiply	by	to get
kWh	306	MJ
BTU/hr	0.000293	kW
hp	0.746	kW

		Power						
Item no.	Equipment description	Name plate rating (BTU/hr,kW, hp,etc,)	Conv. factor to kW	Est.% load (100%,50%, etc.)	Est. hrs use per period	kWh	Conv. factor	Total energy use per period (MJ)

FIGURE 10.3 Energy audit data sheet.

XYZ corporation

Lighting energy savings summary				Prepare by:	
Existing annual kWh:	181,828 kWh	Existing kW draw:	36.37 kW	Annual energy $$ saved:	$12,094
Proposed annual kWh:	95,234 kWh	Proposed kW draw:	18.52 kW	Estimated PRE-REBATE cost.	$13,592
Annual kWh savings:	86,594 kWh	KW savings:	17.85 kW		
% kWh savings:	47.6 %	% kW savings:	49.1 %		

Lighting inventory, recommendations, and savings

		Existing				Recommended			Savings			Estimated	
Item #	Location	Weekly hours	Qty	Fixtures	Watts/fix	Qty	Fixtures	Watts/fix	kW	Annual kWh	Annual Energy S	Unit cost	Total cost
1	Presidents office	60	8	75 Watt INC Spotlight	75	8	18 Watt CFL/SI/Ref.	18	0	1368	$219	$22	$176
2	Presidents office	60	6	2-F40T12(40W)/STD	96	6	2-F32T8(32W)/ELEC	61	0	630	$101	$48	$288
3	V.P. office	60	8	75 WATT INC Spotlight	75	8	18 Watt CFL/SI/Ref.	18	0	1,368	$219	$22	$176
4	V.P. office	60	4	2-F40T12(40W)/STD	96	4	2-F32T8(32W)/ELEC	61	0	420	$67	$48	192
5	Night lighting	168	4	2-F40T12(40W)/STD	96	4	2-F32T8(32W)/ELEC	61	0	1176	$145	$48	$192
6	Women's restroom mirror	60	12	25 Watt INC	25		None						
7	Women's restroom	60	6	2-F40T12(34W)/U/STD	94	6	2-F40T12(34W)/U/ELEC	60	0	612	$98	$48	$288
8	Men's restroom	60	6	2-F40T12(34W)/U/STD	94	6	2-F40T12(34W)/U/ELEC	60	0	612	$98	$48	$288
9	Main office area	60	56	3-F40T12(40W)/2-Class 1 11	136	56	3-F32T8(32W)/1-ELEC	90	2	7728	$1,240	$48	$2,688
10	Storage room	25	1	100 Watt INC	100	1	28 Watt PL CFL/SI	30	0	88	$19	$32	$32
11	Parking garage	168	26	100 Watt Quartz	350	26	175 Watt MH	205	3	3,166	$3,906	$200	$5,200
12	Parking garage	168	8	2-F96T12(75W)/STD	173	8	20F96T8(50W)/ELEC	104	0	4637	$572	$60	$480
13	Physical plant	80	28	2-F96T12(215W)/WHO/STD	450	28	2-F96T12(95W)/HO/ELEC	166	8	3,180	$4,741	$110	$3,080
14	Physical plant	80	162	100 Watt INC	100	16	28 Watt PL CFL/SI	30	1	4480	$668	$32	$512
	Total								1	8,659	$12,094	N/A	$13,592

FIGURE 10.4 Sample energy audit results.

Another way to perform energy audits is to use a microcomputer and a commercially available database or spreadsheet program to record the data and make the calculations. If the workload is extensive, the program can include "lookup" tables of frequently used electrical loads, utility rates, and other essential information to automate the process. We have used teams of engineers with portable computers to rapidly survey and collect the energy data from large commercial facilities. See Figure 10.4 for an example of an audit result.

Still another way of making an energy audit is to use commercially available computer software that estimates energy use for typical building occupancies based on size, type, climate zone, and other identifying parameters.

These programs are not as accurate as an actual survey, but can be used as a preliminary screening criterion to select the buildings worthy of more detailed investigations.

10.2.1.3 Identify Energy Management Opportunities

An overall estimate should be made of how effectively the facility uses its energy resources. This is difficult to do in many cases because so many operations are unique. An idea can be obtained, however, by comparing similar buildings located in similar climates. Table 10.2 shows representative values and indicates the range in performance factors that is possible.

Next, areas or equipment that use the greatest amounts of electricity should be examined. Each item should be reviewed and these questions should be asked:

- Is this actually needed?
- How can the same equipment be used more efficiently?
- How can the same purpose be accomplished with less energy?
- Can the equipment be modified to use less energy?
- Would new, more efficient equipment be cost-effective?

10.2.1.4 Implement Changes

After certain actions to save energy have been identified, an economic analysis will be necessary to establish the economic benefits, and to determine if the cost of the action is justified (refer to Chapter 3 for guidance). Those changes that satisfy the economic criteria of the building owner (or occupant) will then be implemented. Economic criteria might include a minimum return on investment (e.g., 25%), a minimum payback period (e.g., 2 years) or a minimum benefit–cost ratio (e.g., 2.0).

10.2.1.5 Monitor the Program, Establish Goals

This is the final and perhaps most important step in the program. A continuing monitoring program is necessary to ensure that energy savings do not gradually disappear as personnel return to their old ways of operation, equipment gets out of calibration, needed maintenance is neglected, etc. Also, setting goals (they should be realistic) provides energy management personnel with targets against which they can gauge their performance and the success of their programs.

TABLE 10.2 Typical Energy Use Performance Factors [EUPFs]

Type of Facility	EUPF #1		EUPF #2	
	$kBTU/ft.^2yr$	MJ/m^2yr	kBTU/person·yr	GJ/person·yr
Small Office Building (300 m^2)	40	455	12	12.7
Engineering Office (1,000 m^2)	30	341	14	14.8
Elementary School (4,000 m^2)	70	796	5	5.3
Office Building (50,000 m^2)	100	1,138	50	53

10.2.1.6 Summary of Energy Management Programs

The way to move forward has been outlined in two tables to provide a step-by-step procedure for electrical energy management in buildings. Table 10.3 is directed at the homeowner or apartment manager, while Table 10.4 has been prepared for the commercial building owner or operator. Industrial facilities are treated separately (refer to Chapter 14).

One problem in performing the energy audit is determining the energy used by each item of equipment. In many cases, published data are available—as in Table 10.5 for residential appliances. In other cases, engineering judgments must be made, the manufacturer consulted, or instrumentation must be provided to actually measure energy use.

TABLE 10.3 An Energy Management Plan for the Homeowner or Apartment Manager

First Step: Review Historical Data

1. Collect utility bills for a recent 12-month period.
2. Add up the bills and calculate total kWh, total $, average kWh (divide total by 12), average $, and note the months with the lowest and highest kWh.
3 Calculate a seasonal variation factor (svf) by dividing the kWh for the greatest month by the kWh for the lowest month.

Second Step: Perform Energy Audits

4. Identify all electrical loads greater than 1 kW (1000 W). Refer to Table 5 for assistance. Most electrical appliances have labels indicating the wattage. If not, use the relation $W = V \times A$.
5. Estimate the number of hours per month each appliance is used.
6. Estimate the percentage of full load (pfl) by each device under normal use. For a lamp, it is 100%; for water heaters and refrigerators, which cycle on and off, about 30%, for a range, about 25% (only rarely are *all* burners *and* the oven used), etc.
7. For each device, calculate kWh by multiplying: kW x hours/month x pfl = kWh/month.
8. Add up all kWh calculated by this method. The total should be smaller than the average monthly kWh calculated in (2).
9. Note: if the svf is greater than 1.5, the load shows strong seasonal variation, e.g., summer air conditioning, winter heating, etc. If this is the case, make two sets of calculation, one for the lowest month (when the fewest loads are operating) and one for the highest month.
10. Make a table listing the wattage of each lamp and the estimated numbers of hours of use per month for each lamp. Multiply watts times hours for each, sum, and divide by 1000. This gives kWh for the lighting loads. Add this to the total shown.
11. Add the refrigerator, television, and all other appliances or tools that use 5 kWh per month or more.
12. By this process you should now have identified 80 to 90% of electricity using loads. Other small appliances that are used infrequently can be ignored. The test is to now compare with the average month (high or low month if svf is greater than 1.5). If your total is too high, you have over estimated the pfl or the hours of use.
13. Now rank each appliance in descending order of kWh used per month. Your list should read approximately like this:

First:	Heating (in cold climates); air conditioning would be first in hot climates.
Second:	Water heating
Third:	Lighting
Fourth:	Refrigeration
Fifth:	Cooking
Sixth:	Television
Seventh to last:	All others

Third Step: Apply Energy Management Principles

14. Attack the highest priority loads first. There are three general things that can be done: (1) reduce kW (smaller lamps, more efficient appliances); (2) reduce pfl ("oven cooked" meals, change thermostats, etc.); (3) reduce hours of use (turn lights off, etc.). Refer to the text for detailed suggestions.

Fourth Step: Monitor Program, Calculate Savings

15. After the energy management program has been initiated, examine subsequent utility bills to determine if you are succeeding.
16. Calculate savings by comparing utility bills. Note: since utility rates are rising, your utility bills may not be any lower. In this case it is informative to calculate what your bill would have been without the energy management program.

TABLE 10.4 An Energy Management Plan for Commercial Building Operator

First Step: Review Historical Data

1. Collect utility bills for a recent 12-month period.
2. Add up the bills and calculate total kWh, total $, average kWh (divide total by 12), average $, and note the months with the lowest and highest kWh.
3. Calculate a seasonal variation factor (svf) by dividing the kWh for the greatest month by the kWh for the lowest month.
4. Prepare a graph of historical energy use (see Figure 10.1).

Second Step: Perform Energy Audits

5. Evaluate major loads. In commercial buildings loads can be divided into four categories:
   a) HVAC (fans, pumps, chillers, heaters, cooling towers)
   b) Lighting
   c) Office equipment and appliances (elevators, typewriters, cash registers, copy machines, hot water heaters, etc.)
   d) Process equipment (as in laundries, restaurants, bakeries, shops, etc.) Items a, b, and c are common to all commercial operations and will be discussed here. Item d overlaps with industry and the reader should also refer to Chapters 11, 12, and 14. Generally items a, b, and d account for the greatest use of electricity and should be examined in that order.
6. In carrying out the energy audit, focus on major loads. Items that together comprise less than 1% of the total connected load in kW can often be ignored with little sacrifice in accuracy.
7. Use the methodology described above and in Chapter 16 for making the audit.
8. Compare audit results with historical energy use. If 80 to 90% of the total (according to the historical records) has been identified, this is generally adequate.

Third Step: Formulate the Energy Management Plan

9. Secure management commitment. The need for this varies with the size and complexity of the operation. However, any formal program will cost something, in terms of salary for the energy coordinator as well as (possibly) an investment in building modifications and new equipment. At this stage it is very important to project current energy usage and costs ahead for the next 3 to 5 years, make a preliminary estimate of potential savings (typically 10 to 50% per year), and establish the potential payback or return on investment in the program.
10. Develop a list of energy management opportunities (EMOs); e.g., install heat recovery equipment in building exhaust air), estimate the cost of each EMO, and also the payback. Methods for economic analysis are given in Chapter 3. For ideas and approaches useful for identifying EMO's, refer to the text.
11. Communicate the plan to employees, department heads, equipment operators, etc. Spell out who will do what, why there is a need, what are the potential benefits and savings. Make the point (if appropriate) that "the energy you save may save your job". If employees are informed, understand the purpose, and realize that the plan applies to everyone, including the President, cooperation is increased.
12. Set goals for department managers, building engineers, equipment operators, etc., and provide monthly reports so they can measure their performance.
13. Enlist the assistance of all personnel in: (1) better "housekeeping and operations", e.g., turning off lights, keeping doors closed: (2) locating obvious wastes of electricity; e.g., equipment operating needlessly, better methods of doing jobs.

Fourth Step: Implement Plan

14. Implementation should be done in two parts. First, carry out operational and housekeeping improvements with a goal of, say, 10% reduction in electricity use at essentially no cost and no reduction in quality of service or quantity of production. Second, carry out those modifications (retrofitting of buildings, new equipment, process changes) that have been shown to be economically attractive.
15. As changes are made it is important to continue to monitor electricity usage to determine if goals are being realized. Additional energy audits may be justified.

Fifth Step: Evaluate Progress, Management Report

16. Compare actual performance to the goals established in Item 12. Make corrections for weather variations, increases or decreases in production or number of employees, addition of new buildings, etc.
17. Provide a summary report of energy quantities and dollars saved, and prepare new plans for the future.

10.2.2 Electricity-Saving Techniques by Category of End Use

This section discusses strategies for saving energy that can be implemented in a short time at zero or low capital cost. Retrofit and new design strategies are then described. The ordering of topics corresponds

TABLE 10.5 Residential Energy Usage–Typical Appliances

Electric Appliances	Power (watts)	Typical Use (kWh/year)
	Home Entertainment	
Radio (solid-state)	10	10
Stereo	90	90
Color Television (solid-state)	100	229
Compact disc player	12	6
Video cassette recorder	30	15
Micro computer	125	63
Computer printer	250	25
Facsimile	65	75
	Food Preparation	
Blender	300	1
Broiler	1,140	85
Carving knife	92	8
Coffee maker	1,200	140
Deep fryer	1,448	83
Dishwasher	1,201	363
Egg cooker	516	14
Frying pan	1,196	100
Hot plate	1,200	90
Mixer	127	2
Microwave oven	1,300	170
Range		
Oven bake unit	3,200	288
Broil unit	3,600	168
Self-cleaning feature	4,000	192
Roaster	1,333	60
Sandwich grill	1,161	33
Toaster	1,146	39
Trash compactor	400	50
Waffle iron	1,200	20
Waste dispenser	445	7
	Refrigerator/freezer	
Top freezer (18.5 to 20.4 cubic ft.) Energy star July 2001 or newer	—	444
Side-by-side (20.5 to 22.4 cubic ft.) Energy star July 2001 or newer	—	612
	Laundry	
Electric clothes dryer	—	1,020
Iron (hand)	1,100	60
Washing machine (Energy star)	—	286
Water heater	2,475	4,219
	Housewares	
Clock	2	17
Floor Polisher	305	15
Sewing machine	75	11
Vacuum cleaner	630	46
	Comfort Conditioning	
Air cleaner	50	216
Air conditioner (room)	600	600
Bed covering	177	147
Dehumidifier	257	377
Fan (attic)	370	291

(continued)

TABLE 10.5 ***(Continued)***

Electric Appliances	Power (watts)	Typical Use (kWh/year)
Fan (circulating)	88	43
Fan (roll away)	171	138
Fan (window)	200	170
Heater (portable)	1,322	176
Heating pad	65	10
Humidifier	177	163
	Health and Beauty	
Germicidal lamp	20	141
Hair dryer	1,000	40
Heat lamp (infrared)	250	13
Shaver	15	0.5
Sun lamp	279	16
Tooth brush	1.1	1.0

approximately to their importance in terms of building energy use. Energy used specifically for a process (e.g., heating) is excluded except as it relates to buildings and their occupants.

10.2.2.1 Residential HVAC

Residential HVAC units using electricity are generally heat pumps, refrigeration systems, and electrical resistance heaters. Heaters range from electric furnace types, small radiant heaters, duct heaters, and strip or baseboard heaters to embedded floor or ceiling heating systems. Efficiency for heating is usually high because there are no stack or flue losses, and the heater transfers heat directly into the living space.

Cooling systems range from window air conditioning to central refrigeration or heat pump systems. Evaporative coolers are also used in some climates.

Principal operational and maintenance strategies for existing equipment include:

- System maintenance and cleanup
- Thermostat calibration and setback
- Microprocessor controls, time clocks, night cool down
- Improved controls and operating procedures
- Heated or cooled volume reduction
- Reduction of infiltration and exfiltration losses

System maintenance is an obvious but often neglected energy-saving tool. Dirty heat transfer surfaces decrease efficiency. Clogged filters increase pressure drops and pumping power. Inoperable or malfunctioning dampers can waste energy and prevent proper operation of the system.

In residential systems, the room thermostat generally controls heating and cooling. Thermostats should be set to 24°C (75°F) or higher for cooling and approximately 18°C (65°F) during the daytime for heating. As a first step, the calibration of the thermostat should be checked because these low-cost devices can be inaccurate by as much as ± 5°C. Several manufacturers now offer "smart" thermostats with microprocessor controls that can be programmed to set back or set forward the temperature depending on the time of day and day of week. By eliminating the need for manual control, they ensure that the settings will indeed be changed, whereas manual resetting of thermostats depends on occupant diligence. Some utilities have setup load control programs in which they can also communicate with smart thermostats and turn them down (or up) during high peak periods. A general rule of thumb is that for every 1°F of thermostat set back (heating) or set forward (cooling) during an 8-h period, there is a 1% savings in annual heating or cooling energy costs (the energy savings are generally lower in more severe climates).

Sometimes simple changes in controls or operating procedures will save energy. In cooling, use night air for summer cool down. When the outside air temperature is cool, turn off the refrigeration unit and circulate straight outside air. If fan units have more than one speed, use the lowest speed that provides satisfactory operation. Check the balance of the system and the operation of dampers and vents to insure that heating and cooling is provided in the correct quantities where needed.

Energy savings can be achieved by reducing the volume of the heated or cooled space. This can be accomplished by closing vents, doors, or other appropriate means. Usually it is not necessary to heat or cool an entire residence; the spare bedroom is rarely used, halls can be closed off, etc.

A major cause of energy wastage is air entering or leaving a home. Unintentional air transfer toward the inside is referred to as *infiltration*, and unintentional air transfer toward the outside is referred to as *exfiltration*. However, *infiltration* is often used to imply air leakage both into and out of a home, and this is the terminology used in this chapter. In a poorly "sealed" residence, infiltration of cold or hot air will increase heating or cooling energy use. According to Energy star, a typical home loses 25%–40% of its HVAC energy through infiltration. Infiltration also affects concentrations of indoor pollutants and can cause uncomfortable drafts. Air can infiltrate through numerous cracks and spaces created during building construction, such as those associated with electrical outlets, pipes, ducts, windows, doors, and gaps between ceilings, walls, floors, and so on. Infiltration results from temperature and pressure differences between the inside and outside of a home caused by wind, natural convection, and other forces. Major sources of air leakage are attic bypasses (paths within walls that connect conditioned spaces with the attic), fireplaces without dampers, leaky ductwork, window and door frames, and holes drilled in framing members for plumbing, electrical, and HVAC equipment. According to the U.S. DOE(Department of Energy) Energy Efficiency and Renewable Energy (EERE) program, the most significant source for infiltration is the combination of walls, ceilings, and floors that comprise 31% of the total infiltration in a typical home. Ducts (15%), fireplaces (14%), plumbing penetrations (13%), doors (11%), and windows (10%) are also substantial contributors to infiltration. Of lesser consequence are fans and vents (4%) and electrical outlets (2%).

To combat infiltration, builders of energy efficient homes use house wraps, caulking, foam insulation, tapes, and other seals. Sealing ducts in the home is also important to prevent the escape of heated or cooled air. Homeowners should also check for open doors and windows, open fireplace dampers, inadequate weather stripping around windows and doors, and any other openings that can be sealed. However, caution must be exercised to provide adequate ventilation. Standards vary, depending on the type of occupancy. Ventilation rates specified in the builder guidelines for the American Lung Association's Health House program state that for healthy homes "continuous general ventilation should be at least 1.0 cfm per 100 sq ft. of floor area plus at least 15 cfm for the first bedroom and 7.5 cfm for each additional bedroom." In addition, intermittent ventilation for the kitchen should be at least 100 cfm. For the bathrooms, rates should be 50 cfm intermittent or 20 cfm continuous. The Health House ventilation rates comply with ASHRAE standard 62.2.

In retrofit or new design projects the following techniques will save energy:

- Site selection and building orientation
- Building envelope design
- Selection of efficient heating/cooling equipment

Site selection and building orientation are not always under the control of the owner/occupant. Where possible, select a site sheltered from temperature extremes and wind. Orient the building (in cold climates) with a maximum southerly exposure to take advantage of direct solar heating in winter. Use earth berms to reduce heat losses on northerly exposed parts of the building. Deciduous trees provide summer shading but permit winter solar heating.

Building envelope design can improve heat absorption and retention in winter, and summer coolness. The first requirement is to design a well-insulated, thermally tight structure. Insulation made out of synthetic fibers reduces heating and cooling loads by resisting the transfer of heat through ceilings, walls, floors, and ducts. Reductions are usually proportionately higher for heating than for cooling because of generally larger indoor-to-outdoor temperature differences in winter than in summer. Insulation is available in bat, board, and loose-fill forms. The appropriate insulation material is selected on a basis of climate, building type, and recommended R-value. Higher R-values indicate better insulating properties. It is typically cost-effective to use greater-than-recommended R-values to improve energy efficiency above and beyond standard building practice.

Windows are an important source of heat gain and loss. The heat loss for single-pane glazing is around 5–7 W/m$^{2\circ}$C. For double glazing, the comparable value is in the range of 3–4 W/m$^{2\circ}$C, whereas for triple glazing it is 2–3 W/m$^{2\circ}$C. Window technology is constantly improving. Newer windows often have low emissivity (low-E), or spectrally selective coatings to prevent heat gain and/or loss. Low-E windows filled with argon gas have a heat loss rate of about 2 W/m$^{2\circ}$C. They have a higher visible transmittance, and are available with a low solar heat gain coefficient to reduce cooling loads in the summer. Low-E windows are available with an internal plastic film that essentially makes them triple glazed. The heat loss rate for these windows is on the order of 1 W/m$^{2\circ}$C.

Windows equipped with vinyl, wood, or fiberglass frames, or aluminum frames with a thermal barrier, provide the best insulation. It is also important to seal windows to prevent infiltration, as well as to use window coverings to minimize heat loss by radiation to the exterior during the evening. The appropriate placement of windows can also save energy by providing daylighting.

In general, the most efficient electric heating and cooling system is the heat pump. Common types are air-to-air heat pumps, either a single-package unit (similar to a window air conditioner), or a split system where the air handling equipment is inside the building and the compressor and related equipment are outdoors. Commercially available equipment demonstrates a wide range of efficiency. Heating performance is measured in terms of a heating seasonal performance factor (HSPF), in BTUs of heat added per Watt-hour of electricity input. Typical values are 6.8–9.0 and higher for the most efficient heat pumps. Cooling performance of residential heat pumps, air conditioners and packaged systems is measured in terms of a seasonal energy efficiency ratio (SEER), which describes the ratio of cooling capacity to electrical power input. Typical values are 10.0–14.5 and higher for the most efficient systems. The federal standards set in 1992 for air conditioners, heat pumps, and residential packaged units require a minimum SEER of 10.0 and a minimum HSPF of 6.8. New standards that will take effect in 2006 require a minimum SEER of 13 and a minimum HSPF of 7.7. Many existing older units have SEERs of 6–7, or roughly half the new minimum requirement. Therefore, substantial efficiency improvements are possible by replacing older equipment. In purchasing new equipment, selection of equipment with the highest HSPF and SEER should be considered. The higher initial cost of these units is almost always justified by operating savings. In addition, many utilities offer rebates for installing the more efficient units.

Sizing of equipment is important because the most efficient operation generally occurs at or near full load. Selection of oversized equipment is thus initially more expensive, and will also lead to greater operating costs.

The efficiency of heat pumps declines as the temperature difference between the heat source and heat sink decreases. Because outside air is generally the heat source, heat is most difficult to get when it is most needed. For this reason, heat pumps often have electrical backup heaters for extremely cold weather.

An alternate approach is to design the system using a heat source other than outside air. Examples include heated air (such as is exhausted from a building), a deep well (providing water at a constant year-round temperature), the ground, or a solar heat source. There are a great many variations on solar heating and heat pump combinations.

10.2.2.2 Nonresidential HVAC

HVAC systems in nonresidential installations may involve package rooftop or ground mounted units, or a central plant. Although the basic principles are similar to those discussed above in connection with residential systems, the equipment is larger and control more complex.

Efficiency of many existing HVAC systems can be improved. Modifications can reduce energy use by 10%–15%, often with building occupants unaware that changes have been made.

The basic function of HVAC systems is to heat, cool, dehumidify, humidify, and provide air mixing and ventilation. The energy required to carry out these functions depends on the building design, its duty cycle (e.g., 24 h/day use as in a hospital vs. 10 h/day in an office), the type of occupancy, the occupants' use patterns and training in the use of the HVAC system, the type of HVAC equipment installed, and finally, daily and seasonal temperature and weather conditions to which the building is exposed.

A complete discussion of psychometrics, HVAC system design, and commercially available equipment types is beyond the scope of this chapter.

Energy management strategies will be described in three parts:

- Equipment modifications (control, retrofit, and new designs)
    - Fans
    - Pumps
    - Packaged air conditioning units
    - Chillers
    - Ducts and dampers
    - Systems
- Economizer systems and enthalpy controllers
- Heat recovery techniques

10.2.2.2.1 Equipment Modifications (Control, Retrofit, and New Designs)

10.2.2.2.1.1 Fans

All HVAC systems involve some movement of air. The energy needed for this motion can make up a large portion of the total system energy used. This is especially true in moderate weather when the heating or cooling load drops off, but the distribution systems often operate at the same level.

Control. Simple control changes can save electrical energy in the operation of fans. Examples include turning off large fan systems when relatively few people are in the building, or stopping ventilation 30 min before the building closes. The types of changes that can be made will depend upon the specific facility. Some changes involve more sophisticated controls, which may already be available in the HVAC system.

Retrofit. The capacity of the building ventilation system is usually determined by the maximum cooling or heating load in the building. This load has been changing due to reduced outside air requirements, lower lighting levels, and wider acceptable comfort ranges. As a result, it is now feasible to decrease airflow in many existing commercial buildings as long as adequate indoor air quality is maintained.

The volume rate of airflow through a centrifugal fan, Q, varies directly with the speed of the impeller's rotation. This is expressed as follows for a fan whose speed is changed from N_1 to N_2:

$$Q_2 = (N_2/N_1) \times Q_1. \tag{10.1}$$

The pressure developed by the fan, P, (either static or total) varies as the square of the impeller speed:

$$P_2 = (N_2/N_1)^2 \times P_1. \tag{10.2}$$

The power needed to drive the fan, H, varies as the cube of the impeller speed:

$$H_2 = (N_2/N_1)^3 \times H_1. \tag{10.3}$$

The result of these laws is that for a given air distribution system (specified ducts, dampers, etc.), if the airflow is to be doubled, eight (2^3) times the power is needed. Conversely, if the airflow is to be cut in half, one-eighth ($\frac{1}{2}^3$) of the power is required. This is useful in HVAC systems because even a small reduction in airflow (e.g., 10%) can result in significant energy savings (27%).

The manner in which the airflow is reduced is critical in realizing these savings. Maximum savings are achieved by sizing the motor exactly to the requirements. Simply changing pulleys to provide the desired speed will also result in energy reductions according to the cubic law. The efficiency of existing fan motors tends to drop off below the half-load range.

If variable volume air delivery is required, it may be achieved through inlet vane control, outlet dampers, variable speed drives (VSDs), controlled pitch fans, or cycling. Energy efficiency in a retrofit design is best obtainable with variable speed drives on motors, or controlled pitch fans. This can be seen by calculating the power reduction that would accompany reduced flow using different methods of control, as noted below. Numbers in the table are the percent of full-flow input power:

	Fans			Pumps	
% Flow	Inlet Vanes	Dampers	VSDs	Throttle Valve	VSDs
100	102	103	102	101	103
90	86	98	76	96	77
80	73	94	58	89	58
70	64	88	43	83	41
60	56	81	31	77	30
50	50	74	22	71	19
40	46	67	15	65	13
30	41	59	9	59	8

New design. The parameters for new design are similar to those for fan retrofit. It is desirable, when possible, to use a varying ventilation rate that will decrease as the load decreases. A system such as *variable air volume* incorporates this in the interior zones of a building. In some cases, there will be a trade-off between power saved by running the fan slower and the additional power needed to generate colder air. The choices should be determined on a case-by-case basis.

10.2.2.2.1.2 Pumps

Pumps are found in a variety of HVAC applications such as chilled water, heating hot water, and condenser water loops. They are another piece of peripheral equipment that can use a large portion of HVAC energy, especially at low system loads.

Control. The control of pumps is often neglected in medium and large HVAC systems where it could significantly reduce the demand. A typical system would be a three-chiller installation where only one chiller is needed much of the year. Two chilled water pumps in parallel are designed to handle the maximum load through all three chillers. Even when only one chiller is on, both pumps are used. By manual adjustments, two chillers could be bypassed and one pump turned off. All systems should be reviewed in this manner to ensure that only the necessary pumps operate under normal load conditions.

Retrofit. Centrifugal pumps follow laws similar to fan laws, the key being the cubic relationship of power to the volume pumped through a given system. Small decreases in flow rate can save significant portions of energy.

In systems in which cooling or heating requirements have been permanently decreased, flow rates may also be reduced. A simple way to do this is by trimming the pump impeller. The pump curve must be checked first, however, because pump efficiency is a function of the impeller diameter, flow rate, and pressure rise. After trimming, one should ensure that the pump will still be operating in an efficient region. This is roughly the equivalent of changing fan pulleys in that the savings follow the cubic law of power reduction.

Another common method for decreasing flow rates is to use a “throttle” (pressure-reducing) valve. The result is equivalent to that of the discharge damper in the air-side systems. The valve creates an artificial use of energy that can be responsible for much of the work performed by the pumps. VSDs are more efficient.

New design. In a variable load situation common to most HVAC systems, more efficient systems with new designs are available, rather than the standard constant-volume pump (these may also apply to some retrofit situations).

One option is the use of several pumps of different capacities so that a smaller pump can be used when it can handle the load and a larger pump used the rest of the time. This can be a retrofit modification as well when a backup pump provides redundancy. Its impeller would be trimmed to provide the lower flow rate.

Another option is to use variable speed drive pumps. Although their initial cost is greater, they offer an improvement in efficiency over the standard pumps. The economic desirability of this or any similar change can be determined by estimating the number of hours the system will operate under various loads. Some utilities also offer rebates for installing variable speed pumps.

10.2.2.2.1.3 Package Air Conditioning Units

The most common space conditioning systems for commercial buildings are unitary equipment, either single-package systems or split systems. These are used for cooling approximately two-thirds of the air-conditioned commercial buildings in the U.S. For very large buildings or building complexes, absorption chillers or central chiller plants are used. Chillers are described in the following section.

Air conditioner efficiency is rated by one or more of three parameters: the energy efficiency ratio (EER), the SEER as described previously, and the integrated part load value (IPLV). The EER is easy to understand: it is the ratio of cooling capacity, expressed in BTU/h (kJ/h) to the power input required, in Watts. The SEER is a calculated ratio of the total annual cooling produced per annual electrical energy input in Watt-hours for units rated at less than 65,000 BTU/h. The IPLV is used for commercial loads on units rated at more than 65,000 BTU/h.

Great improvements in packaged air conditioner efficiency have been made in the last few decades, and new standards will increase efficiency even further. This is illustrated by the 30% increase in minimum SEER requirement that will take place in 2006 for split systems and single-package units under 65,000 BTU/h. As mentioned in the residential HVAC discussion earlier in the chapter, the current standard set in 1992 is a minimum SEER of 10.0 and the new standard will require a minimum SEER of 13.0. Standards for larger units ($>$65,000 BTU/h) will also continue to increase. As of October 29, 2001, the minimum federal standards for larger units are as follows:

	2001 Minimum Standard	
Equipment Size (BTU/h)	EER	IPLV
65,000 to $<$135,000	10.3	
135,000 to $<$240,000	9.7	
240,000 to $<$760,000	9.5	9.7
760,000 and larger	9.2	9.4

The Air Conditioning and Refrigeration Institute (ARI) is the basic rating agency in the U.S. In addition, there are both federal and state regulations, many of which either adopt or are similar to the ARI standards. Foreign manufacturers have similar rating systems.

Manufacturers sell systems with a broad range of efficiencies. For example, in the 65,000–135,000 BTU/h capacity range, it is possible to buy units with an EER as high as 12.5, even though the current federal standard is 10.3. Units with high EERs are typically more expensive, as the greater efficiency is achieved with larger heat exchange surface, more efficient motors, and so on.

To evaluate the economic benefit of the more efficient units, it is necessary to determine an annual operating profile that depends, in part, on the nature of the load and on the weather and temperature conditions at the site where the equipment will be installed. Or, an approximate method can be used. The American Society of Heating, Refrigeration, and Air Conditioning Engineers (ASHRAE) publishes tables that show typical "equivalent full-load operating hours" for different climate zones. These can be used to estimate the savings in electrical energy use over a year, and thereby determine if the added cost of a more efficient unit is justified (it almost always is).

Because the more efficient unit is almost always more cost-effective (except in light or intermittent load conditions), one might wonder why the less efficient units are sold. The reason is that many commercial buildings are constructed and sold by developers whose principal concern is keeping the initial cost of the building as low as practicable. They do not have to bear the annual operating expense of the building after it is sold, and therefore have no incentive to minimize operating expenses.

10.2.2.2.1.4 Chillers

Chillers are often the largest single energy user in the HVAC system. The chiller cools the water used to extract heat from the building and outside ventilation air. By optimizing chiller operation the performance of the whole system is improved.

Two basic types of chillers are found in commercial and industrial applications: absorption and vapor compression (mechanical) chillers. Absorption units boil water, the refrigerant, at a low pressure through absorption into a high-concentration lithium bromide solution. Mechanical chillers cool through evaporation of a refrigerant, such as Freon, at a low pressure after it has been compressed, cooled, and passed through an expansion valve.

There are three common types of mechanical chillers. They have similar thermodynamic properties, but use different types of compressors. Reciprocating and screw-type compressors are both positive displacement units. The centrifugal chiller uses a rapidly rotating impeller to pressurize the refrigerant.

All of these chillers must reject heat to a heat sink outside the building. Some use air-cooled condensers, but most large units operate with evaporative cooling towers. Cooling towers have the advantage of rejecting heat to a lower temperature heat sink because the water approaches the ambient wet-bulb temperature, whereas air-cooled units are limited to the dry-bulb temperature. As a result, air-cooled chillers have a higher condensing temperature that lowers the efficiency of the chiller. In full-load applications, air-cooled chillers require about 1–1.3 kW, or more, per ton of cooling, whereas water-cooled chillers usually require between 0.4 and 0.9 kW per ton. Air-cooled condensers are sometimes used because they require much less maintenance than cooling towers and have lower installation costs. They can also be desirable in areas of the country where water is scarce and/or water and water treatment costs are high because they do not depend on water for cooling.

Mechanical cooling can also be performed by direct expansion (DX) units. These are similar to chillers except that they cool the air directly. They eliminate the need for chilled water pumps, and also reduce efficiency losses associated with the transfer of the heat to and from the water. DX units must be located close ($\sim$30 m) to the ducts they are cooling, so they are typically limited in size to the cooling required for a single air handler. A single large chiller can serve a number of distributed air handlers. Where the air handlers are located close together, it can be more efficient to use a DX unit.

Control. Mechanical chillers operate on a principle similar to the heat pump. The objective is to remove heat from a low-temperature building and deposit it in a higher temperature atmosphere. The lower the temperature rise that the chiller has to face, the more efficiently it will operate. It is useful, therefore, to maintain as warm a chilled water loop and as cold a condenser water loop as possible.

Energy can be saved by using lower temperature water from the cooling tower to reject the heat. However, as the condenser temperature drops, the pressure differential across the expansion valve drops, starving the evaporator of refrigerant. Many units with expansion valves, therefore, operate at a constant condensing temperature, usually 41°C (105°F), even when more cooling is available from the cooling tower. Field experience has shown that in many systems, if the chiller is not fully loaded, it can be operated with a lower cooling tower temperature.

Retrofit. Where a heat load exists and the wet-bulb temperature is low, cooling can be done directly with the cooling tower. If proper filtering is available, the cooling tower water can be piped directly into the chilled water loop. Often a direct heat exchanger between the two loops is preferred to protect the coils from fouling. Another technique is to turn off the chiller but use its refrigerant to transfer heat between the two loops. This "thermocycle" uses the same principles as a heat pipe, and only works on chillers with the proper configuration.

A low wet-bulb temperature during the night can also be utilized. It requires a chiller that handles low condensing temperatures and a cold storage tank. This thermal energy storage (TES) technique is particularly desirable for consumers with access to time-of-use electricity rates that reward peak-shaving or load-shifting.

New design. In the purchase of a new chiller, an important consideration should be the load control feature. Because the chiller will be operating at partial load most of the time, it is important that it can do so efficiently.

In addition to control of single units, it is sometimes desirable to use multiple compressor reciprocating chillers. This allows some units to be shut down at partial load. The remaining compressors operate near full load, usually more efficiently.

If a new chiller is being installed to replace an old unit, or to retire equipment that uses environmentally unacceptable chlorinated fluorocarbon (CFC) refrigerants, it is a good opportunity to install a high-efficiency chiller. The Environmental Protection Agency (EPA) lists acceptable non-CFC substitutes for various types of chillers on their website. Examples include HCFC-123, HCFC-22, HFC-134a, HFC-227ea, HFC-245fa, and ammonia.

Commonly, in commercial and industrial buildings, a convenient source of heat for a heat pump is the building exhaust air. This is a constant source of warm air available throughout the heating season. A typical heat pump design could generate hot water for space heating from this source at around 32°C–35°C (90°F–95°F). Heat pumps designed specifically to use building exhaust air can reach 66°C (150°F).

Another application of the heat pump is a continuous loop of water traveling throughout the building with small heat pumps located in each zone. Each small pump can both heat and cool, depending upon the needs of the zone. This system can be used to transfer heat from the warm side of a building to the cool side. A supplemental cooling tower and boiler are included in the loop to compensate for net heating or cooling loads.

A double bundle condenser can be used as a retrofit design for a centralized system. This creates the option of pumping the heat higher to the cooling tower or into the heating system hot duct. Some chillers can be retrofitted to act as heat pumps. Centrifugal chillers will work much more effectively with a heat source warmer than outside air (exhaust air, for example). The compression efficiency of the centrifugal chiller falls off as the evaporator temperature drops.

Because they are positive displacement machines, reciprocating and screw-type compressors operate more effectively at lower evaporator temperatures. They can be used to transfer heat across a larger temperature differential. Multistage compressors increase this capacity even further.

10.2.2.2.1.5 Ducting-Dampers

Control. In HVAC systems using dual ducts, static pressure dampers are often placed near the start of the hot or cold plenum run. They control the pressure throughout the entire distribution system and can be indicators of system operation. Often in an overdesigned system, the static pressure dampers may never open more than 25%. Fan pulleys can be changed to slow the fan and open the dampers fully, eliminating the previous pressure drop. The same volume of air is delivered with a significant drop in fan power.

Retrofit. Other HVAC systems use constant volume mixing boxes for balancing, which create their own pressure drops as the static pressure increases. An entire system of these boxes could be over pressurized by several inches of water without affecting the airflow, but the required fan power would increase (one inch of water pressure is about 250 N/m^2 or 250 Pa). These systems should be monitored to ensure that static pressure is controlled at the lowest required value. It may also be desirable to replace the constant volume mixing boxes with boxes without volume control to eliminate their minimum pressure drop of approximately 1 in. of water. In this case, static pressure dampers will be necessary in the ducting.

Leakage in any dampers can cause a loss of hot or cold air. Neoprene seals can be added to blades to slow leakage considerably. If a damper leaks more than 10% it can be less costly to replace the entire damper assembly with effective positive closing damper blades rather than to tolerate the loss of energy.

New design. In the past, small ducts were installed because of their low initial cost despite the fact that the additional fan power required offset the initial cost on a life cycle basis. ASHRAE 90.1 guidelines now set a maximum limit on the fan power that can be used for a given cooling capacity. As a result, the air system pressure drop must be low enough to permit the desired airflow. In small buildings this pressure drop is often largest across filters, coils, and registers. In large buildings the duct runs may be responsible for a significant fraction of the total static pressure drop, particularly in high velocity systems.

10.2.2.2.1.6 Systems

The use of efficient equipment is only the first step in the optimum operation of a building. Equal emphasis should be placed upon the combination of elements in a system and the control of those elements.

Control. Many systems use a combination of hot and cold to achieve moderate temperatures. Included are dual duct, multizone, and terminal reheat systems, and some induction, variable air volume and fan coil units. Whenever combined heating and cooling occurs, the temperatures of the hot and cold ducts or water loops should be brought as close together as possible, while still maintaining building comfort.

This can be accomplished in a number of ways. Hot and cold duct temperatures are often reset on the basis of the temperature of the outside air or the return air. A more complex approach is to monitor the demand for heating and cooling in each zone. For example, in a multizone building, the demand of each zone is transferred back to the supply unit by electric or pneumatic signals. At the supply end hot and cold air are mixed in proportion to this demand. The cold air temperature should be just low enough to cool the zone calling for the most cooling. If the cold air were any colder, it would be mixed with hot air to achieve the right temperature. This creates an overlap in heating and cooling not only for the zone, but also for all the zones because they would all be mixing in the colder air.

If no zone calls for total cooling, then the cold air temperature can be increased gradually until the first zone requires cooling. At this point, the minimum cooling necessary for that multizone configuration is performed. The same operation can be performed with the hot air temperature until the first zone is calling for heating only.

Note that simultaneous heating and cooling is still occurring in the rest of the zones. This is not an ideal system, but it is a first step towards improving operating efficiency.

The technique for resetting hot and cold duct temperatures can be extended to the systems that have been mentioned. It may be performed automatically with pneumatic or electric controls, or manually. In some buildings, it will require the installation of more monitoring equipment (usually only in the zones of greatest demand), but the expense should be relatively small and the payback period short.

Nighttime temperature setback is another control option that can save energy without significantly affecting the comfort level. Energy is saved by shutting off, or cycling fans. Building heat loss may also be reduced because the building is cooler and no longer pressurized.

In moderate climates complete night shutdown can be used with a morning warm-up period. In colder areas where the overall night temperature is below 4°C (40°F), it is usually necessary to provide some heat during the night. Building setback temperature is partially dictated by the capacity of the heating system to warm the building in the morning. In some cases, it may be the mean radiant temperature of the building rather than air temperature that determines occupant comfort.

Some warm-up designs use "free" heating from people and lights to help attain the last few degrees of heat. This also provides a transition period for the occupants to adjust from the colder outdoor temperatures.

In some locations during the summer, it is desirable to use night air for a cool-down period. This "free cooling" can decrease the temperature of the building mass that has accumulated heat during the day. In certain types of massive buildings (such as libraries or buildings with thick walls), a long period of night cooling may decrease the building mass temperature by a degree or two. This represents a large amount of cooling that the chiller will not have to perform the following day.

Retrofit. Retrofitting HVAC systems may be an easy or difficult task depending upon the possibility of using existing equipment in a more efficient manner. Often retrofitting involves control or ducting changes that appear relatively minor, but which greatly increase the efficiency of the system. Some of these common changes, such as decreasing airflow, are discussed elsewhere in this chapter. This section will describe a few changes appropriate to particular systems.

Both dual duct and multizone systems mix hot and cold air to achieve the proper degree of heating or cooling. In most large buildings, the need for heating interior areas is essentially nonexistent, due to internal heat generation. A modification that adjusts for this is simply shutting off air to the hot duct. The mixing box then acts as a variable air volume box, modulating cold air according to room demand as relayed by the existing thermostat (it should be confirmed that the low volume from a particular box meets minimum air requirements).

Savings from this modification come mostly from the elimination of simultaneous heating and cooling. Since fans in these systems are likely to be controlled by static pressure dampers in the duct after the fan, they do not unload very efficiently and represent only a small portion of the savings.

10.2.2.2.2 Economizer Systems and Enthalpy Controllers

The economizer cycle is a technique for introducing varying amounts of outside air to the mixed air duct. Basically it permits mixing warm return air at 24°C (75°F) with cold, outside air to maintain a preset temperature in the mixed air plenum (typically 10°C–15°C, 50°F–60°F). When the outside temperature is slightly above this set point, 100% outside air is used to provide as much of the cooling as possible. During very hot outside weather, minimum outside air will be added to the system.

A major downfall of economizer systems is poor maintenance. The failure of the motor or dampers may not cause a noticeable comfort change in the building because the system is often capable of handling the additional load. Since the problem is not readily apparent, corrective maintenance may be put off indefinitely. In the meantime, the HVAC system will be working harder than necessary for any economizer installation.

Typically, economizers are controlled by the dry-bulb temperature of the outside air, rather than its enthalpy (actual heat content). This is adequate most of the time, but can lead to unnecessary cooling of air. When enthalpy controls are used to measure wet-bulb temperatures, this cooling can be reduced. However, enthalpy controllers are more expensive and less reliable.

The rules that govern the more complex enthalpy controls for cooling-only applications are as follows:

- When outside air enthalpy is greater than that of the return air, or when outside air dry-bulb temperature is greater than that of the return air, use minimum outside air.
- When the outside air enthalpy is below the return air enthalpy, and the outside dry-bulb temperature is below the return air dry-bulb temperature but above the cooling coil control point, use 100% outside air.
- When outside air enthalpy is below the return air enthalpy and the outside air dry-bulb temperature is below the return air dry-bulb temperature and below the cooling coil controller setting, the return and outside air are mixed by modulating dampers according to the cooling set point.

These points are valid for the majority of cases. When mixed air is to be used for heating and cooling, a more intricate optimization plan will be necessary, which is based on the value of the fuels used for heating and cooling.

10.2.2.2.3 Heat Recovery

Heat recovery is often practiced in industrial processes that involve high temperatures. It can also be employed in HVAC systems.

Systems are available that operate with direct heat transfer from the exhaust air to the inlet air. These are most reasonable when there is a large volume of exhaust air, for example, in once-through systems, and when weather conditions are not moderate.

Common heat recovery systems are broken down into two types: regenerative and recuperative. Regenerative units use alternating airflow from the hot and cold stream over the same heat storage/transfer medium. This flow may be reversed by dampers, or the whole heat exchanger may rotate between streams. Recuperative units involve continuous flow; the emphasis is upon heat transfer through a medium with little storage.

The rotary regenerative unit, or heat wheel, is one of the common heat recovery devices. It contains a corrugated or woven heat storage material that gains heat in the hot stream. This material is then rotated into the cold stream where the heat is given off again. The wheels can be impregnated with a desiccant to transfer latent as well as sensible heat. Purge sections for HVAC applications can reduce carry-over from the exhaust stream to acceptable limits for most installations.

The heat transfer efficiency of heat wheels generally ranges from 60 to 85% depending upon the installation, type of media, and air velocity. For easiest installation, the intake and exhaust ducts should be located near each other.

Another system that can be employed with convenient duct location is a plate type air-to-air heat exchanger. This system is usually lighter though more voluminous than heat wheels. Heat transfer efficiency is typically in the 60%–75% range. Individual units range from 1000 to 11,000 SCFM and can be grouped together for greater capacity. Almost all designs employ counterflow heat transfer for maximum efficiency.

Another option to consider for nearly contiguous ducts is the heat pipe. This is a unit that uses a boiling refrigerant within a closed pipe to transfer heat. Since the heat of vaporization is utilized, a great deal of heat transfer can take place in a small space.

Heat pipes are often used in double-wide coils which look very much like two steam coils fastened together. The amount of heat transferred can be varied by tilting the tubes to increase or decrease the flow of liquid through capillary action. Heat pipes cannot be "turned off," so bypass ducting is often desirable. The efficiency of heat transfer ranges from 55 to 75%, depending upon the number of pipes, fins per inch, air face velocity, etc.

Runaround systems are also popular for HVAC applications, particularly when the supply and exhaust plenums are not physically close. Runaround systems involve two coils (air-to-water heat exchangers) connected by a piping loop of water or glycol solution and a small pump. The glycol solution is necessary if the air temperatures in the inlet coils are below freezing.

Standard air conditioning coils can be used for the runaround system, and some equipment manufacturers supply computer programs for size optimization. Precaution should be used when the exhaust air temperature drops below 0°C (32°F) because this can cause condensed water on the system's fins to freeze. A three-way bypass valve will maintain the temperature of the solution entering the coil at just above 0°C (32°F). The heat transfer efficiency of this system ranges from 60 to 75% depending upon the installation.

Another system similar to the runaround in layout is the desiccant spray system. Instead of using coils in the air plenums, it uses spray towers. The heat transfer fluid is a desiccant (lithium chloride) which transfers both latent and sensible heat; this is desirable in many applications. Tower capacities range from 7700 to 92,000 SCFM; multiple units can be used in large installations. The enthalpy recovery efficiency is in the range of 60%–65%.

10.2.2.2.4 Thermal Energy Storage

TES systems are used to reduce the on-peak electricity demand caused by large cooling loads. TES systems utilize several different storage media, with chilled water, ice, or eutectic salts being most common. Chilled water requires the most space, with the water typically being stored in underground tanks. Ice storage systems can be aboveground, insulated tanks with heat exchanger coils that cause the water to freeze, or can be one of several types of ice-making machines.

In a typical system, a chiller operates during off-peak hours to make ice (usually at night). Because the chiller can operate for a longer period of time than during the daily peak, it can have a smaller capacity. Efficiency is greater at night, when the condensing temperature is lower than it is during the day. During daytime operation, chilled water pumps circulate water through the ice storage system and extract heat. Systems can be designed to meet the entire load, or to meet a partial load, with an auxiliary chiller as a backup.

This system reduces peak demand and can also reduce energy use. With ice storage, it is possible to deliver water at a lower temperature than is normally done. This means that the chilled water piping can be smaller and the pumping power reduced. A low-temperature air distribution system will allow smaller ducts and lower capacity fans to deliver a given amount of cooling. Careful attention must be paid to the system design to insure occupant comfort in conditioned spaces. Some government agencies and electric utilities offer incentives to customers installing TES systems. The utility incentives could be in the form of a set rebate per ton of capacity, per kW of deferred peak demand, or a time-of-use pricing structure that favors TES.

10.2.2.3 Water Heating

10.2.2.3.1 Residential Systems

Residential water heaters typically range in size from 76 L (20 gal) to 303 L (80 gal). Electric units generally have one or two immersion heaters, each rated at 2–6 kW, depending on tank size. Energy input for water heating is a function of the temperature at which water is delivered, the supply water temperature, and standby losses from the water heater, storage tanks, and piping.

The efficiency of water heaters is referred to as the energy factor (EF). Higher EF values equate to more efficient water heaters. Typical EF values range from about 0.8–0.95 for electric resistance heaters, 0.5–0.8 for natural gas units, 0.7–0.85 for oil units, and 1.5–2.0 for heat pump water heaters.

In single-tank residential systems, major savings can be obtained by:

- Thermostat temperature setback to 60°C (140°F)
- Automated control
- Supplementary tank insulation
- Hot water piping insulation

The major source of heat loss from electric water heaters is standby losses through the tank walls and from piping because there are no flame or stack losses in electric units. The heat loss is proportional to the temperature difference between the tank and its surroundings. Thus lowering the temperature to 60°C will result in two savings: (1) a reduction in the energy needed to heat water, and (2) a reduction in the amount of heat lost. Residential hot water uses do not require temperatures in excess of 60°C; for any special use which does, it would be advantageous to provide a booster heater to meet this requirement when needed, rather than maintain 100–200 L of water continuously at this temperature with associated losses.

When the tank is charged with cold water, both heating elements operate until the temperature reaches a set point. After this initial rise, one heating element thermostatically cycles on and off to maintain the temperature, replacing heat that is removed by withdrawing hot water, or that is lost by conduction and convection during standby operation.

Experiments indicate that the heating elements may be energized only 10%–20% of the time, depending on the ambient temperature, demand for hot water, water supply temperature, etc. By carefully scheduling hot water usage, this time can be greatly reduced. In one case, a residential water heater was operated for 1 h in the morning and 1 h in the evening. The morning cycle provided sufficient hot water for clothes washing, dishes, and other needs. Throughout the day the water in the tank, although gradually cooling, was still sufficiently hot for incidental needs. The evening heating cycle provided sufficient water for cooking, washing dishes, and bathing. Standby losses were reduced. Electricity use was cut to a fraction of the normal amount.

This method requires the installation of a time clock or other type of control to regulate the water heater. A manual override can be provided to meet special needs.

Supplementary tank insulation can be installed at a low cost to reduce standby losses. The economic benefit depends on the price of electricity and the type of insulation installed. However, paybacks of a few months up to a year are typical on older water heaters. Newer units have better insulation and reduced losses. Hot water piping should also be insulated, particularly when hot water tanks are located outside or when there are long piping runs. If copper pipe is used, it is particularly important to insulate the pipe for the first 3–5 m where it joins the tank because it can provide an efficient heat conduction path.

Because the energy input depends on the water flow rate and the temperature difference between the supply water temperature and the hot water discharge temperature, reducing either of these two quantities reduces energy use. Hot water demand can be reduced by cold water clothes washing, and by providing hot water at or near the use temperature, so as to avoid the need for dilution with cold water. Supply water should be provided at the warmest temperature possible. Because reservoirs and underground piping systems are generally warmer than the air temperature on a winter day in a cold climate, supply piping should be buried, insulated, or otherwise kept above the ambient temperature.

Solar systems are available today for heating hot water. Simple inexpensive systems can preheat the water, reducing the amount of electricity needed to reach the final temperature. Alternatively, solar heaters (some with electric backup heaters) are also available, although initial costs may be prohibitively high depending on the particular installation.

Heat pump water heaters may save as much as 25%–30% of the electricity used by a conventional electric water heater. Some utilities have offered rebates of several thousand dollars to encourage customers to install heat pump water heaters.

Tankless water heaters, also called *on-demand* or *instantaneous* water heaters, have been used for decades in Asia, Europe, and South America, but only recently have become common in the United States. Both electric and gas versions are available. The basic operating principal is that the flow of water actuates the gas-fired or electric heating element. A microcomputer senses water flow rate and temperature and controls the heat input accordingly. There are two key advantages. First, standby losses are virtually eliminated because the unit only produces hot water when it is required. Second, the units are very compact and require a fraction of the space required by tank-type water heaters. Efficiencies (EF) are in the range of 0.80–0.82 for gas units and around 0.95 for electric units. For gas units, if equipped with an electric igniter, the pilot light losses are also eliminated. The larger units match the performance of tank-type water heaters. Tests performed by the authors in Southern California show that hot water supply is adequate for a typical three- or four-bedroom residence.

The microwave water heater is another interesting technology that is just beginning to emerge in both residential and commercial applications. Microwave water heaters are also tankless systems that produce hot water only when needed, thereby avoiding the energy losses incurred by conventional water heaters during the storage of hot water. These heaters consist of a closed, stainless steel chamber with a silica-based flexible coil and a magnetron. When there is a demand for hot water, either because a user has opened a tap or because of a heater timing device, water flows into the coil and the magnetron bombards it with microwave energy at a frequency of 2450 MHz. The microwave energy excites the water molecules, heating the water to the required temperature.

Heat recovery is another technique for preheating or heating water, although opportunities in residences are limited. This is discussed in more detail under commercial water heating.

Apartments and larger buildings use a combined water heater/storage tank, a circulation loop, and a circulating pump. Cold water is supplied to the tank, which thermostatically maintains a preset temperature, typically 71°C (160°F). The circulating pump maintains a flow of water through the circulating loop, so hot water is always available instantaneously upon demand to any user. This method is also used in hotels, office buildings, etc.

Adequate piping and tank insulation is even more important here because the systems are larger and operate at higher temperature. The circulating hot water line should be very well insulated. Otherwise, it will dissipate heat continuously.

10.2.2.3.2 Heat Recovery in Nonresidential Systems

Commercial/industrial hot water systems offer many opportunities for employing heat recovery. Examples of possible sources of heat include air compressors, chillers, heat pumps, refrigeration systems, and water-cooled equipment. Heat recovery permits a double energy savings in many cases. First, recovery of heat for hot water or space heating reduces the direct energy input needed for heating. The secondary benefit comes from reducing the energy used to dissipate waste heat to a heat sink (usually the atmosphere). This includes using energy and energy expended to operate cooling towers and heat exchangers.

Solar hot water systems are also finding increasing use. Interestingly, the prerequisites for solar hot water systems also permit heat recovery. After the hot water storage capacity and backup heating capability has been provided for the solar hot water system, it is economical to tie in other sources of waste heat (e.g., water jackets on air compressors).

10.2.2.4 Lighting

There are seven basic techniques for improving the efficiency of lighting systems:

- Delamping
- Relamping
- Improved controls
- More efficient lamps and devices
- Task-oriented lighting
- Increased use of daylight
- Room color changes, lamp maintenance

The first two techniques and possibly the third are low cost and may be considered operational changes. The last four items generally involve retrofit or new designs.

The first step in reviewing lighting electricity use is to perform a lighting survey. An inexpensive handheld light meter can be used as a first approximation; however, distinction must be made between raw intensities (lux or foot-candles) and illumination quality recorded in this way.

Many variables can affect the "correct" lighting values for a particular task: task complexity, age of employee, glare, etc. For reliable results, consult a lighting specialist, or refer to the literature and publications of the Illuminating Engineering Society.

The lighting survey indicates those areas of the building where lighting is potentially inadequate or excessive. Deviations from illumination levels that are adequate can occur for several reasons: over design, building changes, change of occupancy, modified layout of equipment or personnel, more efficient lamps, improper use of equipment, dirt buildup, etc.

After the building manager has identified areas with potentially excessive illumination levels, he or she can apply one or more of the seven techniques previously listed. Each of these will be described briefly.

Delamping refers to the removal of lamps to reduce illumination to acceptable levels. With incandescent lamps, bulbs are removed. With fluorescent or high intensity discharge lamps, ballasts account for 10%–20% of total energy use and should be disconnected after lamps are removed. However, delamping often changes the distribution of the lighting, as well as the overall level, and is therefore generally not recommended.

Relamping refers to the replacement of existing lamps with lamps of lower wattage or increased efficiency. Low wattage fluorescent tubes that require 15%–20% less wattage, but produce 10–15% less light, are available. In Chapter 12, Section 2, Energy Efficient Lighting Technologies, there is a discussion on the idea of using lower-power-factor electronic ballasts, as well as further details on energy efficient lighting technologies. In some types of high-intensity discharge (HID) lamps, a more efficient lamp can be substituted directly (it too will have a lower lumen output). However, in most cases, ballasts must also be changed.

Improved controls permit lamps to be used only when and where needed. For example, certain office buildings have all lights for one floor on a single contactor. These lamps will be switched on at 6:00 a.m. before work begins, and are not turned off until 10:00 p.m. when maintenance personnel finish their cleanup duties.

Energy usage can be cut by as much as 50% by installing individual switches for each office or work area, installing timers, occupancy sensors, photocell controls, and/or dimmers, or by instructing custodial crews to turn lights on as needed and turn them off when work is complete. Sophisticated building-wide lighting control systems are also available, and are being implemented at an increasing rate, particularly in commercial buildings.

There is a great variation in the efficacy (a measure of light output per unit of electricity input) of various lamps. Incandescent lamps have the lowest efficacy, typically 5–20 lm/W. Wherever possible, fluorescent lamps should be substituted for incandescent lamps. This not only saves energy, but offers substantial economic savings as well because fluorescent lamps last 10–50 times longer. Conventional fluorescent lamps have efficacies in the range of 30–70 lm/W; high-efficiency fluorescent systems currently yield about 85–100 lm/W.

Compact fluorescent lamps can also be used as substitutes for a wide range of incandescent lamps. They are available in wattages of 5–70 W and will replace incandescent lamps with about one-fourth of the energy consumption. In addition to the energy savings, they have a 10,000-h rated life, and thus do not need to be replaced as often as incandescent lamps.

Still greater improvements are possible with HID lamps such as metal halide and high-pressure sodium lamps. Although they are generally not suited to residential use (high light output and high capital cost), they are increasingly being used in commercial and industrial buildings for their high efficiency and long life. It should be noted that HID lamps are not suited for any area where lamp switching is required, as they still take several minutes to restart after being switched off.

Improving ballasts is another way of saving lighting energy. One example of a significant increase in efficacy in the commercial sector is illustrated in the transition from T12 (1.5-in. diameter) fluorescent lamps with magnetic ballasts to T8 (1-in. diameter) fluorescent lamps with electronic ballasts. This transition began to occur in the late 1970s and early 1980s. Now T8 electronic ballast systems are the standard for new construction and retrofits (note that competition from new T5 systems is increasing). The efficacy improvement, depending on the fixture, is roughly 20%–40%, or even more. For example, a two-lamp F34T12 fixture (a fixture with two 1.5-in. diameter, 34-W lamps) with energy-saving magnetic ballast requires 76 W, whereas a two-lamp F32T8 fixture (a fixture with two 1-in. diameter, 32-W lamps) with electronic ballast requires only 59 W, which is an electricity savings of 22%. The savings is attributable to the lower wattage lamps as well as the considerably more efficient ballast.

In addition to the lighting energy savings, there are additional savings from the reduced air conditioning load due to less heat output from the ballasts.

Task-oriented lighting is another important lighting concept. In this approach, lighting is provided for work areas in proportion to the needs of the task. Hallways, storage areas, and other nonwork areas receive less illumination. Task lighting can replace the so-called "uniform illumination" method sometimes used in office buildings. The rationale for uniform illumination was based on the fact that the designer could never know the exact layout of desks and equipment in advance, so uniform illumination was provided. This also accommodates revisions in the floor plan. Originally, electricity was cheap and any added cost was inconsequential.

Daylighting was an important element of building design for centuries before the discovery of electricity. In certain types of buildings and operations today, daylighting can be utilized to at least reduce (if not replace) electric lighting. Techniques include windows, an atrium, skylights, etc. There are obvious limitations, such as those imposed by the need for privacy, 24-h operation, and building core locations with no access to natural light.

The final step is to review building and room color schemes and decor. The use of light colors can substantially enhance illumination without modifying existing lamps.

An effective lamp maintenance program also has important benefits. Light output gradually decreases over lamp lifetime. This should be considered in the initial design and when deciding on lamp replacement. Dirt can substantially reduce light output; simply cleaning lamps and luminaries more frequently can gain up to 5%–10% greater illumination, permitting some lamps to be removed.

10.2.2.5 Refrigeration

The refrigerator, at roughly 40–140 kWh/month depending on the size and age of the model, is among the top six residential users of electricity. In the last 50 years, the design of refrigerator/freezers has changed considerably, with sizes increasing from 5–10 to 20–25 ft.^{3} today. At the same time, the energy input per unit increased up until the oil embargo, after which efforts were made that led to a steady decline in energy use per unit between the mid-1970s through today, despite increases in average refrigerator size. The net result is that current refrigerators use about as much energy as those half their size did 40 years ago.

Energy losses in refrigerators arise from a variety of sources. The largest losses are due to heat gains through the walls and door openings. Because much of the energy used by a refrigerator depends on its design, care should be used in selection. Look for Energy star models to maximize efficiency.

Purchase of a new, more efficient unit is not a viable option for many individuals who have a serviceable unit and do not wish to replace it. In this case, the energy management challenge is to obtain the most effective operation of the existing equipment.

More efficient operation of refrigeration equipment can be achieved by:

- Better insulation (although this is not generally a user option for refrigerators)
- Disconnecting or reducing operation of automatic defrost and anti-sweat heaters
- Providing a cool location for the refrigerator coils (or reduce room temperature); also clean coils frequently
- Reducing the number of door openings
- Increasing temperature settings
- Precooling foods before refrigerating

Commercial refrigeration systems are found in supermarkets, liquor stores, restaurants, hospitals, hotels, schools, and other institutions—about one-fifth of all commercial facilities. Systems include walk-in dairy cases, open refrigerated cases, and freezer cases. In a typical supermarket, lighting, HVAC, and miscellaneous uses account for half the electricity use, while refrigerated display cases, compressors, and condenser fans account for the other half. Thus commercial refrigeration can be an important element of electric energy efficiency.

It is common practice in some types of units to have the compressor and heat exchange equipment located remotely from the refrigerator compartment. In such systems, a cool location should be selected, rater than locating the compressor next to other equipment which gives off heat. Many of the newer commercial refrigerators now come equipped with heat recovery systems, which recover compressor heat for space conditioning or water heating.

Walk-in freezers and refrigerators lose energy though door openings; refrigerated display cases have direct transfer of heat. Covers, strip curtains, air curtains, glass doors, or other thermal barriers can help mitigate these problems. The most efficient light sources should be used in large refrigerators and freezers; every 1 W eliminated saves 3 W. Elimination of 1 W of electricity to produce light also eliminates an additional 2 W required to extract the heat. Other improvements can reduce energy use, including high-efficiency motors, variable speed drives, more efficient compressors, and improved refrigeration cycles and controls.

10.2.2.6 Cooking

Cooking accounted for about 7% of residential electricity use in 2001 and about 2.1% of commercial electricity use in 1999. Since the last edition of this handbook, consumer behavior toward cooking has changed. Consumers today are cooking less in the home and dining out or picking up prepared food more often. Figure 10.5 compares the number of hot meals cooked in the home during the years of 1993 and 2001. The data show that fewer households cooked one or more hot meals per day in 2001 than in 1993; in addition, more households cooked a few meals or less per week in 2001 than in 1993. Consumers are also purchasing more foods that are easier to prepare—convenience is key to the modern family. In response, the food processing industry offers a wide variety of pre-prepared, ready-to-eat products.

Though habits are changing and more cooking is occurring outside the home (in restaurants and food processing facilities), reductions in energy use are still important. In general, improvements in energy use efficiency for cooking can be divided into three categories:

- More efficient use of existing appliances
- Use of most efficient existing appliances
- More efficient new appliances

The most efficient use of existing appliances can lead to substantial reductions in energy use. Although slanted towards electric ranges and appliances, the following observations also apply to cooking devices using other sources of heat.

First, select the right size equipment for the job. Do not heat excessive masses or large surface areas that will needlessly radiate heat. Second, optimize heat transfer by ensuring that pots and pans provide good thermal coupling to the heat sources. Flat-bottomed pans should be used on electric ranges. Third, be sure that pans are covered to prevent heat loss and to shorten cooking times. Fourth, when using the oven, plan meals so several dishes are cooked at once. Use small appliances (electric fry pans, "slow" cookers, toaster ovens, etc.) whenever they can be substituted efficiently for the larger appliances such as the oven.

Different appliances perform similar cooking tasks with widely varying efficiencies. For example, the electricity used and cooking time required for common food items can vary as much as ten to one in energy use and five to one in cooking times, depending on the method. As an example, four baked potatoes require 2.3 kWh and 60 min in an oven (5.2 kW) of an electric range, 0.5 kWh and 75 min in a

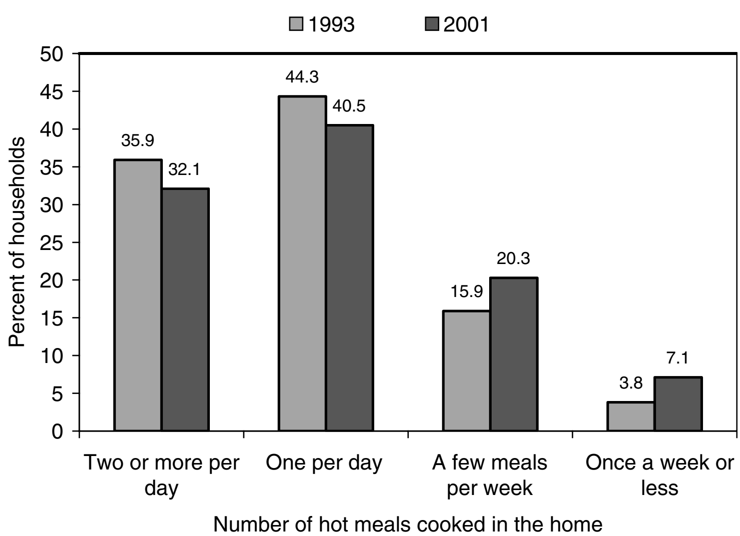


FIGURE 10.5 Number of hot meals cooked in the home for all U.S. households, 1993 and 2001.

toaster oven (1.0 kW), or 0.3 kWh and 16 min in a microwave oven (1.3 kW). Small appliances are generally more efficient when used as intended. Measurements in a home indicated that a pop-up toaster cooks two slices of bread using only 0.025 kWh. The toaster would be more efficient than using the broiler in the electric range oven, unless a large number of slices of bread (more than 17 in this case) were to be toasted at once.

Cooking several dishes at once can save energy. A complete meal, consisting of a ham (2.3 kg), frozen peas (0.23 kg), four yams, and a pineapple upside-down cake (23×23 cm^2) were cooked separately using a toaster oven and an electric range, together in an electric oven, and separately in a microwave oven. Cooked separately using the toaster oven and range required 5.2 kWh; together in the oven took 2.5 kWh, whereas the microwave required 1.2 kWh.

If new appliances are being purchased, select the most efficient ones available. Heat losses from a conventional oven approach 1 kW, with insulation accounting for about 50%; losses around the oven door edge and through the window are next in importance. These losses are reduced in certain models. Self-cleaning ovens are normally manufactured with more insulation. Careful design of heating elements can also contribute to better heat transfer. Typically, household electric ranges require around 3200 W for oven use, 3600 W for broiler use, and 4000 W for use of the self-cleaning feature.

Microwave cooking is highly efficient for many types of foods because the microwave energy is deposited directly in the food. Energy input is minimized because there is no need to heat the cooking utensil. Although many common foods can be prepared effectively using a microwave oven, different methods must be used, as certain foods are not suitable for microwave cooking. A typical microwave oven requires about 1300 W. Convection ovens and induction cook tops are two newer developments that may also reduce cooking energy use.

Commercial cooking operations range from small restaurants and cafes, where methods similar to those described above for residences are practiced, to large institutional kitchens in hotels, hospitals, and finally, to food processing plants.

Many of the same techniques apply. Microwave heating is finding increasing use in hotel and restaurant cooking. Careful scheduling of equipment use, and provision of several small units rather than a single large one, will save energy. For example, in restaurants, grills, soup kettles, bread warmers, etc., often operate continuously. Generally it is unnecessary to have full capacity during off-peak hours; one small grill might handle mid-morning and mid-afternoon needs, permitting the second and third units to be shut down. The same strategy can be applied to coffee warming stations, hot plates, etc.

In food processing plants where food is cooked and canned, heat recovery is an important technique. Normally, heat is rejected via cooling water at some step in the process. This heat can be recovered and used to preheat products entering the process, decreasing the amount of heating that eventually must be done.

10.2.2.7 Residential Appliances

A complete discussion of energy management opportunities associated with all the appliances found in homes is beyond the scope of this chapter. However, several of the major ones will be discussed and general suggestions applicable to the others will be given.

10.2.2.7.1 Clothes Drying

Clothes dryers typically use about 2.5 kWh per load. A parameter called the *energy factor*, which is a measure of the pounds of clothing dried per kWh of electricity consumed, can be used to quantify the efficiency of clothes drying. The minimum EF for a standard capacity electric dryer is 3.01. In the U.S., new dryers are not required to display energy use information, so it is difficult to compare models. In fact, most electric dryers on the market are comparable in their construction and the basic heating technology. However, the actual energy consumption of the dryer varies with the types of controls it has, and how the operator uses those controls. Models with moisture-sensing capability can result in the most energy savings—savings on the order of 15% compared to conventional operation are common.

In addition, electric clothes dryers operate most efficiently when fully loaded. Operating with one-third to one-half load costs roughly 10%–15% in energy efficiency.

Locating clothes dryers in heated spaces could save 10%–20% of the energy used by reducing energy needed for heating up. Another approach is to save up loads and do several loads sequentially, so the dryer does not cool down between loads.

The heavier the clothes the greater the amount of water they hold. Mechanical water removal (pressing, spinning, wringing) generally requires less energy than electric heat. Therefore, be certain the washing machine goes through a complete spin cycle (0.1 kWh) before putting clothes in the dryer.

Solar drying, which requires a clothesline (rope) and two poles or trees, has been practiced for millennia and is very sparing of electricity. The chief limitation is, of course, inclement weather. New technologies such as microwave or heat pump clothes dryers may help reduce clothes drying energy consumption in the future.

10.2.2.7.2 Clothes Washing

The modified energy factor (MEF) can be used to compare different models of clothes washers. It is a measure of the machine energy required during washing, the water heating energy, and the dryer energy needed to remove the remaining moisture. A higher MEF value indicates a more efficient clothes washer. All new clothes washers manufactured or imported after January 1, 2004 are required to have an MEF of at least 1.04, and after January 1, 2007, an MEF of at least 1.26. In addition, as of January 2004, a minimum MEF of 1.42 is required for a clothes dryer to be qualified as an Energy star unit. A typical Energy star unit uses about 0.7 kWh per load.

Electric clothes washers are designed for typical loads of 3–7 kg. Surprisingly, most of the energy used in clothes washing is for hot water; the washer itself only requires a few percent of the total energy input. Therefore, the major opportunity for energy management in clothes washing is the use of cold or warm water for washing. Under normal household conditions it is not necessary to use hot water. Clothes are just as clean (in terms of bacteria count) after a 20°C wash as after a 50°C wash. If there is concern for sanitation (e.g., a sick person in the house), authorities recommend use of chlorine bleach. If special cleaning is required, such as removing oil or grease stains, hot water (50°C) and detergent will emulsify oil and fat. There is no benefit in a hot rinse.

A secondary savings can come from using full loads. Surveys indicate that machines are frequently operated with partial loads, even though a full load of hot water is used.

10.2.2.7.3 Dishwashers

The two major energy uses in electric dishwashers are the hot water and the dry cycle. Depending on the efficiency of the model and operation, dishwashers use between 2 and 5 kWh per load.

The water heating often accounts for 80% of the total energy requirement of a dishwasher. The volume of hot water used ranges from about 5 gal for the more efficient units to more than double that for less efficient models. The water volume can be varied on some machines depending on the load.

Some models also allow for a no-heat drying option. If not available, stop the cycle prior to the drying step and let the dishes air-dry. Operating the dishwasher with a full load and using a cold water prerinse are additional ways to minimize energy use.

10.2.2.7.4 General Suggestions for Residential Appliances and Electrical Equipment

Many electrical appliances (pool pumps, televisions, stereos, DVD and CD players, electronic gaming systems, aquariums, blenders, floor polishers, hand tools, mixers, etc.) perform unique functions that are difficult to duplicate. This is their chief value.

Attention should be focused on those appliances that use more than a few percent of annual electricity use. General techniques for energy management include:

- Reduce use of equipment where feasible (e.g., turn off entertainment systems when not in use)
- Perform maintenance to improve efficiency (e.g., clean pool filters to reduce pumping power)
- Schedule use for off-peak hours (evenings)

The last point requires further comment and applies to large electric appliances such as washers, dryers, and dishwashers. Some utilities now offer "time-of-use" rates that include a premium charge for usage occurring "on-peak" (when the greatest demand for electricity takes place), and lower energy costs for "off-peak" electricity use. By scheduling energy-intensive activities for off-peak hours (clothes washing and drying in the evening for example) the user helps the utility reduce its peaking power requirement, thereby reducing generating costs. The utility then returns the favor by providing lower rates for off-peak use.

10.2.2.8 Computers and Office Equipment

Computers are fairly ubiquitous in both the residential and commercial sectors at this point. There are significant energy savings available from going to LCD displays in place of CRT displays, and in enabling the sleep or power saver routines available with the machines. The use of laptop machines in place of desktop machines also offers substantial savings.

A wide variety of equipment in addition to computers can be found in commercial buildings, depending on the size and function. Process equipment within buildings (e.g., clothes washers in laundries, printing presses, refrigerated display cases, etc.) will not be discussed due to the great diversity of these items.

Excluding process equipment, major energy-using equipment in commercial buildings generally includes HVAC systems, lighting, and "other" equipment. Because energy management options for HVAC and lighting have already been described, the discussion here will be directed at "other" equipment, including:

- Computers, local area networks and peripherals
- Facsimile machines, electronic mail
- Vending machines and water coolers
- Copy machines
- Elevators and escalators

The energy management opportunities with these types of equipment are more restricted. One obvious strategy is to insure that all equipment is turned off when not needed or not in use. Another is to size equipment with the right capacity to do the job, avoiding overcapacity, which increases both energy and demand charges.

In general, the most likely opportunity with elevators, escalators, and similar equipment is to shut them down during off hours or other times when they are not needed.

There is a host of miscellaneous equipment in buildings that contributes only a small percentage of total energy use, is used infrequently, and is an unlikely candidate for improved efficiency.

10.3 Closing Remarks

This chapter has discussed the management of electrical energy in buildings. Beginning with a discussion of energy use in buildings, next outlined was the major energy-using systems and equipment, along with a brief description of their particular features that influence energy use and waste. A systematic methodology for implementing an energy management program was then described. The procedure has been implemented in a wide variety of situations including individual homes, commercial buildings, institutions, multinational conglomerates, and cities, and has worked well in each case. Following the discussion of how to set up an energy management program, a series of techniques for saving electricity in each major end use was discussed. The emphasis has been on currently available, cost-effective technology. Undoubtedly, there are other techniques available, but this chapter has concentrated on those that are known to work in today's economy for the typical energy consumer.

The first edition of this book was published in 1980. Much of the data in the first edition dated from the 1975–1980 timeframe, when the initial response to the oil embargo of 1973 was gathering momentum and maturing. It is remarkable to return to those data and look at the progress that has been made. In 1975, total U.S. energy use was 71.2 quads (75.1 GJ). Most projections at that time predicted U.S. energy use in excess of 100 quads (106 GJ) by 1992; instead, it only reached 85 quads (90 GJ) by 1992. In fact, by 2004, total U.S. energy use was close, but still had not reached the 100 quad mark; the preliminary estimate was 99.7 quads (105 GJ). During this same period, U.S. gross domestic product (GDP) increased from $4.31 trillion (1975) to $7.34 trillion (1992) to $10.8 trillion (2004). In other words, energy usage increased by 40% while enabling GDP to increase 2.5-fold. Some of this is due to a decrease in domestic energy-intensive industry, but for the most part it represents a remarkable improvement in overall energy efficiency.

As noted earlier in this chapter, a significant growth of the residential sector has occurred in the intervening decades, but efficiency improvements have held down the growth rate of energy consumption. The improvement in energy efficiency in lighting, refrigerators, air conditioning, and other devices has been truly remarkable. Today, the local hardware or homebuilder's store has supplies of energy efficient devices that were beyond imagination in 1973.

This is a remarkable accomplishment, technically and politically, given the diversity of the residential/commercial market. Besides the huge economic savings this has meant to millions of homeowners, apartment dwellers, and businesses, think of the environmental benefits associated with avoiding the massive additional amounts of fuel, mining, and combustion which otherwise would have been necessary.

11

Heating, Ventilating, and Air Conditioning Control Systems

Jan F. Kreider
University of Colorado

David E. Claridge and
Charles H. Culp
Texas A&M University

11.1 Introduction

This chapter describes the essentials of control systems for heating, ventilating, and air conditioning (HVAC) of buildings designed for energy conserving operation. Of course, there are other renewable and energy conserving systems that require control. The principles described herein for buildings also apply with appropriate and obvious modification to these other systems. For further reference, the reader is referred to several standard references in the list at the end of this chapter.

HVAC system controls are the information link between varying energy demands on a building's primary and secondary systems and the (usually) approximately uniform demands for indoor environmental conditions. Without a properly functioning control system, the most expensive, most thoroughly designed HVAC system will be a failure. It simply will not control indoor conditions to provide comfort.

The HVAC designer must design a control system that

- Sustains a comfortable building interior environment
- Maintains acceptable indoor air quality

- Is as simple and inexpensive as possible and yet meets HVAC system operation criteria reliably for the system lifetime
- Results in efficient HVAC system operation under all conditions
- Commissions the building, equipment and control systems
- Documents the system operation so that the building staff can successfully operate and maintain the HVAC system

It is a considerable challenge for the HVAC system designer to design a control system that is energy efficient and reliable. Inadequate control system design, inadequate commissioning, and inadequate documentation and training for the building staff often create problems and poor operational control of HVAC systems. This chapter describes the basics of HVAC control and the operational needs for successfully maintained operation. The reader is encouraged to review the following references on the subject: ASHRAE (2002, 2003, 2004, 2005), Haines (1987), Honeywell (1988), Levine (1996), Sauer, Howell, and Coad (2001), Stein and Reynolds (2000), and Tao and Janis (2005).

To achieve proper control based on the control system design, the HVAC system must be designed correctly and then constructed, calibrated and commissioned according to the mechanical and electrical systems drawings. These must include properly sized primary and secondary systems. In addition, air stratification must be avoided, proper provision for control sensors is required, freeze protection is necessary in cold climates, and proper attention must be paid to minimizing energy consumption, subject to reliable operation and occupant comfort.

The principle and final controlled variable in buildings is zone temperature (and to a lesser extent humidity and/or air quality in some buildings). This chapter will therefore focus on methods to control temperature. Supporting the zone temperature control, numerous other control loops exist in buildings within the primary and secondary HVAC systems, including boiler and chiller control, pump and fan control, liquid and air flow control, humidity control, and auxiliary system control (for example, thermal energy storage control). This chapter discusses only automatic control of these subsystems. Honeywell (1988) defines an automatic control system as "a system that reacts to a change or imbalance in the variable it controls by adjusting other variables to restore the system to the desired balance."

Figure 11.1 defines a familiar control problem with feedback. The water level in the tank must be maintained under varying outflow conditions. The float operates a valve that admits water to the tank as the tank is drained. This simple system includes all the elements of a control system:

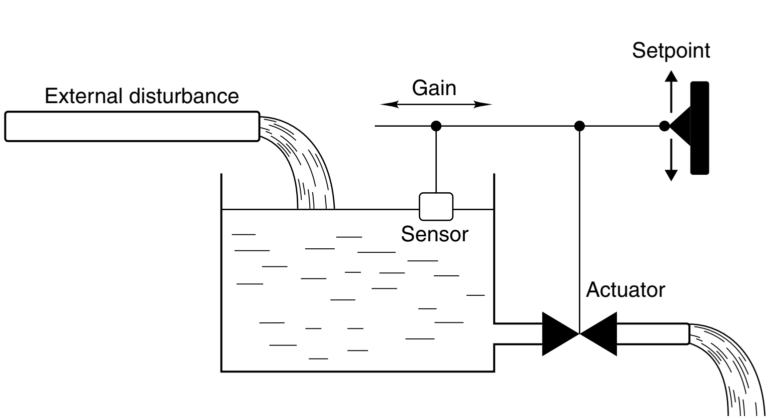


FIGURE 11.1 Simple water level controller. The setpoint is the full water level; the error is the difference between the full level and the actual level.

- Sensor—float; reads the controlled variable, the water level
- Controller—linkage connecting float to valve stem; senses difference between full tank level and operating level and determines needed position of valve stem
- Actuator (controlled device)—internal valve mechanism; sets valve (the final control element) flow in response to level difference sensed by controller
- Controlled system characteristic—water level; this is often termed the controlled variable

This system is called a *closed loop* or *feedback* system because the sensor (float) is directly affected by the action of the controlled device (valve). In an open loop system the sensor operating the controller does not directly sense the action of the controller or actuator. An example would be a method of controlling the valve based on an external parameter such as time of day, which may have an indirect relation to water consumption from the tank.

There are four common methods of control, of which Figure 11.1 shows but one. In the next section, each method will be described in relation to an HVAC system example.

11.2 Modes of Feedback Control

Feedback control systems adjust an output control signal based on feedback. The feedback is used to generate an error signal, which then drives a control element. Figure 11.1 illustrates a basic control system with feedback. Both off–on (two-position) control and analog (variable) control can be used. Numerous methodologies have been developed to implement analog control. These include proportional, proportional-integral (PI), proportional-integral-differential (PID), fuzzy logic, neural networks, and auto-regressive moving average (ARMA) control systems. Proportional and PI control systems are used for most HVAC control applications.

Figure 11.2a shows a steam coil used to heat air in a duct. The simple control system shown includes an air temperature sensor, a controller that compares the sensed temperature to the setpoint, a steam valve controlled by the controller, and the coil itself. This example system will be used as the point of reference when discussing the various control system types. Figure 11.2b is the control diagram corresponding to the physical system shown in Figure 11.2a.

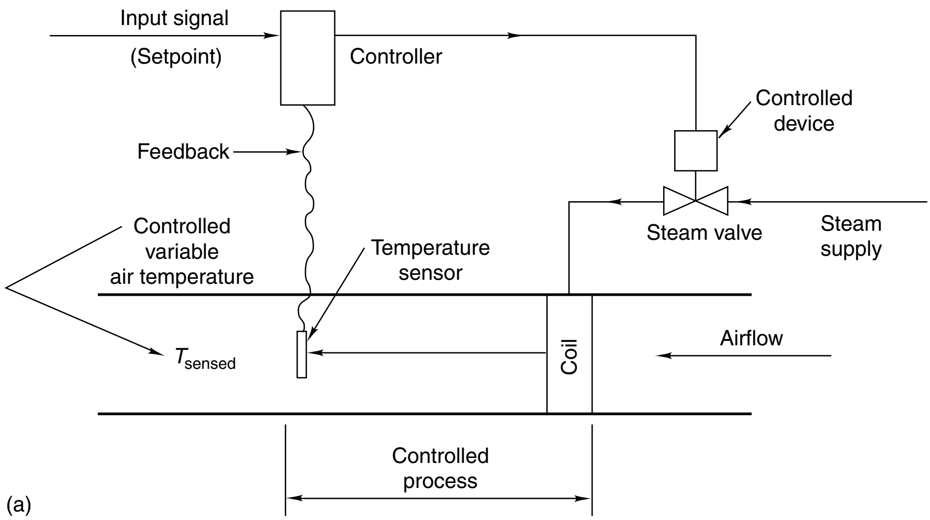
Two-position control applies to an actuator that is either fully open or fully closed. In Figure 11.2a, the valve is a two-position valve if two-position control is used. The position of the steam valve is determined by the value of the coil outlet temperature. Figure 11.3 depicts two-position control of the valve. If the air temperature drops below 95°F, the valve opens and remains open until the air temperature reaches 100°F. The differential is usually adjustable, as is the temperature setting itself. Two-position control is the least expensive method of automatic control and is suitable for control of HVAC systems with large time constants. Examples include residential space and water heating systems. Systems that are fast-reacting should not be controlled using this approach because overshoot and undershoot may be excessive.

Proportional control adjusts the controlled variable in proportion to the difference between the controlled variable and the setpoint. For example, a proportional controller would increase the coil heat rate in Figure 11.2 by 10% if the coil outlet air temperature dropped by an amount equal to 10% of the temperature range specified for the heating to go from off to fully on. Equation 11.1 defines the behavior of a proportional control loop:

$$T = T_{set} + K_p e \tag{11.1}$$

where T is the controller output, T_{set} is the set temperature corresponding to a constant value of controller output when no error exists, K_p is the loop gain that determines the rate or proportion at which the control signal changes in response to the error, and e is the error. In the case of the steam coil, the error is the difference between the air temperature setpoint and the sensed supply air temperature:

$$e = T_{set} - T_{sa} \tag{11.2}$$



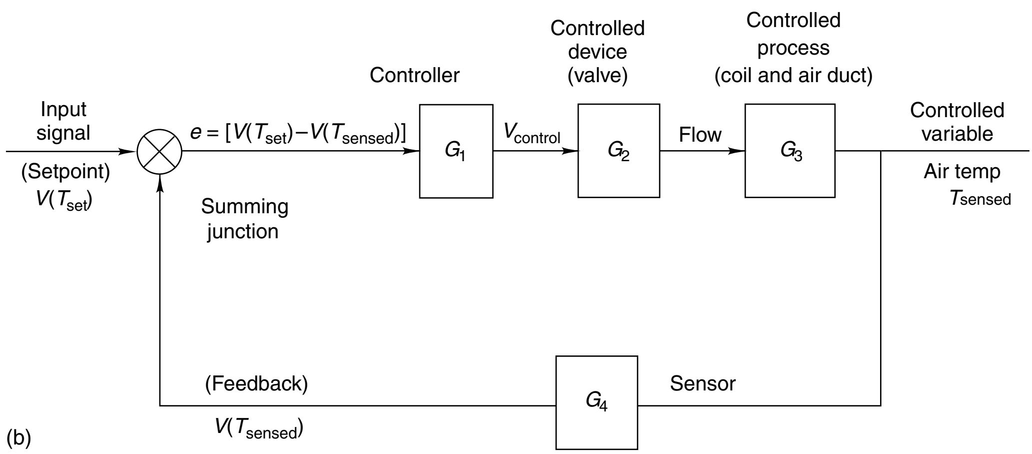


FIGURE 11.2 (a) Simple heating coil control system showing the process (coil and short duct length), controller, controlled device (valve and its actuator) and sensor. The setpoint entered externally is the desired coil outlet temperature. (b) Equivalent control diagram for heating coil. The G's represent functions relating the input to the output of each module. Voltages, V, represent both temperatures (setpoint and coil outlet) and the controller output to the valve in electronic control systems.

As coil air-outlet temperature drops farther below the set temperature, error increases, leading to increased control action—an increased steam flow rate. Note that the temperatures in Equation 11.1 and Equation 11.2 are often replaced by voltages or other variables, particularly in electronic controllers.

The throttling range (ΔT_{max}) is the total change in the controlled variable that is required to cause the actuator or controlled device to move between its limits. For example, if the nominal temperature of a zone is 72°F and the heating controller throttling range is 6°F, then the heating control undergoes its full travel between a zone temperature of 69°F and 75°F. This control, whose characteristic is shown in Figure 11.4, is reverse acting; i.e., as temperature (controlled variable) increases, the heating valve position decreases.

The throttling range is inversely proportional to the gain as shown in Figure 11.4. Beyond the throttling range, the system is out of control. In actual hardware, one can set the setpoint and either the

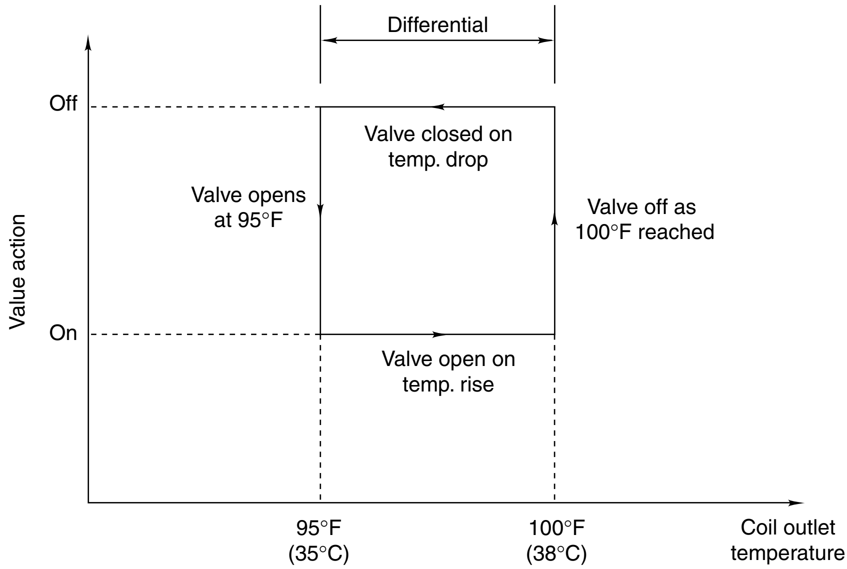


FIGURE 11.3 Two position (on-off) control characteristic.

gain or the throttling range (most common), but not both of the latter. Proportional control by itself is not capable of reducing the error to zero because an error is needed to produce the capacity required for meeting a load, as will be discussed in the following example. This unavoidable value of the error in proportional systems is called the *offset*. It is easy to see from Figure 11.4 that the offset is larger for systems with smaller gains. There is a limit to which one can increase the gain to reduce offset, because high gains can produce control instability.

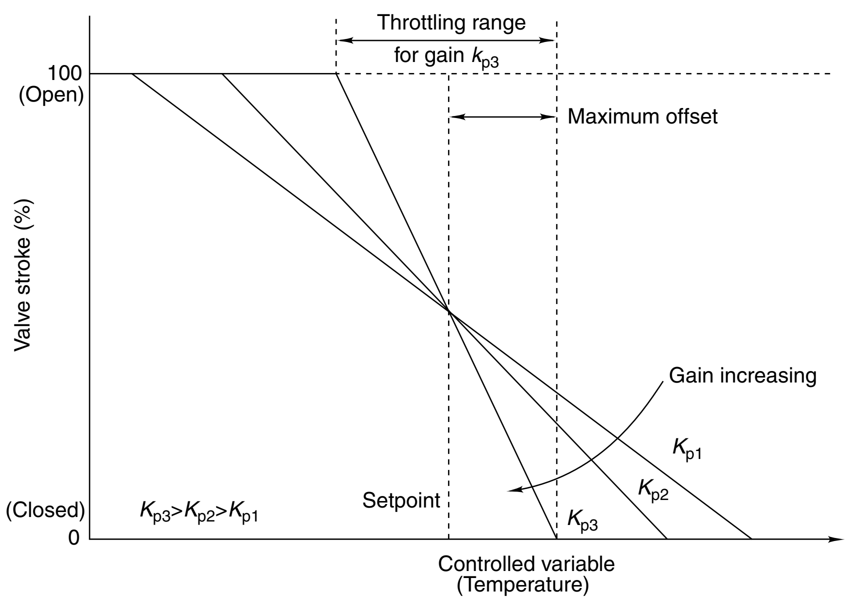


FIGURE 11.4 Proportional control characteristic showing various throttling ranges and the corresponding proportional gains, K_p. This characteristic is typical of a heating coil temperature controller.

Example 11.1 Proportional Gain Calculation

Problem: If the steam heating coil in Figure 11.2a has a heat output that varies from 0 to 20 kW as the outlet air temperature varies from 35 to 45°C in an industrial process, what is the coil gain and what is the throttling range? Find an equation relating the heat rate at any sensed air temperature to the maximum rate in terms of the gain and setpoint.

Solution: Given that $\dot{Q}_{max} = 20$ kW, $\dot{Q}_{min} = 0$ kW, $T_{max} = 45°C$, and $T_{min} = 35°C$ for the system in Figure 11.2a, and assuming steady-state operation, the problem is to determine K_p and ΔT_{max}.

The throttling range is the range of the controlled variable (air temperature) over which the controlled system (heating coil) exhibits its full capacity range. The temperature varies from 35 to 45°C; therefore the throttling range is

$$\Delta T_{max} = 45°C - 35°C = 10°C \tag{11.3}$$

The proportional gain is the ratio of the controlled system (coil) output to the throttling range. For this example, the controller output is $\dot{Q}$ and the gain is

$$K_p = \frac{\dot{Q}_{max} - \dot{Q}_{min}}{\Delta T_{max}} = \frac{(20-0)\ \text{kW}}{10\ \text{K}} = 2.0(\text{kW/K}). \tag{11.4}$$

The controller characteristic can be found by inspecting Figure 11.4. It is assumed that the average air temperature (40°C) occurs at the average heat rate (10 kW). The equation of the straight line shown is

$$\dot{Q} = K_p(T_{set} - T_{sensed}) + \frac{\dot{Q}_{max}}{2} = K_p e + \frac{\dot{Q}_{max}}{2}. \tag{11.5}$$

Note that the quantity $(T_{set} - T_{sensed})$ is the error, e, and a nonzero value indicates that the set temperature is not met. However, the proportional control system used here requires presence of an error signal to fully open or fully close the valve.

Inserting the numerical values:

$$\dot{Q} = 2.0\frac{\text{kW}}{\text{K}}(40 - T_{sensed}) + 10\ \text{kW}. \tag{11.6}$$

Comments: In an actual steam-coil control system, it is the steam valve that is controlled directly to indirectly control the heat rate of the coil. This is typical of many HVAC system controls, in that the desired control action is achieved indirectly by controlling another variable that in turn accomplishes the desired result. This is why the controller and controlled device are often shown separately as in Figure 11.2b.

This example illustrates with a simple system how proportional control uses an error signal to generate an offset, and how that offset controls an output quantity. Using a bias value, the error can be set to be zero at one value in the control range. Proportional control requires a nonzero error over the remainder of the control range.

Real systems also have a time response called a *loop time constant*. This limits proportional control applications to slow response systems, where the throttling range can be set so that the system achieves stability. Typically, slow-responding mechanical systems include pneumatic thermostats for zone control and air handler unit damper control. Fast-acting systems like duct pressure control must be artificially slowed down to be appropriate for proportional control.

Integral control is often added to proportional control to eliminate the offset inherent in proportional-only control. The result—proportional plus integral control—is identified by the acronym PI. Initially, the corrective action produced by a PI controller is the same as for a proportional-only controller. After the initial period, a further adjustment due to the integral term reduces the offset to zero. The rate at which this occurs depends on the time scale of the integration. In equation form, the PI controller is modeled by

$$V = V_0 + K_p e + K_i \int e \mathrm{d}t, \tag{11.7}$$

in which K_i is the integral gain constant. It has units of reciprocal time and is the number of times that the integral term is calculated per unit time. This is also known as the *reset rate; reset control* is an older term used by some to identify integral control.

Today, most PI control implementations use electronic sensors, analog-to-digital converters (A/Ds), and digital logic to implement the PI control. Integral windup must be taken into account when using PI control. Integral windup occurs when the control first starts or comes out of a reset condition. The integral term increases by integrating the error signal, and can have a full output. When the control becomes enabled, a large offset error can occur, driving the output variable to an undesired state. Various methods exist to minimize or eliminate the windup problem.

The integral term in Equation 11.7 has the effect of adding a correction to the output signal V for as long as the error term exists. The continuous offset produced by the proportional-only controller can thereby be reduced to zero because of the integral term. For HVAC systems, the time scale (K_P/K_i) of the integral term is often in the range of 10+ s to 10+ min. PI control is used for fast-acting systems for which accurate control is needed. Examples include mixed air controls, duct static pressure controls, and coil controls. Because the offset is eventually eliminated with PI control, the throttling range can be set rather wide to insure stability under a wider range of conditions than good control would permit with proportional-only control. Therefore, PI control is also used on almost all electronic thermostats.

Derivative control is used to speed up the action of PI control. When derivative control is added to PI control, the result is called *PID control*. The derivative term added to Equation 11.7 generates a correction signal proportional to the time rate of change of error. This term has little effect on a steady proportional system with uniform offset (time derivative is zero) but initially, after a system disturbance, produces a larger correction more rapidly. Equation 11.8 includes the derivative term in the mathematical model of the PID controller

$$V = V_0 + K_p e + K_i \int e \mathrm{d}t + K_d \frac{\mathrm{d}e}{\mathrm{d}t}, \tag{11.8}$$

in which K_d is the derivative gain constant. The time scale (K_d/K_p) of the derivative term is typically in the range of 0.2–15 minutes. Because HVAC systems do not often require rapid control response, the use of PID control is less common than use of PI control. Because a derivative is involved, any noise in the error (i.e., sensor) signal must be avoided to maintain stable control. One effective application of PID control in buildings is in duct static pressure control, a fast-acting subsystem that otherwise has a tendency to be unstable.

Derivative control has limited application in HVAC systems because the PID control loops can easily become unstable. PID loops require correct tuning for each of the three gain constants (Ks) over the performance range that the control loop will need to operate. Another serious limitation centers on the fact that most facility operators lack training and skills in tuning PID control loops.

Figure 11.5 illustrates the loop response for three correctly configured systems when a step function change, or disturbance, occurs. Note that the PI loop achieves the same final control as the PID, only the PI error signal is larger. An improperly configured PID loop can oscillate from the high value to the low value continuously.

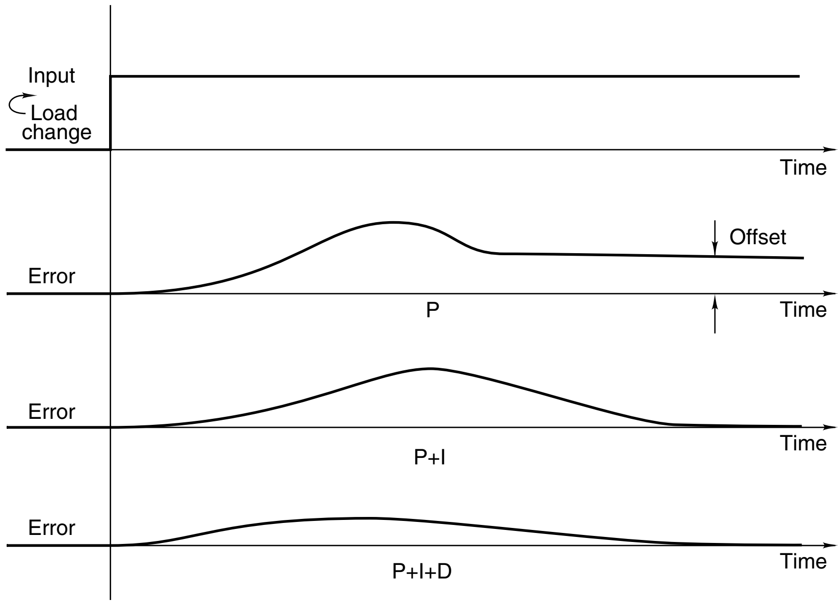


FIGURE 11.5 Performance comparison of P, PI, and PID controllers when subjected to a uniform, input step change.

11.3 Basic Control Hardware

In this section, the various physical components needed to perform the actions required by the control strategies of the previous section are described. Because there are two fundamentally different control approaches—pneumatic and electronic—the following material is so divided. Sensors, controllers, and actuators for principal HVAC applications are described.

11.3.1 Pneumatic Systems

The first widely adopted automatic control systems used compressed air as the signaling transmission medium. Compressed air had the advantages that it could be "metered" through various sensors, and it could power large actuators. The fact that the response of a normal pneumatic loop could take several minutes often worked as an advantage, as well. Pneumatic controls use compressed air (approximately 20 psig in the US) for operation of sensors and actuators. Though most new buildings use electronic controls, many existing buildings use pneumatic controls. This section provides an overview of how these devices operate.

Temperature control and damper control comprise the bulk of pneumatic loop controls. Figure 11.6 shows a method of sensing temperature and producing a control signal. Main supply air, supplied by a compressor, enters a branch line through a restriction. The zone thermostat bleeds out a variable amount of air, depending on the position of the flapper, controlled by the temperature sensor bellows. As more air bleeds out, the branch line pressure (control pressure) drops. This reduction in the total pressure to the control element changes the output of the control element. This control can be forward-acting or reverse-acting. The restrictions typically have hole diameters on the order of a few thousandths of an inch, and consume very little air. Typical pressures in the branch lines range between 3 and 13 psig (20–90 kPa). In simple systems this pressure from a thermostat could operate an actuator such as a control valve for a room heating unit. In this case, the thermostat is both the sensor and the controller—a rather common configuration.

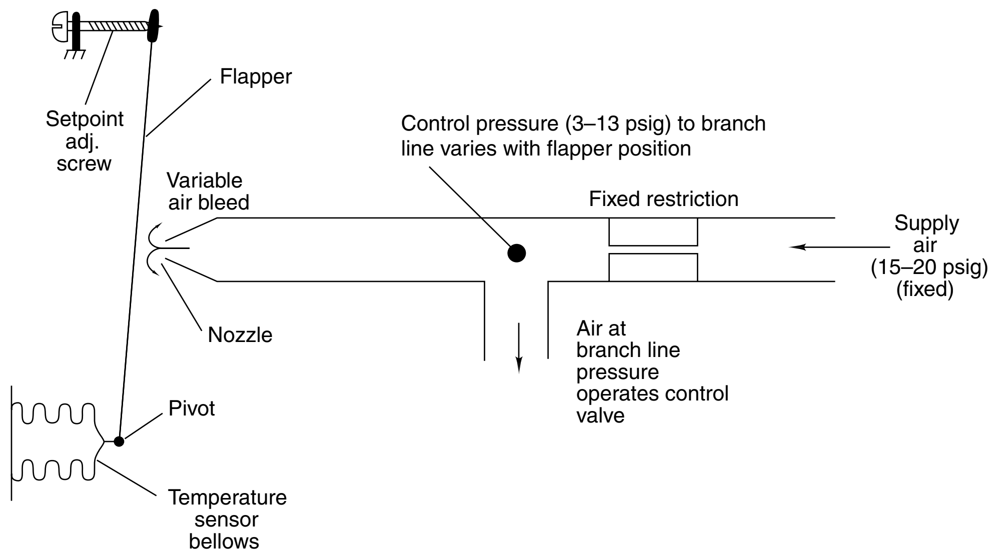


FIGURE 11.6 Drawing of pneumatic thermostat showing adjustment screw used to change temperature setting.

Many other temperature sensor approaches can be used. For example, the bellows shown in Figure 11.6 can be eliminated, and the flapper can be made of a bimetallic strip. As temperature changes, the bimetallic strip changes curvature, opening or closing the flapper/nozzle gap. Another approach uses a remote bulb filled with either liquid or vapor that pushes a rod (or a bellows) against the flapper to control the pressure signal. This device is useful if the sensing element must be located where direct measurement of temperature by a metal strip or bellows is not possible, such as in a water stream or high-velocity ductwork. The bulb and connecting capillary size may vary considerably by application.

Pressure sensors may use either bellows or diaphragms to control branch line pressure. For example, the motion of a diaphragm may replace that of the flapper in Figure 11.6 to control the bleed rate. A bellows similar to that shown in the same figure may be internally pressurized to produce a displacement that can control air bleed rate. A bellows produces significantly greater displacements than a single diaphragm.

Humidity sensors in pneumatic systems are made from materials that change size with moisture content. Nylon or other synthetic hygroscopic fibers that change size significantly (i.e., 1%–2%) with humidity are commonly used. Because the dimensional change is relatively small on an absolute basis, mechanical amplification of the displacement is used. The materials that exhibit the desired property include nylon, hair, and cotton fibers. Human hair exhibits a much more linear response with humidity than nylon; however, because the properties of hair vary with age, nylon has much wider use (Letherman 1981). Humidity sensors for electronic systems are quite different and are discussed in the next section.

An actuator converts pneumatic energy to motion—either linear or rotary. It creates a change in the controlled variable by operating control devices such as dampers or valves. Figure 11.7 shows a pneumatically operated control valve. The valve opening is controlled by the pressure in the diaphragm acting against the spring. The spring is essentially a linear device. Therefore, the motion of the valve stem is essentially linear with air pressure. However, this does not necessarily produce a linear effect on flow, as discussed later. Figure 11.8 shows a pneumatic damper actuator. Linear actuator motion is converted into rotary damper motion by the simple mechanism shown.

Pneumatic controllers produce a branch line (see Figure 11.6) pressure that is appropriate to produce the needed control action for reaching the setpoint. Such controls are manufactured by a number of control firms for specific purposes. Classifications of controllers include the sign of the output (direct- or reverse-acting) produced by an error, by the control action (proportional, PI, or two-position), or by number of inputs or outputs. Figure 11.9 shows the essential elements of a dual-input, single-output

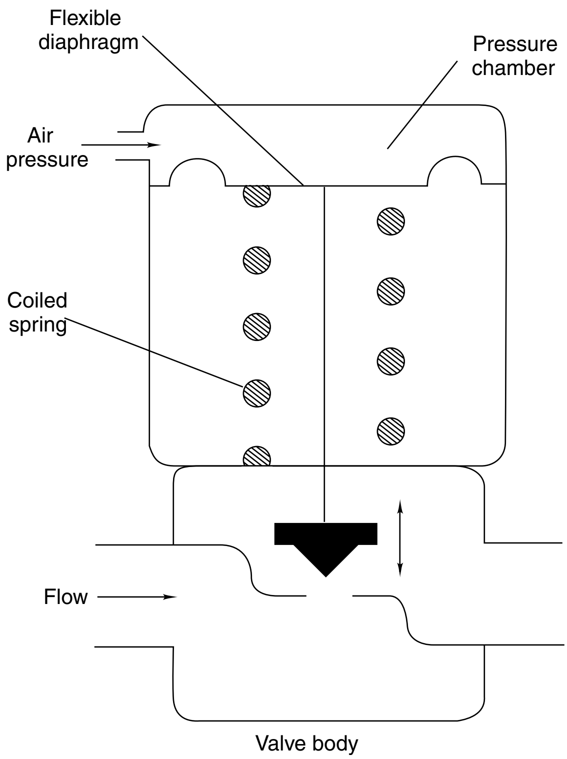


FIGURE 11.7 Pneumatic control valve showing counterforce spring and valve body. Increasing pressure closes the valve.

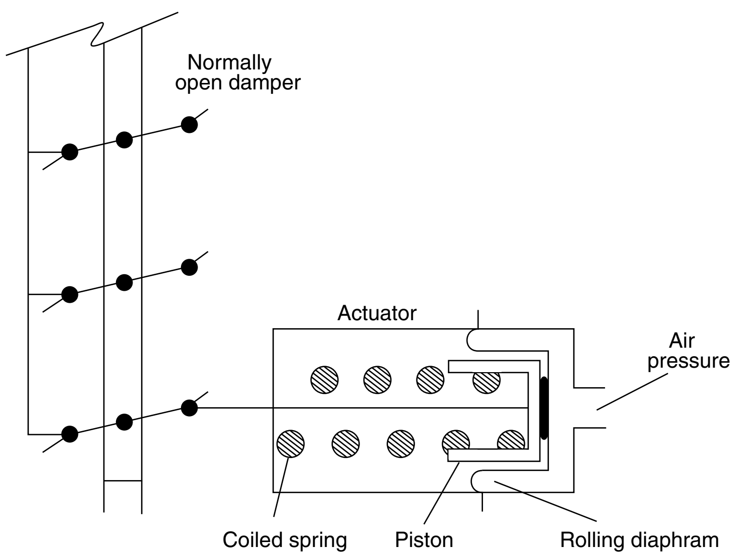


FIGURE 11.8 Pneumatic damper actuator. Increasing pressure closes the parallel blade damper.

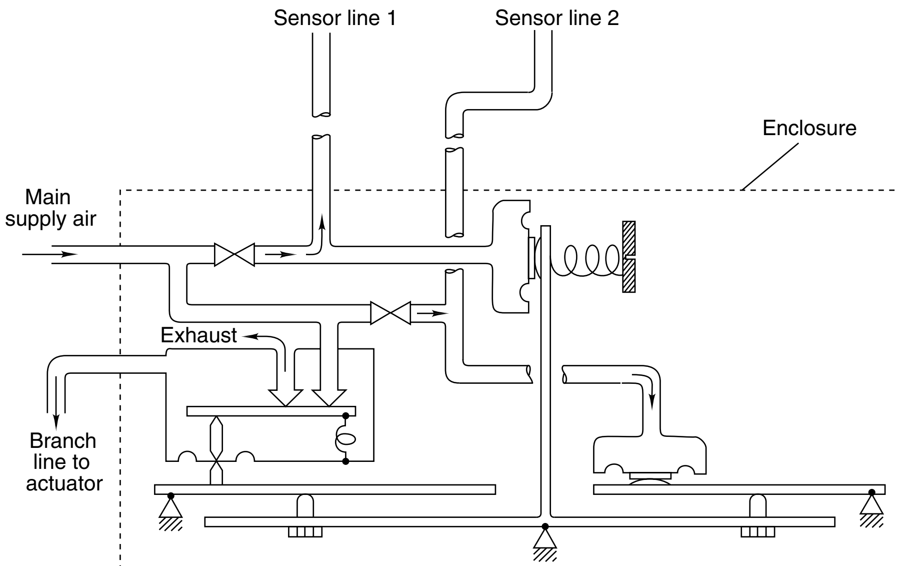


FIGURE 11.9 Example pneumatic controller with two inputs and one control signal output.

controller. The two inputs could be heating system supply temperature and outdoor temperature sensors, used to control the output water temperature setting of a boiler in a building heating system. This is essentially a boiler temperature reset system that reduces heating water temperature with increasing ambient temperature for better system control and reduced energy use.

The air supply for pneumatic systems must produce very clean, oil-free, dry air. A compressor producing 80–100 psig is typical. Compressed air is stored in a tank for use as needed, avoiding continuous operation of the compressor. The air system should be oversized by 50%–100% of estimated nominal consumption. The air is then dried to prevent moisture freezing in cold control lines in air handling units and elsewhere. Dried air should have a dew point of -30°F or less in severe heating climates. In deep cooling climates, the lowest temperature to which the compressed air lines are exposed may be the building cold air supply. Next, the air is filtered to remove water droplets, oil (from the compressor), and any dirt. Finally, the air pressure is reduced in a pressure regulator to the control system operating pressure of approximately 20 psig. Control air piping uses either copper or, in accessible locations, nylon.

11.3.2 Electronic Control Systems

Electronic controls comprise the bulk of the controllers for HVAC systems. Direct digital control systems (DDCs) began to make inroads in the early 1990s and now make up over 80% of all controller sales. Low-end microprocessors now cost under \$0.50 each and are thus very economical to apply. Along with the decreased cost, increased functionality can be obtained with DDC controls. BACnet has emerged as the standard communication protocol (ASHRAE 2001), and most control vendors offer a version of the BACnet protocol. In this section, the sensors, actuators and controllers used in modern electronic control systems for buildings are surveyed.

Direct digital control (DDC) enhances the previous analog-only electronic system with digital features. Modern DDC systems use analog sensors (converted to digital signals within a computer) along with digital computer programs to control HVAC systems. The output of this microprocessor-based system can be used to control electronic, electrical, or pneumatic actuators or a combination. DDC systems have the advantage of reliability and flexibility that others do not. For example, it is easier to accurately set control constants in computer Software than by making adjustments at a control panel with a

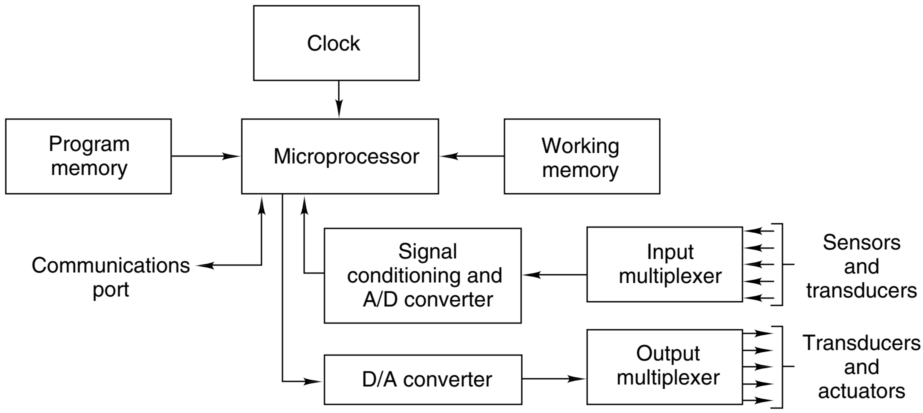


FIGURE 11.10 Block diagram of DDC controller.

screwdriver. DDC systems offer the option of operating energy management systems (EMS) and HVAC diagnostic, knowledge-based systems because the sensor data used for control is very similar to that used in EMSs. Pneumatic systems do not offer this ability. Figure 11.10 shows a schematic diagram of a DDC controller. The entire control system must include sensors and actuators not shown in this controller-only drawing.

Temperature measurements for DDC applications are made by three principal methods:

- Thermocouples
- Resistance temperature detectors (RTDs)
- Thermistors

Each has its advantages for particular applications. Thermocouples consist of two dissimilar metals chosen to produce a measurable voltage at the temperature of interest. The voltage output is low (in the millivolt range) but a well-established function of the junction temperature. Except for flame temperature measurements, thermocouples produce voltages too small to be useful in most HVAC applications (for example, a type-J thermocouple produces only 5.3 mV at 100°C).

RTDs use small, responsive sensing sections constructed from metals whose resistance-temperature characteristic is well established and reproducible. To first order,

$$R = R_0(1 + kT), \tag{11.9}$$

where R is the resistance (Ω), R_0 is the resistance at the reference temperature of 0°C (Ω), k is the temperature coefficient of resistance (°C^{-1}), and T is the RTD temperature (°C). This equation is easy to invert to find the temperature as a function of resistance. Although complex higher-order expressions exist, their use is not needed for HVAC applications.

Two common materials for RTDs are platinum and Balco (a 40%-nickel, 60%-iron alloy). The nominal values of k, respectively, are 3.85×10^{-3} and 4.1×10^{-3} °C^{-1}.

Modern electronics measure current and voltage and then determine the resistance using Ohm's law. The measurement causes power dissipation in the RTD element, raising the temperature and creating an error in the measurement. This Joule self-heating can be minimized by reducing the power dissipated in the RTD. Raising the resistance of the RTD helps reduce self-heating, but the most effective approach requires pulsing the current and making the measurement in a few milliseconds. Because one measurement per second will generally satisfy the most demanding HVAC control loop, the power dissipation can be reduced by a factor of 100 or more. Modern digital controls can easily handle the calculations necessary to implement piecewise linearization and other curve-fitting methods to improve the accuracy of the RTD measurements. In addition, lead wire resistance can cause lack of accuracy for the class of platinum RTDs whose nominal resistance is only 100 Ω, because the lead resistance of 1–2 Ω is not negligible by comparison to that of the sensor itself.

Thermistors are semiconductors that exhibit a standard exponential dependence for resistance versus temperature given by

$$R = A\mathrm{e}^{(\mathrm{B/T})}. \tag{11.10}$$

A is related to the nominal value of resistance at the reference temperature (77°F) and is on the order of several thousands of ohms. The exponential coefficient B (a weak function of temperature) is on the order of 5400–7200 R (3000–4000 K). The nonlinearity inherent in a thermistor can be reduced by connecting a properly selected fixed resistor in parallel with it. The resulting linearity is desirable from a control system design viewpoint. Thermistors can have a problem with long-term drift and aging; the designer and control manufacturer should consult on the most stable thermistor design for HVAC applications. Some manufacturers provide linearized thermistors that combine both positive and negative resistive dependence on temperature to yield a more linear response function.

Humidity measurements are needed for control of enthalpy economizers, or may also be needed to control special environments as such as clean rooms, hospitals and areas housing computers. Relative humidity, dew point, and humidity ratio are all indicators of the moisture content of air. An electrical, capacitance-based approach using a polymer with interdigitated electrodes has become the most common sensor type. The polymer material absorbs moisture and changes the dielectric constant of the material, changing the capacitance of the sensor. The capacitance of the sensor forms part of a resonant circuit so that when the capacitance changes, the resonant frequency changes. This frequency can then be correlated to the relative humidity and provide reproducible readings, if not saturated by excessive exposure to high humidity levels (Huang 1991). The response times of tens of seconds easily satisfy most HVAC application requirements. These humidity sensors need frequent calibration, generally yearly. If a sensor becomes saturated or has condensation on the surface, it becomes uncalibrated and exhibits an offset from its calibration curve. Older technologies used ionic salts on gold grids. These expensive sensors failed frequently.

Pressure measurements are made by electronic devices that depend on a change of resistance or capacitance with imposed pressure. Figure 11.11 shows a cross-sectional drawing of each. In the resistance type, stretching of the membrane lengthens the resistive element, thereby increasing resistance. This resistor is an element in a Wheatstone bridge; the resulting bridge voltage imbalance is linearly related to the imposed pressure. The capacitative-type unit has a capacitance between a fixed and a flexible metal that decreases with pressure. The capacitance change is amplified by a local amplifier that produces an output signal proportional to pressure. Pressure sensors can burst from overpressure or a water-hammer effect. Installation must carefully follow the manufacturer's requirements.

DDC systems require flow measurements to determine the energy flow for air and water delivery systems. Pitot tubes (or arrays of tubes) and other flow measurement devices can be used to measure either air or liquid flow in HVAC systems. Air flow measurements allow for proper flow in variable air volume (VAV) system control, building pressurization control and outside air control. Water flow measurements enable chiller and boiler control, and monitoring of various water loops used in the HVAC system. Some controls require only the knowledge of flow being present. Open–closed sensors fill this need and typically have a paddle that makes a switch connection in the presence of flow. These types of switches can also be used to detect "end of range," i.e., fully open or closed for dampers and other mechanical control elements.

Temperature, humidity and pressure transmitters are often used in HVAC systems. They amplify signals produced by the basic devices described in the preceding paragraphs and produce an electrical signal over a standard range, thereby permitting standardization of this aspect of DDC systems. The standard ranges are

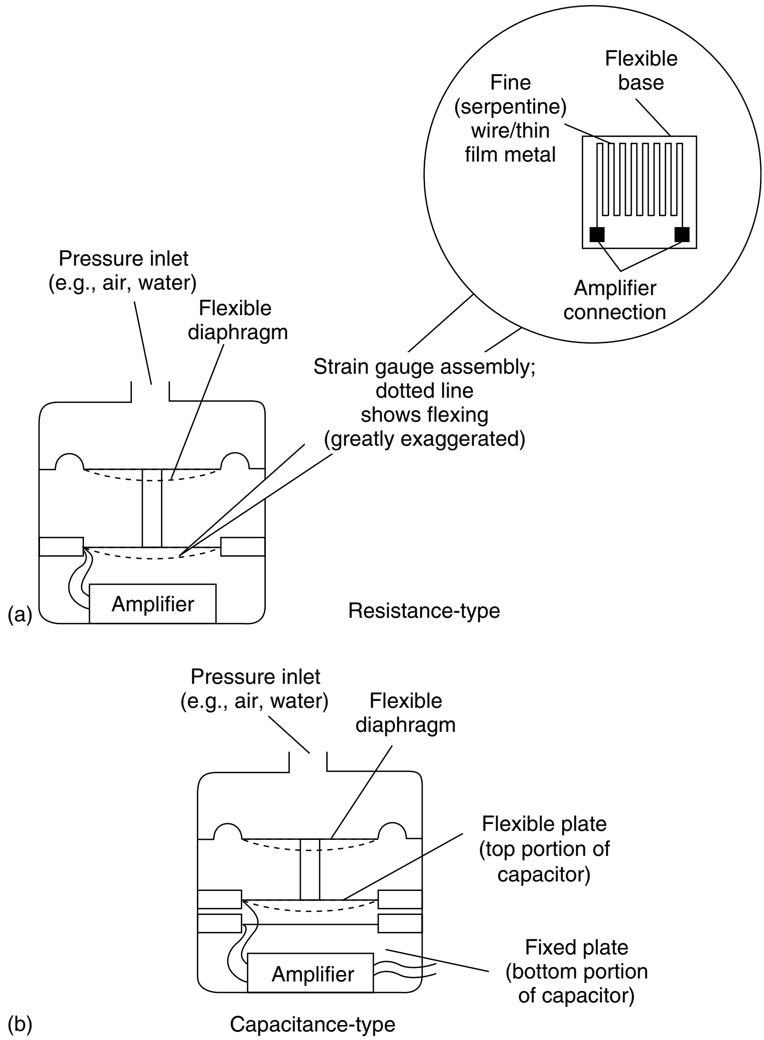


FIGURE 11.11 Resistance and capacitance type pressure sensors.

Current: 4–20 ma (DC)

Voltage: 0–10 volts (DC)

Although the majority of transmitters produce such signals, the noted values are not universally used.

Figure 11.10 shows the elements of a DDC controller. The heart of the controller is a microprocessor that can be programmed in either a standard or system-specific language. Control algorithms (linear or not), sensor calibrations, output signal shaping, and historical data archiving can be programmed as the user requires. A number of firms have constructed controllers on standard personal computer platforms. It is beyond the scope of this chapter to describe the details of programming HVAC controllers because each manufacturer uses a different approach. The essence of any DDC system, however, is the same as shown in the figure. Honeywell (1988) discusses DDC systems and their programming in more detail. Actuators for electronic control systems include

- Motors—operate valves, dampers
- Variable speed controls—pump, fan, chiller drives

- Relays and motor starters—operate other mechanical or electrical equipment (pumps, fans, chillers, compressors), electrical heating equipment
- Transducers—for example, convert electrical signal to pneumatic (EP transducer)
- Visual displays—not actuators in the usual sense, but used to inform system operator of control and HVAC system function

Pneumatic and DDC systems have their own advantages and disadvantages. Pneumatic systems possess increasing disadvantages of cost, hard to find replacements, requiring an air compressor with clean oil-free air, sensor drift, and imprecise control. The retained advantages include explosion-proof operation and a fail-soft degradation of performance. DDC systems have emerged and have taken the lead over pneumatic controls for HVAC systems because of the ability to integrate the control system into a large energy management and control system (EMCS), the accuracy of the control, and the ability to diagnose problems remotely. Systems based on either technology require maintenance and skilled operators.

11.4 Basic Control System Design Considerations

This section discusses selected topics in control system design, including control system zoning, valve and damper selection, and control logic diagrams. The following section shows several HVAC system control design concepts. Bauman (1998) may be consulted for additional information.

The ultimate purpose of an HVAC control system is to control zone temperature (and secondarily air motion and humidity) to conditions that assure maximum comfort and productivity of the occupants. From a controls viewpoint, the HVAC system is assumed to be able to provide comfort conditions if controlled properly. A zone is any portion of a building having loads that differ in magnitude and timing sufficiently from those of other areas, such that separate portions of the secondary HVAC system and control system are needed to maintain comfort.

Having specified the zones, the designer must select the location for the thermostat (and other sensors, if used). Thermostat signals are either passed to the central controller or used locally to control the amount and temperature of conditioned air or coil water introduced into a zone. The air is conditioned either locally (e.g., by a unit ventilator or baseboard heater) or centrally (e.g., by the heating and cooling coils in the central air handler). In either case, a flow control actuator is controlled by the thermostat signal. In addition, airflow itself may be controlled in response to zone information in VAV systems. Except for variable speed drives used in variable volume air or liquid systems, flow is controlled by valves or dampers. The design selection of valves and dampers is discussed next.

11.4.1 Steam and Liquid Flow Control

The flow through valves such as that shown in Figure 11.7 is controlled by valve stem position, which determines the flow area. The variable flow resistance offered by valves depends on their design. The flow characteristic may or may not be linear with position. Figure 11.12 shows flow characteristics of the three most common valve types. Note that the plotted characteristics apply only for constant valve pressure drop. The characteristics shown are idealizations of actual valves. Commercially available valves will resemble, but not necessarily exactly match, the curves shown.

The linear valve has a proportional relation between volumetric flow, $\dot{V}$, and valve stem position, z.

$$\dot{V} = kz. \tag{11.11}$$

The flow in equal percentage valves increases by the same fractional amount for each increment of opening. In other words, if the valve is opened from 20 to 30% of full travel, the flow will increase by the same percentage as if the travel had increased from 80 to 90% of its full travel. However, the absolute volumetric flow increase for the latter case is much greater than for the former. The equal percentage

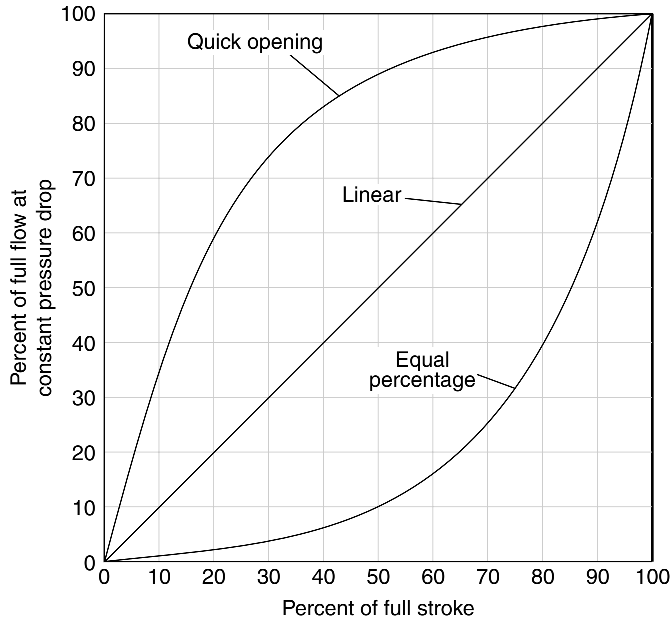


FIGURE 11.12 Quick opening, linear and equal percentage valve characteristics.

valve flow characteristic is given by

$$\dot{V} = Ke^{(kz)}, \tag{11.12}$$

in which k and K are proportionality constants for a specific valve. Quick-opening valves do not provide good flow control, but are used when rapid action is required with little stem movement for on/off control.

Example 11.2 Equal Percentage Valve

Problem: A valve at 30% of travel has a flow of 4 gal/min. If the valve opens another 10% and the flow increases by 50% to 6 gal/min, what are the constants in Equation 11.12? What will be the flow at 50% of full travel? (See Figure 11.12.)

Assumptions: Pressure drop across the valve remains constant.

Solution: The problem is to determine k, K, $\dot{V}_{50}$. Equation 11.12 can be evaluated at the two flow conditions. If the results are divided by each other, then

$$\frac{\dot{V}_2}{\dot{V}_1} = \frac{6}{4} = e^{k(z_2 - z_1)} = e^{k(0.4-0.3)}. \tag{11.13}$$

In this expression, the travel, z, is expressed as a fraction of the total travel and is dimensionless. Solving this equation for k gives the result

$$k = 4.05 \text{ (no units)}$$

From the known flow at 30% travel, the second constant, K, can be determined:

$$K = \frac{4 \text{ gal/min}}{e^{4.05 \times 0.3}} = 1.19 \text{ gal/min}. \tag{11.14}$$

Finally, the flow is given by

$$\dot{V} = 1.19e^{4.05z}. \tag{11.15}$$

At 50% travel, the flow can be found from

$$\dot{V}_{50} = 1.19e^{4.05\times0.5} = 9.0\ \text{gal/min}. \tag{11.16}$$

Comments: This result can be checked because the valve is an equal percentage valve. At 50% travel, the valve has moved 10% beyond its 40% setting, at which the flow was 6 gal/min. Another 10% stem movement will result in another 50% flow increase from 6 gal/min to 9 gal/min, confirming the solution.

The plotted characteristics of all three valve types assume constant pressure drop across the valve. In an actual system, the pressure drop across a valve will not remain constant, but if the valve is to maintain its control characteristics, the pressure drop across it must be the majority of the entire loop pressure drop. If the valve is designed to have a full open-pressure drop equal to that of the balance of the loop, good flow control will exist. This introduces the concept of valve authority, defined as valve pressure drop as a fraction of total system pressure drop:

$$A \equiv \frac{\Delta p_{\text{v,open}}}{(\Delta p_{\text{v,open}} + \Delta p_{\text{system}})}. \tag{11.17}$$

For proper control, the full-open valve authority should be at least 0.5. If the authority is 0.5 or more, control valves will have installed characteristics not much different from those shown in Figure 11.12. If not, the valve characteristic will be distorted upward because the majority of the system pressure drop will be dissipated across the valve.

Valves are further classified by the number of connections or ports. Figure 11.13 shows sections of typical two-way and three-way valves. Two-port valves control flow through coils or other HVAC equipment by varying valve flow resistance as a result of flow area changes. As shown, the flow must oppose the closing of the valve. If not, near closure the valve would slam shut or oscillate, both of which cause excessive wear and noise. The three-way valve shown in the figure is configured in the diverting mode. That is, one stream is split into two, depending on the valve opening. The three-way valve shown is double seated (single-seated three-way valves are also available); it is therefore easier to close than a single-seated valve, but tight shutoff is not possible.

Three-way valves can also be used as mixing valves. In this application two streams enter the valve, and one leaves. Because their internal design is different, mixing and diverting valves cannot be used interchangeably, to ensure that they can each seat properly. Particular attention is needed by the installer to be sure that connections are made properly; arrows cast in the valve body show the proper flow direction. Figure 11.14 shows an example of three-way valves for both mixing and diverting applications.

Valve flow capacity is denoted in the industry by the dimensional flow coefficient, C_v, defined by

$$\dot{V}(\text{gal/min}) = C_v[\Delta p(\text{psi})]^{0.5}. \tag{11.18}$$

Δp is the pressure drop across the fully open valve, so C_v is specified as the flow rate of 60°F water that will pass through the fully open valve if a pressure difference of 1.0 psi is imposed across the valve. If SI units (m^3/s and Pa) are used, the numerical value of C_v is 17% larger than in USCS units. After the designer has determined a value of C_v, manufacturer's tables can be consulted to select a valve for the known pipe size. If a fluid other than water is to be controlled, the C_v found from Equation 11.18 should be multiplied by the square root of the fluid's specific gravity.

Steam valves are sized using a similar dimensional expression

$$\dot{m}(\text{lb/h}) = 63.5C_v[\Delta p(\text{psi})/v(\text{ft.}^3/\text{lb})]^{0.5}, \tag{11.19}$$

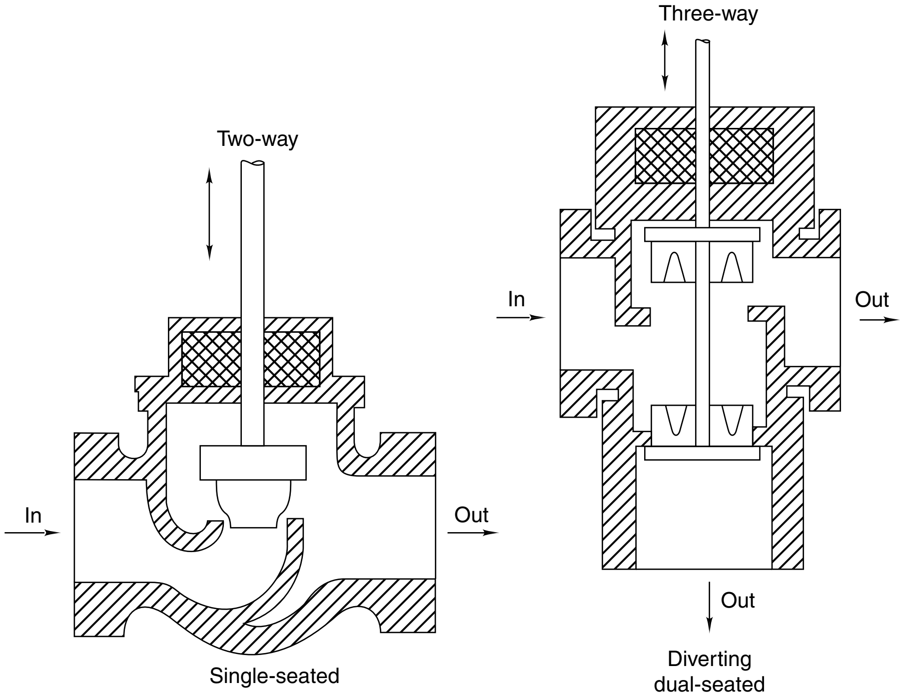


FIGURE 11.13 Cross-sectional drawings of direct-acting, single-seated, two-way valve and dual-seated, three-way, diverting valve.

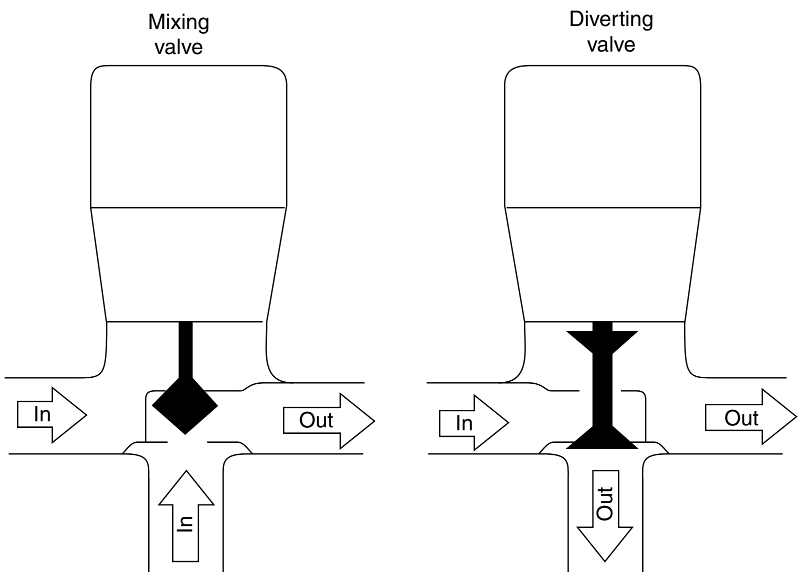


FIGURE 11.14 Three-way mixing and diverting valves. Note the significant difference in internal construction. Mixing valves are more commonly used.

in which v is the steam specific volume. If the steam is highly superheated, multiply C_v found from Equation 11.19 by 1.07 for every 100°F of superheat. For wet steam, multiply C_v by the square root of the steam quality. Honeywell (1988) recommends that the pressure drop across the valve to be used in the equation be 80% of the difference between steam supply and return pressures (subject to the sonic flow limitation discussed below). Table 11.1 can be used for preliminary selection of control valves for either steam or water.

The type of valve (linear or not) for a specific application must be selected so that the controlled system is as nearly linear as possible. Control valves are very commonly used to control the heat transfer rate in coils. For a linear system, the combined characteristic of the actuator, valve, and coil should be linear. This will require quite different valves for hot water and steam control, for example.

Figure 11.15 shows the part load performance of a hot water coil used for air heating; at 10% of full flow the heat rate is 50% of its peak value. The heat rate in a cross-flow heat exchanger increases roughly in exponential fashion with flow rate—a highly nonlinear characteristic. This heating coil nonlinearity follows from the longer water residence time in a coil at reduced flow, and the relatively large temperature difference between air being heated and the water heating it.

However, if one were to control the flow through this heating coil by an equal percentage valve (positive exponential increase of flow with valve position), the combined valve plus coil characteristic would be roughly linear. Referring to Figure 11.15, 50% of stem travel corresponds to 10% flow. The third graph in the figure is the combined characteristic. This near-linear subsystem is much easier to control than if a linear valve were used with the highly nonlinear coil. Hence the rule: use equal percentage valves for heating coil control.

Linear, two-port valves are to be used for steam flow control to coils because the transfer of heat by steam condensation is a linear, constant temperature process—the more steam supplied, the greater the heat rate, in exact proportion. Note that this is a completely different coil flow characteristic than for hot-water coils. However, steam is a compressible fluid and the sonic velocity sets the flow limit for a given valve opening when the pressure drop across the valve is more than 60% of the steam supply line absolute pressure. As a result, the pressure drop to be used in Equation 11.19 is the *smaller* of (1) 50% of the absolute stream pressure upstream of the valve or (2) 80% of the difference between the steam supply and return line pressures. The 80% rule gives good valve modulation in the subsonic flow regime (Honeywell 1988).

Chilled water control valves should also be linear because the performance of chilled water coils (smaller air–water temperature difference than in hot-water coils) is more similar to steam coils than to hot-water coils.

Either two- or three-way valves can be used to control flow at part load through heating and cooling coils as shown in Figure 11.16. The control valve can be controlled from either coil outlet water or air temperature. Two- or three-way valves achieve the same local result at the coil when used for part load control. However, the designer must consider effects on the balance of the secondary system when selecting the valve type.

In essence, the two-way valve flow control method results in variable flow (tracking variable loads) with constant coil water temperature change, whereas the three-way valve approach results in roughly constant secondary loop flow rate, but smaller coil water temperature change (beyond the local coil loop itself). In large systems, a primary/secondary design with two-way valves is preferred, unless the primary equipment can handle the range of flow variation that will result without a secondary loop. Because chillers and boilers require that flow remain within a restricted range, the energy and cost savings that could accrue due to the two-way valve, variable volume system are difficult to achieve in small systems unless a two-pump, primary/secondary loop approach is employed. If this dual-loop approach is not used, the three-way valve method is required to maintain required boiler or chiller flow.

The location of the three-way valve at a coil must also be considered by the designer. Figure 11.16b shows the valve used downstream of the coil in a mixing, bypass mode. If a balancing valve is installed in the bypass line, and set to have the same pressure drop as the coil, the local coil loop will have the same pressure drop for both full and zero coil flow. However, at the valve mid-flow position, overall flow resistance is less,

TABLE 11.1 Quick sizing Chart for Control Valves

Cv	Steam Capacity, lb/h						Water Capacity, gal/min						
	Vacuum Return Systems[a]			Atmospheric Return Systems			Differential Pressure, psig						
	2-psi Supply Press. 3.2-psi Press. Drop[b]	5-psi Supply Press. 5.6-psi Press. Drop[b]	10-psi Supply Press. 9.6-psi Press. Drop[b]	2-psi Supply Press. 1.6-psi Press. Drop[b]	5-psi Supply Press. 4.0-psi Press. Drop[b]	10-psi Supply Press. 8.0-psi Press. Drop[b]	2	4	6	8	10	15	20
0.33	7.7	11.0	16.0	5.4	9.3	14.6	0.41	0.66	0.81	0.93	1.04	1.27	1.47
0.63	14.6	20.9	30.5	10.4	17.7	27.8	0.89	1.26	1.54	1.78	1.99	2.4	2.81
0.73	17.0	24.3	35.4	12	20.5	32.2	1.0	1.46	1.78	2.06	2.3	2.8	3.25
1.0	23.0	33.2	48.5	16.4	28	44	1.4	2.0	2.44	2.82	3.16	3.9	4.46
1.6	37.09	53.1	77.6	26.8	45	70.6	2.25	3.2	3.9	4.51	5.06	6.2	7.13
2.5	58.25	82.9	121.2	41.9	70.25	110.25	3.53	5.0	6.1	7.05	7.9	9.68	11.15
3.0	69.9	99.5	145.5	50.2	84.3	132.3	4.23	6.0	7.32	8.46	9.48	11.61	13.38
4.0	93.2	132.2	194.0	67	112.4	177.4	5.6	8.0	9.76	11.28	12.6	15.5	17.87
5.0	116.2	165.2	242.5	82.7	140.5	220.5	7.1	10.0	12.2	14.1	15.8	19.4	22.3
6.0	139	200	291.0	99	168	265	8.5	12.0	14.6	16.92	18.9	23.2	27.0
6.3	146	209	311.5	104	177	278	8.9	12.6	15.4	17.78	19.9	24.4	28.1
7.0	162	233	339.5	115	196	309	9.9	14.0	17.1	19.74	22.1	27.1	31
8.0	186.5	264.4	388.0	131.2	224.8	352.8	11.3	16.0	19.5	22.56	25.3	31.6	35.7
10.0	232	332	485.0	164	281	441	14.1	20	24.4	28.2	31.6	38.7	44.6
11.0	256	366	533.5	181	309	486	15.5	22	27	31.02	34.4	42.5	49
13.0	303	434	630.5	213.7	365.3	373.3	18.3	27	31.7	36.7	41.1	50.3	58
14.0	326	465	679.0	232	393	617	19.7	28	34	39	44	54	62
15.0	349.3	497.6	727.5	246	421.5	661.5	21.1	30	36.6	42.3	47.4	58	66.9
16.0	370.9	531	776.0	268	450	706	22.5	32	39	45.1	50.6	62	71.3
18.0	419	597	873.0	301	505	794	25	36	44	51	57	70	80
20.0	466	664	970.0	335	562	882	28	40	49	56	63	77	89
23.0	541	763	1,115	385	646	1,014	32	46	56	65	73	89	103
25.0	582.5	829	1,212	419	702.5	1,102.5	35.3	50	61	70.5	79	96.8	111.5
27.0	628.2	896	1,309	452.5	758.7	1,190.7	38.1	54	65.9	76.1	85.3	104.5	120.4
30.0	699	995	1,455	502	843	1,323	42.3	60	73.2	84.6	94.8	116.1	133.8
38.0	885	1,257	1,833	636	1,069	1,676	53	76	93	107	120	147	169
40.0	932	1,322	1,940	670	1,124	1,764	56	80	97.6	112.8	126	155	178.7
50.0	1,162	1,652	2,425	827	1,405	2,205	71	100	122	141	158	194	223
56.0	1,305	1,851	2,716	938	1,574	2,469	79	112	137	158	177	217	250

63.0	1,460	2,090	3,056	1,043	1,770	2,778	89	126	154	178	199	244	281
75.0	1,748	2,481	3,637	1,230	2,107	3,307	106	150	183	212	237	290	335
80.0	1,865	2,644	3,880	1,312	2,248	3,528	113	160	195	225.6	253	316	357
90.0	2,096	2,980	4,365	1,476	2,529	3,969	127	180	220	254	284	348	401
97.0	2,229	3,204	4,703	1,590	2,725	4,277	137	196	231	274	307	375	432
100.0	2,330	3,319	4,850	1,640	2,816	4,410	141	200	244	282	316	387	446
105.0	2,442	3,481	5,092	1,722	2,950	4,630	148	210	256	296	332	406	468
130.0	3,030	4,340	6,305	2,137	3,653	5,733	183	270	317	367	411	503	580
150.0	3,493	4,976	7,275	2,460	4,215	6,615	211	300	366	423	474	280	699
160.0	3,709	5,310	7,760	2,680	4,500	7,060	225	320	390	451	560	620	713
170.0	3,960	5,642	8,245	2,788	4,777	7,497	240	340	415	479	537	658	758
190.0	4,450	6,310	9,215	3,116	5,339	8,379	268	360	464	536	600	735	847
244.0	5,670	7,930	11,834	4,001	6,856	10,760	344	488	595	688	771	944	1,088
250.0	5,825	8,290	12,125	4,190	7,025	11,025	353	500	610	705	790	968	1,115
270.0	6,282	8,960	13,095	4,525	7,587	11,907	381	540	659	761	853	1,045	1,204
300.0	6,990	9,950	14,550	5,025	8,430	13,230	423	600	732	846	948	1,161	1,338
350.0	8,160	11,590	16,975	5,860	9,835	15,435	494	700	854	987	1,106	1,355	1,561
360.0	8,380	11,910	17,460	6,030	10,116	15,876	508	720	878	1,015	1,137	1,393	1,606
430.0	10,010	14,225	20,855	7,200	12,083	18,963	606	860	1,049	1,213	1,359	1,664	1,918
480.0	11,180	15,860	23,280	8,045	13,408	21,168	677	960	1,171	1,353	1,517	1,858	2,141
640.0	14,910	21,180	31,040	10,496	17,984	28,224	902	1,280	1,561	1,805	2,022	2,477	2,854
760.0	17,70	25,120	36,860	12,464	21,356	33,516	1,071	1,520	1,854	2,143	2,401	2,941	3,390
1,000.0	23,300	33,190	48,500	16,400	28,160	44,100	1,410	2,000	2,440	2,820	3,160	3,870	4,460
1,200.0	27,150	39,790	58,200	19,680	33,720	52,920	1,692	2,400	2,928	3,384	2,792	4,644	5,352
1,440.0	33,290	47,160	69,840	23,616	40,464	63,504	2,030	2,880	3,514	4,061	4,550	5,573	6,422

[a] Assuming a 4-in. through 8-in. vacuum.
[b] Pressure drop across fully open valve taking 80% of the pressure difference between supply and return main pressures.
Source: From Honeywell, Inc., *Engineering Manual of Automatic Control*, Honey well, Inc., Minneapolis, MN, 1988.

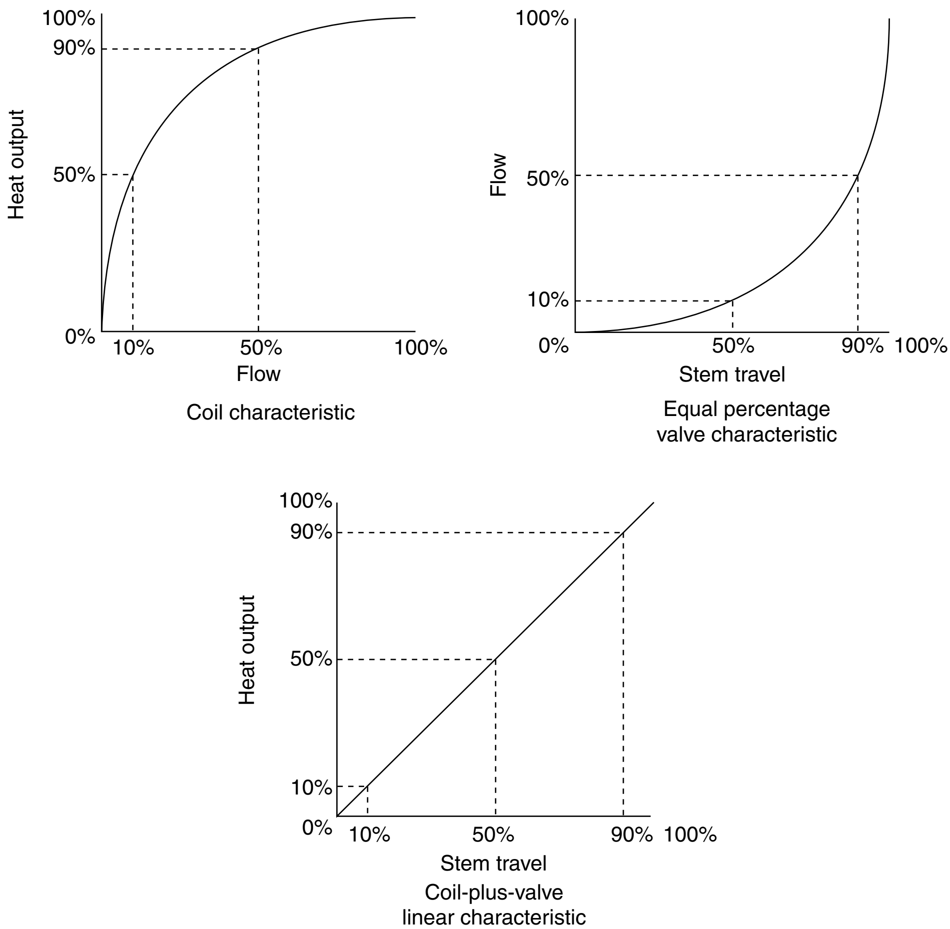


FIGURE 11.15 Heating coil, equal percentage valve, and combined coil + valve linear characteristic.

because two parallel paths are involved, and the total loop flow increases to 25% more than that at either extreme.

Alternatively, the three-way valve can also be used in a diverting mode, as shown in Figure 11.16c. In this arrangement, essentially the same considerations apply as for the mixing arrangement discussed earlier.[1] However, if a circulator (small pump) is inserted, as shown in Figure 11.16d, the direction of flow in the branch line changes and a mixing valve is used. The reason that pumped coils are used is that control is improved. With constant coil flow, the highly nonlinear coil characteristic shown in Figure 11.15 is reduced because the residence time of hot water in the coil is constant, independent of load. However, this arrangement appears to the external secondary loop the same as a two-way valve. As load is decreased, flow into the local coil loop also decreases. Therefore, the uniform secondary loop flow normally associated with three-way valves is not present unless the optional bypass is used.

[1]A little-known disadvantage of three-way valve control has to do with the conduction of heat from a closed valve to a coil. For example, the constant flow of hot water through two ports of a closed three-way heating coil control valve keeps the valve body hot. Conduction from the closed, hot valve mounted close to a coil can cause sufficient air heating to actually decrease the expected cooling rate of a downstream cooling coil during the cooling season. Three-way valves have a second practical problem; installers often connect three-way valves incorrectly, given the choice of three pipe connections and three pipes to be connected. Both of these problems are avoided by using two-way valves.

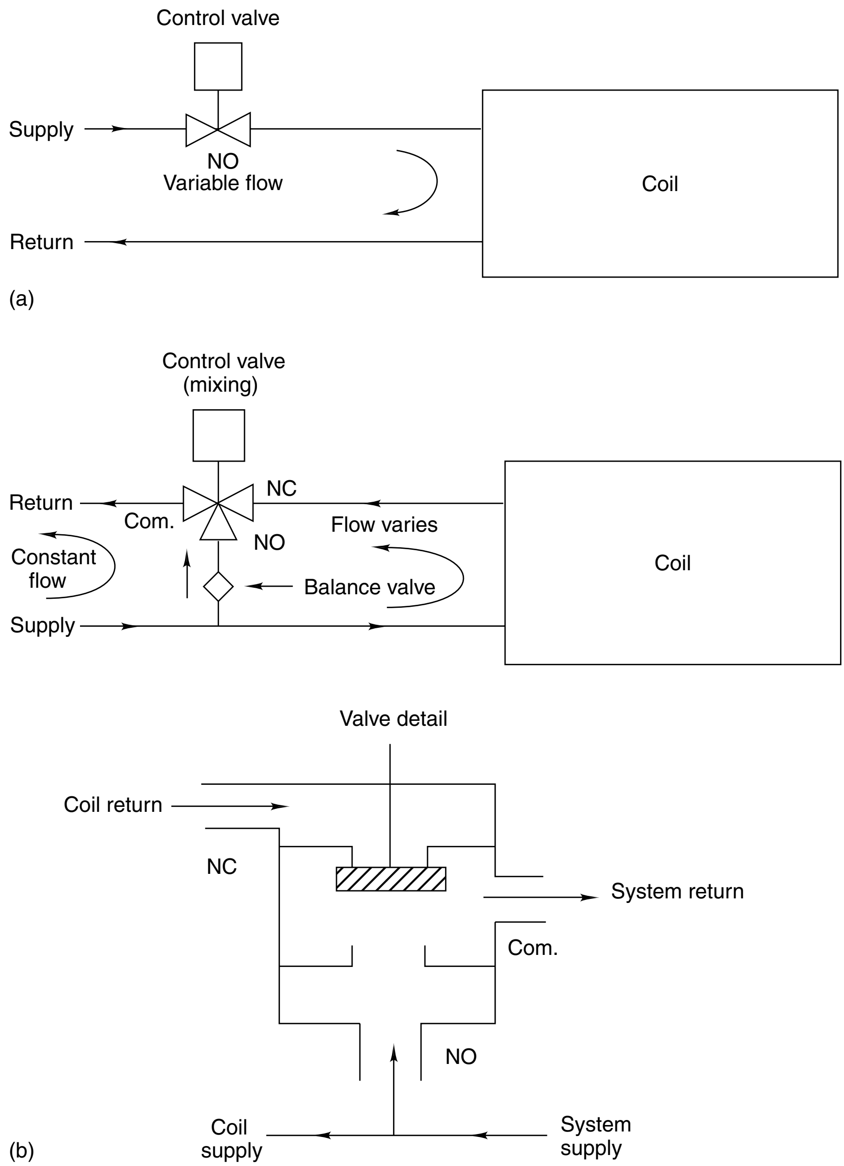


FIGURE 11.16 Various control valve piping arrangements (a) two-way valve; (b) three-way mixing valve; (c) three-way diverting valve; (d) pumped coil with three-way mixing valve.

For HVAC systems requiring precise control, high-quality control valves are required. The best controllers and valves are of "industrial quality." The additional cost for these valves compared to conventional building hardware results in more accurate control and longer lifetime.

11.4.2 Air Flow Control

Dampers are used to control airflow in secondary HVAC air systems in buildings. In this section, the characteristics of dampers used for flow control in systems where constant speed fans are involved are discussed. Figure 11.17 shows cross sections of the two common types of dampers used in commercial buildings. Parallel blade dampers use blades that all rotate in the same direction. They are most often

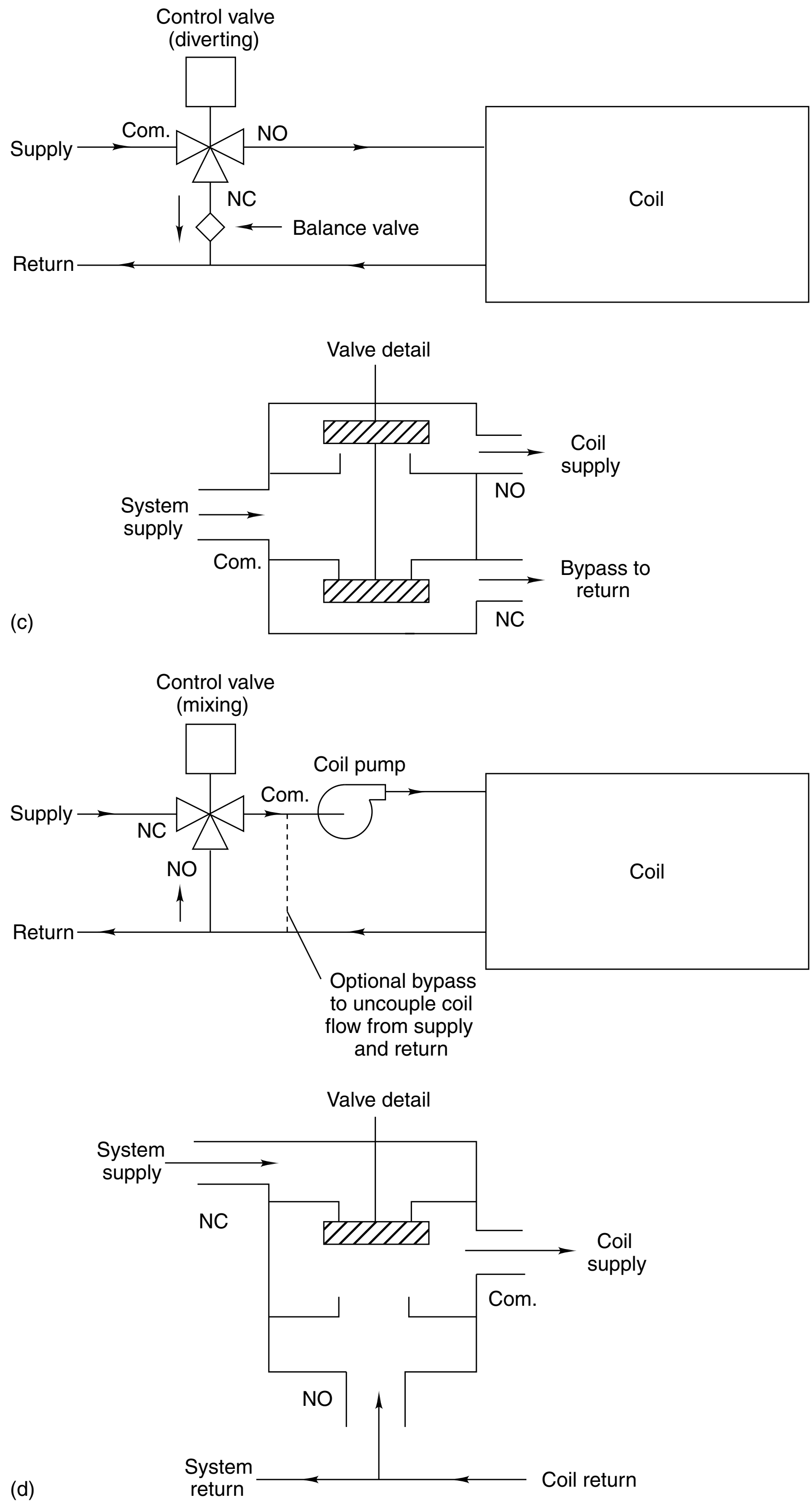


FIGURE 11.16 (*continued*)

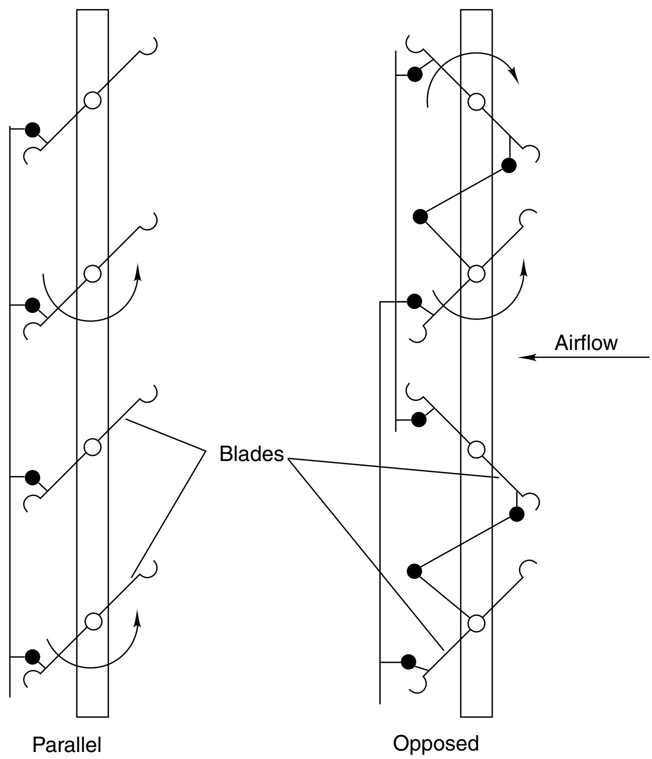


FIGURE 11.17 Diagram of parallel and opposed blade dampers.

applied to two position locations—open or closed. Use for flow control is not recommended. The blade rotation changes airflow direction, a characteristic that can be useful when airstreams at different temperatures are to be effectively blended.

Opposed blade dampers have adjacent counter-rotating blades. Airflow direction is not changed with this design, but pressure drops are higher than for parallel blading. Opposed blade dampers are preferred for flow control. Figure 11.18 shows the flow characteristics of these dampers to be closer to the desired linear behavior. The parameter α on the curves is the ratio of system pressure drop to fully open damper pressure drop.

A common application of dampers controlling the flow of outside air uses two sets in a face and bypass configuration as shown in Figure 11.19. For full heating, all air is passed through the coil and the bypass dampers are closed. If no heating is needed in mild weather, the coil is bypassed (for minimum flow resistance and fan power cost, flow through fully open face and bypass dampers can be used if the preheat coil water flow is shut off). Between these extremes, flow is split between the two paths. The face and bypass dampers are sized so that the pressure drop in full bypass mode (damper pressure drop only) and full heating mode (coil plus damper pressure drop) is the same.

11.5 Example HVAC Control Systems

Several widely used control configurations for specific tasks are described in this section. These have been selected from the hundreds of control system configurations that have been used for buildings. The goal of this section is to illustrate how control components described above are assembled into systems, and what design considerations are involved. For a complete overview of HVAC control system configurations see Honeywell (1988), Grimm and Rosaler (1990), Sauer, Howell, and Coad (2001), ASHRAE

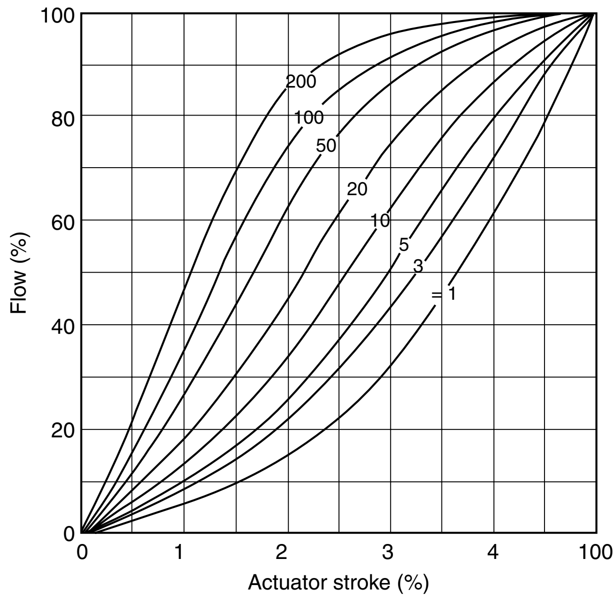


FIGURE 11.18 Flow characteristics of opposed blade dampers. The parameter α is the ratio of system resistance (not including the damper) to damper resistance. An approximately linear damper characteristic is achieved if this ratio is about 10 for opposed blade dampers.

(2002, 2003, 2004), and Tao and Janis (2005). The illustrative systems in this section are drawn in part from the first of these references.

In this section, seven control systems in common use will be discussed. Each system will be described using a schematic diagram, and its operation and key features will be discussed in the accompanying text.

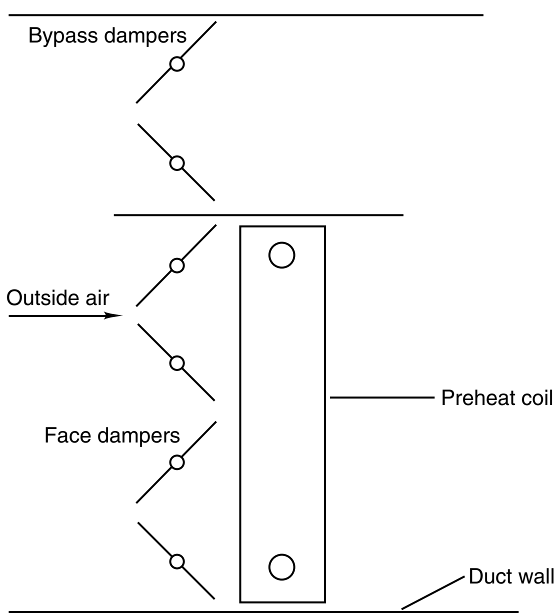


FIGURE 11.19 Face and bypass dampers used for preheating coil control.

11.5.1 Outside Air Control

Figure 11.20 shows a system for controlling outside and exhaust air from a central air handling unit equipped for economizer cooling when available. In this and the following diagrams, the following symbols are used:

C—cooling coil
DA—discharge air (supply air from fan)
DX—direct-expansion coil
E—damper controller
EA—exhaust air
H—heating coil
LT—low-temperature limit sensor or switch, must sense the lowest temperature in the air volume being controlled
M—motor or actuator (for damper or valve), variable speed drive
MA—mixed air
NC—normally closed
NO—normally open
OA—outside air
PI—proportional plus integral controller
R—relay
RA—return air
S—switch
SP—static pressure sensor used in VAV systems

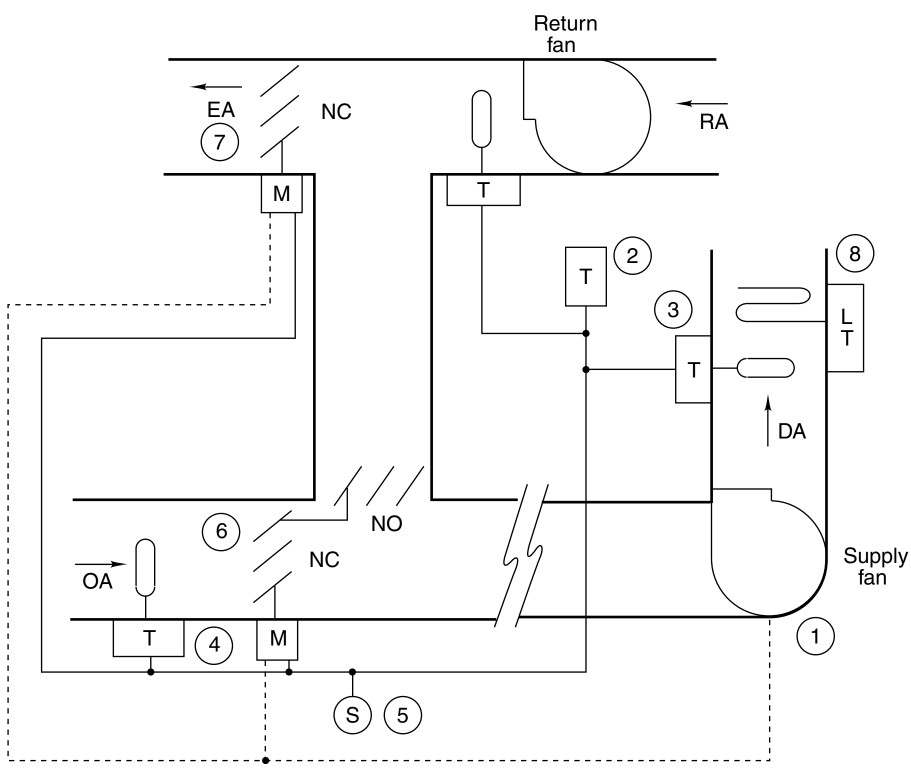


FIGURE 11.20 Outside-air-control system with economizer capability.

T—temperature sensor; must be located to read the average temperature representative of the air volume being controlled

This system is able to provide the minimum outside air during occupied periods; to use outdoor air for cooling when appropriate, by means of a temperature-based economizer cycle; and to operate fans and dampers under all conditions. The numbering system used in the figure indicates the sequence of events as the air handling system begins operation after an off period:

1. The fan control system turns on when the fan is turned on. This may be by a clock signal or a low- or high-temperature space condition.
2. The space temperature signal determines whether the space is above or below the setpoint. If above, the economizer feature will be activated if the OA temperature is below the upper limit for economizer operation, and will control the outdoor and mixed air dampers. If below, the outside air damper is set to its minimum position.
3. The discharge air PI controller controls both sets of dampers (OA/RA and EA) to provide the desired mixed air temperature.
4. When the outdoor temperature rises above the upper limit for economizer operation, the outdoor air damper is returned to its minimum setting.
5. Switch S is used to set the minimum setting on outside and exhaust air dampers manually. This is ordinarily done only once, during building commissioning and flow testing.
6. When the supply fan is off, the outdoor air damper returns to its NC position and the return air damper returns to its NO position.
7. When the supply fan is off, the exhaust damper also returns to its NC position.
8. Low temperature sensed in the duct will initiate a freeze-protect cycle. This may be as simple as turning on the supply fan to circulate warmer room air. Of course, the OA and EA dampers remain tightly closed during this operation.

11.5.2 Heating Control

If the minimum air setting is large in the preceding system, the amount of outdoor air admitted in cold climates may require preheating. Figure 11.21 shows a preheating system using face and bypass dampers. (A similar arrangement is used for direct-expansion [DX] cooling coils.) The equipment shown is installed upstream of the fan in Figure 11.20. This system operates as follows:

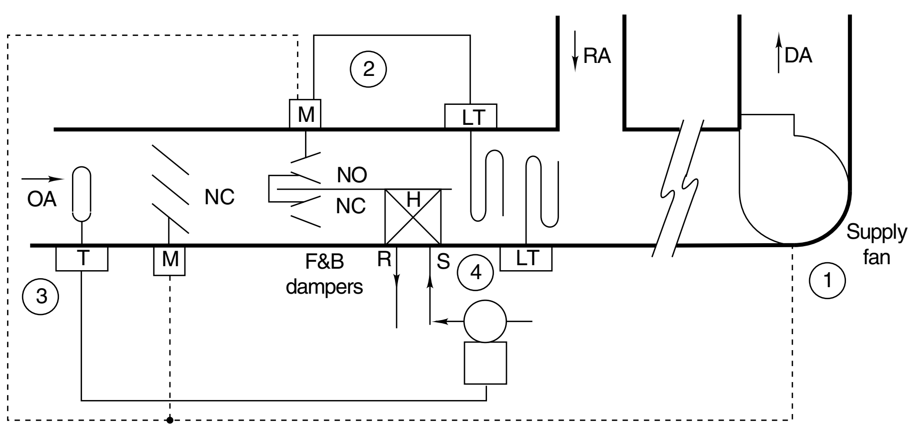


FIGURE 11.21 Preheat control system. Counter flow of air and hot water in the preheat coil results in the highest heat transfer rate.

1. The preheat subsystem control is activated when the supply fan is turned on.
2. The preheat PI controller senses temperature leaving the preheat section. It operates the face and bypass dampers to control the exit air temperature between 45 and 50°F.
3. The outdoor air sensor and associated controller controls the water valve at the preheat coil. The valve may be either a modulating valve (better control) or an on–off valve (less costly).
4. The low-temperature sensors (LTs) activate coil freeze protection measures, including closing dampers and turning off the supply fan.

Note that the preheat coil (as well as all other coils in this section) is connected so that the hot water (or steam) flows counter to the direction of airflow. Counter flow provides a higher heating rate for a given coil than does parallel flow. Mixing of heated and cold bypass air must occur upstream of the control sensors. Stratification can be reduced by using sheet metal air blenders or by propeller fans in the ducting. The preheat coil should be located in the bottom of the duct. Steam preheat coils must have adequately sized traps and vacuum breakers to avoid condensate buildup that could lead to coil freezing at light loads.

The face and bypass damper approach enables air to be heated to the required system supply temperature without endangering the heating coil. (If a coil were to be as large as the duct—no bypass area—it could freeze when the hot water control valve cycles open and closed to maintain discharge temperature.) The designer should consider pumping the preheat coil as shown in Figure 11.19d to maintain water velocity above the 3 ft./s needed to avoid freezing. If glycol is used in the system, the pump is not necessary, but heat transfer will be reduced.

During winter in heating climates, heat must be added to the mixed air stream to heat the outside air portion of mixed air to an acceptable discharge temperature. Figure 11.22 shows a common heating subsystem controller used with central air handlers. (It is assumed that the mixed air temperature is kept above freezing by action of the preheat coil, if needed.) This system has the added feature that coil discharge temperature is adjusted for ambient temperature because the amount of heat needed decreases with increasing outside temperature. This feature, called *coil discharge reset*, provides better control and can reduce energy consumption. The system operates as follows:

1. During operation the discharge air sensor and PI controller controls the hot water valve.
2. The outside air sensor and controller resets the setpoint of the discharge air PI controller up as ambient temperature drops.

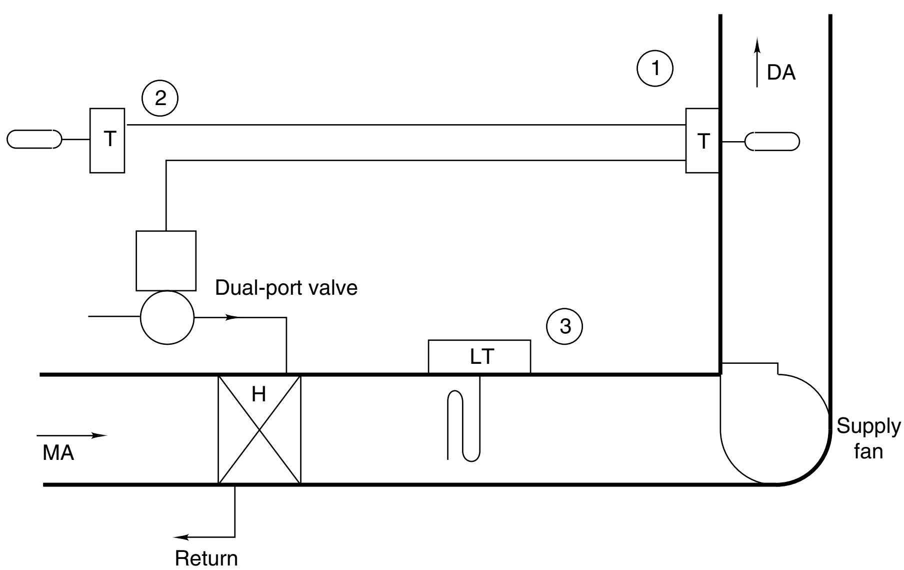


FIGURE 11.22 Heating-coil control subsystem using two-way valve and optional reset sensor.

3. Under sensed low-temperature conditions, freeze-protection measures are initiated as discussed earlier.

Reheating at zones in VAV or other systems uses a system similar to that just discussed. However, boiler water temperature is reset and no freeze protection is normally included. The air temperature sensor is the zone thermostat for VAV reheat, not a duct temperature sensor.

11.5.3 Cooling Control

Figure 11.23 shows the components in a cooling coil control system for a single-zone system. Control is similar to that for the heating coil discussed above, except that the zone thermostat (not a duct temperature sensor) controls the coil. If the system were a central system serving several zones, a duct sensor would be used. Chilled water supplied to the coil partially bypasses and partially flows through the coil, depending on the coil load. The use of three- and two-way valves for coil control has been discussed in detail previously. The valve NC connection is used as shown so that valve failure will not block secondary loop flow.

Figure 11.24 shows another common cooling coil control system. In this case the coil is a direct-expansion (DX) refrigerant coil and the controlled medium is refrigerant flow. DX coils are used when precise temperature control is not required because the coil outlet temperature drop is large whenever refrigerant is released into the coil because refrigerant flow is not modulated; it is most commonly either on or off. The control system sequences as follows:

1. The coil control system is energized when the supply fan is turned on.
2. The zone thermostat opens the two-position refrigerant valve for temperatures above the setpoint and closes it in the opposite condition.
3. At the same time, the compressor is energized or de-energized. The compressor has its own internal controls for oil control and pumpdown.
4. When the supply fan is off, the refrigerant solenoid valve returns to its NC position and the compressor relay to its NO position.

At light loads, bypass rates are high and ice may build up on coils. Therefore, control is poor at light loads with this system.

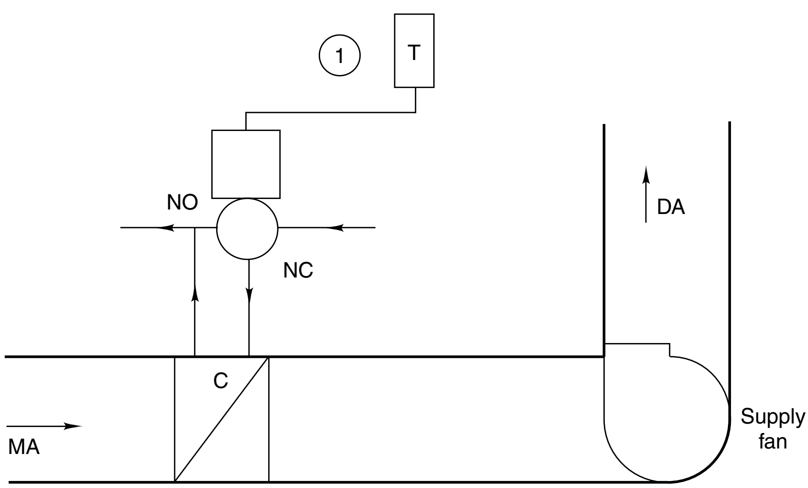


FIGURE 11.23 Cooling-coil control subsystem using three-way diverting valve.

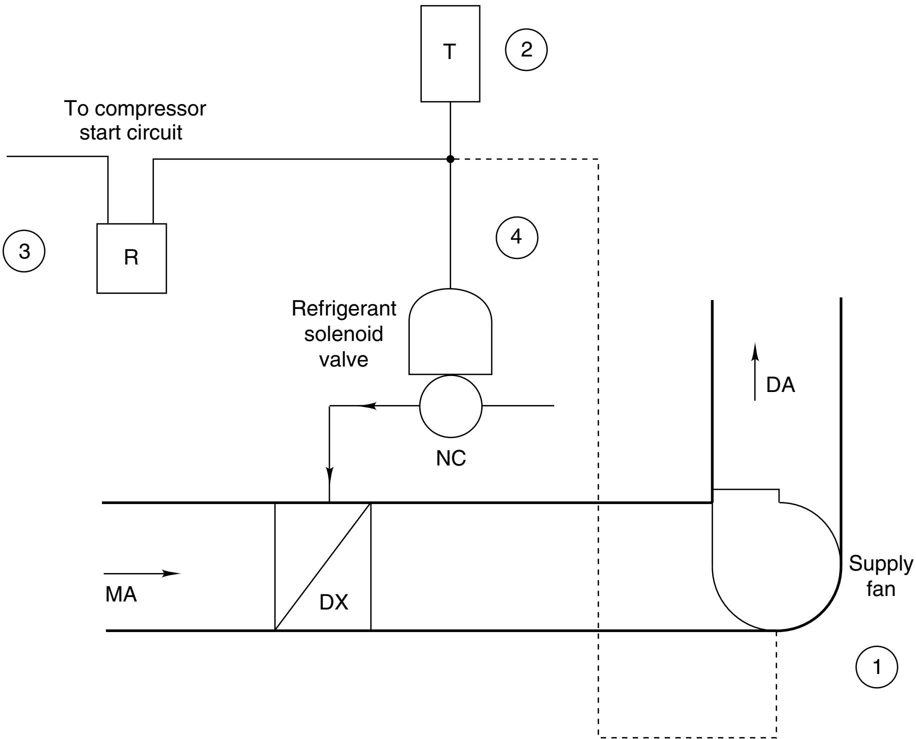


FIGURE 11.24 DX cooling coil control subsystem (on-off control).

11.5.4 Complete Systems

The preceding five example systems are actually control subsystems that must be integrated into a single control system for the HVAC system's primary and secondary systems. In the remainder of this section, two complete HVAC control systems widely used in commercial buildings will be briefly described. The first is a constant volume system, and the second is a VAV system.

Figure 11.25 shows a constant volume, central air-handling system equipped with supply and return fans, heating and cooling coils, and economizer for a single-zone application. If the system were to be used for multiple zones, the zone thermostat shown would be replaced by a discharge air temperature sensor. This constant volume system operates as follows:

1. When the fan is energized, the control system is activated.
2. The minimum outside air setting is set (usually only once, during commissioning, as described above).
3. The OA temperature sensor supplies a signal to the damper controller.
4. The RA temperature sensor supplies a signal to the damper controller.
5. The damper controller positions the dampers to use outdoor or return air, depending on which is cooler.
6. The mixed-air low-temperature controller controls the outside air dampers to avoid excessively low-temperature air from entering the coils. If a preheating system were included, this sensor would control it.
7. The space temperature sensor resets the coil discharge air PI controller.
8. The discharge air controller controls the
   a. Heating coil valve
   b. Outdoor air damper

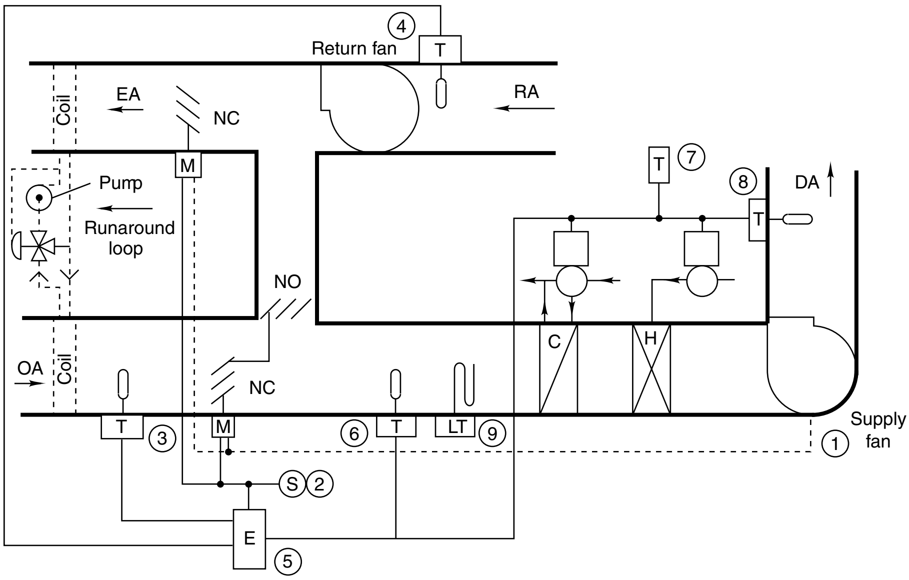


FIGURE 11.25 Control for a complete, constant volume HVAC system. Optional runaround heat recovery system is shown to left in dashed lines.

c. Exhaust air damper
d. Return air damper
e. Cooling coil valve (after the economizer cycle upper limit is reached)

9. The low-temperature sensor initiates freeze protection measures as described previously.

A method for reclaiming either heating or cooling energy is shown by dashed lines on the left side of Figure 11.25. This so-called "runaround" system extracts energy from exhaust air and uses it to precondition outside air. For example, the heating season exhaust air may be at 75°F, while outdoor air is at 10°F. The upper coil in the figure extracts heat from the 75°F exhaust and transfers it through the lower coil to the 10°F intake air. To avoid icing of the air intake coil, the three-way valve controls this coil's liquid inlet temperature to a temperature above freezing. In heating climates, the liquid loop should also be freeze protected with a glycol solution. Heat reclaiming systems of this type can also be effective in the cooling season, when outdoor temperatures are well above indoor temperatures.

A VAV system has additional control features including a motor speed control (or inlet vanes in some older systems) and a duct static pressure control. Figure 11.26 shows a VAV system serving both perimeter and interior zones. It is assumed that the core zones always require cooling during the occupied period. The system shown has a number of options, and does not include every feature present in all VAV systems. However, it is representative of VAV design practice. The sequence of operation during the heating season is as follows:

1. When the fan is energized, the control system is activated. Prior to activation, during unoccupied periods the perimeter zone baseboard heating is under control of room thermostats.
2. Return and supply fan interlocks are used to prevent pressure imbalances in the supply air ductwork.
3. The mixed air sensor controls the outdoor air dampers or preheat coil (not shown) to provide proper coil air inlet temperature. The dampers will typically be at their minimum position at about 40°F.
4. The damper minimum position controls the minimum outdoor airflow.

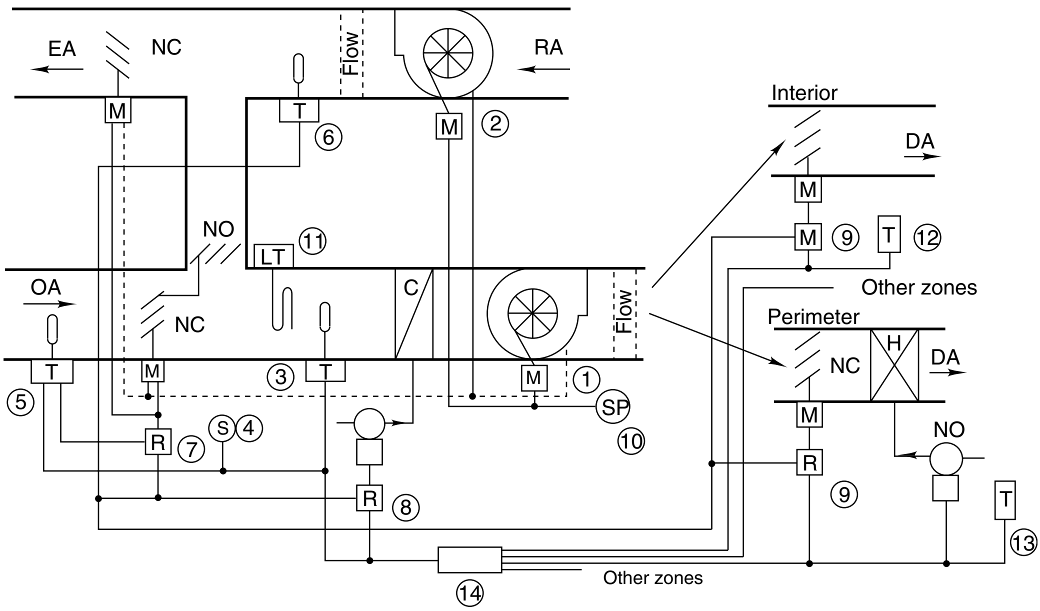


FIGURE 11.26 Control for complete, VAV system. Optional supply and return flow stations shown with dashed lines.

5. As the upper limit for economizer operation is reached, the OA dampers are returned to their minimum position.
6. The return air temperature is used to control the morning warmup cycle after night setback (option present only if night setback is used).
7. The outdoor air damper is not permitted to open during morning warmup, by action of the relay shown.
8. Likewise, the cooling coil valve is de-energized (NC) during morning warmup.
9. All VAV box dampers are moved full open during morning warmup by action of the relay override. This minimizes warmup time. Perimeter zone coils and baseboard units are under control of the local thermostat.
10. During operating periods, the PI static pressure controller controls both supply and return fan speeds (or inlet vane positions) to maintain approximately 1.0 in. WG of static pressure at the pressure sensor location (or, optionally, to maintain building pressure). An additional pressure sensor (not shown) at the supply fan outlet will shut down the fan if fire dampers or other dampers should close completely and block airflow. This sensor overrides the duct static pressure sensor shown.
11. The low-temperature sensor initiates freeze-protection measures.
12. At each zone, room thermostats control VAV boxes (and fans, if present); as zone temperature rises the boxes open more.
13. At each perimeter zone room, thermostats close VAV dampers to their minimum settings and activate zone heat (coil or perimeter baseboard) as zone temperature falls.
14. The controller, using temperature information for all zones (or at least for enough zones to represent the characteristics of all zones), modulates outdoor air dampers (during economizer operation) and the cooling control valve (above the economizer cycle cutoff), to provide air sufficiently cooled to maintain acceptable zone humidity and meet the load of the warmest zone.

The duct static pressure controller is critical to the proper operation of VAV systems. The static pressure controller must be of PI design because a proportional-only controller would permit duct pressure to drift upward as cooling loads drop due to the unavoidable offset in P-type controllers. In addition, the control system should position inlet vanes (if present) closed during fan shutdown, to avoid overloading on restart.

Return fan control is best achieved in VAV systems by an actual flow measurement in supply and return ducts as shown by dashed lines in the figure. The return airflow rate is the supply rate less local exhausts (fume hoods, toilets, etc.) and exfiltration needed to pressurize the building.

VAV boxes are controlled locally, assuming that adequate duct static pressure exists in the supply duct and that supply air is at an adequate temperature to meet the load (this is the function of the controller described in item 11.14). Figure 11.27 shows a local control system used with a series-type, fan-powered VAV box. This particular system delivers a constant flow rate to the zone by action of the airflow controller, to assure proper zone air distribution. Primary air varies with cooling load, as shown in the lower part of the figure. Optional reheating is provided by the coil shown.

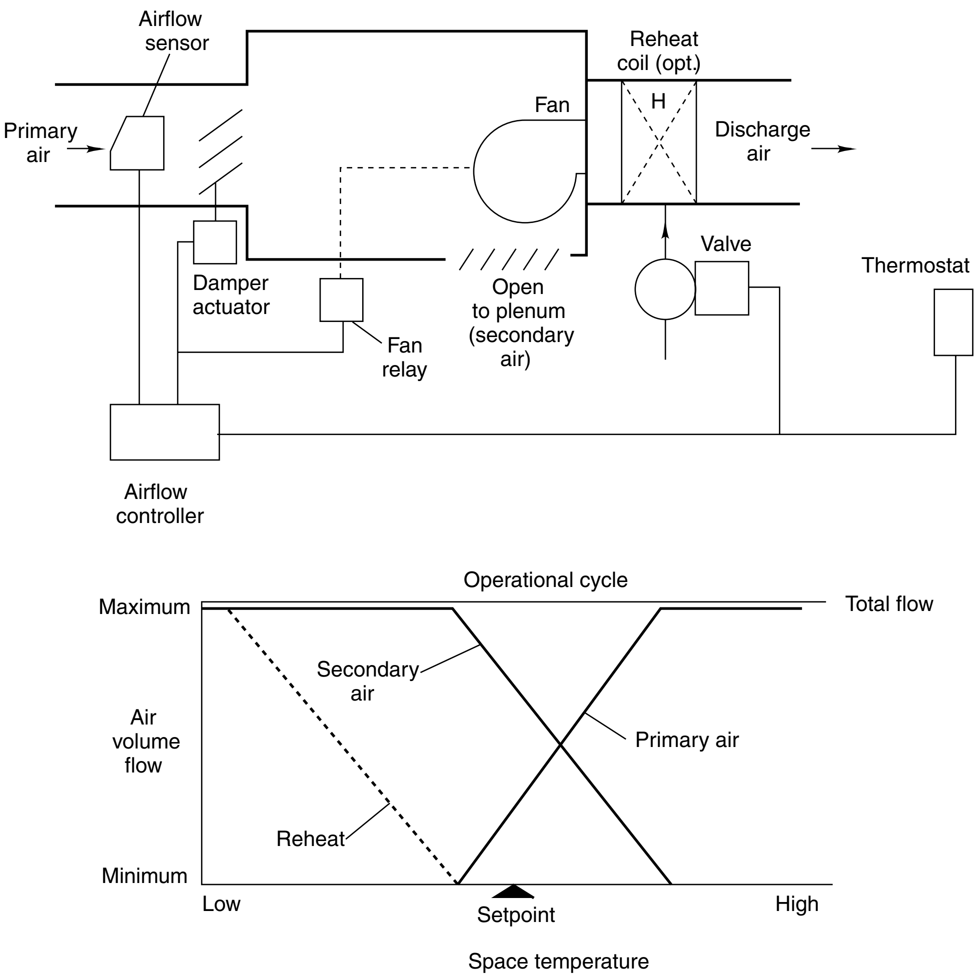


FIGURE 11.27 Series type, fan powered VAV box control subsystem and primary flow characteristic. The total box flow is constant at the level identified as "maximum" in the figure. The difference between primary and total air flow is secondary air recirculated through the return air grille. Optional reheat coil requires air flow shown by dashed line.

11.5.5 Other Systems

This section has not covered the control of central plant equipment such as chillers and boilers. Most primary system equipment controls are furnished with the equipment and as such do not offer much flexibility to the designer. However, Braun et al. (1989) have shown that considerable energy savings can be made by properly sequencing cooling tower stages on chiller plants, and by properly sequencing chillers themselves in multiple chiller plants.

Fire and smoke control are important for life safety in large buildings. The design of smoke control systems is controlled by national codes. The principal goal is to eliminate smoke from the zones where it is present, while keeping adjacent zones pressurized to prevent smoke infiltration. Some components of space conditioning systems (e.g., fans) can be used for smoke control, but HVAC systems are generally not designed to be smoke control systems.

Electrical systems are primarily the responsibility of the electrical engineer on a design team. However, HVAC engineers must make sure that the electrical design accommodates the HVAC control system. Interfaces between the two occur where the HVAC controls activate motors on fans or chiller compressors, pumps, electrical boilers, or other electrical equipment.

In addition to electrical specifications, the HVAC engineer often conveys electrical control logic using a ladder diagram. An example, for the control of the supply and return fans in a central system, is shown in Figure 11.28. The electrical control system, shown at the bottom, operates on low voltage (24 or 48 VAC) from the control transformer shown. The supply fan is started manually by closing the "start" switch. This activates the motor starter coil labeled 1M, thereby closing the three contacts labeled 1M in the supply fan circuit. The fourth 1M contact (in parallel with the start switch) holds the starter closed after the start button is released.

The hand-off auto switch is typical, and allows both automatic and manual operation of the return fan. When switched to the "hand" position, the fan starts. In the "auto" position, the fans will operate only when the adjacent contacts 3M are closed. Either of these actions activates the relay coil 2M, which in turn closes the three 2M contacts in the return fan motor starter. When either fan produces actual airflow, a flow switch is closed in the ducting, thereby completing the circuit to the pilot lamps L. The fan motors are protected by fuses and thermal overload heaters. If motor current draw is excessive, the heaters shown in the figure produce sufficient heat to open the normally closed thermal overload contacts.

FIGURE 11.28 The Brooke Army Medical Center (BAMC) in San Antonio, Texas.

This example ladder diagram is primarily illustrative, and is not typical of an actual design. In a fully automatic system, both fans would be controlled by 3M contacts actuated by the HVAC control system. In a fully manual system, the return fan would be activated by a fifth 1M contact, not by the 3M automatic control system.

11.6 Commissioning and Operation of Control Systems

This chapter emphasizes the importance of making sound decisions in the design of HVAC control systems. It is also extremely important that the control system be commissioned and used properly. The design process requires many assumptions about the building and its use. The designer must be sure that the systems will provide comfort under extreme conditions, and the sequence of design decisions and construction decisions often leads to systems that are substantially oversized. Operation at loads far below design conditions is generally much less efficient than at larger loads. Normal control practice can be a major contributor to this inefficiency. For example, it is quite common to see variable volume air handler systems operating at minimum flow as constant volume systems almost all the time, due to design flows that are sometimes twice as large as the maximum flow used in the building.

Thus, it is very important that following construction, the control system and the rest of the HVAC system be commissioned. This process (ASHRAE 2005) normally seeks to ensure that the control system operates according to design intent. This is really a minimum requirement to be sure that the system functions as designed. However, after construction, the control system setup can be modified to meet the loads actually present in the building, and to fit the way the building is actually being used, rather than basing these decisions on the design assumptions. If the VAV system is designed for more flow than is required, minimum flow settings of the terminal boxes can be reduced below the design value, ensuring that the system will operate in the VAV mode most of the time. Numerous other adjustments may be made as well. Such adjustments, commonly made during the version of commissioning known as *Continuous Commissioning*®[2](CC®), can frequently reduce the overall building energy use by 10% or more (Liu, Claridge, and Turner 2002). If the process is applied to an older building where control practices have drifted away from design intent and undetected component failures have further eroded system efficiency, energy savings often exceed 20% (Claridge et al. 2004).

11.6.1 Control Commissioning Case Study

A case study in which this process was applied to a major Army hospital facility located in San Antonio, Texas is reported in Zhu et al. (2000a, 2000b, 2000c). The Brooke Army Medical Center (BAMC) was a relatively new facility when the CC® process was begun. The facility was operated for the Army by a third-party company, and it was operated in accordance with the original design intent.

BAMC is a large, multifunctional medical facility with a total floor area of 1,349,707 ft.^{2}. The complex includes all the normal inpatient facilities, as well as outpatient and research areas. The complex is equipped with a central energy plant, which has four 1,200 ton water-cooled electric chillers. Four primary pumps (75 hp each) are used to pump water through the chillers. Two secondary pumps (200 hp each) equipped with VFDs supply chilled water from the plant to the building entrance. Fourteen chilled water risers equipped with 28 pumps totaling 557 hp are used to pump chilled water to all of the AHUs and small fan coil units. All of the chilled water riser pumps are equipped with VFDs. There are four natural gas-fired steam boilers in this plant. The maximum output of each boiler is 20 MMBtu/h. Steam is supplied to each building, where heating water was generated, at 125 psi (prior to commissioning).

There are 90 major AHUS serving the whole complex, with a total fan power of 2570 hp. VFDs are installed on 65 AHUs while the others are constant volume systems. There are 2,700 terminal boxes in the

[2]Continuous Commissioning and CC are registered trademarks of the Texas Engineering Experiment Station.

complex, of which 27% are dual duct variable volume (DDVAV) boxes, 71% are dual duct constant volume (DDCV) boxes, and 2% are single duct variable volume (SDVAV) boxes.

The HVAC systems (chillers, boilers, AHUs, pumps, terminal boxes and room conditions) are controlled by a DDC control system. Individual controller-field panels are used for the AHUs and water loops located in the mechanical rooms. The control program and parameters can be changed by either the central computers or the field panels.

11.6.1.1 Design Conditions

The design control program was being fully utilized by the EMCS. It included the following features:

1. Hot deck reset control for AHUS
2. Cold deck reset during unoccupied periods for some units
3. Static pressure reset between high and low limits for VAV units
4. Hot water supply temperature control with reset schedule
5. VFD control of chilled water pumps with ΔP setpoint (no reset schedule)
6. Terminal box level control and monitoring

It was also determined that the facility was being well maintained by the facility operator, in accordance with the original design intent. The building is considered energy efficient for a large hospital complex.

The commissioning activities were performed at the terminal box level, AHU level, loop level and central plant level. Several different types of improved operation measures and energy solutions were implemented in different HVAC systems, due to the actual function and usage of the areas and rooms. Each measure will be discussed briefly, starting with the air-handling units.

11.6.1.2 Optimization of AHU Operation

EMCS trending, complemented by site measurements and use of short-term data loggers, found that many supply fans operated above 90% of full speed most of the time. Static pressures were much higher than needed. Wide room temperature swings due to AHU shutoff led to hot and cold complaints in some areas. Through field measurements and analysis, the following possible means of improving the operation of the two AHUs were identified.

- Improve zone air balancing and determine new static pressure setpoints for VFDs
- Optimize the cold deck temperature setpoints with reset schedules
- Optimize the hot deck temperature reset schedules
- Improve control of outside air intake and relief dampers during unoccupied periods to reduce ventilation during these periods
- Optimize time schedule for fans to improve room conditions
- Improve the preheat temperature setpoint to avoid unnecessary preheating

Implementation of these measures improved comfort and reduced heating, cooling, and electric use.

11.6.1.3 Optimization at the Terminal Box Level

Field measurements showed that many VAV boxes had minimum flow settings that were higher than necessary, and that some boxes were unable to supply adequate hot air due to specific control sequences. New control logic was developed that increased hot air capacity by 30% on average, in the full heating mode, and reduced simultaneous heating and cooling. During unoccupied periods, minimum flow settings on VAV boxes were reduced to zero and flow settings were reduced in constant volume boxes.

During commissioning, it was found that some terminal boxes could not provide the required airflow either before or after the control program modification. Specific problems were identified in about 200 boxes, with most being high flow resistance due to kinked flex ducts.

11.6.1.4 Water Loop Optimization

There are 14 chilled water risers, equipped with 28 pumps, which provide chilled water to the entire complex. During the commissioning assessment phase, the following were observed:

- All the riser pumps were equipped with VFDs, which were running at 70%–100% of full speed.
- All of the manual balancing valves on the risers were only 30%–60% open.
- The ΔP sensor for each riser was located 10–20 ft. from the far-end coil of the AHU on the top floor.
- Differential pressure setpoints for each riser ranged from 13 to 26 psi.
- There was no control valve on the return loop.
- Although most of the cold deck temperatures were holding well, there were 13 AHUs whose cooling coils were 100% open, but which could not maintain cold deck temperature setpoints.

Because the risers are equipped with VFDs, traditional manual balancing techniques are not appropriate. All the risers were rebalanced by initially opening all of the manual balancing valves. The actual pressure requirements were measured for each riser, and it was determined that the ΔP for each riser could be reduced significantly. Pumping power requirements were reduced by more than 40%.

11.6.1.5 Central Plant Measures

1. *Boiler System*: Steam pressure was reduced from 125 psi to110 psi, and one boiler operated instead of two during summer and swing seasons.
2. *Chilled Water Loop*: Before the commissioning, the blending valve separating the primary and secondary loops at the plant was 100% open. The primary and secondary pumps were both running. The manual valves were partially open for the secondary loop, although the secondary loop pumps are equipped with VFDs. After the commissioning assessment and investigations, the following were implemented:

- Open the manual valves for the secondary loop
- Close the blending stations
- Shut down the secondary loop pumps

As a result, the primary loop pumps provide required chilled water flow and pressure to the building entrance for most of the year, and the secondary pumps stay offline most of the time. The operator drops the online chiller numbers according to the load conditions, and the minimum chilled water flow can be maintained to the chillers. At the same time, the chiller efficiency is increased.

11.6.1.6 Results

For the fourteen-month period following initial CC® implementation, measured savings were nearly \$410,000, or approximately \$30,000/month, for a reduction in both electricity and gas use of about 10%. The contracted cost to meter, monitor, commission, and provide a year's follow-up services was less than \$350,000. This cost does not include any time for the facilities operating staff who repaired kinked flex ducts, replaced failed sensors, implemented some of the controls and subroutines, and participated in the commissioning process.

11.6.2 Commissioning Existing Buildings

The savings achieved from commissioning HVAC systems in older buildings are even larger. In addition to the opportunities for improving efficiency similar to those in new buildings, opportunities come from:

- Control changes that have been made to "solve" problems, often resulting in lower operating efficiency

- Component failures that compromise efficiency without compromising comfort
- Deferred maintenance that lowers efficiency

Mills et al. (2004 and 2005) surveyed 150 existing buildings that had been commissioned and found median energy cost savings of 15%, with savings in one-fourth of the buildings exceeding 29%. Over 60% of the problems corrected were control changes, and another 20% were related to faulty components that prevented proper control. This suggests that relatively few control systems actually achieve the efficiency they are capable of providing.

11.7 Advanced Control System Design Topics: Neural Networks

Neural networks offer considerable opportunity to improve the control possible in standard PID systems. This section provides a short introduction to this novel approach to control.

11.7.1 Neural Network Introduction

An artificial neural network is a massively parallel, dynamic system of interconnected, interacting parts based on some aspects of the brain. Neural networks are considered to be intuitive because they learn by example rather than by following programmed rules. The ability to "learn" is one of the key aspects of neural networks. A neural network consists of several layers of neurons that are connected to each other. A *connection* is a unique information transport link from one sending to one receiving neuron. The structure of part of an NN is schematically shown in Figure 11.29. Any number of input, output, and "hidden layer" neurons can be used (only one hidden layer is shown). One of the challenges of this technology is to construct a net with sufficient complexity to learn accurately without imposing a burden of excessive computational time.

The neuron is the fundamental building block of a network. A set of inputs is applied to each. Each element of the input set is multiplied by a weight, indicated by the W in the figure, and the products are summed at the neuron. The symbol for the summation of weighted inputs is termed $INPUT$ and must be calculated for each neuron in the network. In equation form this process for one neuron is

$$INPUT = \sum_i O_i W_i + B \tag{11.20}$$

where O_i are inputs to a neuron, i.e., outputs of the previous layer, W_i are weights, and B is the bias. After $INPUT$ is calculated, an activation function, F, is applied to modify it, thereby producing the neuron's output as described shortly.

Artificial networks have been trained by a wide variety of methods (McClelland and Rumelhart 1988). Back-propagation is one systematic method for training multilayer neural networks. The weights of a net are initiated with small random numbers. The objective of training the network is to adjust the weights iteratively so that application of a set of inputs produces the desired set of outputs matching a training data set. Usually a network is trained with a data set that consists of many input–output pairs; these data are called a *training set*. Training the net using back-propagation requires the following steps:

1. Select a training pair from the training set and apply the input vector to the network input layer.
2. Calculate the output of the network, OUT_i.
3. Calculate the error, $ERROR_i$, the network output, and the desired output (the target vector from the training pair).
4. Adjust the weights of the network in a way that minimizes the error.
5. Repeat steps 1 through 4 for each vector in the training set until the error for the entire set is lower than the user specified, preset training tolerance

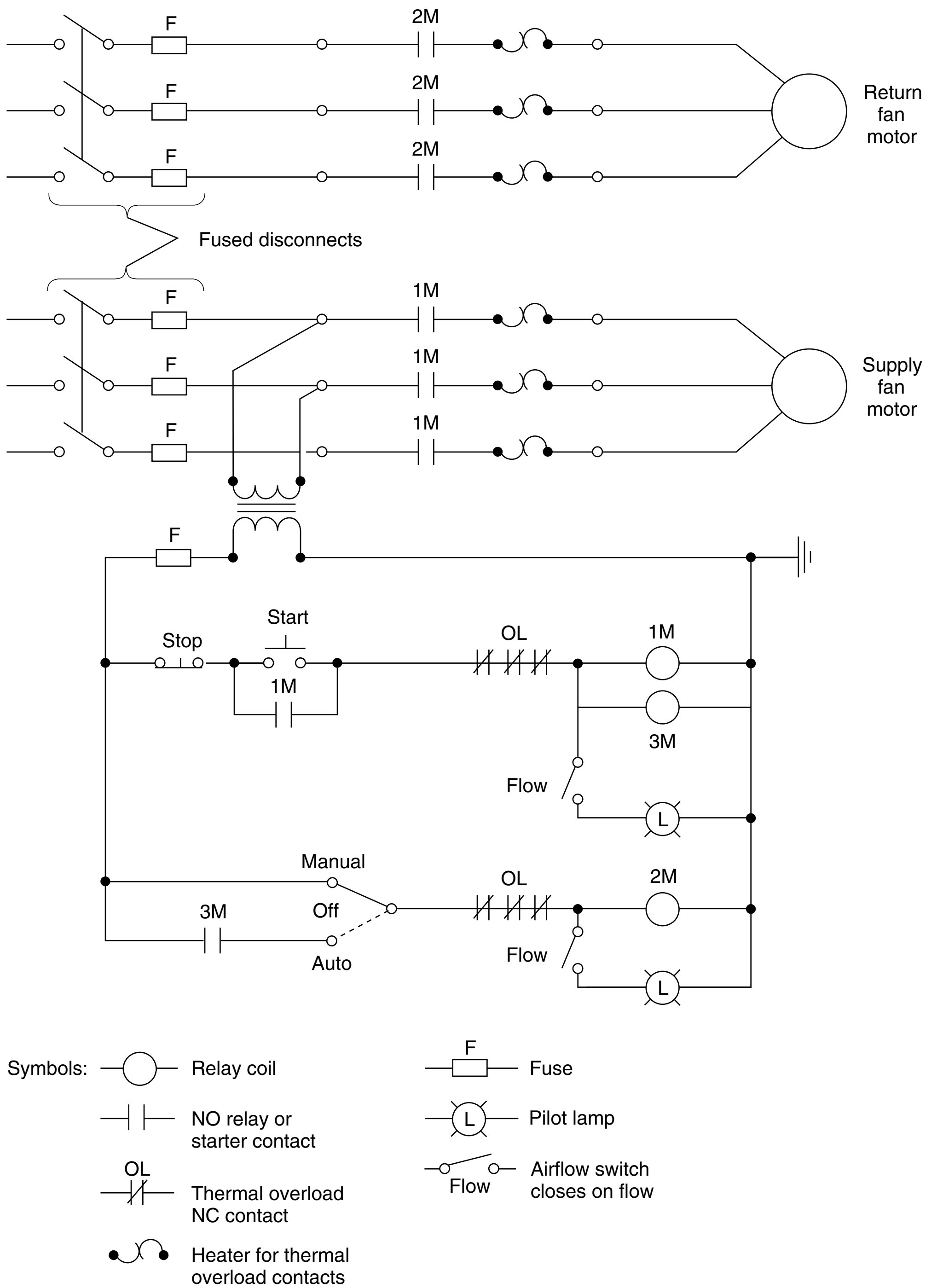


FIGURE 11.29 Ladder diagram for supply and return fan control. Hand-off auto switch permits manual or automatic control of the return fan.

Steps 1 and 2 are the "forward pass." The following expression describes the calculation process in which an activation function, *F*, is applied to the weighted sum of inputs, *INPUT*, as follows.

$$OUT = F(INPUT) = F\left(\sum_i O_i W_i + B\right), \tag{11.21}$$

where F is the activation function and B is the bias of each neuron.

The activation function used for this work was selected to be

$$F(INPUT) = \frac{1}{1 + \mathrm{e}^{-INPUT}}. \tag{11.22}$$

This is referred to as a *sigmoid function* and is shown in Figure 11.30. It has a value of 0.0 when *INPUT* is a large negative number and a value of 1.0 for large and positive *INPUT*, making a smooth transition between these limiting values. The bias, B, is the activation threshold for each neuron. The bias avoids the tendency of a sigmoid function to get "stuck" in the saturated, limiting value area.

Steps 3 and 4 comprise the "reverse pass" in which the delta rule is used as follows: for each neuron in the output layer, the previous weight $W(n)$ is adjusted to a new value $W(n+1)$ to reduce the error by the following rule:

$$W(n+1) = W(n) + (\eta\delta)OUT, \tag{11.23}$$

where $W(n)$ is the previous value of a weight, $W(n+1)$ is the weight after adjusting, η is the training rate coefficient. δ is calculated from

$$\delta = \left(\frac{\partial INPUT}{\partial OUT}\right)(TARGET - OUT) = OUT(1 - OUT)(TARGET - OUT), \tag{11.24}$$

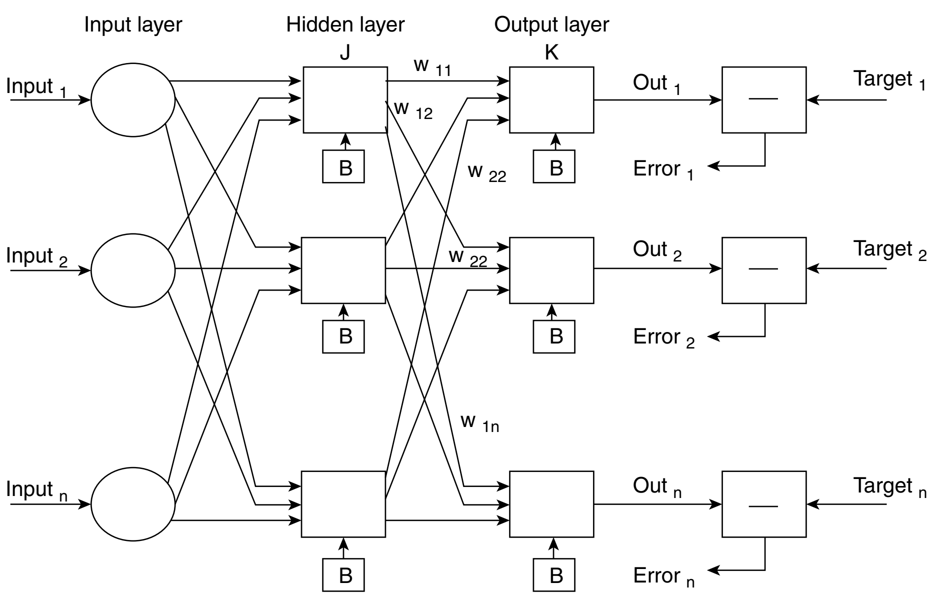


FIGURE 11.30 Schematic diagram of a neural network showing input layer, hidden layers, and output along with target training values. Hidden and output layers consist of connected neurons; the input layer does not contain neurons.

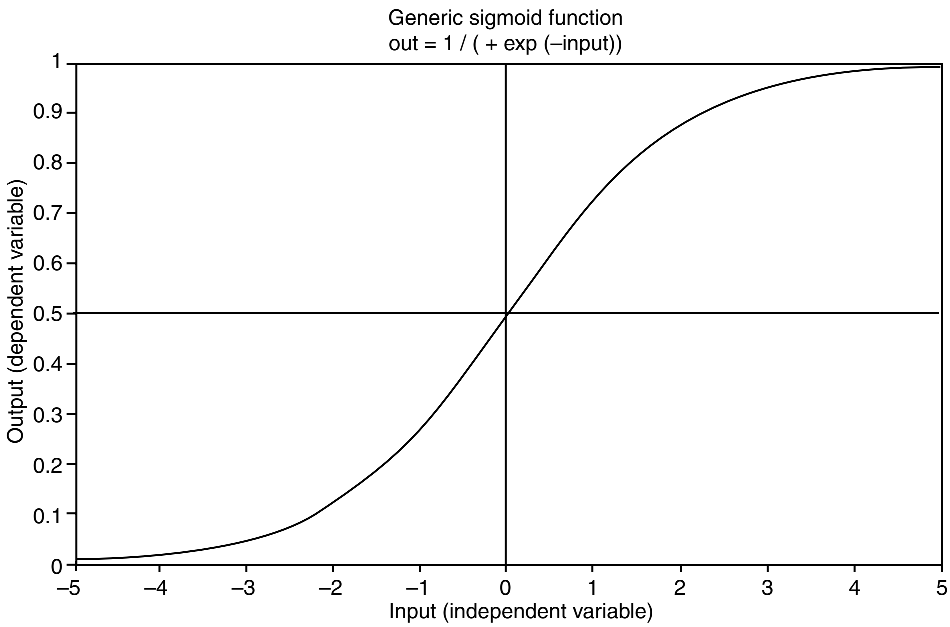


FIGURE 11.31 Sigmoid function used to process the weighted sum of network inputs.

in which the derivative has been calculated from Equation 11.21 and Equation 11.22, and *TARGET* (see Figure 11.29) is the training-set target value. This method of correcting weights bases the magnitude of the correction on the error itself.

Of course, hidden layers have no target vector; therefore, back-propagation trains these layers by propagating the output error back through the network layer by layer, adjusting weights at each layer. The delta rule adjustment δ is calculated from

$$\delta_j = OUT(1 - OUT) \sum (\delta_{j+1} W_{j+1}) \tag{11.25}$$

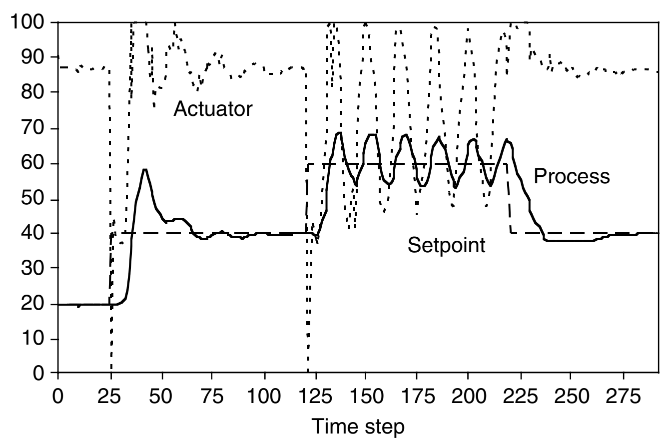


FIGURE 11.32 PID controller response to step changes in coil load. Proportional gain of 2.0. (From Curtiss, P. S., Kreider, J. F., and Brandemuehl, M. J., *ASHRAE Transactions*, 99 (1) 1993.)

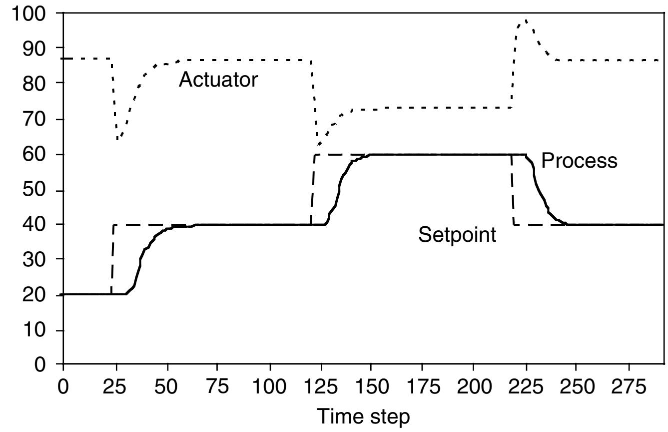


FIGURE 11.33 NN controller with learning rate of 1.0 and window of 15 time steps. (From Curtiss, P. S., Kreider, J. F., and Brandemuehl, M. J., *ASHRAE Transactions,* 99 (1) 1993.)

where δ_j and δ_{j+1} belong to the jth and $(j+1)$th hidden layers, respectively (being numbered with increasing values from left to right in Figure 11.29). This overall method of adjusting weights belongs to the general class of steepest descent algorithms. The weights and biases after training contain meaningful system information; before training, the initial, random biases and random weights have no physical meaning.

11.7.2 Commercial Building Adaptive Control Example

A proof of concept experiment in which neural networks (NNs) were used for both local and global control of a commercial building HVAC system was conducted in the JCEM laboratory in which full-scale, repeatable testing of multizone HVAC systems can be done. Data collected in the laboratory were used to train NNs for both the components and the full systems involved (Curtiss, Brandemuehl, and Kreider 1993; Curtiss, Kreider, and Brandemuehl 1993). Any neural network-based controller will be useful only if it can perform better than a conventional PID controller. Figure 11.31 and Figure 11.32 show typical results for the PID and NN control of a heating coil. The difficulty that the PID controller experienced is due to the highly nonlinear nature of the heating coil. A PID controller tuned at one level of load is unable to control acceptably at another, while the NN controller does not have this difficulty. With the NN controller, excellent control is demonstrated—minimal overshoot and quick response to the setpoint changes.

In an affiliated study, Curtiss, Brandemuehl, and Kreider (1993) showed that NNs offered a method for global control of HVAC systems as well. The goal of such controls could be to reduce energy consumption as much as possible, while meeting comfort conditions as a constraint. Energy savings of over 15% were achieved by the NN method vs. standard PID control (Figure 11.33).

11.8 Summary

This chapter has introduced the important features of properly designed control systems for HVAC applications. Sensors, actuators, and control methods have been described. Methods for determining control system characteristics, either analytically or empirically, have been discussed.

The following rules (adapted from ASHRAE 1987) should be followed to ensure that the control system is as energy efficient as possible. Neural networks offer one method for achieving energy efficient control.

1. Operate HVAC equipment only when the building is occupied or when heat is needed to prevent freezing.
2. Consider the efficacy of night setback vis à vis building mass. Massive buildings may not benefit from night setback due to the overcapacity needed for the morning pickup load.
3. Do not supply heating and cooling simultaneously. Do not supply humidification and dehumidification at the same time.
4. Reset heating and cooling air or water temperature to provide only the heating or cooling needed.
5. Use the most economical source of energy first, the most costly last.
6. Minimize the use of outdoor air during the deep heating and cooling seasons, subject to ventilation requirements.
7. Consider the use of "dead-band" or "zero-energy" thermostats.
8. Establish control settings for stable operation to avoid system wear and to achieve proper comfort.
9. Commission the control system and HVAC system for optimum efficiency based on actual building conditions and use.

References

ASHRAE. 2001. *BACnet: A Data Communication Protocol for Building Automation and Control Networks*. ASHRAE., ANSI/ASHRAE Standard 135-2001.

ASHRAE. 2002. *HVAC Applications*, ASHRAE, Atlanta, GA.

ASHRAE. 2003. *HVAC Systems and Equipment*, ASHRAEA, Atlanta, GA.

ASHRAE. 2004. Fundamentals of control. In *Handbook of Fundamentals*, ASHRAE, Atlanta, GA.

ASHRAE. 2005. *Guideline 0, Commissioning,* ASHRAE, Atlanta, GA.

Bauman, H. D. 1998. *Control Value Primer, 2nd Ed.,* Instrument Society of America..

Braun, J. E., Klein, S. A., Mitchell, J. W., and Beckman, W. A. 1989. Applications of optimal control to chilled water systems without storage, *ASHRAE Transactions*, 95, 1.

Claridge, D. E., Turner, W. D., Liu, M., Deng, S., Wei, G., Culp, C., Chen, H., and Cho, S. Y. 2004. Is commissioning once enough? *Energy Engineering*, 101, 4, 7–19.

Curtiss, P. S., Brandemuehl, M. J., and Kreider, J. F. 1993. Energy management in central HVAC plants using neural networks. *ASHRAE Transactions*, 100, 1.

Curtiss, P. S., Kreider, J. F., and Brandemuehl, M. J., 1993 Adaptive control of HVAC processes using predictive neural networks. *ASHRAE Transactions*, Vol. 99, Pt. 1.

Grimm, N. R. and Rosaler, R. C. 1990. Handbook of HVAC Design, McGraw-Hill, New York.

Haines, R. W. 1987. *Control Systems for Heating, Ventilating and Air Conditioning, 4th Ed*., NewYork, Van Nostrand Reinhold.

Honeywell, Inc. 1988. *Engineering Manual of Automatic Control,* Honeywell, Inc., Minneapolis, MN.

Huang, P. H. 1991. Humidity measurements and calibration standards. *ASHRAE Transactions*, 97, 2.

Letherman, K. M. 1981. *Automatic Controls for Heating and Air Conditioning,* Pergamon Press, New York.

Levine, W. S. 1996. *The Control Handbook,* CRC Press, Boca Raton, FL.

Liu, M., Claridge, D. E., and Turner, W. D. 2002. *Continuous Commissioning^{SM} Guidebook: Maximizing Building Energy Efficiency and Comfort*. Federal Energy Management Program, US Department of Energy. Available at http://www.eere.energy.gov/femp/operations_maintenance/commissioning_-guidebook.cfm (accessed on October 16, 2006).

McClelland, J. L. and Rumelhart, D. E. 1988. *Exploration in Parallel Distributed Processing,* MIT Press, Cambridge, MA.

Mills, E., Friedman, H., Powell, T., Bourassa, N., Claridge, D., Haasl, T., and Piette, M. A. 2004. The cost-effectiveness of commercial-buildings commissioning: A meta-analysis of energy and non-energy impacts in existing buildings and new construction in the United States, Lawrence Berkeley National Laboratory Report No. 56637. Available at http://eetd.lbl.gov/emills/PUBS/Cx-Costs-Benefits.html (accessed on October 16, 2006).

Mills, E., Bourassa, N., Piette, M. A., Friedman, H., Haasl, T., Powell, T., Claridge, D. 2005. The cost-effectiveness of commissioning new and existing commercial buildings: Lessons from 224 buildings. In *Proceedings of 13th National Conference on Building Commissioning*, Portland Energy Conservation, Inc., Portland, OR.

Sauer, H. J., Howell, R. H., and Coad, W. J. 2001. *Principles of Heating, Ventilating and Air Conditioning,* ASHRAE, Atlanta, GA.

Stein, B. and Reynolds, J. S. 2000. *Mechanical and Electrical Equipment for Buildings,* Wiley, New York.

Tao, W. K. Y. and Janis, R. R. 2005. *Mechanical and Electrical Systems in Buildings,* Pearson Prentice Hall, Upper Saddle River, NJ.

Zhu, Y., Liu, M., Batten, T., Noboa, H., Claridge, D. E., and Turner, W.D. 2000a. Optimization control strategies for HVAC terminal boxes. In *Proceedings of Twelfth Symposium on Improving Building Systems in Hot and Humid Climates*, Energy Systems Laboratory, Texas A & M University, College Station, TX.

Zhu, Y., Liu, M., Batten, T., Turner, W. D., and Claridge, D. E. 2000b. A simple and quick chilled water loop balancing for variable flow systems. In *Proceedings of Twelfth Symposium on Improving Building Systems in Hot and Humid Climates*, Energy Systems Laboratory, Texas A & M University, College Station, TX.

Zhu, Y., Liu, M., Batten, T., Turner, W. D., Claridge, D. E., Noboa, H., and Zhou, J. 2000c. Integrated commissioning for a large medical facility. In *Proceedings of Twelfth Symposium on Improving Building Systems in Hot and Humid Climates*, Energy Systems Laboratory, Texas A & M University, College Station, TX.

12

Energy Efficient Technologies

Frank Kreith
University of Colorado

Aníbal T. de Almeida
Universidade de Coimbra

Katherine Johnson
Market Development Group

James E. McMahon,
Barbara Atkinson,
Peter Biermayer,
Robert Clear,
Andrea Denver,
Steve Greenberg,
Alex Lekov,
James Lutz,
Stephen Meyers, and
Greg Rosenquist
Lawrence Berkeley National Laboratory

12.0 Introduction

Frank Kreith

Energy efficiency is defined as the ratio of energy required to perform a specific service to the amount of primary energy used for the process. Improving energy efficiency increases the productivity of basic energy sources by providing given services with less energy resources. For example, space conditioning, lighting or mechanical power can be provided with less input of coal, solar, wind, or uranium in a more energy efficient system.

If energy efficiency improvement and energy conservation in the United States were pursued vigorously and consistently with realistic energy price signals, the total cumulative total energy savings from higher energy efficiency standards for residential and commercial equipment that would be effective in the years 2010–2030 amounts to just under 26 quads [1]. Annual savings amounting to one and a half and three quads in 2025 have been estimated by Lawrence Berkeley National Labarotory (LBNL) and American Council of Energy Efficiency (ACEE) respectively for improved appliances [2]. An additional savings potential from improved building technologies amounting to 4 quads per year has been estimated to be possible for 2025 by the Commission on Energy Policy. Large energy savings are associated with standards for residential electronics products followed by higher efficiency standards for commercial refrigeration, lighting, and air conditioning. The next largest savings in the residential sectors could come from higher standards for electric water heaters and lighting [1].

Achieving this savings potential could increase national energy security and help improve the nations international balance of payments. Hence improving energy efficiency across all sectors of the economy is an important national objective [3]. It should be noted, however, that free market price signals may not always be sufficient to effect energy efficiency. Hence, legislation on the state and/or national level for energy efficiency standards for equipment in the residential and commercial sector may be necessary. There is considerable debate whether incentives or mandates are the preferred way to improved energy efficiency. Such measures may be necessary because national surveys indicate that consumers consistently rank energy use and operating costs quite low on the lists of attributes they consider when purchasing an appliance or construct a building. Incentives may be the preferred option provided they induce decision makers to take appropriate action. Unfortunately, in the case of buildings and appliances, the long-term economical benefits of conservation do not rank as high as the initial investment costs. Hence, to achieve increased energy efficiency, mandates may be necessary. Mandates are politically acceptable when the required actions are inexpensive, noncontroversial, and simple to perform. When properly enforced, mandates have predictable results and may be the preferred method of achieving energy efficiency [3].

Every energy conservation measure requires an up front capital investment and given usual economic constraints, the initial costs of an energy conservation measure is very important. One of the criteria by which to judge an energy conservation measure is its benefit-to-cost ratio.

There are two aspects to the benefit-to-cost ratio of energy conservation: the total value of BTU's or kilowatt hours saved during the lifetime of the system to the total system cost (investment, operating, maintenance), and the value of yearly net energy savings (i.e., the difference between the energy saved and the energy used for operation and maintenance) divided by the annual levelized cost of the capital equipment. When the value of the saved energy and the cost of installing conservation measure are known, the simple payback period (PP) can be calculated from:

$$\mathrm{PP} = \frac{\text{Cost of installation in \$}}{\text{Net value of energy saved per year in (\$/yr)}}$$

The net value of the saved energy equals the amount of energy saved times the cost per unit of the energy (dollars/BTU or dollars/kwh). This approach is acceptable for preliminary estimates if the payback period is short, say less than four years. For a more precise estimate, the time value of money, the inflation rate and the escalation in fuel costs must be considered as shown in Chapter 3.

By the year 2004, national efficiency standards were in effect for a variety of residential and commercial appliances. Updated standards will take effect in the next few years for several products. Outside the U.S. over 30 countries have adopted minimum energy performance standards. A study conducted for the U.S. National Commission on Energy Policy in 2004 estimated the national impact of new and upgraded energy standards for residential and commercial equipment [1].

The results of this study indicated that the cost of conserved energy (CCE) for the proposed 2010 standards are well below the relevant energy prices for most products as shown in Figure 12.1 for equipment in the residential sector, and in Figure 12.2 for equipment in the commercial sector. The result of this study indicates that most of the standards for improving energy efficiency would be cost-effective. Figure 12.3 shows the cumulative primary energy savings that could be attained by upgrading efficiency standards for such equipment installed in the 2010–2030 period in the residential and commercial sectors.

Traditional energy prices are understated because they do not include the health, social, and environmental costs of using fuels. For example, gasoline prices do not take into account the costs associated with military requirements to protect access to oil sources, global warming, acid rain, and adverse health effects. This is an institutional barrier to increasing energy efficency. Procedures for estimating these social costs are discussed in Chapter 4. Some of the key barriers to achieving increased efficiency are listed below [3]:

Lack of objective consumer information. Efficiency claims in the market place are often made by competing manufacturers, without an objective third party to evaluate the actual efficiency claims.

Failure of consumers to make optimal energy-efficiency decisions. Consumers often choose the least expensive appliance, rather than the appliance that will save them money over the long term; consumers are also often confused about efficiency ratings and efficiency improvements.

Replacement market decisions based on availability rather than efficiency. Decisions concerning replacement of worn out or broken equipment are made without energy efficiency as a high priority.

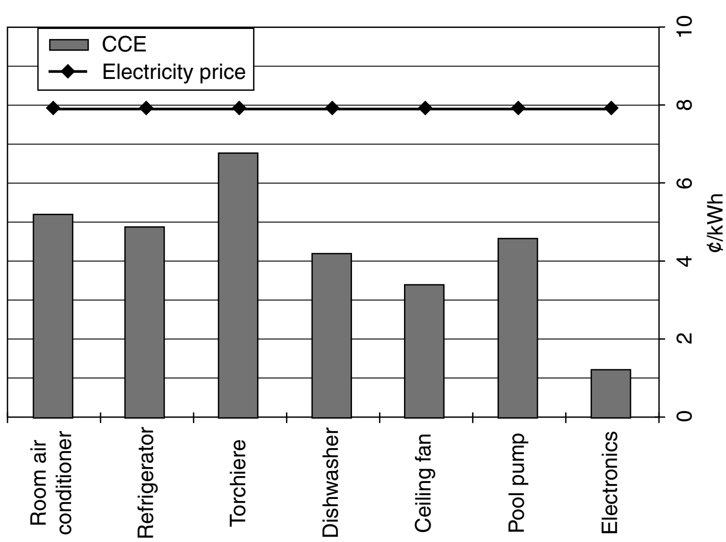


FIGURE 12.1 Comparison of cost of conserved energy for 2010 standards to projected electricity price in the residential sector. (From Ending the Energy Stalemate: A Bipartisan Strategy to Meet America's Energy Challenges, The National Commission On Energy Policy, December 2004, www.energycommission.org).

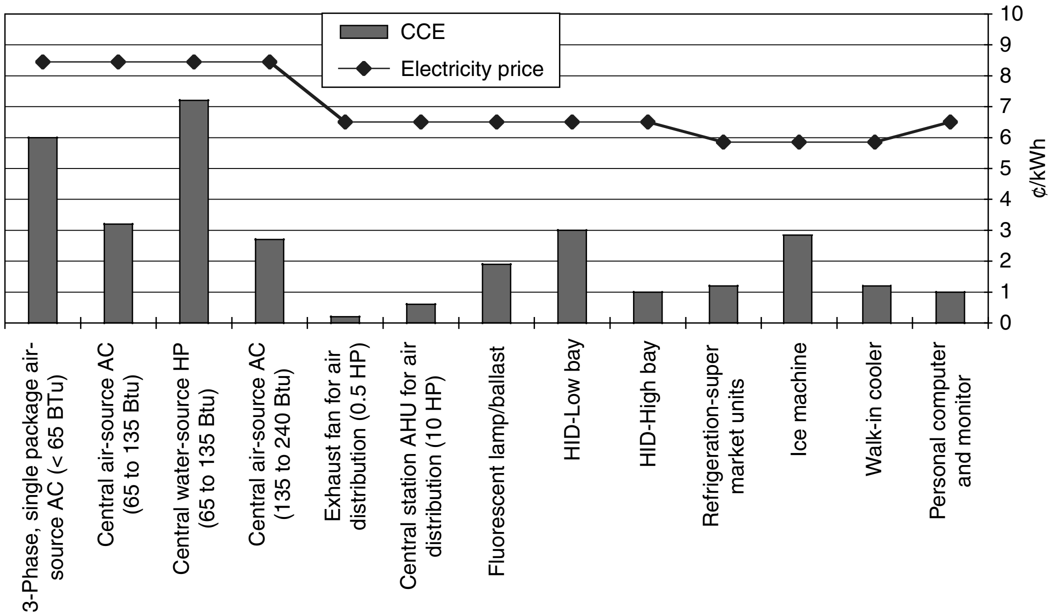


FIGURE 12.2 Comparison of cost of conserved energy for representative 2010 standards to marginal electricity price in the commercial sector. (From Ending the Energy Stalemate: A Bipartisan Strategy to Meet America's Energy Challenges, The National Commission On Energy Policy, December 2004, www.energycommission.org).

Usually the primary concern for the consumer is restoring service as quickly as possible. This requires buying whatever equipment the plumbing or heating contractor may have on hand.

Energy prices do not take into account the full environmental or societal costs. External costs associated with public health, energy production, global warming, acid rain, air pollution, energy security, or reliability of supply are usually ignored.

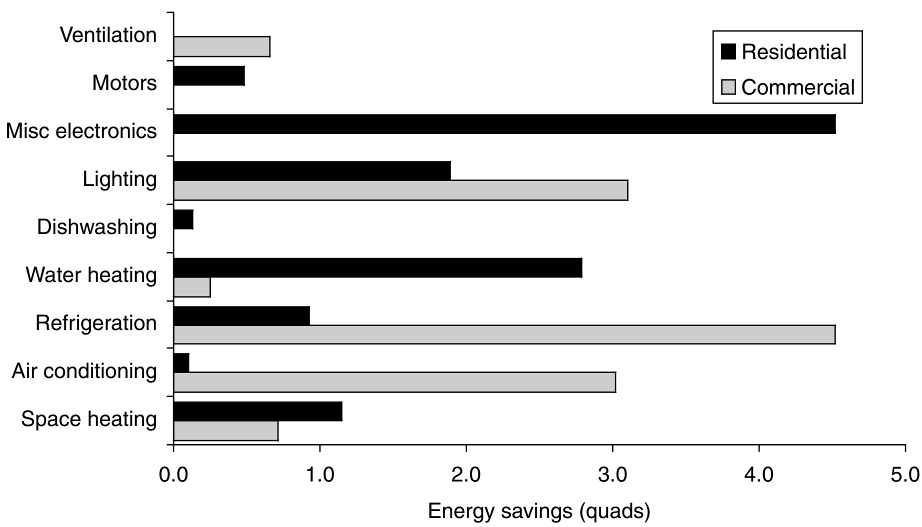


FIGURE 12.3 Cumulative primary energy savings from upgraded standards for products installed in 2010–2030 period. (From Ending the Energy Stalemate: A Bipartisan Strategy to Meet America's Energy Challenges, The National Commission On Energy Policy, December 2004, www.energycommission.org).

Competition for capital to make energy-efficiency investments. Energy-efficiency investments in the commercial and industrial sectors often must compete with other business investments; therefore, efficiency investments with a payback of more than three years are avoided.

The separation of building ownership from utility bill responsibility. Renters will rarely make energy-efficiency investments in buildings that they do not own, especially when the utilities are included in the rent.

Commercial buildings and retail space are usually built on speculation with low first-cost a priority. The building's long-term operation cost, which is usually paid by the tenant(s) rather than the owner, is not important to the speculator/builder.

Overcoming these barriers may require legislation on the state and/or federal level. This type of legislation is likely to be more successful if it considers criteria such as fiscal fairness, probability of success, and ease of implementation in addition to the economic benefits that will eventually ensue. This chapter presents energy efficient technologies for electric motors, lighting, home appliances, and space conditioning which offer the potential for substantial energy conservation in the near future according to experts to their respective fields.

References

1. Rosenquist, G., McNeil, M., Lyer, M., Meyers, S., and McMahon, J. 2004. *Energy Efficiency Standards for Residential and Commercial Equipment: Additional Opportunities*, Lawrence Berkeley National Laboratory.
2. Ending the Energy Stalemate: A Bipartisan Strategy to Meet America's Energy Challenges, The National Commission on Energy Policy, December 2004, www.energycommission.org
3. Kreith, F. and Burmeister, G. 1993. *Energy Management and Conservation*. National Conference of State Legislatures, Denver, CO.

12.1 Electric Motor Systems Efficiency

Anibal T. de Almeida and Steve Greenberg

12.1.1 Introduction

Motor systems are by far the most important type of electrical load, ranging from small fractional hp motors incorporated in home appliances to multi-megawatt motors driving pumps and fans in power plants. Motors consume over half of the total electricity, and in industry they are responsible for about two-thirds of the electricity consumption. In the commercial and residential sectors motors consume slightly less than half of the electricity. The cost of powering motors is immense; roughly $100 billion a year in the U.S. alone. There is a vast potential for saving energy and money by increasing the efficiency of motors and motor systems.

12.1.1.1 Motor Types

Motors produce useful work by causing the shaft to rotate. Motors have a rotating part, the rotor, and a stationary part, the stator. Both parts produce magnetic fields, either through windings excited by electric currents or through the use of permanent magnets. It is the interaction between these two magnetic fields which is responsible for the torque generation, as shown in Figure 12.4.

There are a wide variety of electric motors, based on the type of power supply (AC or DC) that feeds the windings, as well as on different methods and technologies to generate the magnetic fields in the rotor and in the stator. Figure 12.5 presents the most important types of motors.

Because of their low cost, high reliability and fairly high efficiency, most of the motors used in large home appliances, industry, and commercial buildings are induction motors. Figure 12.6 shows the operating principle of a three-phase induction motor.

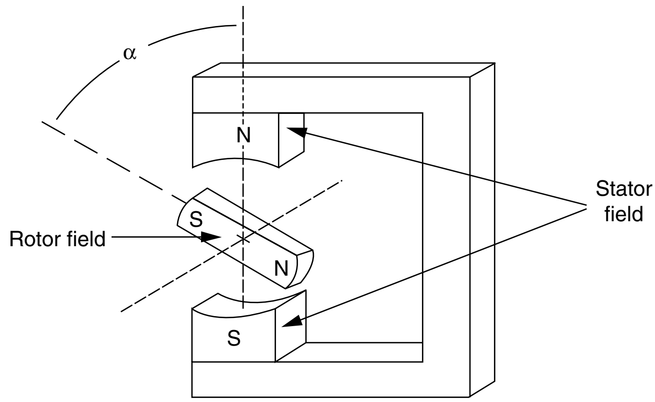


FIGURE 12.4 Torque generation in a motor. The generated torque is proportional to the strength of each magnetic field and depends on the angle α between the two fields. Mathematically, torque equals $|\mathbf{B}_{\text{rotor}}| \times |\mathbf{B}_{\text{stator}}| \times \sin \alpha$, where **B** refers to a magnetic field. (Reprinted from *Energy-Efficient Motor Systems: A Handbook on Technology, Program and Policy Opportunities*, 1992, American Council for an Energy Efficient Economy, Washington, D.C. With permission.)

Synchronous motors are used in applications requiring constant speed, high operating efficiency and controllable power factor. Efficiency and power factor are particularly important above 1000 hp. Although DC motors are easy to control, both in terms of speed and torque, they are expensive to produce and have modest reliability. DC motors are used for some industrial and electric traction applications, but their importance is dwindling.

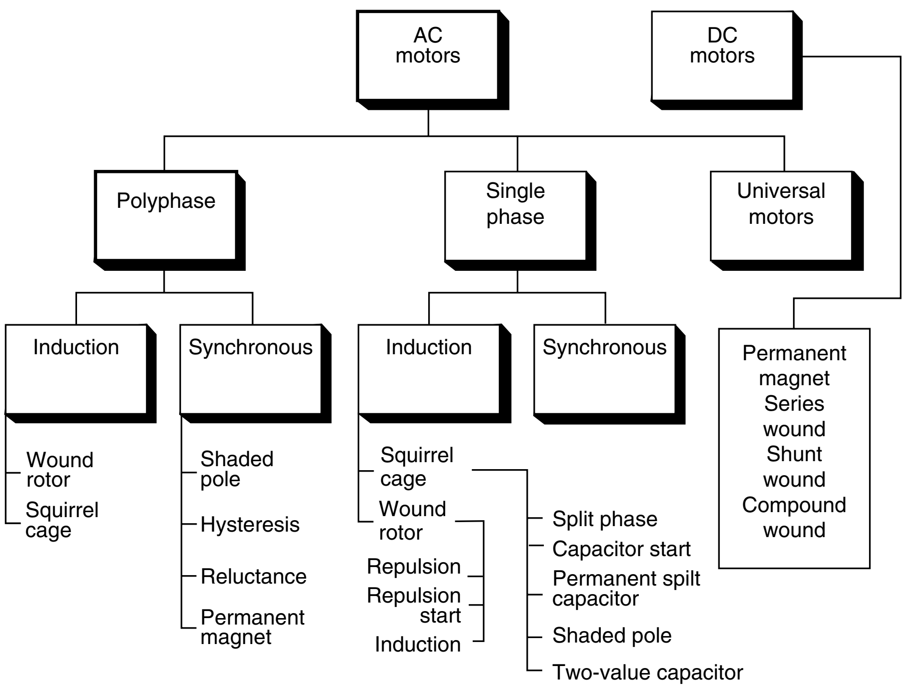


FIGURE 12.5 Motor types. (From EPRI, 1992a.)

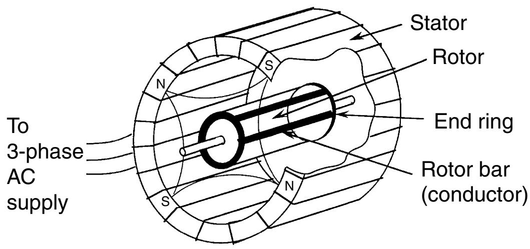


FIGURE 12.6 Operation of a four-pole squirrel-cage induction motor. Rotating magnetic field is created in the stator by AC currents carried in the stator winding. Three-phase voltage source results in the creation of north and south magnetic poles that revolve or "move around" the stator. The changing magnetic field from the stator induces current in the rotor conductors, in turn creating the rotor magnetic field. Magnetic forces in the rotor tend to follow the stator magnetic fields, producing rotary motor action.

12.1.2 Motor Systems Efficiency

The efficiency of a motor-driven process depends upon several factors which may include:

- Motor efficiency
- Motor speed controls
- Proper sizing
- Power supply quality
- Distribution losses
- Transmission
- Maintenance
- Driven equipment (pump, fan, etc.) mechanical efficiency.

It must be emphasized that the design of the process itself influences the overall efficiency (units produced/kWh or service produced/kWh) to a large extent. In fact, in many systems the largest opportunity for increased efficiency is in improved use of the mechanical energy (usually in the form of fluids or solid materials in motion) in the process. Comprehensive programs to address motor-system energy use start with the process and work back toward the power line, optimizing each element in turn, as well as the overall system. Outlining such a program is beyond the scope of this discussion; see (e.g.) Baldwin 1989 for an example of the benefits that propagate all the way back to the power plant.

12.1.2.1 Motor Efficiency

Figure 12.7 shows the distribution of the losses of an induction motor as a function of the load. At low loads the core magnetic losses (hysteresis and eddy currents) are dominant, whereas at higher loads the copper resistive ("Joule" or I^2R) losses are the most important. Mechanical losses are also present in the form of friction in the bearings and windage.

12.1.2.1.1 Energy-Efficient and Premium-Efficiency Motors

After World War II and until the early 1970s, there was a trend to design inefficient motors which minimized the use of raw materials (copper, aluminum, and silicon steel). These induction motors had lower initial costs and were more compact than previous generations of motors, but their running costs were higher. When electricity prices started escalating rapidly in the mid-1970s, most of the large motor manufacturers added a line of higher-efficiency motors to their selection. Such motors feature optimized design, more generous electrical and magnetic circuits and higher quality materials (Baldwin 1989).

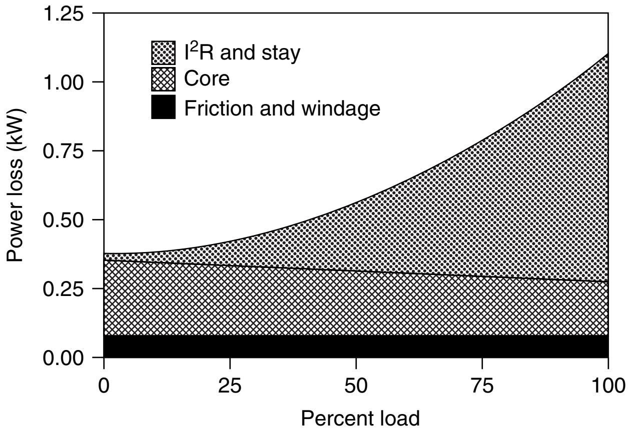


FIGURE 12.7 Variation of losses with load for a 10 hp motor. (Reprinted from *Energy-Efficient Motor Systems: A Handbook on Technology, Program and Policy Opportunities*, 1992, American Council for an Energy Efficient Economy, Washington, D.C. With permission.)

Incremental efficiency improvements are still possible with the use of superior materials (e.g., amorphous silicon steel) and optimized computer-aided design techniques.

In 1997, the U.S. Energy Policy Act (EPAct) put in place mandatory efficiency standards for many general-purpose motors. Also in the 1990s, the Consortium for Energy Efficiency (CEE) developed a voluntary premium-efficiency standard, which evolved into the NEMA Premium designation (NEMA 2000; Nadel et al. 2002). Premium efficiency motors offer an efficiency improvement over EPAct which typically ranges from 4% for a 1 hp motor, to 2% for a 150 hp motor, as shown in Figure 12.8. Due to long motor lives, many motors in use are less efficient than EPAct, and thus there is an even larger

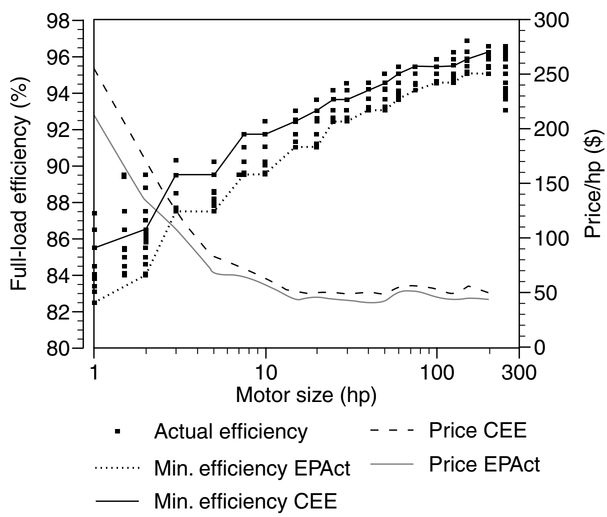


FIGURE 12.8 Ranges for full-load efficiency vs. size, and costs (average per hp trade prices) vs. size for NEMA Design B standard and high-efficiency, 1,800 rpm, three-phase induction motors. Distribution of efficiency data points reflects variation among manufacturers. EPAct and CEE premium-efficiency and minimum-efficiency values are provided for reference. Price per hp values is based on average price for qualifying product in the *MotorMaster* database. (Reprinted from *Energy-Efficient Motor Systems: A Handbook on Technology, Program and Policy Opportunities*, 2002, American Council for an Energy-Efficient Economy, Washington, D.C. With permission.)

difference between them and premium efficiency motors. Premium efficiency motors normally cost around 15%–25% more than standard motors, which translates into a price premium of $8-40/hp. In new applications, and for motors with a large number of operating hours, the paybacks are normally under 4 years for Premium vs. EPAct motors, and under 2 years for Premium vs. older, standard-efficiency motors.

12.1.2.1.2 Efficiency of Rewound Motors

When a motor fails, the user has the options of having the motor rebuilt, buying a new EPAct motor, or buying a NEMA Premium motor. Except for large motors with low annual operating hours, it is typically very cost-effective to replace the failed motor with a Premium motor. Although motor rebuilding is a low-cost alternative, the efficiency of a rebuilt motor can be substantially decreased by the use of improper methods for stripping the old winding. On average, the efficiency of a motor decreases by about 1% each time the motor is rewound.

The use of high temperatures (above 350°C) can damage the interlaminar insulation and distort the magnetic circuit with particular impact on the air gap shape, leading to substantially higher core and stray losses. Before any motor is rewound, it should be checked for mechanical damage and the condition of the magnetic circuit should be tested with an electronic iron core loss meter. There are techniques available to remove the old windings, even the ones coated with epoxy varnish, which do not exceed 350°C (Dreisilker 1987).

12.1.2.2 Recent Motor Developments

In the low horsepower range, the induction motor is being challenged by new developments in motor technology such as permanent magnet and reluctance motors which are as durable as induction motors and have higher efficiency. These advanced motors do not have losses in the rotor and feature higher torque and power/weight ratio. In fractional hp motors, such as the ones used in home appliances, the efficiency improvements can reach 10%–15%, compared with single-phase induction motors. Compared to the shaded-pole motors commonly used in small fans, improved motor types can more than double motor efficiency.

12.1.2.2.1 Permanent-Magnet Motors

Over the last few decades, there has been substantial progress in the area of permanent-magnet materials. Figure 12.9 shows the relative performance of several families of magnetic materials. High-performance permanent-magnet materials such as neodymium–iron–boron alloys, with a large energy density and moderate cost, offer the possibility of achieving high efficiency and compact light-weight motors.

In modern designs, the permanent magnets are used in the rotor. The currents in the stator windings are switched by semiconductor power devices based on the position of the rotor, normally detected by Hall sensors, as shown in Figure 12.10. The rotor rotates in synchronism with the rotating magnetic field created by the stator coils, leading to the possibility of accurate speed control. Because these motors have no brushes, and with suitable control circuits they can be fed from a DC supply, they are sometimes called brushless DC motors.

12.1.2.2.2 Switched Reluctance Motors

Switched reluctance motors are also synchronous motors whose stator windings are commutated by semiconductor power switches to create a rotating field. The rotor has no windings, being made of iron with salient poles. The rotor poles are magnetized by the influence of the stator rotating field. The attraction between the magnetized poles and the rotating field creates a torque that keeps the rotor moving at synchronous speed. Figure 12.11 shows the structure of a switched reluctance motor.

Switched reluctance motors have higher efficiency than induction motors, are simple to build, robust and, if mass-produced, their price can compete with induction motors. Switched reluctance motors can also be used in high-speed applications (above the 3600 rpm possible with induction or synchronous motors operating on a 60-Hz AC supply) without the need for gears.

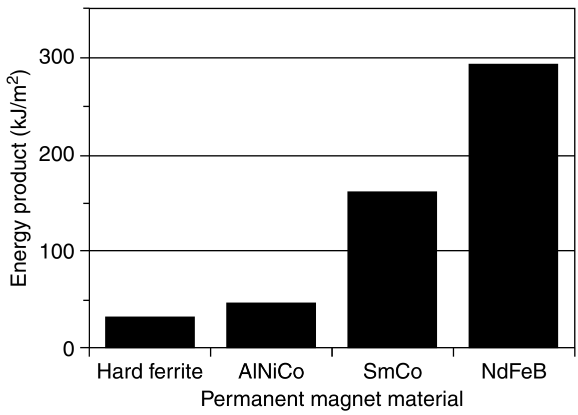


FIGURE 12.9 The evolution of permanent-magnet materials, showing the increasing magnetic energy density ("energy product"). Ferrites were developed in the 1940s; AlNiCos (aluminum, nickel, and cobalt) in the 1930s. The rare-earth magnets were developed beginning in the 1960s (samarium–cobalt) and in the 1980s (neodymium–iron–boron). The higher the energy density, the more compact the motor design can be for a given power rating. (Reprinted from *Energy-Efficient Motor Systems: A Handbook on Technology, Program and Policy Opportunities*, 1992, American Council for an Energy Efficient Economy, Washington, D.C. With permission.)

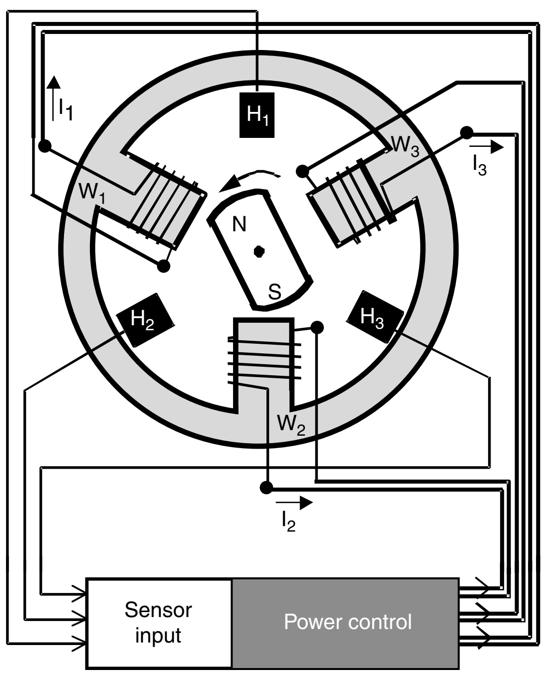


FIGURE 12.10 Control system scheme of a brushless DC motor. This motor is also known as an electronically commutated permanent-magnet (PM) motor. The motor is composed of three sets of stator windings arranged around the permanent-magnet rotor. AC power is first converted to DC, and then switched to the windings by the power control unit, which responds to both an external speed command and rotor position feedback from H_1, H_2, and H_3, which are magnetic position sensors. If a DC power supply is available, it can be used directly by the power control unit in place of the AC supply and converter. The function of the commutator and brushes in the conventional DC motor is replaced by the control unit and power switches. The PM rotor follows the rotating magnetic field created by the stator windings. The speed of the motor is easily changed by varying the frequency of switching.

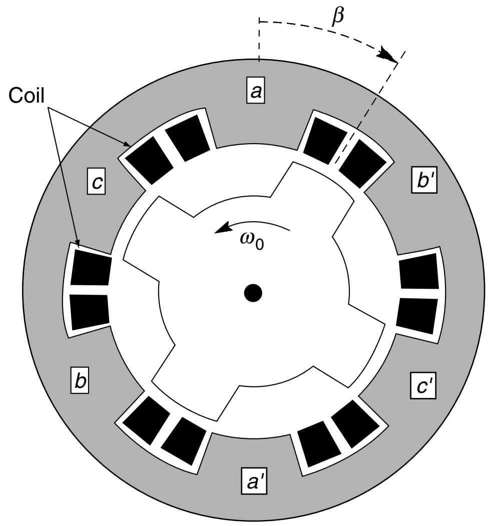


FIGURE 12.11 Schematic view of a switched reluctance motor. The configuration shown is a 6/4 pole. A rotating magnetic field is produced by switching power on and off to the stator coils in sequence, thus magnetizing poles a–a′, b–b′, and c–c′ in sequence. Switching times are controlled by microprocessors with custom programming.

12.1.2.3 Motor Speed Controls

AC induction and synchronous motors are essentially constant-speed motors. Most motor applications would benefit if the speed could be adjusted to the process requirements. This is especially true for new applications where the processes can be designed to take advantage of the variable speed. The potential benefits of speed variation include increased productivity and product quality, less wear in the mechanical components, and substantial energy savings.

In many pump, fan, and compressor applications, the mechanical power grows roughly with the cube of the fluid flow; to move 80% of the nominal flow only half of the power is required. Fluid-flow applications are therefore excellent candidates for motor speed control.

Conventional methods of flow control have used inefficient throttling devices such as valves, dampers and vanes. These devices have a low initial cost, but introduce high running costs due to their inefficiency. Figure 12.12 shows the relative performance of different techniques to control flow produced by a fan.

Motor system operation can be improved through the use of several speed-control technologies, such as those covered in the following three sections.

12.1.2.3.1 Mechanical and Eddy-Current Drives

Mechanical speed control technologies include hydraulic transmissions, adjustable sheaves, and gearboxes. Eddy-current drives work as induction clutches with controlled slip (Magnusson 1984).

Both mechanical drives and eddy drives have relatively low importance. They suffer from low efficiency, bulkiness, limited flexibility, or limited reliability when compared with other alternatives; in the case of mechanical drives they may require regular maintenance.

Mechanical and eddy drives are not normally used as a retrofit due to their space requirements. Their use is more and more restricted to the low horsepower range where their use may be acceptable due to the possible higher cost of ASDs.

12.1.2.3.2 Multi-Speed Motors

In applications where only a few operating speeds are required, multi-speed motors may provide the most cost-effective solution. These motors are available with a variety of torque-speed characteristics

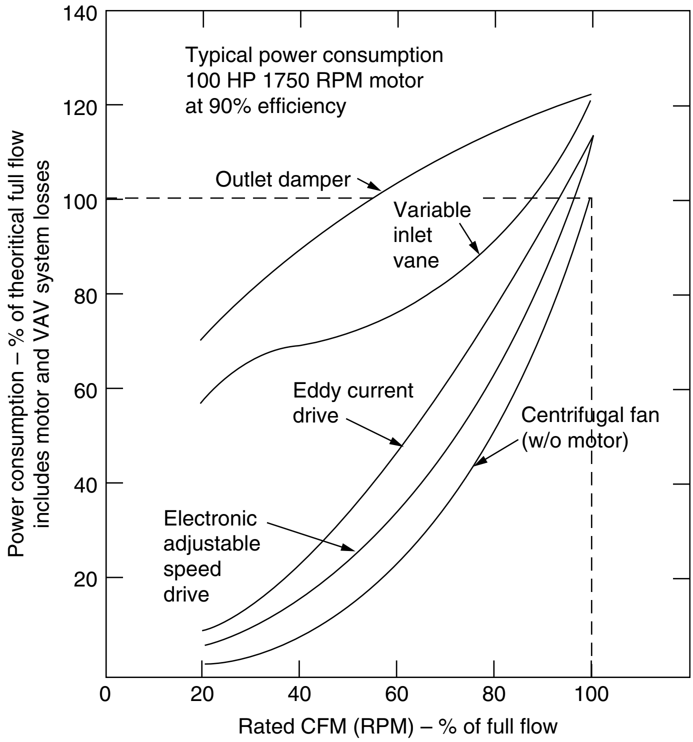


FIGURE 12.12 Comparison of several techniques for varying air flow in a variable-air-volume (VAV) ventilation system. The curve on the lower right represents the power required by the fan itself, not including motor losses. Electronic ASDs are the most efficient VAV control option, offering large savings compared to outlet dampers or inlet vanes, except at high fractions of the rated fan speed. (From Greenberg et al., 1988.)

(variable torque, constant torque, and constant horsepower) (Andreas 1992), to match different types of loads. Two-winding motors can provide up to four speeds but they are normally bulkier (one frame size larger) than single-speed motors for the same horsepower rating. Pole-amplitude modulated (PAM) motors are single winding, two speed, squirrel cage induction motors that provide a wide range of speed ratios (Pastor 1986). Because they use a single winding they have the same frame size of single-speed motors for the same horsepower, and are thus easy to install as a retrofit. PAM motors are available with a broad choice of speed combinations (even ratios close to unity), being especially suited and cost-effective for those fan and pump applications which can be met by a two-speed duty cycle.

12.1.2.3.3 Electronic Adjustable-Speed Drives

Induction motors operate with a torque-speed relation as shown in Figure 12.13. The speed of the motor is very nearly proportional to the frequency of the AC power supplied to it; thus the speed can be varied by applying a variable-frequency input to the motor. Electronic adjustable-speed drives (ASDs) (Bose 1986) achieve this motor input by converting the fixed frequency power supply (50 or 60 Hz), normally first to a DC supply and then to a continuously variable frequency/variable voltage (Figure 12.14). ASDs are thus able to continuously change the speed of AC motors. Electronic ASDs have no moving parts (sometimes with the exception of a cooling fan), presenting high reliability and efficiency and low maintenance requirements. Because ASDs are not bulky and have flexible positioning requirements, they are generally easy to retrofit.

Electronic ASDs are the dominant motor speed control technology at the present and for the foreseeable future. Developments in the past two decades in the areas of microelectronics and power

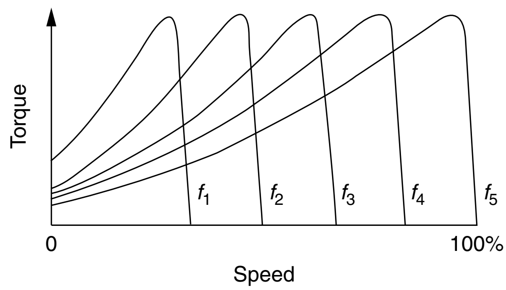


FIGURE 12.13 Speed-torque curves for an induction motor ($f_1 < f_2 < f_3 < f_4 < f_5$, and f_5 = normal line frequency). Normal operation of the motor is in the nearly vertical part of the curves to the right of the "knee" (known as the "breakdown" or "pullout" torque).

electronics make possible the design of efficient, compact and increasingly cost-competitive electronic ASDs. As ASDs control the currents/voltages fed to the motor through power semiconductor switches, it is possible to incorporate motor protection features, soft-start and remote control, at a modest cost. By adding additional power switches and controlling circuitry (Figure 12.15), ASDs can provide regenerative braking, slowing the driven load and feeding power back into the AC supply. Such regenerative braking capability can increase the energy efficiency of such applications as elevators, downhill conveyors, and electric transportation.

Across the range of motor applications, no single ASD technology emerges as a clear winner when compared with other ASD types. Pulse-width modulation (PWM) voltage-source inverters ASDs dominate in the low to medium horsepower range (up to several hundred horsepower) due to their lower cost and good overall performance. Figure 12.16 shows how the variable-frequency/variable-voltage waveform is synthesized by a PWM ASD.

In the range above several hundred horsepower the choice of ASD technology depends on several factors including the type of motor, horsepower, speed range and control requirements (Greenberg et al. 1988). Table 12.1 presents a general classification of the most widely used adjustable-speed motor drive technologies.

12.1.2.4 Motor System Oversizing

Motor systems are often oversized as a result of the compounding of successive safety factors in the design of a system (Smeaton 1988). The magnetic losses, friction, and windage losses are practically constant as a function of the load. Therefore, motors which are oversized (working all the time below 50% of capacity) present not only lower efficiency but also poor power factor (NEMA 1999), as shown in Figure 12.17. The efficiency drops significantly when a motor operates lightly loaded (below 40% for a standard motor). The power factor drops continuously from full load. The decrease in performance is especially noticeable in small motors and standard-efficiency motors.

It is therefore essential to size new motors correctly and to identify motors which run grossly underloaded all the time. In the last case, the economics of replacement by a correctly sized motor should

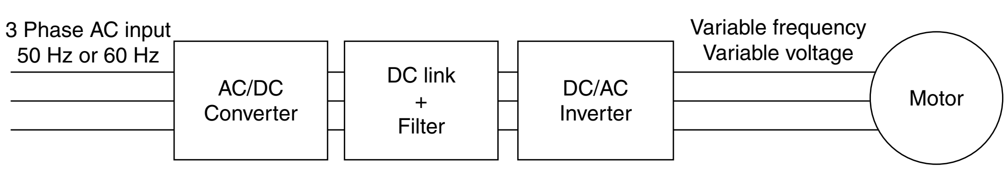


FIGURE 12.14 General inverter based ASD power circuit with motor load.

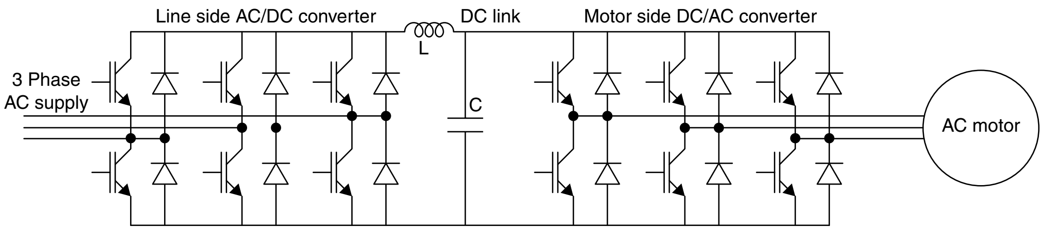


FIGURE 12.15 Power circuitry of a PWM variable speed drive with regenerative capacity and power factor control. Whereas conventional ASDs use diode rectifiers in the input stage, regenerative units use insulated gate bipolar transistors (IGBTs) at both the input and output stages to enable bidirectional power flow.

be considered. In medium or large industrial plants, where a stock of motors is normally available, oversized motors may be exchanged for the correct size versions.

12.1.2.5 Power Quality

Electric motors, and in particular induction motors, are designed to operate with optimal performance when fed by symmetrical three-phase sinusoidal waveforms with the nominal voltage value. Deviations from these ideal conditions may cause significant deterioration of the motor efficiency and lifetime. Possible power quality problems include voltage unbalance, undervoltage or overvoltage, and harmonics and interference. Harmonics and interference can be caused by, as well as affect, motor systems.

12.1.2.5.1 Voltage Unbalance

Induction motors are designed to operate at their best with three-phase balanced sinusoidal voltages. When the three-phase voltages are not equal, the losses increase substantially. Phase unbalance is normally caused by an unequal distribution of the single-phase loads (such as lighting) on the three phases or by faulty conditions. An unbalanced supply can be mathematically represented by two balanced systems rotating in opposite directions. The system rotating in the opposite direction to the motor induces currents in the rotor which heat the motor and decrease the torque. Even a modest phase unbalance of 2% can increase the losses by 25% (Cummings, Dunki-Jacobs, and Kerr 1985).

When a phase unbalance is present, the motor must be derated according to Figure 12.18.

12.1.2.5.2 Voltage Level

When an induction motor is operated above or below its rated voltage, its efficiency and power factor change. If the motor is underloaded, a voltage reduction may be beneficial, but for a properly sized motor the best overall performance is achieved at the rated voltage. The voltage fluctuations are normally

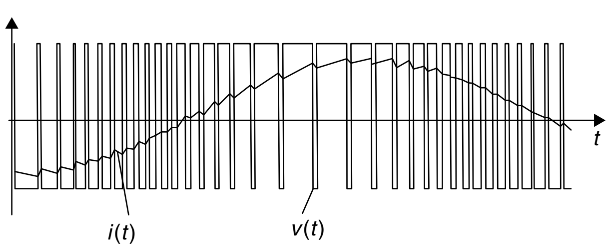


FIGURE 12.16 Pulse-width modulation for synthesizing a sinusoidal output. Output voltage, the average of which resembles the current waveform $i(t)$, is varied by changing the width of the voltage pulses $v(t)$. Output frequency is varied by changing the length of the cycle.

TABLE 12.1 Adjustable-Speed Motor Drive Technologies

Technology	Applicability (R= Retrofit; N= New)	Cost[b]	Comments
		Motors	
Multispeed (incl PAM[a]) motors	Fractional-500 hp PAM; fractional-2,000 + hp R,N	1.5 to 2 times the price of single-speed motors	Larger and less efficient than 1-speed motors. PAM is more promising than multiwinding. Limited number of available speeds.
Direct-current motors	Fractional-10,000 hp N	Higher than AC induction motors	Easy speed control. More maintenance required.
	Shaft-applied drives (on motor output)		
Mechanical			
Variable-ratio belts	5–125 hp N	\$350–\$50/hp (for 5–125 hp)	High efficiency at part load 3:1 speed range limitation. Requires good maintenance for long life.
Friction dry disks	Up to 5 hp N	\$500–\$300/hp	10:1 speed range. Maintenance required.
Hydraulic drive	5–10,000 hp N	Large variation	5:1 speed range. Low efficiency below 50% speed.
Eddy-current drive	Fractional ~ 2000+ hp N	\$900–\$60/hp (for 1 to 150 hp)	Reliable in clean areas, Relatively long life. Low efficiency below 50% speed.
	Wiring-applied drives (on motor input)		
Electronic adjustable speed drives			
Voltage-source inverter	Fractional-1500 hp R,N	\$300–\$100/hp (for 1 to 300 hp)	Multimotor capability. Can generally use existing motor. PWM[c] appears most promising.
Current-source inverter	100–100,000 hp R,N	\$120–\$50/hp (for 100 to 20,000 hp)	Larger and heavier than VSL Industrial applications, including large synchronous motors.
Others	Fractional-100,000 hp R,N	Large variation	Includes cycloconverters, wound rotor, and variable voltage. Generally for special industrial applications.

[a] PAM means Pole Amplitude Modulated.
[b] The prices are listed from, high to low to correspond with the power rating, which is listed from low to high. Thus, the lower the power rating, the higher the cost per horsepower.
[c] PWM means Pulse Width Modulation.

associated with ohmic (IR) voltage drops or with reactive power (poor power factor) flow in the distribution network (see Section 12.1.2.6).

12.1.2.5.3 Harmonics and Electromagnetic Interference

When harmonics are present in the motor supply, they heat the motor and do not produce useful torque. This in turn affects the motor lifetime and causes a derating of the motor capacity. This is also true

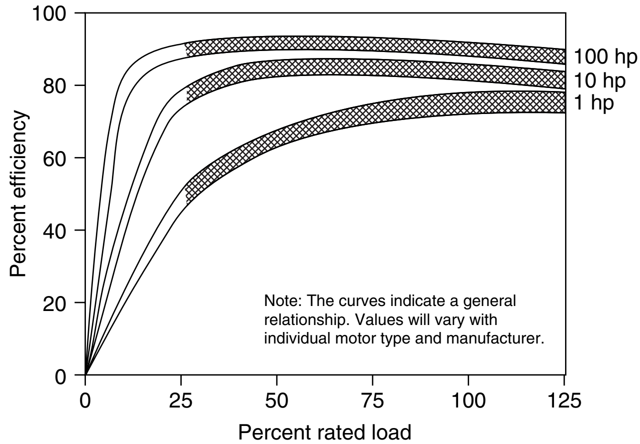


FIGURE 12.17 Typical efficiency vs. load curves for 1800 rpm, three-phase 60 Hz Design B squirrel cage induction motors. (Reprinted from *Energy-Efficient Motor Systems: A Handbook on Technology, Program and Policy Opportunities,* 2002, American Council for an Energy Efficient Economy, Washington, D.C. With permission.)

when motors are supplied by ASDs which generate the harmonics themselves. The use of premium-efficiency motors can alleviate these problems due to their higher efficiency and thermal capacity; there are also motors specially designed for use with ASDs known as inverter-duty motors.

Reduction of harmonics is also important for the benefit of other consumer and utility equipment. Harmonics, caused by nonlinear loads such as the semiconductor switches in ASDs, should be reduced to an acceptable level as close as possible to the source. The most common technique uses inductive/capacitive filters at the ASD input circuit to provide a shunt path for the harmonics and to perform power factor compensation.

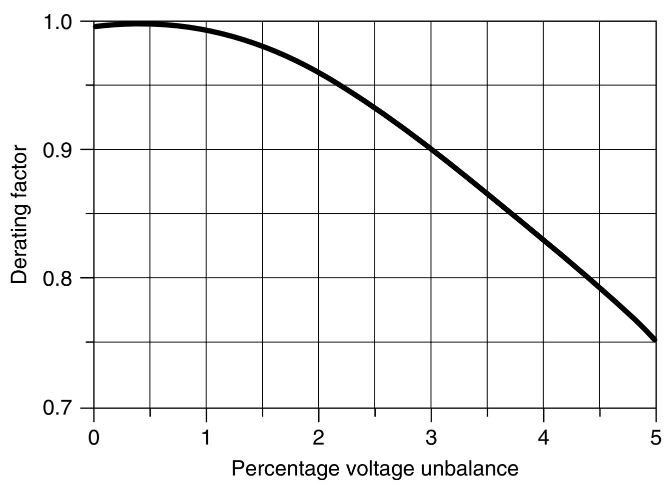


FIGURE 12.18 Derating factor due to unbalanced voltage for integral-horsepower motors. (Reprinted from *Energy-Efficient Motor Systems: A Handbook on Technology, Program and Policy Opportunities,* 1992, American Council for an Energy Efficient Economy, Washington, D.C. With permission.)

IEEE Standard No. 519 (IEEE 1992) contains guidelines for harmonic control and reactive power compensation of power converters. The cost of the harmonic filter to meet this standard is typically around 5% of the cost of the ASD.

ASD power semiconductor switches operate with fast switching speeds to decrease energy losses. The fast transitions in the waveforms contain high-frequency harmonics, including those in the radio-frequency range. These high-frequency components can produce interference through both conduction and radiation. The best way to deal with EMI is to suppress it at the source. Radiated EMI is suppressed through shielding and grounding of the ASD enclosure. Proper ASD design, the use of a dedicated feeder and the use of a low-pass input filter (an inductor; often called a "line reactor"), will normally suppress conducted EMI.

12.1.2.6 Distribution Losses

12.1.2.6.1 Cable Sizing

The currents supplied to the motors in any given installation will produce Joule (I^2R) losses in the distribution cables and transformers of the consumer. Correct sizing of the cables will not only allow a cost-effective minimization of those losses, but also helps to decrease the voltage drop between the transformer and the motor. The use of the National Electrical Code for sizing conductors leads to cable sizes that prevent overheating and allow adequate starting current to the motors, but can be far from an energy-efficient design. For example, when feeding a 100 hp motor located at 150 m from the transformer with a cable sized using NEC, about 4% of the power will be lost in heating the cable (Howe et al. 1999). Considering a 2-year payback, it is normally economical to use a cable one wire size larger than the one required by the NEC.

12.1.2.6.2 Reactive Power Compensation

In most industrial consumers, the main reason for a poor power factor is the widespread application of oversized motors. Correcting oversizing can thus contribute in many cases to a significant improvement of the power factor.

Reactive power compensation, through the application of correction capacitors, not only reduces the losses in the network but also allows full use of the power capacity of the power system components (cables, transformers, circuit breakers, etc.). In addition, voltage fluctuations are reduced, thus helping the motor to operate closer to its design voltage.

12.1.2.7 Mechanical Transmissions

The transmission subsystem transfers the mechanical power from the motor to the motor-driven equipment. To achieve overall high efficiency it is necessary to use simple, properly maintained transmissions with low losses. The choice of transmission is dependent upon many factors including speed ratio desired, horsepower, layout of the shafts, type of mechanical load, etc.

Transmission types available include direct shaft couplings, gearboxes, chains, and belts. Belt transmissions offer significant potential for savings. About one-third of motor transmissions use belts (Howe et al. 1999). Several types of belts can be used such as V-belts, cogged V-belts and synchronous belts.

V-belts have efficiencies in the 90%–96% range. V-belt losses are associated with flexing, slippage, and a small percentage due to windage. With wear, the V-belt stretches and needs retensioning, otherwise the slippage increases and the efficiency drops. Cogged V-belts have lower flexing losses and have better gripping on the pulleys, leading to 2%–3% efficiency improvement when compared with standard V-belts.

Synchronous belts can be 98%–99% efficient as they have no slippage and have low flexing losses; they typically last over twice as long as V-belts, leading to savings in avoided replacements which more than offset their extra cost. Figure 12.19 shows the relative performance of V-belts and synchronous belts. The efficiency gains increase with light loads.

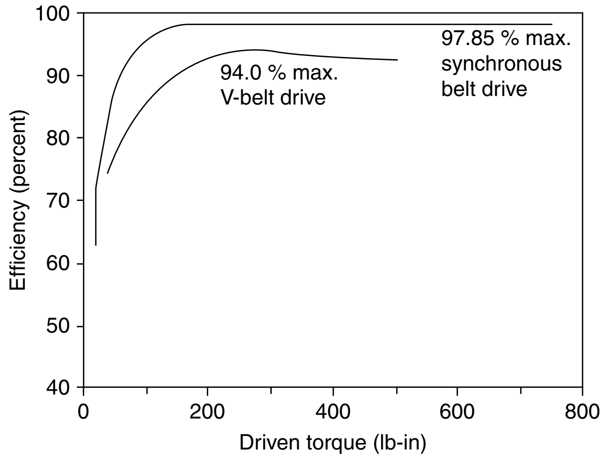


FIGURE 12.19 Efficiency versus torque for V-belts and synchronous belts in a typical application. (Reprinted from *Energy-Efficient Motor Systems: A Handbook on Technology, Program and Policy Opportunities*, 1992, American Council for an Energy Efficient Economy, Washington, D.C. With permission.)

12.1.2.8 Maintenance

Regular maintenance (such as inspection, adjustment, cleaning, filter replacement, lubrication, and tool sharpening) is essential to maintain peak performance of the mechanical parts and to extend their operating lifetime. Both under- and over-lubrication can cause higher friction losses in the bearings and shorten the bearing lifetime. Additionally, overgreasing can cause the accumulation of grease and dirt on the motor windings, leading to overheating and premature failure.

The mechanical efficiency of the driven equipment (pump, fan, cutter, etc.) directly affects the overall system efficiency. Monitoring wear and erosion in this equipment is especially important as its efficiency can be dramatically affected. For example, in chemical process industries the erosion of the pump impeller will cause the pump efficiency to drop sharply; a dull cutter will do the same to a machine tool.

Cleaning the motor casing is also relevant because its operating temperature increases as dust and dirt accumulates on the case. The same can be said about providing a cool environment for the motor. The temperature increase leads to an increase of the windings' resistivity and therefore to larger losses. An increase of 25°C in the motor temperature increases the Joule losses by 10%.

12.1.3 Energy-Saving Applications of ASDs

Typical loads which may benefit from the use of ASDs include those covered in the following four sections.

12.1.3.1 Pumps and Fans

In many pumps and fans where there are variable-flow requirements, substantial savings can be achieved, as the power is roughly proportional to the cube of the flow (and thus speed of the motor). The use of ASDs instead of throttling valves with pumps shows similar behavior to that for fans in Figure 12.12.

12.1.3.2 Centrifugal Compressors and Chillers

Air compressors use 16% of all electricity used to power motor driven processes in U.S. industries (XENERGY 1998). Most industrial compressed air systems have significant savings potential.

Well-engineered efficiency improvements yield verified savings in the range of 15–30 percent of system energy consumption. See Table 12.2 (Fraunhofer Institute 2000).

Centrifugal compressors and chillers can take advantage of motor controls in the same way as other centrifugal loads (pumps and fans). The use of wasteful throttling devices or the on–off cycling of the equipment can be largely avoided, resulting in both energy savings and extended equipment lifetime.

TABLE 12.2 Energy Savings Measures for Compressed-Air Systems

Energy Savings Measure	% Applicability[a]	% Gains[b]	Potential Contribution[c] (%)	Comments
		System Installation or Renewal		
Improvement of drives (high efficiency motors)	25	2	0.5	Most cost-effective in small (< 10 kW) systems
Improvement of drives: (Speed regulation)	25	15	3.8	Applicable to variable load systems. In multi-machine installations, only one machine should be fitted with a variable speed drive. The estimated gain is for overall improvement of systems, be they mono or multi-machine
Upgrading of compressor	30	7	2.1	
Use of sophisticated control systems	20	12	2.4	
Recovering waste heat for use in other functions	20	20	4.0	Note that the gain is in terms of energy, not of electricity consumption, since electricity is converted to useful heat
Improved cooling, drying and filtering	10	5	0.5	This does not include more frequent filter replacement (see below)
Overall system design, including multi-pressure systems	50	9	4.5	
Reducing frictional pressure losses (for example by increasing pipe diameter)	50	3	1.5	
Optimising certain end use devices	5	40	2.0	
		System Operation and Maintenance		
Reducing air leaks	80	20	16.0	Largest potential gain
More frequent filter replacement	40	2	0.8	
Total			32.9	

[a]% of CAS where this measure is applicable and cost-effective
[b]% reduction in annual energy consumption
[c]Potential contribution = applicability × reduction.

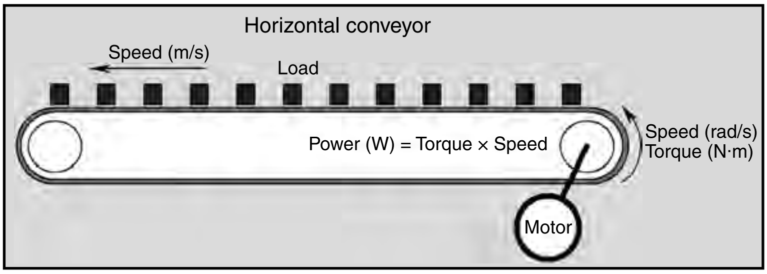


FIGURE 12.20 Power required by a conveyor.

Savings from compressed air measures are mostly coincident with electric system peak periods. Plant air systems normally have a large load factor, typically operating 5000–8000 h per year. Thus energy and demand reductions are very likely to occur at system peaks and contribute to system reliability. Compressed air system efficiency improvements are highly cost-effective and additionally lead to reduced plant downtime. Many projects have identified significant energy and demand reduction projects with paybacks less than 2 years (XENERGY 2001).

12.1.3.3 Conveyors

The use of speed controls, in both horizontal and inclined conveyors, allows the matching of speed to material flow. As the conveyor friction torque is constant, energy savings are obtained when the conveyor is operated at reduced speed. In long conveyors, such as found in power plants and in the mining industry, the benefits of soft-start without the need for complex auxiliary equipment are also significant (De Almeida, Ferreira, and Both 2005).

For horizontal conveyors (Figure 12.20), the torque is approximately independent of the transported load (it is only friction-dependent). Typically, the materials handling output of a conveyor is controlled through the regulation of input quantity, and the torque and speed are roughly constant. But, if the materials input to the conveyor is changed, it is possible to reduce the speed (the torque is the same), and, as it can be seen in Figure 12.21, significant energy savings will be reached, proportional to the speed reduction.

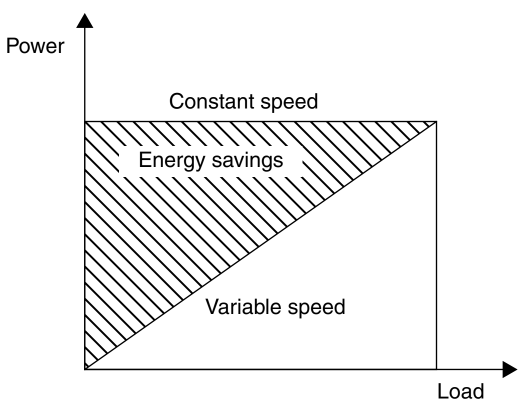


FIGURE 12.21 Energy savings in a conveyor using speed control, in relation to the typical constant speed.

12.1.3.4 High Performance Applications

AC motors have received much attention in recent years as a proposed replacement for DC motors in high-performance speed control applications, where torque and speed must be independently controlled. Induction motors are much more reliable, more compact, more efficient and less expensive than DC motors. As induction motors have no carbon brush commutation, they are especially suitable for corrosive and explosive environments. In the past, induction motors have been difficult to control as they behave as complex nonlinear systems. However, the appearance on the market of powerful and inexpensive microprocessors has made it possible to implement in real time the complex algorithms required for induction motor control.

Field-oriented control, also called vector control, allows accurate control of the speed and torque of induction motors, in a way similar to DC motor control (Leonhard 1984). The motor current and voltage waveforms, together with motor position feedback, are processed in real time, allowing the motor current to be decomposed into a field producing component and into a torque producing component. Vector control operation principle is represented in Figure 12.22 and is being applied to a wide variety of high-performance applications described next.

Rolling mills were one of the strongholds of DC motors, due to the accurate speed and torque requirements. With present ASD technology, AC drives can outperform DC drives in all technical aspects (reliability, torque/speed performance, maximum power, efficiency), and are capable of accurate control down to zero speed.

The availability of large diameter, high torque, and low speed AC drives makes them suitable for use in applications like ball mills and rotary kilns without the need for gearboxes. This area was also a stronghold of DC drives. Again, AC drives have the capability to offer superior performance in terms of reliability, power density, overload capability, efficiency, and dynamic characteristics.

AC traction drives can also feature regenerative braking. AC traction drives are already being used in trains, rapid transit systems, ship propulsion and are the proper choice for the electric automobile.

DC drives have traditionally been used with winders in the paper and steel industry in order to achieve the constant tension requirements as the winding is performed. Sometimes the constant tension is obtained by imposing friction, which wastes energy. AC drives can also be used to replace DC drives, saving energy and achieving better overall performance.

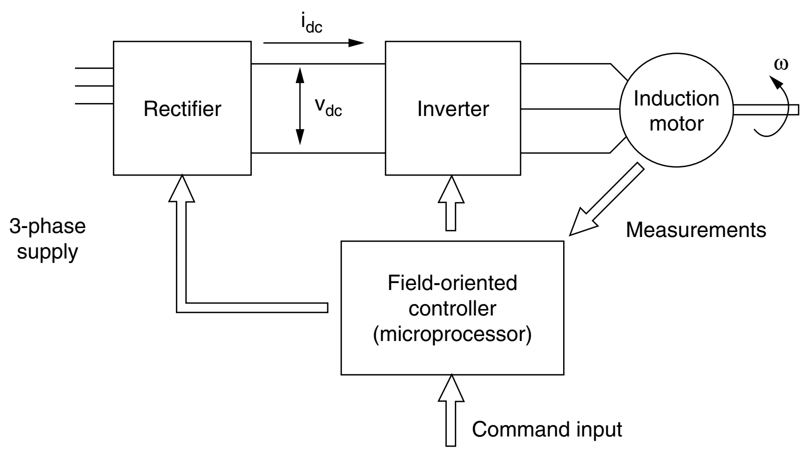


FIGURE 12.22 Schematic of a vector-control drive (also known as a field-oriented control). (Reprinted from *Energy-Efficient Motor Systems: A Handbook on Technology, Program and Policy Opportunities*, 1992, American Council for an Energy Efficient Economy, Washington, D.C. With permission.)

The use of field-oriented AC drives in machine tools and robotics allows the stringent requirements regarding dynamic performance to be met. Positioning drives can produce a peak torque up to 10 times the rated torque and make possible the adjustment of the speed to very low values.

In machine tools, AC drives can provide accurate higher spindle speeds than DC drives without the need for gearboxes and their associated losses. The ASDs can also adjust the voltage level when the spindle drive is lightly loaded, providing further savings. In robotics, the higher power density and superior dynamics of AC drives are important advantages.

12.1.4 Energy and Power Savings Potential; Cost-Effectiveness

The energy and peak power savings potential for any motor-related technology depends on a number of factors, including the characteristics of the motor, the motor drive, the driven equipment, and the load (Nadel et al. 2002). Since all of this information is seldom available, it is difficult to determine the effect of even a single application; it is far more difficult to determine the savings potential for diverse applications, across all sectors, for an entire nation.

This section estimates the energy and power savings that could be realized nationwide in the U.S. through the application of these technologies in the residential, commercial, industrial, and utility sectors. Table 12.3 lists the major motor-driven end uses, estimates of their energy use as a percentage of total national electricity use (based on Howe et al. (1999) and Nadel et al. (2002)) and the potential energy savings from ASDs expressed as a fraction of existing use.

12.1.4.1 Potential Savings in the Residential Sector

About 38% of residential electricity is used in motor systems. The primary motor technology for realizing energy and power savings in the residential sector is the electronic adjustable-speed drive (ASD). Heat pumps, air conditioners, forced-air furnaces and washing machines with ASDs have

TABLE 12.3 Electric Motor Usage and Potential ASD Savings by Sector and End Use

Sector	End Use	Usage (% of Total U.S. Usage)	Potential ASD Savings (% of Usage for Each End Use)
Residential	Refrigeration	3.0	10
	Space heating	2.1	20
	Air conditioning	3.4	15
	HVAC Dist. Fan	1.8	25
	Other	1.6	5
	Residential total	11.9	15
Commercial	Refrigeration	2.2	10
	HVAC compressors	3.3	15
	HVAC distribution	4.2	25
	Other	0.7	20
	Commercial total	10.4	18
Industrial	Refrigeration	1.5	10
	Pumps	5.7	20
	Fans	3.2	20
	Compressed air	3.6	15
	Material handling	2.8	15
	Material process	5.2	15
	Other	1.0	0
	Industrial total	23.1	16
Utilities	Pumps and fans	4.9	15
	Material handling and processing	2.2	15
	Utilities total	7.1	19

already been introduced into the market. Other appliances, such as refrigerators, freezers, heat pump water heaters and evaporative coolers, are also potential candidates for adjustable-speed controls. Most of the energy-saving potential of ASDs in the home is associated with the use of refrigerant compressors for cooling or heating (as in heat pumps, air conditioners, refrigerators, and freezers). In all of these applications, ASDs can reduce energy consumption by matching the speed of the compressor to the instantaneous thermal load. Given the assumed savings potential, the overall savings is about 15% of the sector's motor electricity.

Several improvements in home appliance technology that are likely to become common over the next few years will complement the use of ASDs:

- High-efficiency compressor and fan motors. The use of permanent-magnet and reluctance motors can increase the efficiency of the motor by 5%–15%, when compared with conventional squirrel-cage single-phase induction motors; as noted above, even larger savings are possible with many small fan motors. Permanent-magnet AC motors are used in the latest ASD-equipped furnace and heat pump.
- Rotary and scroll compressors. The use of rotary (in small applications) or scroll compressors in place of reciprocating compressors can take full advantage of the speed variation potential of ASDs.
- Larger heat exchangers with improved design. Improved heat exchangers increase the efficiency by decreasing the temperature difference of the compressor thermal cycle.

12.1.4.2 Potential Savings in the Commercial Sector

An estimated 37% of commercial electricity use is for motor-driven end uses (Howe et al. 1999); the percentage for peak power is higher. The savings potential of ASDs in air-conditioning and ventilation applications was estimated by running DOE-2 computer simulations on two representative building types in five U.S. cities (Eto and de Almeida 1988). Comparisons were made between the cooling and ventilation systems with and without ASDs. The results indicate ventilation savings of approximately 25% for energy and 6% for peak power, and cooling savings of about 15% and 0%, respectively. In Table 12.3 we assumed these energy results can be applied nation-wide.

The estimated 10% energy savings for refrigeration are shown in Table 12.3; an estimated 5% savings in peak power should also be attainable. Other motor efficiency measures (discussed in Section 12.1.2) combined, can capture approximately 10% more energy and demand savings, with an overall potential savings of about 18% of the sector's motor system electricity.

12.1.4.3 Potential Savings in the Industrial and Utility Sectors

About 70% of industrial and 89% of the utility sector electricity use is for motor systems. In Table 12.3, the fluid moving end use savings are estimated at 20%, except for utilities, where the system requirements of municipal water works limit the savings. Compressed air and materials applications are assumed to have 15% potential.

As most industries are nonseasonal, with flat load profiles during operating hours, the peak savings are similar to energy savings. When other motor efficiency measures (see Section 12.1.2) are combined, approximately 10% more energy and demand savings can be obtained, resulting in a total of 16% combined savings for the motor systems in these sectors.

12.1.4.4 Cost-Effectiveness of ASDs

The price of ASD equipment, in terms of dollars/horsepower, is a function of the horsepower range, the type of AC motor used, and the additional control and protection facilities offered by the electronic ASD. ASD installation costs vary tremendously depending on whether the application is new or retrofit, available space, weather protection considerations, labor rates, etc. Thus there is a huge

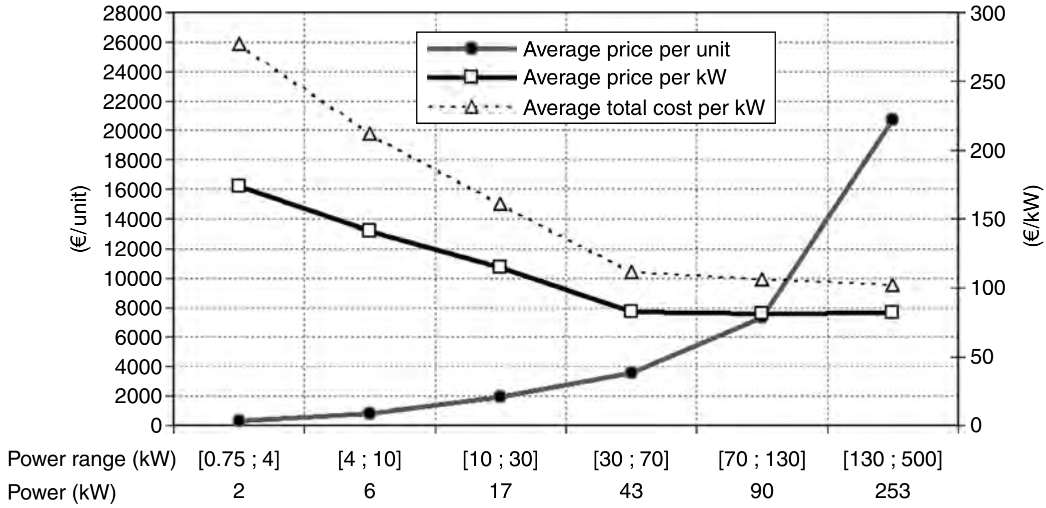


FIGURE 12.23 Average unit costs and average per kW costs, for the different power ranges in the European Union (1 Euro = 1.2 U.S.$, August 2005); 1 kW of power range = 1.3 hp.

range of installed costs possible for any given ASD size, and the costs listed below necessarily have large uncertainties.

A market survey (De Almeida et al. 2001), showed the following typical dollar values for equipment and installation of drives for induction motors (the higher $/hp in each range corresponds to the small end of the hp range):

Installed Cost ($/hp)	Size Range (hp)
300–180	7.5–50
180–120	50–200
120–100	200–1000
100–60	1000–2500
60–50	2500–20,000

These numbers are graphically presented in their original units of Euros and kW in Figure 12.23.

Mass production of smaller motors with built-in ASDs has been very successful in bringing down the cost of Japanese variable-speed heat pumps, with reported incremental costs of $25/hp for the ASD (Abbate 1988).

To determine whether an ASD is cost-effective for any given application, the following need to be taken into account:

- First cost (acquisition and installation)
- System operating load profile (number of hours per year at each level of load)
- Cost of electricity
- Maintenance requirements
- Reliability
- Secondary benefits (less wear on equipment, less operating noise, regeneration capability, improved control, soft-start, and automatic protection features)
- Secondary problems (power factor, harmonics, and interference)

A careful analysis should weigh the value of the benefits offered by each option against the secondary problems, such as power quality, that may impose extra costs for filters and power factor correction capacitors.

Comparing the cost of conserved energy to the cost of electricity is a crude way to assess the cost-effectiveness of energy efficiency measures. More accurate calculations would account for the time at which conservation measures save energy relative to the utility system peak demand, and relate these "load shape characteristics" to baseload, intermediate and peaking supply resources. See Koomey, Rosenfeld, and Gadgil (1990a, 1990b) for more details.

References

1. Abbate, G. 1988. Technology developments in home appliances. In *Demand-Side Management and Electricity End—Use Efficiency,* De Almeidc; A. and Rosenfeld A., eds., Kluwer Academic Publishers, Boston.
2. Andreas, J. 1992. Energy-Efficient Electric Motors: Selection and Application. Marcel Dekker, New York.
3. Arthur D. Little, Inc. 1999. *Opportunities for Energy Savings in the Residential and Commercial Sectors with High-Efficiency Electric Motors, Final Report*. Prepared for the U.S. Department of Energy.
4. Baldwin, S.F. 1989. Energy-efficient electric motor drive systems. In *Electricity: Efficient End- Use and New Generation Technologies, and Their Planning Implications,* Johansson T.B. et al., ed., Lund University Press, Lund, Sweden.
5. Bose, B. 1986. *Power Electronics and AC Drives.* Prentice Hall, Englewood Cliffs, NJ.
6. BPA/EPRI 1993. *Electric Motor Systems Sourcebook.* Bonneville Power Administration, Information Clearinghouse, Olympia, WA.
7. Cummings, P., Dimki-Jacobs, J., and Kerr, R. 1985. Protection of Induction Motors Against Unbalanced Voltage Operation. *IEEE Transactions on Industry Applications,* IA-21 (4), 778–792.
8. De Almeida, A., Ferreira, F., Fonseca, P., Falkner, H., Reichert, J., West, M., Nielsen, S., and Both, D. 2001. *VSDs for Electric Motor Systems*, Institute of Systems and Robotics, University of Coimbra, Portugal, European Commission, Directorate-General for Transport and Energy, SAVE II Programme, 2001.
9. De Almeida, A., Ferreira, F., and Both, D. 2005. Technical and economical considerations in the application of variable speed drives with electric motor systems. *IEEE Industrial Applications Transactions,* 41(1). 188-199.
10. Dreisilker, H. 1987. Modern rewind methods assure better rebuilt motors. *Electrical Construction and Maintenance*, 84(August).
11. EPRI. 1992. *Electric Motors: Markets, Trends and Application*. EPRI Report TR-100423. Electric Power Research Institute, Palo Alto, CA.
12. EPRI 1992b. *Adjustable Speed Drive Directory. 3rd Eel.* Electric Power Research Institute, Palo Alto, CA.
13. EPRI 1993. *Applications of AC Adjustable Speed Drives.* Electric Power Research Institute, Palo Alto, CA.
14. Eto, J. and de Almeida, A. 19SS. Saving electricity in commercial buildings with adjustable speed drives. *IEEE Transactions on Industrial Applications,* 24 (3). 439-443.
15. Fraunhofer Institute. 2000. Compressed Air Systems Market Transformation Study XVII/4.1031/Z/98-266, prepared for the European Commission, Brussels.
16. Greenberg, S., Harris, J. H., Akbari, H., and de Almeida, A. 1988. Technology Assessment: Adjustable Speed Motors and Motor Drives. Lawrence Berkeley Laboratory report LBL-25080, University of California Lawrence Berkeley Laboratory, Berkeley, CA.
17. Hoffman, John. 1992. Building on Success: EPA *Initiatives for Energy Efficiency.* E-Source Members' Forum. E-Source, Snowmass, CO., October 2.

18. Howe, B., Lovins, A., Houghton, D., Shepard, M., and Stickney, B. 1999. *Drivepower Technology Atlas.* E Source, Boulder, CO.
19. IEEE. 1992. *IEEE Guide for Harmonic Control and Reactive Compensation of Static Power Converters*. IEEE Standard 519. Institute of Electrical and Electronics Engineers, New York.
20. Jarc, D. and Schleman, R. 1985. Power line considerations for variable frequency drives. *IEEE Transactions on Industry Applications,* IA-21 (5), 1099–1105.
21. Koomey, J., Rosenfeld, A., and Gadgil, A. 1990. Conservation screening curves to compare efficiency investments to power plants: Applications to commercial sector conservation programs. In *Proceedings of the 1990 ACEEE Summer Study on Energy Efficiency in Buildings*, American Council for an Energy Efficient Economy, Asilomar, CA.
22. Koomey, J., Rosenfeld, A.H., and Gadgil, A.K. 1990b. Conservation screening curves to compare efficiency investments to power plants. *Energy Policy,* 18 (8), 774.
23. Leonhard, W. 1984, *Control of Electrical Drives.* Springer, New York.
24. Lovins, A. and Howe, B. 1992. Switched Reluctance Motor Systems Poised for Rapid Growth. E-Source, Boulder, CO.
25. Magnusson, D. 1984. Energy economics for equipment replacement. *IEEE Transactions on Industry Applications,* IA-20 (2), 402–406.
26. Moore, T. 1988. The advanced heat pump: All the comforts of home…and then some. *EPRI Journal*, March, Electric Power Research Institute, Palo Alto, CA.
27. Nadel, S., Elliot, R.N., Shepard, M., Greenberg, S., Katz, G., and de Almeida, A. 2002. *Energy-Efficient Motor Systems—A Handbook on Technologies, Programs, and Policy Opportunities. 2nd Ed.* American Council for an Energy-Efficient Economy, Washington, DC,. Also 1st Ed., 1992.
28. NEMA. 1999. NEMA Standards Publication MG10-1994 (R1999), *Energy Management Guide for Selection and Use of Polyphase Motors*. National Electrical Manufacturers' Association, Washington, DC.
29. NEMA. 2000. NEMA Standards Publication MG1-1998, Revision 1. *Motors and Generators*. National Electrical Manufacturers' Association, Washington, DC.
30. Pastor, C.E. 1986. *Motor Application Considerations—Single Speed, Multi-Speed, Variable Speed.* Pacific Gas and Electric Company Conference April 30-May 2. Oakland, CA.
31. Smeaton, R. 1988. Motor Application and Maintenance Handbook. McGraw-Hill, New York.
32. Taylor, R.W. 1987. Terminal characteristics of power semiconductor devices. *TechCommentary,* Vol. 1, No. 2, Power Electronics Applications Center, Knoxville.
33. XENERGY Inc. 1998. United Slates Industrial Electric Motor Market Opportunities Assessment. U.S. Department of Energy, Washington DC.
34. XENERGY Inc. 2001. National Market Assessment: Compressed Air System Efficiency Services. U.S. Department of Energy, Washington, DC.

12.2 Energy-Efficient Lighting Technologies and Their Applications in the Commercial and Residential Sectors

Barbara Atkinson, Andrea Denver, James E. McMahon, and Robert Clear

12.2.1 Introduction

Lighting is an important electrical end use in every sector and building type across the United States (U.S.). In 2003, approximately one-third of commercial and almost one-fifth of residential delivered (site) electricity consumption was attributable to lighting, or about one-quarter of all electricity delivered to these two sectors combined [6]. The commercial sector consumes the majority of the electricity used for lighting in the U.S. In 2001, the commercial sector consumed approximately 51% of total lighting electricity, the residential sector consumed 27%, the industrial sector consumed 14%, and outdoor stationary lighting was responsible for 8% [11].

There is great potential for saving electricity, reducing the emission of greenhouse gases associated with electricity production, and reducing consumer energy costs through the use of more efficient lighting technologies as well as advanced lighting design practices and control strategies. New, efficient technologies that enter the market in the future can further reduce energy use and increase financial savings.

In this chapter, we provide an overview of both conventional and newer, more efficient lighting technologies. The discussion includes:

- Design of energy-efficient lighting systems;
- Descriptions, applications, and efficacies of various lighting technologies (including lamps, ballasts, fixtures, and controls);
- Operation of energy-efficient lighting systems;
- Current lighting markets and trends;
- Lighting efficiency standards and incentive programs; and
- Cost-effectiveness of efficient lighting technologies.

12.2.2 Design of Energy-Efficient Lighting Systems

A lighting system is an integral part of a building's architectural design, and interacts with the shape of each room, its furnishings, and the level of natural light. Energy efficiency is an important component of lighting system design; however, lighting designers must also consider economics, productivity, aesthetics, and consumer preference. It is highly important not to compromise lighting *quality* in a new lighting design or energy-efficiency retrofit. To improve the lighting efficiency in a building, a lighting designer must understand the user's lighting needs and tastes, the most efficient technologies available to meet these needs, and the way in which individual lighting components function together as a system.

Efficient, high-quality lighting design includes:

- Attention to task and ambient lighting,
- Effective use of daylighting,
- Effective use of lighting controls, and
- Use of the most cost-effective and efficacious technologies.

Because people require less light in surrounding areas than they do where they perform visual tasks, it is usually both unnecessary and inefficient for an entire space to be lit at a level that is appropriate for visual tasks. For this reason, lighting designers practice task-ambient lighting design. For visual comfort and ease of visual transition between task and ambient spaces, the ambient lighting in a room should be at least one-third as bright as the lighting of the task areas. A common task-ambient lighting strategy is to design the overall lighting system to provide an appropriate ambient level of light and then add task lights (e.g., desk lamps) in areas where people are working.

Effective use of daylighting is also an important component of lighting system design. After decades of overdependence on artificial light, many lighting designers are returning to the use of sunlight to illuminate interior spaces. To make good use of natural light, however, requires more than the simple addition of multiple windows. Light pouring in through windows can create glare and cause other spaces to appear very dark by comparison; in addition, windows that are too large can allow too much heat loss or gain. The challenges to successful daylighting are to admit only as much light as needed, distribute it evenly, and avoid glare. The effective use of daylighting can be greatly enhanced by the overall architectural design of a building. For example, more sunlight is available to a building design that maximizes surface area (e.g., a building that is U-shaped or has an interior courtyard). In addition, skylights, wide windowsills, reflector systems, louvers, blinds, and other innovations can be used to bounce natural light farther into a building. The use of window glazes can limit heat transmission while permitting visible light to pass through a window or skylight.

Efficient lighting design depends on the careful selection of cost-effective and efficient lighting technologies. Lighting control systems are important components of efficient lighting systems. In order to complement other efficiency improvements, lighting designers can use lighting controls to reduce lighting when it is not needed. For example, lighting energy is saved when occupancy sensors turn off the lights after occupants leave a space or daylighting controls dim the fluorescent lamps as the level of natural light in a room increases. Dimming systems can also be used to maintain a constant light level as a system ages, which saves energy when lamps are new. To ensure the persistence of energy savings, lighting designers can install permanent lighting fixtures that are dedicated to efficient lamps. For example, an office retrofit substituting compact fluorescent lamps (CFLs) for screw-in incandescent lamps should utilize hard-wired CFL fixtures, which ensures that the more efficient CFLs will not be replaced with incandescent lamps at a later date.

Lighting design that promotes energy-efficient lighting technologies can also influence the design and energy use of a building's cooling system. Because efficient lighting systems produce less heat, the air conditioning systems installed in new buildings with efficient lighting can have lower capacities. Consequently, less money is spent on air conditioning systems as well as on cooling energy[1] The cost-effectiveness of efficient lighting systems is discussed in Section 12.2.7 below.

12.2.3 Lighting Technologies: Description, Efficacy, Applications

Lighting system components fall into four basic categories:

- Lamps,
- Ballasts,
- Fixtures, and
- Lighting controls.

In this section, we first discuss the properties of light sources and common lighting terms (see also Glossary). We then describe the most common and the efficacious lighting technologies within each category as well as where and how different technologies are used. In addition, we discuss some of the most promising design options for further efficiency improvements in lighting technologies.

12.2.3.1 Properties of Light Sources

Because the purpose of a lamp is to produce light, and not just radiated power, there is no direct measure of lamp efficiency. Instead, a lamp is rated in terms of its *efficacy*, which is the ratio of the amount of light emitted (lumens) to the power (watts) drawn by the lamp. For systems using a ballast, in this chapter we report *system efficacy*, which includes the watts drawn by the lamp and ballast. The unit used to express efficacy is lumens per watt (LPW). The theoretical limit of efficacy is 683 LPW and would be produced by an ideal light source emitting monochromatic radiation with a wavelength of 555 nm. The most efficient white light source in the laboratory provides 275–310 LPW. Of lamps presently in the market, the most efficient practical light source, the T5 fluorescent lamp with electronic ballast, produces about 100 LPW. High-pressure sodium (not a white light source) can produce as high as 130 LPW [9].

The efficacies of various light sources are depicted in Figure 12.24. Lamps also differ in terms of their cost, size, color, lifetime, optical controllability, dimmability, *lumen maintenance*, reliability, convenience in use, maintenance requirements, disposal, environmental impacts (mercury, lead), and electromagnetic and other emissions (e.g., radio interference, ultraviolet (UV) light, and noise).

[1]More energy may be required to heat a building when lighting electricity consumption is reduced but, in many climate zones and building types, cooling savings offset this heating penalty and net cost savings are accrued.

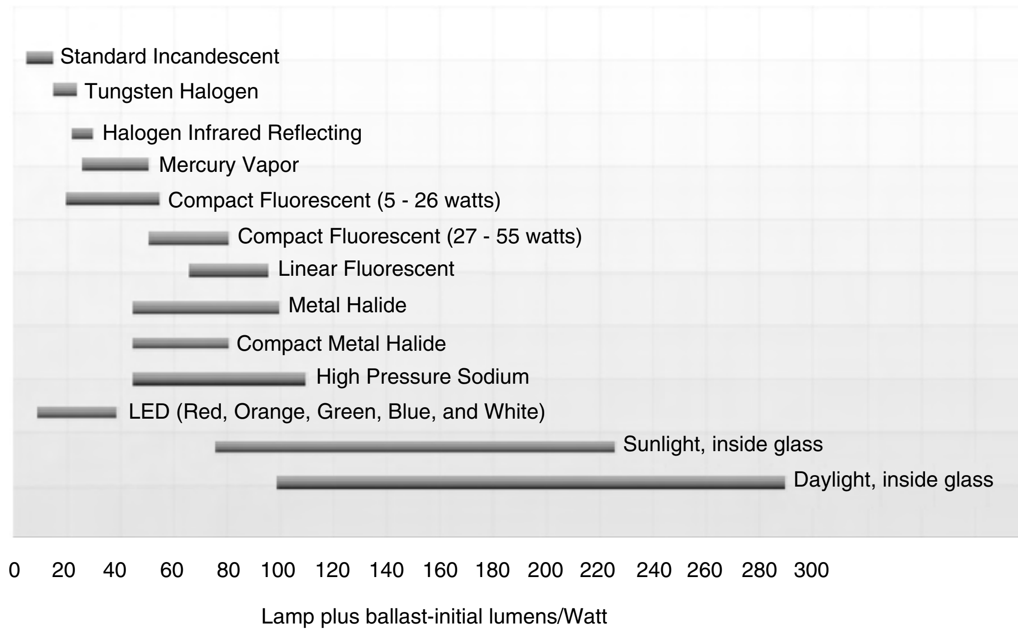


FIGURE 12.24 Efficacy comparison of light sources for general lighting. (From NBI 2003. *Advanced Lighting Guidelines, 2003*, New Buildings Institute, White Salmon, WA. With permission.)

Over time, most lamps continue to draw the same amount of power but produce fewer lumens. The lumen maintenance of a lamp refers to the extent to which the lamp sustains its lumen output, and therefore efficacy, over time. Initial lumens are measured at the beginning of the lamp's life, while mean lumens are measured after a lamp has been used for a percentage of its rated life.

The color properties of a lamp are described by its color temperature and its color rendering index. *Color temperature*, expressed in degrees Kelvin (K), is a measure of the color appearance of the light of a lamp. The concept of color temperature is based on the fact that the emitted radiation spectrum of a blackbody radiator depends on temperature alone. The color temperature of a lamp is the temperature at which an ideal blackbody radiator would emit light that is closest in color to the light of the lamp. Lamps with low color temperatures (3000 K and below) emit warm white light that appears yellowish or reddish in color. Incandescent and warm-white fluorescent lamps have a low color temperature. Lamps with high color temperatures (3500 K and above) emit cool white light that appears bluish in color. Cool-white fluorescent lamps have a high color temperature.

The color rendering index (CRI) of a lamp is a measure of how surface colors appear when illuminated by the lamp compared with how they appear when illuminated by a reference source of the same color temperature. For color temperatures above 5000 K, the reference source is a standard daylight condition of the same color temperature; below 5000 K, the reference source is a blackbody radiator. The CRI of a lamp indicates the difference in the perceived color of objects viewed under the lamp and the reference source. There are 14 differently colored test samples, 8 of which are used in the calculation of the general CRI index. The CRI is measured on a scale that has a maximum value of 100 and is an average of the results for the eight colors observed. A CRI of 100 indicates that there is no difference in perceived color for any of the test objects; a lower value indicates that there are differences. Color rendering indexs of 70 and above are generally considered good, while CRI values of 20 and below are considered poor. Most incandescent lamps have CRI values equal to or approaching 100. Low-pressure sodium (LPS) lamps have the lowest CRI of any common lighting source; their light is essentially monochromatic.

The optical controllability of a lamp describes the extent to which a user can direct the light of the lamp to the area where it is desired. The optical controllability depends on the size of the light-emitting area, which determines the beam spread of the light emitted. In addition, controllability depends on the fixture in which the lamp is used. Incandescent lamps emit light from a small filament area: they are almost point sources of light, and their optical controllability is excellent. In contrast, fluorescent lamps emit light from their entire phosphored bulb wall area; their light is extremely diffuse, and their controllability is poor.

Because of the many different characteristics and the variety of applications, no one light source dominates the lighting market. The types of lamps that are commonly available include incandescent, fluorescent, and high-intensity discharge (HID). Induction lighting systems have come into use as well.

12.2.3.2 Lamps

An electric lamp is a device that converts electric energy into light. Following is a description of common lamp types with discussion on their efficacy and applications.

12.2.3.2.1 Incandescent Lamps

The *incandescent lamp* was invented independently by Thomas Edison in the United States and Joseph Swan in England in the late 1800s. An incandescent lamp produces light when electricity heats the lamp filament to the point of incandescence. In modern lamps the filament is made of tungsten. Because 90% or more of an incandescent lamp's emissions are in the infrared (thermal) rather than the visible range of the electromagnetic spectrum, incandescent lamps are less efficacious than other types of lamps.

The two primary types of standard incandescent lamps are general service and reflector lamps. General-service lamps (also known as A-lamps) are the pear-shaped, common household lamps. A reflector lamp has a reflective coating applied to part of the bulb, this reflective surface (reflector) is specifically contoured for control of the light distribution. Parabolic aluminized reflector (PAR) lamps have optically contoured reflectors. Reflector lamps, such as flood or spotlights, are generally used to illuminate outdoor areas or highlight indoor retail displays and artwork. They are also commonly used to improve the optical efficiency of downlights (recessed can fixtures). Downlights are used where controlling glare or hiding the light source is important. In spite of the fact that they are the least efficacious lamps on the market today, standard incandescent general service lamps are used for almost all residential lighting in the U.S. and are also common in the commercial sector. They have excellent CRI values and a warm color; they are easily dimmed, inexpensive, small, lightweight, and can be used with inexpensive fixtures; and, in a properly designed fixture, they permit excellent optical control. In addition, incandescent lamps make no annoying noises, provide no electro-magnetic interference and contain no toxic chemicals essentially. Incandescent lamps have relatively simple installation, maintenance, and disposal.

Tungsten–halogen and *tungsten–halogen infrared-reflecting (HIR) lamps* are more efficient than standard incandescents. Like standard incandescent lamps, tungsten-halogen lamps produce light when electricity heats the tungsten filament to the point of incandescence. In a standard incandescent lamp, tungsten evaporating from the filament deposits on the glass envelope. Generally, tungsten–halogen lamps use a quartz envelope rather than a glass envelope, which allows the lamp to operate at a much higher temperature. In place of the normal inert gas fill, the tungsten–halogen lamps use a small amount of halogen gas. The halogen gas in the lamp reacts with the tungsten that deposits on the quartz envelope to make a volatile tungsten–halide compound; because tungsten–halide vapor is not stable at the temperature of the filament, the vapor dissociates and deposits the tungsten back onto the filament. The cycle is then repeated. This cycle does not necessarily return the tungsten to the same portion of the filament from which it evaporated, but it does substantially reduce net evaporation of tungsten and thus prolong the life of the filament.

Halogen lamps produce bright white light and have color temperatures and CRI values that are similar to, or slightly higher than, those of standard incandescents. In addition, they have longer rated lives (2000 or more hours vs. 1000 h or less), can be much more compact, are slightly more efficacious, and have better lumen maintenance than standard incandescent lamps. Halogen general service lamps are available but still relatively rare; they offer longer life as well as slightly higher lumen output or lower wattage. Halogen reflector lamp technology has gained the market share for reflector lamps, because this technology meets the efficacy requirements of the Energy Policy Act of 1992, as discussed in Section 12.2.6. Halogen technology is also used in small reflector lamps operated on low-voltage transformers. These lamps, also known as dichroics, are used for accent lighting and sparkle in a variety of applications.

Even more efficacious than the standard tungsten–halogen lamp is the HIR lamp. Because approximately 90% of the energy radiated by incandescent lamps is in the form of heat (infrared radiation), their efficacy can be improved by reflecting the infrared portion of the spectrum back onto the lamp filament. Tungsten–halogen infrared-reflecting lamps use a selective, reflective, thin-film coating on the halogen-filled capsule or on the reflector surface. The coating transmits visible light, but reflects much of the infrared radiation back to the filament hence takes less electricity to heat it. The HIR technology is available in reflector lamps. In general, HIR lamps have a small market share due to their high cost, even though HIR lamps last about 50% longer than regular halogen lamps.

Torchieres, floor lamps that reflect light off the ceiling into the room, became popular in the mid-1990s, especially in the residential sector. The first torchieres were equipped with 300 or 500 W halogen lamps, but these are gradually being phased out. Their high wattages not only caused high energy use, but the lamp temperature was high enough to be a serious fire hazard. The high-wattage halogen lamps can reach almost 1000°F, hot enough to ignite a drape or other combustible material. Because of the resulting fires, some fire departments and college campuses have outlawed halogen torchieres. From an energy perspective, several states have implemented standards to limit the total wattage allowable in the fixture (typically no more than 190 W). As discussed in Section 12.2.6, similar national standards took effect in 2006. Incandescent torchieres presently dominate the market, typically using 150-W bulbs, but more efficient torchiere lamps equipped with various types of fluorescent lamps are available in the market place. A torchiere using a fluorescent lamp drawing 50–70 W costs more than a halogen or incandescent torchiere, but saves energy costs over its lifetime. (See Section 12.2.7).

12.2.3.2.2 Fluorescent Lamps

Fluorescent lamps came into general use in the 1950s. In a fluorescent lamp, gaseous mercury atoms within a phosphor-coated lamp tube are excited by an electric discharge. As the mercury atoms return to their ground state, ultraviolet radiation is emitted. This UV radiation excites the phosphor coating on the lamp tube and causes it to fluoresce, thus producing visible light.

Early fluorescent tubes, and current compact fluorescents lamps as well as some shorter fluorescent tubes, use "preheat start" with an automatic or manual starting switch. "Instant start" lamps were then developed, which uses a high voltage to strike the arc of the lamp. Electronic ballasts (see below) are available that can instant-start most types of fluorescent lamps. "Rapid start" circuits use low-voltage windings for preheating the electrodes and initiating the arc to start the lamps.

Fluorescent lamps are far more efficacious than incandescent lamps. The efficacy of a fluorescent lamp system depends upon the lamp length and diameter, the type of phosphor used to coat the lamp, the type of *ballast* used to drive the lamp, the number of lamps per ballast, the temperature of the lamp (which depends on the fixture and its environment), and a number of lesser factors.

The majority of lighting used in the commercial sector is fluorescent. Fluorescent lighting is also common in the industrial sector. The small amount of full-size fluorescent lighting in the residential sector is primarily found in kitchens, bathrooms, garages and workshops.

Full Size Fluorescent Lamps—The most common fluorescent lamps are tubular and 4 ft. (1.2 m) in length. The next most common length is 8 ft. (2.4 m). Fluorescent lamps are also available in 2-, 3-, 5-, and 6-ft. lengths. Four-feet lamps are also available in U-tube shapes that fit into fixtures with two-feet dimensions. Lamp tubes with a diameter of 1.5 in. (38 mm) are called T12s, tubes that are 1 in (26 mm) in diameter are called T8s, and those that are 5/8 in. (16 mm) are called T5s. (The 12, 8 and 5 refer to the number of eights of an inch in the diameter of the lamp tube.) Lamp tubes are available in other diameters as well. Each type of fluorescent lamp requires a specific ballast, depending on the wattage, length, and current (milliamps) of the lamp.

Fluorescent lamps have long lives and fairly good lumen maintenance. While the reduced wattage halophosphor (cool-white and warm-white) lamps have CRI values in the range of 50–60, rare-earth phosphor lamps have CRI values in the range of 70s and 80s. Fluorescent lamps have rated lifetimes of 12,000 h (8-ft. T12) to 20,000 h (4-ft. T12s and T8s used with rapid start ballasts) When 4-ft. T8s are used with instant start ballasts and frequently switched on and off, their lifetime is decreased to 15,000 h. High-performance (third generation) T8 lamps, also known as "super T8s," with 24,000–30,000 h rated lives and higher CRI values are now available. T5 lamps with better optical control and aesthetics are often used in high ceiling applications and indirect fixtures. T5 lamps are shorter in length than T12 and T8 lamps and require fixtures designed specially for their use. For example, the 28 W T5 lamp is 45.2 in. (1.14 m) in length.

T8 lamps can be operated on magnetic ballasts, but are most commonly used with high-frequency electronic ballasts. Operation with electronic ballasts increases the lamp efficacy by about 9% over operation with magnetic ballasts. T5 lamps operate exclusively with electronic ballasts. The maximum efficacy of a T5 lamp is slightly higher than that of a T8 lamp, and is achievable in ambient conditions about 10°C warmer than those optimal for T8 or T12 lamps. T8 and T5 high output lamps are also available for higher ceiling applications, but this application must be designed to prevent overheating the ballast. Eight-feet lamps have long been available in high-output (HO) and very-high-output (VHO) versions for use in higher ceiling applications.

The most common T12 lamps that meet the present EPAct 1992 lamp standards are reduced-wattage or "energy saver" lamps—for example, the 34-W four-feet lamp. These lamps became popular in the 1970s to retrofit full-wattage lamps (e.g. 40 W lamps). The reduced wattage lamp is similar to its full-wattage predecessor, with krypton added to the gas fill and a conductive coating to lower starting voltage. The lumen output is generally reduced proportionate to the wattage reduction. Previously common 40 W lamps with halophosphors do not meet the EPAct 1992 lamp standards; however, 40 W lamps are available with rare-earth phosphors and higher efficacy that do meet the standards and have a small part of the market share.

For more information about the EPAct 1992 fluorescent lamp standards, see Section 12.2.6, Lighting Efficiency Standards and Incentive Programs.

The specified or nominal wattage of a lamp refers to the power draw of the lamp alone. The ballast typically adds another 10%–20% to the power draw, thus reducing system efficacy. Of the full-size fluorescent lamps available today, *rare-earth phosphor lamps* are the most efficacious. In these lamps, rare-earth phosphor compounds are used to coat the inside of the fluorescent lamp tube. Rare earth phosphor lamps are also called tri-phosphor lamps because they are made with a mixture of three rare earth phosphors that produce visible light of the wavelengths to which the red, green, and blue retinal sensors of the human eye are most sensitive. These lamps have improved color rendition as well as efficacy. Fluorescent T8 and T5 lamps and those with smaller diameters use rare earth phosphors almost exclusively.

The most common efficient fluorescent lamp-ballast systems available today are T8 lamps operating with electronic ballasts. While two T12 34 W halophosphor lamps with an energy saving magnetic ballast have a mean efficacy (with mean lumens) of 54 LPW, two standard 32-W T8 lamps with standard instant-start ballast have a mean efficacy of 75 LPW. *High-performance* T8 systems have evolved over the last few years through a succession of incremental improvements in lamp barrier coatings, phosphors, cathodes, and gas fills, and in ballast circuitry and components.[2] The technical improvements provide higher initial light output, better lumen maintenance, longer life and improved ballast efficiency. High-performance instant start T8 lamp/ballast systems have the highest mean efficacy of any "white" light source at 90 LPW (two-lamp system). The same T8 lamp system with a programmed start ballast has a mean efficacy of 82 LPW [2]. Other T8 lamps offer efficacies above the standard T8 systems, such as four-foot T8 lamps in 25, 28 and 30 W versions that operate on the same electronic ballasts as do 32 W T8 lamps, but these lamps have some operating constraints. There are also T8 lamps with higher light output or longer lamp life than standard T8s, although they don't offer the combined benefits of the high-performance T8 systems.

The consortium for energy efficiency (CEE), a national, nonprofit organization, developed a specification to consistently define high-performance T8 lamps and ballasts. This specification, shown in Table 12.4, was developed for voluntary usage by utilities and energy-efficiency organizations promoting high-performance commercial lighting systems. A listing of qualifying products can be found on the CEE website at www.cee1.org.

Note: The specification and qualifying products list are updated and revised periodically. Please refer to the CEE website for the most current information.

Presently, the installed cost of a high-performance T8 system is more than the of a T12 or a standard T8 system; however, that cost differential is shrinking. The increased efficacy and longer life of the high-performance systems make them the best economic choice for most applications. (See Section 12.2.7).

In spite of their much greater efficiency, fluorescent lamps have several disadvantages when compared with incandescent lamps. Standard and compact fluorescent lamps can be dimmed, but require special dimming ballasts that cost more than the dimming controls used for incandescent lamps. Standard fluorescent lamps are larger than incandescent lamps of equivalent output and are harder to control optically. Fluorescent lamps also emit more UV light than incandescent lamps. Ultraviolet light can cause colors to fade, and fabrics to age, and therefore has to be blocked near sensitive materials like museum displays. Electronic ballasts may interfere with security equipment, such as that used in libraries and with specialized hospital devices.

Fluorescent lamps contain trace amounts of mercury, a toxic metal, and large users are required to either recycle them or dispose of them as hazardous waste. However, mercury is also emitted through the electricity production process, and the net total emission of mercury including the power plant emissions is actually lower for fluorescent lamps than for the incandescent lamps that they replace. Lamp manufacturers have begun to produce fluorescent lamps with lower mercury content that meet U.S. Environmental Protection Agency requirements and allow disposal of small quantities of lamps. Regulations vary by state; for more information, see www.lamprecycle.org or http://www.almr.org/.

Circular Fluorescent Lamps—Circular-shaped fluorescent lamps in 20–40-W sizes have been available for many years, but have had a fairly small market. Essentially, a circular lamp is a standard fluorescent lamp tube (as described earlier) that has been bent into a circle. Although they have a more compact geometry than a straight tube, circular lamps are still moderately large (16.5–41 cm in diameter). Circular lamps are available in several sizes with magnetic or electronic ballasts.

[2]The three major U.S. lamp manufacturers have each developed high-performance T8 lamps: GE High Lumen ECO, Philips Advantage Alto, and Sylvania Xtreme XPS EcoLogic. ("ECO" and "ALTO" identify the lamps as low mercury.) Five manufacturers offer high-performance T8 ballasts: Advance Optanium, GE Ultramax, Howard Industries HEX, Sylvania QHE, and Universal Ultim8.

TABLE 12.4 CEE T8 High-Performance Specification

Performance Characteristics for Lamps				
Mean System Efficacy	≥ 90 MLPW for Instant Start Ballasts ≥ 88 MLPW for Programmed Rapid Start Ballasts			
Color Rendering Index (CRI)	≥ 81			
Minimum Initial Lamp Lumens	≥ 3100 Lumens			
Lamp Life	≥ 24,000 hrs at three hours per start.			
Lumen Maintenance –or – **Minimum Mean Lumens**	≥ 94% –or– ≥ 2900 Mean Lumens			
Performance Characteristics for Ballasts				
Ballast Efficacy Factor (BEF)	**Instant-Start Ballast (BEF)**			
BEF = [BF × 100] / Ballast Input Watts	Lamps	Low BF ≤ 0.85	Norm 0.85 < BF ≤ 1.0	High BF ≥ 1.01
Based on: (1) Type of ballast (2) No. of lamps driven by ballast (3) Ballast Factor	1	≥ 3.08	≥ 3.11	n/a
	2	≥ 1.60	≥ 1.58	≥ 1.55
	3	≥ 1.04	≥ 1.05	≥ 1.04
	4	≥ 0.79	≥ 0.80	≥ 0.77
	Programmed Rapid-Start Ballast (BEF)			
	1	≥ 2.84	≥ 2.84	n/a
	2	≥ 1.48	≥ 1.47	n/a
	3	≥ 0.97	≥ 1.00	n/a
	4	≥ 0.76	≥ 0.75	n/a
Ballast Frequency	20 to 33 kHz or ≥ 40 kHz			
Power Factor	≥ 0.90			
Total Harmonic Distortion	≤ 20%			

Voluntary guidelines for use in energy-efficiency programs. For terms and usage, please see the CEE website (www.w.cee1.org). MLPW = mean lumens per watt.

Compact Fluorescent Lamps—Compact fluorescent lamps, which are substantially smaller than standard fluorescent lamps, were introduced to the U.S. market in the early 1980s. In a CFL, the lamp tube is smaller in diameter and is bent into two to six sections or into a spiral shape. The compact size of the spiral CFLs allow them to substitute for incandescent lamps in many fixtures. Compact fluorescent lamps have much higher power densities per phosphor area than standard fluorescents, and their design was therefore dependent on the development of rare-earth phosphors, which could hold up much better than standard phosphors at high-power loadings. Compact fluorescent lamps (except for very low-wattage CFLs) are much more efficacious than the incandescent lamps they replace, typically drawing one-third to one-quarter of the wattage for similar light output. They are, however, half as efficacious as high-performance T8 systems, hence they are more often cost-effective for applications or retrofits where a smaller light source is needed. Compact fluorescent lamps are also rated to last 6000–10,000 h, with some newer products having even longer rated lifetimes.

Compact fluorescent lamps are available as both screw-in replacements for incandescent lamps and as pin-base lamps for hard-wired fixtures. Common CFLs range from 11 to 26 W, and both higher and lower wattage lamps are available. They may be operated with separate ballasts or purchased as integral lamp/ballast units; integral units with electronic ballasts are the most commonly sold. The electronic ballast provides higher efficacy, eliminates the starting flicker and has a lighter weight. Compact fluorescent lamps have a much higher retail cost than the incandescent lamps they replace, hence consumers have been reluctant to purchase them without discounts or incentives; their prices have decreased with their increasing popularity. Particularly for the residential sector, CFLs have been somewhat limited for use in fixtures because their size and shape relative to incandescent lamps. Manufacturers have developed lamps with smaller, more compatible shapes, as well as adapters for some fixtures. Early CFL users in the residential sector encountered some starting problems at low temperatures for outdoor use and overheating in enclosed fixtures. Compact fluorescent lamps that start at low outdoor temperatures are now available. Standard CFLs may not be used in dimming circuits because of fire hazard, but dimmable CFLs became available during the late 1990s. Compact fluorescent lamps reflector lamps are also available. In the commercial sector, dedicated fixtures with built-in ballasts that accept only pin-based CFLs are often used in downlights, to ensure that screw-in incandescents do not replace the CFLs.

12.2.3.2.3 Induction Lamps

An induction lamp is a fluorescent lamp where the electric discharge is induced by a magnetic field, rather than an electric field as in a fluorescent lamp, and therefore does not have any electrodes. Induction lamps produce light by exciting the same phosphors found in conventional fluorescent lamps. The radio frequency (RF) power supply sends an electric current to an induction coil, generating an electromagnetic field. This field excites the mercury in the gas fill, causing the mercury to emit UV energy. The UV energy strikes and excites the phosphor coating on the inside of the glass bulb, producing light. Electrodeless lamps have efficacies similar to those of CFLs or HID lamps of comparable light output. Electrodeless lamps use rare-earth phosphors, giving them color properties similar to those of higher-end fluorescent lamps. Because the lamp has no electrodes that usually cause lamp failure, the life of this system is limited by the induction coil. Induction lamps are rated at 100,000 h of life. Because of this long life, and the good color rendition, induction technology is coming into use for areas where maintenance to change the lamp is expensive, such as high ceilings in commercial and industrial buildings, atria, tunnels, roadway sign lighting, etc.

Induction lamps are electronic devices, and like all electronic devices they may generate electromagnetic interference (EMI) if unwanted electromagnetic signals, which can travel through wiring or radiate through the air, interfere with desirable signals from other devices. Shielding of the system to protect people and equipment from these emissions is important. Manufacturers must comply with national regulations on EMI to sell products in any country.

12.2.3.2.4 High-Intensity Discharge Lamps

High-intensity discharge (HID) lamps produce light by discharging an electrical arc through a mixture of gases. In contrast to fluorescent lamps, HID lamps use a compact arc tube in which both temperature and pressure are very high. Compared to a fluorescent lamp, the arc tube in an HID lamp is small enough to permit compact reflector designs with good light control. There are presently three common types of HID lamps available: mercury vapor (MV), metal halide (MH), and high-pressure sodium (HPS). Additionally, low pressure sodium (LPS) lamps, while not technically HID lamps, are used in some of the same applications as HPS lamps.

Because of their higher light output levels, HID lamps are most often used for exterior applications such as street and roadway lighting, outdoor area pedestrian and parking lot lighting, commercial, industrial, and residential floodlighting and security lighting, and sports lighting. They are also used in large, high-ceilinged, interior spaces such as industrial facilities and warehouses, where good color rendering is not typically a priority. Occasionally, HID lamps are used for indirect lighting in commercial offices, retail stores, and lobbies. Interior residential applications are rare because of high cost, high light level, and the fact that HID lamps take several minutes to warm up to full light output. If they are turned off or there is a momentary power outage, the lamps must cool down before they restrike. Some HID lamps are now available with dual arc tubes or parallel filaments. Dual arc tubes eliminate the restrike problem and a parallel filament gives instantaneous light output both initially and on restrike, but at a cost of a high initial power draw and higher lamp cost.

The *mercury vapor lamp* was the first HID lamp developed. Including ballast losses, the efficacies of MV lamps range from approximately 25–50 LPW. Uncoated lamps have a bluish tint and very poor color rendering (CRI $\sim$ 15). Phosphor-coated lamps emit more red, but are still bluish, and have a CRI of about 50. Because of their poor color rendition, these lamps are used only where good color is not a priority. MV lamps generally have rated lifetimes greater than 24,000 h. Both MH and HPS HID lamps have higher efficacies than MV lamps and have consequently replaced them in most markets. MV lamps and ballasts are cheaper than the other HID sources and are still often sold as residential security lights. They also persist in some legacy street-lighting applications, landscape lighting, and in some other older systems.

Including ballast losses, *metal halide lamps* range in efficacy from 45 to 100 LPW. They produce a white light and have CRI values ranging from 65 to almost 90. Lamp lifetimes generally range from only 5000 to 20,000 h, depending on the type of MH lamp. Lower-wattage metal halides (particularly the 50, 70 and 100 W) are now available with CRI values of about 65–75 and color temperatures of 2900–4200 K. Reasonably good lumen maintenance, longer life, reduced maintenance costs, and the fact that they blend more naturally with fluorescent sources have made MH lamps a very good replacement in the commercial sector for 300 and 500 W PAR lamps. New fixtures utilizing these lamps, particularly 1-ft. by 1-ft. recessed lensed troffers (downlights), are becoming common in lobbies, shopping malls, and retail stores. Improvements in color stability have made MH systems cost-effective substitutions for high-wattage incandescent lighting in commercial applications.

Metal halide technology is also becoming increasing popular for outdoor lighting, especially in areas where color rendering is important, and because of people's preference for "white light." Pulse-start technology is improving MH lamp performance in almost every aspect. In MH pulse start lamps, a high-voltage pulse (typically 3 kV minimum) applied directly across the main electrodes initiates the arc. Ignitors are used to provide these starting pulses. The average lifetime of pulse-start MH lamps now approaches that of HPS and MV lamps. With the higher efficacy this technology provides, approaching that of HPS lamps, pulse-start MH lamps now compete with HPS lamps in many outdoor applications. Ceramic arc tube MH lamps, with CRI values as high as 90 and better color consistency, now compete with incandescent sources. MH lamps may fail "nonpassively," hence users should always follow the manufacturers' recommended practices for safe operation of the lamps.

At low light levels, such as those found in many outdoor areas at night, the eye's peripheral vision becomes more sensitive to light that is bluish. Although MH lamps are less efficacious than HPS lamps at high, or "photopic" light levels found during daylight hours, they can actually provide higher visual quality and therefore allow lower light levels, making them more efficacious, at least for peripheral vision, than HPS lamps at low or "scotopic" levels. This has led to increased interest in their use for street lighting.

Including ballast losses, *high-pressure sodium lamps* have efficacies ranging from 45 LPW for the smallest lamps to 110 LPW for the largest lamps. Standard HPS lamps emit a yellow–orange light and have poor color rendition in the 20 s; high color rendering versions can have CRI values up to 70 and higher. Like MV lamps, HPS lamps are only used where good color is not a priority. HPS lamps have come to dominate street and roadway lighting because of their high efficacy and long life. The rated lifetimes of HPS lamps rival those of MV lamps and typically exceed 24,000 h.

For more information on the HID lamp market, see Section 12.2.5.

12.2.3.2.5 Low-Pressure Sodium Lamps

Low-pressure sodium (LPS) lamps are discharge lamps that operate at lower arc tube loading pressure than do HID lamps. Low-pressure sodium lamps are monochromatic in the yellow spectral band and have CRI values of 44. They have been used for street and tunnel lighting, especially in cities near astronomical observatories, because the LPS color spectrum can easily be filtered out so as not to interfere with telescopes. (In more recent efforts to limit "sky glow" as well as glare from outdoor lighting, the emphasis has shifted from lamp types to luminaire light control, with various styles of "cutoff" luminaires directing light downward and not upward, or using shielding to the same effect. For street lighting, calculations suggest that the match of the luminaire light distribution to the street, rather than the cutoff classification, is the most critical factor in limiting sky glow.) The LPS lamp has limitations for many applications where color rendering is important for safety and identification.

12.2.3.2.6 Light Emitting Diodes

Light emitting diodes (LEDs) are semiconductor diodes that emit light when current flows through them. They are available as narrowband light sources in "colors" ranging from the infrared to the near UV, and, with the addition of a phosphor coat, as a white light source with color temperatures ranging from 3200 to 12,000°K and CRI values ranging from 60 to over 90. Rated efficacies range up to 30 LPW for white LEDs, which is superior to that of incandescent lamps.

The present major disadvantages of LEDs relative to other light sources are their low maximum wattage per unit (presently 3–10 W) and their high unit cost. The present effective cost per watt (retail) is on the order of $3W. The cost of incandescent lamps is not strongly dependent on lamp wattage, which means that large incandescent lamps are 1000 times less expensive per watt than LEDs, but small ones are more comparable in price.

Light emitting diodes have other advantages over small incandescent lamps that are leading to their increasing dominance for battery operated lighting. Their usable life ranges from 6000 to 50,000 h (at 70% lumen maintenance). Small standard flashlight bulbs have thin filaments and can only manage 8–10 LPW and 15–30 h of lamp life. Light emitting diodes are durable, require less expensive optics for good beam control, and are available with high color temperatures, which lead to higher perceived brightness and better visibility in low light conditions. The power supplies used in the more expensive lights significantly increase the fraction of battery power available to provide useful amounts of light. LED bike lights and flashlights have battery lives that are a factor of 10 greater than those of comparable standard incandescent lights, and a factor of 3–5 greater than those of halogen lamps. The use of a flashing mode can increase this advantage even further.

Light emitting diodes are also displacing incandescent lamps in traffic signal lights and dynamic sign and color displays. Light emitting diodes have the advantage over incandescent lamps in these applications because filtering incandescent lamps to produce colors reduces their efficacy to the 2–3 LPW range, while switching on and off reduces their life. Light emitting diodes are also replacing incandescent and fluorescent light sources in exit signs. As discussed Section 12.2.6, EPAct 2005 contains standards for traffic and pedestrian signals that require LED technology, and for exit signs that require LED technology or technology with similar efficacies.

Manufacturers have announced that 60 LPW white LEDs will be available in the near future. This is comparable to the efficacy of small fluorescent lamps, and the industry target of 150 LPW by 2012 would make LEDs more efficacious than all but monochromatic LPS lamps. Research continues on organic LEDs (OLEDs), which are cheaper, but less efficacious than standard LEDs, and on manufacturing techniques to reduce the cost per unit of standard LEDs while increasing their wattage. The lack of toxic mercury in LEDs could make them preferable to small fluorescent lamps in a few years time, even if they are a bit more expensive, and their efficacy is no higher.

12.2.3.3 Ballasts

Because fluorescent and HID lamps (both discharge lamps) have a low resistance to the flow of electric current once the discharge arc is struck, they require a device to limit current flow. A lamp ballast is an electrical device used to control the current provided to the lamp. In most discharge lamps, a ballast also provides the high voltage necessary to start the lamp. Older preheat fluorescent lamps require a separate starter, but these lamps are becoming increasingly uncommon. In many HID ballasts, the ignitor used for starting the lamp is a replaceable module.

The most common types of fluorescent ballasts are magnetic core-coil and electronic high-frequency ballasts. A *magnetic core-coil ballast* uses a transformer with a magnetic core coiled in copper or aluminum wire to control the current provided to a lamp. Magnetic ballasts operate at an input frequency of 60 Hz and operate lamps at the same 60 Hz. An *electronic high-frequency ballast* uses electronic circuitry rather than magnetic components to control current. Electronic ballasts use standard 60 Hz power but operate lamps at a much higher frequency (20,000–60,000 Hz). Both magnetic and electronic ballasts are available for most fluorescent lamp types.

The *cathode cut-out (hybrid) ballast* is a modified fluorescent magnetic ballast. It uses an electronic circuit to remove the filament power after the discharge has been initiated for rapid-start lamps. Cathode cutout ballasts use approximately 5%–10% less energy than energy-efficient magnetic ballasts.

Of the ballasts that are presently available for fluorescent lamps, the most efficient option is the electronic ballast. Because an electronic ballast is more efficient than a standard core-coil magnetic ballast in transforming the input power to lamp requirements, and because fluorescent lamps are more efficient when operated at frequencies of 20,000 Hz or more, a lamp/ballast system using an electronic rather than magnetic ballast is more efficacious.

In addition, electronic ballasts eliminate flicker, weigh less than magnetic ballasts, and operate more quietly. Since electronic ballasts are packaged in cans that are the same size as magnetic ballasts, they can be placed in fixtures designed to be used with magnetic ballasts. Fluorescent electronic ballasts are available for standard commercial-sector applications. They have become increasingly popular, particularly in new luminaires as well as in energy-efficiency retrofits (see Section 12.2.5).

There are three basic starting modes for electronic ballasts. Rapid start ballasts apply continuous low filament voltage to preheat the cathode and use a higher starting voltage to strike the arc to start the lamp. Instant start ballasts apply high voltage across the lamp without preheating the cathode. This allows lower system wattage, but the instant starting shortens lamp life. Programmed start ballasts, a type of rapid start ballast developed most recently, apply cathode heat prior to starting the lamp and then lower or cease it once the lamp has started. Programmed start ballasts are recommended for use with occupancy sensors and frequently switched applications. For all other applications, instant-start ballasts are the more efficient choice. Advances in dimming fluorescent ballasts allow further energy savings through automatic controls.

Ballast factor indicates the light output of a lamp when operated with a specific ballast relative to the light output of the same lamp when operated with a reference ballast. Electronic ballasts are available with high and low ballast factors for higher or lower light levels where needed for specific applications.

The most commonly-used ballasts for HID lamps are magnetic, and a number of different types are available. The various types differ primarily in how well they tolerate voltage swings and, in the case of HPS lamps, the increased voltage required to operate the lamp as it ages. Electronic ballasts are also available, although the energy savings are less than for fluorescent systems. As with fluorescent systems, HID electronic ballasts provide flicker-free lighting and regulate lamp power, which increases lamp life, maintains constant color, and allows for precise starting, warm-up, and operation. Newer electronic HID ballasts improve system efficiency, are smaller in size and are becoming available in a wider range of wattages.

12.2.3.4 Lighting Fixtures

A lighting fixture is a housing for securing lamp(s) and ballast(s) and for controlling light distribution to a specific area. The function of the fixture is to distribute light to the desired area without causing glare or discomfort. The distribution of light is determined by the geometric design of the fixture as well as the material of which the reflector and/or lens is made. The more efficient a fixture is, the more is the light it emits from the lamp(s) within it. Although the term *luminaire* is sometimes used interchangeably with *fixture*, *luminaire* refers to a complete lighting system including a fixture, lamp(s) and ballast(/s).

Types of fluorescent lighting fixtures that are commonly used in the nonresidential sectors include recessed troffers, pendant-mounted indirect fixtures and indirect/direct fixtures, and surface-mounted fixtures such as wraparound, strip, and industrial fixtures.

Until recently, most offices have been equipped with *recessed lensed troffers*, which are direct (light emitted downward) fixtures and emphasize horizontal surfaces. Many forms of optical control are possible with recessed luminaires. In the past, prismatic lenses were the preferred optical control because they offer high luminaire efficiency and uniform illuminance in the work space. More recently, *parabolic louvered fixtures* have become common in office spaces. These fixtures have reflectors in the form of louvers, often aluminized, with parabolic geometry that directs light downward.

Offices with electronic equipment have become the norm, and until fairly recently there was a trend away from the traditional direct lighting fixtures designed for typing and other horizontal tasks because they tend to cause reflections on video display terminal (VDT) screens. No lighting system reduces glare entirely, but some fixtures and/or components can reduce the amount of glare significantly. Because the glossy, vertical VDT screen can potentially reflect bright spots on the ceiling, and because VDT work is usually done with the head up, existing fixtures are sometimes replaced with indirect or direct/indirect fixtures, which produce light that is considered more comfortable visually. Most indirect lighting systems are suspended from the ceiling. They direct light toward the ceiling where the light is then reflected downward to provide a calm, diffuse light. Some people describe the indirect lighting as similar to the light on an overcast day, with no shadows or highlights. Generally, indirect lighting does not cause bright reflections on VDT screens. A *direct/indirect fixture* is suspended from the ceiling and provides direct light as well as indirect. These fixtures combine the high efficiency of direct lighting systems with the uniformity of light and lack of glare produced by indirect lighting systems. New, flat-panel-liquid crystal display monitors are much less sensitive to reflections than the older cathode-ray tube VDTs, and this may affect the types of fixtures used in offices in the future.

A *wraparound fixture* has a prismatic lens that wraps around the bottom and sides of the lamp, and is always surface-mounted rather than recessed. Wraparound fixtures are less expensive than other commercial fixtures and are typically used in areas where lighting control and distribution are not a priority. *Strip and industrial fixtures* are even less expensive and are typically used in places where light distribution is less important, such as large open areas (grocery stores, for example) and hallways. These are open fixtures in which the lamp is not hidden from view.

The most common incandescent fixture in the nonresidential sector is the *downlight*, also known as a recessed can fixture. Downlights are also becoming increasingly popular for residential lighting. Fixtures designed for CFLs are available to replace incandescent downlight fixtures in areas where lighting control is less critical.

Interior HID luminaires include high bay and low bay fixtures (aisle lighters, parking garage, etc.), downlights, and accent lighting. Exterior HID luminaires include street and roadway fixtures (cobrahead, post-top, architectural), floods of all sizes and types (from sports to landscape), wall mounted, and security fixtures.

Luminaire Efficacy Rating is a single metric that expresses *luminaire efficacy*, the luminaire's light output divided by the input power. The formula is:

$$\text{LER} = \frac{\text{Luminaire efficiency (EFF)} \times \text{Total rated lamp lumens (TLL)} \times \text{Ballast factor (BF)}}{\text{Luminaire Watts input}}$$

Note that the effects of all components of the luminaire system (lamp, ballast and fixture) are included in the LER. The National Electrical Manufacturers Association (NEMA) Standards Publication No. LE5, "Procedure for Determining Luminaire Efficacy Ratings for Fluorescent Luminaires", specifies the major fluorescent luminaire categories covered and the standard industry test procedures [14].

NEMA Publications LE5A and LE5B provide the specifications for commercial downlight luminaires [13] and HID industrial luminaires [12].

12.2.3.5 Lighting Controls

Lighting controls include a wide range of technologies that electronically and/or mechanically control the use of lights in a building. Control systems range from simple light switches and mechanical timeclocks to sophisticated building energy management systems that control the lighting in a building as well as the heating, ventilation, and air conditioning systems. Lighting control systems include programmable timers, occupancy sensors, photosensors, dimmers, switchable or dimmable ballasts, and communications and control systems. In the commercial sector, controls are used to save energy, curtail demand, or tailor the lighting environment to changes in lighting requirements. Both occupancy sensing and scheduling save energy by turning lights off or to a lower level when no one is present. Occupancy sensors, which sense people's presence in the room by means of infrared or ultrasound signals that detect movement and turn lights off a short time after people leave the space, can save from 20 to 40% of lighting energy in the spaces they control. When implemented together with scheduling, where the lights are automatically turned off during known periods of nonoccupancy (with override capability for occupants), savings can be on the order of 50%. Daylighting controls dim or switch lights when there is sufficient daylight, and can save 30%–40% of the energy use for lighting along the building perimeter during the day. Demand curtailment can be performed during certain periods by a building energy management system by dimming or switching lights in response to high peak demand costs, or emergency conditions where there is a risk of power outages. Dimmable systems can be set to turn on at different levels in response to changes in occupancy or work demands. Transitional areas from outdoors to indoors (such as tunnel entrances and lobbies) can use dimming or switching systems to provide high light levels during bright day time conditions and lower levels at night. Older control systems may maintain light levels by increasing power to lamps as they age and lose their lumen maintenance and efficacy, but this is no longer considered economic and is discouraged.

Development of more accurate and reliable sensors, better communication between systems, and better integration of systems continues to increase the utility of control systems and the resulting energy savings. An integrated workstation sensor allows users to control not just lighting, but heating and cooling (space temperature), and other electrical equipment (such as plug loads) for individual workstations or spaces. Communication to a building energy management system allows the main system to set up and change defaults, over-ride individual settings in case of emergency, and monitor equipment status. Improvements in the spectral sensitivity of the photosensors, and improvements in the dimming algorithms, should lead to higher user satisfaction with daylighting systems, and consequently to larger energy savings. Users sometimes deliberately disable poorly functioning control equipment. Improvements in occupancy sensing, and scheduling algorithms, should help produce systems which are more likely to remain functioning, and thus actually save energy.

In the residential sector, the control of exterior lighting through photocells, motion sensors and time clocks is fairly common. Simple motion sensors turn outdoor lights on as occupants approach the home and turn them off a few minutes after they leave the area and also discourage burglars; photocells prevent daytime operation. These sensors sometimes experience false triggering from pets. Automatic controls for residential interior lighting are less common but have come into use. For example, scene controllers let residents set the levels of multiple lamps with a single command. Computer-controlled relays allow residents to use the internet to switch exterior lighting on when they are about to arrive home.

12.2.4 Efficient Lighting Operation

In addition to high quality design and the use of efficient lighting technologies, commissioning and maintenance of lighting systems play important roles in maximizing energy savings.[3] The commissioning

[3]Commissioning involves reviewing design documentation, verifying installation, verifying testing of equipment and system performance, training building operators, and analyzing the operation of an efficient lighting system.

process is necessary because lighting systems (particularly lighting controls) often do not perform exactly as they were designed to perform. Although often ignored, maintenance is essential for maximizing the efficient performance of lighting fixtures.

The presence of dirt on a luminaire can significantly reduce its light output. In a very clean environment (e.g., high-quality offices, clean rooms, laboratories), luminaire dirt is estimated to reduce light output over a period of 3 years by 10%–20%, depending upon the type of luminaire used. In a moderately dirty environment (e.g., mill offices, paper processing, light machining), dirt on luminaires can reduce light output by as much as 45% in 3 years. In a very dirty environment, light output can be reduced by 50% in less than 2 years. Dirt on the surfaces in a room can also reduce the available light by limiting the ability of surfaces to inter-reflect from one to another. Consequently, cleaning not only makes the lighting system and room look better, but also increases the efficiency of the lighting system.

Group relamping (replacing all lamps in an area or building simultaneously rather than as each lamp burns out) is another strategy for maintaining high-efficiency fluorescent and HID lighting systems and often saves time and money as well.[4] As mentioned above, most lamps become less efficacious and produce fewer lumens as they age. Therefore, if a lighting system is designed to maintain a minimum light level over many years, it must be designed to produce more than the minimum light level when first installed. Group relamping is usually done when the lamps have operated for about 75% of their rated lifetimes [16]. Even though the lamps are retired before the end of their useful lives, early relamping can greatly reduce the amount of initial over lighting that is necessary in the design and results in a more efficient lighting system. When compared with the high labor cost of replacing one lamp at a time, group relamping can often result in substantial labor cost savings. In addition, while replacing lamps, a maintenance crew has a convenient opportunity to clean lamp lenses and reflectors.

12.2.5 Current Lighting Markets and Trends

As a whole, the U.S. lamp industry is highly concentrated—almost all of the market is shared among a few firms. Lighting manufacturers are typically multinational corporations serving markets around the world. The U.S ballast companies started out primarily as separate firms from the lamp companies, but in recent years each of the major lamp manufacturers have either purchased or become aligned with a major ballast company U.S. fixture manufacturers have seen a series of acquisitions over the past few decades and many firms are now subsidiaries of larger firms. Five or so fixture manufacturers account for more than one-half of the market share. In contrast to the concentrated markets for lamps, ballasts, and fixtures, there are many firms involved in the lighting controls market. The technologies in this group are highly varied; for example, some firms manufacture very simple timers while others manufacture highly complex control systems for whole buildings.

In this section we present details on market trends for HID lamps and fluorescent lamp ballasts. Details on market trends for most lamp types are difficult to track since the Census Bureau discontinued its Current Industrial Report on electric lamps; however, some data on HID lamps are available from a U.S Department of Energy report [5]. Annual fluorescent ballast data are available from the U.S. Census Bureau's Current Industrial Reports, MQ335C (formerly MQ36C) [3]. The Census Bureau also collected data on electric lighting fixtures through 2001, which are still available on their website as MA335L (formerly MA36L).

Mercury vapor lamps were the first HID lamps, entering the market in the 1930s, and came into common use for outdoor lighting. Even in the early 1970s, nearly all HID lamps were MV. They were replaced by HPS lamps, which were first introduced in the 1960s and became common in street and roadway lighting because of the cost savings from their higher energy efficiency. Metal halide lamps, introduced into the market just after HPS lamps, have more recently become popular for area lighting

[4]Group relamping is less important for incandescent lighting since incandescents lamps have shorter lifetimes and less lumen depreciation than do fluorescent or HID lamps.

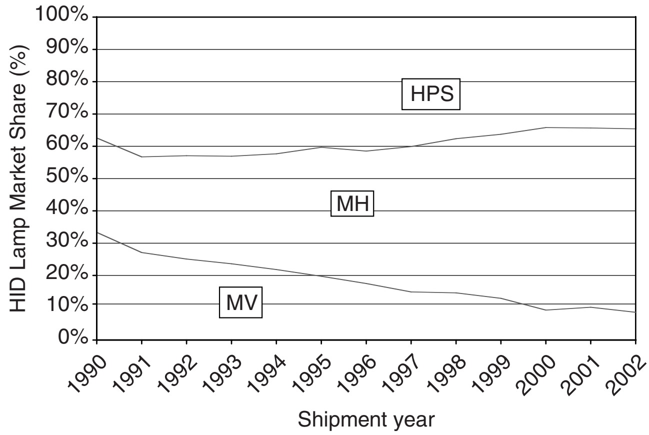


FIGURE 12.25 Total U.S. HID lamp shipments by type, 1990–2002. (From U.S. Department of Energy.)

and some interior applications as a "white light" source with better color rendering than HPS lamps. Mercury vapor lamp usage has declined considerably, although it persists in some areas for street lighting, rural security lighting, and landscape lighting.

As shown in Figure 12.25, the percentage of total HID shipments comprised by MV lamps declined from 33% in 1990 to 9% in 2002, as shown in Figure 12.24. During that period, MH lamp shipments increased, while HPS lamp shipments remained relatively constant.

The fluorescent ballast market has evolved considerably in recent years and has experienced a growing percentage of electronic ballasts. Figure 12.26 shows the evolution of the fluorescent ballast market for ballasts used in the commercial and industrial sectors (electronic and magnetic power-factor corrected ballasts). Electronic ballasts accounted for less than 3% of total ballast sales for these sectors in 1987, but shipments began to grow in the early 1990s, partly through utility rebate and other incentive programs. The percentage of the market comprised by electronic ballasts has grown rapidly, going from 35% in 1997 to 67% in 2004. This strong increase in the sale of electronic ballasts has been accompanied by a significant decrease in price.

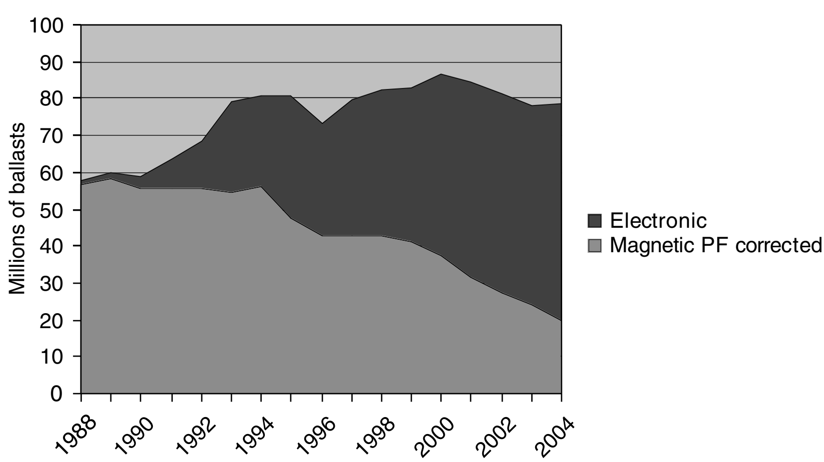


FIGURE 12.26 U.S. fluorescent ballast shipments, 1988–2004. (From U.S. Census Bureau, Current Industrial Reports for Fluorescent Lamp Ballasts, 1988–2004)http://www.census.gov/cir/www/335/mq335c.html.

As discussed in Section 12.2.6, new ballast efficiency standards began to affect the market in 2005, prohibiting the manufacture of magnetic fluorescent ballasts destined for the fluorescent luminaire market, with their production for this market required to cease in 2005 and sales to stop in mid-2006. Magnetic ballasts for the replacement ballast market (for existing fixtures) can still be sold until mid-2010.

12.2.6 Lighting Efficiency Standards and Incentive Programs

In U.S., there are various drivers to development and promotion of efficient lighting technologies. Research and development is an ongoing process for manufacturers of lighting equipment and they regularly refine existing technologies and introduce new products to the market. Developments in the electronics industry are likely to foster innovations in lighting technologies as well. Federal national laboratories have been involved in efficiency-related research and market transformation programs, and have been especially active in developing and promoting technologies such as electronic ballasts and fluorescent torchieres. In the 1990s, electric utility demand-side management (DSM) incentives, such as customer rebates and information programs, actively promoted efficient lighting technologies. Although utility deregulation reduced the involvement of utilities in the promotion of efficient technologies, subsequent electricity supply crises and the slowdown of deregulation have contributed toward the return of some DSM incentives.

Government regulations are important policy tools for encouraging energy efficiency. At the federal level, lamps and ballasts are regulated through equipment standards and labeling requirements. Within states, lighting levels are regulated through building codes. Federal and state governments also sponsor nonregulatory information and incentive programs for efficient lighting. The U.S. Environmental Protection Agency (EPA) and Department of Energy (DOE) sponsor Energy star®, an information and voluntary incentive program that has designed energy-efficiency specifications for a variety of products. The Energy star® program also provides design guidelines and assistance to businesses and homeowners who are interested in reducing their buildings' energy consumption, including lighting energy, and helps them to set savings goals. Federal government and state governments also promote lighting efficiency through procurement policies.

Three U.S. laws specifically mandate the use of efficient lighting components: the National Appliance Energy Conservation Amendments[5] of 1987 (NAECA), the Energy Policy Act of 1992 (EPAct 1992)[6], and the Energy Policy Act of 2005 (EPAct 2005).[7] Some states have also adopted standards on certain lighting products. In addition to standards on specific equipment, states may regulate lighting through building codes that have maximum lighting consumption limits (W/square ft. or W/square meter). EPAct 1992 requires states to adopt certain building code provisions for new construction or substantial renovations, as discussed below. Many states have complied, and a few states have building codes that exceed the EPAct 1992 requirements.

NAECA regulates the sale of ballasts that drive fluorescent lamps. Ballasts sold for use in commercial or industrial lighting installations are required to meet or exceed minimum ballast efficacy factors (BEF) values.[8] The original ballast efficiency standards that took effect in 1991 applied to certain fluorescent lamp ballasts that operate T12 lamps. These standards required ballasts to have efficacies greater than or equal to those of energy-efficient magnetic ballasts. Updated standards, which essentially require the efficiencies of electronic ballasts, apply to ballasts manufactured after July 1, 2005 and incorporated into luminaires by luminaire manufacturers on or after April 1, 2006. Ballasts used for replacement (existing buildings) have until June 30, 2010 to meet the standard levels. The 5-year time delay was designed to allow building

[5]NAECA amended the National Energy Conservation and Policy Act (NECPA) of 1978 that authorized the U.S. Department of Energy (DOE) to set mandatory energy-efficiency standards for various products.

[6]Public Law 102–486.

[7]Public Law 109-58.

[8]BEF is defined as ballast factor $\times$ 100 divided by input watts.

owners to retrofit their spaces with uniform systems rather than replacing them one by one as ballasts burn out. In addition, EPAct 2005 contains provisions that add BEFs for ballasts that drive reduced-wattage T12 lamps, taking effect in 2010. End users who cannot use electronic fluorescent ballasts for technical reasons (primarily electronic interference with specialized equipment) may use T8 lamps with magnetic ballasts. These magnetic ballasts are specifically designed for use with T8 lamps; magnetic ballasts that operate T12 lamps draw different amperage than T8 magnetic ballasts.

EPAct 1992 mandated energy-efficiency standards for linear fluorescent lamps and incandescent reflector lamps sold in the U.S. that became effective in 1994 and 1995. Essentially, they require 4-ft., 8- and 8-ft. high output T12 fluorescent lamps to either be reduced-wattage lamps or use rare-earth phosphors to comply. Incandescent reflector lamp efficacies must essentially use halogen technology (or better) to comply, although there are a variety of exempted reflector lamp categories. Two exempted lamp types, bulged reflector (BR) and elliptical reflector (ER) lamps, have gained in market share since the standards took effect in 1995. BR and ER lamps employ variations on reflector geometry to more efficiently focus light, although their efficacies are not as high as those of reflector lamps using halogen technology. There have been discussions at both the federal and state levels on incorporating ER and BR lamps into existing lamp standards. In 2006, DOE began a rulemaking to amend energy conservation standards for residential, commercial and industrial general service fluorescent lamps, incandescent reflector lamps and general service incandescent lamps.

In addition, EPAct 1992 mandated labeling for linear fluorescent, compact fluorescent, general service incandescent and incandescent reflector lamps. Labeling requirements took effect in 1994 and 1995. Compact fluorescent and incandescent lamps have their labels on the packaging; for linear fluorescents, the label is etched onto the lamp.

EPAct 2005 contains new regulations for several lighting products.

- Medium-based CFLs must meet the Energy star® specifications (version 2, August 2001), for minimum initial efficacy, lumen maintenance at 1000 h, lumen maintenance at 40 % of rated life, rapid cycle stress test, and lamp life. In addition, DOE may add requirements that color quality (CRI), power factor, operating frequency, and maximum allowable start time meet Energy star® version 2 specifications. This regulation took effect in 2006.
- Illuminated exit signs must meet Energy star® specifications (version 2), which require each sign to operate at 5 W or less. Essentially, signs whose light source is LEDs (or another low wattage technology) comply while rather those containing incandescent or fluorescent lamps do not comply. This regulation took effect in 2006.
- Traffic signal and pedestrian modules must meet Energy star®; specifications (version 1.1) that require LED technology. This regulation took effect in 2006.
- Torchieres cannot operate with lamps that total more than 190 W, which eliminates halogen lamps but allows incandescent lamps. This regulation took effect in 2006.
- Beginning in 2007, ceiling fan light kits using screw-base lamps must be packaged with lamps that meet the Energy star® requirements (version 2) for CFLs, or meet the Energy star® requirements (version 4) for residential light fixtures. DOE may also set standards for ceiling fan light kit niche products, such as those with candelabra base or halogen lamps, to take effect in 2007.
- Beginning in 2008, MV lamp ballasts cannot be manufactured or imported.

Information and updates to U.S. lighting equipment standards are available on DOE's website http://www.eere.energy.gov/buildings/appliance_standards.

Some states have enacted standards on lighting products that are not covered by DOE efficiency standards. For information on state standards, see the Appliance Standards Awareness Project website http://www.standardsasap.org/. Note that in the U.S. federal regulatory coverage of a product preempts states from setting standards on that product, hence some state lighting equipment standards are affected by the new EPAct 2005 standards.

EPAct 1992 required states to adopt a building energy code at least as stringent as the ASHRAE/IESNA[9] 90.1-1999 building code by July 15, 2004 [1,4].[10,11] Legislation that provides compliance with this statue varies from state to state. Some states have adopted their own building codes, which may contain provisions more stringent than those in ASHRAE/IESNA 90.1-1999 or 2001, while others have not yet upgraded their building codes. The 2003 International Code Council's energy code contains a reference to ASHRAE/IESNA 90.1-2001. This energy code, named the International Energy Conservation Code (IECC), is used by many states to meet the EPAct building energy code requirement. These codes apply to new construction or additions and to new equipment in existing spaces. For information on DOE's Building Codes Program, see www.energycodes.gov. For more information on the status of state building codes, see the Building Codes Assistance Project website, http://www.bcap-energy.org. The ASHRAE/IESNA 90.1-1999 standard may be ordered from ASHRAE (http://www.ashrae.org/). The IECC 2003 code may be ordered through ICC (http://www.iccsafe.org/).

Various states and local jurisdictions have developed "dark sky" ordinances that regulate outdoor lighting. These regulations may include limits on light trespass, sky glow, pole heights, or energy use, as well requiring automatic shutoff controls. For more information, see the International Dark Sky Association website, http://www.darksky.org/ordsregs/odl-regs.html. ASHRAE/IESNA 90.1, IECC, and California's Title 24 building codes have outdoor lighting provisions related to energy use that have evolved from very simple to more complex as the standards have been revised. One innovation in the latest revision of California Title 24 establishes four exterior lighting zones and sets power limits per zone according to how much light is needed. Lighting zone 1 (parks, recreation areas, wildlife preserves) is designated as "dark"; lighting zone 2 (rural areas) is designated for low ambient illumination; lighting zone 3 (urban areas) is designated for medium ambient illumination; and lighting zone 4 is designed for high ambient illumination. Methods for calculating the alllowed lighting power levels are specified in the code for each zone.

Programs that promote consumer purchase of energy-efficient technology through setting voluntary energy-efficiency targets, notably the Energy star® program conducted by the U.S. Environmental Protection Agency (EPA) and DOE, have been an effective incentive for energy-efficient lighting. As noted above, some Energy star® qualifications have become the basis for lighting standards in EPAct 2005, for the FEMP procurement guidelines described below, and for various state standards and public utility incentive programs. Information on product qualifications may be found on the Energy star® website www.energystar.gov.

Lighting products for federal buildings often must meet specific standards or procurement guidelines. DOE's Federal Energy Management Program (FEMP) has energy efficiency recommendations for government procurement for several lighting products, including lamps, ballasts, exit signs, luminaires, and controls. Many of FEMP's recommendations for energy using products match the qualifications for the Energy star® products, and must be selected when life-cycle costs warrant such selection. For more information on the equipment procurement product energy-efficient recommendations, see http://www.eere.energy.gov/femp/.

12.2.7 Cost-Effectiveness of Efficient Lighting Technologies

The energy savings from energy-efficient lighting systems provide energy cost savings over the lifetime of the equipment. In some cases, maintenance savings also accrue because the lighting systems have longer useful lives. Consumers and decision-makers weigh these costs against the higher first cost of

[9]ASHRAE = American Society of Heating Refrigerating and Air Conditioning Engineers. IESNA = Illuminating Engineering Society of North America.

[10]As updated by the Department of Energy's reference in the Federal Register, July 15, 2002.

[11]The original ASHRAE Standard 90 was published in 1975 and revised editions were published in 1980, 1989, 1999, 2001, and 2004.

TABLE 12.5 Cost-Effectiveness of Efficient Lighting Technologies

Product	Baseline Technology	Efficient Technology	Cost of Conserved Energy (CCE) in c/kWh
Torchiere (residential)	Incandescent	Fluorescent	6.8
Fluorescent lamp/ballast (commercial)	Current practice[a]	High performance T8 with high performance ballast	1.1

[a] F40T12 reduced wattage lamp (34 W) with magnetic ballast (15% market share); standard T8 lamp with electronic ballast (85% share).

the efficient technologies. Consumer choices regarding what types of efficient lighting technologies to purchase also depend on the lighting quality provided by the alternative system.

To fully assess the financial costs and benefits of an alternative lighting system, required sophisticated economic analysis is required. In addition to the costs of equipment and installation, one should consider energy, relamping, and maintenance costs over the life of the lighting system, as well as disposal costs. Evaluation of lighting quality and its potential economic impacts on user satisfaction and productivity is also important.

A study for the National Commission on Energy Policy (NCEP) estimated their costs and energy savings of appliances and lighting equipment relative to baseline designs [15]. For a number of products, it determined the efficiency level that provides users with the lowest life-cycle cost over the equipment lifetime. Table 12.5 lists the selected efficiency levels along with their associated cost of conserved energy (CCE) for consumers. CCE is an expression of the extra first cost incurred to save a unit of energy, with units of cents per 1 kWh. The CCEs in Table 12.5 are calculated assuming that equipment is purchased in the year 2010. The CCE may be compared to expected electricity prices. Long term forecasts for average residential electricity prices for the period studied are 8.5–9.0 cents/kWh [7], which is higher than the CCE for fluorescent torchieres, making the investment cost-effective. For the commercial sector lighting, marginal electricity prices are forecast between 6.8 and 7.5 cents/kWh [7], much higher than the CCE for high-performance T8 lamp/ballast systems.

12.2.8 Conclusion

In this chapter, we have provided an overview of energy-efficient lighting design practices as well as both traditional and newer, more efficient lighting technologies. Lighting is an important electrical end use in all sectors in the United States, and accounts for approximately one-quarter of national electricity use for commercial plus residential buildings. Through the use of more efficient lighting technologies as well as advanced lighting design practices and control strategies, there is significant potential for saving electricity, reducing consumer energy costs, and reducing the emission of greenhouse gases associated with electricity production. In addition, efficient lighting technologies and design can improve the quality of light in both the workplace and the home.

Glossary

Ballast: A lamp ballast is an electrical device used to control the current provided to the lamp. In most discharge lamps, a ballast also provides the high voltage necessary to start the lamp.

Ballast Factor: the fractional (luminous) flux of a fluorescent lamp operated on a ballast compared to the flux when operated on the standard (reference) ballast specified for rating lamp lumens.

Color Rendering Index (CRI): A measure of how surface colors appear when illuminated by a lamp compared to how they appear when illuminated by a reference source of the same color temperature. For color temperature above 5000 K, the reference source is a standard daylight condition of the same color temperature; below 5000 K, the reference source is a blackbody radiator.

Color Temperature: The color of a lamp's light expressed in degrees Kelvin (K). The concept of color temperature is based on the fact that the emitted radiation spectrum of a black-body radiator depends on temperature alone. The color temperature of a lamp is the temperature at which an ideal blackbody radiator would emit light that is the same color as the light of the lamp.

Efficacy: The ratio of the amount of light emitted (lumens) to the power (watts) drawn by a lighting system. The unit used to express efficacy is lumens per watt (LPW). Efficacy may be expressed as *lamp efficacy*, using the nominal wattage of the lamp, or as *system efficacy*, using the system watts that include the ballast losses.

Lumen Maintenance: The extent to which a lamp sustains its lumen output (and therefore efficacy) over time.

Acknowledgments

The authors would like to thank Francis Rubinstein and Steve Johnson of the Lighting Systems Research Group, Building Technologies Program at the Lawrence Berkeley National Laboratory for providing information that was useful in the preparation of this chapter.

References

1. ASHRAE. 1999. *ASHRAE/IESNA Standard 90.1-1999, Energy Standard for Buildings Except Low-Rise Residential Buildings, I-P Edition*, American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc., approved by the ASHRAE Standards Committee June 19, 1999, and by the ASHRAE Board of Directors June 24, 1999, Atlanta, GA.
2. CEE. 2005. *CEE High-Performance T8 Specification*, Consortium for Energy Efficiency, Inc. (CEE), Boston, MA. http://www.cee1.org/com/com-lt/com-lt-specs.pdf
3. Census. 1988–2004. *Current Industrial Reports, Fluorescent Lamp Ballasts (MQ36C, MQ335C)*. http://www.census.gov/cir/www/335/mq335c.html
4. DOE. 2002. *Determination Regarding Energy Efficiency Improvements in the Energy Standard for Buildings, Except Low-Rise Residential Buildings, ASHRAE/IESNA Standard 90.1-1999; Final Rule.* U.S. Department of Energy—Office of Energy Efficiency and Renewable Energy, Building Energy Standards Program. Washington, DC, Federal Register, 67(135), pp. 46464–46488.
5. DOE. 2004. *Technical Support Document: Energy Efficiency Program for Commercial and Industrial Equipment. High-Intensity Discharge Lamps, Analysis of Potential Energy Savings*. Draft. U.S. Department of Energy—Energy Efficiency Program for Commercial and Industrial Equipment, Washington, D.C. http://www.eere.energy.gov/buildings/appliance_standards/commercial/high_intensity_lamps.html
6. DOE/EIA. 2005a. *Buildings Energy Databook, August 2005.* U.S. Department of Energy/Energy Information Administration (EIA). Washington, D.C. http://buildingsdatabook.eere.energy.gov/ (Tables 1.1, 1.2, 1.3.)
7. DOE/EIA. 2005b. *Annual Energy Outlook 2005*. U.S. Department of Energy/Energy Information Administration. Washington, D.C.
8. IESNA. 2002. Rea, M.S. *The IESNA Lighting Handbook: Reference and Application, 9th Edition*, Illuminating Engineering Society of North America, New York.

9. LRG. 2003. Personal communication, Lighting Research Group, Building Technologies Department, Environmental Energy Technologies Division. Lawrence Berkeley National Laboratory, Berkeley, CA.
10. NBI. 2003. *Advanced Lighting Guidelines.* New Buildings Institute, White Salmon, WA.
11. NCI. 2002. *U.S. Lighting Market Characterization Volume 1: National Lighting Inventory and Energy Consumption Estimate*. Navigant Consulting, Inc. for the U.S. Department of Energy, Washington, D.C.
12. NEMA. 1998. *Procedure for Determining Luminaire Efficacy Ratings for High-Intensity Discharge Industrial Luminaires*. National Electrical Manufacturers Association, Washington, D.C. Standards Publication LE 5B-1998.
13. NEMA. 1999. *Procedure for Determining Luminaire Efficacy Ratings for Commercial, Non-Residential Downlight Luminaires*. Standards Publication LE 5A-1999.
14. NEMA. 1999. *Procedure for Determining Luminaire Efficacy Ratings for Fluorescent Luminaires*. National Electrical Manufacturers Association, Washington, D.C. Standards Publication LE 5A-2001.
15. LBNL. 2004. Rosenquist, G., Mcneil, M., Iyer, M., Meyers, S. and McMahon, J. E. *Energy Efficiency Standards for Residential and Commercial Equipment: Additional Opportunities*. Lawrence Berkeley National Laboratory, Berkeley, CA. LBNL-56207.
16. Walerczyk, S. 2005. *So Watt if Already T8*. Pacific Energy Center, Lighting Class PowerPoint Presentation, October 2005, San Francisco, CA.

Other Information Sources

E Source 2003. Sandinsky, R. and Benya, J. 2003. *Super T8s: Super Lamps, Super Ballasts*. ESource, Boulder, CO. ER-03-16.

National Lighting Product Information Program, Lighting Research Center, Troy, New York. http://www.lrc.rpi.edu/programs/NLPIP/index.asp

12.3 Energy Efficient Technologies: Major Appliances and Space Conditioning Equipment

James E. McMahon, Greg Rosenquist, James Lutz, Alex Lekov, Peter Biermayer, and Stephen Meyers

12.3.1 Introduction

Residential energy consumption accounts for approximately 37% of electricity use and 23% of natural gas use in the United States. In total, U.S. households spent more than $140 billion for home energy in 2001 and the average U.S. household spent almost $1,400 on its energy bill (Energy Information Administration (EIA) 2005a).

Major appliances and space conditioning equipment (which includes furnaces, boilers, heat pumps, and air conditioners) accounted for approximately 70% of U.S. residential energy consumption in 2003 (see Table 12.6).[1]

Reducing the energy consumption of residential appliances and space conditioning equipment depends on replacing older equipment with the much more efficient models that are now available and on continuing to design even more energy-efficient appliances. National energy efficiency standards for appliances have driven efficiency improvements over the last 20 years, and appliances have become significantly more efficient as a result. Further improvement of appliance efficiency represents a significant untapped technological opportunity.

[1]This chapter covers the equipment listed in Table 12.6. It does not cover the many other small appliances that have come to play a significant role in U.S. homes.

TABLE 12.6 U.S. Residential Energy Consumption by End Use in 2003

End Use	Electricity Use (EJ/yr)	Natural Gas Use (EJ/yr)	Other Energy Use (EJ/yr)[a]	Total Primary Energy Use (EJ/yr)	Share of Total Residential Primary Energy Use (%)
Space heating	0.46[b]	3.70	1.54	6.81	32.0
Space cooling	0.67[c]	0.00	0.00	2.18	10.2
Water heating	0.37	1.17	0.17	2.53	11.9
Refrigeration	0.40	0.00	0.00	1.30	6.1
Cooking	0.10	0.21	0.03	0.57	2.7
Clothes dryers	0.24	0.07	0.00	0.85	4.0
Freezers	0.13	0.00	0.00	0.42	2.0
Clothes washers	0.03	0.00	0.00	0.10	0.5
Dishwashers	0.02	0.00	0.00	0.08	0.4
Subtotal	2.42	5.15	1.74	14.8	69.6
Lighting	0.78	0.00	0.00	2.60	12.2
All other uses	1.17	0.10	0.17	3.89	18.3
Total	4.37	5.25	1.90	21.3	100

[a] Includes distillate, liquefied petroleum gas (LPG), and wood.
[b] Includes 3/4 of total furnace fan electricity use.
[c] Includes 1/4 of total furnace fan electricity use.
Source: From Energy Information Administration (EIA). 2005a. *Household Energy Consumption and Expenditures 2001*, U.S. Department of Energy, Washington, DC; Energy Information Administration (EIA). 2005b. *Annual Energy Outlook 2005*, U.S. Department of Energy, Washington, DC.

12.3.2 Description of Major Appliances and Space Conditioning Equipment

12.3.2.1 Refrigerator-Freezers and Freezers

Refrigerators, refrigerator-freezers, and freezers keep food cold by transferring heat from the air in the appliance cabinet to the outside. A refrigerator is a well-insulated cabinet used to store food at 0°C or above; a refrigerator-freezer is a refrigerator with an attached freezer compartment that stores food below −13°C; and a standalone freezer is a refrigerated cabinet to store and freeze foods at −18°C or below. Almost all refrigerators are fueled by electricity. The refrigeration system includes an evaporator, a condenser, and a compressor. The system uses a vapor compression cycle, in which the refrigerant changes phase (from liquid to vapor and back to liquid again) while circulating in a closed system. The refrigerant absorbs or discharges heat as it changes phase. Although most refrigerants and insulating materials once contained chlorofluorocarbons (CFCs), all U.S. models sold after January 1, 1996, are CFC-free.

Almost all households have a refrigerator-freezer and many have more than one. Thirty-two percent of households had stand alone freezers in 2001. Ambient temperature is a significant determinant of the energy consumed by these appliances, but user behavior has relatively little effect on the energy consumption of these appliances (Meier 1995).

12.3.2.2 Water Heaters

A water heater is an appliance that is used to heat potable water for use outside the heater upon demand. Water heaters supply water to sinks, bathtubs and showers, dishwashers, and clothes washing machines. Most water heaters in the United States are storage water heaters, which continuously maintain a tank of water at a thermostatically controlled temperature. The most common storage water heaters consist of a cylindrical steel tank that is lined with glass in order to prevent corrosion. Most hot water tanks manufactured today are insulated with polyurethane foam and wrapped in a steel jacket. Although some use oil, almost all storage water heaters are fueled by natural gas (or LPG) or electricity.

Rather than storing water at a controlled temperature, instantaneous water heaters heat water as it is being drawn through the water heater. Both gas-fired and electric instantaneous water heaters are available. Instantaneous water heaters are quite popular in Europe and Asia. Although they are not commonly used in the United States, their presence does seem to be increasing. No heaters of this type are manufactured in the United States.

Like refrigerators, water heaters are present in almost all U.S. households. Approximately 54% of households have gas-fired water heaters, and approximately 38% have electric water heaters. Hot water use varies significantly from household to household, mostly due to differences in household size and occupant behavior.

12.3.2.3 Furnaces and Boilers

Furnaces and boilers are major household appliances used to provide central space heating. Both fuel-burning and electric furnaces and boilers are available. A typical gas furnace installation is composed of the following basic components: (1) a cabinet or casing; (2) heat exchangers; (3) a system for obtaining air for combustion; (4) a combustion system including burners and controls; (5) a venting system for exhausting combustion products; (6) a circulating air blower and motor; and (7) an air filter and other accessories. (Furnaces that burn oil and liquid petroleum gas (LPG) are also available, though not as common.) In an electric furnace, the casing, air filter, and blower are very similar to those used in a gas furnace. Rather than receiving heat from fuel-fired heat exchangers, however, the air in an electric furnace receives heat from electric heating elements. Controls include electric overload protection, contactor, limit switches, and a fan switch. Furnaces provide heated air through a system of ducts leading to spaces where heat is desired. In a batter system, hot water or steam is piped to terminal heating units placed throughout the household. The boiler itself is typically a pressurized heat exchanger of cast iron, steel, or copper in which water is heated.

Natural gas was the primary space-heating fuel in approximately 55% of U.S. households in 2001, and electricity was the primary space-heating fuel in 29% of households. The amount of energy consumed by a household for space heating varies significantly with climate and various household characteristics.

12.3.2.4 Central and Room Air Conditioners

A central air conditioning (AC) system is an appliance designed to provide cool air to an enclosed space. Typically, central AC systems consist of an indoor unit and an outdoor unit. The outdoor unit contains a compressor, condenser (outdoor heat exchanger coil), condenser fan, and condenser fan motor; the indoor unit consists of an evaporator (indoor conditioning coil) and a flow control device (a capillary tube, thermostatic expansion valve, or orifice) residing either in a forced-air furnace or an air handler. Refrigerant tubing connects the two units. A central AC system provides conditioned air by drawing warm air from the space and blowing it through the evaporator; as it is passing through the evaporator, the air gives up its heat content to the refrigerant. The conditioned air is then delivered back to the space (via a ducted system) by the blower residing in the furnace or air handler. The compressor takes the vaporized refrigerant aiming out of the evaporator and raises it to a temperature exceeding that of the outside air. The refrigerant then passes on to the condenser (outside coil), where the condenser fail blows outside air over it, gives up its heat to the cooler outside air, and condenses. The liquid refrigerant is then taken by the flow control device and its pressure and temperature are reduced. The refrigerant reenters the evaporator, where the refrigeration cycle is repeated.

Unlike the two-unit, central AC system, a room air conditioner is contained within one cabinet and is mounted in a window or a wall so that part of the unit is outside and part is in. The two sides of the cabinet are typically separated by an insulated divider wall in order to reduce heat transfer. The components in the outdoor portion of the cabinet are the compressor, condenser, condenser fan, fan motor, and capillary tube. The components in the indoor portion of the cabinet are the evaporator and evaporator fan. The fan motor drives both the condenser and evaporator fans. A room AC provides conditioned air in the same manner described for a central AC system.

Approximately 78% of U.S. households had an AC system in 2001. Central AC systems (including heat pumps) are used in 54% of households; room ACs are used in 23% of households.

12.3.2.5 Heat Pumps

Unlike air conditioners, which provide only air source space cooling, heat pumps use the same equipment to provide both space heating and cooling. An air source heat pump draws heat from the outside air into a building during the heating season and removes heat from a building to the outside during the cooling season. An air source heat pump contains the same components and operates in the same way as a central AC system but is able to operate in reverse as well, in order to provide space heating. In providing space heat, the indoor coil acts as the condenser while the outdoor coil acts as the evaporator. When the outside air temperature drops below 2°C during the heating season, a heat pump will utilize supplementary electric-resistance backup heat. Heat pumps were used in approximately 11% of U.S. households in 2001. The energy consumption of heat pumps varies according to the same user characteristics discussed above for AC systems. A more detailed analysis of heat pumps is presented in Section 12.4.

12.3.2.6 Clothes Washers

A clothes washer is an appliance that is designed to clean fabrics by using water, detergent, and mechanical agitation. The clothes are washed, rinsed, and spun within the insulated cabinet of the washer. Top-loading washers move clothes up and down, and back and forth, typically about a vertical axis. Front-loading machines move clothes around a horizontal axis. Electricity is used to power an electric motor that agitates and spins the clothes, as well as a pump that is used to circulate and drain the water in the washer tub. A separate water heater is used to heat the water used in the washer (some clothes washers also have an internal electric water heater).

Approximately 79% of households had clothes washers in 2001. Most of the clothes washers sold in the United States are top-loading, vertical-axis machines. The majority of energy used for clothes washing (85%–90%) is used to heat the water. User behavior significantly affects the energy consumption of clothes washers. The user can adjust the amount of water used by the machine to the size of the load, and thereby save water and energy. Choosing to wash with cold water rather than hot water reduces energy consumption by the water heater. Similarly, rinsing with cold water rather than warm can reduce energy consumption. Energy consumption depends on how frequently the washer is used. The Department of Energy's test procedure assumes clothes washers are used 392 times a year on average.

12.3.2.7 Clothes Dryers

A clothes dryer is an appliance that is designed to dry fabrics by tumbling them in a cabinet like drum with forced-air circulation. The source of heated air may be powered either by electricity or natural gas. The motors that rotate the drum and drive the fan are powered by electricity. Approximately 57% of U.S. households had electric clothes dryers, and 16% had gas dryers in 2001.

12.3.2.8 Dishwashers

A dishwasher is an appliance that is designed to wash and dry kitchenware by using water and detergent. Typically in North America hot water is supplied to the dishwasher by an external water heater. In addition an internal electric heater further raises the water temperature within the dishwasher. Electric motors pump water through spray arms impinging on the kitchenware in a series of wash and rinse cycles. An optional drying function is also enabled by electric heaters and sometimes a fan. In recent years, some dishwashers incorporate soil sensors that determine when the dishes are clean and the washing cycle can be stopped. Approximately 53% of U.S. households had dishwashers in 2001.

12.3.2.9 Cooktops and Ovens

A cooktop is a horizontal surface on which food is cooked or heated from below; a conventional oven is an insulated, cabinet like appliance in which food is surrounded by heated air. When a cooktop and an oven are combined in a single unit, the appliance is referred to as a range. Both gas and electric ranges are

TABLE 12.7 Shipments of Major Appliances and Space Conditioning Equipment in the U.S.

Product	1995 (Million Units)	2003 (Million Units)
Refrigerators[a]	7.65	10.02
Freezers	1.46	2.52
Water heaters	8.37	9.58
Clothes washers	6.08	8.15
Clothes dryers	4.99	7.34
Dishwashers	4.30	6.28
Gas furnaces	2.60	3.27
Heat pumps	1.03	1.63
Central air conditioners	3.11	3.75
Room air conditioners	3.96	8.22

[a] Standard size; reported shipments of compact refrigerators in 2003 were 1.46 million.

available. Cooktops and ovens are present in almost all households. Almost 60% of households use electric cooktops and ovens, and the remaining 40% of households use gas cooktops and ovens.

In a microwave oven, microwaves directed into the oven cabinet cause water molecules inside the food to vibrate. Movement of the water molecules heats the food from the inside out. The fraction of households with microwave ovens has increased dramatically in recent years.

12.3.3 Current Production

Table 12.7 shows the number of various appliances that was shipped by manufacturers in 1995 and 2003. Shipments have been increasing for most of the major appliances.

12.3.4 Efficient Designs

State and federal standards requiring increased efficiency for residential appliances, utility programs, and labels (such as Energy star) have improved appliance efficiency dramatically since the late 1970s (Meyers et al. 2004). For example, the annual energy consumption (according to the DOE test procedure) of a new refrigerator in 2003 was less than half the consumption in 1980. Because of the slow turnover rate of appliances, however, the older, less efficient equipment remains in use for a long time. Promising design options for further improving the efficiency of residential appliances are discussed next.

12.3.4.1 Refrigerators and Freezers

Relative to the 2001 U.S. federal efficiency standard, achieving a 15% energy use reduction for refrigerator-freezers is possible with the use of a high-efficiency compressor, high-efficiency motors for the evaporator and condenser fans, and adaptive defrost control. Models at this level of efficiency account for a modest market share in the U.S. achieving a 25% energy use reduction generally would require a reduction in load transmitted through the unit's walls and doors, which might require the use of vacuum panels.

12.3.4.1.1 Improved Fan Motors

The evaporator and condenser fans of large refrigerators are powered by motors. The most common motor used for this purpose is a shaded-pole motor. Large efficiency gains are possible in refrigerators and freezers by switching to electronically commutated motors (ECMs), also known as brushless permanent-magnet motors, which typically demand less than half as much power as shaded-pole motors.

12.3.4.1.2 Vacuum Insulation Panels

The use of vacuum insulation panels (VIP) can significantly reduce heat gain in a refrigerated cabinet and thereby decrease the amount of energy necessary to maintain a refrigerator or freezer at a low temperature. When using VIP, a partial vacuum is created within the walls of the insulation panels. Because air is conductive, the amount of heat transfer from the outside air to the refrigerated cabinet is reduced as the amount of air within the panels is reduced. Evacuated panels are filled with low-conductivity powder, fiber, or aerogel in order to prevent collapse. Energy savings associated with the use of vacuum panel insulation range from 10 to 20%. Vacuum panel technology still faces issues regarding cost and reliability before it can come into widespread use in refrigeration applications (Malone and Weir 2001).

12.3.4.2 Water Heaters

12.3.4.2.1 Gas-Fired Storage Water Heater

The current models of gas-fired storage water heaters have a central flue that remains open when the water heater is not firing. This leads to large off-cycle standby losses. It should be possible to dramatically reduce off-cycle losses with relatively inexpensive technical modifications to the water heater. Energy savings derived from models with these modifications are expected to be about 25% compared to the 2004 U.S. standards.

The amount of heat extracted from the fuel used to fire a gas appliance can be increased by condensing the water vapor in the flue gases. In a condensing storage water heater, the flue is lengthened by coiling it around inside the tank. The flue exit is located near the bottom of the tank where the water is coolest. Because the flue gases are relatively cool, a plastic venting system may be used. A drain must be installed in condensing systems. Energy savings associated with the use of a gas-fired condensing water heater are approximately 40% compared to the 2004 efficiency standard. At this time, the high cost of the water heater results in a payback time that exceeds the typical lifetime of a water heater, but it is reasonable to assume that the cost could be reduced to the point that the water heater would be cost-effective. In applications with heavy hot water use, such as laundromats or hotels, they may already be cost-effective. Currently, a few companies produce condensing storage water heaters for commercial markets, and they are sometimes sold as combined water heater/space heating systems for residential use.

12.3.4.2.2 Heat Pump Water Heaters

Heat pumps used with water heaters capture heat from the surrounding air or recycle waste heat from AC systems and then transfer the heat to the water in the storage tank. In this way, less energy is used to bring the water to the desired temperature. The heat pump can be a separate unit that can be attached to a standard electric water heater. Water is circulated out of the water heater storage tank, through the heat pump, and back to the storage tank. The pump is small enough to sit on the top of a water heater but could be anywhere nearby. Alternatively, the heat pump can be directly integrated into the water heater. Research indicates that this technology uses 60%–70% less energy than conventional electric resistance water heaters. Field and lab tests have been completed for several prototypes. A few models are currently available for sale.

12.3.4.2.3 Solar Water Heaters

Technological improvements in the last decade have improved the quality and performance of both passive and active solar water heaters. Research indicates that, in general, solar water heaters use 60% less energy than conventional electric resistance water heaters. There are several types of solar water heaters commercially available today.

12.3.4.3 Furnaces

12.3.4.3.1 Condensing Furnaces

The efficiency of a conventional gas furnace can be increased by using an additional heat exchanger to capture the heat of the flue gases before they are expelled to the outside The secondary heat exchanger is typically located at the outlet of the circulating air blower, upstream of the primary heat exchanger. A floor drain is required for the condensate. A condensing furnace has an efficiency up to 96% annual fuel utilization efficiency (AFUE), well above the 80% AFUE rating of a standard noncondensing gas furnace. Condensing furnaces have been on the market since the 1980s and now constitute one-third of all gas furnace sales. They are particularly popular in colder areas of the United States, where the cost of heating is high. An early technical problem, corrosion of the secondary heat exchanger, has been resolved by the industry.

12.3.4.3.2 Integrated Water Heaters and Furnaces

Traditionally, water heating and space heating have required two separate appliances—a hot water heater and a furnace. Combining a water heater and a furnace into a single system can potentially provide both space heating and hot water at a lower overall cost. Integrated water and space heating is most cost-effective when installed in new buildings because gas connections are necessary for only one appliance rather than two.

Combination space- and water-heating appliances fall into two major classes: (1) boiler/tankless-coil combination units, and (2) water-heater/fancoil combination units. A great majority of boiler/tankless-coil combination units are fired with oil, whereas most water-heater/fancoil combination units are fired with natural gas. In the latter units, the primary design function is domestic water heating. Domestic hot water is circulated through a heating coil of an air-handling system for space heating. Usually the water heater is a tank-type gas-fired water heater, but instantaneous gas-fired water heaters can be used as well.

The efficiencies of these integrated systems are determined largely by the hot water heating component of the system. Compared to a system using a standard water heater or boiler, an integrated system using a condensing water heater or condensing boiler can reduce energy consumption by as much as 25%.

12.3.4.4 Central and Room Air Conditioners

12.3.4.4.1 Electric Variable-Speed Air Conditioning

Variable-speed central air conditioners use ECMs, which are more efficient than the induction motors used in a single-speed system. In addition, the speed of the ECM can be varied to match system capacity more precisely to a building load. Cycling losses, which are associated with a system that is continually turned off and on in order to meet building load conditions, are thus reduced. Unlike induction motors, ECMs retain their efficiency at low speeds; consequently, energy use is also reduced at low-load conditions. A variable-speed AC system uses approximately 40% less energy than a standard single-speed AC system. Although these AC systems are now available from major manufacturers, they account for a small fraction of sales.

12.3.4.4.2 Electric Two-Speed Air Conditioning

Two-speed induction motors are not as efficient as variable-speed ECMs, but they are less expensive. Like variable-speed air conditioners, two-speed air conditioners reduce cycling losses. When two-speed induction motors are used to drive compressor and fans, the system can operate at two distinct capacities. Cycling losses are reduced because the air conditioner can operate at a low speed lo meet low building loads. In some models, two-speed compressors are coupled with variable-speed indoor blowers to improve system efficiency further. A two-speed AC system reduces energy consumption by approximately one-third. Although these AC systems are available from several major manufacturers, they account for a small fraction of sales.

12.3.4.4.3 Room Air Conditioners

The most efficient room air conditioners have relatively large evaporator and condenser heat-exchanger coils, high-efficiency rotary compressors, and permanent split-capacitor fan motors. Compared to standard room ACs, highly-efficient room ACs reduce energy consumption by approximately 25%. Such room ACs are available from several manufacturers. Units with Energy star designation, which must exceed federal minimum efficiency standards by 10%, accounted for 35% of sales in 2004.

12.3.4.5 Clothes Washers and Dryers

12.3.4.5.1 Horizontal-Axis Washers

Although horizontal-axis clothes washers dominate the European market, the vast majority of clothes washers sold in the United States are top-loading, vertical-axis machines. However, this is expected to change by 2007 when more stringent minimum efficiency regulations on clothes washers will take effect. Horizontal-axis washers, in which the tub spins around a horizontal axis, use much less water than their vertical-axis counterparts, and less hot water is therefore required from water heaters. As mentioned earlier, the majority of energy used for clothes washing is used for heating water, so a significant amount of energy can be saved by using horizontal-axis washers. Research has indicated that horizontal-axis washers are more than twice as efficient as vertical-axis washers of comparable size (U.S. Department of Energy 2000).

12.3.4.5.2 High-Spin-Speed Washers

Clothes washers can be designed so that less energy is required to dry clothes after they have been washed. Extracting water from clothes mechanically in a clothes washer uses approximately 70 times less energy than extracting the water with thermal energy in an electric clothes dryer. Thus, by increasing the speed of a washer's spin cycle, one can reduce the energy required to dry clothes.

Because gas clothes dryers require so much less energy than electric dryers, based on energy consumption including that consumed at the electric power station, energy savings are much more significant when a high-spin-speed washer is used with an electric dryer. In a vertical-axis clothes washer, an increase in spin speed from 550 to 850 rpm reduces moisture retention from 65 to 41%. In an electric dryer, this reduces the energy consumption by more than 40%. In a horizontal-axis washer, an increase in spin speed from 550 to 750 rpm reduces moisture retention from 65 to 47%; in an electric dryer, energy consumption is reduced by more than 30%. High-spin-speed washers have been common on European horizontal axis machines for some time, and are becoming more common in the U.S.

12.3.4.5.3 Microwave Dryers

In conventional clothes dryers, hot air passes over wet clothes and vaporizes the surface water. During the later stages of drying, the surface dries out and heat from the hot air must be transferred to the interior, where the remaining moisture resides. In contrast, in microwave drying, water molecules in the interior of a fabric absorb electromagnetic energy at microwave wavelengths, thereby heating the water and allowing it to vaporize. Several U.S. appliance manufacturers have experimented with microwave clothes dryers, and a few small companies have built demonstration machines. Issues of small metal objects heating, such as rivets, would need to be resolved before microwave clothes dryers are likely to be sold.

12.3.4.5.4 Heat Pump Dryers

A heat pump dryer is essentially a clothes dryer and an air conditioner packaged as one appliance. In a heat pump dryer, exhaust heat energy is recovered by recirculating all the exhaust air back to the dryer; the moisture in the recycled air is removed by a refrigeration–dehumidification system. A drain is required to remove the condensate; because washers and dryers are usually located side by side, a drain is generally easily accessible. Heat pump dryers can be 50%–60% more efficient than conventional electric dryers. Introduced in Europe in 1999, heat pump dryers are available both in the U.S. and Europe.

12.3.4.6 Cost-Effectiveness of Energy-Efficient Designs

The cost-effectiveness of energy-efficient designs for major appliances and space conditioning equipment depends on consumer energy prices and also on per-unit costs when new technologies are manufactured on a large scale. Therefore, the cost-effectiveness of such designs varies

A U.S study, conducted in 2004 for the National Commission on Energy Policy (NCEP) reviewed the literature on energy efficient designs for major appliances and space conditioning equipment and estimated their costs (in the year 2010) and energy savings relative to baseline designs (Rosenquist et al. 2004). For a number of products, it determined the efficiency level that provides users with the lowest

TABLE 12.8 Cost-Effectiveness of Efficient Designs for Selected Appliances and Space Conditioning Equipment

Product	Baseline Technology	Efficient Technology	Cost of Conserved Energy[a]
Refrigerator	484 kWh/yr	426 kWh/yr	4.9 c/kWh
Room air conditioner	9.85 EER	10.11 EER	5.2 c/kWh
Electric water heater	92 EF	Heat pump	3.9 c/kWh
Clothes washer[b]	3.23 kWh/cycle	Horizontal-axis and high spin speed (1.87 kWh/cycle)	5.0 c/kWh
Gas furnace[c]	80% AFUE	Condensing furnace (90% AFUE)	$9.30/MBtu

[a] The CCE may be compared to expected residential energy prices. In the U.S., long-term forecasts from the Department of Energy (as of 2005) place average prices in the range of 8.5–9.0 cents/kWh for electricity and $7.7–8.4/million Btu for natural gas.

[b] Includes the energy for water heating and drying associated with clothes washing. Refers to electric water heater and dryer. Baseline refers to the 1994 DOE standard level. Additional consumer benefits derive from reduction in water usage.

[c] The CCE value is a U.S. national average. In cold climates, the amount of saved energy is higher, so the CCE is lower by a third or more.

life-cycle cost over the appliance lifetime. Table 12.8 lists the selected efficiency levels (expressed in terms of annual energy consumption in some cases), along with their associated cost of conserved energy (CCE) for consumers. The CCE in terms of dollars per kWh or MMBtu (gas) is an expression of the extra first cost incurred to save a unit of energy. Calculation of CCE requires application of a present worth factor (PWF) to spread the initial incremental cost over the lifetime of the equipment. The PWF uses a discount rate to effectively amortize costs over time. The CCEs for electric appliances in Table 12.8 are all below the average U.S. residential electricity price.

12.3.5 Conclusion

Residential appliances consume significant amounts of electricity and natural gas. This chapter described the basic engineering principles of the major appliances and space conditioning equipment as well as promising energy-efficient designs.

Significant potential energy savings are possible beyond the models typically sold in the marketplace today. Among the various end uses, energy savings range from 10 to 50%. Many of these efficient appliances appear to be cost-effective at currently projected manufacturing costs, with simple payback times that are shorter than the typical appliance lifetimes of 10–20 years. If costs of efficiency improvements decrease, or if future energy prices increase, more of the potential energy savings that are already technically possible will become economically attractive. In addition, future research is likely to identify additional technological opportunities to save energy.

Acknowledgments

This work was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Buildings Technology, Office of Code and Standards, of the U.S. Department of Energy, under Contract No. DE-AC03-76SF00098. The opinions expressed in this paper are solely those of the authors and do not necessarily represent those of Ernest Orlando Lawrence Berkeley National Laboratory, or the U.S. Department of Energy.

References

1. Energy Information Administration (EIA), 2005a. Household Energy Consumption and Expenditures 2001. U.S. Department of Energy, Washington, DC
2. Energy Information Administration (EIA) 2005(b). *Annual Energy Outlook 2005*. U.S. Department of Energy, Washington, DC

3. Malone, B. and Weir, K. 2001. State of the Art for VIP Usage in Refrigeration Applications, *International Appliance Manufacturing*.
4. Meier, Alan 1995. Refrigerator energy use in the laboratory and the field. *Energy and Buildings*, 22(3), 233–243.
5. Meyers, S., McMahon, J., and McNeil, M. 2005. *Realized and Prospective Impacts of U.S. Energy Efficiency Standards for Residential Appliances: 2004 Update*. Lawrence Berkeley National Laboratory, Berkeley, CA. Report No. LBNL-56417.
6. Rosenquist, G., McNeil, M., Iyer, M., Meyers, S., and McMahon, J. 2004. *Energy Efficiency Standards for Residential and Commercial Equipment: Additional Opportunities*. Lawrence Berkeley National Laboratory, Berkeley, CA. Report No. LBNL-56207.
7. U.S. Department of Energy-Office of Building Research and Standards. 2000. Final Rule Technical Support Document (TSD): Energy Efficiency Standards for Consumer Products: Clothes Washers. U.S. Department of Energy, Washington, DC. Report No. LBNL-47462. http://www.eren.doe.gov/buildings/codes_standards/reports/cwtsd/index.html

12.4 Heat Pumps

Katherine Johnson and Frank Kreith

12.4.1 Basic Principles

Heat pumps are one of the most underused means of conserving energy for heating and cooling of buildings. The basic components of a heat pump are a working fluid or refrigerant, a gas compressor, two heat exchangers, piping, controls and accessories which can provide either heating or cooling to a building space. In the heating mode, heat is extracted from a natural or waste heat source and transferred to the space while in the cooling mode, heat is removed from the building and discharged to a heat sink.

There are four basic types of heat pumps, named air-to-air, water-to-air, water-to-water, and earth-to-air. In an air-to-air heat pump system, heat is removed from indoor air and rejected to the outdoor air of a building during the cooling cycle, while the reverse happens during the heating cycle. Water can replace the outdoors air as the source or sink for the heat, depending on whether the unit is in the heating or cooling mode. Air-to-air heat pumps or air source heat pumps (ASHP) is typically roof top units either completely packaged or split packaged systems. Split package heat pumps are designed with an air-handling unit located inside the conditioned space while the condenser and compressor are packaged for outdoor installation on the roof. ASHP are best suited for mild climates such as the southeastern part of the U.S. and areas where natural gas is either unavailable or expensive. For details on sizing, as well as protecting heat pumps operating in the heating mode when the ambient air temperatures drops below freezing see Ref. [1]

In water source heat pumps, instead of air, water is used to transfer the heat between the building and the outside. Geothermal heat pumps (discussed in more details in Section 12.4.2) use energy from the ground soil or ground water as the source or sink. In winter a geothermal heat pump transfers thermal energy from the ground to provide space heating. In the summer the energy transfer process is reversed. The ground absorbs thermal energy from the conditioned space and cools the air in the building. A GHP benefits from a nearly constant ground temperature year round, which is higher on average than winter air temperatures and lower on average than summer air temperatures. The energy efficiency of a GHP is thus higher than that of conventional ASHP's and some are also more efficient than fossil fuel furnaces in the heating mode. The primary difference between an ASHP and a GHP is the investment in a ground loop for heat collection and rejection required for the GHP system. Whether or not the GHP is cost-effective relative to a conventional ASHP depends upon generating annual energy cost savings that are high enough for the extra cost of the ground loop.

Figure 12.27a and Figure 12.27b illustrate schematically the basic mode of operation for cooling and heating a building. Figure 12.28a shows the different steps of operation on a pressure–enthalpy diagram for an ideal cycle in the refrigeration mode. An ideal cycle consists of the following four steps:

1. An isothermal evaporation process during which the refrigerant is completely evaporated and produces a refrigeration effect given by $Q_{rf} = (h_1 - h_4)$ where h_1 and h_4 are the enthalpies of the refrigerant at state point 1 and 4 respectively.
2. An isotropic compression process from 1 to 2 during which a compressor isotropically compresses the vapor refrigerant from state point 1–2. The work input for this step is given by $W_c = (h_2 - h_1)$
3. A constant pressure condensation process from state point 2–3 during which the hot gaseous refrigerant is first cooled to its saturation point and then condensed into liquid. The heat during the process is rejected to water, ambient air or the ground. The heat rejection is given by $Q_{23} = (h_2 - h_3)$
4. A throttling process from 3 to 4 during which the refrigerant flows through a throttling device, such as an orifice, and its pressure is reduced to the evaporation pressure. A portion of the liquid flashes into vapor and enters the evaporator at state point 4. This is a constant enthalpy irreversible process and in the ideal cycle $h_3 = h_4$.

The *coefficient of performance* (COP) is a dimensionless index used to indicate the performance of a refrigerator or a heat pump.

- In the refrigeration mode

$$\mathrm{COP_{ref}} = \frac{\text{Refrigeration effect}}{\text{Work input}} = \frac{h_1 - h_4}{h_2 - h_1}$$

- The COP of a heat pump that produces a useful heat input is given by

$$\mathrm{COP} = \frac{Q_{23}}{\text{Work input}}$$

For details on the design of individual components in a heat pump see refs. [1], [2], and [3] as well as publications of the American Society of Heating Refrigeration and Air Conditioning Engineering (ASHRAE e.g. Handbook of Refrigeration, 1998).

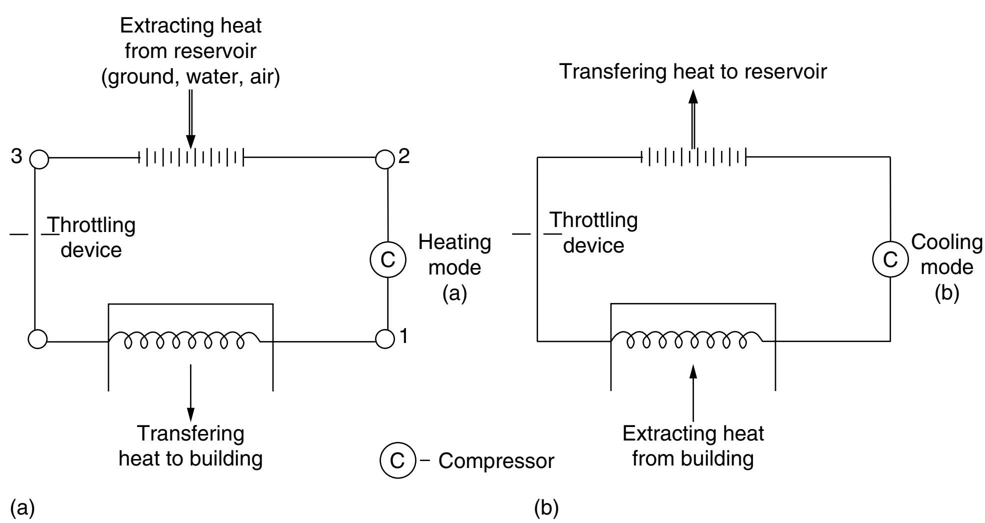


FIGURE 12.27 Schematic diagram of a heat pump in the heating mode (a) and cooling mode (b).

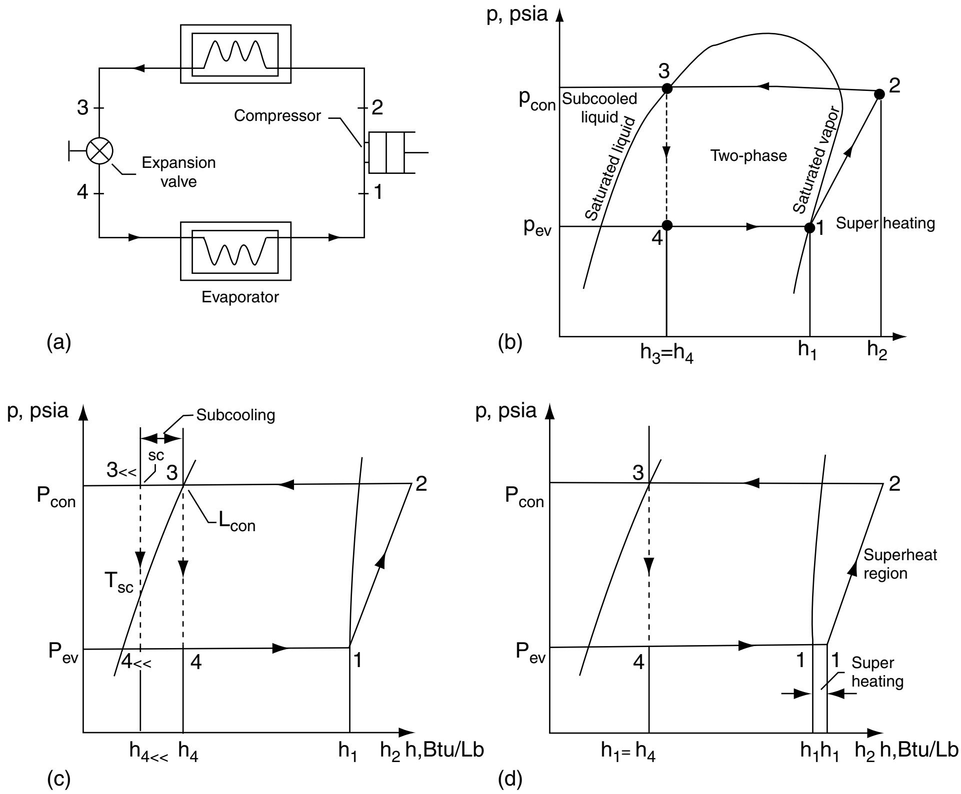


FIGURE 12.28 A single-stage ideal vapor compression refrigeration or heat pump cooling cycle; (a) schematic diagram (b) pressure enthalpy (p-h) diagram (c) subcooling, and (d) superheating.

At present, R22 is the most widely used working fluid in the air conditioning and heat pump industry, especially in residential unitary and air conditioning systems. However, the Montreal and Kyoto Protocols, as well as the U.S. Clean Air Act Amendments of 1990, which were based on the Montreal Protocol, forced the climate control industry to change its use of CFCs to HFCs. The phase-out of R22 compelled manufacturers to find alternatives. The air conditioning and refrigeration industry (ARI) started an R22 Alternative Refrigerants Evaluation Program to find and evaluate promising alternatives to the refrigerants containing ozone-depleting CFCs. The alternatives were to satisfy the requirements that refrigerants be: environmentally benign, nonflammable, chemically inert to avoid corrosion, non-toxic, and thermodynamically suitable. So far, no single refrigerant has been identified that can meet all of these requirements, but refrigerant mixtures are gaining importance as acceptable working fluids for future equipment. Ref. 4 presents the characteristics of mixtures and applications to heat pumps and other climate control equipment. The reader is referred to this book for detailed information on thermo-physical properties of mixtures of refrigerants and their applications in climate control systems.

The COP of real heat pumps is less than that of the ideal OCP because there must be a temperature difference between the working fluid and the heat sink a s well as the heat source. Also, the work input to the compressor is larger than that shown for the ideal cycle. Methods to increase the COP are discussed below.

Variable-Speed and Two-Speed Heat Pumps. Like central air conditioners, heat pumps can be made more efficient by the use of two-speed and variable-speed motors (see the earlier discussion of efficient central air conditioners). Both two-speed and variable speed air-source heat pumps are available. Compared to standard models, two-speed air-source heat pumps reduce energy consumption by approximately 27%, variable-speed air-source heat pumps reduce energy consumption by 35%, and two-speed ground-source heat pumps reduce energy consumption by 46%. Variable-speed and two-speed air-source heat pumps are made by the same companies that make variable speed and two-speed AC systems. Several manufacturers produce efficient air-source and two-speed ground-source heat pumps, but they account for a small fraction of all heat pump sales.

Gas-Fired Heat Pumps. Currently, all residential heat pumps are electric, but researchers have been developing gas heat pumps. The Gas Research Institute (GRI) and a private corporation jointly developed a natural gas, engine-driven, variable-speed heat pump in which the compressor is driven by an internal combustion spark-ignition engine and heat is recovered in the space-heating mode. This engine-driven heat pump was put on the market in 1994, but was withdrawn in the late 1990s due to lack of a maintenance infrastructure. In addition, the DOE has been funding the development of a gas-fired ammonia–water absorption-cycle heat pump. Gas-driven heat pumps have the potential to reduce heat pump energy consumption by approximately 35%–45%.

Distribution Systems. When assessing the efficiency of a space conditioning system, it is important to consider the efficiency of the distribution system as well as the appliance. It is not uncommon for air ducts to have distribution losses of 20%–40% due to conduction as well as leakage. Better insulation as well as more careful duct sealing can reduce these losses, In general, the most effective strategy for reducing distribution losses is to include the distribution system in the conditioned space so that any losses due to conduction or air leakage go directly into the space to be conditioned. This requires careful attention by the architect in the design of new buildings and is typically expensive as a retrofit measure.

12.4.2 Solar-Assisted Heat Pump Systems

A solar collector can improve the performance of either a liquid or an air system, but it adds complexity and expense. Low-temperature solar heat applied to one side of the heat-pump system (Figure 12.29) evaporates the low-pressure refrigerant liquid. The compressor then raises the pressure and temperature of the vapor which, when it condenses, gives off heat at a higher temperature than that at which heat was provided. When this temperature difference is less than 30°F–40°F (20°C), a good heat pump can provide heat at 105°F–115°F with a COP of about 3.5.

Solar-assisted heat-pump systems can be of several types[5]. One type uses liquid in the solar collector loop, water storage, and a water-to-water heat pump. Another type uses an air-to-air heat pump in

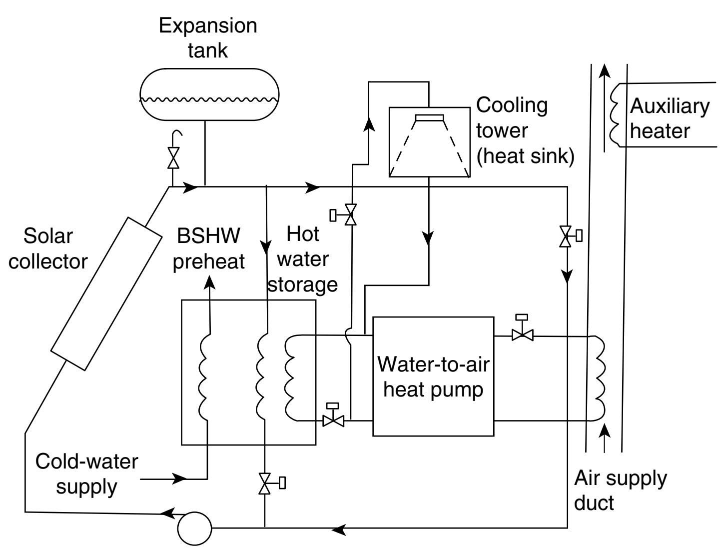


FIGURE 12.29 Schematic diagram for a solar collector assisted heat pump.

conjunction with liquid-to-air heat exchangers and a liquid solar collector loop. Although solar-assisted heat pumps are normally used for heating only, the concept of a solar-assisted, double-bundle condenser, heat recovery chiller can be included in the heat-pump category[5]. In the summer, this system would operate as a conventional central station, chilled-water air-conditioning system; however, it could also operate at night under more favorable ambient conditions and possible reduced off-peak power rates, and store the "coolness" in its storage tank. In the winter, solar heat, as well as heat from people and lights in the interior zone of a large building, would create the load on the chiller evaporator. This system, suitable for large buildings, has the advantages of a heat pump without the complications associated with the reversing cycle features.

Heat pumps enhance the efficiency and lower the costs of the solar energy system by permitting the collectors to operate at low fluid temperatures with the heat pump boosting the air or water temperatures for delivery to the space at 105°F–115°F. The space-heating system must be designed for those low utilization temperatures. In large buildings the heat pump always utilizes the solar-heated water for a heat source to maintain high COP and capacity. A computer analysis is required to optimize the operating mode, collector and storage size, number of storage vessels, and control sequence. In some cases it may be more efficient to use all of the solar-heated hot water at low temperature as a source for the heat pump.

12.4.3 Geothermal Heat Pumps

Geothermal heat pumps (GHPS) are also called ground source heat pumps, earth source/earth coupled heat pumps, geothermal heating and cooling systems, direct exchange, and "geo" among others. Developing a common name was an early mission of the Geothermal Heat Pump Consortium (GHPC), a nonprofit industry association based in Washington, D.C. Their efforts led to creating a national "brand" or identity for this technology, now referred to as "GeoExchange."

Every geothermal system consists of three major elements: 1) a *geothermal heat pump* to move heat between the building and the fluid in the earth connection, 2) an *earth connection* for transferring heat between its fluid and the earth, and 3) a *distribution subsystem* for delivering heating or cooling to the building. To heat a building, the heat is extracted from the fluid in the earth connection by the geothermal heat pump and distributed through a system of air ducts. Cooler air from the building is returned to the geothermal heat pump, where it cools the fluid flowing to the earth connection Figure 12.30. The fluid is warmed again as it flows through the earth connection. The process is reversed to cool the building (Figure 12.31).

The ground loop may be installed either directly in the ground or through a well. The fluid could be either water or a refrigerant. Geothermal heat pumps can provide heating, cooling, and even hot water, at a significantly lower cost compared to conventional systems. However, geothermal systems installations usually cost two to three times more compared to conventional systems. These costs are higher because geothermal installations requiring working with three different kinds of specialists:

- Geothermal contractor
- Driller
- Loop installer.

Three types of configurations are used in installing a geothermal heat pump system. The type selected depends upon a number of factors including soil type, availability of a water source, the size of the installation, and the amount of land available. Loops may be installed horizontally or vertically in the ground, or submersed in a body of water. The type of loop configurations includes:

- Horizontal closed ground loop (Figure 12.32a)
- Vertical closed ground loop (Figure 12.32b)
- Pond or lake closed loop
- Open loop system
- Standing column well system

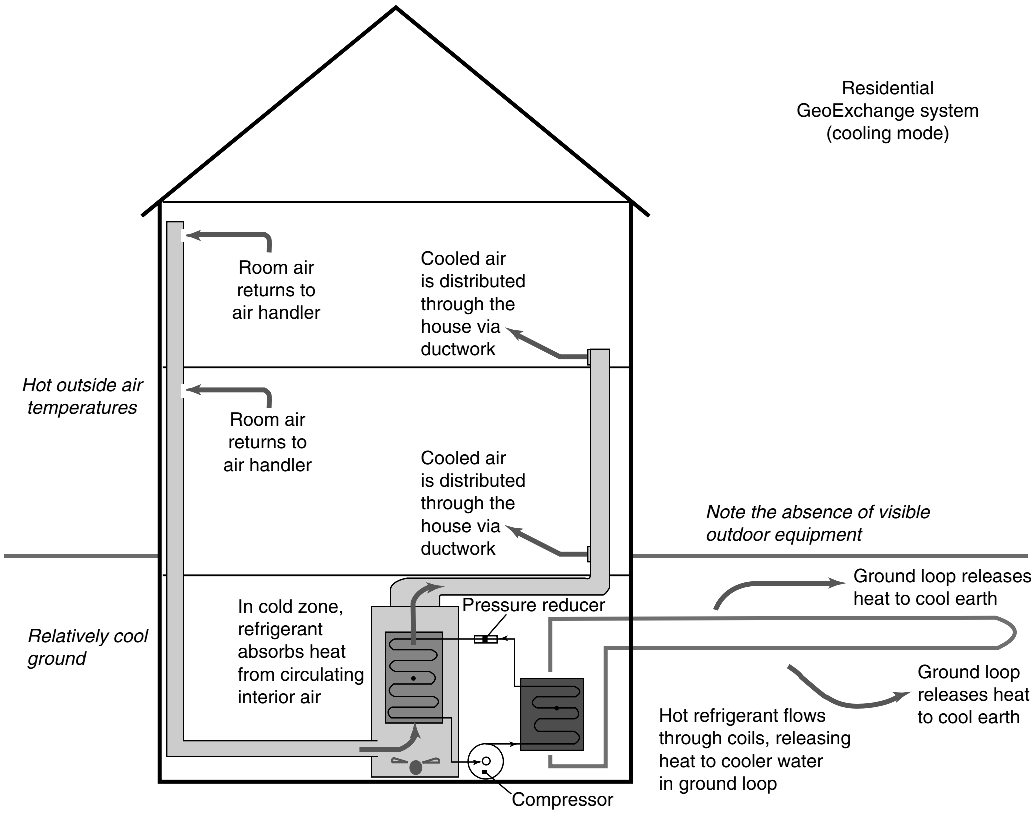


FIGURE 12.30 Example of a Residential Geothermal Heat Pump Configuration in the Summer Months (Drawing courtesy of the Geothermal Heat Pump Consortium.)

Geothermal heat pump systems offer a great deal of flexibility and according to the Geothermal Heat Pump Consortium (GHPC), these systems have been installed in thousands of commercial applications from guard shacks to high-rise office buildings, not to mention in thousands of houses and schools across the United States.

Geothermal heat pumps also offer significant benefits to customers willing to pay the higher upfront installation costs (Figure 12.33 through Figure 12.35). These benefits include:

- *Substantial cost savings*: Geothermal systems can save as much as 50% compared to air-source heat pumps and up to 45% over fossil-fuel (gas, propane, or oil) furnaces.
- *Economical rates*: Some utilities, such as Kansas City Power and Light, offer special, lower winter rates for geothermal customers, offering even more savings.
- *Environmentally Friendly*: Geothermal systems are a "renewable" energy source that encourages conservation of natural resources.
- *Financing*: Many utilities offer financing through either private financing or utility-sponsored loop leases.

Even though geothermal heat pumps have been around for decades, this still remains a niche market. This technology still faces numerous challenges including reducing first costs, raising awareness, and developing a sustainable infrastructure.

Geothermal heat pumps take advantage of the natural constant temperature of the Earth. Just three to five feet below ground, the temperature remains between 50 and 60°F (10°C–16°C) year round.

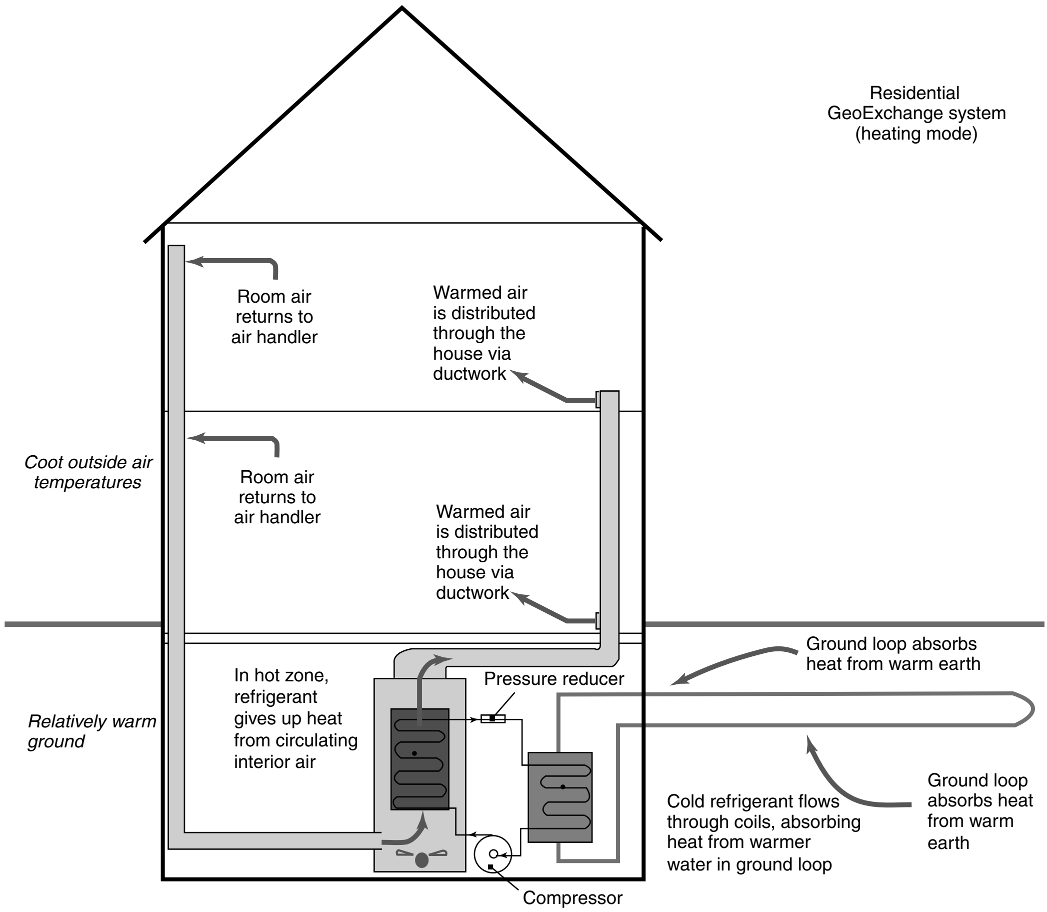


FIGURE 12.31 Example of a Residential Geothermal Heat Pump Configuration in the Winter Months (Drawing courtesy of the Geothermal Heat Pump Consortium.)

The ground temperature is warmer than the air above it in the winter and cooler than the air in the summer. Geothermal heat pumps take advantage of this difference to heat and cool buildings [6,7].

The heat pump must compensate for the differences between where the heat is absorbed (e.g., the source) and the temperature where the heat is delivered (e.g., the "sink." This difference is called the "lift" and the larger the "lift," the more power is required from the heat pump. An air-source heat pump

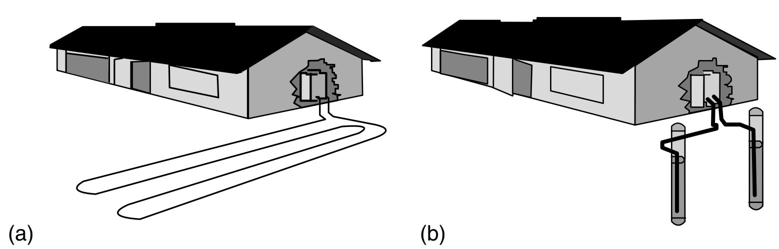


FIGURE 12.32 (a) Example of a Horizontal Loop Configuration in a Residential Installation. (b) Example of a Vertical Loop Configuration in a Residential Installation: Drawings courtesy of the Geothermal Heat Pump Consortium.

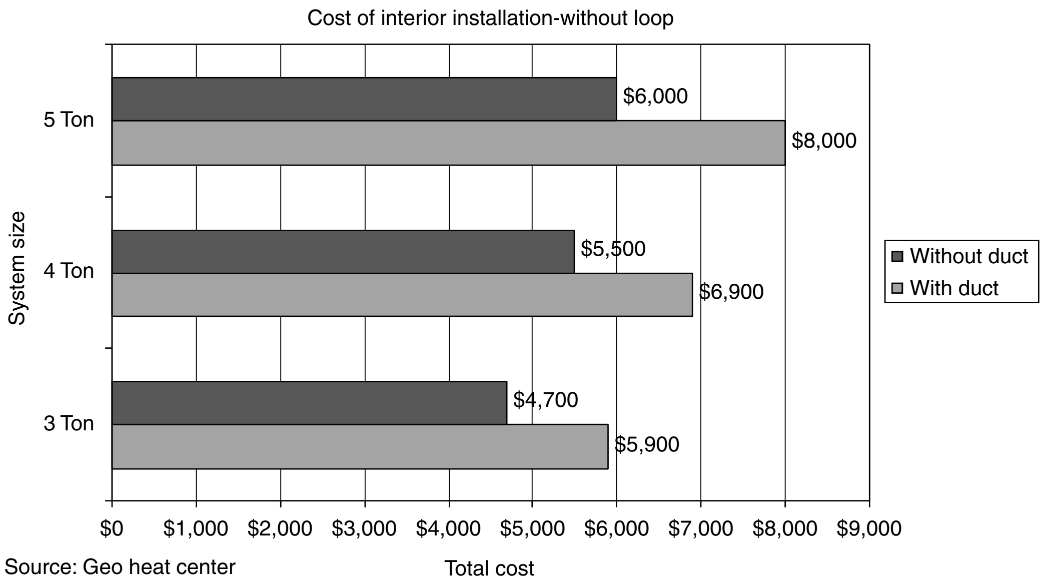


FIGURE 12.33 Average Interior Installation Costs for GHPs by System Size.

removes heat from cold outside air in the winter and delivers heat to hot outside air in the summer. The geothermal heat pump, on the other hand, recovers heat from relatively warm soil (or groundwater) in the winter and delivers heat to the same relatively cool soil (or groundwater) in the summer. As a result, the geothermal heat pump is pumping the heat over a smaller temperature difference than the air-source heat pump. This leads to higher efficiency and lower energy use [7].

The geothermal heat pump system has three major parts: the ground heat exchanger, the heat pump unit, and the air delivery system (ductwork). The heat exchanger is a system of pipes called a loop, which

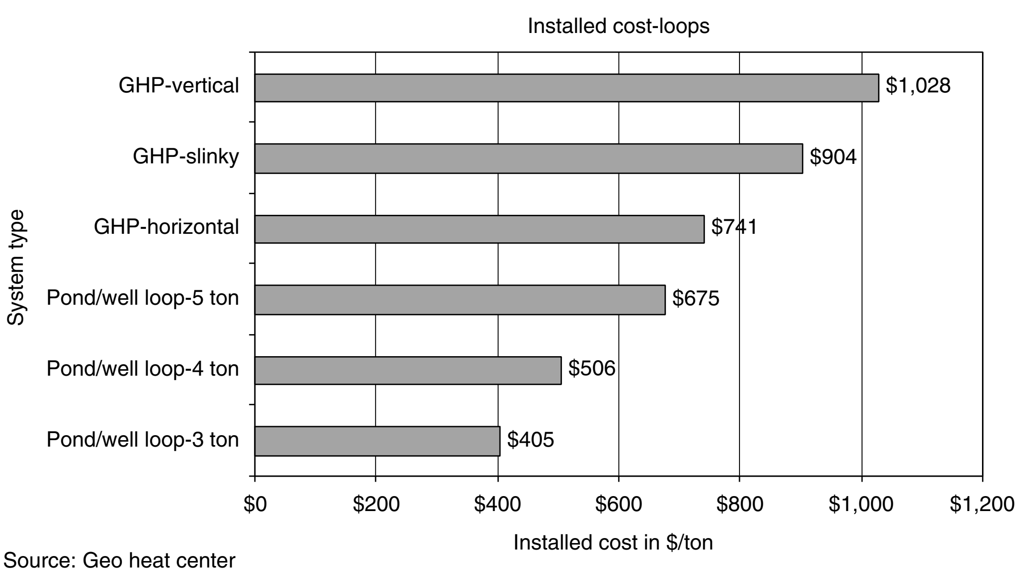


FIGURE 12.34 Average Loop Costs by GHP System Size.

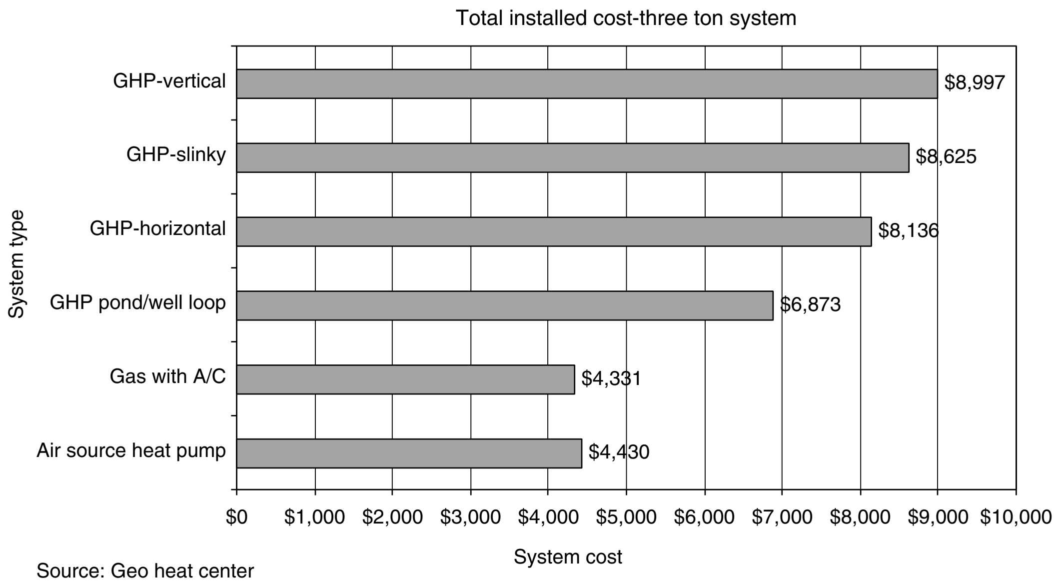


FIGURE 12.35 Total Installed Cost for Three Ton GHP Systems by Type.

is buried in the shallow ground near the building. A fluid (usually water or a mixture of water and antifreeze) circulates through the pipes to absorb or deposit heat within the ground.

In the winter, the heat pump removes heat from the heat exchanger and transfers it into the indoor air delivery system. In the summer, the process is reversed, and the heat pump moves heat from the indoor air into the heat exchanger. The heat removed from the indoor air during the summer can also be used to heat water.

In most standard installations, especially in the residential market, the geothermal heat pump consists of a single package water-to-air heat pump. All components are contained in a single enclosure that is about the same size as a small gas furnace. The unit includes a refrigerant-to-water heat exchanger, refrigerant piping and control valve, compressor, air coil (heats in winter; cools and dehumidifies in summer), fan and controls. Nearly all geothermal heat pump units use refrigerant R-22.

Manufacturers also offer split systems, water-to-water heat pumps, multispeed compressors, dual compressor, and rooftop versions of this equipment to suit various applications [8].

12.4.3.1 The Loop

The ground connection, or loop, provides the means of transferring heat to the earth in summer, and extracting heat from the earth in winter. Physically, the "ground loop" consists of several lengths of plastic pipe typically installed either in horizontal trenches or vertical holes that are covered up with earth.

Fluid inside the ground loop, either water or refrigerant, is pumped through a heat exchanger in the geothermal heat pump. In the summer, it absorbs heat from the refrigerant hot zone and carries it to the ground through the ground loop piping. In winter, it absorbs heat from the earth through the ground loop, and then transfers that heat to the refrigerant cold zone.

The length of the ground loop is determined by the heating and cooling loads, which are determined in turn by the home or building, its design and construction, orientation, and climate.

Once installed, the loop remains out of sight beneath the surface. The ground loop consists of polyethylene piping which is the same kind used for cross-country natural gas lines. It will not degrade, corrode, or break down in ground or water contact, so proper installations are projected to last at least 50 years [9].

Sometimes geothermal heat pumps are installed using a variation on this practice called direct exchange (DX). This method uses copper piping placed underground and as the refrigerant is pumped through the loop, the heat is transferred directly through the copper to the earth.

The loop field should be designed and installed by professionals who follow the guidelines established by the International Ground Source Heat Pump Association (IGSHPA). Installers should be certified by IGSHPA or demonstrate equivalent training by manufacturers or other recognized authorities.

Most loops are installed either horizontally or vertically in the ground, or submersed in water in a pond or lake. In most cases, the fluid runs through the loop in a closed system, but open-loop systems may be used where local codes permit. Each type of loop configuration has advantages and disadvantages.

Horizontal Ground Closed Loops. This configuration is the most cost-effective when adequate land available and trenches are easy to dig. Workers use trenchers or backhoes to dig the trenches three to six feet below the ground and then lay a series of parallel plastic pipes. They backfill the trench, taking care not to allow sharp rocks or debris to damage the pipes. A typical horizontal loop will be 400 to 600 ft. long per ton of heating and cooling capacity.

Vertical Ground Closed Loops. This type of loop configuration is used under the following conditions:

- There is insufficient land to permit horizontal buildings with large heating and cooling loads,
- The earth is rocky close to the surface, or
- In retrofit applications where minimum disruption of the landscaping is desired.

Contractors bore vertical holes in the ground 150 to 450 ft. deep. Each hole contains a single loop of pipe with a U-bend at the bottom. After the pipe is inserted, the hole is backfilled or grouted. Each vertical pipe is then connected to a horizontal pipe, which is also concealed underground. The horizontal pipe then carries fluid in a closed system to and from the geothermal heat pump system.

Vertical loops are generally more expensive to install, but require less piping than horizontal loops because the earth deeper down is cooler in summer and warmer in winter [9].

The major advantage of the vertical design is that it places the loop in a much more thermally stable zone. Soil at 100 ft is not subject to the same temperature fluctuations as soil at a 4 or 5 ft depth. Vertical loop configurations are most appropriate in extreme climate zones.

Open Loop Systems. Open loops may be cost-effective if ground water is plentiful. They are the simplest to install and have been used successfully for decades in areas where local codes permit. In this type of system, ground water from an aquifer is piped directly from the well to the building, where it transfers its heat to a heat pump. After it leaves the building, the water is pumped back into the same aquifer via a second well-called a discharge well-located at a suitable distance from the first.

This configuration requires a slightly larger well pump is installed to provide for the water required by the heat pump. A major concern is water disposal. Open loops have used ponds, lakes, rivers, irrigation ditches, and return (or injection) wells. It is also important to maintain the water quality. Since the water either absorbs or gives up heat, but is not altered in any other way, it leaves the geothermal heat pump unit as pure as it was when it entered it.

Pond Closed Loops. If the building is near a body of surface water, such as a pond or lake, this type of loop design may be the most economical. The fluid circulates through polyethylene piping in a closed system. Typically, workers run the pipe to the water and then submerge long sections under water. The pipe may be coiled in a slinky shape to fit more of it into a given amount of space. Geothermal heat pump experts recommend using a pond loop only if the water level never drops below six to eight feet at its lowest level to assure sufficient heat-transfer capability. Pond loops used in a closed system have no adverse impacts on the aquatic system.

Standing Column Well System. Standing column wells have been used in installations in the northeast United States. Standing wells are typically six inches in diameter and may be as deep as 1,500 ft..

TABLE 12.9 Comparison of Costs for Residential Heating Systems

Compare	Safety	Installation Cost	Operating Cost	Maintenance Cost	Life-Cycle Cost
Combustion-based	A Concern	Moderate	Moderate	High	Moderate
Heat pump	Excellent	Moderate	Moderate	Moderate	Moderate
Geothermal Heat Pump	Excellent	High	Low	Low	Low

Temperate water from the bottom of the well is withdrawn, circulated through the heat pump's heat exchanger, and returned to the top of the water column in the same well.

A standing well requires plentiful ground water. In normal circumstances, the water diverted for building (potable) use is replaced by constant-temperature ground water, which makes the system act like a true open-loop system [9].

12.4.3.2 Costs

When comparing heating systems, safety, installation cost, operating costs, and maintenance costs must be considered. The table below compares the various types of central heating systems using the lifecycle cost analysis Table 12.9.

A study by the Environmental Protection Agency compared six locations representing major climate zones in the U.S. The cities (Burlington, VT; Chicago, IL; upper New York City; Portland, OR; Atlanta, GA; and Phoenix, AZ) were chosen to compare the [10,11] performance and costs of emerging high-efficiency space-conditioning equipment with equipment already on the market.

In all locations, geothermal heat pumps were the most efficient heating and cooling systems over other types of space-conditioning equipment including high-efficiency gas furnaces and air conditioners. Geothermal heat pump installations in both new and existing homes may reduce energy consumption 25 to 75 percent compared to older or conventional replacement systems. Geothermal heat pumps also had the lowest annual operating costs.

Geothermal heat pumps have also been installed in hundreds of commercial and institutional buildings throughout the United States. Some of the largest installations include 24,000 square foot administration building at Fort Polk, LA in 1993 and an 80,000 square foot office building in Allentown, PA as well as numerous schools throughout the United States.

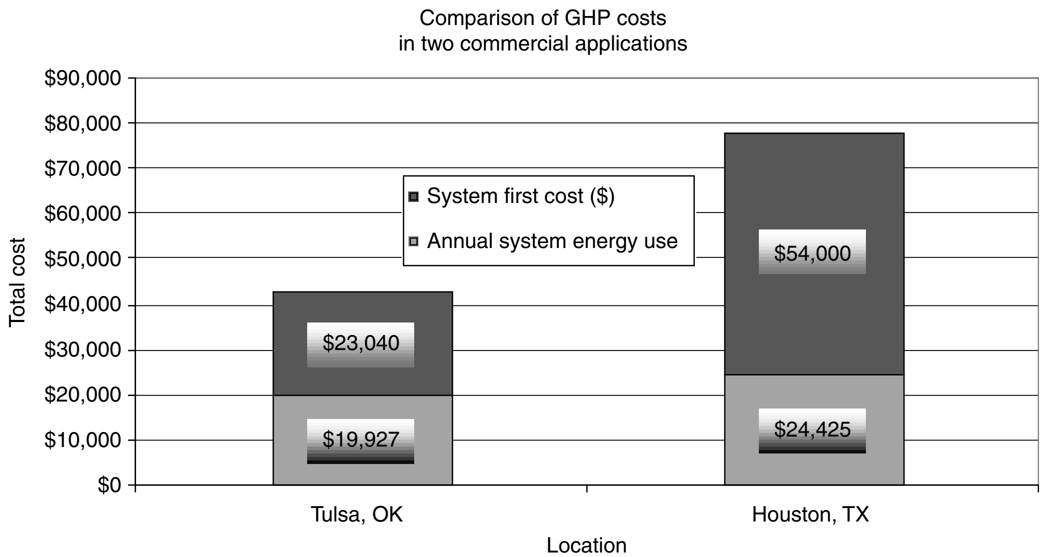
The Fort Polk Project illustrates the diversity of applications in which geothermal heat pumps have been used in commercial settings. This $18 million project, completed in 1996, was projected to reduce annual maintenance and energy costs by more than $3 million annually—a projected $44 million savings over the life of the project. The 300 mile facility houses 23,000 military personnel and also includes a military hospital, administration buildings, training centers, and storage facilities.

The geothermal installations included replacing more than 3,000 air source heat pumps with geothermal heat pumps. This replacement effectively raised the installed SEER levels from 7 to 15.5, a significant energy savings and performance improvement. The total installation involved vertical drilling of more than 1.8 million ft., or approximately 686 square miles. This project is the largest geothermal heat pump installation in the world and involved more than nine separate drilling contractors [12].

An evaluation of a 4000-home comprehensive GHP retrofit at the U.S. Army's Fort Polk in Louisiana showed that the GHPs reduced summer peak electric demand on military base by 7.5 MW, or 43 percent, and reduced electricity consumption in post housing by 33 percent, while eliminating natural gas consumption completely [13].

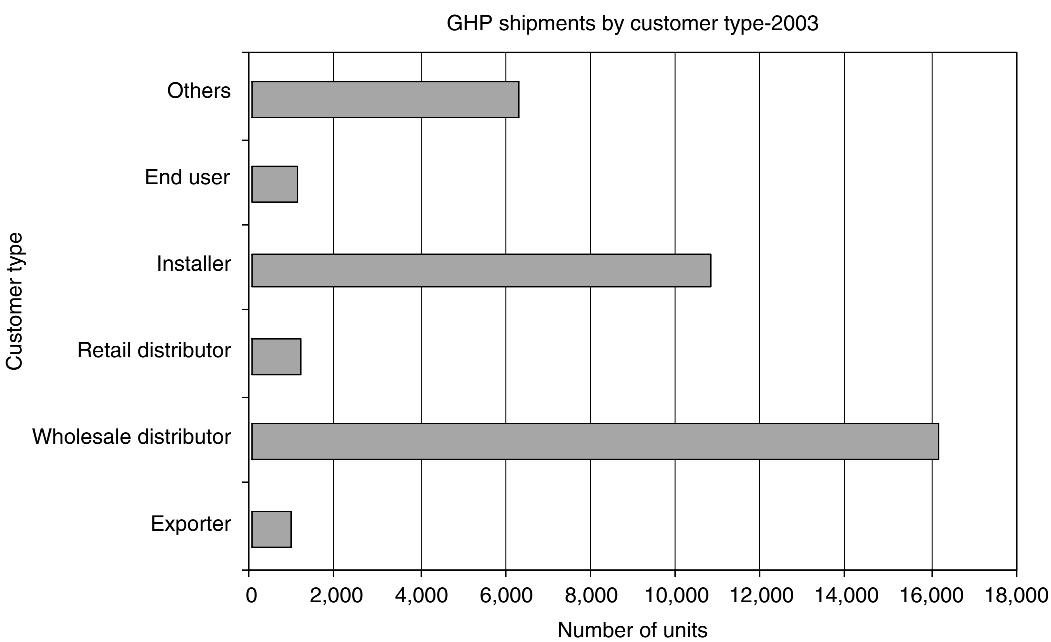
Figure 12.36 illustrates the average installed costs for two geothermal heat pump installations in two locations. It also demonstrates that the total installation costs will vary based upon a number of factors, including climate, soil type, and drilling costs.

Installation Trends. The Energy Information Administration (EIA) conducts an annual survey of geothermal heat pump shipments on a per ton basis. The next two figures illustrate the most recent



Source: Analysis by S.P. Kavanaugh (1998), for the U.S. Department of Energy-Oak Ridge National Laboratory, December 2001

FIGURE 12.36 Comparison of Commercial GHP Installation Costs by Location.



Source: Energy information administration, from EIA-902 "Annual geothermal heat pump manufacturers survey."

FIGURE 12.37 GHP Shipments by Customer Type in 2003.

data collected regarding GHPS. The total ton capacity of geothermal heat pumps peaked in 1999 but it has been declining steadily ever since.

Figure 12.37 illustrates that the dominant customer group receiving geothermal heat pump shipments are the key trade allies of the wholesale distributors and installers.

12.4.4 Conclusions

According to the most recent government estimates, there are more than 900,000 geothermal heat pump installations in the United States. These installations have led to significant energy savings and reductions in carbon emissions:

- Elimination of more than 5.2 million metric tons of CO2 annually
- Elimination of more than 1.4 million metric tons of carbon equivalent annually
- Annual savings of more than 7 billion kWh!
- Annual savings of more than 36 trillion Btus of fossil fuels
- Reduced electricity demand of 2.3 million kW

The following organizations are able to provide more detailed information about both the costs and benefits of purchasing and installing a geothermal heat pump system:

- DOE Geothermal Technologies Program (www.eere.energy.gov/geothermal)
- Geo-Heat Center (www.geoheat.oit)
- Geothermal Heat Pump Consortium (www.geoexchange.org)
- International Ground Source Heat Pump Association (IGHSHPA) (www.ighspa.okstate.edu)

Definition of Terms and Abbreviations

ACCA: Air Conditioning Contractors Association, a trade association that is working closely with several geothermal organizations to educate HVAC installers about geothermal heat pumps.
ARI: Air-conditioning and Refrigeration Institute
ASHRAE: American Society of Heating, Refrigeration, and Air Conditioning Engineers
COP: Coefficient of Performance—a measure of energy usage and efficiency used in heating and air conditioning equipment.
GHPC: Geothermal Heat Pump Consortium
GHPS: Geothermal heat pumps, also known as: ground source heat pumps, earth source/earth coupled heat pumps, geothermal heating and cooling systems, direct exchange. or “geo”among others.
HVAC: Heating, ventilation and air conditioning
IGSHPA: International Ground Source Heat Pump Association

References

1. CRC Handbook of Mechanical Engineering, *2nd Ed.*, 2005, Chapter 9: Air Conditioning and Refrigeration, Herbert A. Ingley and Shan K Wang
2. Handbook of Heating and Air Conditioning, 2002, ASHRAE Handbook: Refrigeration by Robert Parson, American Society of Heating, Refrigeration and Airconditioning.
3. Handbook of Heating, Ventilation and Air Conditioning, Jan F. Kreider, ed., 2001, CRC Press, Boca Raton, FL.
4. Radermacher, R. and Yunho Hwang, *Vapor Compression Heat Pumps with Refrigerant Mixtures,* CRC Press, Taylor and Francis Group, Boca Raton, FL. 2005.

5. Solar Energy Handbook, Jan F. Kreider and F. Kreith, ed., McGraw-Hill, 1981, Chapter 13: Liquid based Solar Systems for Space Heating, Fred Dubin and S. Bloome
6. National Renewable Energy Labs, http://www.nrel.gov/clean_energy/geoheatpumps.html
7. Geo-Heat Center, http://geoheat.oit.edu/ghp/hpfund.htm
8. Geo-Heat Center, Klamath Falls, OR 97601, http://geoheat.oit.edu/ghp/hpfund.htm
9. *Geoexchange Systems* — Outstanding Home Comfort through Advanced Technology Fact Sheet, GHPC, Washington, D.C.
10. *Geoexchange Systems* — Outstanding Home Comfort through Advanced Technology Fact Sheet, GHPC, Washington, D.C.
11. *Geoexchange Systems* — Outstanding Home Comfort through Advanced Technology Fact Sheet, GHPC, Washington, D.C.
12. Comparing Heating Systems Fact Sheet — 1997, GHPC, Washington, D.C.
13. Space Conditioning: The Next Frontier, EPA 430-R-93-004, April 1993.
14. Geoexchange Savings Millions at Fort Polk, LA, GHPC Case Study, Washington, D.C.
15. Hughes, P. J., and J. A. Shonder, 1998. "The Evaluation of a 4000-Home Geothermal Heat Pump Retrofit at Fort Polk, Louisiana: Final Report," ORNL/CON-460, Oak Ridge National Laboratory, 1998. (Available at www.ornl.gov/femp)

13

Compact Heat Exchangers—Recuperators and Regenerators

Ramesh K. Shah
Subros Limited

13.1 Introduction

A heat exchanger is a device to transfer thermal energy (enthalpy) between two or more fluids, between a solid surface and a fluid, or between solid particulates and a fluid at different temperatures and in thermal contact without external heat and work interactions. The fluids may be single compounds or mixtures. Typical applications involve heating or cooling of a fluid stream of concern, evaporation or condensation of single or multicomponent fluid stream, and heat recovery or heat rejection from a system. In other applications, the objective may be to heat, cool, condense, vaporize, sterilize, pasteurize, fractionate, distill, concentrate, crystallize, or control process fluid. In some heat exchangers, the fluids transferring heat are in direct contact. In other heat exchangers, heat transfer between fluids takes place through a separating wall or into and out of a wall in a transient manner. In most heat exchangers, the fluids are separated by a heat transfer surface, and ideally they do not mix. Such exchangers are referred to as *direct transfer type* or simply *recuperators*. In contrast, exchangers in which there is an intermittent flow of heat from the hot to cold fluid—via heat storage and heat rejection through the exchanger surface or matrix—are referred to as *indirect transfer type* or *storage type exchangers*, or simply *regenerators*.[1]

A heat exchanger consists of heat exchanging elements, such as a core or a matrix containing the heat transfer surface, and fluid distribution elements such as headers, manifolds, tanks, inlet, and outlet nozzles or pipes, or seals. Usually there are no moving parts in a heat exchanger; however, there are exceptions such as a rotary regenerator, in which the matrix is mechanically driven to rotate at some design speed.

The heat transfer surface is a surface of the exchanger core that is in direct contact with fluids and through which heat is transferred by conduction in a recuperator, and by heat storage and rejection in a regenerator. The portion of the surface which also separates the fluids in a recuperator is referred to as *primary* or *direct surface*. To increase heat transfer area, appendages known as fins may be intimately connected to the primary surface to provide *extended*, *secondary*, or *indirect surface*. Thus, the addition of fins reduces the thermal resistance on that fluid side and thereby increases the net heat transfer from the surface for the same temperature difference. Note that in a regenerator, all heat transfer surface acts as a primary surface.

Heat exchangers may be classified according to transfer process, construction, flow arrangement, surface compactness, number of fluids, and heat transfer mechanisms as shown in Figure 13.1 (Shah 1981; Shah and Muller 1988) or according to the process function as shown in Figure 13.2 (Shah and Muller 1988). Further general description of heat exchangers is provided by Walker (1990); Saunders (1988), and Hewitt (1989).

A gas-to-fluid heat exchanger is referred to as compact heat exchanger if it incorporates heat transfer surface having a surface area density above about 700 m^2/m^3 (213 $ft.^2/ft.^3$) or a hydraulic diameter $D_h \leq 6$ mm (0.25 in.) on at least one of the fluid sides that usually has gas flow. It is referred to as a laminar flow (or meso) heat exchanger if the surface area density is above about 3000 m^2/m^3 (914 $ft.^2/ft.^3$) or 100 $\mu m \leq D_h \leq 1$ mm, and as a micro heat exchanger if the surface area density is above about 15,000 m^2/m^3 (4570 $ft.^2/ft.^3$) or 1 $\mu m \leq D_h \leq 100$ μm. A liquid/two-phase heat exchanger is referred to as compact heat exchanger if the surface area density on anyone fluid side is above about 400 m^2/m^3 (122 $ft.^2/ft.^3$). A typical process industry shell-and-tube exchanger has a surface area density

[1]In vehicular gas turbines, a stationary heat exchanger is usually referred to as a recuperator, and a rotating heat exchanger as a regenerator. However, in industrial gas turbines, by long tradition and from a thermodynamic sense, a stationary heat exchanger is generally referred to as a regenerator. Hence, a gas turbine regenerator could be either a recuperator or a regenerator in a strict sense depending on the application.

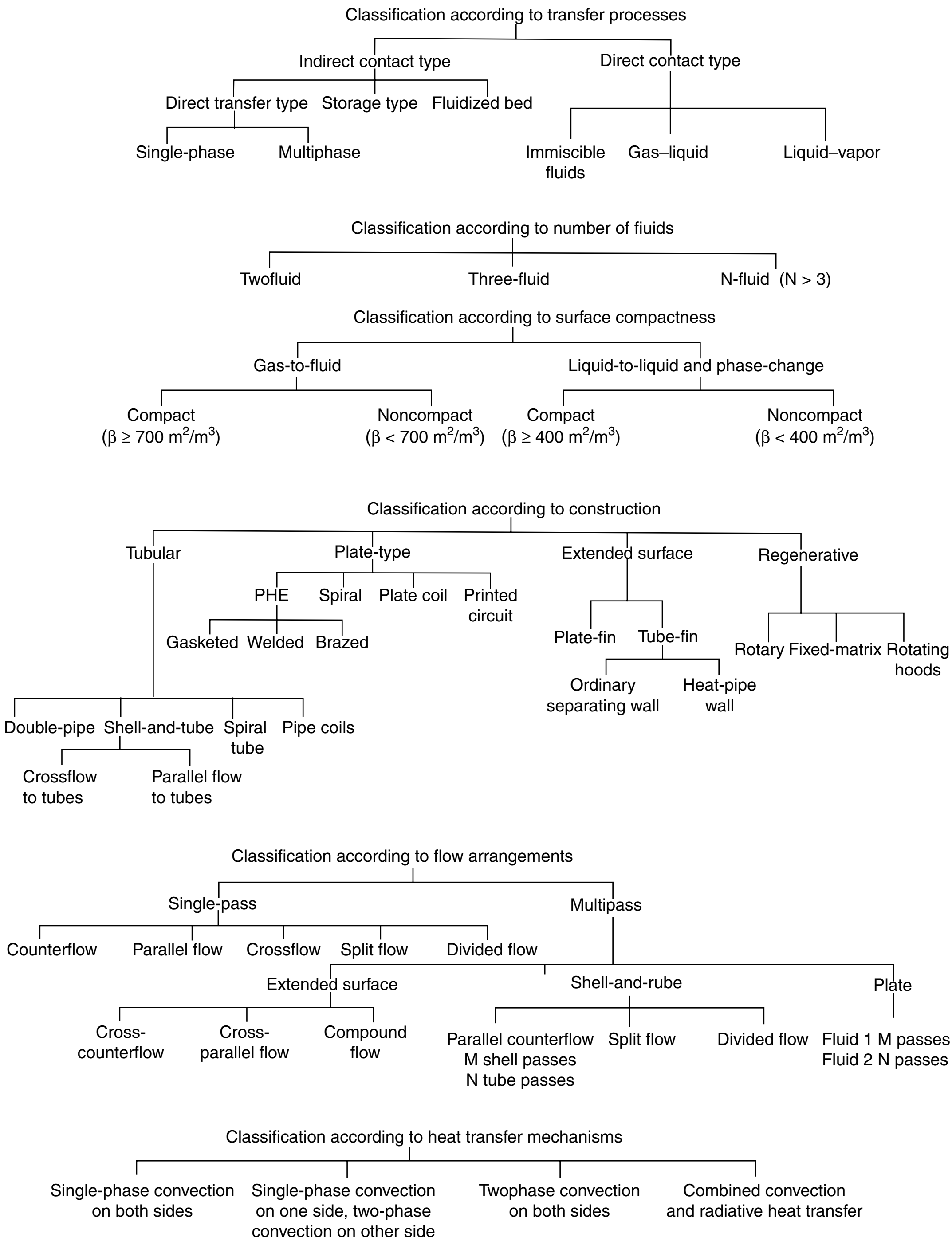


FIGURE 13.1 General classification of heat exchangers. (Modified from Shah, R. K. and Muller, A. C., *Ullmann's Encyclopedia of Industrial Chemistry, Unit Operations II, Vol. B3*. VCH Publishers, Weinheim, Germany, Chap. 2, 1988.)

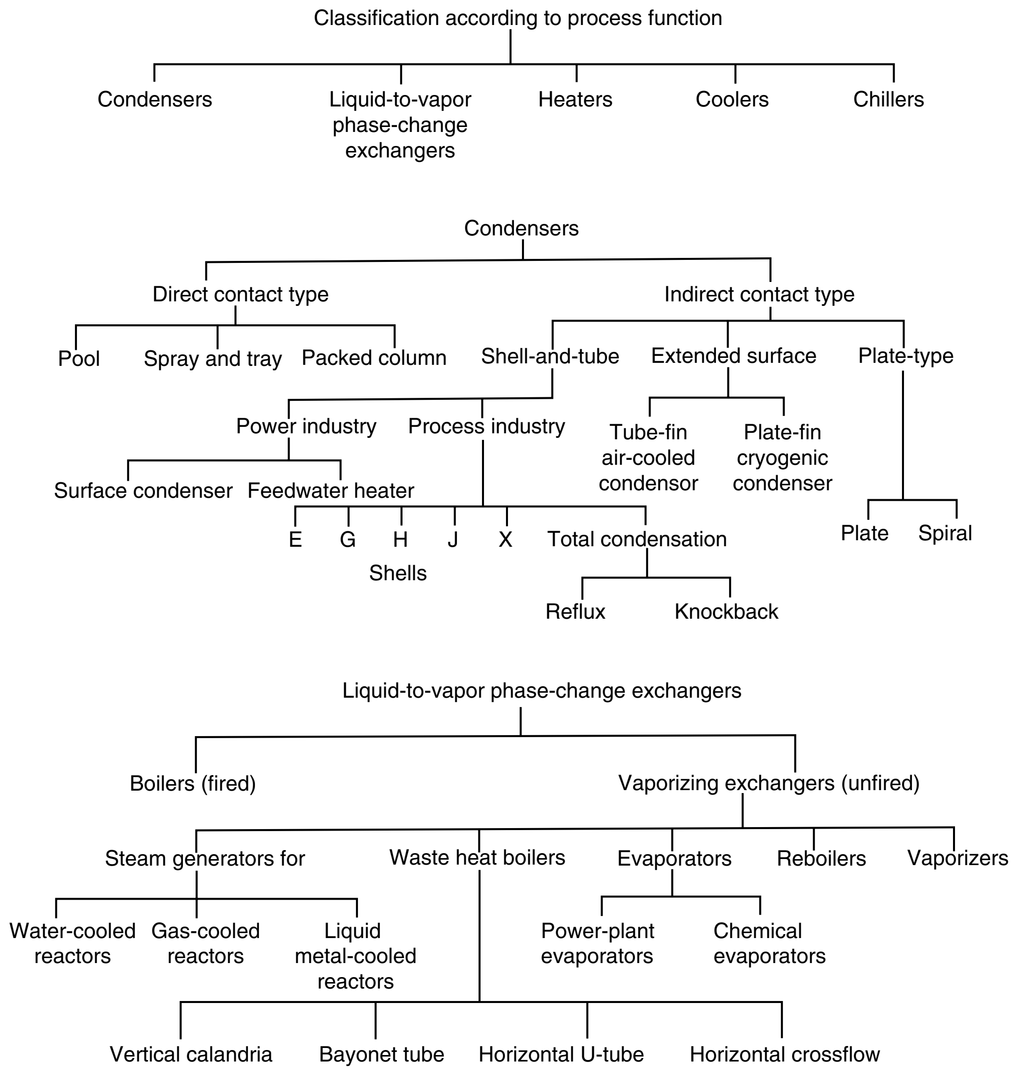


FIGURE 13.2 (a) Classification according to the process functions, (b) classification of condensers, (c) classification of liquid-to-vapor phase-change exchangers. (Modified from Shah, R. K. and Muller, A. C., *Ullmann's Encyclopedia of Industrial Chemistry, Unit Operations II, Vol. B3*. VCH Publishers, Weinheim, Germany, Chap. 2, 1988.)

of less than 100 m^2/m^3 on one fluid side with plain tubes, and two or three times that with the high-fin-density, low-finned tubing. Plate-fin, tube-fin, and rotary regenerators are examples of compact heat exchangers for gas flows on one or both fluid sides, and gasketed and welded plate heat exchangers are examples of compact heat exchangers for liquid flows.

In this chapter, exchanger heat transfer and pressure drop analysis is presented first. Next, theoretical results/insights and empirical correlations for nondimensional heat transfer and flow friction characteristics of exchanger surfaces are presented. Overall design methodology and step-by-step design procedures for the exchanger rating and sizing problems are then outlined. Finally, flow maldistribution and fouling problems and design considerations are summarized.

There is ever-increasing use of compact heat exchangers for advanced power cycles, and hence the focus in this chapter is on compact heat exchangers. Readers are referred to excellent works of Singh and Soler (1984); Palen (1987); Saunders (1988); Hewitt (1989); Yokell (1990), and Hewitt, Shires, and Bott (1994) for design information on shell-and-tube and other heat exchangers; Reay (1979) and Hesselgreaves (2001) for compact heat exchangers, and Kakaç and Liu (1998) for single-phase and phase-change heat exchangers. Refer to Shah and Sekulić (2003) for an extensive list of publications on heat exchanger books.

13.2 Recuperator Heat Transfer and Pressure Drop Analysis

In this section, starting with the thermal circuit associated with a two-fluid exchanger, ε-NTU, P-NTU, and MTD methods used for an exchanger analysis are presented, followed by the fin efficiency concept and various expressions. Finally, expressions for computation of pressure drop are outlined for various single-phase exchangers.

13.2.1 Thermal Circuit

In order to develop relationships among variables for various exchangers, consider the counterflow exchanger of Figure 13.3 as an example. Two energy conservation differential equations for a two-fluid exchanger with *any* flow arrangement are

$$\mathrm{d}q = q''\mathrm{d}A = -C_\mathrm{h}\mathrm{d}T_\mathrm{h} = \pm C_\mathrm{c}\mathrm{d}T_\mathrm{c} \tag{13.1}$$

where the $\pm$ sign depends on whether $\mathrm{d}T_\mathrm{c}$ is increasing or decreasing with increasing $\mathrm{d}A$. The overall rate equation on a local basis is

$$\mathrm{d}q = q''dA = U(T_\mathrm{h} - T_\mathrm{c})_\mathrm{local}\mathrm{d}A = U\Delta T\ \mathrm{d}A \tag{13.2}$$

Integration of Equation 13.1 and Equation 13.2 across the exchanger surface area results in

$$q = C_\mathrm{h}(T_{\mathrm{h},i} - T_{\mathrm{h,0}}) = C_\mathrm{c}(T_{\mathrm{c,0}} - T_{\mathrm{c,i}}) \tag{13.3}$$

and

$$q = \mathrm{UA}\Delta T_\mathrm{m} = \Delta T_\mathrm{m}/R_0 \tag{13.4}$$

Here, ΔT_m is the true mean temperature difference dependent on the exchanger flow arrangement and degree of fluid mixing within each fluid stream. The inverse of the overall thermal conductance UA is referred to as the overall thermal resistance R_0 as follows (see Figure 13.4).

$$R_\mathrm{o} = R_\mathrm{h} + R_\mathrm{h,s} + R_\mathrm{w} + R_\mathrm{c,s} + R_\mathrm{c} \tag{13.5}$$

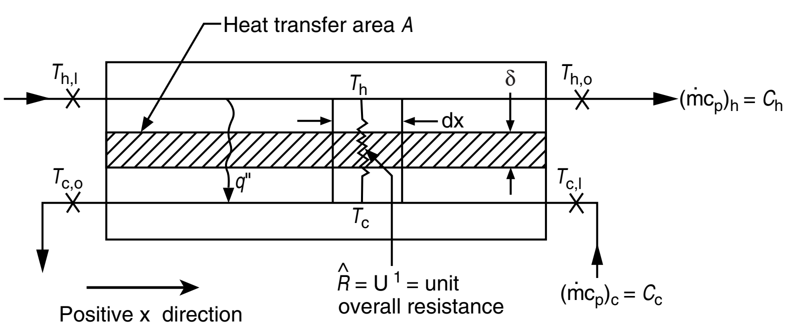


FIGURE 13.3 Nomenclature for heat exchanger variables.

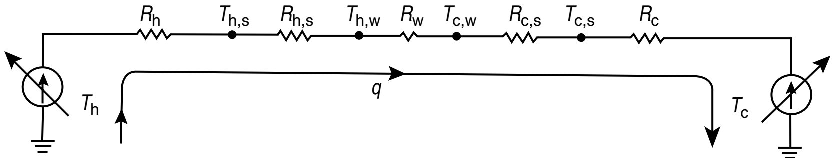


FIGURE 13.4 Thermal circuit for heat transfer in an exchanger.

where the subscripts h, c, s, and w denote hot, cold, scale or fouling, and wall, respectively. In terms of the overall and individual heat transfer coefficients, Equation 13.5 is represented as

$$\frac{1}{\mathrm{UA}} = \frac{1}{(\eta_0 hA)_\mathrm{h}} + \frac{1}{(\eta_0 h_\mathrm{s} A)_\mathrm{h}} + R_\mathrm{w} + \frac{1}{(\eta_0 h_\mathrm{s} A)_\mathrm{c}} + \frac{1}{(\eta_0 hA)_\mathrm{c}} \tag{13.6}$$

where η_0 is the extended surface efficiency and is related to the fin efficiency η_f, fin surface area A_f, and the total heat transfer surface area A as follows:

$$\eta_0 = 1 - \frac{A_\mathrm{f}}{A}(1 - \eta_\mathrm{f}) \tag{13.7}$$

The wall thermal resistance R_w of Equation 13.5 is given by

$$R_\mathrm{w} = \begin{cases} \dfrac{\delta}{k_\mathrm{w} A_\mathrm{w}} \text{ for flat walls with a single layer wall} \\ \sum_j \left(\dfrac{\delta}{k_\mathrm{w} A_\mathrm{w}}\right)_j \text{ for flat walls with } j \text{ multiple layer wall} \end{cases} \tag{13.8a}$$

and

$$R_\mathrm{w} = \begin{cases} \dfrac{\ln(d_0/d_\mathrm{i})}{2\pi k_\mathrm{w} \mathrm{LN_t}} \text{ for } N_\mathrm{t} \text{ circular tubes with a single layer wall} \\ \dfrac{1}{2\pi \mathrm{LN_t}} \sum_j \dfrac{\ln(d_{j+1}/d_j)}{k_{\mathrm{w},j}} \text{ for } N_\mathrm{t} \text{ circular tubes with } j \text{ multiple layer wall} \end{cases} \tag{13.8b}$$

If there is any contact or bond resistance present between the fin and tube or plate on the hot or cold fluid side, it is included as an added thermal resistance on the right-hand side of Equation 13.5 or Equation 13.6. For a heat-pipe heat exchanger, additional thermal resistances associated with the heat pipe should be included on the right-hand side of Equation 13.5 or Equation 13.6; these resistances are evaporator resistance at the evaporator section of the heat pipe, viscous vapor flow resistance inside the heat pipe (very small), internal wick resistance at the condenser section of the heat pipe, and condensation resistance at the condenser section.

If one of the resistances on the right-hand side of Equation 13.5 or Equation 13.6 is significantly higher than the other resistances, it is referred to as the *controlling resistance*. A reduction in the controlling thermal resistance will have much more impact in reducing the exchanger surface area A requirement compared to the reduction in A due to the reduction in other thermal resistances.

UA of Equation 13.6 may be defined in terms of hot or cold fluid side surface area or wall conduction area as

$$\mathrm{UA} = U_\mathrm{h} A_\mathrm{h} = U_\mathrm{c} A_\mathrm{c} = U_\mathrm{w} A_\mathrm{w} \tag{13.9}$$

When R_w is negligible, $T_{w,h} = T_{w,c} = T_w$ of Figure 13.4 is computed from

$$T_w = \frac{T_h + [(R_h + R_{h,s})/(R_c + R_{c,s})]T_c}{1 + [(R_h + R_{h,s})/(R_c + R_{c,s})]} \tag{13.10}$$

When $R_{w,h} = R_{w,c} = 0$, Equation 13.10 reduces to

$$T_w = \frac{T_h/R_h + T_c/R_c}{1/R_h + 1/R_c} = \frac{(\eta_0 hA)_h T_h + (\eta_0 hA)_c T_c}{(\eta_0 hA)_h + (\eta_0 hA)_c} \tag{13.11}$$

13.2.2 ε-NTU, *P*-NTU, and MTD Methods

If we consider the fluid outlet temperatures or heat transfer rate as dependent variables, they are related to independent variables/parameters of Figure 13.3 as follows.

$$T_{h,0}, T_{c,0}, \text{ or } q = \phi\{T_{h,i}, T_{c,i}, C_h, C_c, U, A, \text{flow arrangement}\} \tag{13.12}$$

Six independent and three dependent variables of Equation 13.12 for a given flow arrangement can be transferred into two independent and one dependent dimensionless groups; three different methods are presented in Table 13.1 based on the choice of three dimensionless groups. The relationship among three dimensionless groups is derived by integrating Equation 13.1 and Equation 13.2 across the surface area for a specified exchanger flow arrangement. Such expressions are presented later, in Table 13.3, for industrially most important flow arrangements. Now we will briefly describe the three methods.

13.2.3 The ε-NTU Method

In this method, the heat transfer rate from the hot fluid to the cold fluid in the exchanger is expressed as

$$q = \varepsilon C_{\min}(T_{h,i} - T_{c,i}) \tag{13.13}$$

TABLE 13.1 General Functional Relationships and Dimensionless Groups for ε-NTU, *P*-NTU, and LMTD Methods

ε-NTU Method	*P*-NTU Method
$q = \varepsilon C_{\min}(T_{h,i} - T_{c,i})$	$q = P_1 C_1 \lvert T_{1,i} - T_{2,i}\rvert$
$\varepsilon = \phi(\text{NTU}, \text{C}^*, \text{flow arrangement})$	$P_1 = \phi(\text{NTU}, \text{R}_1, \text{flow arrangement})$
$\varepsilon = \dfrac{C_h(T_{h,i} - T_{h,o})}{C_{\min}(T_{h,i} - T_{c,i})} = \dfrac{C_c(T_{c,o} - T_{c,i})}{C_{\min}(T_{h,i} - T_{c,i})}$	$P = \dfrac{T_{1,o} - T_{1,i}}{T_{2,i} - T_{1,i}}$
$\text{NTU} = \dfrac{UA}{C_{\min}} = \dfrac{1}{C_{\min}} \displaystyle\int_A U dA$	$\text{NTU}_1 = \dfrac{UA}{C_1} = \dfrac{\lvert T_{1,o} - T_{1,i}\rvert}{\Delta T_m}$
$C^* = \dfrac{C_{\min}}{C_{\max}} = \dfrac{(\dot{m}c_p)_{\min}}{(\dot{m}c_p)_{\max}}$	$R = \dfrac{C_1}{C_2} = \dfrac{T_{2,i} - T_{2,o}}{T_{1,o} - T_{1,i}}$
MTD method[a]	
$Q = UAF\,\Delta T_{lm}$	
$\text{LMTD} = \Delta T_{lm} = \dfrac{\Delta T_1 - \Delta T_2}{In(\Delta T_1/\Delta T_2)}$	
$\Delta T_1 = T_{h,i} - T_{c,o} \qquad \Delta T_2 = T_{h,o} - T_{c,i}$	
$F = \phi(P, R, \text{flow arrangement})$	
$F = \dfrac{\Delta T_m}{\Delta T_{lm}}$	
P and *R* are defined in the *P*-NTU method	

[a] Although *P*, *R*, and NTU are defined on Fluid 1 side. It must be emphasized that all the results of the *P-NTU* and MTD methods are valid if the definitions of *P*, NTU, and *R* are consistently based on C_t, C_s, C_h, or C_c.

Here, the exchanger effectiveness ε is an efficiency factor. It is a ratio of the actual heat transfer rate from the hot fluid to the cold fluid in a given heat exchanger of any flow arrangement to the maximum possible heat transfer rate q_{max} thermodynamically permitted by the second law of thermodynamics. The q_{max} is obtained in a *counterflow* heat exchanger (recuperator) of *infinite surface area* operating with the fluid flow rates (heat capacity rates) and fluid inlet temperatures equal to those of an actual exchanger (constant fluid properties are idealized). Refer to Shah and Sekulić (2003) for further details. As noted in Table 13.1, the exchanger effectiveness ε is a function of NTU and C^* in this method. The number of transfer units NTU is a ratio of the overall conductance UA to the smaller heat capacity rate C_{min}. NTU designates the dimensionless "heat transfer size" or "thermal size" of the exchanger. Other interpretations of NTU are given by Shah (1983); Shah and Sekulić (2003). The heat capacity rate ratio C^* is simply a ratio of the smaller to the larger heat capacity rate for the two fluid streams. Note that $0 < \varepsilon < 1$, $0 < \text{NTU} < \infty$ and $0 \leq C^* \leq 1$.

13.2.4 The *P*-NTU Method

This method represents a variant of the ε-NTU method. The ε-NTU relationship is different depending on whether the shell fluid is the C_{min} or C_{max} fluid in the (stream asymmetric) flow arrangements commonly used for shell-and-tube exchangers. In order to avoid possible errors and confusion, an alternative is to present the temperature effectiveness P as a function of NTU and R, where P, NTU, and R are defined consistently for either Fluid 1 side or Fluid 2 side; in Table 13.1, they are defined for the Fluid 1 side (regardless of whether that side is hot or cold fluid side), and Fluid 1 side is clearly defined for each flow arrangement in Table 13.3; it is the shell side in a shell-and-tube exchanger. Note that

$$q = P_1 C_1 \left| T_{1,i} - T_{2,i} \right| = P_2 C_2 \left| T_{2,i} - T_{1,i} \right| \tag{13.14}$$

$$P_1 = P_2 R_2 \quad \text{or} \quad P_2 = P_1 R_1 \tag{13.15}$$

$$\text{NTU}_1 = \text{NTU}_2 R_2 \quad \text{or} \quad \text{NTU}_2 = \text{NTU}_1 R_1 \tag{13.16}$$

$$R_1 = \frac{1}{R_2} \quad \text{or} \quad R_2 = \frac{1}{R_1} \tag{13.17}$$

13.2.5 The MTD Method

In this method, the heat transfer rate from the hot fluid to the cold fluid in the exchanger is given by

$$q = \text{UAF}\Delta T_{\text{lm}} \tag{13.18}$$

Here, the log-mean temperature difference correction factor F is a ratio of mean (actual) temperature difference (MTD) to the log-mean temperature difference (LMTD), where

$$\text{LMTD} = \Delta T_{\text{lm}} = \frac{\Delta T_1 - \Delta T_2}{\ln\left(\Delta T_1 / \Delta T_2\right)} \tag{13.19}$$

Here, ΔT_1 and ΔT_2 are defined as

$$\Delta T_i = \begin{cases} T_{h,i} - T_{c,o} \quad \Delta T_2 = T_{h,o} - T_{c,i} & \text{for all flow arrangements except for parallel flow} \\ T_{h,i} - T_{c,i} \quad \Delta T_2 = T_{h,o} - T_{c,o} & \text{for parallel flow exchanger} \end{cases} \tag{13.20}$$

The LMTD represents a true mean temperature difference for a counterflow arrangement under the idealizations listed next. Thus, the LMTD correction factor F represents a degree of departure for the

MTD from the counterflow LMTD; it does not represent the effectiveness of a heat exchanger. It depends on two dimensionless groups P_1 and R_1 or P_2 and R_2 for a given flow arrangement.

The relationships among the dimensionless groups of the ε-NTU, P-NTU, and MTD methods are presented in Table 13.2. The closed-form formulas for industrially important exchangers are presented in terms of P_1, NTU_1, and R_1 in Table 13.3. These formulas are valid under the following idealizations.

1. The heat exchanger operates under steady-state conditions, that is, constant fluid temperatures (at the inlet and within the exchanger) independent of time.
2. Heat losses to the surroundings are negligible.
3. There are no thermal energy sources and sinks in the exchanger walls or fluids, such as electric heating, chemical reaction, or nuclear processes.
4. In counterflow and parallel flow exchangers, the temperature of each fluid is uniform over every flow cross section. From the temperature distribution point of view, in crossflow exchangers each fluid is considered mixed or unmixed at every cross section depending on the surface geometries used. For a multipass exchanger, the foregoing statements apply to each pass depending on the basic flow arrangement of the passes; the fluid is considered mixed or unmixed between passes.
5. Either there are no phase changes in the fluid streams flowing through the exchanger or the phase changes (condensation or boiling) occur under one of the following conditions: (a) phase change occurs at a constant temperature for a single component fluid at constant pressure; the effective specific heat for the phase-changing fluid is infinity in this case, and hence $C_{\max} \to \infty$. (b) The temperature of the phase-changing fluid varies linearly with heat transfer during the condensation or boiling. In this case, the effective specific heat is constant and finite for the phase-changing fluid.
6. The specific heat of each fluid is constant throughout the exchanger so that the heat capacity rate on each fluid side is treated as constant.
7. The velocity and temperature at the entrance of the heat exchanger on each fluid side are uniform.
8. For an extended surface exchanger, the overall extended surface temperature effectiveness η_0 is considered uniform and constant.
9. The individual and overall heat transfer coefficients are constant throughout the exchanger, including the case of phase-changing fluid in idealization 5.

TABLE 13.2 Relationships between Dimensionless Groups of the P-NTU and MTD Methods and Those of the ε-NTU Method

$$P_1 = \frac{C_{\min}}{C_1}\varepsilon = \begin{cases} \varepsilon & \text{for } C_1 = C_{\min} \\ \varepsilon C^* & \text{for } C_1 = C_{\max} \end{cases}$$

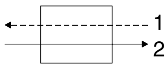
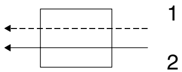
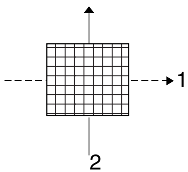
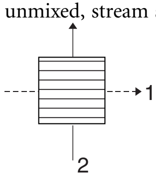
$$R_1 = \frac{C_1}{C_2} = \begin{cases} C^* & \text{for } C_1 = C_{\min} \\ 1/C^* & \text{for } C_1 = C_{\max} \end{cases}$$

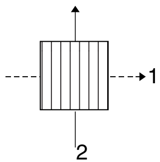
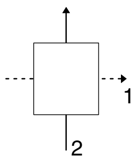
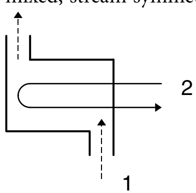
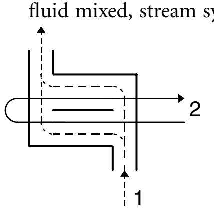
$$\text{NTU}_1 = \text{NTU}\frac{C_{\min}}{C_1} = \begin{cases} \text{NTU} & \text{for } C_1 = C_{\min} \\ \text{NTU}\, C^* & \text{for } C_1 = C_{\max} \end{cases}$$

$$F = \frac{\text{NTU}_{cf}}{\text{NTU}} = \frac{1}{\text{NTU}(1-C^*)}\ln\frac{1-C^*\varepsilon}{1-\varepsilon} \xrightarrow[C^*=1]{} \frac{\varepsilon}{\text{NTU}(1-\varepsilon)}$$

$$F = \frac{1}{\text{NTU}_1(1-R_1)}\ln\left[\frac{1-R_1P_1}{1-P_1}\right] \xrightarrow[R_1=1]{} \frac{P_1}{\text{NTU}_1(1-P_1)}$$

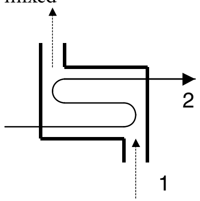
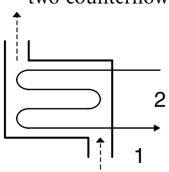
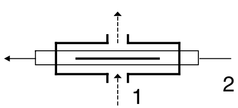
TABLE 13.3 P_1-NTU_1 Formulas and Limiting Values of P_1 for $R_1 = 1$ and $NTU_1 \to \infty$ for Various Exchanger Flow Arrangements

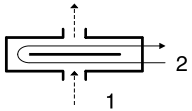
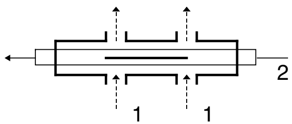
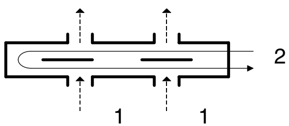
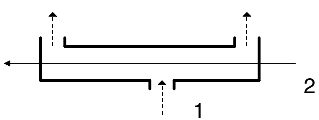
Flow Arrangement	Eq. No.	General Formula[a]	Value for $R_1 = 1$	Value of $NTU_1 \to \infty$
	1.1.1	$P_1 = \dfrac{1-\exp[-NTU_1(1-R_1)]}{1-R_1\exp[-NTU_1(1-R_1)]}$	$P_1 = \dfrac{NTU_1}{1+NTU_1}$	$P_1 \to 1$ for $R_1 \le 1$ $P_1 \to 1/R_1$ for $R_1 \ge 1$
	1.1.2	$NTU_1 = \dfrac{1}{(1-R_1)}\ln\left[(1-R_1P_1)/(1-P_1)\right]$	$NTU_1 = \dfrac{P_1}{1-P_1}$	$NTU_1 \to \infty$
Counterflow exchanger, stream symmetric[b]	1.1.3	$F = 1$	$F = 1$	$F = 1$
	1.2.1	$P_1 = \dfrac{1-\exp[-NTU_1(1+R_1)]}{1+R_1}$	$P_1 = \dfrac{1}{2}[1-\exp(-2NTU_1)]$	$P_1 \to \dfrac{1}{1+R_1}$
	1.2.2	$NTU_1 = \dfrac{1}{1+R_1}\ln\left[\dfrac{1}{1-P_1(1+R_1)}\right]$	$NTU_1 = \dfrac{1}{2}\ln\left[\dfrac{1}{1-2P_1}\right]$	$NTU_1 \to \infty$
Parallel flow exchanger, stream symmetric	1.2.3	$F = \dfrac{(R_1+1)\ln\left[1-R_1P_1/(1-P_1)\right]}{(R_1-1)\ln[1-P_1(1+R_1)]}$	$F = \dfrac{2P_1}{(P_1-1)\ln(1-2P_1)}$	$F \to 0$
	2.1	$P_1 = 1-\exp(-NTU_1)$	Same as Equation 2.1 with $R_1 = 1$	$P_1 \to 1$ for $R_1 \le 1$
		$-\exp[-(1+R_1)NTU_1]$		$P_1 \to \dfrac{1}{R_1}$ for $R_1 \ge 1$
		$\times\displaystyle\sum_{n=1}^{\infty}\dfrac{1}{(n+1)!}R_1^n$		$P_1 \approx \exp\left[\dfrac{NTU_1^{0.22}}{R_1}(e^{-P_1 NTU_1^{0.78}}-1)\right]$
Single-pass crossflow exchanger, both fluids unmixed, stream symmetric		$\times\displaystyle\sum_{j=1}^{n}\dfrac{(n+1-j)}{j!}(NTU_1)^{n+j}$		This approximate equation is accurate within $\pm 1\%$ for $1 < NTU_1 < 7$
	2.2.1	$P_1 = [1-\exp(-KR_1)]/R_1$ $K = 1-\exp(-NTU_1)$	$P_1 = 1-\exp(-K)$	$P_1 \to \dfrac{1-\exp(-R_1)}{R_1}$
	2.2.2	$NTU = \ln\left[\dfrac{1}{1+(1/R_1)\ln(1-R_1P_1)}\right]$	$NTU_1 = \ln\left[\dfrac{1}{1+\ln(1-P_1)}\right]$	$NTU_1 \to \infty$
Single-pass crossflow exchanger, Fluid 1 unmixed, Fluid 2 mixed	2.2.3	$F = \dfrac{\ln[(1-R_1P_1)/(1-P_1)]}{(R_1-1)\ln\left[1+(1/R_1)\ln(1-R_1P_1)\right]}$	$F = \dfrac{P_1}{(P_1-1)\ln[1+\ln(1-P_1)]}$	$F \to 0$
	2.3.1	$P = 1-\exp(-K/R_1)$	$P = 1-\exp(-K)$	$P_1 \to 1-\exp(-1/R_1)$

		$K = 1 - \exp(-R_1 \mathrm{NTU}_1)$	$K = 1 - \exp(-\mathrm{NTU}_1)$	
	2.3.2	$\mathrm{NTU}_1 = \frac{1}{R_1} \ln 1/(1 + R_1) \ln(1 - P_1)$	$\mathrm{NTU}_1 = \ln\left[\frac{1}{1 + \ln(1 - P_1)}\right]$	$\mathrm{NTU}_1 \to \infty$
Single-pass crossflow exchanger, Fluid 1 mixed, Fluid 2 unmixed	2.3.3	$F = \frac{\ln[(1 - R_1 P_1)/(1 - P_1)]}{(1 - 1/R_1)\ln[1 + R_1 \ln(1 - P_1)]}$	$F = \frac{P_1}{(P_1 - 1)\ln[1 + \ln(1 - P_1)]}$	$P_1 \to \frac{1}{1 + R_1}$
	2.4	$P_1 = \left[\frac{1}{K_1} + \frac{R_1}{K_2} - \frac{1}{\mathrm{NTU}_1}\right]^{-1}$ $K_1 = 1 - \exp(-\mathrm{NTU}_1)$ $K_2 = 1 - \exp(-\mathrm{R}_1 \mathrm{NTU}_1)$	$P_1 = \left[\frac{2}{K_1} - \frac{1}{\mathrm{NTU}_1}\right]^{-1}$	$P_1 \to \frac{1}{1 + R_1}$
Single-pass crossflow exchanger, both fluids mixed, stream symmetric				
	3.1.1	$P_1 = \frac{2}{1 + R_1 + \mathrm{E} \coth(\mathrm{E}\,\mathrm{NTU}_1/2)}$ $\mathrm{E} = [1 + R_1^2]^{1/2}$	$P_1 = \frac{1}{1 + \coth(\mathrm{NTU}_1/\sqrt{2})/\sqrt{2}}$	$P_1 \to \frac{2}{1 + R_1 + \mathrm{E}}$
	3.1.2	$\mathrm{NTU}_1 = \frac{1}{E} \ln\left[\frac{2 - P_1(1 + R_1 - E)}{2 - P_1(1 + R_1 + E)}\right]$	$\mathrm{NTU}_1 = \ln\left[\frac{2 - P_1}{2 - 3P_1}\right]$	$\mathrm{NTU}_1 \to \infty$
1-2 TEMA E shell-and-tube exchanger, shell fluid mixed, stream symmetric	3.1.3	$F = \frac{E \ln[(1 - R_1 P_1)/(1 - P_1)]}{(1 - R_1)\ln\left[\frac{2 - P_1(1 + R_1 - E)}{2 - P_1(1 + R_1 + E)}\right]}$	$F = \frac{P_1/(1 - P_1)}{\ln[(2 - P_1)/(2 - 3P_1)]}$	$F \to 0$
	3.2	$P_1 = \frac{1}{R_1}\left[1 - \frac{(2 - R_1)(2\mathrm{E} + R_1 B)}{(2 + R_1)(2\mathrm{E} - R_1/B)}\right]$ $\mathrm{E} = \exp(\mathrm{NTU}_1)$ $\mathrm{B} = \exp(-\mathrm{NTU}_1 R_1/2)$ Same as 1-1 J shell, Equation 3.10	$P_1 = \frac{1}{2}\left[1 - \frac{1 + \mathrm{E}^{-2}}{2(1 + \mathrm{NTU}_1)}\right]$ for $R_1 = 2$	$P_1 \to \frac{2}{2 + R_1}$ for $R_1 \le 2$ $P_1 \to \frac{1}{R_1}$ for $R_1 \ge 2$

(continued)

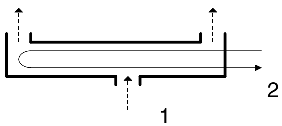
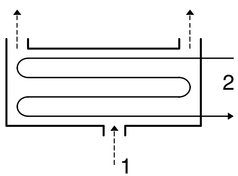
TABLE 13.3 ***(Continued)***

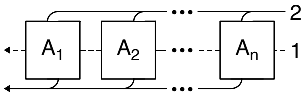
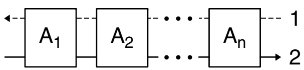
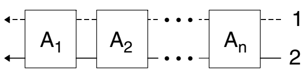
Flow Arrangement	Eq. No.	General Formula[a]	Value for $R_1 = 1$	Value of $\mathrm{NTU}_1 \to \infty$
1-2 TEMA E shell-and-tube exchanger, shell fluid divided into two streams individually mixed 	3.3	$P_1 = \frac{1}{R_1}\left[1 - \frac{C}{AC + B^2}\right]$ $A = X_1(R_1 + \lambda_1)(R_1 - \lambda_2)/2\lambda_1 - X_3\delta$ $\quad -X_2(R_1 + \lambda_2)(R_1 - \lambda_1)/2\lambda_2 + 1/(1 - R_1)$ $B = X_1(R_1 - \lambda_2) - X_2(R_1 - \lambda_1) + X_3\delta$ $C = X_2(3R_1 + \lambda_1) - X_1(3R_1 + \lambda_2) + X_3\delta$ $X_i = \exp(\lambda_1 \mathrm{NTU}_1/3)/2\delta, \quad i = 1,2,3$ $\delta = \lambda_1 - \lambda_2$ $\lambda_1 = -\frac{3}{2} + \left[\frac{9}{4} + R_1(R_1 - 1)\right]^{1/2}$ $\lambda_2 = -\frac{3}{2} - \left[\frac{9}{4} + R_1(R_1 - 1)\right]^{1/2}$ $\lambda_3 = R_1$	Same as Equation 3.3 with $R_1 = 1$ $A = -\exp(-\mathrm{NTU}_1)/18$ $\quad -\exp(\mathrm{NTU}_1/3)/2$ $\quad +(\mathrm{NTU}_1 + 5)/9$	$P_1 \to 1$ for $R_1 \leq 1$ $P_1 \to \frac{1}{R_1}$ for $R_1 \geq 1$
1-3 TEMA E shell-and-tube exchanger, shell and tube fluids mixed, one parallel flow and two counterflow passes 	3.4	$P_1 = 4[2(1 + R_1) + \mathrm{D}A + R_1 B]^{-1}$ $A = \coth(\mathrm{D}\,\mathrm{NTU}_1/4)$ $B = \tanh(R_1 \mathrm{NTU}_1/4)$ $\mathrm{D} = [4 + R_1^2]^{1/2}$	$P_1 = 4[4 + \sqrt{5}A + B]^{-1}$ $A = \coth(\sqrt{5}\mathrm{NTU}_1/4)$ $B = \tanh(\mathrm{NTU}_1/4)$	$P_1 \to \frac{4}{2(1 + 2R_1) + \mathrm{D} - R_1}$
1-4 TEMA E shell-and-tube exchanger, shell and tube fluids mixed				
1-4 TEMA E shell-and-tube exchanger, shell and tube fluids mixed with $n \to \infty$	3.5	Equation 2.4 applies in this limit with $n \to \infty$	Same as for Equation 2.4	Same as for Equation 2.4
	3.6	$P_1 = A + B - AB(1 + R_1) + R_1 AB^2$ $A = \frac{1}{1 + R_1}\{1 - \exp[-\mathrm{NTU}_1(1 + R_1)/2\}$ $B = \frac{1 - D}{1 - R_1 D}$	Same as Equation (3.6) with $B = \mathrm{NTU}_1/(2 + \mathrm{NTU}_1)$ for $R_1 = 1$	$P_1 \to 1$ for $R_1 \leq 1$ $P_1 \to \frac{1}{R_1}$ for $R_1 \geq 1$

		$D=\exp[-\mathrm{NTU}_1(1-R_1)/2]$		
1-1 TEMA G shell-and-tube exchanger, tube fluid split into two streams individually mixed, shell fluid mixed. Stream symmetric. 	3.7	$P_1=(B-\alpha^2)/(A+2+R_1B)$ $A=-2R_1(1-\alpha)^2/(2+R_1)$ $B=[4-\beta(2+R_1)]/(2-R_1)$ $\alpha=\exp[-\mathrm{NTU}_1(2+R_1)/4]$ $\beta=\exp[-\mathrm{NTU}_1(2-R_1)/2]$	$P_1=\dfrac{1+2\mathrm{NTU}_1-\alpha^2}{4+4\mathrm{NTU}_1-(1-\alpha)^2}$ for $R_1=2$ $\alpha=\exp(-\mathrm{NTU}_1)$	$P_1\to\dfrac{2+R_1}{R_1^2+R_1+2}$ for $R_1\le 2$ $P_1\to\dfrac{1}{R_1}$ for $R_1\ge 2$
Overall counterflow 1-2 TEMA G shell-and-tube exchanger. Shell and tube fluids mixed in each pass at a cross section. 	3.8	$P_1=E[1+(1-\mathrm{B}R_1/2)(1-\mathrm{A}R_1/2+\mathrm{AB}R_1)]-\mathrm{AB}(1-\mathrm{B}R_1/2)$ $A=\dfrac{1}{1+R_1/2}\{1-\exp[-\mathrm{NTU}_1(1+R_1/2)/2]\}$ $B=(1-\mathrm{D})/(1-R_1\mathrm{D}/2)$ $D=\exp[-\mathrm{NTU}_1(1-R_1/2)/2]$ $\mathrm{E}=(A+B-\mathrm{AB}R_1/2)/2$	Same as Equation (3.8) with $B=\mathrm{NTU}_1/(2+\mathrm{NTU}_1)$ for $R_1=2$	$P_1\to\dfrac{4(1+R_1)-R_1^2}{(2+R_1)^2}$ for $R_1\le 2$ $P_1\to\dfrac{1}{R_1}$ for $R_1\ge 2$
1-1 TEMA H shell-and-tube exchanger, tube fluid split into two streams individually mixed, shell fluid mixed. 	3.9	$P_1=\dfrac{1}{R_1}\left[1-\dfrac{(1-D)^4}{B-4G/R_1}\right]$ $B=(1+H)\ (1+E)^2$ $G=(1-D)^2\ (D^2+E^2)+D^2(1+E)^2$ $H=[1-\exp(-2\beta)]/(4/R_1-1)$ $E=[1-\exp(-\beta)]/(4/R_1-1)$ $D=[1-\exp(-\alpha)]/(4/R_1+1)$ $\alpha=\mathrm{NTU}_1\ (4+R_1)/8$ $\beta=\mathrm{NTU}_1\ (4-R_1)/8$	Same as Equation 3.11 with $H=\mathrm{NTU}_1+1$ $E=\mathrm{NTU}_1/2$ for $R_1=4$	$P_1\to\left[R_1+\dfrac{(4-R_1)^3}{(4+R_1)R_1^3+16)}\right]^{-1}$ $P_1\to\frac{1}{R_1}$ for $R_1\ge 4$
Overall counterflow 1-2 TEMA H shell-and-tube exchanger, shell and tube fluids mixed in each pass at a cross section 	3.10	$P_1=\dfrac{1}{R_1}\left[1-\dfrac{(2-R_1)(2A+R_1B)}{(2+R_1)(2A-R_1/B)}\right]$ $A=\exp(\mathrm{NTU}_1)$ $B=\exp(-\mathrm{NTU}_1R_1/2)$	$P_1=\dfrac{1}{2}\left[1-\dfrac{1+A^{-2}}{2(1+\mathrm{NTU}_1)}\right]$ for $R_1=2$	$P_1\to\dfrac{2}{2+R_1}$ for $R_1\le 2$ $P_1\to\frac{1}{R_1}$ for $R_1\ge 2$
1-1 TEMA J shell-and-tube exchanger, shell and tube fluids mixed		Same as Equation 3.2		

(continued)

TABLE 13.3 ***(Continued)***

Flow Arrangement	Eq. No.	General Formula[a]	Value for $R_1 = 1$	Value of $\text{NTU}_1 \to \infty$
 1-2 TEAM J shell-and-tube exchanger, shell and tube fluids mixed; the same; the same results of Fluid 2 reversed	3.11	$P_1 = \left[1 + \frac{R_1}{2} + \lambda B - 2\lambda CD\right]^{-1}$ $B = (A^{\lambda} + 1)/(A^{\lambda} - 1)$ $C = \frac{A^{(1+\lambda)/2}}{\lambda - 1 + (1+\lambda)A^{2}}$ $D = 1 + \frac{\lambda A^{(\lambda-1)/2}}{A^{\lambda} - 1}$ $A = \exp(\text{NTU}_1)$ $\lambda = (1 - R_1^2/4)^{1/2}$	Same as Equation 3.13 with $R_1 = 1$	$P_1 \to \left[1 + \frac{R_1}{2} + \lambda\right]^{-1}$
 1-4 TEMA J shell-and-tube exchanger, shell and tube fluids mixed	3.12	$P_1 = \left[1 + \frac{R_1}{4}\left(\frac{1+3E}{1+E}\right) + \lambda B - 2\lambda CD\right]^{-1}$ $B = \frac{A^{\lambda} + 1}{A^{\lambda} - 1}$ $C = \frac{A^{(1+\lambda)/2}}{\lambda - 1 + (1+\lambda)A^{\lambda}}$ $D = 1 + \frac{\lambda A^{(\lambda-1)/2}}{A^{\lambda} - 1}$ $A = \exp(\text{NTU}_1)$ $E = \exp(R_1\ \text{NTU}_1/2)$ $\lambda = (1 + R_1^2/16)^{1/2}$	Same as Equation 3.14 with $R_1 = 1$	$P_1 \to \left[1 + \frac{3R_1}{4} + \lambda\right]^{-1}$
Limit of $1-n$ TEMA J shell-and-tube exchangers for $n \to \infty$. Shell and tube fluids mixed, stream symmeteric	3.13	Equation 2.4 applies in this limit	Same as for Equation 2.4	Same as for Equation 2.4

	4.1.1	$P_1 = 1 - \prod_{i=1}^{n}(1 - P_{1,\,A_i})$	Same as Equation 4.1.1	Same as Equation 4.1.1
	4.1.2	$\frac{1}{R_1} = \sum_{i=1}^{n} R_{1,\,A_i}$	$1 = \sum_{i=1}^{n} R_{1,\,A_i}$	Same as Equation 4.1.2
Parallel coupling of *n* exchangers. Fluid 2 split arbitrarily into *n* streams	4.1.3	$\mathrm{NTU}_1 = \sum_{i=1}^{n} \mathrm{NTU}_{1,\,A_i}$	Same as Equation 4.1.3	$\mathrm{NTU}_1 \to \infty$
	4.2.1	$P_1 = \dfrac{\prod_{i=1}^{n}(1 - R_1 P_{1,A_i}) - \prod_{i=1}^{n}(1 - P_{1,\,A_i})}{\prod_{i=1}^{n}(1 - R_1 P_{1,\,A_i}) - R_1 \prod_{i=1}^{n}(1 - P_{1,\,A_i})}$	$P_1 = \dfrac{\sum_{i=1}^{n} \frac{P_{1,\,A_i}}{1-P_{1,\,A_i}}}{1 + \sum_{i=1}^{n} \frac{P_{1,\,A_i}}{1-P_{1,\,A_i}}}$	Same as Equation 1.2.1 counterflow
	4.2.2	$R_1 = R_{1,\,A_i} \quad i = 1,\ldots,n$	$1 = R_{1,\,A_i}, \quad i = 1,\ldots,n$	Same as Equation 4.2.2
	4.2.3	$\mathrm{NTU}_1 = \sum_{i=1}^{n} \mathrm{NTU}_{1,\,A_i}$	Same as Equation 4.2.3	Same as Equation 4.2.3
Series coupling of *n* exchangers, overall counterflow arrangement, stream symmetric if all A_i are steam symmetric	4.2.4	$F = \frac{1}{\mathrm{NTU}_1} \sum_{i=1}^{n} \mathrm{NTU}_{1,\,A_i} F_{A_i}$	Same as Equation 4.2.4	Same as Equation 4.2.4
	4.3.1	$P = \frac{1}{1+R_1}\left\{1 - \prod_{i=1}^{n}[1 - (1 + R_1)P_{1,\,A_i}]\right\}$	$P_1 = \dfrac{1 - \prod_{i=1}^{n}[1 - 2P_{1,\,A_i}]}{2}$	Same as Equation 4.3.1
	4.3.2	$R_1 = R_{1,A_i} \quad i = 1,\ldots,n$	$1 = R_{1,\,A_i} \quad i = 1,\ldots,n$	Same as Equation 4.3.2
Series coupling of *n* exchangers, overall parallel flow arrangement; stream symmetric if all A_i are stream symmetric	4.3.3	$\mathrm{NTU}_1 = \sum_{i=1}^{n} \mathrm{NTU}_{1,\,A_i}$	Same as Equation 4.3.3	$\mathrm{NTU}_1 \to \infty$

In this table all variables except P_1, R_1, NTU_1, and F are local or dummy variables not necessarily related to similar ones defined in nomenclature and the text.

Reprinted from Shah, R. K. and Muller, A. C., *Ullmann's Encyclopedia of Industrial Chemistry, Unit Operations II, Vol. B3*. VCH Publishers, Weinheim, Germany, Chap. 2, 1988. With permission.

[a] All the formulas in this table are based on the Fluid 1 side. They can be converted to the Fluid side 2 using the following relations: (1) for stream symmetric exchangers, change P_1, NTU_1 and R_1 to P_2, NTU_2 and R_2. (2) For stream asymmetric exchangers, convert P_1-NTU_1-R_1, expressions to P_2-NTU_2-R_2 expressions using the following relationships: $P_1 = P_2R_2$, $\mathrm{NTU}_1 = \mathrm{NTU}_2R_2$, and $R_1 = 1/R_2$.

[b] For those flow arrangements where "stream symmetric" is not explicitly mentioned, they are asymmetric.

10. The heat transfer surface area, A is distributed uniformly on each fluid side. In a multipass unit, heat transfer surface area is distributed uniformly in each pass, although different passes can have different surface areas.
11. For a plate-baffled shell-and-tube exchanger, the temperature rise per baffle pass is small compared to the overall temperature rise along the exchanger; that is, the number of baffles is large. Thus, the shell fluid is treated as mixed at its every flow cross section.
12. The fluid flow rate is uniformly distributed through the exchanger on each fluid side in each pass. No stratification, flow bypassing, or flow leakages occur in any stream. The flow condition is characterized by the bulk (or mean) velocity at any cross section.
13. Longitudinal heat conduction in the fluid and the wall is negligible.
14. Wall thermal resistance is distributed uniformly in the entire heat exchanger.

Idealizations 1–4 are necessary in a theoretical analysis of steady-state heat exchangers. Idealization 5 essentially restricts the analysis to single-phase flow on both fluid sides or on one fluid side with a dominating thermal resistance. For two-phase flows on both fluid sides, many of the foregoing idealizations are not valid, since mass transfer in phase change results in variable properties and variable flow rates of each phase, and the heat transfer coefficients vary significantly. As a result, the heat exchanger cannot be analyzed using the theory presented here.

If idealization 6 is not valid, divide the exchanger into small segments until the specific heats can be treated as constant. Idealizations 7 and 8 are primarily important for compact heat exchangers and will be discussed, respectively, in Sections on Flow Maldistribution, and Fin Efficiency and Extended Surface Efficiency.

Idealization 9 has been discussed in detail by Shah (1993). The overall heat transfer coefficient can vary because of variations in local heat transfer coefficients due to two effects: (1) change in heat transfer coefficients in the exchanger as a result of changes in the fluid properties or radiation due to rise or drop of fluid temperatures, and (2) change in heat transfer coefficients in the exchanger due to developing thermal boundary layers; this is referred to as the *length effect*. The first effect due to fluid property variations (or radiation) consists of two components: (i) distortion of velocity and temperature profiles at a given flow cross section due to fluid property variations; this effect is usually taken into account by the so-called property ratio method, with the correction scheme of Equation 13.103 and Equation 13.104, and (ii) variations in the fluid temperature along the axial and transverse directions in the exchanger depending on the exchanger flow arrangement; this effect is referred to as the *temperature effect*. The resultant axial changes in the overall mean heat transfer coefficient can be significant; the variations in U_{local} could be nonlinear depending on the type of the fluid. The effect of varying U_{local} can be taken into account by evaluating U_{local} at a few points in the exchanger and subsequently integrating U_{local} values by Simpson or Gauss method (Shah 1993). The temperature effect can increase or decrease mean U slightly or significantly depending on the fluids and applications. The length effect is important for developing laminar flows for which high heat transfer coefficients are obtained in the thermal entrance region. However, in general, it will have less impact on the overall heat transfer coefficient because the other thermal resistances in series in an exchanger may be controlling. The length effect reduces then overall heat transfer coefficient compared to the mean value calculated conventionally (assuming uniform mean heat transfer coefficient on each fluid side). It is shown that this reduction is up to about 14% for the worst case (Shah 1993).

Idealization 10 is generally true for individual passes due to the manufacturing considerations. However, in a multipass unit, it is possible to have different surface areas in different passes due to the design considerations. This effect has been investigated by Roetzel and Spang (1987, 1989) for 1-2, 1-3, and 1-2N TEMA E exchangers, by Spang, Xuan, and Roetzel (1991) for 1-N split flow TEMA G exchangers, by Xuan, Spang, and Roetzel (1991) for 1-N divided flow TEMA J exchangers, and by Bačlić, Romie, and Herman (1988) for two-pass cross-counterflow exchangers.

For Idealization 11, Shah and Pignotti (1997) have shown that the following are the specific number of baffles beyond which the influence of the finite number of baffles on the exchanger effectiveness is not significantly larger than 2%: $N_b \geq 10$ for the 1-1 TEMA E counterflow exchanger; $N_b \geq 6$ for the 1-2 TEMA E exchanger for $NTU_s \leq 2$, $R_s \leq 5$; $N_b \geq 9$ for the 1-2 TEMA J exchanger for $NTU_s \leq 2$, $R_s \leq 5$; $N_b \geq 5$ for the 1-2 TEMA G exchanger for $NTU_s \leq 3$, all R_s; $N_b \geq 11$ for the 1-2 TEMA H exchanger for $NTU_s \leq 3$, all R_s.

If any of these idealizations are not valid for a particular exchanger application, the best solution is to work directly with either Equation 13.1 and Equation 13.2 or their modified forms by including a particular effect, and to integrate them over a small exchanger segment (i.e., conduct finite difference type analysis) in which all of the idealizations are valid.

13.2.6 Fin Efficiency and Extended Surface Efficiency

Extended surfaces have fins attached to the primary surface on one or both fluid sides of a two-fluid or a multifluid heat exchanger. Fins can be of a variety of geometries—plain, wavy, or interrupted, and can be attached to the inside, outside, or both sides of circular, flat, or oval tubes, or parting sheets. Fins are primarily used to increase the surface area (when the heat transfer coefficient on that fluid side is relatively low) and consequently to increase the total heat transfer rate. In addition, enhanced fin geometries also increase the heat transfer coefficient compared to that for a plain fin. Fins may also be used on the high heat transfer coefficient fluid side in a heat exchanger primarily for structural strength purpose (for example, for high-pressure water flow through a flat tube) or to provide a thorough mixing of a highly viscous liquid (such as for laminar oil flow in a flat or a round tube). Fins are attached to the primary surface by brazing, soldering, welding, adhesive bonding, or mechanical expansion, or they are extruded or integrally connected to the tubes. Major categories of extended surface heat exchangers are plate-fin (Figure 13.5) and tube-fin (Figure 13.6). Note that shell-and-tube exchangers sometimes employ individually finned tubes—low finned tubes (similar to Figure 13.6a but with low height fins; Shah 1985).

The concept of fin efficiency takes into account the reduction in temperature potential between the fin and the ambient fluid due to conduction along the fin and convection from or to the fin surface depending on fin cooling or heating situation. The *fin efficiency* is defined as the ratio of the actual heat transfer rate through the fin base divided by the maximum possible heat transfer rate through the fin base that would be obtained if the entire fin were at the base temperature (i.e., its material thermal conductivity were infinite). Since most of the real fins are "thin," they are treated as one-dimensional

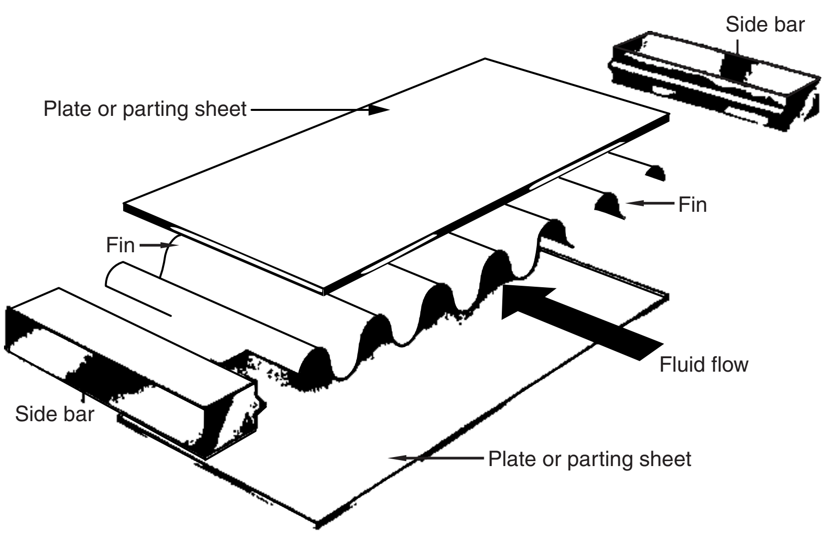


FIGURE 13.5 Basic components of a plate-fin exchanger.

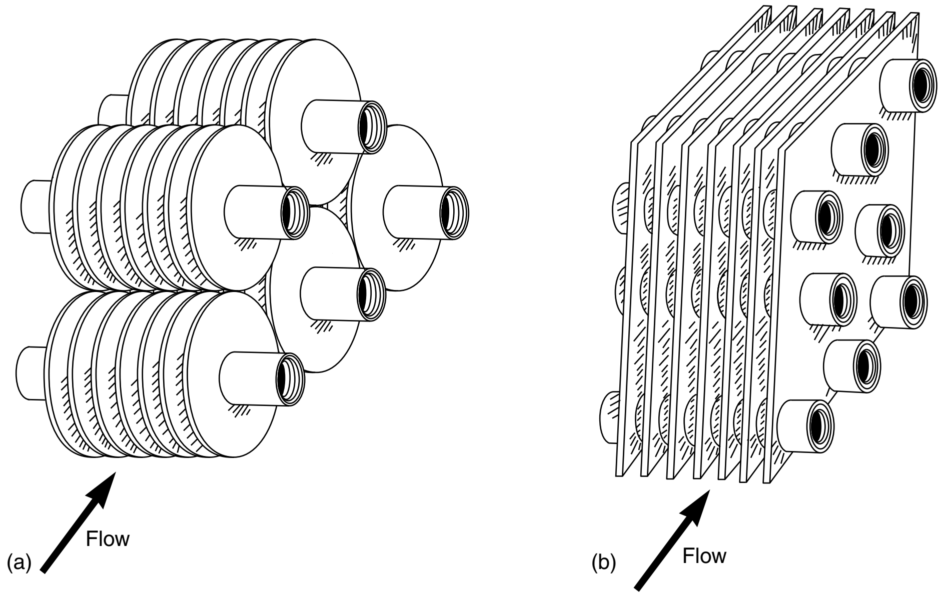


FIGURE 13.6 Types of tube-fin exchangers: (a) individually finned tubes, (b) flat (continuous) fins on an array of tubes; flat fins shown as plain fins, but can be wavy, louvered or interrupted.

(1D) with standard idealizations used for the analysis (Huang and Shah 1992). This 1D fin efficiency is a function of the fin geometry, fin material thermal conductivity, heat transfer coefficient at the fin surface, and the fin tip boundary condition; it is not a function of the fin base or fin tip temperature, ambient temperature, and heat flux at the fin base or fin tip in general. Fin efficiency formulas for some common fins are presented in Table 13.4 (Shah 1985).

The fin efficiency for flat plain fins on inline and staggered tube arrangements is obtained by the sector method. The smallest representative segment of the fin of Figure 13.7c is divided into two parts, OAB and OBC; the OAB part (having the subtended angle θ_0) is then divided into m equal angle ($\Delta\theta = \theta_0/m$) segments. Similarly, the OBC part (having the subtended angle ϕ_0) is then divided into n equal angle ($\Delta\phi = \phi_0/n$) segments. The outer radius of each circular sector is determined by equating the area of the sector with the area of the equivalent annular sector (Kundu and Das 2000). Thus for the inline tube arrangement of Figure 13.7a and c, it is given by

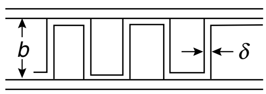
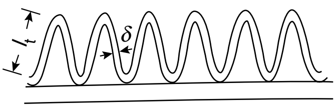
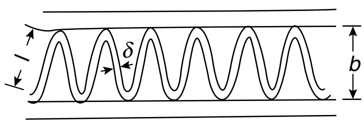
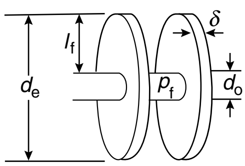
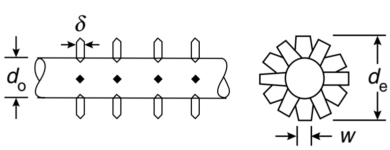
$$r_{e,i} = \frac{X_t}{2}\left(\frac{\tan(i\Delta\theta) - \tan[(i-1)\Delta\theta]}{\Delta\theta}\right)^{1/2} \qquad r_{e,j} = \frac{X_\ell}{2}\left(\frac{\tan(j\Delta\phi) - \tan[(j-1)\Delta\phi]}{\Delta\phi}\right)^{1/2} \tag{13.21a}$$

The smallest representative segment of the staggered tube arrangement of Figure 13.7b is shown in Figure 13.7d, which is divided into two parts, OAD and ODF; and ODF is divided into two equal parts ODE and OEF as shown in Figure 13.7d. Here again $r_{e,i}$ and $r_{e,j}$ of ith and jth segments of OAD and ODE are given by

$$\begin{aligned} r_{e,i} &= X_\ell\left(\frac{\tan(i\Delta\theta) - \tan[(i-1)\Delta\theta]}{\Delta\theta}\right)^{1/2} \\ r_{e,j} &= \frac{\left[X_\ell^2 + (X_t/2)^2\right]^{1/2}}{2}\left(\frac{\tan(j\Delta\phi) - \tan[(j-1)\Delta\phi]}{\Delta\phi}\right)^{1/2} \end{aligned} \tag{13.21b}$$

The fin efficiency of each sector is then determined by the circular fin of constant cross section of Table 13.4. Once η_f for each sector is determined, η_f for the whole fin is the surface area weighted average

TABLE 13.4 Fin Efficiency Expressions for Plate-Fin and Tube-Fin Geometries of Uniform Fin Thickness

Geometry	Fin efficiency formula
	$m_1 = \left[\frac{2h}{k_r\delta_i}\left(1+\frac{\delta_i}{l_f}\right)\right]^{1/2} E_1 = \frac{\tanh(m_i l_i)}{m_i l_i}\; i = 1, 2, 3, 4$
 Plain, wavy, or offset strip fin of rectangular cross section	$\eta_f = E_1$ $l_1 = \frac{b}{2} - \delta_1 \quad \delta_1 = \delta$
 Triangular fin heated/cooled from one side	$\eta_f = \frac{hA_1(T_0 - T_a)\frac{\sinh(m_1 l_1)}{m_1 l_1} q_e}{\cosh(m_1 l_1)\left[hA_1(T_0 - T_a) + q_e\frac{T_0 - T_a}{T_1 - T_b}\right]}, \delta_1 = \delta$
 Plain, wavy, or louver fin of triangular cross section	$\eta_f = E_1$ $l_1 = \frac{l}{2} \quad \delta_1 = \delta$
 Circular fin	$\eta_f = \begin{cases} a(ml_e)^{-b} & \text{for } \Phi > 0.6 + 2.257(r^*)^{-0.445} \\ \frac{\tanh\Phi}{\Phi} & \text{for } \Phi \le 0.6 + 2.257(r^*)^{-0.445} \end{cases}$ $a = (r^*)^{-0.246}\; \Phi = ml_e(r^*)^{\exp\{0.13ml_e - 1.3863\}}$ $b = \begin{cases} 0.9107 + 0.0893r^* & \text{for } r^* \le 2 \\ 0.9706 + 0.17125\ln r^* & \text{for } r^* > 2 \end{cases}$ $m = \left(\frac{2h}{k_f\delta}\right)^{1/2}$ $l_e = l_f + \frac{\delta}{2}\; r^* = \frac{r_e}{r_o}$
 Studded fin	$\eta_f = \frac{\tanh(ml_e)}{ml_e}$ $m = \left[\frac{2h}{k_f\delta}\left(1+\frac{\delta}{w}\right)\right]^{1/2}\; l_e = l_f + \frac{\delta}{2} \cdot l_f = \frac{(d_e - d_o)}{2}$
Rectangular fin over circular tubes	See the text

of η_fs for each sector.

$$\eta_f = \frac{\sum_{i=1}^{m} \eta_{f,i}A_{f,i} + a\sum_{j=1}^{n} \eta_{f,j}A_{f,j}}{\sum_{i=1}^{m} A_{f,i} + a\sum_{j=1}^{n} A_{f,j}} \tag{13.22}$$

Here, $a = 1$ for inline arrangement (Figure 13.7c) for segment OBC and $a = 2$ for staggered arrangement (Figure 13.7d) for two equal segments ODE and OEF. This approximation improves as the number of sectors $m, n \rightarrow \infty$. However, in reality, only a few sectors m and n will suffice to provide η_f within the

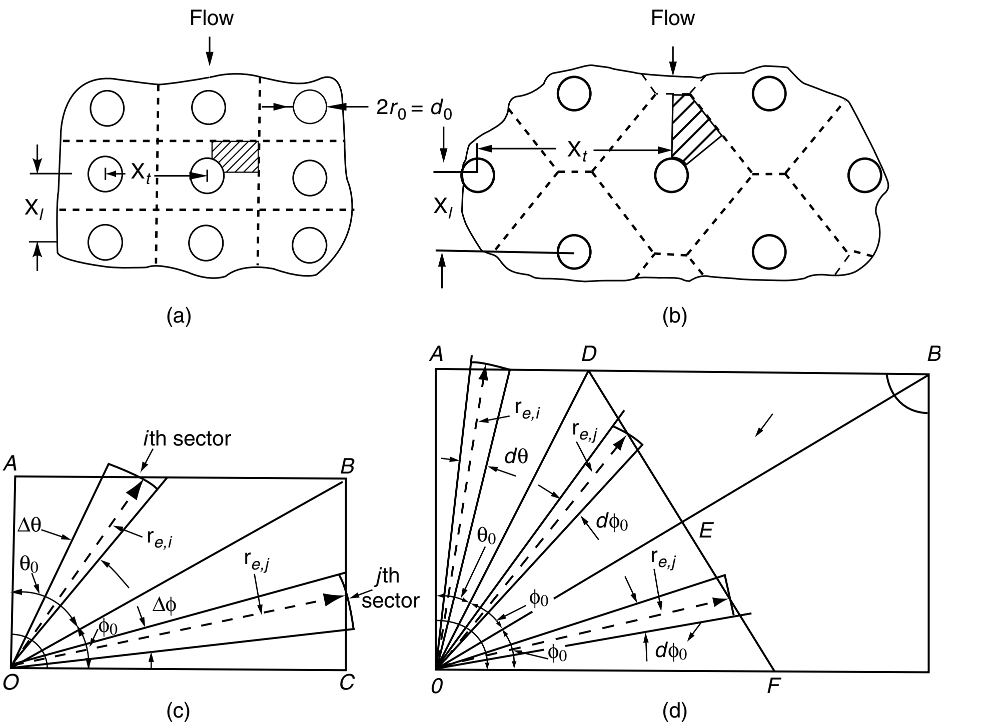


FIGURE 13.7 Flat fin over (a) an inline and (b) staggered arrangement. Smallest representative shaded segment of the fin for (c) an inline and (d) a staggered tube arrangement.

desired accuracy (such as 0.1%). An implicit idealization made in this method is that the heat flow is only in the radial direction, and not in the path of least thermal resistance. Hence, η_f calculated by the sector method will be lower than that for the actual flat fin, a conservative value.

The η_f values of Table 13.4 or Equation 13.22 are not valid in general when the fin is thick, is subject to variable heat transfer coefficients or variable ambient fluid temperature, or has a considerable temperature depression at the base. For a thin rectangular fin of constant cross section, the fin efficiency as presented in Table 13.4 is given by

$$\eta_f = \frac{\tanh(m\ell)}{m\ell} \tag{13.23}$$

where $m = [2h(1+\delta/\ell_i)/k_f\delta_f]^{1/2}$. For a thick rectangular fin of constant cross section, the fin efficiency (a counterpart of Equation 13.23) is given by (Huang and Shah 1992)

$$\eta_f = \frac{(\mathrm{Bi}^+)^{1/2}}{K\mathrm{Bi}} \tanh\left[K\left(\mathrm{Bi}^+\right)^{1/2}\right] \tag{13.24}$$

where $\mathrm{Bi}^+ = Bi/(1+\mathrm{Bi}/4)$, $\mathrm{Bi} = h\delta_f/2k_f$, $K = 2\ell/\delta_f$. Equation 13.23 is accurate (within 0.3%) for a "thick" rectangular fin having $\eta_f > 80\%$; otherwise use Equation 13.24 for a thick fin.

The nonuniform heat transfer coefficient h over the fin surface can lead to a significant error in η_f (Huang and Shah 1992) compared to that for a uniform h over the fin surface. However, generally h is obtained experimentally by considering a constant (uniform) value of h over the fin surface and also associated primary surface area. Hence, such experimental h will not introduce significant errors in η_f while designing a heat exchanger, particularly for $\eta_f > 80\%$. However, one needs to be aware of the impact of nonuniform h on η_f if the heat exchanger test conditions and design conditions are significantly different. Nonuniform ambient temperature has less than a 1% effect on the fin efficiency for $\eta_f > 60\%$ and hence can be neglected. The longitudinal heat conduction effect on the fin efficiency is less than 1% for $\eta_f > 10\%$ and hence can be neglected. The fin base temperature depression increases the total heat flow rate through the extended surface compared to that having no fin base temperature depression, and hence neglecting this effect provides a conservative approach for the extended surface heat transfer. Refer to Huang and Shah (1992) for further details on the foregoing effects and modifications to η_f for rectangular fins of constant cross sections. Refer to Srinivasan and Shah (1995) for fin efficiency of extended surfaces for two-phase flows.

In an extended surface heat exchanger, heat transfer takes place from both the fins ($\eta_f < 100\%$) and the primary surface ($\eta_f = 100\%$). In that case, the total heat transfer rate is evaluated through a concept of *extended surface efficiency* η_0 defined as

$$\eta_0 = \frac{A_p}{A} + \eta_f \frac{A_f}{A} = 1 - \frac{A_f}{A}\left(1 - \eta_f\right) \tag{13.25}$$

where A_f is the fin surface area, A_p is the primary surface area, and $A = A_f + A_p$ is the total heat transfer area. In Equation 13.25, heat transfer coefficients over the finned and primary surfaces are idealized to be equal as noted earlier. Note that $\eta_0 \geq \eta_f$, and η_0 is always required for the determination of thermal resistances of Equation 13.5 in heat exchanger analysis.

13.2.7 Pressure Drop Analysis

Usually a fan, blower, or pump is used to flow fluid through individual fluid sides of a heat exchanger. Due to potential initial and operating high cost, low fluid pumping power requirement is highly desired for gases and viscous liquids. The fluid pumping power $\mathcal{P}$ is approximately related to the core pressure drop in the exchanger as (Shah 1985).

$$\mathcal{P} = \frac{\dot{m}\Delta p}{\rho\eta_p} \approx \begin{cases} \dfrac{1}{2g_c\eta_p}\dfrac{\mu}{\rho^2}\dfrac{4L}{D_h}\dfrac{\dot{m}^2}{D_hA_o} f\,Re \\[2ex] \dfrac{0.046}{2g_c\eta_p}\dfrac{\mu^{0.2}}{\rho^2}\dfrac{4L}{D_h}\dfrac{\dot{m}^{2.8}}{D_h^{0.2}A_o^{1.8}} \end{cases} \tag{13.26}$$

where η_p is the pump/fan efficiency.

It is clear from Equation 13.26 and Equation 13.26 that the fluid pumping power is strongly dependent on the fluid density ($\mathcal{P} \propto 1/\rho^2$), particularly for low-density fluids in laminar and turbulent flows, and on the viscosity μ in laminar flow. In addition, the pressure drop itself can be an important consideration when blowers and pumps are used for the fluid flow, since they are head limited. Also, for condensing and evaporating fluids, the pressure drop affects the heat transfer rate. Hence, the pressure drop determination in the exchanger is important.

The pressure drop associated with a heat exchanger consists of (1) core pressure drop and (2) the pressure drop associated with the fluid distribution devices, such as inlet and outlet manifolds, headers, tanks, nozzles, ducting, and so on, which may include bends, valves, and fittings. This second Δp component is determined from Idelchik (1994); Miller (1990). The core pressure drop may consist of one or more of the following components depending on the exchanger construction: (i) friction losses associated with fluid flow over heat transfer surface, which usually consists of skin friction, form (profile) drag, and internal contractions and expansions, if any; (ii) the momentum effect (pressure drop or rise due to fluid density changes) in the core, (iii) pressure drop associated with sudden contraction and expansion at the core inlet and outlet, and (iv) the gravity effect due to the change in elevation between the inlet and outlet of the exchanger. The gravity effect is generally negligible for gases. For vertical flow through the exchanger, the pressure drop or rise ("static head") due to the elevation change is given by

$$\Delta p = \pm \frac{\rho_m g L}{g_c} \tag{13.27}$$

Here, the "+" sign denotes vertical upflow (i.e., pressure drop) and the "−" sign denotes vertical downflow (i.e., pressure rise). The first three components of the core pressure drop are now presented for plate-fin, tube-fin, plate, and regenerative heat exchangers.

13.2.7.1 Plate-Fin Heat Exchangers

For the plate-fin exchanger (Figure 13.5), all three components are considered in the core pressure drop evaluation as follows.

$$\frac{\Delta p}{p_i} = \frac{G^2}{2 g_c \rho_i p_i} \left[\underbrace{(1 - \sigma^2 + K_c)}_{\text{entrance effect}} + \underbrace{\left(f \frac{L}{r_h} \rho_i \left(\frac{1}{\rho}\right)_m \right)}_{\text{core friction}} + 2 \underbrace{\left(\frac{\rho_i}{\rho_o} - 1\right)}_{\text{momentum effect}} - \underbrace{(1 - \sigma^2 - K_e)}_{\text{exit effect}} \frac{\rho_i}{\rho_o} \right] \tag{13.28}$$

where f is the Fanning fraction factor, K_c and K_e are flow contraction (entrance) and expansion (exit) pressure loss coefficients, respectively (see Figure 13.8), and σ is a ratio of minimum free flow area to frontal area. K_c and K_e for four different entrance flow passage geometries are presented by Kays and London (1998). The entrance and exit losses are important at low values of σ and L (short cores) and for high values of Re for gases; they are negligible for liquids. The values of K_c and K_e apply to long tubes for which flow is fully-developed at the exit. For partially developed flows, K_c is lower and K_e is higher than that for fully-developed flows. For interrupted surfaces, flow is never fully-developed boundary layer type. For highly interrupted fin geometries, the entrance and exit losses are generally small compared to the core pressure drop and the flow is well mixed; hence, K_c and K_e for Re → ∞ should represent a good approximation. The mean specific volume v_m or $(l/\rho)_m$ in Equation 13.28 is given as follows: for liquids with any flow arrangement, or for a perfect gas with $C^* = 1$ and any flow arrangement (except for parallelflow),

$$\left(\frac{1}{\rho}\right)_m = v_m = \frac{v_i + v_o}{2} = \frac{1}{2}\left(\frac{1}{\rho_i} + \frac{1}{\rho_o}\right) \tag{13.29}$$

where v is the specific volume in m^3/kg.

For a perfect gas with $C^* = 0$ and any flow arrangement,

$$\left(\frac{1}{\rho}\right)_m = \frac{\tilde{R}}{p_{avg}} T_{lm} \tag{13.30}$$

Here, $\tilde{R}$ is the gas constant for a particular gas in J/(kg K) (e.g., $\tilde{R} = 287.04$ J/kg K for air), $p_{avg} = (p_i + p_o)/2$, and $T_{\ell m} = T_{const} + \Delta T_{\ell m}$, where T_{const} is the mean temperature of the fluid on the other side of the exchanger; the log-mean temperature difference ΔT_{lm} is defined in Table 13.1. The core frictional pressure drop in Equation 13.28 may be approximated as

$$\Delta p \approx \frac{4 f L G^2}{2 g_c D_h} \left(\frac{1}{\rho}\right)_m \approx \frac{4 f L G^2}{2 g_c \rho D_h} = f \frac{4L}{D_h} \frac{\rho_m u_m^2}{2 g_c} = f \frac{4L}{D_h} \frac{G^2}{2 g_c \rho} \tag{13.31}$$

13.2.7.2 Tube-Fin Heat Exchangers

The pressure drop inside a circular tube is computed using Equation 13.28 with proper values of f factors (see equation in Table 13.6 and Table 13.7) and K_c and K_e from Figure 13.8 for circular tubes.

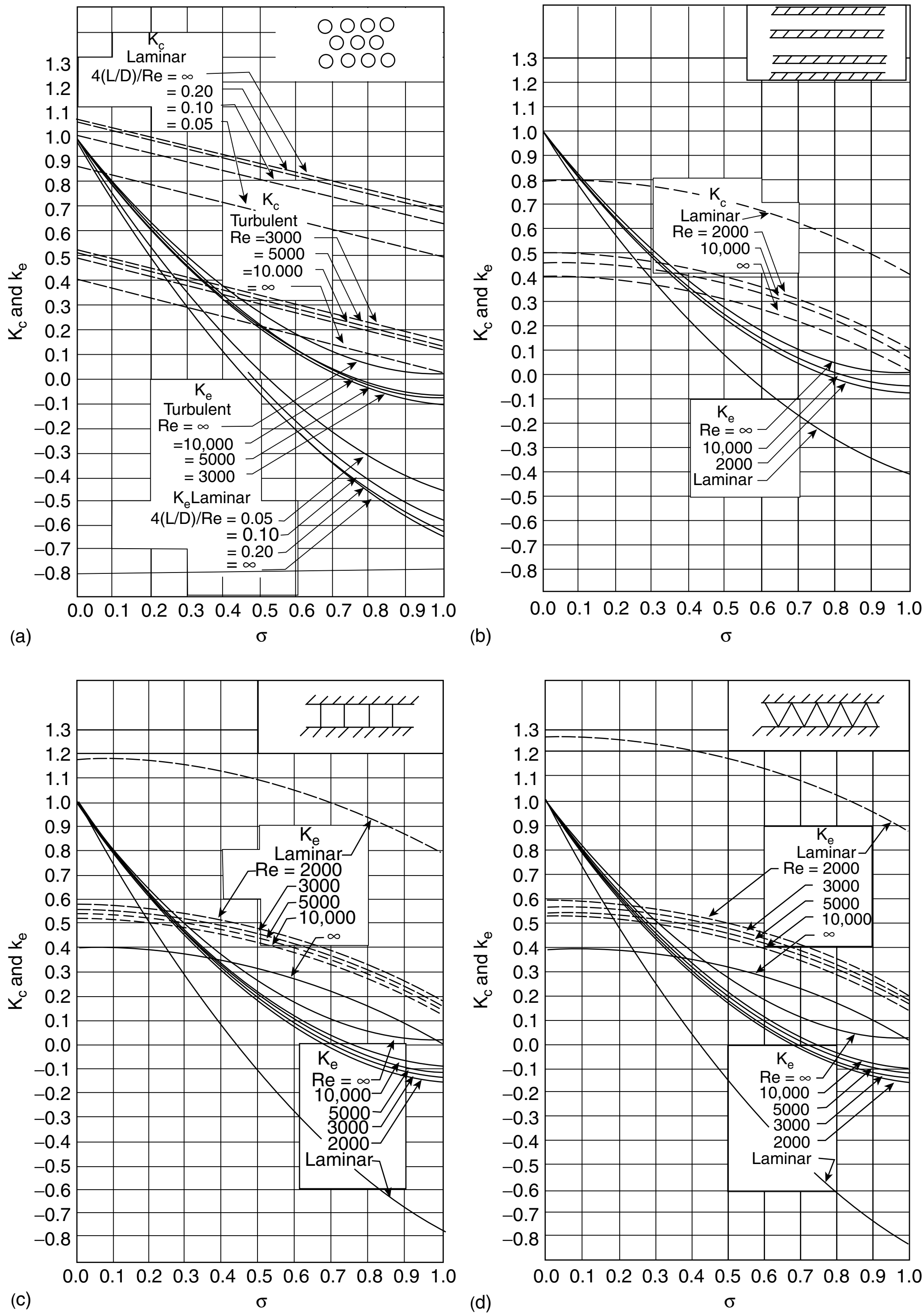


FIGURE 13.8 Entrance and exit pressure loss coefficients: (a) circular tubes, (b) parallel plates, (c) square passages, and (d) triangular passages. For each of these flow passages shown in the inset, the fluid flows perpendicular to the plane of the paper into the flow passages. (From Kays and London 1998. With permission.)

For flat fins on an array of tubes (see Figure 13.6b), the components of the core pressure drop, such as those in Equation 13.28, are the same with the following exception: the core friction and momentum effect take place within the core with $G = \dot{m}/A_0$, where A_0 is the minimum free flow area within the core, and the entrance and exit losses occur at the leading and trailing edges of the core with the associated flow area A'_0 so that

$$\dot{m} = GA_0 = G'A'_0 \text{ or } G'\sigma' = G\sigma \tag{13.32}$$

where $\sigma'$is the ratio of free flow area to frontal area at the fin leading edges. The pressure drop for flow normal to a tube bank with flat fins is then given by

$$\begin{aligned}\frac{\Delta p}{p_i} = {} & \frac{G^2}{2g_c\rho_i p_i}\left[f\frac{L}{r_h}\rho_i\left(\frac{1}{\rho}\right)_m + 2\left(\frac{\rho_i}{\rho_0} - 1\right)\right] \\ & + \frac{G'^2}{2g_c\rho_i p_i}\left[\left(1 - \sigma'^2 + K_c\right) - \left(1 - \sigma'^2 - K_e\right)\frac{\rho_i}{\rho_0}\right]\end{aligned} \tag{13.33}$$

For individually finned tubes as shown in Figure 13.6a, flow expansion and contraction take place along each tube row, and the magnitude is of the same order as that at the entrance and exit. Hence, the entrance and exit losses are generally lumped into the core friction factor. Equation 13.28 for individually (plain and longitudinally) finned tubes then reduces to

$$\frac{\Delta p}{p_i} = \frac{G^2}{2g_c\rho_i\rho_i}\left[f\frac{L}{r_h}\rho_i\left(\frac{1}{\rho}\right)_m + 2\left(\frac{\rho_i}{\rho_0} - 1\right)\right] \tag{13.34}$$

13.2.7.3 Plate Heat Exchangers

Pressure drop in a plate heat exchanger consists of three components: (1) pressure drop associated with the inlet and outlet manifolds and ports, (2) pressure drop within the core (plate passages), and (3) pressure drop due to the elevation change. The pressure drop in the manifolds and ports should be kept as low as possible (generally $<10\%$ of the total pressure drop, but it is found as high as 25%–30% or higher in some designs). Empirically, it is calculated as approximately 1.5 times the inlet velocity head per pass. Since the entrance and exit losses in the core (plate passages) cannot be determined experimentally, they are included in the friction factor for the given plate geometry. The pressure drop (rise) caused by the elevation change for liquids is given by Equation 13.27. Hence, the pressure drop on one fluid side in a plate heat exchanger is given by

$$\Delta p = \frac{1.5\,G_p^2 N_p}{2g_c\rho_i} + \frac{4fLG^2}{2g_c D_e}\left(\frac{1}{\rho}\right)_m + \left(\frac{1}{\rho_0} - \frac{1}{\rho_i}\right)\frac{G^2}{g_c} \pm \frac{\rho_m g L}{g_c} \tag{13.35}$$

where $G_p = \dot{m}/(\pi/4)D_p^2$ is the fluid mass velocity in the port, D_p is the port /manifold diameter, N_p is the number of passes on the given fluid side, and D_e is the equivalent diameter of flow passages (usually twice the plate spacing). Note that the third term on the right-hand side of the equality sign of Equation 13.35 is for the momentum effect which is generally negligible for liquids.

13.3 Regenerator Heat Transfer and Pressure Drop Analysis

A regenerator is a storage type heat exchanger in which both fluids flow alternatively through the same flow passages, and hence heat transfer is from the hot gas to the cold gas via thermal energy storage and release through the exchanger surface or matrix wall. For a continuous operation, either the matrix is moved periodically in and out of the fixed gas streams as in a rotary regenerator or the gas flows are diverted through the valves to and from the fixed matrices as in a fixed-matrix regenerator. Seal leakage in rotary regenerators or valve leakage in fixed-matrix regenerators cannot be avoided and should be kept minimum (less than about 5%) so that its impact on the regenerator performance is minimum. We will focus only on the rotary regenerators because of their use as compact exchangers. Refer to Shah and Sekulić (2003) for further details on rotary and fixed-matrix regenerators.

The heat transfer rate in a regenerator is computed from Equation 13.13 as is for the recuperators.

$$q = \varepsilon C_{\min}(T_{\mathrm{h},i} - T_{\mathrm{c},i}) \tag{13.13b}$$

Two methods have been used for the regenerator design theory: ε-NTU$_0$ and Λ-Π methods. Also, the most common flow arrangement is counterflow, and parallel flow is also possible. Other flow arrangements of recuperators are not possible due to construction features. We will now briefly describe these methods.

13.3.1 ε-NTU$_0$ Method for Heat Transfer

Since heat transfer is through the storage in the matrix wall and is of transient nature, it adds two additional dimensionless groups compared to those for the recuperators for the ε-NTU method for recuperators. They are matrix heat capacity rate ratio C_r^* and thermal conductance ratio $(hA)^* = (hA)$ on $C_{\min}$ side divided by (hA) on $C_{\max}$ side. In most practical regenerator uses, the influence of $(hA)^*$ on the regenerator effectiveness is insignificant and is neglected. Hence, the regenerator effectiveness is given by

$$\varepsilon = \phi(\mathrm{NTU}_0, C^*, C_r^*) \tag{13.36}$$

where

$$\mathrm{NTU}_0 = \frac{1}{C_{\min}}\left[\frac{1}{1/(hA)_{\mathrm{h}} + 1/(hA)_{\mathrm{c}}}\right] = \frac{U_0 A}{C_{\min}} \tag{13.37}$$

$$C^* = \frac{C_{\min}}{C_{\max}} \tag{13.38}$$

$$C_{\mathrm{r}}^* = \frac{C_{\mathrm{r}}}{C_{\min}} \tag{13.39}$$

For a counterflow regenerator, the following approximate procedure is proposed by Razelos (1980) to calculate the regenerator effectiveness ε. For the known values of NTU$_0$, C^* and C_{r}^*, calculate "equivalent" values of NTU$_0$ and C_{r}^* for a balanced regenerator ($C^* = 1$), designated with a subscript m, as follows.

$$\mathrm{NTU}_{0,\mathrm{m}} = \frac{2\mathrm{NTU}_0 C^*}{1 + C^*} \tag{13.40}$$

$$C_{\mathrm{r,m}}^* = \frac{2C_{\mathrm{r}}^* C^*}{1 + C^*} \tag{13.41}$$

With these values of NTU$_{0,\mathrm{m}}$ and $C_{\mathrm{r,m}}^*$, obtain the value of ε_{r} from the following equation.

$$\varepsilon_{\mathrm{r}} = \frac{\mathrm{NTU}_{0,\mathrm{m}}}{1 + \mathrm{NTU}_{0,\mathrm{m}}}\left[1 - \frac{1}{9(C_{\mathrm{r,m}}^*)^{1.93}}\right] \tag{13.42}$$

The regenerator effectiveness ε is then determined from the following equation.

$$\varepsilon = \frac{1 - \exp\{\varepsilon_r (C^{*^2} - 1)/[2C^* (1 - \varepsilon_r)]\}}{1 - C^* \exp\{\varepsilon_r (C^{*^2} - 1)/[2C^* (1 - \varepsilon_r)]\}} \tag{13.43}$$

This formula yields regenerator effectiveness within 1% accuracy for $\varepsilon > 80\%$, and lower effectiveness that accuracy is somewhat lower. The following qualitative observations can be made from the ε-NTU$_0$ results for regenerators:

1. For specified C_r^* and C^*, the heat exchanger effectiveness increases with increasing NTU$_0$. For all C_r^* and C^*, $\varepsilon \to 1$ as NTU$_0 \to \infty$.
2. For specified NTU$_0$ and C^*, ε increases with increasing values of C_r^* and asymptotically approaches the value for a counterflow recuperator.
3. For specified NTU$_0$ and C_r^*, ε increases with decreasing values of C^*. The percentage change in ε is the largest in the lower NTU$_0$ range, and this percentage change in ε increases with increasing values of C_r^*.
4. For $\varepsilon < 40\%$ and $C_r^* > 0.6$, C^* and C_r^* do not have a significant influence on the exchanger effectiveness.

The ε-NTU$_0$ results for a parallel flow regenerator are provided in Shah and Sekulić (2003).

The alternative method for regenerator analysis is referred to as Λ-Π method as summarized in detail by Shah and Sekulić (2003).

Longitudinal wall heat conduction can become important and can *reduce* the regenerator effectiveness significantly if the regenerator is designed over about 85%. In that case, the effect of longitudinal heat conduction in the wall can be taken into account by the following procedure to compute the regenerator effectiveness.

1. Use the Razelos method to compute NTU$_{0,m}$, $C_{r,m}^*$ and $\varepsilon_{r,\lambda=0}$ for an equivalent balanced regenerator for known NTU$_0$ and C^* using Equation 13.40 through Equation 13.42. The ε_r computed from Equation 13.42 is designated as $\varepsilon_{r,\lambda=0}$ here.
2. Compute $\varepsilon_{r,\lambda\neq 0}$ as follows using computed NTU$_{0,m}$ and given λ.

$$\varepsilon_{r,\lambda\neq 0} = \varepsilon_{r,\lambda=0}\left(1 - \frac{C_\lambda}{2 - C^*}\right) \tag{13.44}$$

where

$$C_\lambda = \frac{1}{1 + \frac{\mathrm{NTU}_0(1+\lambda\Phi)}{1+\lambda\mathrm{NTU}_0}} - \frac{1}{1 + \mathrm{NTU}_0} \tag{13.45}$$

and

$$\Phi = \left[\frac{\lambda\,\mathrm{NTU}_0}{1 + \lambda\,\mathrm{NTU}_0}\right]^{1/2} \tanh\left[\frac{\mathrm{NTU}_0}{\{\lambda\,\mathrm{NTU}_0/(1 + \lambda\,\mathrm{NTU}_0)\}^{1/2}}\right] \tag{13.46a}$$

$$\approx \left[\frac{\lambda\,\mathrm{NTU}_0}{1 + \lambda\,\mathrm{NTU}_0}\right]^{1/2} \quad \text{for } \mathrm{NTU}_0 \geq 3 \tag{13.46b}$$

3. Finally determine ε from Equation 13.43 with ε_r replaced by $\varepsilon_{r,\lambda\neq 0}$ calculated from Equation 13.44.

Influence of transverse wall heat conduction may be important for walls of matrix passages either being thick or made up of ceramic materials. In that case, wall thermal resistance term is added to two convective thermal resistances associated with hot and cold fluid flows within the matrix. Also the pressure and carryover leakages in the regenerator can reduce the regenerator effectiveness considerably, and must be taken into account for the desired performance level. Refer to Shah and Sekulić (2003) for the details.

13.3.2 Pressure Drop

For regenerator matrices having cylindrical passages, the pressure drop is computed using Equation 13.28 with appropriate values of f, K_c, and K_e. For regenerator matrices made up of any porous material (such as checkerwork, wire, mesh, spheres, copper wools, etc.), the pressure drop is calculated using Equation 13.34 since the entrance and exit losses are lumped into the friction factor f.

13.4 Heat Transfer and Flow Friction Correlations—Single-Phase Flows

Accurate and reliable surface heat transfer and flow friction characteristics are a key input to the exchanger heat transfer and pressure drop analyses or to the rating and sizing problems (Shah and Sekulić 1998). After presenting the associated nondimensional groups, we will present important analytical solutions and empirical correlations for some important exchanger geometries.

13.4.1 Dimensionless Groups

Heat transfer characteristics of an exchanger surface are presented in terms of the Nusselt number Nu, Stanton number St, or Colburn factor j vs. the Reynolds number Re, dimensionless axial length x^*, or the Graetz number Gz. Flow friction characteristics are presented in terms of the Fanning friction factor f vs. Re or dimensionless axial length x^+. These and other important dimensionless groups used in presenting and correlating internal flow forced convection heat transfer are summarized in Table 13.5 with their definitions and physical meanings. Where applicable, the hydraulic diameter D_h is used as a characteristic dimension in all dimensionless groups. A number of different definitions are used in the literature for some of the dimensionless groups; the user should pay particular attention to the specific definitions used in any research paper before using specific results. This is particularly true for the Nusselt number, where many different temperature differences are used in the definition of h, and for f, Re, and other dimensionless groups having characteristic dimensions different from D_h.

13.4.2 Analytical Solutions

Flow passages in most compact heat exchangers are complex with frequent boundary layer interruptions; some heat exchangers (particularly the tube side of shell-and-tube exchangers and highly compact regenerators) have continuous flow passages. The velocity and temperature profiles across the flow cross section are generally fully-developed in the continuous flow passages, whereas they develop at each boundary layer interruption in an interrupted surface and may reach a periodic fully-developed flow. The heat transfer and flow friction characteristics are generally different for fully-developed flows and developing flows. Analytical results for developed and developing flows for simple flow passage geometries follow. For complex surface geometries, the basic surface characteristics are primarily obtained experimentally (Shah and Sekulić 2003); the pertinent correlations are presented in the next subsection.

Analytical solutions for developed and developing velocity/temperature profiles in constant cross section circular and noncircular flow passages are important when no empirical correlations are available, extrapolations are needed for empirical correlations, or in the development of empirical correlations. Fully-developed-laminar-flow solutions are applicable to highly compact regenerator surfaces or highly compact plate-fin exchangers, with plain uninterrupted fins. Developing laminar flow solutions are applicable to interrupted fin geometries and plain uninterrupted fins of "short" lengths, and turbulent flow solutions to not so compact heat exchanger surfaces.

The heat transfer rate in laminar duct flow is very sensitive to the thermal boundary condition. Hence, it is essential to carefully identify the thermal boundary condition in laminar flow. The heat transfer rate in turbulent duct flow is insensitive to the thermal boundary condition for most common fluids

TABLE 13.5 Important Dimensionless Groups for Internal Flow Forced Convection Heat Transfer and Flow Friction, Useful in Heat Exchanger Design

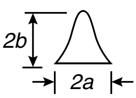
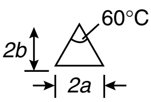
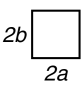
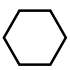
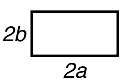
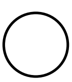
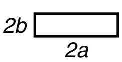
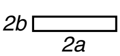
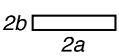
Dimensionless Groups	Definitions and Working Relationships	Physical Meaning and Comments
Reynolds number	$\mathrm{Re} = \dfrac{\rho u_m D_h}{\mu} = \dfrac{G D_h}{\mu}$	a flow modulus, proportional to the ratio of flow momentum rate ("inertia force") to viscous force
Fanning friction factor	$f = \dfrac{\tau_w}{(\rho u_m^2/2g_c)}$ $f = \Delta p^* \dfrac{r_h}{L} \dfrac{\Delta_p}{(\rho u_m^2/2g_c)} \dfrac{r_h}{L}$	the ratio of wall shear (skin frictional) stress to the flow kinetic energy per unit volume; commonly used in heat transfer literature
Apparent fanning friction factor	$f_{app} = \Delta p^* \dfrac{r_h}{L}$	Includes the effects of skin friction and the change in the momentum rates in the entrance region (developing flows)
Incremental pressure drop number	$K(x) = (f_{\mathrm{app}} - f_{fd}) \dfrac{L}{r_h}$	Represents the excess dimensionless pressure drop in the entrance region over that for fully-developed flow
Darcy friction factor	$K(\infty) = \text{constant for } x \to \infty$ $f_D = 4f = \Delta p^* \dfrac{D_h}{L}$	four times the Fanning friction factor; commonly used in fluid mechanics literature
Euler number	$Eu = \Delta p^* = \dfrac{\Delta p}{(\rho u_m^2/2g_c)}$	the pressure drop normalized with respect to the dynamic velocity head; commonly used in the Russian literature
Dimensionless axial distance for the fluid flow problem	$x^+ = \dfrac{x}{D_h \mathrm{Re}}$	the ratio of the dimensionless axial distance (x/D_h) to the Reynolds number; useful in the hydrodynamic entrance region
Nusselt number	$\mathrm{Nu} = \dfrac{h}{k/D_h} = \dfrac{q'' D_h}{k(T_w - T_m)}$	the ratio of the convective conductance h to the pure molecular thermal conductance k/D_h
Stanton number	$St = \dfrac{h}{Gc_p} = \dfrac{\mathrm{Nu}}{\mathrm{Pe}} = \dfrac{\mathrm{Nu}}{\mathrm{RePr}}$	the ratio of convection heat transfer (per unit duct surface area) to amount virtually transferable (per unit of flow cross-sectional area); no dependence upon any geometric characteristic dimension
Colburn factor	$j = \mathrm{St}\,\mathrm{Pr}^{2/3} = (\mathrm{NuPr}^{-1/3})/\mathrm{Re}$	a modified Stanton number to take into account the moderate variations in the Prandtl number for $0.5 \leq \mathrm{Pr} \leq 10.0$ in turbulent flow
Prandtl number	$\mathrm{Pr} = \dfrac{\nu}{\alpha} = \dfrac{\mu c_p}{k}$	a fluid property modulus representing the ratio of momentum diffusivity to thermal diffusivity of the fluid
Péclet number	$\mathrm{Pe} = \dfrac{\rho c_p u_m D_h}{k} = \dfrac{u_m D_h}{\alpha} = \mathrm{RePr}$	proportional to the ratio of thermal energy convected to the fluid to thermal energy conducted axially within the fluid; the inverse of Pe indicates relative importance of fluid axial heat conduction
Dimensionless axial distance for the heat transfer problem	$x^* = \dfrac{x}{D_h \mathrm{Pe}} = \dfrac{x}{D_h \mathrm{RePr}}$	axial coordinate for describing the thermal entrance region heat transfer results
Graetz number	$Gz = \dfrac{\dot{m} c_p}{kL} = \dfrac{\mathrm{Pe}P}{4L} = \dfrac{P}{4D_h} \dfrac{1}{x^*}$ $Gz = \pi/(4x^*)$ for a circular tube	conventionally used in the chemical engineering literature related to x^* as shown when the flow length in Gz is treated as a length variable

Source: From Shah, R. K. and Muller, A. C. Heat exchange. In *Ullmann's Encyclopedia of Industrial Chemistry, Unit Operations II*, *Vol. B3*. VCH Publishers, Weinheim, Germany, Chap. 2, 1988. With permission.

(Pr≥0.7); the exception is liquid metals (Pr<0.03). Hence, there is generally no need to identify the thermal boundary condition in turbulent flow for all fluids except liquid metals.

Fully-developed-laminar-flow analytical solutions for some duct shapes of interest in compact heat exchangers are presented in Table 13.6 for three important thermal boundary conditions denoted by the subscripts H1, H2, and T (Shah and London 1978; Shah and Bhatti 1987). Here, H1 denotes constant axial wall heat flux with constant peripheral wall temperature, H2 denotes constant axial and peripheral wall heat flux, and T denotes constant wall temperature. The following observations may be made from this table: (1) there is a strong influence of flow passage geometry on Nu and *fRe*. Rectangular passages approaching a small aspect ratio exhibit the highest Nu and *fRe*. (2) The thermal boundary conditions have a strong influence on the Nusselt numbers. (3) As $\mathrm{Nu}=hD_h/k$, a constant Nu implies the convective heat transfer coefficient h independent of the flow velocity and fluid type (Prandtl number). (4) An increase in h can be best achieved either by reducing D_h or by selecting a geometry with a low aspect ratio rectangular flow passage. Reducing the hydraulic diameter is an obvious way to increase exchanger compactness and heat transfer, or D_h can be optimized using well-known heat transfer correlations based on design problem specifications. (5) Since fRe=constant, $f \propto 1/Re \propto 1/u_m$. In this case, it can be shown

TABLE 13.6 Solutions for Heat Transfer and Friction for Fully-Developed-Laminar-Flow Through Specified Ducts

Geometry ($L/D_h > 100$)		Nu_{H1}	Nu_{H2}	Nu_T	$f\,Re$	$\frac{j_{H1}}{f}$[a]	$K(\infty)$[b]	L^+_{hy}[c]
	$\frac{2b}{2a}=\frac{\sqrt{3}}{2}$	3.014	1.474	2.39	12.630	0.269	1.739	0.040
	$\frac{2b}{2a}=\frac{\sqrt{3}}{2}$	3.111	1.892	2.47	13.333	0.263	1.818	0.040
	$\frac{2b}{2a}=1$	3.608	3.091	2.976	14.227	0.286	1.433	0.090
		4.002	3.862	3.34	15.054	0.299	1.335	0.086
	$\frac{2b}{2a}=\frac{1}{2}$	4.123	3.017	3.391	15.548	0.299	1.281	0.085
		4.364	4.364	3.657	16.000	0.307	1.25	0.056
	$\frac{2b}{2a}=\frac{1}{4}$	5.331	2.94	4.439	18.233	0.329	1.001	0.078
	$\frac{2b}{2a}=\frac{1}{6}$	6.049	2.93	5.137	19.702	0.346	0.885	0.070
	$\frac{2b}{2a}=\frac{1}{8}$	6.490	2.94	5.597	20.585	0.355	0.825	0.063
	$\frac{2b}{2a}=0$	8.235	8.235	7.541	24.000	0.386	0.674	0.011

[a] $j_{H1}/f=Nu_{H1}\,Pr^{-1/3}/(fRe)$ with $Pr=0.7$. Similarly, values of j_{H2}/f and j_T/f may be computed.

[b] $K(\infty)$ for sine and equilateral triangular channels may be too high (Shah and London 1978); $K(\infty)$ for some rectangular and hexagonal channels is interpolated based on the recommended values in Shah and London (1978).

[c] L^+_{hy} for sine and equilateral triangular channels is too low (Shah and London 1978), so use with caution. L^+_{hy} for rectangular channels is based on the faired curve drawn through the recommended value in Shah and London (1978). L^+_{hy} for a hexagonal channel is an interpolated value.

Source: From Shah and London, 1978. With permission.

that $\Delta p \propto u_\mathrm{m}$. Many additional analytical results for fully-developed-laminar-flow ($Re \le 2000$) are presented by Shah and London (1978); Shah and Bhatti (1987). The entrance effects, flow maldistribution, free convection, property variation, fouling, and surface roughness all affect fully-developed analytical solutions. In order to account for these effects in real plate-fin plain fin geometries having fully-developed flows, it is best to reduce the magnitude of the analytical Nu by at least 10% and to increase the value of the analytical fRe by 10% for design purposes.

The initiation of transition flow (the lower limit of the critical Reynolds number, Re_crit) depends on the type of entrance (e.g., smooth vs. abrupt configuration at the exchanger flow passage entrance). For a sharp square inlet configuration, Re_crit is about 10%–15% lower than that for a rounded inlet configuration. For most exchangers, the entrance configuration would be sharp. Some information on Re_crit is provided by Ghajar and Tam (1994).

Transition flow and fully-developed turbulent flow Fanning friction factors are given by Bhatti and Shah (1987) as

$$f = A + BRe^{-1/m} \tag{13.47}$$

where

$$\begin{aligned}
&A = 0.0054, \quad B = 2.3 \times 10^{-8}, \quad m = -2/3 \quad \text{for } 2100 \le Re \le 4000\\
&A = 0.00128 \quad B = 0.1143 \quad m = 3.2154 \quad \text{for } 4000 \le Re \le 10^7
\end{aligned}$$

Equation 13.47 is accurate within $\pm 2\%$ (Bhatti and Shah 1987). The transition flow and fully-developed turbulent flow Nusselt number correlation for a circular tube is given by Gnielinski as reported in Bhatti and Shah (1987) as

$$\mathrm{Nu} = \frac{(f/2)(Re - 1000)\mathrm{Pr}}{1 + 12.7(f/2)^{1/2}(\mathrm{Pr}^{2/3} - 1)} \tag{13.48}$$

which is accurate within about $\pm 10\%$ with experimental data for $2300 \le Re \le 5 \times 10^6$ and $0.5 \le \mathrm{Pr} \le 2000$. For higher accuracies in turbulent flow, refer to the correlations by Petukhov et al. reported by Bhatti and Shah (1987); Shah and Sekulić (2003). Churchill (1977) (see also Bhatti and Shah 1987) provides a correlation for laminar, transition, and turbulent flow regimes in a circular tube for $2100 < Re < 10^6$ and $0 < \mathrm{Pr} < \infty$.

A careful observation of accurate experimental friction factors for all noncircular smooth ducts reveals that ducts with laminar $fRe < 16$ have turbulent f factors lower than those for the circular tube, whereas ducts with laminar $fRe > 16$ have turbulent f factors higher than those for the circular tube (Shah and Bhatti 1988). Similar trends are observed for the Nusselt numbers. If one is satisfied within $\pm 15\%$ accuracy, Equation 13.47 and Equation 13.48 for f and Nu can be used for noncircular passages with the hydraulic diameter as the characteristic length in f, Nu, and Re; otherwise, refer to Table 13.7 for more accurate results for turbulent flow.

For hydrodynamically and thermally developing flows, the analytical solutions are boundary condition dependent (for laminar flow heat transfer only) and geometry dependent. The reader may refer to Shah and London (1978); Shah and Bhatti (1987), and Bhatti and Shah (1987) for specific solutions. The hydrodynamic entrance lengths L_hy for developing laminar and turbulent flow are given by Shah and Bhatti (1987); Bhatti and Shah (1987) as

$$\frac{L_\mathrm{hy}}{D_\mathrm{h}} = \begin{cases} 0.0565\, Re & \text{for laminar flow } (Re \le 2100) \\ 1.359\, Re^{1/4} & \text{for turbulent flow } (Re > 10^4) \end{cases} \tag{13.49}$$

TABLE 13.7 Fully-Developed Turbulent Flow Friction Factors and Nusselt Numbers ($\mathrm{Pr} > 0.5$) for Technically Important Smooth-Walled Ducts

Duct Geometry and Characteristic Dimension	Recommended Correlations[a]
$D_h = 2a$	Friction factor correlation for $2300 < Re < 10^7$ [41]: Where $A = 0.0054$, $B = 2.3 \times 10^{-8}$, $m = -2/3$ for $2100 < Re < 4000$ and $= 1.28 \times 10^{-3}$. $B = 0.1143$, $m = 3.2154$ for $4000 < Re < 10^7$ [41]. Nusselt number correlation by Gnielinski for $2300 < Re < 5\text{x}10^6$: $Nu = \dfrac{(f/2)(\mathrm{Re} - 1000)\mathrm{Pr}}{1 + 12.7(f/2)^{1/2}(\mathrm{Pr}^{2/3} - 1)}$
$D_h = 4b$	Use circular duct f and Nu correlations. Predicted f are up to 12.5% lower and predicted Nu are within $\pm 9\%$ of the most reliable experimental results.
$D_h = \dfrac{4ab}{a+b},\ \alpha^* \dfrac{2b}{2a}$ $\dfrac{D_l}{D_h} = \dfrac{2}{3} + \dfrac{11}{24}\alpha^*(2 - \alpha^*)$	f factors: (1) substitute D_1, for D_h in the circular duct correlation, and calculate f from the resulting equation. (2) Alternatively, calculate f from $f = (1.0875 - 0.1125\ \alpha^*)\ f_c$, where f_c is the friction factor for the circular duct using D_h. In both cases, predicted f factors are within $\pm 5\%$ of the experimental results. Nusselt numbers: (1) With uniform heating at four walls, use ciruclar duct Nu correlation for an accuracy of $\pm 9\%$ for $0.5 \leq Pr \leq 100$ and $10^4 \leq Re \leq 10^6$. (2) With equal heating at two long walls, use circular duct correlation for an accuracy of $\pm 10\%$ for $0.5 \leq Pr \leq 10$ and $10^4 \leq Re \leq 10^5$. (3) With heating at one long wall only, use circular duct correlation to get approximate Nu values for $0.5 < Pr < 10$ and $10^4 \leq Re \leq 10^6$. These calculated values may be up to 20% higher than the actual experimental values
$D_h = 2\sqrt{3}a = 4b/3$ $D_l = \sqrt{3}a = 2b/3\sqrt{3}$	Use circular duct f and Nu correlations with D_h replaced by D_l, predicted f are within $+3\%$ and (11% and predicted Nu within $+9\%$ of the experimental values.
$D_h = \dfrac{4ab}{a + \sqrt{a^2 + 4b^2}}$ $\dfrac{D_g}{D_h} = \dfrac{1}{2\pi}\left[3\ln\cot\dfrac{\theta}{2} + 2\ln\tan\dfrac{\phi}{2} - \ln\tan\dfrac{\theta}{2}\right]$	Here $\theta = (90° - \phi)/2$. For $0 < 2\phi < 60°$, use circular duct f and Nu correlations with D_h replaced by D_g; for $2\phi = 60°$, replace D_h by D_l (see above); and for $60° < 2\phi \leq 90°$ use circular duct correlations directly with D_h. Predicted f and Nu are within $+9\%$ and 11% of the experimental values. No recommentations can be made for $2\phi > 90°$ due to lack of the experimental data.
$D_h = 2(r_o - r_i)$ $r^* = \dfrac{r_i}{r_o}$ $\dfrac{D_l}{D_h} = \dfrac{1 + r^{*2} + (1 - r^{*2})/\ln r^*}{(1 - r^*)^2}$	f factors: (1) substitute D_l for D_h in the circular duct correlation, and calculate f from the resulting equation. (2) Alternatively, calculate f from $f = (1 + 0.0925\ r^*)\ f_c$, where f_c is the friction factor for the circular duct using D_h. In both cases, predicted f factors are within $\pm 5\%$ of the experimental results. Nusselt numbers: In all the following recommendations, use D_h with a wetted perimeter in Nu and Re: (1) Nu at the outer wall can be determined from the circular duct correlation within the accuracy of about $\pm 10\%$ regardless of the condition at the inner wall. (2) Nu at the inner wall cannot be determined accurately using the circular tube correlation regardless of the heating/cooling condition at the outer wall.

[a] The friction factor and Nusselt number correlations for the circular duct are the most reliable and agree with a large amount of the experimental data within $\pm 2\%$ and $\pm 10\%$, respectively. The correlations for all other duct geometries are not as good as those for the circular duct on an absolute basis.

Source: From Bhatti, M. S. and Shah, R. K., *Handbook of Single-Phase Convective Heat Transfer*, Kakaç, S. Shah, R. K. and Aung, W. ed. Wiley, New York, 1987. With permission.

13.4.3 Experimental Correlations

Analytical results presented in the preceding section are useful for well-defined constant cross-sectional surfaces with essentially unidirectional flows. The flows encountered in heat exchangers are generally very complex, having flow separation, reattachment, recirculation, and vortices. Such flows significantly affect Nu and f for the specific exchanger surfaces. Since no analytical or accurate numerical solutions are available, the information is derived experimentally; Kays and London (1998) and Webb (1994) present many of the experimental results reported in the open literature. Many of the geometries reported in these references have been fine-tuned, refined and made more compact with the advancement of manufacturing technologies. Hence the j (Nu) and f vs. Re data reported in these references have a limited use now since they cannot be used for accurate performance evaluation/prediction of modern surface geometries. In the following, empirical correlations for only some important surfaces are summarized due to space limitations.

13.4.3.1 Bare Tubebanks

Zukauskas (1987) has presented extensive experimental results for flow normal to inline and staggered plain and finned tube bundles in graphical form. Comprehensive pressure drop correlations for flow normal to inline and staggered *plain* tube bundles have been developed by Gaddis and Gnielinski (1985) and recast in terms of the Hagen number by Martin (2002) as follows.

$$\mathrm{Hg} = \begin{cases} \mathrm{Hg}_{\mathrm{lam}} + \mathrm{Hg}_{\mathrm{turb},i}\left[1-\exp\left(1-\dfrac{Re_{\mathrm{d}}+1000}{2000}\right)\right] & \text{inline tube bundles} \\ \mathrm{Hg}_{\mathrm{lam}} + \mathrm{Hg}_{\mathrm{turb,s}}\left[1-\exp\left(1-\dfrac{Re_{\mathrm{d}}+200}{1000}\right)\right] & \text{staggered tube bundles} \end{cases} \tag{13.50}$$

where

$$\mathrm{Hg}_{\mathrm{lam}} = 140Re_{\mathrm{d}}\frac{(X_\ell^{*0.5}-0.6)^2+0.75}{X_{\mathrm{t}}^{*1.6}(4X_{\mathrm{t}}^*X_\ell^*/\pi-1)} \tag{13.51}$$

Equation 13.51 is valid for all inline tube bundles. It is also valid for staggered tube bundles except that the term $X_{\mathrm{t}}^{*1.6}$needs to be changed to $X_{\mathrm{d}}^{*1.6}$ when the minimum free flow area occurs in the diagonal planes of the staggered tube bundle, i.e., $X_\ell^* < 0.5(2X_{\mathrm{t}}^*+1)^{1/2}$.

$$\begin{aligned}\mathrm{Hg}_{\mathrm{turb},i} = {} & \left[\left\{0.11+\frac{0.6(1-0.94/X_\ell^*)^{0.6}}{(X_{\mathrm{t}}^*-0.85)^{1.3}}\right\}\times 10^{0.47(X_\ell^*/X_t^*-1.5)} + 0.015(X_{\mathrm{t}}^*-1)(X_\ell^*-1)\right] \\ & \times Re_{\mathrm{d}}^{2-0.1(X_t^*/X_\ell^*)} + \phi_{\mathrm{t,n}}Re_{\mathrm{d}}^2\end{aligned} \tag{13.52}$$

$$\begin{aligned}\mathrm{Hg}_{\mathrm{turb,s}} = {} & \left[\left\{1.25+\frac{0.6}{(X_{\mathrm{t}}^*-0.85)^{1.08}}\right\}+0.2\left(\frac{X_\ell^*}{X_{\mathrm{t}}^*}-1\right)^3-0.005\left(\frac{X_{\mathrm{t}}^*}{X_\ell^*}-1\right)^3\right]Re_{\mathrm{d}}^{1.75} \\ & + \phi_{\mathrm{t,n}}Re_{\mathrm{d}}^2\end{aligned} \tag{13.53}$$

Equation 13.53 is valid for $Re_d \leq 250{,}000$. For higher Re_d, correct $\mathrm{Hg}_{\mathrm{turb,s}}$ of Equation 13.53 as follows.

$$\mathrm{Hg}_{\mathrm{turb,s,corr}} = \mathrm{Hg}_{\mathrm{turb,s}}\left(1 + \frac{Re_d - 250{,}000}{325{,}000}\right) \quad \text{for } Re_d > 250{,}000 \tag{13.54}$$

$$\phi_{t,n} = \begin{cases} \dfrac{1}{2X_t^{*2}}\left(\dfrac{1}{N_r} - \dfrac{1}{10}\right) & \text{for } 5 \leq N_r \leq 10 \quad \text{and} \quad X_\ell^* \geq 0.5(2X_t^* + 1)^{1/2} \\ 2\left[\dfrac{(X_d^* - 1)}{X_t^*(X_t^* - 1)}\right]^2\left(\dfrac{1}{N_r} - \dfrac{1}{10}\right) & \text{for } 5 \leq N_r \leq 10 \quad \text{and} \quad X_\ell^* < 0.5(2X_t^* + 1)^{1/2} \end{cases} \tag{13.55}$$

and $\phi_{t,n} = 0$ for $N_r > 10$ and N_r is the number of tube rows in the flow direction. It should be emphasized that $\phi_{t,n} Re_d^2$ takes into account the influence of tube bundle inlet and outlet pressure losses while the first bracketed [ ] terms on the right-hand sides of Equation 13.52 and Equation 13.53 take into account the frictional pressure loss in the tube bundle. For the total pressure loss of a tube bundle, the complete Equation 13.52 or Equation 13.53 should be used, i.e., both terms on the right-hand sides of these equations should be included.

The foregoing correlation of Equation 13.50 is valid for $1 < Re_d < 300{,}000$ and $N_r \geq 5$ for both inline and staggered tube bundles: $1.25 \leq X_t^* \leq 3.0$, and $1.2 \leq X_\ell^* \leq 3.0$ for inline tube bundle, and $1.25 \leq X_t^* \leq 3.0$, $0.6 \leq X_\ell^* \leq 3.0$ and, $X_d^* \geq 1.25$ for staggered tube bundles. The experimental data for this correlation had $7.9 \leq d_0 \leq 73$ mm.

The pressure drop for the flow normal to the tube bundle is then computed from

$$\Delta p = \frac{\mu^2}{\rho g_c}\frac{N_r}{d_0^2}\mathrm{Hg} \tag{13.56}$$

Note that only when the minimum flow area occurs in the diagonals for a staggered tube bundle, the term N_r should be replaced by $N_r - 1$ considering the number of flow resistances in the diagonal flow area.

In all correlations for heat transfer and pressure drop for tube bundles, $Re_d = \rho u_m d_0/\mu$, where

$$u_m = \begin{cases} u_\infty \dfrac{X_t^*}{X_t^* - 1} & \text{inline tube bundles} \\ u_\infty \dfrac{X_t^*}{X_t^* - 1} & \text{staggered tube bundles with } X_\ell^* \geq 0.5(2X_t^* + 1)^{1/2} \\ u_\infty \dfrac{X_t^*}{2(X_d^* - 1)} & \text{staggered tube bundles with } X_\ell^* < 0.5(2X_t^* + 1)^{1/2} \end{cases} \tag{13.57}$$

Martin (2002) developed comprehensive correlations for heat transfer with flow normal to inline and staggered plain tube bundles as follows.

$$\mathrm{Nu} = \begin{cases} 0.404\,\mathrm{Lq}^{1/3}\left(\dfrac{Re_d + 1}{Re_d + 1000}\right)^{0.1} & \text{inline tube bundles} \\ 0.404\,\mathrm{Lq}^{1/3} & \text{staggered tube bundles} \end{cases} \tag{13.58}$$

where

$$\mathrm{Lq} = \begin{cases} 1.18\ \mathrm{Hg\ Pr}\left(\dfrac{(4X_t^*/\pi)-1}{X_\ell^*}\right) & \text{inline tube bundles} \\ 0.92\ \mathrm{Hg\ Pr}\left(\dfrac{(4X_t^*/\pi)-1}{X_d^*}\right) & \text{staggered tube bundles with } X_\ell^* \geq 1 \\ 0.92\ \mathrm{Hg\ Pr}\left(\dfrac{(4X_t^* X_\ell^*/\pi)-1}{X_\ell^* X_d^*}\right) & \text{staggered tube bundles with } X_\ell^* < 1 \end{cases} \tag{13.59}$$

where Hg is obtained from Equation 13.50. Note that no explicit definition of the Lévêque number Lq is required here. Refer to Shah and Sekulić (2003) for further details.

The foregoing heat transfer correlation of Equation 13.58 is valid for $1 < Re_d < 2{,}000{,}000$, $0.7 \leq \mathrm{Pr} \leq 700$ (validity expected for $\mathrm{Pr} > 700$, but not for $\mathrm{Pr} < 0.6$) and $2 \leq N_r \leq 15$ for inline tube bundles (90° tube bundles) and $4 \leq N_r \leq 80$ for staggered tube bundles (30, 45, 60° tube bundles and other staggered tube bundles within the correlation tube pitch ranges); the tube pitch ratio ranged $1.02 \leq X_t^* \leq 3.0$ and $0.6 \leq X_\ell^* \leq 3.0$ for both inline and staggered tube bundles. The experimental data for this correlation had $7.9 \leq d_0 \leq 73$ mm. The Nusselt numbers are predicted accurately within $\pm 20\%$ for inline tube bundles and $\pm 14\%$ for staggered tube bundles using Equation 13.58 where the Hagen number Hg of Equation 13.59 is determined from the Gaddis and Gnielinski correlation of Equation 13.50. The Nusselt number prediction may be better if the experimental friction factors are used. Note that when the Gaddis and Gnielinski correlation is extrapolated outside their ranges of Re_d and N_r for Nu calculations, it has predicted Nu within the accuracy mentioned.

13.4.3.2 Plate-Fin Extended Surfaces

Offset Strip Fins. This is one of the most widely used enhanced fin geometries Figure 13.9 in aircraft, cryogenics, and many other industries that do not require automotive type mass production. This surface has one of the highest heat transfer performance relative to the pressure drop. Extensive analytical, numerical, and experimental investigations have been conducted over the last 50 years. The most comprehensive correlations for *j* and *f* factors for the *offset* strip fin geometry are provided by Manglik and Bergles (1995) as follows:

$$j = 0.6522\ \mathrm{Re}^{-0.5403}\left(\frac{s}{h'}\right)^{-0.1541}\left(\frac{\delta_f}{\ell_f}\right)^{0.1499}\left(\frac{\delta_f}{s}\right)^{-0.0678} \left[1 + 5.269 \times 10^{-5}\ \mathrm{Re}^{1.340}\left(\frac{s}{h'}\right)^{0.504}\left(\frac{\delta_f}{\ell_f}\right)^{0.456}\left(\frac{\delta_f}{s}\right)^{-1.055}\right]^{0.1} \tag{13.60}$$

$$f = 9.6243\ \mathrm{Re}^{-0.7422}\left(\frac{s}{h'}\right)^{-0.1856}\left(\frac{\delta_f}{\ell_f}\right)^{0.3053}\left(\frac{\delta_f}{s}\right)^{-0.2659} \left[1 + 7.669 \times 10^{-8}\ \mathrm{Re}^{4.429}\left(\frac{s}{h'}\right)^{0.920}\left(\frac{\delta_f}{\ell_f}\right)^{3.767}\left(\frac{\delta_f}{s}\right)^{0.236}\right]^{0.1} \tag{13.61}$$

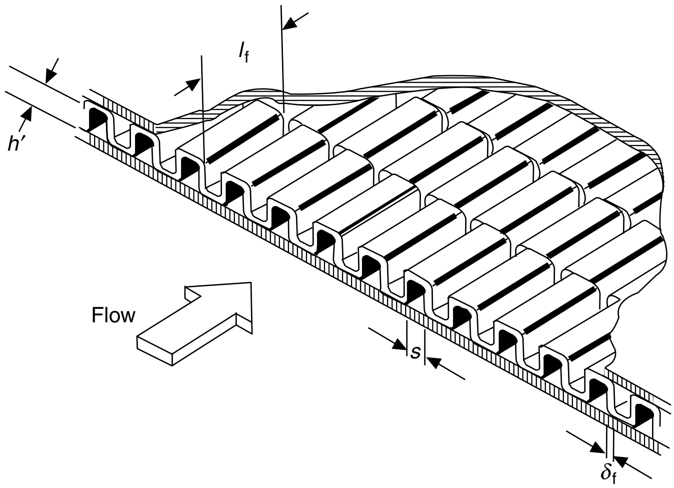


FIGURE 13.9 An offset strip fin geometry.

where

$$D_h = \frac{4A_{0,\text{cell}}}{A_{\text{cell}}/\ell_f} = \frac{4sh'\ell_f}{2(s\ell_f + h'\ell_f + h'\delta_f) + s\delta_f} \tag{13.62}$$

Geometrical symbols in Equation 13.62 are shown in Figure 13.9.

These correlations predict the experimental data of 18 test cores within $\pm 20\%$ for $120 \leq Re \leq 10^4$. Although all the experimental data for these correlations are obtained for air, the j factor takes into consideration minor variations in the Prandtl number, and the correlations should be valid for $0.5 < Pr < 15$.

Louver Fins. Louver or multilouver fins are extensively used in the auto industry due to their mass production manufacturability and hence lower cost. It has generally higher j and f factors than those for the offset strip fin geometry, and the increase in the friction factors is in general higher than the increase in the j factors. However, the exchanger can be designed for higher heat transfer and the same pressure drop compared to that with the offset strip fins by a proper selection of exchanger frontal area, core depth, and fin density. Published literature and correlations on the louver fins are summarized by Webb (1994); Cowell, Heikal, and Achaichia (1995), and the understanding of flow and heat transfer phenomena is summarized by Cowell, Heikal, and Achaichia (1995). The correlation for the Colburn factors for the corrugated louver fins (see Figure 13.10), based on an extensive database for *airflow* over louver fins, is obtained by Chang and Wang (1997) and Wang (2000) as follows.

$$j = Re_{\ell_p}^{-0.49}\left(\frac{\theta}{90}\right)^{0.27}\left(\frac{p_f}{\ell_p}\right)^{-0.14}\left(\frac{b}{\ell_p}\right)^{-0.29}\left(\frac{W_t}{\ell_p}\right)^{-0.23}\left(\frac{\ell_\ell}{\ell_p}\right)^{0.68}\left(\frac{p_t}{\ell_p}\right)^{-0.28}\left(\frac{\delta}{\ell_p}\right)^{-0.05} \tag{13.63}$$

where $Re_{\ell p} = G\ell_p/\mu$ represents the Reynolds number based on the louver pitch ℓ_p. Also, θ is the louver angle, deg; p_f is the fin pitch, mm; b is the vertical fin height, mm; W_t is the tube outside width (= total fin length in the airflow direction if no overhangs), mm; ℓ_ℓ is the louver cut length, mm; p_t is the tube pitch, mm; δ is the fin thickness, mm. These symbols are shown in Figure 13.10.

Equation 13.63 and Equation 13.65 are valid for the following ranges of the parameters: $0.82 \leq D_h \leq 5.02$ mm, $0.51 \leq p_f \leq 3.33$ mm, $0.5 \leq \ell_p \leq 3$, $2.84 \leq b \leq 20$ mm, $15.6 \leq W_t \leq 57.4$ mm, $2.13 \leq \ell_\ell \leq 18.5$ mm, $7.51 \leq p_t \leq 25$ mm, $0.0254 \leq \delta \leq 0.16$ mm, $1 \leq N \leq 2$, and $8.4 \leq \theta \leq 35°$. This

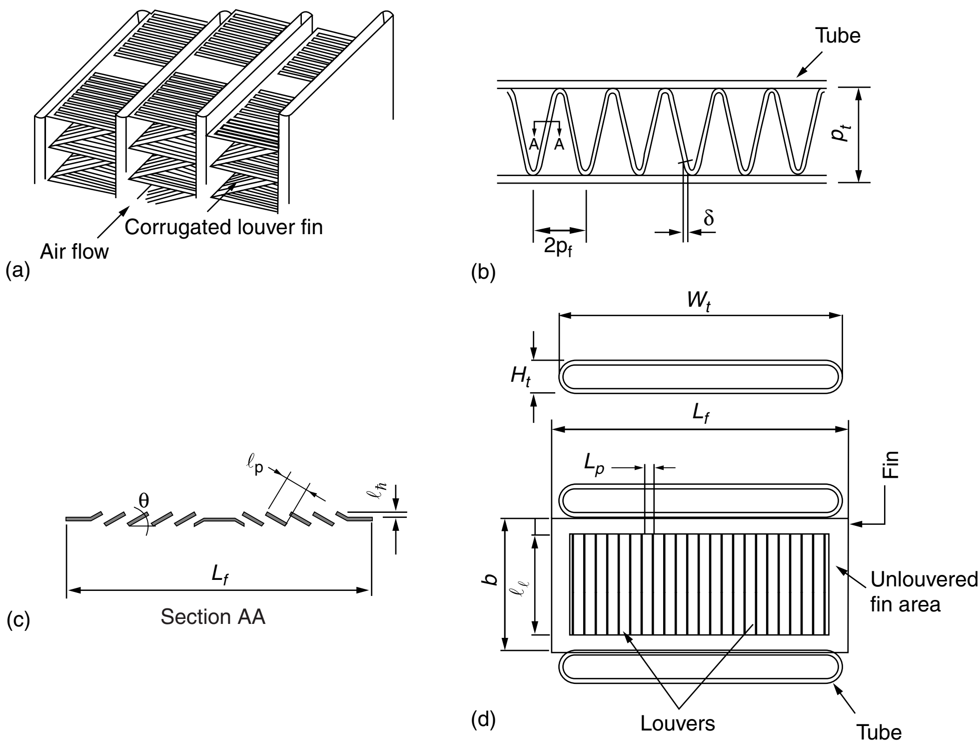


FIGURE 13.10 Definition of geometrical parameters of corrugated louver fins. (From Chang and Wang 1997. With permission.)

correlation predicts 89% of experimental j factors for 91 test cores within $\pm 15\%$ for $30 < Re_{\ell p} < 5000$ with a mean deviation of 8%.

Chang and Wang (1997) also presented a simplified correlation for Equation 13.63 as

$$j = 0.425\, Re_{\ell_p}^{-0.496} \tag{13.64}$$

They report that this correlation predicts 88% of data points within $\pm 25\%$ with the mean deviation of 13%.

The correlation for the Fanning friction factor based on the same database by Chang et al. (2000) as follows.

$$f = f_1 f_2 f_3 \tag{13.65}$$

where

$$f_1 = \begin{cases} 14.39\mathrm{Re}_{\ell_p}^{(-0.805 p_f/b)} \{\ln[1.0 + (p_f/\ell_p)]\}^{3.04} & \mathrm{Re}_{\ell_p} < 150 \\ 4.97\mathrm{Re}_{\ell_p}^{(0.6049 - 1.064/\theta^{0.2})} \left\{\ln\left[(\delta/p_f)^{0.5} + 0.9\right]\right\}^{-0.527} & 150 < \mathrm{Re}_{\ell_p} < 5000 \end{cases} \tag{13.66}$$

$$f_2 = \begin{cases} \left\{\ln\left[(\delta/p_f)^{0.48} + 0.9\right]\right\}^{-1.435} (D_h/\ell_p)^{-3.01} \left[\ln\left(0.5\ \mathrm{Re}_{\ell_p}\right)\right]^{-3.01} & \mathrm{Re}_{\ell_p} < 150 \\ \left[(D_h/\ell_p)\ln(0.3\ Re_{\ell_p})\right]^{-2.966} (p_f/\ell_\ell)^{-0.7931(p_t/b)} & 150 < \mathrm{Re}_{\ell_p} < 5000 \end{cases} \quad (13.67)$$

$$f_3 = \begin{cases} (p_f/\ell_\ell)^{-0.308} (L_f/\ell_\ell)^{-0.308} \left(e^{-0.1167 p_t/H_t}\right) \theta^{0.35} & \mathrm{Re}_{\ell p} < 150 \\ (p_t/H_t)^{-0.0446} \left\{\ln\left[1.2 + (\ell_p/p_f)^{1.4}\right]\right\}^{-3.553} \theta^{-0.477} & 150 < \mathrm{Re}_{\ell_p} < 5000 \end{cases} \quad (13.68)$$

Additional parameters for the friction factor correlations are: D_h is the hydraulic diameter of the fin geometry, mm; H_t is the tube outside height, mm; and L_f is the fin length in the airflow direction, mm. Note that θ in Equation 13.66 and Equation 13.68 are in degrees. The conventional definition of the hydraulic diameter ($D_h = 4A_0/\tilde{P}$) is used in Equation 13.67 considering as if the louver fins were plain fins (without cuts) for the calculation of A_0 and A which is an excellent approximation considering the path of heat flow ideally not obstructed by the louvers; the effect of the braze fillets has also been neglected since no such information is available in the open literature. The correlation of Equation 13.65 predicts 83% of the experimental friction factor data points within ±15% with a mean deviation of 9% for the parameter ranges the same as those for Equation 13.63.

Note that the authors included the tube width W_t in the heat transfer correlation (to take into account the correct surface area). However, they used the fin length L_f for the friction factor correlation since the friction factor f of Equation 13.65 is for the fin friction component only (excluding entrance and exit pressure losses from the measured pressure drops).

13.4.3.3 Tube-Fin Extended Surfaces

Two major types of tube-fin extended surfaces are (1) individually finned tubes and (2) flat fins (also sometimes referred to as plate fins) with or without enhancements/interruptions on an array of tubes as shown in Figure 13.6. Extensive coverage of the published literature and correlations for these extended surfaces are provided by Webb (1994), Kays and London (1998), and Rozenman (1976). Empirical correlations for some important geometries are summarized next.

Individually Finned Tubes. This fin geometry, helically wrapped (or extruded) circular fins on a circular tube as shown in Figure 13.6a, is commonly used in process and waste heat recovery industries. The following correlation for j factors is recommended by Briggs and Young (see Webb 1994) for individually finned tubes on staggered tubebanks.

$$j = 0.134\ \mathrm{Re}_d^{-0.319} \left(\frac{s}{\ell_f}\right)^{0.2} \left(\frac{s}{\delta_f}\right)^{0.11} \quad (13.69)$$

where ℓ_f [$=(d_e - d_0)/2$] is the radial height of the fin, δ_f is the fin thickness, $s = p_f - \delta_f$ is the distance between adjacent fins, and p_f is the fin pitch. Equation 13.69 is valid for the following ranges: $1100 \le \mathrm{Re}_d \le 18{,}000$, $0.13 \le s/\ell_f \le 0.63$, $1.01 \le s/\delta_f \le 7.62$, $0.09 \le \ell_f/d_0 \le 0.69$, $0.011 \le \delta_f/d_0 \le 0.15$, $1.54 \le X_t/d_0 \le 8.23$, fin root diameter d_0 between 11.1 and 40.9 mm, and fin density N_f ($=1/p_f$) between 246 and 768 fins per meter. All data have been obtained on equilateral triangular tube bundle (Shah and Sekulić 2003). The standard deviation of Equation 13.69 with experimental results is 5.1%.

For friction factors, Robinson and Briggs (see Webb 1994) recommended the following correlation.

$$f_{tb} = 9.465\ \mathrm{Re}_d^{-0.316} \left(\frac{X_t}{d_0}\right)^{-0.927} \left(\frac{X_t}{X_d}\right)^{0.515} \quad (13.70)$$

Here, $X_d = \left(X_t^2 + X_\ell^2\right)^{1/2}$ is the diagonal pitch and X_t and X_λ are the transverse and longitudinal tube pitches, respectively. The correlation is valid for the following ranges: $2000 \leq \mathrm{Re_d} \leq 50{,}000$, $0.15 \leq s/\ell_f \leq 0.19$, $3.75 \leq s/\delta_f \leq 6.03$, $0.35 \leq \ell_f/d_0 \leq 0.56$, $0.011 \leq \delta_f\ /d_0 \leq 0.025$, $1.86 \leq X_t/d_0 \leq 4.60$, $18.6 \leq d_0 \leq 40.9$ mm, and $311 \leq N_f \leq 431$ fins per meter. The standard deviation of Equation 13.44 with correlated data is 7.8%.

For crossflow over low-height finned tubes, Rabas and Taborek (1987); Ganguli and Yilmaz (1987), and Chai (1988) have assessed the pertinent literature. A simple but accurate correlation for heat transfer is given by Ganguli and Yilmaz (1987) as

$$j = 0.255\ \mathrm{Re_d^{-0.3}} \left(\frac{d_e}{s}\right)^{-0.3} \tag{13.71}$$

A more accurate correlation for heat transfer is given by Rabas and Taborek (1987). Chai (1988) provides the best correlation for friction factors:

$$f_{tb} = 1.748\ \mathrm{Re_d^{-0.233}} \left(\frac{\ell_f}{s}\right)^{0.552} \left(\frac{d_0}{x_t}\right)^{0.599} \left(\frac{d_0}{x_t}\right)^{0.1738} \tag{13.72}$$

This correlation is valid for $895 < \mathrm{Re_d} < 713{,}000$, $20 < \theta < 40°$, $X_t/d_0 < 4$, and $N \geq 4$, and θ is the tube layout angle. It predicts 89 literature data points within a mean absolute error of 6%; the range of actual error is from -16.7 to 19.9%. An alternative correlation for friction factor is given by Ganguli and Yilmaz (1987), and is summarized by Shah and Sekulić (2003).

Plain Flat Fins on a Staggered Tubebank. This geometry, as shown in Figure 13.6b, is used in the air-conditioning/refrigeration industry as well as where the pressure drop on the fin side prohibits the use of enhanced flat fins. An inline tubebank is generally not used unless fin side very low pressure drop is the essential requirement. Heat transfer correlation for Figure 13.6b flat plain fins on staggered tubebanks is provided by Wang and Chi (2000) and summarized by Wang (2000) as follows.

$$j = \begin{cases} 0.108\ \mathrm{Re_{dc}^{-0.29}} \left(\dfrac{X_t}{X_\ell}\right)^{c_1} \left(\dfrac{p_f}{d_c}\right)^{-1.084} \left(\dfrac{p_f}{D_h}\right)^{-0.786} \left(\dfrac{p_f}{X_t}\right)^{c_2} & \text{for } N_r = 1 \\ 0.086\ Re_{dc}^{c_3} N_r^{c_4} \left(\dfrac{p_f}{d_c}\right)^{c_5} \left(\dfrac{p_f}{D_h}\right)^{c_6} \left(\dfrac{p_f}{X_t}\right)^{-0.93} & \text{for } N_r \geq 2 \end{cases} \tag{13.73}$$

where

$$c_1 = 1.9 - 0.23 \ln(\mathrm{Re_{dc}}) \qquad c_2 = -0.236 + 0.126 \ln(\mathrm{Re_{dc}}) \tag{13.74a}$$

$$c_3 = -0.361 - \frac{0.042\ N_r}{\ln(\mathrm{Re_{dc}})} + 0.158 \ln\left[N_r \left(\frac{p_f}{d_c}\right)^{0.41}\right] \qquad c_4 = -1.224 - \frac{0.076(X_\ell/D_h)^{1.42}}{\ln(\mathrm{Re_{dc}})} \tag{13.74b}$$

$$c_5 = -0.083 + \frac{0.058\ N_r}{\ln(\mathrm{Re_{dc}})} \qquad c_6 = -5.735 + 1.21 \ln\left(\frac{\mathrm{Re_{dc}}}{N_r}\right) \tag{13.74c}$$

where p_f is the fin pitch, d_c is the collar diameter of the fin, and $\mathrm{Re_{dc}} = \rho u_m d_c/\mu$. This j factor correlation predicts 89% of the test points of 74 cores within $\pm 15\%$ with a mean deviation of 8%. Wang and Chi (2000) also provided the following correlation for the friction factors.

$$f = 0.0267\ \mathrm{Re}_{\mathrm{dc}}^{c_7}\left(\frac{X_{\mathrm{t}}}{X_{\ell}}\right)^{c_8}\left(\frac{p_{\mathrm{f}}}{d_{\mathrm{c}}}\right)^{c_9} \tag{13.75}$$

where

$$c_7 = -0.764 + 0.739\left(\frac{X_{\mathrm{t}}}{X_{\ell}}\right) + 0.177\left(\frac{p_{\mathrm{f}}}{d_{\mathrm{c}}}\right) - \frac{0.00758}{N_{\mathrm{r}}} \tag{13.76a}$$

$$c_8 = -15.689 + \frac{64.021}{\ln(\mathrm{Re}_{\mathrm{dc}})} \qquad c_9 = 1.696 - \frac{15.695}{\ln(\mathrm{Re}_{\mathrm{dc}})} \tag{13.76b}$$

Equation 13.73 and Equation 13.75 are valid for the following ranges of the parameters: $300 \leq \mathrm{Re}_{\mathrm{dc}} \leq 20{,}000$, $6.9 \leq \mathrm{d}_{\mathrm{c}} \leq 13.6$ mm, $1.30 \leq D_{\mathrm{h}} \leq 9.37$ mm, $20.4 \leq X_{\mathrm{t}} \leq 31.8$, $12.7 \leq X_{\ell} \leq 32$ mm, $1.0 \leq p_{\mathrm{f}} \leq 8.7$ mm, and $1 \leq N_{\mathrm{r}} \leq 6$. This friction factor correlation of Equation 13.75 predicts 85% of experimental friction factors of 74 test cores within $\pm 15\%$ with a mean deviation of 8%.

Corrugated Flat Fins on a Tube Array. There are a number of variations available for flat fins with a sharp vs. smooth wave. The specific flat-fin geometry shown in Figure 13.11 is designated as a corrugated (herringbone or sharp-wave) fin. The heat transfer and flow friction correlations are developed by Wang (2000) and presented separately for large and small diameter tubes as follows. For larger tube diameters ($d_0 = 12.7$ and 15.88 mm, before tube expansion), the following are the correlations.

$$j = 1.7910\mathrm{Re}_{\mathrm{dc}}^{c_1}\left(\frac{X_{\ell}}{\delta}\right)^{-0.456} N_{\mathrm{r}}^{-0.27}\left(\frac{p_{\mathrm{f}}}{d_{\mathrm{c}}}\right)^{-1.343}\left(\frac{p_{\mathrm{d}}}{x_{\mathrm{f}}}\right)^{0.317} \tag{13.77}$$

$$f = 0.05273\ \mathrm{Re}_{\mathrm{dc}}^{c_2}\left(\frac{p_{\mathrm{d}}}{x_{\mathrm{f}}}\right)^{c_3}\left(\frac{p_{\mathrm{f}}}{X_{\mathrm{t}}}\right)^{c_4}\left[\ln\left(\frac{A}{A_{\mathrm{p,t}}}\right)\right]^{-2.726}\left(\frac{D_{\mathrm{h}}}{d_{\mathrm{c}}}\right)^{0.1325} N_{\mathrm{r}}^{0.02305} \tag{13.78}$$

where

$$c_1 = -0.1707 - 1.374\left(\frac{X_{\ell}}{\delta}\right)^{-0.493}\left(\frac{p_{\mathrm{f}}}{d_{\mathrm{c}}}\right)^{-0.886} N_{\mathrm{r}}^{-0.143}\left(\frac{p_{\mathrm{d}}}{x_{\mathrm{f}}}\right)^{-0.0296} \tag{13.79a}$$

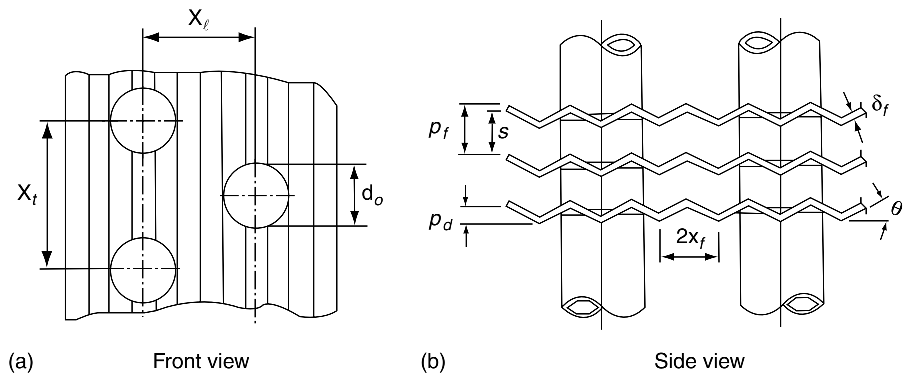


FIGURE 13.11 Corrugated fins on a staggered tube array.

$$c_2 = 0.1714 - 0.07372\left(\frac{p_f}{X_\ell}\right)^{0.25}\left(\ln\frac{A}{A_{p,t}}\right)\left(\frac{p_d}{x_f}\right)^{-0.2} \tag{13.79b}$$

$$c_3 = 0.426\left(\frac{p_f}{X_t}\right)^{0.3}\ln\left(\frac{A}{A_{p,t}}\right) \qquad c_4 = -\frac{10.2192}{\ln(\mathrm{Re}_{dc})} \tag{13.79c}$$

Here, x_f is the projected fin pattern length for one-half wave length, p_d is the fin pattern depth, peak-to-valley distance excluding fin thickness (as shown in Figure 13.11), and $A_{p,t}$ is the tube outside surface area when there are no fins. Equation 13.77 and Equation 13.78 are valid for the following ranges of the parameters: $500 \le \mathrm{Re}_{dc} \le 10{,}000$, $3.63 \le D_h \le 7.23$ mm, $13.6 \le d_c \le 16.85$ mm, $31.75 \le X_t \le 38.1$, $27.5 \le X_\ell \le 33$ mm, $2.98 \le p_f \le 6.43$ mm, $1 \le N_r \le 6$, $12.3 \le \theta \le 14.7°$, $6.87 \le x_f \le 8.25$ mm, and $p_d =$ 1.8 mm. The correlation of Equation 13.77 predicts 93% of experimental Colburn factors for 18 test cores within $\pm 10\%$ with a mean deviation of 4%. Similarly, the correlation of Equation 13.78 predicts 92% of experimental friction factors for 18 test cores within $\pm 10\%$ with a mean deviation of 5%.

For smaller diameter tubes ($d_0 = 7.94$ and 9.53 mm before tube expansion), the following are the correlations for j and f factors.

$$j = 0.324\,\mathrm{Re}_{dc}^{c_1}\left(\frac{p_f}{X_\ell}\right)^{c_2}(\tan\theta)^{c_3}\left(\frac{X_\ell}{X_t}\right)^{c_4}N_r^{0.428} \tag{13.80}$$

$$f = 0.01915\,\mathrm{Re}_{dc}^{c_5}(\tan\theta)^{c_6}\left(\frac{p_f}{X_\ell}\right)^{c_7}\left[\ln\left(\frac{A}{A_{p,t}}\right)\right]^{-5.35}\left(\frac{D_h}{d_c}\right)^{1.3796}N_r^{-0.0916} \tag{13.81}$$

where

$$c_1 = -0.229 + 0.115\left(\frac{p_f}{D_c}\right)^{0.6}\left(\frac{X_\ell}{D_h}\right)^{0.54}N_r^{-0.284}\ln(0.5\tan\theta) \tag{13.82a}$$

$$c_2 = -0.251 + \frac{0.232N_r^{1.37}}{\ln(\mathrm{Re}_{dc}) - 2.303} \qquad c_3 = -0.439\left(\frac{p_f}{D_h}\right)^{0.09}\left(\frac{X_\ell}{X_t}\right)^{-1.75}N_r^{-0.93} \tag{13.82b}$$

$$c_4 = 0.502\left[\ln(\mathrm{Re}_{dc}) - 2.54\right] \qquad c_5 = 0.4604 - 0.01336\left(\frac{p_f}{X_\ell}\right)^{0.58}\left(\ln\frac{A}{A_{p,t}}\right)(\tan\theta)^{-1.5} \tag{13.82c}$$

$$c_6 = 3.247\left(\frac{p_f}{X_t}\right)^{1.4}\ln\left(\frac{A}{A_{p,t}}\right) \qquad c_7 = -\frac{20.113}{\ln(\mathrm{Re}_{dc})} \tag{13.82d}$$

Equation 13.80 and Equation 13.81 are valid for the following ranges of the parameters: $300 \le \mathrm{Re}_{dc} \le 8000$, $1.53 \le D_h \le 4.52$ mm, $8.58 \le d_0 \le 10.38$ mm, $X_t = 25.4$ mm, $19.05 \le X_\ell \le 25.04$ mm, $1.21 \le p_f \le 3.66$ mm, $1 \le N_r \le 6$, $14.5 \le \theta \le 18.5°$, $4.76 \le x_f \le 6.35$ mm, and $1.18 \le p_d \le 1.68$ mm. The correlation of Equation 13.80 predicts 95% of experimental Colburn factors for 27 test cores within $\pm 15\%$ with a mean deviation of 6%. Similarly, the correlation of Equation 13.81 predicts 97% of experimental friction factors for data points of 27 test cores within $\pm 15\%$ with a mean deviation of 5%.

13.4.3.4 Regenerator Surfaces

Two most common types of regenerator surfaces are: (1) continuous cylindrical passages for rotary regenerators and some compact fixed-matrix regenerators and (2) randomly packed woven screens, crossed rods, and packed beds using a variety of materials.

For compact regenerators, the continuous cylindrical flow passages have simple geometries, such as triangular, rectangular, and hexagonal passages. The Nu and f factors of Table 13.6 are a valuable baseline for such passages. Due to the differences in the ideal geometries and boundary conditions compared to the actual ones, London, Young, and Stang (1970) presented the following correlations for air flow through triangular passages ($40 < \mathrm{Re} < 800$).

$$f = \frac{14.0}{\mathrm{Re}} \qquad j = \frac{3.0}{\mathrm{Re}} \tag{13.83}$$

London and Shah (1973) presented the following correlations for air flow through hexagonal passages ($80 < \mathrm{Re} < 800$).

$$f = \frac{17.0}{\mathrm{Re}} \qquad j = \frac{4.0}{\mathrm{Re}} \tag{13.84}$$

Refer to Shah and Sekulić (2003) for the correlations for crossed rod geometries.

13.5 Heat Transfer and Pressure Drop Correlations—Two-Phase Flows

13.5.1 Two-Phase Pressure Drop Correlations

In most compact heat exchanger applications, the pressure drop per unit length or pressure gradient along the flow length is practically constant in single-phase flows. Hence, we generally work directly with the pressure drop in single-phase flows. However, due to the phase change during condensation or vaporization, the pressure gradient within the fluid changes considerably along the flow path or axial length. The pressure drop in the phase-change fluid can then be computed by integrating the nonlinear pressure gradient along the flow path. Hence, the pressure drop computation is somewhat more complicated in two-phase flows.

The total local pressure gradient in two-phase flow through a constant cross-sectional duct can be calculated, based on a homogeneous model, as follows:

$$\frac{\mathrm{d}p}{\mathrm{d}z} = \frac{\mathrm{d}p_{\mathrm{fr}}}{\mathrm{d}z} + \frac{\mathrm{d}p_{\mathrm{mo}}}{\mathrm{d}z} + \frac{\mathrm{d}p_{\mathrm{gr}}}{\mathrm{d}z} \tag{13.85}$$

where the three terms on the right-hand side correspond to the contributions by friction, momentum rate change, and gravity denoted by the subscripts fr, mo, and gr, respectively. The entrance and exit pressure loss terms of single-phase flow (see Equation 13.28) are lumped into the Δp_{fr} term since the information about these contributions is not available, due to the difficulty in measurements and small contribution. The in-tube two-phase frictional pressure drop is computed from the corresponding pressure drop for single-phase flow as follows using the two-phase friction multiplier denoted as ϕ^2:

$$\left(\frac{\mathrm{d}p}{\mathrm{d}z}\right)_{\mathrm{fr}} = \frac{4 f_{\mathrm{lo}} G^2}{2 g_{\mathrm{c}} \rho_{\mathrm{l}} D_{\mathrm{h}}} \varphi_{\mathrm{lo}}^2 \quad \text{where} \quad \varphi_{\mathrm{lo}}^2 = \frac{(\mathrm{d}p/\mathrm{d}z)_{\mathrm{fr}}}{(\mathrm{d}p/\mathrm{d}z)_{\mathrm{fr,lo}}} \tag{13.86}$$

where f_{lo} is the single-phase Fanning friction factor (see Table 13.6 and Table 13.7) based on the total mass flow rate as liquid and G is also based on the total mass flow rate as liquid; this means that the subscript "lo" indicates the two-phase flow considered as all liquid flow.

Alternatively, $(dp/dz)_{fr}$ is determined using the liquid or vapor-phase pressure drop multiplier as follows:

$$\left(\frac{dp}{dz}\right)_{fr} = \left(\frac{dp}{dz}\right)_{fr,l} \varphi_l^2 = \left(\frac{dp}{dz}\right)_{fr,g} \varphi_g^2 \tag{13.87}$$

where

$$\varphi_l^2 = \frac{(dp/dz)_{fr}}{(dp/dz)_{fr,l}} \qquad \varphi_g^2 = \frac{(dp/dz)_{fr}}{(dp/dz)_{fr,g}} \qquad \left(\frac{dp}{dz}\right)_{fr,l} = \frac{4f_l G^2}{2g_c \rho_l D_h} \qquad \left(\frac{dp}{dz}\right)_{fr,g} = \frac{4f_g G^2}{2g_c \rho_g D_h} \tag{13.88}$$

where the subscripts l and g denote liquid and gas/vapor phases, respectively. φ_l^2 and φ_g^2 are functions of the parameter X (*Martinelli parameter*). φ_{go}^2 (defined similar to φ_{lo}^2 of Equation 13.86, with the subscript lo replaced by go) is a function of Y (*Chisholm parameter*). The X and Y are defined as follows:

$$X^2 = \frac{(dp/dz)_{fr,l}}{(dp/dz)_{fr,g}} \qquad Y^2 = \frac{(dp/dz)_{fr,go}}{(dp/dz)_{fr,lo}} \tag{13.89}$$

Here, the subscript go means the total two-phase flow considered as all gas flow. The correlations to determine the two-phase frictional pressure gradient are presented in Table 13.8 for various ranges of G and μ_l/μ_g (Kandlikar, Shoji, and Dhir 1999, p. 228).

TABLE 13.8 Frictional Multiplier Correlations Used for Determining the Two-Phase Frictional Pressure Gradient in Equation 13.86

Correlation	Parameters
Friedel correlation (1979) for $\mu_l/\mu_g > 1000$ and all values of G:	$E = (1-x)^2 + x^2 \frac{\rho_l}{\rho_g}\frac{f_{go}}{f_{lo}}$
$\varphi_{go}^2 = E + \frac{3.24FH}{Fr^{0.045}We^{0.035}}$	$F = x^{0.78}(1-x)^{0.24}$
Accuracy for annular flow: ±21% Ould Didi et al. (2002)	$H = \left(\frac{\rho_l}{\rho_g}\right)^{0.91}\left(\frac{\mu_g}{\mu_l}\right)^{0.19}\left(1-\frac{\mu_g}{\mu_l}\right)^{0.7}$
	$Fr = \frac{G^2}{gd_i\rho_{hom}^2}$ $\quad We = \frac{G^2 d_i}{\rho_{hom}\sigma}$
	$\frac{1}{\rho_{hom}} = \frac{x}{\rho_g} + \frac{1-x}{\rho_l}$ $\quad \sigma$ = surface tension (N/m)
Chisholm correlation (1973) for $\mu_l/\mu_g > 1000$ and $G > 100$ kg/m^{2}s:	Y defined in Equation 13.89; n = 1/4(exponent in $f = C\,Re^n$)
	G = total mass velocity, kg/m^{2}s
Accuracy for annular flow: ±38% Ould Didi et al. (2002)	$B = \begin{cases} 4.8 & G < 500 \\ 2400/G & 500 \le G \le 1900 \\ 55/G^{1/2} & G \ge 1900 \end{cases}$ for $0 < Y < 9.5$
	$B = \begin{cases} 520/YG^{1/2} & G \le 600 \\ 21/G & G > 600 \end{cases}$ for $9.5 < Y \le 28$
	$B = 15{,}000/(Y^2G^{1/2})$ for $Y > 28$
Lockhart–Martinelli correlation (1949) for μ_l/μ_g (1000 and $G <$ 100 kg/m^{2}s:	Correlation constant by Chisholm (1967):
$\varphi_l^2 = \frac{(dp/dz)_{fr}}{(dp/dz)_l} = 1 + \frac{c}{X} + \frac{1}{X^2}$ $\qquad \varphi_g^2 = \frac{(dp/dz)_{fr}}{(dp/dz)_g} = 1 + cX + X^2$	c = 20 for liquid and vapor both turbulent c = 10 for liquid-turbulent, vapor-laminar c = 12 for liquid-laminar, vapor-turbulent c = 5 for liquid and vapor both laminar
Accuracy for annular flow: ±29% Ould Didi et al. (2002)	

Source: From Shah, R. K. and Sekulić, D. P., *Fundamentals of Heat Exchanger Design*. Wiley, Hoboken, NJ, 2003.

The momentum pressure gradient can be calculated by integrating the momentum balance equation (Collier and Thome 1994), thus obtaining

$$\left(\frac{\mathrm{d}p}{\mathrm{d}z}\right)_{\mathrm{mo}} = \frac{\mathrm{d}}{\mathrm{d}z}\left[\frac{G^2}{g_c}\left(\frac{x^2}{\alpha\rho_g} + \frac{(1-x)^2}{(1-\alpha)\rho_l}\right)\right] \tag{13.90}$$

where α represents the void fraction of the gas (vapor) phase (a ratio of volumetric flow rate of the gas/vapor phase divided by the total volumetric flow rate of the two-phase mixture) and x is the mass quality (a ratio of the mass flow rate of the vapor/gas phase divided by the total mass flow rate of the two-phase mixture). Equation 13.90 is valid for constant cross-sectional (flow) area along the flow length. For the homogeneous model, the two-phase flow behaves like a single phase and the vapor and liquid velocities are equal. A number of correlations for the void fraction α are given by Carey (1992); Kandlikar, Shoji, and Dhir (1999). An empirical correlation for the void fraction whose general form is valid for several frequently used models is given by Butterworth (Carey 1992) as

$$\alpha = \left[1 + A\left(\frac{1-x}{x}\right)^p \left(\frac{\rho_g}{\rho_l}\right)^q \left(\frac{\mu_l}{\mu_g}\right)^r\right]^{-1} \tag{13.91}$$

where the constants A, p, q, and r depend on the two-phase model and/or empirical data chosen. These constants for a nonhomogeneous model, based on steam-water data, are $A=1$, $p=1$, $q=0.89$, and $r=0.18$. For the homogeneous model, $A=p=q=1$ and $r=0$. For the Lockhart and Martinelli model, $A=0.28$, $p=0.64$, $q=0.36$, and $r=0.07$. For engineering design calculations, the homogeneous model yields the best results when the slip velocity between the gas and liquid phases is small (for bubbly or mist flows).

Finally, the pressure gradient due to the gravity (hydrostatic) effect is

$$\left(\frac{\mathrm{d}p}{\mathrm{d}z}\right)_{\mathrm{gr}} = \pm\frac{g}{g_c}\sin\theta\left[\alpha\rho_g + (1-\alpha)\rho_l\right] \tag{13.92}$$

Note that the negative sign (i.e., the pressure recovery) stands for downward flow in inclined or vertical tubes/channels and the positive sign (i.e., pressure drop) represents upward flow in inclined or vertical tubes/channels. θ represents the angle of tube/channel inclination measured from the horizontal axis.

13.5.2 Heat Transfer Correlations for Condensation

Condensation represents a vapor–liquid phase-change phenomenon that usually takes place when vapor is cooled below its saturation temperature at a given pressure. The heat transfer rate per unit heat transfer surface area from the pure condensing fluid to the wall is given by

$$q'' = h_{\mathrm{con}}(T_{\mathrm{sat}} - T_{\mathrm{w}}) \tag{13.93}$$

where h_{con} is the condensation heat transfer coefficient, T_{sat} is the saturation temperature of the condensing fluid at a given pressure, and T_{w} is the wall temperature. We summarize here the correlations for filmwise (convective) in-tube condensation, a common condensation mode in most industrial applications, for two most common flow patterns—annular film flow in horizontal and vertical tubes and stratified flow in horizontal tubes. For annular film flow, the correlation for the local heat transfer coefficient h_{loc} [$h_{\mathrm{con}}=h_{\mathrm{loc}}$ in Equation 13.93] is given in Table 13.9; and also for stratified flow, the correlation for mean condensation heat transfer coefficient $h_{\mathrm{con}}=h_{\mathrm{m}}$ is given in Table 13.9. Shah, Zhou, and Tagavi (1999) provide condensation correlations for a number of noncircular flow passage geometries.

TABLE 13.9 Heat Transfer Correlations for Internal Condensation in Horizontal Tubes

Stratification Conditions	Correlation
Annular flow[a] (film condensation) Shah (1979) Accuracy: ±14.4% Kandlikar, Shoji, and Dhir 1999	$h_{loc} = 0.023\frac{k_l}{d_i}\mathrm{Re}_l^{0.8}\mathrm{Pr}_l^{0.4}\left[(1-x)^{0.8} + \frac{3.8x^{0.76}(1-x)^{0.04}}{(p_{sat}/p_{cr})^{0.38}}\right]$ $\mathrm{Re}_l = \frac{Gd_i}{\mu_l}$, xG = total mass velocity, kg/m^2 s $0.002 \le p_{sat}/p_{cr} \le 0.44$ $11 \le G \le 1599$ kg/m^{2}s $21 \le T_{sat} \le 310^\circ$C, $0 \le x \le 1$, $\mathrm{Pr}_l > 0.5$ $3 \le u_{vap} \le 300$ m/s, no limit on q $7 \le d_i \le 40$ mm $\mathrm{Re}_l > 350$ for circular tubes
Stratified flow Carey (1992) Accuracy: ±18% Ould Didi et al. (2002)	$h_m = 0.728\left[1 + \frac{1-x}{x}\left(\frac{\rho_g}{\rho_l}\right)^{2/3}\right]^{-3/4}\left[\frac{k_l^3\rho_l(\rho_l-\rho_g)gh'_{lg}}{\mu_l(T_{sat}-T_w)d_i}\right]^{1/4}$ where $h'_{lg} = h_{lg} + 0.68c_{p,l}(T_{sat} - T_w)$

13.5.3 Heat Transfer Correlations for Boiling

Vaporization (boiling and evaporation) phenomena have been investigated and reported extensively in the literature. In this case, the heat transfer rate per unit heat transfer surface area from the wall to the pure vaporizing fluid is given by

$$q'' = h_{tp}(T_w - T_{sat}) \tag{13.94}$$

where h_{tp} is the two-phase heat transfer coefficient during the vaporization process. We present here a general in-tube force convective boiling correlation proposed by Kandlikar (1991) and further modified by Kandlikar and Steinke (2003); Kandlikar and Balasubramanian (2004). It is based on empirical data for water, refrigerants, and cryogens. The correlation consists of two parts, the convective and nucleate boiling terms and utilizes a fluid-surface parameter. The Kandlikar correlation for the two-phase heat transfer coefficient is as follows:

$$\frac{h_{tp}}{h_{lo}} = \text{larger of}\begin{cases}\left[0.6683\mathrm{Co}^{-0.2}f_2(\mathrm{Fr}_{lo}) + 1058\mathrm{Bo}^{0.7}F_{fl}\right](1-x)^{0.8}\\ \left[1.136\mathrm{Co}^{-0.9}f_2(\mathrm{Fr}_{lo}) + 667.2\mathrm{Bo}^{0.7}F_{fl}\right](1-x)^{0.8}\end{cases} \tag{13.95}$$

where

$$h_{lo} = \begin{cases}\dfrac{\mathrm{Re}_{lo}\mathrm{Pr}_l(f/2)(k_l/d_i)}{1.07 + 12.7(\mathrm{Pr}^{2/3} - 1)(f/2)^{0.5}} & 10^4 \le \mathrm{Re}_{lo} \le 5\times10^6\\ \dfrac{(\mathrm{Re}_{lo} - 1000)\mathrm{Pr}_l(f/2)(k_l/d_i)}{1.00 + 12.7(\mathrm{Pr}_l^{2/3} - 1)(f/2)^{0.5}} & 3000 \le \mathrm{Re}_{lo} \le 10^4\\ \dfrac{Nuk_l}{d_i} & 100 \le \mathrm{Re}_{lo} \le 1600\end{cases} \tag{13.96}$$

Use a linear interpolation for h_{lo} for Re between 1600 and 3000 from h_{lo} calculated from the last two equations of Equation 13.96. For Re_{lo} below 100, use Equation 13.95 where h_{lo} is calculated from Equation 13.96. This is because for Re ≤ 100, the nucleate boiling mechanism governs. Note that Nu in Equation 13.96 is calculated based on the appropriate boundary condition for the single-phase flow.

TABLE 13.10 F_{fl} Recommended by Kandlikar 1991

Fluid	F_{fl}	Fluid	F_{fl}
Water	1.00	R-114	1.24
R-11	1.30	R-134a	1.63
R-12	1.50	R-152a	1.10
R-13B1	1.31	R-32/R-132 (60%−40% wt.)	3.30
R-22	2.20	Kerosene	0.488
R-113	1.30		

$$f_2(\mathrm{Fr_{lo}}) = \begin{cases} (25\ \mathrm{Fr_{lo}})^{0.3} & \text{for } \mathrm{Fr_{lo}} < 0.04 \text{ in horizontal tubes} \\ 1 & \text{for vertical tubes and for } \mathrm{Fr_{lo}} \geq 0.04 \text{ in horizontal tubes} \end{cases} \tag{13.97}$$

$$f = \frac{1}{[1.58\ \ln(\mathrm{Re_{lo}}) - 3.28]^2} \tag{13.98}$$

Here, h_{lo} is the single-phase heat transfer coefficient for the entire flow as liquid flow. Also, the convection number Co, the nucleate boiling number Bo, and the Froude number Fr for the entire flow as liquid are defined as follows:

$$\mathrm{Co} = \left(\frac{\rho_g}{\rho_l}\right)^{0.5}\left(\frac{1-x}{x}\right)^{0.8} \qquad \mathrm{Bo} = \frac{q\prime}{Gh_{fg}} \qquad \mathrm{Fr} = \frac{G^2}{\rho_l^2 g d_i} \tag{13.99}$$

F_{fl} is a fluid-surface parameter and depends on the fluid and the heat transfer surface. F_{fl} values for several fluids in copper tubes are presented in Table 13.10. F_{fl} should be taken as 1.0 for stainless tubes. This correlation is valid for either vertical (upward and downward) or horizontal in-tube flow. A mean deviation of slightly less than 16% with water and 19% with refrigerants has been reported by Kandlikar (1991).

Note that being fluid specific, F_{fl} cannot be used for other fluids (new refrigerants) and mixtures. It is also not accurate for stratified wavy flows and at high vapor qualities since it is not based on the onset of dry out. The Thome model (Kattan, Thome, and Favrat 1998; Zrcher, Thome, and Favrat 1999), based on a flow pattern map, is recommended for those cases.

13.6 Exchanger Design Methodology

The problem of heat exchanger design is complex and multidisciplinary (Shah 1991b). The major design considerations for a new heat exchanger include process/design specifications, thermal and hydraulic design, mechanical design, manufacturing and cost considerations, and trade-offs and system-based optimization, as shown in Figure 13.12 (with possible strong interactions among these considerations as indicated by double-sided arrows). The thermal and hydraulic design has mainly analytical solutions; the structural design has also to some extent analytical/FEA solutions. Most of the other major design considerations involve qualitative and experience-based judgments, trade-offs, and compromises. Therefore, there is no unique solution to designing a heat exchanger for given process specifications. Further details on this design methodology are given by Shah and Sekulić (2003).

Two of the most important heat exchanger design problems are basic rating and sizing problems. Refer to Shah and Sekulić (2003) for a total of 21 rating and sizing problems. Determination of heat transfer and pressure drop performance of either an existing exchanger or an already sized exchanger is referred

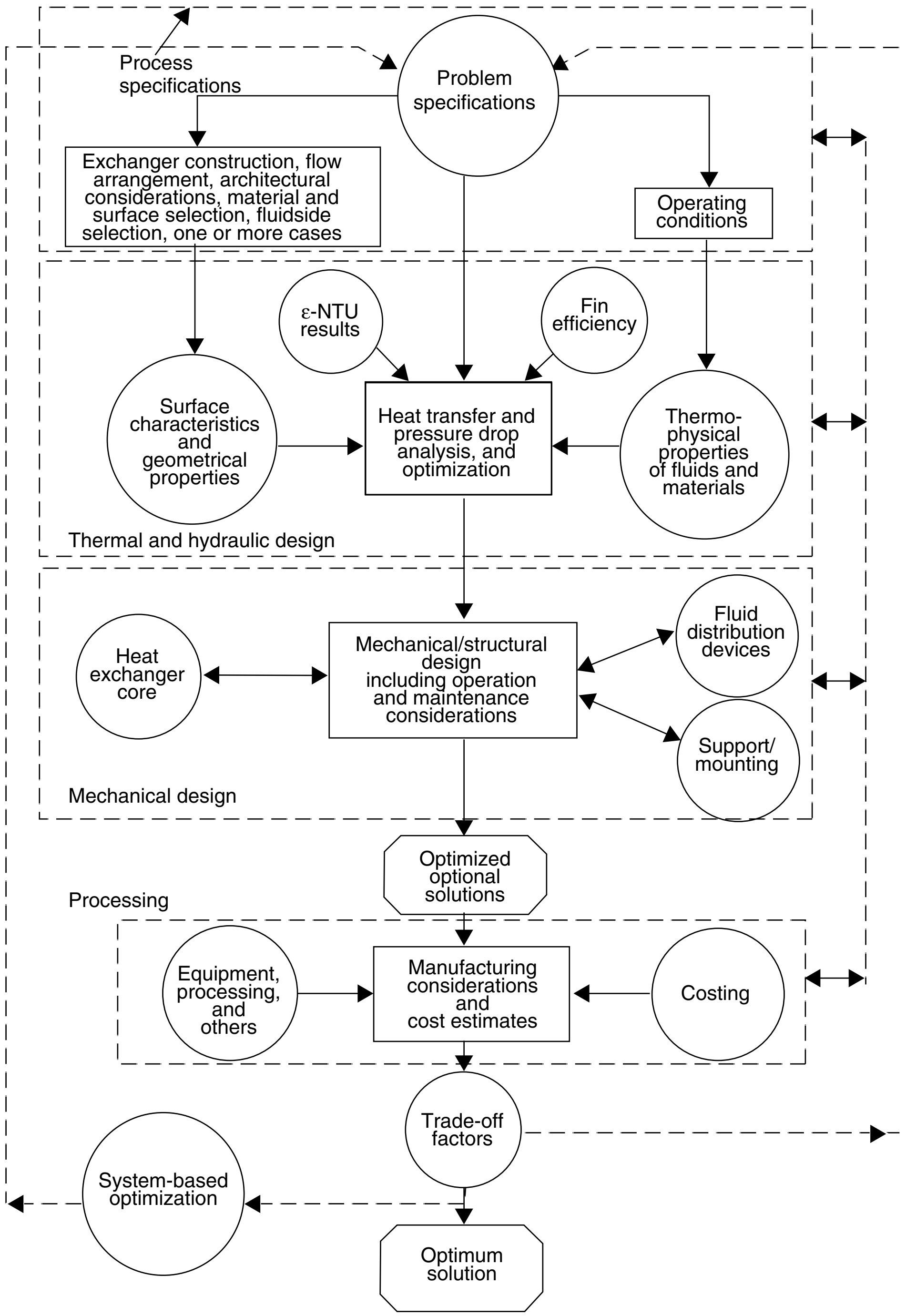


FIGURE 13.12 Heat exchanger design methodology.

to as the *rating problem*. The objective here is to verify the vendor's specifications or to determine the performance at off-design conditions. The rating problem is also sometimes referred to as the performance problem. In contrast, the design of a new or existing type exchanger is referred to as the *sizing problem*. In a broad sense, this means the determination of the exchanger construction type, flow arrangement, heat transfer surface geometries and materials, and physical size of an exchanger to meet the specified heat transfer and pressure drops. However, from the viewpoint of quantitative thermal–hydraulic analysis, we will consider that the selection of the exchanger construction type, flow arrangement, and materials has already been made. Thus in the sizing problem, we will determine the physical size (length, width, and height) and surface areas on each fluid side of the exchanger. The sizing problem is also sometimes referred to as the *design problem*.

The step-by-step solution procedures for the rating and sizing problems for counterflow and crossflow single-pass plate-fin heat exchangers have been presented with a detailed illustrative example by Shah (1981). Shah (1988a) presented further refinements in these procedures as well as step-by-step solution procedures for two-pass cross-counterflow plate-fin exchangers, and single-pass crossflow and two-pass cross-counterflow tube-fin exchangers. Also, step-by-step solution procedures for the rating and sizing problems for rotary regenerators (Shah 1988b), heat-pipe heat exchangers (Shah and Giovannelli 1988), and plate heat exchangers (Shah and Wanniarachchi 1991; Shah and Sekulić 2003) are available. As an illustration, the step-by-step solution procedures will be covered here for a single-pass crossflow plate-fin exchanger.

13.6.1 Rating Problem for a Single-Phase Crossflow Plate-Fin Exchanger

We will present here a step-by-step solution procedure for the rating problem for a crossflow plate-fin exchanger. Inputs to the rating problem for a two-fluid exchanger are the exchanger construction, flow arrangement and overall dimensions, complete details on the materials and surface geometries on both fluid sides including their nondimensional heat transfer and pressure drop characteristics (j and f vs. Re), fluid flow rates, inlet temperatures, and fouling factors. The fluid outlet temperatures, total heat transfer rate, and pressure drops on each fluid side of the exchanger are then determined as the rating problem solution.

1. Determine the surface geometrical properties on each fluid side. This includes the minimum free flow area A_0, heat transfer surface area A (both primary and secondary), flow lengths L, hydraulic diameter D_{h}, heat transfer surface area density β, the ratio of minimum free flow area to frontal area σ, fin geometry (ℓ, δ, etc.) for fin efficiency determination, and any specialized dimensions used for heat transfer and pressure drop correlations.
2. Compute the fluid bulk mean temperature and fluid thermophysical properties on each fluid side. Since the outlet temperatures are not known for the rating problem, they are estimated initially. Unless it is known from the past experience, assume an exchanger effectiveness of 60%–75% for most single-pass crossflow exchangers or 80%–85% for single-pass counterflow and two-pass cross-counterflow exchangers. For the assumed effectiveness, calculate the fluid outlet temperatures.

$$T_{\mathrm{h,o}} = T_{\mathrm{h},i} + \varepsilon(C_{\min}/C_{\mathrm{c}})(T_{\mathrm{h},i} - T_{\mathrm{c},i}) \tag{13.100}$$

$$T_{\mathrm{c,o}} = T_{\mathrm{c},i} + \varepsilon(C_{\min}/C_{\mathrm{c}})(T_{\mathrm{h},i} - T_{\mathrm{c},i}) \tag{13.101}$$

Initially, assume $C_{\mathrm{c}}/C_{\mathrm{h}} \approx \dot{m}_{\mathrm{c}}/\dot{m}_{\mathrm{h}}$ for a gas-to-gas exchanger or $C_{\mathrm{c}}/C_{\mathrm{h}} \approx \dot{m}_{\mathrm{c}}c_{\mathrm{p,c}}/\dot{m}_{\mathrm{h}}c_{\mathrm{p,h}}$ for a gas-to-liquid exchanger with very approximate values of c_{p}s for the fluids in question. For exchangers with $C^* > 0.5$ (usually gas-to-gas exchangers), the bulk mean temperatures on each fluid side will be the arithmetic mean of the inlet and outlet temperatures on each fluid side (Shah 1981). For exchangers with $C^* < 0.5$ (usually gas-to-gas exchangers), the bulk mean temperature on the $C_{\max}$ side will be the arithmetic mean of inlet and outlet temperatures; the bulk mean temperature on the $C_{\min}$ side will be the log-mean average temperature obtained as follows.

$$T_{\mathrm{m},C_{\min}} = T_{\mathrm{m},C_{\max}} \pm \Delta T_{\ell\mathrm{m}} \tag{13.102}$$

where ΔT_{lm} is the log-mean temperature difference based on the terminal temperatures (see Equation 13.19); use the plus sign if the C_{min} side is hot, otherwise use the negative sign. Once the bulk mean temperature is obtained on each fluid side, obtain the fluid properties from thermophysical property books or from handbooks. The properties needed for the rating problem are μ, c_p, k, Pr, and ρ. With this c_p, one more iteration may be carried out to determine $T_{h,0}$ or $T_{c,0}$ from Equation 13.100 or Equation 13.101 on the C_{max} side, and subsequently T_m on the C_{max} side, and refine fluid properties accordingly.

3. Calculate the Reynolds number $\mathrm{Re} = GD_h/\mu$ and/or any other pertinent dimensionless groups (from the basic definitions) needed to determine the nondimensional heat transfer and flow friction characteristics (e.g., j or Nu and f) of heat transfer surfaces on each fluid side of the exchanger. Subsequently, compute j or Nu and f factors. Correct Nu (or j) for variable fluid property effects (Shah and Sekulić 2003) in the second and subsequent iterations from the following equations.

$$\text{For gases,} \quad \frac{\mathrm{Nu}}{\mathrm{Nu}_{cp}} = \left[\frac{T_w}{T_m}\right]^{n'} \qquad \frac{f}{f_{cp}} = \left[\frac{T_w}{T_m}\right]^{m'} \tag{13.103}$$

$$\text{For liquids,} \quad \frac{\mathrm{Nu}}{\mathrm{Nu}_{cp}} = \left[\frac{\mu_w}{\mu_m}\right]^{n'} \qquad \frac{f}{f_{cp}} = \left[\frac{\mu_w}{\mu_m}\right]^{m'} \tag{13.104}$$

where the subscript cp denotes constant fluid properties and m' and n' are empirical constants provided in Table 13.11. Note that T_w and T_m in Equation 13.103 and in Table 13.11a and Table 13.11b are absolute temperatures.

4. From Nu or j, compute the heat transfer coefficients for both fluid streams.

$$h = \mathrm{Nu}k/D_h = jGc_p\mathrm{Pr}^{-2/3} \tag{13.105}$$

Subsequently, determine the fin efficiency η_f and the extended surface efficiency η_0 as follows.

$$\eta_f = \frac{\tanh m\ell}{m\ell} \quad \text{where} \quad m^2 = \frac{h\tilde{P}}{k_f A_k} \tag{13.106}$$

$$\eta_0 = 1 - (1 - \eta_f)\frac{A_f}{A} \tag{13.107}$$

where $\tilde{P}$ is the wetted perimeter of the fin surface.

5. From the known heat capacity rates on each fluid side, compute $C^* = C_{min}/C_{max}$. From the known UA, determine $\mathrm{NTU} = \mathrm{UA}/C_{min}$. Also calculate the longitudinal conduction parameter λ. With the known NTU, C^*, λ, and flow arrangement determine the exchanger effectiveness ε from either closed-form equations of Table 13.3 or tabular/graphical results from Kays and London (1998).

TABLE 13.11a Property Ratio Method Exponents of Equation 13.103 and Equation 13.104 for Laminar Flow

Fluid	Heating	Cooling
Gases	$n' = 0.0,\ m' = 1.00$ for $1 < T_w/T_m < 3$	$n' = 0.0,\ m' = 0.81$ for $0.5 < T_w/T_m < 1$
Liquids	$n' = -0.14,\ m' = 0.58$ for $\mu_w/\mu_m < 1$	$n' = -0.14, m' = 0.54$ for $\mu_w/\mu_m > 1$

TABLE 13.11b Property Ratio Method Correlations or Exponents of Equation 13.103 and Equation 13.104 for Turbulent Flow

Fluid	Heating	Cooling
Gases	$Nu = 5 + 0.012\,Re^{0.83}(Pr + 0.29)\,(T_w/T_m)^n$	$n' = 0$
	$n = -[\log_{10}(T_w/T_m)]^{1/4} + 0.3$	
	for $1 < T_w/T_m < 5$, $0.6 < Pr < 0.9$,	
	$10^4 < Re < 10^6$ and $L/D_h > 40$	
	$m' = -0.1$	$m' = -0.1$ (tentative)
	for $1 < T_w/T_m < 2.4$	
Liquids	$n' = -0.11$[a]	$n' = -0.25$[a]
	for $0.08 < \mu_w/\mu_m < 1$	for $1 < \mu_w/\mu_m < 40$
	$f/f_{cp} = (7 - \mu_m/\mu_w)/6$[b] or $m' \simeq 0.25$	$m' = 0.24$[b]
	for $0.35 < \mu_w/\mu_m < 1$	for $1 < \mu_w/\mu_m < 2$

[a] Valid for $2 \le Pr \le 140$, $10^4 \le Re \le 1.25 \times 10^5$.
[b] Valid for $1.3 \le Pr \le 10$, $10^4 \le Re \le 2.3 \times 10^5$.
Source: From Shah, 1981.

6. With this ε, finally compute the outlet temperatures from Equation 13.100 and Equation 13.101. If these outlet temperatures are significantly different from those assumed in Step 2, use these outlet temperature in Step 2 [and do not use Equations (13.100) and (13.101)] and continue iterating Steps 2–6, until the assumed and computed outlet temperatures converge within the desired degree of accuracy. For a gas-to-gas exchanger, one iteration may be sufficient.
7. Finally, compute the heat duty from

$$q = \varepsilon C_{min}(T_{h,i} - T_{c,i}) \tag{13.108}$$

8. For the pressure drop calculations, we first need to determine the fluid densities at the exchanger inlet and outlet, ρ_i and ρ_0, for each fluid. The mean specific volume on each fluid side is then computed from Equation 13.29. Next, the entrance and exit loss coefficients, K_c and K_e, are obtained from Figure 13.9 for known σ, Re, and the flow passage entrance geometry. The friction factor on each fluid side is corrected for variable fluid properties using Equation 13.103 or Equation 13.104. Here, the wall temperature T_w is computed from

$$T_{w,h} = T_{m,h} - (R_h + R_{h,s})q \tag{13.109}$$

$$T_{w,c} = T_{m,c} + (R_c + R_{c,s})q \tag{13.110}$$

where the various resistance terms are defined by Equation 13.6. The core pressure drops on each fluid side are then calculated from Equation 13.28. This then completes the procedure for solving the rating problem.

13.6.2 Sizing Problem for a Single-Phase Crossflow Plate-Fin Exchangers

As defined earlier, we will concentrate here to determine the physical size (length, width, and height) of a single-pass counterflow and crossflow exchangers for specified heat duty and pressure drops. More specifically inputs to the sizing problem are surface geometries (including their nondimensional heat transfer and pressure drop characteristics), fluid flow rates, inlet and outlet fluid temperatures, fouling factors, and pressure drops on each fluid side.

For the solution to sizing of counterflow problem, there are four unknowns: two flow rates or Reynolds numbers (to determine correct heat transfer coefficients and friction factors) and two surface areas for the two-fluid crossflow exchanger. The following four equations (Equation 13.111, Equation 13.113 for hot and cold fluids, and Equation 13.115) are used to solve iteratively the surface areas on each fluid side: UA in Equation 13.111 is determined from NTU computed from the known heat duty or ε and C^*; G in

Equation 13.113 represents two equations, for Fluid 1 and 2 (Shah 1988a); and the volume of the exchanger in Equation 13.115 is the same based on the surface area density of Fluid 1 and 2 sides.

$$\frac{1}{UA} \approx \frac{1}{(\eta_0 hA)_{\rm h}} + \frac{1}{(\eta_0 hA)_{\rm c}} \tag{13.111}$$

Here, we have neglected the wall and fouling thermal resistances. This equation in nondimensional form is given by

$$\frac{1}{\rm NTU} = \frac{1}{{\rm ntu_h}(C_{\rm h}/C_{\rm min})} + \frac{1}{{\rm ntu_c}(C_{\rm c}/C_{\rm min})} \tag{13.112}$$

$$G = \left[\frac{2g_c \Delta p}{\rm Deno}\right]^{1/2} \tag{13.113}$$

$${\rm Deno} = \frac{f}{j}\frac{\rm ntu}{\eta_0}{\rm Pr}^{2/3}\left(\frac{1}{\rho}\right)_{\rm m} + 2\left(\frac{1}{\rho_0} - \frac{1}{\rho_{\rm i}}\right) + (1-\sigma^2+K_{\rm c})\frac{1}{\rho_{\rm i}} - (1-\sigma^2-K_{\rm e})\frac{1}{\rho_0} \tag{13.114}$$

$$V = \frac{A_1}{\alpha_1} = \frac{A_2}{\alpha_2} \tag{13.115}$$

In the iterative solutions, the first time one needs $\rm ntu_h$ and $\rm ntu_c$ to start the iterations. These can be determined either from the past experience or by estimations. If both fluids are gases or liquids, one could consider that the design is "balanced," that is, the thermal resistances are distributed approximately equally on the hot and cold fluid sides. In that case, $C_{\rm h} = C_{\rm c}$, and

$${\rm ntu_h} \approx {\rm ntu_c} \approx 2{\rm NTU} \tag{13.116}$$

Alternatively, if we have liquid on one fluid side and gas on the other fluid side, consider 10% thermal resistance on the liquid side; that is,

$$0.10\left(\frac{1}{UA}\right) = \frac{1}{(\eta_0 hA)_{\rm liq}} \tag{13.117}$$

Then from Equation 13.111 and Equation 13.112 with $C_{\rm gas} = C_{\rm min}$, we can determine ntu on each fluid side as follows.

$${\rm ntu_{gas}} = 1.11{\rm NTU},\ {\rm ntu_{liq}} = 10C^*{\rm NTU} \tag{13.118}$$

Also note that initial guesses of $\eta_{\rm o}$ and j/f are needed for the first iteration to solve Equation 13.113. For a good design, consider $\eta_{\rm o} = 0.80$ and determine approximate value of j/f from the plot of j/f vs. *Re* curve for the known j and f vs. *Re* characteristics of each fluid side surface. The specific step-by-step design procedure is as follows.

1. To compute the fluid bulk mean temperature and the fluid thermophysical properties on each fluid side, determine the fluid outlet temperatures from the specified heat duty.

$$q = (\dot m c_{\rm p})_{\rm h}(T_{\rm h,i} - T_{\rm h,o}) = (\dot m c_{\rm p})_{\rm c}(T_{\rm c,o} - T_{\rm c,i}) \tag{13.119}$$

   or from the specified exchanger effectiveness using Equation 13.100 and Equation 13.101. For the first time, estimate the values of $c_{\rm p}$s. For exchangers with $C^* \geq 0.5$, the bulk mean temperature on each fluid side will be the arithmetic mean of inlet and outlet temperatures on each fluid side. For exchangers with $C^* < 0.5$, the bulk mean temperature on the $C_{\rm max}$ side will be the arithmetic mean of the inlet and outlet temperatures on each fluid side, the bulk mean temperature on the C_{min} side will be the log-mean average as given by Equation 13.102. With these bulk mean temperatures, determine c_p and iterate one more time for the outlet temperatures if warranted. Subsequently, determine μ, c_p, k, Pr, and ρ on each fluid side.

2. Calculate C^* and ε (if q is given), and determine NTU from the ε-NTU expression, tabular or graphical results for the selected flow arrangement (in this case, it is unmixed–unmixed crossflow, Table 13.3). The influence of longitudinal heat conduction, if any, is ignored in the first iteration since we do not know the exchanger size yet.
3. Determine ntu on each fluid side by the approximations discussed with Equation 13.116 and Equation 13.118 unless it can be estimated from the past experience.
4. For the selected surfaces on each fluid side, plot j/f vs. Re curve from the given surface characteristics, and obtain an approximate mean value of j/f. If fins are employed, assume $\eta_o = 0.80$ unless a better value can be estimated.
5. Evaluate G from Equation 13.113 on each fluid side using the information from Steps 1 to 4 and the input value of Δp.
6. Calculate Reynolds number Re, and determine j and f on each fluid side from the given design data for each surface.
7. Compute h, η_f, and η_o using Equation 13.105 through Equation 13.107. For the first iteration, determine U_1 on Fluid 1 side from the following equation derived from Equation 13.6 and Equation 13.115.

$$\frac{1}{U_1} = \frac{1}{(\eta_0 h)_1} + \frac{1}{(\eta_0 h_s)_1} + \frac{\alpha_1/\alpha_2}{(\eta_0 h_s)_2} + \frac{\alpha_1/\alpha_2}{(\eta_0 h)_2} \tag{13.120}$$

where $\alpha_1/\alpha_2 = A_1/A_2$, $\alpha = A/V$, V is the exchanger total volume, and subscripts 1 and 2 denote Fluid 1 and 2 sides. For a plate-fin exchanger, αs are related to βs as follows (Shah 1981; Kays and London 1998).

$$\alpha_1 = \frac{b_1 \beta_1}{b_1 + b_2 + 2\delta} \qquad \alpha_2 = \frac{b_2 \beta_2}{b_1 + b_2 + 2\delta} \tag{13.121}$$

Here, β is the total surface area on one fluid side to the volume between plates on that fluid side. Note that the wall thermal resistance in Equation 13.120 is ignored in the first iteration. In second and subsequent iterations, compute U_1 from

$$\frac{1}{U_1} = \frac{1}{(\eta_0 h)_1} + \frac{1}{(\eta_0 h_s)_1} + \frac{\delta A_1}{k_w A_w} + \frac{A_1/A_2}{(\eta_0 h_s)_2} + \frac{A_1/A_2}{(\eta_0 h)_2} \tag{13.122}$$

where the necessary geometry information A_1/A_2 and A_1/A_w are determined from the geometry calculated in the previous iteration.

8. Now calculate the core dimensions. In the first iteration, use NTU computed in Step 2. For subsequent iterations, calculate longitudinal conduction parameter λ (and other dimensionless groups for a crossflow exchanger). With known ε, C^*, and λ, determine the correct value of NTU using either a closed-form equation or tabulated/graphical results (Shah and Mueller 1985). Determine A_1 from NTU using U_1 from the previous step and known C_{min},

$$A_1 = \frac{\mathrm{NTU}\,C_{min}}{U_1} \tag{13.123}$$

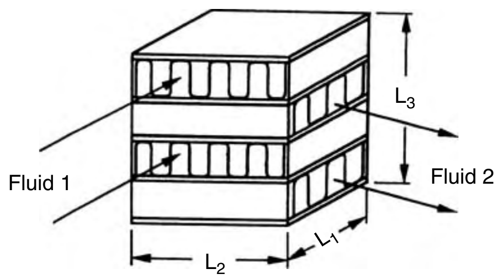


FIGURE 13.13 A single-pass crossflow exchanger.

and hence

$$A_2 = \frac{A_2}{A_1} A_1 = \frac{\alpha_2}{\alpha_1} A_1 \tag{13.124}$$

The free flow area A_0 from known $\dot{m}$ and G is given by

$$A_{0,1} = \left(\frac{\dot{m}}{G}\right)_1 \qquad A_{0,2} = \left(\frac{\dot{m}}{G}\right)_2 \tag{13.125}$$

so that

$$A_{\mathrm{fr},1} = \frac{A_{0,1}}{\sigma_1} \qquad A_{\mathrm{fr},2} = \frac{A_{0,2}}{\sigma_2} \tag{13.126}$$

where σ_1 and σ_2 are generally specified for the surface or can be computed for plate-fin surfaces from (Kays and London 1998; Shah and Sekulić 2003):

$$\sigma_1 = \frac{b_1 \beta_1 D_{\mathrm{h},1}/4}{b_1 + b_2 + 2\delta} = \frac{\alpha_1 D_{\mathrm{h},1}}{4} \qquad \sigma_2 = \frac{b_2 \beta_2 D_{\mathrm{h},2}/4}{b_1 + b_2 + 2\delta} = \frac{\alpha_2 D_{\mathrm{h},2}}{4} \tag{13.127}$$

Now compute the fluid flow lengths on each fluid side (see Figure 13.13) from the definition of the hydraulic diameter of the surface employed on each fluid side.

$$L_1 = \left(\frac{D_{\mathrm{h}} A}{4A_0}\right)_1 \qquad L_2 = \left(\frac{D_{\mathrm{h}} A}{4A_0}\right)_2 \tag{13.128}$$

Since $A_{\mathrm{fr},1} = L_2/L_3$ and $A_{\mathrm{fr},2} = L_1/L_3$, we can obtain

$$L_3 = \frac{A_{\mathrm{fr},1}}{L_2} \quad \text{or} \quad L_3 = \frac{A_{\mathrm{fr},2}}{L_1} \tag{13.129}$$

Theoretically, L_3s calculated from both expressions of Equation 13.129 should be identical. In reality, they may differ slightly due to the round-off error. In that case, consider an average value for L_3.

9. Now compute the pressure drop on each fluid side, after correcting f factors for variable property effects, in a manner similar to Step 8 of the rating problem for a crossflow plate-fin exchanger.
10. If the values calculated for Δps are within input specifications and close to them, the solution to the sizing problem is completed. Finer refinements in the core dimensions, such as integer numbers of flow passages, may be carried out at this time. Otherwise, compute the new value of G on each fluid side using Equation 13.28 in which Δp is the input specified value, and f, K_c, K_e, and the geometrical dimensions are from the previous iteration.
11. Repeat (iterate) Steps 6–10 until both heat transfer and pressure drops are met as specified. It should be emphasized that since we have imposed no constraints on the exchanger dimensions, the procedure will yield L_1, L_2, and L_3, for the selected surfaces such that the design will meet the heat duty and pressure drops on both fluid sides exactly.

Refer to Shah and Sekulić (2003) for taking into account the effect of longitudinal conduction in the exchanger as well as rating and sizing problem solution for other exchanger flow arrangements.

13.6.3 Rating and Sizing Problem for Condensers and Evaporators

Generally, the heat transfer coefficients and friction factors for two-phase flows are so much changing along the flow length that the lumped parameter approach for rating and sizing of the exchanger is not followed as is done for single-phase exchangers and discussed in the preceding section. The heat exchanger is divided into a large number of small elements and the mass, momentum, and energy equations are solved for each element along with the evaluation of quality of the phase-change fluid after each element. Because of complicated two-fluid flow arrangements for many phase-change exchangers, such calculations may not be possible even for the starting element since only one fluid side inlet conditions are known on that fluid side. Hence, for most phase-change exchangers, the conditions for the fluid side with unknown parameters are assumed and the finite difference calculations are pursued. After the first complete iteration for the exchanger, when the inlet conditions for the second fluid are computed and do not match, iterative calculations are done with the appropriate iterative scheme to get the converged solution. In the case of a sizing problem, first the exchanger size with the appropriate two-fluid flow arrangement is chosen and the rating calculations are done as mentioned earlier. If the converged solution does not provide the desired performance, a new exchanger size is chosen and iterative calculations are performed to get the new converged solution. This process of assuming the exchanger size is continued and iterative calculations are performed till the desired performance is achieved. Due to the complexity of the procedure and the length limitations for the article, the details on the rating and sizing of the two-fluid condensers and evaporators are not being covered here.

13.7 Flow Maldistribution

In previously presented heat transfer methods (ε-NTU, MTD, etc.) and pressure drop analyses, it is presumed that the fluid is uniformly distributed through the core. In practice, flow maldistribution occurs to some extent and often severely, and may result in a significant reduction in exchanger heat transfer performance and an increase in the pressure drop. Hence, it may be necessary for the designer to take into account the effect of flow maldistribution, causing undesirable performance deterioration upfront while designing a heat exchanger.

Some maldistributions are geometry induced (i.e., the result of exchanger fabrication conditions, such as header design or manufacturing tolerances, or the duct geometry/structure upstream of the exchanger); others are the result of exchanger operating conditions. Gross, passage-to-passage, and manifold-induced flow maldistributions are the examples of the former category; flow maldistributions induced by viscosity, natural convection, and density difference are of the latter category. Flow maldistributions associated with two-phase and multiphase flow are too complex and beyond the scope of this chapter. The analysis methods and results for some of these flow maldistributions for single-phase flows are summarized next.

13.7.1 Gross Flow Maldistribution

In gross flow maldistribution, the nonuniform distribution of the fluid flow is on the macroscopic level. It results due to the shape of the inlet header, upstream flow conditions, or gross blockage (caused by brazing, soldering, or other manufacturing considerations) of any part of the exchanger. Therefore, gross flow maldistribution is independent of the exchanger heat transfer surface geometry configuration and its microscopic nonuniformity. Gross flow maldistribution generally results in a *significant increase* in exchanger pressure drop and *some reduction* in exchanger heat transfer.

A design procedure for calculating the reduction in heat exchanger effectiveness due to one-dimensional nonuniform flow distribution on only one fluid side of the exchanger has been introduced by Shah (1981). The flow maldistribution in this procedure is characterized by an *n-step* velocity function

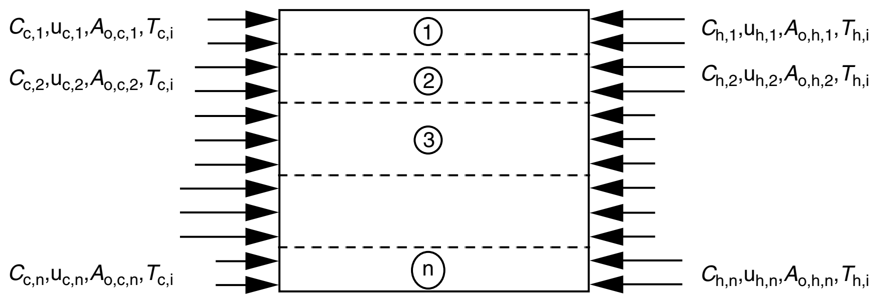


FIGURE 13.14 Idealized flow nonuniformity in a counterflow exchanger.

relating the flow nonuniformity as n adjacent subexchangers, each having a uniform flow distribution unique from the rest. This method has been applied to a counterflow, a parallel flow, and a single-pass unmixed–mixed crossflow exchanger (for the case where the nonuniform fluid side is the unmixed side) with the maldistributed fluid side having either the hot or cold fluid. For all other exchanger flow arrangements, a numerical analysis is essential.

In the case of a counterflow or parallel flow exchanger, the flow nonuniformity can be on hot or cold fluid sides as shown in Figure 13.14. The analysis for the exchanger of Figure 13.14 is straightforward, computing temperature effectiveness of individual subexchangers based on one fluid side only, and computing the total heat transfer rate by adding the qs of the subexchangers (Shah and Sekulić 2003).

A typical influence of gross flow maldistribution on the counterflow exchanger effectiveness is shown in Figure 13.15 (Shah 1985). The subject exchanger contains a two-step velocity maldistribution on the hot fluid side with $C^* = 1$, $u_{max}/u_m = 1.5$ and 2.0, and $A_{0,1}/A_0 = 0.50$, 0.35, and 0.25. It is idealized that the flows are fully-developed laminar; hence, h and U are constant. Otherwise, the variations in h should be evaluated at respective flow velocities (i.e., Reynolds numbers) and the subsequent new values of U

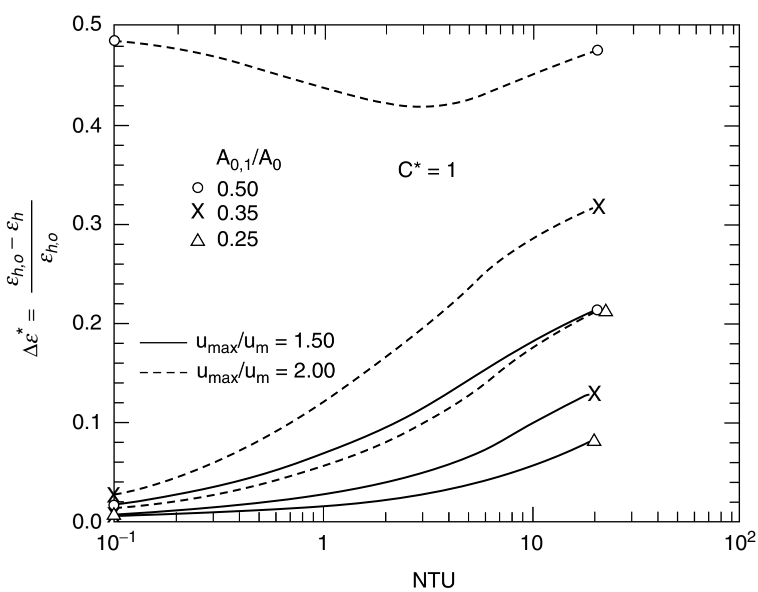


FIGURE 13.15 Influence of gross flow maldistribution on the counterflow exchanger effectiveness. (From Shah, R. K., *Handbook of Heat Transfer Applications, 2nd Ed.*, Rohsenow, W. M. Hartnett, J. P. and Ganić, E. N. Part 3. McGraw-Hill, New York, Chap. 4, 1985. With permission.)

(or UA) should be calculated. From Figure 13.16, it can be seen that for particular values of u_{max}/u_m and exchanger NTU, the greatest reduction in exchanger effectiveness occurs when the velocity function is equally distributed over flow area (e.g., $A_{0,1}/A_0 = A_{0,2}/A_0 = 0.5$). Also, as expected, the greater the maldistribution, the greater is the loss in the exchanger effectiveness.

Although not demonstrated here, it can be shown that an n-step ($n > 2$) velocity maldistribution results in less deterioration in ε compared to that for the two-step velocity maldistribution for the same u_{max}/u_m ratio for the same exchanger (Shah 1981).

In a series of papers, Chiou investigated numerically the effect of various flow maldistributions on a single-pass unmixed–unmixed crossflow exchanger, and some of these results are summarized by Mueller and Chiou (1987). Chiou (1982) reported a lower reduction in the exchanger effectiveness for nonuniform inlet temperature than for nonuniform inlet velocity for a single-pass unmixed–unmixed crossflow exchanger. Cichelli and Boucher (1956); Fleming (1966), and Chowdhury and Sarangi (1985) analyzed the case of tube-side maldistribution in a counterflow shell-and-tube exchanger, while Mueller (1977) considered flow maldistributions on both tube and shell sides. For a rotary regenerator, Kutchey and Julien (1974) investigated gross flow maldistribution, while Kohler (1974) investigated the influence of flow path geometry and manufacturing tolerances.

No analysis is available for an increase in the pressure drop due to gross flow maldistribution. Since such nonuniform flow is associated with poor header design or gross core blockage, the static pressure distributions at the core inlet and outlet faces may not be uniform. Hence, no simple modeling for the pressure drop evaluation is possible. As a conservative approach, it is recommended that the core pressure drop on flow maldistributed side is evaluated based on the highest velocity component of that fluid side.

13.7.2 Passage-to-Passage Flow Maldistribution

Neighboring passages in a compact heat exchanger are geometrically never identical because of manufacturing tolerances. It is especially difficult to control precisely the passage size when small dimensions are involved (e.g., a rotary regenerator with $D_h = 0.5$ mm or 0.020 in.) in hundreds of neighboring passages. Since differently sized and shaped passages exhibit different flow resistances in laminar flow and the flow seeks a path of least resistance, a nonuniform flow results through the matrix. This passage-to-passage flow nonuniformity can result in a *significant penalty* in heat transfer performance with only a small compensating effect of *reduced* pressure drop. This effect is especially important for laminar flows in continuous cylindrical passages, but is of lesser importance for interrupted surfaces (where transverse flow mixing can occur) or for turbulent flow.

The theoretical analysis for this flow maldistribution for low-*Re* laminar flow surfaces has been carried out by London (1970) for a two-passage model and extended by Shah and London (1980) for an n-passage model. In the latter model, there are n different-size passages of the same basic shape, either rectangular or triangular. In Figure 13.16, a reduction in ntu for rectangular passages is shown when 50% of the flow passages are large ($c_2 > c_r$) and 50% of the passages are small ($c_1 < c_r$) compared to the reference or nominal passages. The results are presented for the passages having a nominal aspect ratio α^* of 1, 0.5, 0.25, and 0.125 for the H1 and T boundary conditions and for a reference ntu_r of 5.0. Here, ntu_{cost}, a percentage loss in ntu, and the channel deviation parameter δ_c are defined as:

$$\text{ntu}^*_{\cos\text{t}} = \left(1 - \frac{\text{ntu}_{\text{eff}}}{\text{ntu}_r}\right) \tag{13.130}$$

$$\delta_c = 1 - \frac{c_i}{c_r} \tag{13.131}$$

where ntu_{eff} is the effective ntu when two-passage model passage-to-passage nonuniformity is present, and ntu_r is the reference or nominal NTU. It can be seen from Figure 13.16 that a 10% channel deviation (which is common for a highly compact surface) results in 10 and 21% reduction in NTU_{H1} and NTU_{T}, respectively, for $\alpha^* = 0.125$ and $\text{ntu}_r = 5.0$. In contrast, a gain in the pressure drop due to

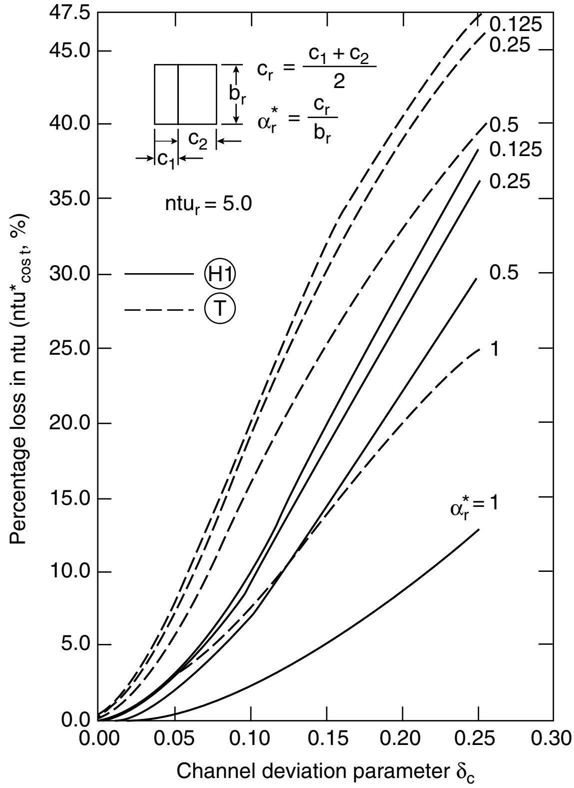


FIGURE 13.16 Percentage loss in NTU for two-passage nonuniformities in rectangular passages.

passage-to-passage nonuniformity is only 2.5% for $\delta_c = 0.10$ and $\alpha_r^* = 0.125$, as found from Figure 13.17. Here, Δp^*_{gain} is defined as:

$$\Delta p^*_{\text{gain}} = \left(1 - \frac{\Delta p_{\text{actual}}}{\Delta p_{\text{nominal}}}\right) \tag{13.132}$$

The results of Figure 13.16 and Figure 13.17 are also applicable to an n-passage model in which there are n different-size passages in a normal distribution about the nominal passage size. The channel deviation parameter needs to be modified for this case to

$$\delta_c = \left[\sum_{i=1}^{n} \chi_i \left(1 - \frac{c_i}{c_r}\right)^2\right]^{1/2} \tag{13.133}$$

Here, χ_i is the fractional distribution of the ith shaped passage. For $n=2$, Equation 13.133 reduces to Equation 13.131. For triangular passages, c in Equation 13.133 is replaced by r_h. The following observations may be made from Figure 13.16 and additional results presented by Shah and London (1980): (1) the loss in ntu is more significant for the T boundary condition than for the H1 boundary condition. (2) The loss in ntu increases with higher nominal ntu_r. (3) The loss in ntu is much more significant compared to the gain in Δp at a given δ_c. (4) The deterioration in performance is the highest for the two-passage model compared to the n-passage model with $n > 2$ for the same value of δ_c. Refer to Shah and Sekulić (2003) for details on additional analyses.

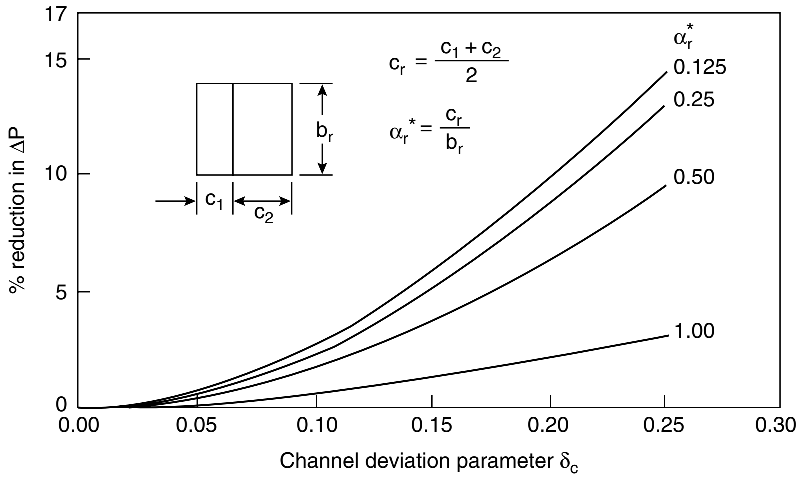


FIGURE 13.17 Percentage reduction in Δp for two-passage nonuniformities in rectangular passages.

13.7.3 Manifold-Induced Flow Maldistribution

Two of the most commonly used manifold systems connecting manifolds to the branches (particularly for plate heat exchangers) are the parallel flow (S or Z) and reverse-flow (U) systems as shown in Figure 13.18 for a single-pass exchanger. Since the flow length (and subsequently the flow resistance) is different for fluid particles going through different branches and inlet/outlet manifolds, and the imposed pressure drop (between inlet and outlet of the exchanger) is the same, it results in flow maldistribution regardless of perfect flow passages (branch geometry) and the type of the flow through the manifolds and branches (such as laminar, transition, or turbulent). Bajura and Jones (1976); Majumdar (1980); Shen (1992) have investigated this flow maldistribution, and qualitative conclusions of the first two investigators are summarized below due to space limitation; for some quantitative results, refer to Shah (1985).

1. A design rule of thumb is to limit the ratio of the flow area of lateral branches (exchanger core) to the flow area of the inlet header (area of pipe before lateral branches) A_0^* to less than unity to minimize flow maldistributions.
2. More uniform flow distribution through the core is achieved by a reverse-flow (U) manifold system in comparison to a parallel flow (S or Z) manifold system.
3. In a parallel flow (S or Z) manifold system, the maximum flow occurs through the last port, and in the reverse-flow (U) manifold system the first port.
4. The influence of the manifold pipe friction parameter F is less significant than that of the flow area ratio A_0^*.
5. Flow distribution becomes more uniform with a higher branch pressure loss coefficient K.
6. The flow area of a combining-flow manifold (outlet manifold/pipe in Figure 13.18a and b) should be larger than that for the dividing-flow manifold for a more uniform flow distribution through the core in the absence of heat transfer within the core. If there is heat transfer in the lateral branches (core), the fluid densities will be different in the inlet and outlet manifolds, the flow areas should be adjusted for this density change, and then the flow area of the combing manifold should be larger than that calculated previously.
7. Flow reversal is more likely to occur in parallel flow (S flow) systems that are subject to poor flow distribution.

Refer to Shah and Sekulić (2003) for further details.

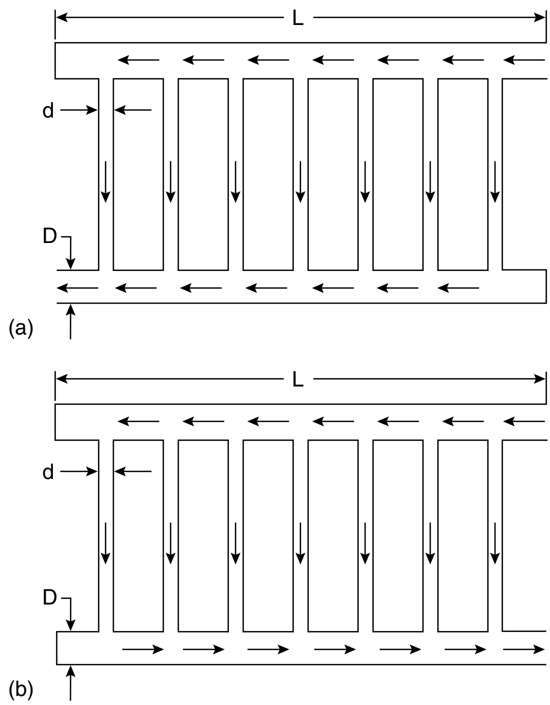


FIGURE 13.18 (a) Parallel flow or *S* or *Z* flow and (b) reverse-flow or U flow manifold systems.

13.7.4 Viscosity-Induced Flow Maldistribution

Whenever one or both fluids flowing in a heat exchanger are liquids and operate in the laminar flow region, there exists a possibility of viscosity-induced flow instability and maldistribution, and thus exchanger performance deterioration. The viscosity-induced flow instability is a result of large changes in fluid viscosity within the exchanger and is found to be present when the viscous liquid is being cooled; it is not present for the case of liquid being heated, although flow maldistribution will be present. For the case of liquid cooling, the liquid viscosity increases significantly with decreasing temperature. Thus, moderate temperature differences between parallel passages (or tubes) within an exchanger can result in large viscosity differences between these adjacently flowing streams. Because of the direct proportionality between the pressure drop and the product of flow rate and viscosity ($\Delta p \propto \dot{m}\mu$) in laminar flow, flow will be maldistributed in tubes having different heat transfer rates (and hence resultant different viscosities for the tube fluid), but will have the same pressure drop across each tube. The end result is that slower fluid streams are forced to flow slower and faster fluid streams to flow faster. This trend normally continues until a new equilibrium is reached, but unfortunately the resulting flow nonuniformity of this new exchanger condition will generally cause undesirable performance deterioration.

Mueller (1974); Putman and Rohsenow (1985) addressed the problem of viscosity-induced flow maldistribution from the flow instability point of view and not from the heat transfer performance deterioration point of view. Mueller (1974) provided a simple analysis of the pressure drop–mass flow rate relationship in a single tube laminar flow cooler. From an extension of this analysis, he proposed a method for determining a point (maximum pressure drop or flow rate) beyond which it would be safe to assume flow instability of this kind will not occur in a multitubular exchanger.

When the exchanger is operating outside the region of viscous instability, the deterioration in heat transfer performance and the pressure drop "gain" can be calculated using the same procedure as that for the passage-to-passage flow maldistribution. In this case, the flow rate ratio depends on the viscosity ratio, whereas it depends on the ratios of the friction factor, hydraulic diameter, and flow area of two different flow passages in the case of passage-to-passage flow maldistribution (London 1970). Refer to Shah and Sekulić (2003) for more details and insight into this flow maldistribution.

13.8 Fouling in Heat Exchangers

13.8.1 Fouling, Its Effect and Mechanisms

Fouling refers to undesired accumulation of solid material (by-products of the heat transfer processes) on heat exchanger surfaces, which results in additional thermal resistance to heat transfer, thus reducing exchanger performance. The fouling layer also blocks the flow passage/area and increases surface roughness, which either reduces the flow rate in the exchanger or increases the pressure drop or both. The foulant deposits may be loose, such as magnetite particles, or hard and tenacious, such as calcium carbonate scale; other deposits may be sediment, polymers, cooking or corrosion products, inorganic salts, biological growth, and so on. Depending on the fluids, operating conditions, and heat exchanger construction, the maximum fouling layer thickness on the heat transfer surface may result in a few hours to a number of years.

Fouling in heat exchangers is a very complex phenomena and its negative impact translates into billions of dollars every year worldwide and is being investigated with considerable effort. Significant research activity in the fouling of heat exchanger equipment is going on worldwide. Some of the recent references for further studies are: Bott (1995); Panchal et al. (1997); Bott et al. (1999), and Müller-Steinhagen (2000), Bott (1995), Hewitt and Müller-Steinhagen (2000, 2003).

Fouling could be very costly depending on the nature of fouling and the applications. It increases capital costs: (a) oversurfacing the heat exchanger, (b) provisions for cleaning, and (c) use of special materials and constructions/surface features. It increases maintenance costs: (a) cleaning techniques, (b) chemical additives, and (c) troubleshooting. It may cause a loss of production: (a) reduced capacity, and (b) shutdown. It increases energy losses: (a) reduced heat transfer, (b) increased pressure drop, and (c) dumping dirty streams. Fouling promotes corrosion, severe plugging, and eventual failure of uncleaned heat exchangers. In a fossil-fired exhaust environment, gas-side fouling produces a potential fire hazard in heat exchangers.

The following are the major fouling mechanisms:

- *Crystallization or Precipitation Fouling*. This results from the deposition/formation of crystals of dissolved substances from the liquid onto the heat transfer surface due to solubility changes with temperatures beyond the saturation point. If the deposited layer is hard and tenacious, it is often referred to as scaling. If it is porous and mushy, it is called sludge.
- *Particulate Fouling*. This results from the accumulation of finely divided substances suspended in the fluid stream onto the heat transfer surface. If the settling occurs due to gravity, it is referred to as *sedimentation fouling*.
- *Chemical Reaction Fouling*. This is defined as the deposition of material produced by chemical reaction (between reactants contained in the fluid stream) in which the heat transfer surface material does not participate.
- *Corrosion Fouling*. This results from corrosion of the heat transfer surface that produces products fouling the surface and/or roughens the surface, promoting attachment of other foulants.
- *Biological Fouling*. This results from the deposition, attachment, and growth of biological organisms from liquid onto a heat transfer surface. Fouling due to microorganisms refers to *microbial fouling* and due to macroorganisms refers to *macrobial fouling*.
- *Freezing Fouling*. This results from the freezing of a single component liquid or higher-melting point constituents of a multicomponent liquid onto a subcooled heat transfer surface.

Biological fouling occurs only with liquids, since there are no nutrients in the gases. Also, crystallization fouling is not too common with gases, since most gases contain a few dissolved salts (mainly in mists) and even fewer inverse-solubility salts. All other types of fouling occur in both liquid and gas. More than one mechanism is usually present in many fouling situations, often with synergetic results. Liquid-side fouling generally occurs on the exchanger side where the liquid is being heated, and gas-side fouling occurs where the gas is being cooled; however, reverse examples can be found.

13.8.2 Importance of Fouling

Fouling in liquids and two-phase flows has a significant detrimental effect on heat transfer with some increase in pressure drop. In contrast, fouling in gases reduces heat transfer somewhat (5%–10% in general) in compact heat exchangers but increases pressure drop significantly (up to several hundred percent). For example, consider $U = 1400$ W/m^2 K as in a process plant liquid-to-liquid heat exchanger. Hence, $R = 1/U = 0.00072$ m^2 K/W. If the fouling factors ($r_{h,s} + r_{c,s}$) together amount to 0.00036 (considering a typical TEMA value of the fouling factor as 0.00018), 50% of the heat transfer area A requirement for given q is chargeable to fouling. However, for gas flows on both fluid sides of an exchanger, $U \approx 280$ W/m^2 K, and the same fouling factor of 0.00036 would represent only about 10% of the total surface area. Thus, one can see a significant impact on heat transfer surface area requirement due to fouling in heat exchangers having high U values (such as having liquids or phase-change flows).

Considering the core frictional pressure drop, Equation 13.31, as the main pressure drop component, the ratio of pressure drops of fouled and cleaned exchanger is given by

$$\frac{\Delta p_F}{\Delta p_C} = \frac{f_F}{f_C}\left(\frac{D_{h,C}}{D_{h,F}}\right)\left(\frac{u_{m,F}}{u_{m,C}}\right)^2 = \frac{f_F}{f_C}\left(\frac{D_{h,C}}{D_{h,F}}\right)^5 \tag{13.134}$$

where the term after the second equality sign is for a circular tube and the mass flow rates under fouled and clean conditions remain the same. Generally $f_F > f_C$ due to the fouled surface being rough. Thus although the effect of fouling on the pressure drop is usually neglected, it can be significant, particularly for compact heat exchangers with gas flows. If we consider $f_F = f_C$ and the reduction in the tube inside diameter due to fouling by only 10 and 20%, the resultant pressure drop increase will be 69 and 205%, respectively, according to Equation 13.134, regardless of whether the fluid is liquid or gas! However, such a significant increase in Δp will have a small effect on fluid pumping power for liquids compared to gases since $\mathcal{P} \propto \Delta p/\rho$ for a given flow rate and liquid density in general is considerably higher than the gas density.

13.8.3 Accounting of Fouling in Heat Exchangers

Fouling is an extremely complex phenomena characterized by a combined heat, mass, and momentum transfer under transient condition. Fouling is affected by a large number of variables related to heat exchanger surfaces, operating conditions, and fluids. Fouling is time (τ) dependent and zero at $\tau = 0$. After the induction or delay period τ_d, the fouling resistance is pseudolinear, a falling rate, or asymptotic, as shown by the smooth lines in Figure 13.19, and the actual transient behavior is as represented by the dashed line for one asymptotic behavior.

Fouling is characterized by all or some of the following sequential events: initiation, transport, attachment, removal, and aging (Epstein 1983). Research efforts are concentrated on quantifying these events by semi-theoretical models (Epstein 1978) with a very limited success on specific fouling situations. Hence, the current heat exchanger design approach is to use a constant (supposedly an asymptotic) value of the fouling factor $r_s = 1/h_s$. Equation 13.6 presented earlier includes the fouling resistances on the hot and cold fluid sides for a nontubular extended surface exchanger. Here $1/h_s = r_s$

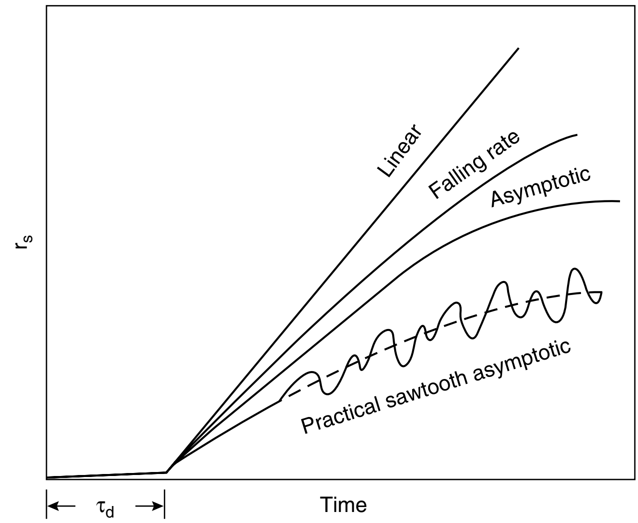


FIGURE 13.19 Characteristic fouling curves.

is generally referred to as the *fouling factor*. Fouling factors for some common fluids are presented in Table 13.12 and Table 13.13.

The specification of fouling effects in a process heat exchanger is usually represented in the following form, wherein the combined fouling factor $r_{s,t}$ is the sum of the fouling factors on the hot and cold fluid sides.

$$\textit{Combined fouling factor } r_{s,t} = \frac{1}{U_F} - \frac{1}{U_C} \tag{13.135}$$

$$\text{Cleanliness factor CF} = \frac{U_F}{U_C} = \frac{1}{1 + r_{s,t} U_C} \tag{13.136}$$

$$\text{Percentage oversurface \%OS} = \left(\frac{A_F}{A_C} - 1\right) 100 \tag{13.137}$$

Here, the subscripts F and C denote fouled and clean exchanger values, respectively. From Equation 13.6 with $A_h = A_c = A$, $\eta_o = 1$, $\Delta T_{m,F} = \Delta T_{m,C}$, it can be shown that

$$\frac{A_F}{A_c} = \frac{U_c}{U_F} = 1 + U_c r_{s,t} \tag{13.138}$$

where $r_{s,t} = r_{s,h} + r_{s,c}$. In heat exchanger design, constant (supposedly and asymptotic) values of $r_{s,h}$ and $r_{s,c}$ are used. Accordingly, extra heat transfer surface area is provided to take into account the deleterious effect of fouling. Thus, the heat exchanger will be "oversized" for the initial clean condition, "correctly sized" for asymptotic fouling (if it occurs in practice), and "undersized" just before the cleaning operation for nonasymptotic fouling.

13.8.4 Influence of Operating and Design Variables

Based on operational experience and research over the last several decades, many variables have been identified that influence fouling significantly. Most important variables are summarized as follows.

TABLE 13.12 Fouling Factors for Various Fluid Streams Used in Heat Exchangers

Fluid	Fouling Factors (m^2 K/W)
Water Type	
Seawater (43°C maximum outlet)	0.000275–0.00035
Brackish water (43°C maximum outlet)	0.00035–0.00053
Treated cooling tower water (49°C maximum outlet)	0.000175–0.00035
Artificial spray pond (49°C maximum outlet)	0.000175–0.00035
Closed loop treated water	0.000175
River water	0.00035–0.00053
Engine jacket water	0.000175
Distilled water or closed cycle condensate	0.00009–0.000175
Treated boiler feedwater	0.00009
Boiler blowdown water	0.00035–0.00053
Liquids	
No. 2 fuel oil	0.00035
No. 6 fuel oil	0.0009
Transformer oil	0.000175
Engine lube oil	0.000175
Refrigerants	0.000175
Hydraulic fluid	0.000175
Industrial organic HT fluids	0.000175–0.00035
Ammonia	0.000175
Ammonia (oil bearing)	0.00053
Methanol solutions	0.00035
Ethanol solutions	0.00035
Ethylene glycol solutions	0.00035
MEA and DEA solutions	0.00035
DEG and TEG solutions	0.00035
Stable side draw and bottom products	0.000175–0.00035
Caustic solutions	0.00035
Gas or Vapor	
Steam (non-oil-bearing)	0.0009
Exhaust steam (oil-bearing)	0.00026–0.00035
Refrigerant (oil-bearing)	0.00035
Compressed air	0.000175
Ammonia	0.000175
Carbon dioxide	0.00035
Coal flue gas	0.00175
Natural gas flue gas	0.00090
Acid gas	0.00035–0.00053
Solvent vapor	0.000175
Stable overhead products	0.000175
Natural Gas & Petroleum Streams	
Natural gas	0.000175–0.00035
Overhead products	0.000175–0.00035
Lean oil	0.00035
Rich oil	0.000175–0.00035
Natural gasoline and liquefied petroleum gases	0.000175–0.00035
Oil Refinery Streams	
Crude and vacuum unit gases and vapors	
Atmospheric tower overhead vapors	0.00017
Light naphthas	0.00017
Vacuum overhead vapors	0.00035

(continued)

TABLE 13.12 ***(Continued)***

Fluid	Fouling Factors (m^2 K/W)
Crude and Vacuum Liquids	
Crude oil	
Gasoline	0.00035
Naphtha and light distillates	0.00035–0.00053
Kerosene	0.00035–0.00053
Light gas oil	0.00035–0.00053
Heavy gas oil	0.00053–0.0009
Heavy fuel oil	0.00053–0.00123
Vacuum tower bottoms	0.00176
Atmospheric tower bottoms	0.00123
Cracking and Coking Unit Streams	
Overhead vapors	0.00035
Light cycle oil	0.00035–0.00053
Heavy cycle oil	0.00053–0.0007
Light coker gas oil	0.00053–0.0007
Heavy coker gas oil	0.00070–0.0009
Bottoms slurry oil (1.5 m/s minimum)	0.00053
Light liquid products	0.00035
Catalytic Reforming, Hydrocracking, and Hydrodesulfurization Streams	
Reformer charge	0.00026
Reformer effluent	0.00026
Hydrocharger charge and effluent	0.00035
Recycle gas	0.000175
Liquid product over 50°C (API)	0.000175
Liquid product 30–50°C (API)	0.00035
Light Ends Processing Streams	
Overhead vapors and gases	0.000175
Liquid products	0.000175
Absorption oils	0.00035–0.00053
Alkylation trace acid streams	0.00035
Reboiler streams	0.00035–0.00053

Source: From Chenoweth, J. M. Final report, HTRI/TEMA joint committee to review the fouling section of TEMA standards. *Heat Transfer Eng*., 11 (1), 73–107, 1988.

13.8.4.1 Flow Velocity

Flow velocity is one of the most important variables affecting fouling. Higher velocities increase fluid shear stress at the fouling deposit–fluid interface and increase the heat transfer coefficient, but at the same time increase pressure drop and fluid pumping power, may erode the surface, and may accelerate the corrosion of the surface by removing the protective oxide layer. The fouling build up in general is inversely proportional to flow velocity u. For water, the velocity should be kept above 2 m/s to suppress fouling, and the absolute minimum should be above 1 m/s to minimize fouling.

13.8.4.2 Surface Temperature

Higher surface temperatures promote chemical reaction, corrosion, crystal formation (with inverse solubility salts), and polymerization, but reduce biofouling for temperatures above the optimum growth, avoid potential freezing fouling, and avoid precipitation of normal solubility salts. It is highly recommended that the surface temperature is maintained below the reaction temperature; it should be kept below 60°C for cooling tower water.

TABLE 13.13 Fouling Factors and Design Parameters for Finned Tubes in Fossil-Fuel Exhaust Gases

Type of Flue Gas	Fouling Factor, m^2 K/W	Minimum Spacing Between Fins, m	Maximum Gas Velocity to Avoid Erosion, m/s
	Clean Gas (cleaning devices not required)		
Natural gas	0.0000881–0.000528	0.00127–0.003	30.5–36.6
Propane	0.000176–0.000528	0.00178	
Butane	0.000176–0.000528	0.00178	
Gas turbine	0.000176		
	Average Gas (provisions for future installation of cleaning devices)		
No. 2 oil	0.000352–0.000704	0.00305–0.00384	25.9–30.5
Gas turbine	0.000264		
Diesel engine	0.000528		
	Dirty Gas (cleaning devices required)		
No. 6 oil	0.000528–0.00123	0.00457–0.00579	18.3–24.4
Crude oil	0.000704–0.00264	0.00508	
Residual oil	0.000881–0.00352	0.00508	
Coal	0.000881–0.00881	0.00587–0.00864	15.2–21.3

Source: From Weierman, R. C. In *Design of Heat Transfer Equipment for Gas-Side Fouling Service, Workshop on an Assessment of Gas-Side Fouling in Fossil Fuel Exhaust Environments*, Marner, W. J., and Webb, R. L., eds., pp. 853–861, 1982. Jet Propulsion Lab., Calif. Inst. of Technology, Pasadena, CA., JPL Publ. 82–67. 1982.

13.8.4.3 Tube Material

The selection of the tube material is important from the corrosion point of view, which in turn could increase crystallization and biological fouling. Copper alloys can reduce certain biofouling, but its use is limited by environmental concerns for river, ocean, and lake waters. Many other variables affect fouling. Discussing them is beyond our scope here, but the reader may refer to TEMA (1999).

13.8.5 Fouling Control and Cleaning Techniques

Control of fouling should be attempted first before any cleaning method is attempted. For gas-side fouling, one should verify that fouling exists, identify the sequential event that dominates the foulant accumulation, and characterize the deposit. For liquid-side fouling, fouling inhibitors/additives should be employed while the exchanger is in operation, such as using antidispersant polymers to prevent sedimentation fouling, "stabilizing" compounds to prevent polymerization and chemical reaction fouling, corrosion inhibitors to prevent corrosion fouling, biocide/germicides to prevent biofouling, and softeners, acids, and polyphosphates to prevent crystallization fouling.

If the foulant control is not effective, the exchanger must be cleaned either on-line or off-line. On-line cleaning includes: flow-driven brushes/sponge balls inside tubes, power-driven rotating brushes inside tubes, acoustic horns/mechanical vibrations for tube banks with gases, sootblowers, and shutting the cold gas supply, flowing hot gas, or reversing the fluids. Off-line cleaning methods without dismantling the exchanger include chemical cleaning (circulate acid/detergent solutions), circulating particulate slurry (such as sand and water), and thermal cleaning to melt frost layers. Off-line cleaning with a heat exchanger opened includes high-pressure stream or water cleaning, thermal baking of an exchanger, and rinsing of small heat exchanger modules removed from the container of modular exchangers.

13.9 Concluding Remarks

Heat exchangers play a critical and dominant role in energy conservation, conversion, recovery, and utilization, and in the economic development of new energy sources as well as in many solutions to environmental problems. Compact heat exchangers have the major advantages of having high heat fluxes, small volumes and packaging, and total lower cost. With the advancement of manufacturing technologies, many innovative and new designs of compact heat exchangers have emerged in recent years that are capable of withstanding either extreme high pressures or high temperatures, but generally cannot handle heavy fouling due to small flow passages that cannot be cleaned mechanically. Hence, compact heat exchangers are used for relatively clean gas flows, refrigerant flows and some liquid flows in many mass production, and niche applications replacing shell-and-tube and other heat exchangers. In this article, a comprehensive review is made of thermal and hydraulic design aspects of single-phase compact heat exchangers. A good number of references are also provided for further study of many types of heat exchangers. Refer to Shah (2006) for comprehensive review on compact heat exchangers that includes CHEs for fuel cell systems and microturbines as well as other topics not covered here.

13.10 Nomenclature

A total heat transfer area (primary + fin) on one fluid side of a heat exchanger—A_p, primary surface area; A_f, fin surface area (m^2)

A_f fin or extended surface area on one fluid side of the exchanger (m^2)

A_{fr} frontal area on one fluid side of an exchanger (m^2)

A_k total wall cross-sectional area for heat conduction in fin or for longitudinal conduction in the exchanger (m^2)

A_0 minimum free flow area on one fluid side of a heat exchanger (m^2)

A_p primary surface area on one fluid side of an exchanger (m^2)

A_w total wall area for heat conduction from the hot fluid to the cold fluid, or total wall area for transverse heat conduction (in the matrix wall thickness direction) (m^2)

Bi Biot number, $\mathrm{Bi} = h(\delta_w/2)/k_w$ for the regenerator analysis, dimensionless

b plate spacing, $h' + \delta_{f'}$ (m)

C flow stream heat capacity rate with a subscript c or h, $\dot{m}c_p$ (W/°C)

C^* heat capacity rate ratio, C_{min}/C_{max}, dimensionless

C_r heat capacity rate of a regenerator, $M_w c_w N$ (W/°C)

C_r^* total matrix heat capacity rate ratio, C_r/C_{min}, $C_{r,h}^* = C_{r,h}/C_h$, $= C_{r,c}/C_c$, dimensionless

c_p specific heat of fluid at constant pressure (J/kg K)

c_w specific heat of wall material (J/kg K)

D_h hydraulic diameter of flow passages, $4A_0L/A$ (m)

d_e fin tip diameter of an individually finned tube (m)

d_i, d_o tube inside and outside diameters, respectively (m)

Eu N-row average Euler number, $\Delta p/(\rho u_m^2\ N/2g_c)$, $\Delta p/(G^2N/2g_c\rho)$, dimensionless

F log-mean temperature difference correction factor, dimensionless

f Fanning friction factor, $\rho\Delta p g_c D_h/(2LG^2)$, dimensionless

f_{tb} average Fanning friction factor per tube row for crossflow over a tubebank outside, $\Delta p/(4G^2N/2g_c\rho)$, Eu/4, dimensionless

G mass velocity based on the minimum free flow area, $\dot{m}/A_0$ (kg/m^2 s)

g gravitational acceleration (m^2/s)

g_c proportionality constant in Newton's Second Law of Motion, $g_c = 1$ and dimensionless in SI units, $g_c = 32.174$ lbm ft./lbf s^2

Hg Hagen number, defined by Equation 13.56, dimensionless

H thermal boundary condition referring to constant axial as well as peripheral wall heat flux, also constant peripheral wall temperature; boundary condition valid only for the circular tube, parallel plates, and concentric annular ducts when symmetrically heated

H1 thermal boundary condition referring to constant axial wall heat flux with constant peripheral wall temperature

H2 thermal boundary condition referring to constant axial wall heat flux with constant peripheral wall heat flux
h heat transfer coefficient (W/m^2 K)
h' height of the offset strip fin (see Figure 13.9) (m)
h_{fg} specific enthalpy of phase change (see Table 13.9) (J/kg)
j Colburn factor, $NuPr^{-1/3}/Re$, $StPr^{2/3}$, dimensionless
K_c contraction loss coefficient for flow at heat exchanger entrance, dimensionless
K_e expansion loss coefficient for flow at heat exchanger exit, dimensionless
k fluid thermal conductivity (W/m K)
k_f thermal conductivity of the fin material (W/m K)
k_w thermal conductivity of the matrix (wall) material (W/m K)
L fluid flow (core or tube) length on one fluid side of an exchanger (m)
ℓ fin length for heat conduction from primary surface to the midpoint between plates for symmetric heating, see Table 13.4 for other definitions of ℓ (m)
ℓ_f offset strip fin length or fin height for individually finned tubes, ℓ_f represents the fin length in the fluid flow direction for an uninterrupted fin with $\ell_f = L$ in most cases (m)
M_w mass of the regenerator (kg)
m fin parameter, $[2h(1+\delta_i/\ell_i)/k_f\delta_f]^{1/2} \approx [2h/k_f\delta_f]^{1/2}$ (1/m)
$\dot{m}$ fluid mass flow rate, $\rho u_m A_o$ (kg/s, 1 bm/h)
N number of tube rows
N regenerator rotational speed (rev/s)
N_f number of fins per meter (1/m)
N_r number of tube rows in the flow direction
N_t total number of tubes in an exchanger
NTU number of heat transfer units, UA/C_{min}, it represents the total number of transfer units in a multipass unit, $NTU_s = UA/C_{shell}$, dimensionless
Nu Nusselt number, hD_h/k, dimensionless
ntu_c number of heat transfer units based on the cold fluid side, $(\eta_0 hA)_c/C_c$, dimensionless
ntu_h number of heat transfer units based on the hot fluid side, $(\eta_0 hA)_h/C_h$, dimensionless
$\dot{m}$ mass flow rate (kg/s)
$\tilde{P}$ temperature effectiveness of one fluid, dimensionless
P wetted perimeter of exchanger passages on one fluid side, $\tilde{P} = A/L$ (m)
$\mathcal{P}$ fluid pumping power (W)
Pr fluid Prandtl number, $\mu c_p/k$, dimensionless
p fluid static pressure (Pa)
$\triangle p$ fluid static pressure drop on one fluid side of heat exchanger core (Pa)
p_f fin pitch (m)
q exchanger heat transfer rate or heat duty (W)
q_e heat transfer rate (leakage) at the fin tip (W)
q' heat flux, q/A (W/m^2)
R heat capacity rate ratio used in the P-NTU method, $R_1 = C_1/C_2$, $R_2 = C_2/C_1$, dimensionless
R thermal resistance based on the surface area A, compare Equation 13.5 and Equation 13.6 for definitions of specific thermal resistances, generally with the subscript noted in Eq. (13.5) (K/W)
Re Reynolds number, GD_h/μ, dimensionless
Re_d Reynolds number, $\rho u_m d_o/\mu$, dimensionless
r_h hydraulic radius, $D_h/4$, A_0L/A (m)
r_s fouling factor, l/h_s (m^2 K/W)
St Stanton number, h/Gc_p, dimensionless
s distance between adjacent fins, $p_f - \delta_f$ (m)
T fluid static temperature to a specified arbitrary datum (°C)
T_s ambient temperature (°C)

T_o fin base temperature (°C)
T_λ fin tip temperature (°C)
U overall heat transfer coefficient (W/m^2 K)
u_m fluid mean axial velocity in the exchanger minimum free flow area on one fluid side (m/s)
V heat exchanger total volume (m^3)
X Martinelli parameter defined by Equation 13.89, dimensionless
X_d diagonal tube pitch (m)
X_ℓ longitudinal tube pitch (m)
X_t transverse tube pitch (m)
X_l^* a ratio of the longitudinal pitch to the tube outside diameter in a circular tube bank, X_l/d_o, dimensionless
X_t^* a ratio of the transverse pitch to the tube diameter in a circular tube bank, X_t/d_o, dimensionless
x mass quality (a ratio of mass flow rate of the vapor (or gas) phase divided by the total mass flow rate of two-phase mixture, dimensionless
Y Chisholm parameter, defined by Equation 13.89, dimensionless
z axial coordinate in Section 13.5 (m)
α void fraction of the vapor phase in Section 13.5, dimensionless
α ratio of total heat transfer area on one fluid side of an exchanger to the total volume of an exchanger, A/V (m^2/m^3)
β heat transfer surface area density, a ratio of total transfer area on one fluid side of a plate-fin heat exchanger to the volume between the plates on that fluid side (m^2/m^3)
ε heat exchanger effectiveness, represents an overall exchanger effectiveness for a multipass unit, dimensionless
δ wall thickness (m)
δ_f fin thickness (m)
η_f fin efficiency, dimensionless
η_o extended surface efficiency, dimensionless
η_p pump/fan efficiency, dimensionless
λ longitudinal wall heat conduction parameter based on the total conduction area, $\lambda = k_w A_{k,t}/C_{min} L$, $\lambda_c = k_w A_{k,c}/C_c L$, $\lambda_h = k_w A_{k,h}/C_h L$, dimensionless
μ fluid dynamic viscosity (Pa s)
ρ fluid density (kg/m^3)
σ ratio of free flow area within the core to frontal area A_0/A_{fr}, dimensionless
θ, ϕ angular coordinate shown in Figure 13.7 (rad)

13.10.1 Subscripts

C clean surface value
c cold fluid side
con condensation
F fouled surface value
f fin
fr friction
g gas
go gas only
h hot fluid side
i inlet to the exchanger
l liquid
lam laminar
lo liquid only

loc local
o outlet to the exchanger or overall
s scale or fouling, or staggered tubebank
sat saturated
tp two phase
turb turbulent
w wall or properties at the wall temperature
1 one section (inlet or outlet) of the exchanger
2 other section (outlet or inlet) of the exchanger

References

Bačlić, B. S., Romie, P. E., and Herman, C. Y. 1988. The galerkin method for two-pass crossflow heat exchanger problem. *Chem. Eng. Comm.*, 70 177–198.

Bajura, R. A. and Jones, E. H. 1976. Flow distribution manifolds. *J. Fluid Eng.-Trans. ASME*, 98 654–666.

Bhatti, M. S. and Shah, R. K. 1987. Turbulent and transition flow convective heat transfer in ducts. In *Handbook of Single-Phase Convective Heat Transfer*, S. Kakaç, R.K. Shah, and W. Aung, eds., pp. 95–101. Wiley, New York. Chap. 4.

Bott, T. R. 1995. *Fouling of Heat Exchangers*. Elsevier Science Publishers Ltd., Amsterdam, The Netherlands.

Bott, T. R., Melo, L. F., Panchal, C. B., and Somerscales, E. F. C. 1999. *Understanding Heat Exchanger Fouling and Its Mitigation*. Begell House, Inc., New York.

Carey, V. P. 1992. *Liquid-Vapor Phase Change Phenomena*. Taylor & Francis, Bristol, PA.

Chai, H. C. 1988. A simple pressure drop correlation equation for low finned tube crossflow heat exchangers. *Int. Commun. Heat Mass Transfer*, 15 95–101.

Chang, Y. J., and Wang, C. C. 1997. A generalized heat transfer correlation for louver fin geometry. *Int. J. heat and Mass Transfer*, Vol. 40, 533–544.

Chang, Y. J., Hsu, K. C., Lin, Y. T., and Wang, C. C. 2000. A generalized friction correlation for louver fin geometry. *Int. J. Heat Mass Transfer*, Vol. 43, 2237–2243.

Chenoweth, J. M. 1988. Final report, HTRI/TEMA joint committee to review the fouling section of TEMA standards. *Heat Transfer Eng.*, 11 (1), 73–107.

Chiou, J. P. 1982. The effect of nonuniformities of inlet temperatures of both fluids on the thermal performance of crossflow heat exchanger. *Heat Transfer 1982. 7th International Heat Transfer Conference*, pp. 179–184.

Chowdhury, K. and Sarangi, S. 1985. The effect of flow maldistribution on multipassage heat exchanger performance. *Heat Transfer Eng.*, 6 (4), 45–54.

Churchill, S. W. 1977. Friction-factor equation spans all fluid flow regimes. *Chem. Eng.*, 84 (24), 91–92.

Cichelli, M. T. and Boucher, D. E. 1956. Design of heat exchanger heads for low holdup, *AIChE Chem. Eng. Prog.*, 52 (5), 213–218.

Collier, J. G. and Thome, J. R. 1994. *Convective Boiling and Condensation*. *3rd Ed.* McGraw-Hill, New York.

Cowell, T. A., Heikal, M. R., and Achaichia, A. 1995. Flow and heat transfer in compact louvered fin surfaces. *Exp. Therm. Fluid Sci.*, 10 192–199.

Epstein, N. 1978. Fouling in heat exchangers. In *Heat Transfer 1978*, *Vol. 6*, pp. 235–254. Hemisphere, New York.

Epstein, N. 1983. Thinking about heat transfer fouling: a 5 x 5 matrix. *Heat Transfer Eng.*, 4 (1), 43–56.

Fleming, R. B. 1966. The effect of flow distribution in parallel channels of counterflow heat exchangers. *Adv. Cryogenic Eng.* 352–362.

Friedel, L. 1979. Improved friction pressure drop correlations for horizontal and vertical two-phase pipe flow. European Two-Phase Flow Group Meeting, Ispra, Italy, Paper E2.

Gaddis, E. S. and Gnielinski, V. 1985. Pressure drop in cross flow across the tube bundles. *Int. Chem. Eng.*, 25 1–15.

Ganguli, A. and Yilmaz, S. B. 1987. New heat transfer and pressure drop correlations for crossflow over low-finned tube banks. *AIChE Symp. Ser.*, 83 9–14.

Ghajar, A. J. and Tam, L. M. 1994. Heat transfer measurements and correlations in the transition region for a circular tube with three different inlet configurations. *Exp. Therm. Fluid Sci.*, 8 79–90.

Hesselgreaves, J. G. 2001, *Compact Heat Exchangers-Selection Design, and Operation*, Elsevier Science, Ltd., Oxford, UK.

Hewitt, G. E. 1989. *Hemisphere Handbook of Heat Exchanger Design*. *Hemisphere Handbook of Heat Exchanger Design*. Hemisphere, New York.

Hewitt, G. and Müller-Steinhagen, H., 2000. *Heat Exchanger Fouling in the Pre-heat Train of a Crude Oil Distillation Units*. ESDU Data Item 00016, International Ltd., London, U.K.

Hewitt, G. and Müller-Steinhagen, H. 2003. *Fouling in Cooling Water Systems Using Seawater*. ESDU Data Item 03004, International Ltd., London, U.K.

Hewitt, G. E., Shires, G. L., and Bott, T. R. 1994. *Process Heat Transfer*. CRC Press, Boca Raton, Florida.

Huang, L. J. and Shah, R. K. 1992. Assessment of calculation methods for efficiency of straight fins of rectangular profiles. *Int. J. Heat Fluid Fl.*, 13 282–293.

Idelchik, I. E. 1994. *Handbook of Hydraulic Resistance*. *3rd Ed.* CRC Press, Boca Raton, FL.

Kakaç, S., and Liu, H., 1998, *Heat Exchangers: Selection, Rating and Thermal Design*, CRC Press Boca Raton, FL.

Kandlikar, S. G. 1991. Development of a flow boiling map for subcooled and saturated flow boiling of different fluids in circular tubes. *ASME J. Heat Transfer*, 113 190–200.

Kandlikar, S. G. and Steinke, M. E. 2003. Predicting heat transfer during flow boiling in minichannels and microchannels. *ASHRAE Trans.*, 109 (1), 1–9.

Kandlikar, S. G. and Balasubramanian, P. 2004. An extension of the flow boiling correlation to transition, laminar and deep laminar flows in minichannels and microchannels. *Heat Transfer Eng.*, 25 (3), 86–93.

Kandlikar, S. G., Shoji, M., and Dhir, V. K. eds. 1999. *Handbook of Phase Change: Boiling and Condensation*, pp. 395–401. Taylor & Francis, New York.

Kattan, N., Thome, J. R., and Favrat, D. 1998. Flow boiling in horizontal tubes, Part I; Development of diabatic two-phase flow pattern map. *ASME J. Heat Transfer*, 120, 140–147; Part 2; New heat transfer data for five refrigerants. *ASME J. Heat Transfer*, 120, 148–155; Part 3; Development of a new heat transfer model based on flow patterns. *ASME J. Heat Transfer*, 120, 156–165.

Kays, W. M. and London, A. L. 1998. *Compact Heat Exchangers*. *Reprint 3rd Ed.* Krieger Publishing, Malabar, Fl.

Kohler, M. 1974. *The Influence of Flow Path Geometry and Manufacturing Tolerances on Gas Turbine Regenerator Efficiency*, SAE Paper No. 740183.

Kundu, B. and Das, P. K. 2000. Performance of symmetric polygonal fins with and without tip loss—a comparison of different methods of prediction. *Can. J. Chem. Eng.*, 78 395–401.

Kutchey, J. A. and Julien, H. L. 1974. The measured influence of flow distribution on regenerator performance, *SAE Trans.*, 83, SAE Paper No. 740164.

London, A. L. 1970. Laminar flow gas turbine regenerators—the influence of manufacturing tolerances. *ASME J. Eng. Power*, 92A 45–56.

London, A. L. and Shah, R. K. 1973. Glass-ceramic hexagonal and circular passage surfaces—heat transfer and flow friction design characteristics. *SAE Trans.*, 82 (Sec. 1), 425–434.

London, A. L., Young, M. B. O., and Stang, J. H. 1970. Glass ceramic surfaces, straight triangular passages—heat transfer and flow friction characteristics. *ASME J. Eng. Power*, 92A 381–389.

Majumdar, A. K. 1980. Mathematical modeling of flows in dividing and combining flow manifold. *Appl. Math. Model.*, 4 424–432.

Manglik, R. M. and Bergles, A. E. 1995. Heat transfer and pressure drop correlations for the rectangular offset-strip-fin compact heat exchanger. *Exp. Therm. Fluid Sci.*, 10 171–180.

Martin, H. 2002. The generalized Lévêque equation and its practical use for the prediction of heat and mass transfer rates from pressure drop. *Chem. Eng. Sci.*, 57 3217–3223.

Miller, D. S. 1990. *Internal Flow Systems. 2nd Ed. BHRA Fluids Engineering Series., Vol. 5*, British Hydromechanics Research Association, Cranfield, United Kingdom.

Mueller, A. C. 1974. Criteria for maldistribution in viscous flow coolers. *Heat Transfer 1974*, 5 170–174.

Mueller, A. C. 1977. An inquiry of selected topics on heat exchanger design. *AIChE Symp. Ser. 164*, 73 273–287.

Mueller, A. C. and Chiou, J. P. 1987. Review of various types of flow maldistribution in heat exchangers. Book No. H00394, HTD-*Vol. 75*, pp. 3–16. ASME, New York.

Müller-Steinhagen, H. ed. 2000. *Heat Exchanger Fouling—Mitigation and Cleaning Technologies*, pp. 935–947. Publico Publications, Essen, Germany.

Ould Didi, M. B., Kattan, N., and Thome, J. R. 2002. Predication of two-phase pressure gradients of refrigerants in horizontal tubes. *Int. J. Refrig.*, 25 935–947.

Palen, J. W. ed. 1987. *Heat Exchanger Sourcebook*, pp. 1031–1038. Hemisphere Publishing Corp., Washington, DC.

Panchal, C. B., Bott, T. R., Somerscales, E. F. C., and Toyama, S. 1997. *Fouling Mitigation of Industrial Heat Exchange Equipment*. Begell House, New York.

Putnam, G. R. and Rohsenow, W. M. 1985. Viscosity induced nonuniform flow in laminar flow heat exchangers. *Int. J. Heat Mass Tran.*, 28 1031–1038.

Rabas, T. J. and Taborek, J. 1987. Survey of turbulent forced-convection heat transfer and pressure drop characteristics of low-finned tube banks in cross flow. *Heat Transfer Eng.*, 8 (2), 49–62.

Razelos, P. 1980. *Personal Communication, Department of Applied Science*. City University of New York, Staten Island, NY.

Reay, D. A. 1979. *Heat Recovery Systems*. E. & EN. Spon, London; UK.

Roetzel, W. and Spang, B. 1987. Analytisches verfahren zur thermisen berechnung mehrgängiger rohrbünädelwreübertrager. Fortschr-Ber. VDI, Reihe. 19, Nr. 18.

Roetzel, W. and Spang, B. 1989. Thermal calculation of multipass shell and tube heat exchangers. *Chem. Eng. Res. Des.*, 67 115–120.

Rozenman, T. 1976. Heat transfer and pressure drop characteristics of dry cooling tower extended surfaces. Part I: Heat Transfer and Pressure Drop Data, Report BNWL-PFR 7–100; Part II: Data Analysis and Correlation, Report BNWL-PFR 7–102, Battelle Pacific Northwest Laboratories, Richland, WA.

Saunders, E. A. D. 1988. *Heat Exchangers: Selection, Design* & *Construction*. Longman Scientific & Technical, Essex, UK.

Shah, M. M. 1977. A general correlation for heat transfer during subcooled boiling in pipes and annuli. *ASHRAE Trans.*, 83 (1), 205–215; also, Shah, M. M. 1982. Chart correlation for saturated boiling heat transfer: equations and further study. *ASHRAE Trans.*, 88 (1), 185–196.

Shah, M. M. 1979. A general correlation for heat energy during film considensation inside pipe, *Int. J. Heat Mass Transfer*, Vol. 22, pp. 547–556.

Shah, R. K. 1981. Compact heat exchangers. In *Heat Exchangers: Thermal-Hydraulic Fundamentals and Design*, Kakaç, S., Bergles, A. E., and Mayinger, E., eds., pp. 111–151. Hemisphere Publishing Corp., Washington, DC.

Shah, R. K. 1983. Heat exchanger basic design methods. In *Low Reynolds Number Flow Heat Exchangers*, Kakaç, S., Shah R. K., and Bergles, A. E., eds., pp. 21–72. Hemisphere Publishing Corp., Washington, DC.

Shah, R. K. 1985. Compact heat exchangers. In *Handbook of Heat Transfer Applications*, *2nd Ed.*, Rohsenow, W. M., Hartnett, J. P., and Ganić, E. N., eds., Chap. 4, Part 3, pp. 255–266. McGraw-Hill, New York.

Shah, R. K. 1988. Plate-fin and tube-fin heat exchanger design procedures. In *Heat Transfer Equipment Design*, Shah, R. K., Subbarao, E. C., and MasheIkar, R. A., eds., pp. 255–266. Hemisphere Publishing Corp., Washington, DC.

Shah, R. K. 1988. Counterflow rotary regenerator thermal design procedures. In *Heat Transfer Equipment Design*, Shah, R. K., Subbarao, E. C., and MasheIkar R. A., eds., pp. 267–296. Hemisphere Publishing Corp., Washington, DC.

Shah, R. K. 1991a. Industrial heat exchangers—functions and types. In *Industrial Heat Exchangers*, Buchlin, J-M., ed., von Kármán Institute for Fluid Dynamics, Belgium. Lecture Series No. 1991-04.

Shah, R. K. 1991b. Multidisciplinary approach to heat exchanger design. In *Industrial Heat Exchangers*, Buchlin, J-M., ed., von Kármán Institute for Fluid Dynamics, Belgium. Lecture Series No. 1991-04.

Shah, R. K. 1993. Nonuniform heat transfer coefficients for heat exchanger thermal design. In *Aerospace Heat Exchanger Technology*, Shah, R. K. and Hashemi, A., eds., pp. 417–445. Elsevier Science, Amsterdam, Netherlands.

Shah, R. K. and Bhatti, M. S. 1987. Laminar convective heat transfer in ducts. In *Handbook of Single-Phase Convective Heat Transfer*, Kakaç, S., Shah, R. K., and Aung, W., eds., pp. 81–122. Wiley, New York,. Chap. 3.

Shah, R. K. and Bhatti, M. S. 1988. Assessment of correlations for single-phase heat exchangers. In *Two-Phase Flow Heat Exchangers: Thermal-Hydraulic Fundamentals and Design*, Kakaç, S., Bergles, A. E., and Fernandes, E. O., eds., pp. 81–122. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Shah, R. K. and Giovannelli, A. D. 1988. Heat pipe heat exchanger design theory. In *Heat Transfer Equipment Design*, Shah, R. K., Subbarao, E. C., and Mashelkar, R. A., eds., pp. 609–653. Hemisphere Publishing Corp., Washington, DC.

Shah, R. K. and London, A. L. 1978. *Laminar Flow Forced Convection in Ducts*. Supplement 1 to Advances in Heat Transfer. Academic Press, New York.

Shah, R. K. and London, A. L. 1980. Effects of nonuniform passages on compact heat exchanger performance. *ASME J. Eng. Power*, 102A, 653–659.

Shah, R. K. and Mueller, A. C. 1985. Heat exchanger basic thermal design methods. In *Handbook of Heat Transfer and Applications*, *2nd Ed.*, Rohsenow, W. M. Hartnett, J. P., and Ganić, E. N., eds., pp. 4-1–4-77. McGraw-Hill, New York. Chap. 4, Part I.

Shah, R. K. and Muller, A. C. 1988. Heat exchange. In *Ullmann's Encyclopedia of Industrial Chemistry, Unit Operations II*, *Vol. B3*. VCH Publishers, Weinheim, Germany, Chap. 2.

Shah, R. K. and Pignotti, A. 1997. The influence of a finite number of baffles on the shell-and-tube heat exchanger performance. *Heat Transfer Eng.*, 18 (1), 82–94.

Shah, R. K., and Sekulic, D. P. 1998. Heat Exchangers, In *Handbook of Heat Transfer, 3rd Ed.*, Rohsenow, W. M., Hartnett, J. P., and Cho, Y. I., eds., Chap. 17, 169 pages, McGraw-Hill, New York.

Shah, R. K. and Sekulić, D. P. 2003. *Fundamentals of Heat Exchanger Design*. Wiley, Hoboken, NJ.

Shah, R. K. and Wanniarachchi, A. S. 1991. Plate heat exchanger design theory. In *Industrial Heat Exchangers*, J. M. Buchlin, ed., pp. 179–216. von Kármán Institute for Fluid Dynamics, Belgium,. Lecture Series No. 1991-04.

Shah, R. K., Zhou, S. Q., and Tagavi, K. 1999. The role of surface tension in film condensation in extended surface passages. *J. Enhanc. Heat Transf.*, 6 179–216.

Shah, R. K., 2006, Advances in science and Technology of compact heat exchangers, *Heat Transfer Engineering*, Vol. 27, No. 5, pp. 3–22.

Shen, P. I. 1992. The effect of friction on flow distribution in dividing and combining flow manifolds. *ASME J. Fluids Eng.*, 114 121–123.

Singh, K. P. and Soler, A. I. 1984. *Mechanical Design of Heat Exchangers and Pressure Vessel Components*. Arcturus Publishers, Cherry Hill, NJ.

Spang, B., Xuan, Y., and Roetzel, W. 1991. Thermal performance of split-flow heat exchangers. *Int. J. Heat Mass Trans.*, 34 863–874.

Srinivasan, V. and Shah, R. K. 1995. Fin Efficiency of Extended Surfaces in Two-Phase Flow, Celata, G. P., and Shah, R. K., eds., In *Two Phase Flow Modeling and Experimentation 1995*, Vol. 2, pp. 1107–1118. Edizioni ETS, Pisa, ltaly.

TEMA 1999. *Standards of the Tubular Exchanger Manufacturers Association*. *8th Ed.* Tubular Exchanger Manufacturers Association, New York.

Walker, G. 1990. *Industrial Heat Exchangers-A Basic Guide*. *2nd Ed.* Hemisphere, Washington, DC.

Wang, C. C. 2000. Recent progress on the air-side performance of fin-and-tube heat exchangers. *Int. J. Heat Exchangers*, 1 49–76.

Wang, C. C. and Chi, K. U. 2000. Heat transfer and friction characteristics of plain fin-and-tube heat exchangers: part 2: correlation. *Int. J. Heat Mass Tran.*, 43 2692–2700.

Webb, R. L. 1994. *Principles of Enhanced Heat Transfer*. Wiley, New York.

Weierman, R. C. 1982. In *Design of Heat Transfer Equipment for Gas-Side Fouling Service, Workshop on an Assessment of Gas-side Fouling in Fossil Fuel Exhaust Environments*, Marner, W. J., and Webb, R. L., eds., pp. 853–861. Jet Propulsion Lab., Calif. Inst. of Technology, Pasadena, CA,. JPL Publ. 82-67.

Xuan, Y., Spang, B., and Roetzel, W. 1991. Thermal analysis of shell and tube exchangers with divided-flow pattern. *Int. J. Heat Mass Trans.*, 34 853–861.

Yokell, S. 1990. *A Working Guide to Shell-and-Tube Heat Exchangers*. McGraw-Hill, New York.

Zrcher, O., Thome, J. R., and Favrat, D. 1999. Evaporation of ammonia in a smooth horizontal tube: heat transfer measurements and predictions. *ASME J. Heat Transfer*, 121 89–101.

Zukauskas, A. 1987. Convective heat transfer in cross flow. In *Handbook of Single-Phase Convective Heat Transfer*, Kakaç, S., Shah, R. K., and Aung, W., eds., Wiley, New York, Chap. 6.

14

Industrial Energy Efficiency and Energy Management

Craig B. Smith
Dockside Consultants, Inc.

Barney L. Capehart
University of Florida

Wesley M. Rohrer Jr.✠

14.1 Introduction

The industrial sector in the United States is highly diverse—consisting of manufacturing, mining, agriculture and construction activities—and consumes one-third of the nation's primary energy use, at an annual cost of around $100 billion.[1] The industrial sector encompasses more than 3 million

✠Deceased

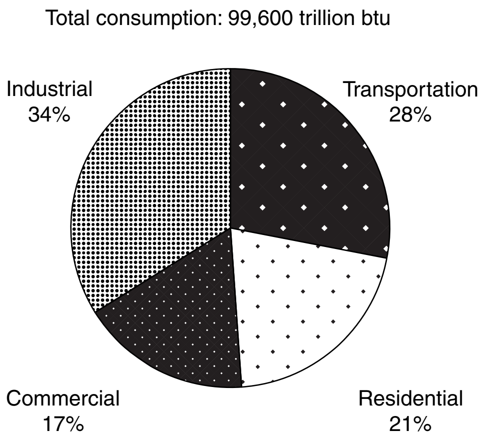


FIGURE 14.1 U.S. energy use by sector, 2004. (From U.S. DOE Energy Information Agency, *Industrial Sector Energy Price and Expenditure Estimates for 2001*, U.S. Department of Energy, Office of Industrial Technologies, Industrial Technology Program, Washington, DC, 2005.)

establishments engaged in manufacturing, agriculture, forestry, construction, and mining. These industries require energy to light, heat, cool, and ventilate facilities (end uses characterized as energy needed for comfort). They also use energy to harvest crops, process livestock, drill and extract minerals, power various manufacturing processes, move equipment and material, raise steam, and generate electricity. Some industries require additional energy fuels for use as raw materials—or feedstocks—in their production processes. Many industries use by-product fuels to satisfy part or most of their energy requirements. In the more energy-intensive manufacturing and nonmanufacturing industries, energy used by processes dwarfs the energy demand for comfort.

U.S. sector energy use for 2004 is shown in Figure 14.1, industrial energy use is shown in Figure 14.2, and industrial electricity use is shown in Figure 14.3.

Manufacturing companies, which use mechanical or chemical processes to transform materials or substances to new products, account for about 80% of the total industrial sector use. The "big three" in energy use are petroleum, chemicals, and primary metals; these manufacturers together consume over one-half of all industrial energy. The "big six," which adds the pulp and paper group, the food and kindred products group, as well as the stone, clay and glass group, together account for 88% of manufacturing energy use, and over 70% of all industrial sector energy consumption.[2]

According to the U.S. Energy Information Administration, energy efficiency in the manufacturing sector improved by 25% over the period 1980–1985.[3] During that time, manufacturing energy use declined 19%, and output increased 8%. These changes resulted in an overall improvement in energy efficiency of 25%. However, the "big five" did not match this overall improvement; although their energy use declined 2%, their output decreased by 5%, resulting in only a 17% improvement in energy efficiency during 1980–1985. This five-year record of improvement in energy efficiency of the manufacturing sector

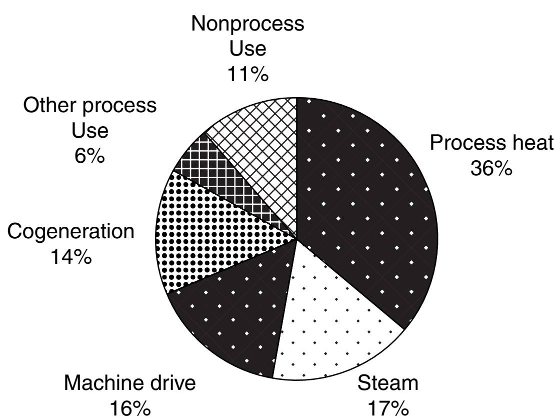


FIGURE 14.2 Manufacturing energy use 1998 (end use basis). (From U.S. Department of Energy EIA MECS, 2002.)

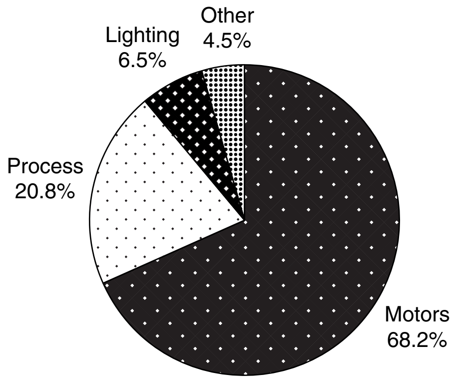


FIGURE 14.3 Manufacturing electrical energy use 1998 (end use basis). (From Manufacturer Energy Consumption Survey, 1998; U.S. Depatment of Energy; Energy Information Agency, 2002.)

came to an end, with total energy use in the sector growing by 10% from 1986 to 1988, and overall manufacturing energy intensity stagnated during 1985–1994 due to falling and low energy prices, economic recession during part of this period, and a recovery in some of the more energy-intensive groups such as steel and aluminum production. However, industrial energy intensity again declined significantly during the late 1990s due to high capital investment, rapid industrial modernization, and explosive growth of "high-tech" industries that are not energy-intensive.

Since 1980, the overall value of industrial output has increased through 2003, while the total energy consumed by the industrial sector has fallen overall throughout 2003.[4] This relationship is shown in Figure 14.4, where the consumption index for both primary and site energy is greater than the output index before 1980, and less afterward, with the gap consistently widening in the late 1980s. New energy-efficient technology and the changing production mix from the manufacture of energy-intensive products to less intensive products are responsible for this change.

Continuing this overall record of energy-efficiency improvements in industry will require emphasis on energy-management activities, as well as making capital investments in new plant processes and facilities improvements. Reducing the energy costs per unit of manufactured product is one way that the United States can become more competitive in the global industrial market. The U.S. Department of Energy has

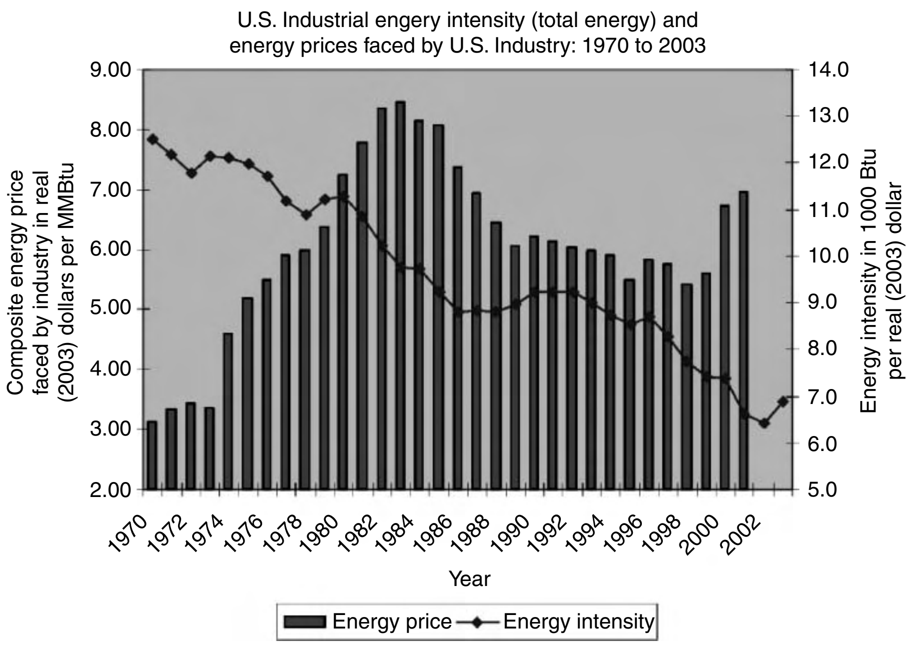


FIGURE 14.4 Industrial energy intensity 1980–2003. (From National Association of Manufacturers, 2005. With permission.)

formally recognized these multiple benefits to the country by including the following statement in its 2004 Industrial Technology Program Report: "By developing and adopting more energy-efficiency technologies, U.S. industry can boost its productivity and competitiveness while strengthening national energy security, improving the environment, and reducing emissions linked to global climate change."[5] Additionally, it is interesting to note that Japan—one of the U.S.' major industrial competitors—has a law that says every industrial plant must have a full-time energy manager.[6]

Several studies of industrial energy efficiency have been performed in the last few years, and the results from the studies have been fairly consistent—that there is a readily achievable, cost-effective, 20% reduction in industrial consumption using good energy management practices and energy-efficient equipment. The most recent, highly credible study that has been done of the potential for industrial energy-efficiency improvement in the U.S. is the "Energy Use, Loss and Opportunities Analysis: U.S. Manufacturing and Mining" report issued by the U.S. Department of Energy, Office of Industrial Technologies, in December 2004.[1] This report, together with the together with the report on "Efficiency and innovation in U.S. Manufacturing Energy Use" from the National Association of Manufacturers (NAM), conducted with the Alliance to Save Energy, make a strong and very credible case for the achievement of a 20% savings in energy for the industrial sector.[4]

Earlier studies by the national laboratories in the U.S. produced a composite report, "Scenarios for a Clean Energy Future," in 2000 that also supported the achievement of a 20% reduction in industrial energy use. One element of this study was to examine data from the Industrial Assessment Center Program (discussed in more detail later in this chapter), from 12,000 plant energy audits making 82,000 recommendations for actions to increase energy efficiency in the facilities audited.[7] Results from these real-world energy audits also supported the 20% reduction estimate.

Other groups, such as Lawrence Berkeley Laboratory, have performed studies of industrial energy efficiency, and their results from their studies show improvements of 20% or greater in industrial energy efficiency.[8] A recent study from the American Council for an Energy Efficient Economy (ACEEE) showed the potential for a 40% reduction in electricity use for fan and pumping applications in industry.[9] Specific details and recommendations from these studies are presented in Section 14.5.

14.2 Industrial Energy Management and Efficiency Improvement

14.2.1 Setting Up an Energy-Management Program

The effectiveness of energy utilization varies with specific industrial operations because of the diversity of the products and the processes required to manufacture them. The organization of personnel and operations, involved also varies. Consequently, an effective energy management program should be tailored for each company and its plant operations. There are some generalized guidelines, however, for initiating and implementing an energy management program. Many large companies have already instituted energy management programs and have realized substantial savings in fuel and electric costs. Smaller industries and plants, however, often lack the technical personnel and equipment to institute and carry out effective programs. In these situations, reliance on external consultants may be appropriate to initiate the program. Internal participation, however, is essential for success. A well-planned, organized, and executed energy management program requires a strong commitment by top management.

Assistance also can be obtained from local utilities. Utility participation would include help in getting the customer started on an energy management program, technical guidance, or making information available. Most electric and gas utilities today have active programs that include training of customer personnel or provision of technical assistance. Table 14.1 summarizes the elements of an effective energy management program. These will now be discussed in more detail.

TABLE 14.1 Elements of an Energy Management Program

	Phase 1: Management Commitment
1.1	Commitment by management to an energy management program
1.2	Assignment of an energy management coordinator
1.3	Creation of an energy management committee of major plant and department representatives
	Phase 2: Audit and Analysis
2.1	Review of historical patterns of fuel and energy use
2.2	Facility walk-through survey
2.3	Preliminary analyses, review of drawings, data sheets, equipment specifications
2.4	Development of energy audit plans
2.5	Conduct facility energy audit, covering
	a. Processes
	b. Facilities and equipment
2.6	Calculation of annual energy use based on audit results
2.7	Comparison with historical records
2.8	Analysis and simulation step (engineering calculations, heat and mass balances, theoretical efficiency calculations, computer analysis and simulation) to evaluate energy management options
2.9	Economic analysis of selected energy management options (life cycle costs, rate of return, benefit–cost ratio)
	Phase 3: Implementation
3.1	Establish energy effectiveness goals for the organization and individual plants
3.2	Determine capital investment requirements and priorities
3.3	Establish measurement and reporting procedures, install monitoring and recording instruments and submeters as required.
3.4	Institute routine reporting procedures ("energy-tracking" charts) for managers and publicize results
3.5	Promote continuing awareness and involvement of personnel
3.6	Provide for periodic review and evaluation of overall energy management program

14.2.1.1 Phase I: Management Commitment

A commitment by the directors of a company to initiate and support a program is essential. An energy coordinator is designated and an energy management committee is formed. The committee should include personnel representing major company activities utilizing energy. A plan is formulated to set up the programs with a commitment of funds and personnel. Realistic overall goals and guidelines in energy savings should be established based on overall information in the company records, projected activities, and future fuel costs and supply. A formal organization as described earlier is not an absolute requirement for the program; smaller companies will simply give the energy management coordination task to a staff member.

Organizing for Energy Conservation Programs. The most important organizational step which will effect the success of an energy management (e.m.) program is the appointment of one person who has full responsibility for its operation. Preferably that person should report directly to the top management position and be given substantial authority in directing technical and financial resources within the bounds set by the level of management commitment. It is difficult to stress enough the importance of making the position of plant energy manager a full-time job. Any diversion of interest and attention to other aspects of the business is bound to badly affect the e.m. program. One reason is that the greatest opportunity for energy cost control and energy-efficiency gains is in improved operational and maintenance practices. Implementing and sustaining good operational and maintenance procedures is an exceedingly demanding job and requires a constant attention and a dedication to detail that is rarely found in corporate business life. The energy manager should be energetic, enthusiastic, dedicated, and political.

The second step is the appointment of the plant e.m. committee. This should consist of one group of persons who are able to and have some motivation for cutting fuel and electric costs and a second group who have the technical knowledge or access to data needed for the program department managers or their assistants. Thus, the e.m. committees should include labor representatives, the maintenance department head, a manager of finance or data storage, some engineers, and a public relations person. The energy manager should keep up to date on the energy situation daily, convene the committee weekly, and present a definitive report to top management at least monthly and at other times when required by circumstance. It is suggested also that several subcommittees be broken out of the main committee to consider such important aspects as capital investments, employee education, operator-training programs, external public relations, and so on. The committee will define strategy, provide criticism, publish newsletters and press releases, carry out employee programs, argue for the acceptance of feasible measures before management, represent the program in the larger community, and be as supportive as possible to the energy coordinator. This group has the most to risk and the most to gain. They must defend their own individual interests against the group but at the same time must cooperate in making the program successful and thus be eligible for rewards from top management for their good work and corporate success.

As the e.m. program progresses to the energy audit and beyond, it will be necessary to keep all employees informed as to its purposes, its goals, and how its operation will impact plant operations and employee routine, comfort, and job performance. The education should proceed through written and oral channels as best benefits the organizational structure. Newsletters, posters, and employee meetings have been used successfully.

In addition to general education about energy conservation, it may prove worthwhile to offer specialized courses for boiler, mechanical and electrical equipment operators and other workers whose jobs can affect energy utilization in the plant. The syllabuses should be based on thermodynamic principles applied to the systems involved and given on an academic level consistent with the workers backgrounds. Long-range attempts to upgrade job qualifications through such training can have very beneficial effects on performance. The courses can be given by community colleges, by private enterprises, professional societies or by in-house technical staff, if available.

The material presented here on organization is based on the presumption that a considerable management organization already exists and that sufficient technical and financial resources exist for support of the energy management program as outlined. Obviously, very small businesses cannot operate on this scale, however, we have found many small companies that have carried out effective energy management efforts.

Setting Energy Conservation Goals. It is entirely appropriate and perhaps even necessary to select an energy reduction goal for the first year of the program very early in the program. The purpose is to gain the advantage of the competitive spirit of those employees that can be aroused by a target goal. Unfortunately, the true potential for conservation and the investment costs required to achieve it are not known until the plant energy audit is completed and a detailed study made of the data. Furthermore, a wide variety of energy-use patterns exists even with a single industry.

However, looking at the experience of other industries that have set goals and met them can provide some useful guidance.

An excellent example of a long-term successful energy management program in a large industrial corporation is that of the 3M Company, headquartered in St. Paul, Minnesota.[10] 3M is a large, diversified manufacturing company with more than 50 major product lines; it makes some 50,000 products at over 50 different factory locations around the country. The corporate energy management objective is to use energy as efficiently as possible in all operations; the management believes that all companies have an obligation to conserve energy and all other natural resources.

Energy productivity at 3M improved 63% from 1973 to 2004. They saved over $70 million in 2004 because of their energy management programs, and saved a total of over $1.5 billion in energy expenses from 1973 to 2004. From 1998 through 2004, they reduced their overall energy use by 27% in their worldwide operations.[11] Their program is staffed by three to six people who educate and motivate all levels of personnel on the benefits of energy management. The categories of programs implemented by 3M include conservation, maintenance procedures, utility operation optimization, efficient new designs, retrofits through energy surveys, and process changes.

Energy-efficiency goals at 3M are set and then the results are measured against a set standard to determine the success of the programs. The technologies that have resulted in the most dramatic improvement in energy efficiency include heat-recovery systems, high-efficiency motors, variable-speed drives, computerized facility management systems, steam-trap maintenance, combustion improvements, variable-air-volume systems, thermal insulation, cogeneration, waste-steam utilization, and process improvements. Integrated manufacturing techniques, better equipment utilization, and shifting to nonhazardous solvents have also resulted in major process improvements.

The energy management program at 3M has worked very well, but the company's management is not yet satisfied. They have set a goal of further improving energy efficiency at a rate of 2%–3% per year for the next 5 years. This goal will produce a 10%–15% reduction in energy use per pound of product or per square foot of building space. They expect to substantially reduce their emissions of waste gases and liquids, to increase the energy recovered from wastes, and to constantly increase the profitability of their operations. 3M continues to stress the extreme importance that efficient use of energy can have on their industrial productivity.

14.2.1.2 Phase 2: Audit and Analysis

Energy Audit of Equipment and Facilities. Historical data for the facility should be collected, reviewed, and analyzed. The review should identify gross energy uses by fuel types, cyclic trends, fiscal year effects, dependence on sales or work load, and minimum energy-use ratios. Historical data are graphed in a form similar to the one shown in Chapter 10. Historical data assist in planning a detailed energy audit and alert the auditors as to the type of fuel and general equipment to expect. A brief facility walk-through is recommended to establish the plant layout, major energy uses, and primary processes or functions of the facility.

The energy audit is best performed by an experienced or at least trained team, since visual observation is the principal means of information gathering and operational assessment. A team would have from three to five members, each with a specific assignment for the audit. For example, one auditor would check the lighting, another the HVAC system, another the equipment and processes, another the building structure (floor space, volume, insulation, age, etc.), and another the occupancy use schedule, administration procedures, and employees' general awareness of energy management.

The objectives of the audit are to determine how, where, when, and how much energy is used in the facility. In addition, the audit helps to identify opportunities to improve the energy-use efficiency and its operations. Some of the problems encountered during energy audits are determining the rated power of equipment, determining the effective hours of use per year, and determining the effect of seasonal, climatic, or other variable conditions on energy use. Equipment ratings are often obscured by dust or grease (unreadable nameplates). Complex machinery may not have a single nameplate listing the total capacity, but several giving ratings for component equipment. The effect of load is also important because energy use in a machine operating at less than full load may be reduced along with a possible loss in operating efficiency.

The quantitative assessment of fuel and energy use is best determined by actual measurements under typical operational conditions using portable or installed meters and sensing devices. Such devices include light meters, ammeters, thermometers, airflow meters, recorders, and so on. In some situations, sophisticated techniques such as infrared scanning or thermography are useful. The degree of measurement and recording sophistication naturally depends on available funds and the potential savings anticipated. For most situations, however, nameplate and catalog information are sufficient to estimate power demand. Useful information can be obtained from operating personnel and their supervisors—particularly as it relates to usage patterns throughout the day. A sample form that can be used for recording audit data is shown in Chapter 10 (Figure 10.3).

The first two columns of the form are self-explanatory. The third column is used for the rated capacity of the device, (e.g., 5kW). The sixth column is used if the device is operated at partial load. Usage hours (column 7) are based on all work shifts, and are corrected to account for the actual operating time of the equipment. The last three columns are used to convert energy units to a common basis (e.g., MJ or Btu).

Data recorded in the field are reduced easily by the use of specialized software or spread sheets that provide uniform results and summaries in a form suitable for review or for further analysis. Computer analysis also provides easy modification of the results to reflect specific management reporting requirements or to present desired comparisons for different energy use, types of equipment, and so on.

In-Plant Metering. Submetering reduces the work and time required for an energy audit; indeed, it does much more that that. Because meters are tools for assessing production control and for measuring equipment efficiency, they can contribute directly to energy conservation and cost containment. Furthermore submetering offers the most effective way of evaluating the worth of an energy-efficiency measure. Too many managers accept a vendor's estimate of fuel savings after buying a recuperator. They may scan the fuel bills for a month or two after the purchase to get an indication of savings—usually in vain—and then relax and accept the promised benefit without ever having any real indication that it exists. It may well be that, in fact, it does not yet exist. The equipment without directly metering the fuel input. It is estimated that at least 2.5% waste is recoverable by in-plant metering.

Oil meters are just as effective as gas meters used in the same way and are even less expensive on an energy-flow basis. Electric meters are particularly helpful in monitoring the continued use of machines or lighting during shutdown periods and for evaluating the efficacy of lubricants and the machineability of feed stock. The use of in-plant metering can have its dark side too. The depressing part is the requirement for making periodic readings. It does not stop even there. Someone must analyze the readings so that something can be done about them. If full use is to be made of the information contained in meter readings, it must be incorporated into the energy information portion of the management information system. At the very least each subreading must be examined chronologically to detect malfunctions or losses of efficiency. Better still, a derived quantity such as average energy per unit of production should be examined.

A Special Case: Energy Audit of a Process. In some manufacturing and process industries it is of interest to determine the energy content of a product. This can be done by a variation of the energy audit techniques described earlier. Since this approach resembles classical financial accounting, it is sometimes called energy accounting. In this procedure the energy content of the raw materials is determined in a consistent set of energy units. Then, the energy required for conversion to a product is accounted for in the same units. The same is done for energy in the waste streams and the by-products. Finally, the net energy content per unit produced is used as a basis for establishing efficiency goals.

In this approach, all materials used in the product or used to produce it are determined. Input raw materials used in any specific period are normally available from plant records. Approximations of specific energy content for some materials can be found in the literature or can be obtained from the U.S. Department of Commerce or other sources. The energy content of a material includes that due to extraction and refinement as well as an inherent heating value it would have as a fuel prior to processing. Consequently, nonfuel-type ores in the ground are assigned zero energy and petroleum products are assigned their alternate value as a fuel prior to processing in addition to the refinement energy. The energy of an input metal stock would include the energy due to extraction, ore refinement to metal, and any milling operations.

Conversion energy is an important aspect of the energy audit, since it is under direct control of plant management. All utilities and fuels coming into the plant are accounted for. They are converted to consistent energy units (joules or Btu) using the actual data available on the fuels or using approximate conversions.

Electrical energy is assigned the actual fuel energy required to produce the electricity. This accounts for power conversion efficiencies. A suggested approach is to assume (unless actual values are available from your utility) that 10.8 MJ (10,200 Btu) is used to produce 3.6 MJe (1 kWh), giving a fuel conversion efficiency of 3.6 + 10.8 = 0.33 or 33%.

The energy content of process steam includes the total fuel and electrical energy required to operate the boiler as well as line and other losses. Some complexities are introduced when a plant produces both power and steam, since it is necessary to allocate the fuel used to the steam and power produced. One suggested way to make this allocation is to assume that there is a large efficient boiler feeding steam to a totally condensing vacuum turbine. Then, one must determine the amount of extra boiler fuel that would be required to permit the extraction of steam at whatever pressure while maintaining the constant load on the generator. The extra fuel is considered the energy content of the steam being extracted.

Waste disposal energy is that energy required to dispose of or treat the waste products. This includes all the energy required to bring the waste to a satisfactory disposal state. In a case where waste is handled by a contractor or some other utility service, it would include the cost of transportation and treatment energy.

If the plant has by-products or coproducts, then energy credit is allocated to them. A number of criteria can be used. If the by-product must be treated to be utilized or recycled (such as scrap), then the credit would be based on the raw material less the energy expended to treat the by-product for recycle. If the by-product is to be sold, the relative value ratio of the by-product to the primary product can be used to allocate the energy.

Analysis of Audit Results, Identification of Energy Management Opportunities. Often the energy audit will identify immediate energy management opportunities, such as unoccupied areas that have been inadvertently illuminated 24 hours per day, equipment operating needlessly, and so on. Corrective housekeeping and maintenance action can be instituted to achieve short-term savings with little or no capital investment.

An analysis of the audit data is required for a more critical investigation of fuel waste and identification of the potential for conservation. This includes a detailed energy balance of each process, activity, or facility. Process modification and alternatives in equipment design should be formulated, based on technical feasibility and economic and environmental impact. Economic studies to determine payback, return on investment, and net savings are essential before making capital investments.

14.2.1.3 Phase 3: Implementation and Submeters

At this point goals for saving energy can be established more firmly and priorities set on the modification and alterations to equipment and the process: Effective measurement and monitoring Procedures are essential in evaluating progress in the energy management program. Routine reporting procedures between management and operations should be established to accumulate information on plant performance and to inform plant supervisiors of the effectiveness of their operation. Time-tracking charts of energy use and costs can be helpful involving employees and recognizing their contributions facilitate the achievement of objectives. Finally, the program must be continually reviewed and analyzed with regard to established goals and procedures.

14.2.2 The Energy Audit Report

Energy audits do not save money and energy for companies unless the recommendations are implemented. Audit reports should be designed to encourage implementation. The goal in writing an audit report should not be the report itself; rather, it should be to achieve implementation of the report recommendations and thus achieve increased energy efficiency and energy cost savings for the customer. In this section, the authors discuss their experience with writing industrial energy audit reports and suggest some ways to make the reports more successful in terms of achieving a high rate of the recommendations.[12]

- Present information visually. The authors present their client's energy-use data visually with graphs showing the annual energy and demand usage by month. These graphs give a picture of use patterns. Any discrepancies in use show up clearly.

- Make calculation sections helpful. The methodology and calculations used to develop specific energy management opportunity recommendations are useful in an audit report. Including the methodology and calculations gives technical personnel the ability to check the accuracy of one's assumptions and one's work. However, not every reader wants to wade through pages describing the methodology and showing the calculations. Therefore, the authors provide this information in a technical supplement to the audit report. Because this section is clearly labeled as the technical supplement, other readers are put on notice as to the purpose of this section.
- Use commonly understood units. When preparing the report, be sure to use units that the client will understand. Discussing energy savings in terms of Btus (British thermal units) may only be meaningful to the engineers and more technical readers. For management and operating personnel, kilowatt-hours (for electricity) or therms (for natural gas) are better units, because most energy bills use these units.
- Explain your assumptions. A major problem with many reports is a failure to explain the assumptions underlying the calculations. For example, when the authors use operating hours in a calculation, it is always carefully shown how the number was figured; for example, "Your facility operates from 7:30 am to 8:00 pm, five days a week, 51 weeks per year. Therefore, we will use 3188 h in our calculations."

  When basic assumptions and calculations are shown, the reader can make adjustments if those facts change. In the example above, if the facility decided to operate 24 h/day, the reader would know where and how to make changes in operating hours, because that calculation had been clearly labeled.

  The authors use one section of their report to list the standard assumptions and calculations. Thus, explanations for each of the recommendations do not have to be repeated. Some of the standard assumptions/calculations included in this section are operating hours, average cost of electricity, demand rate, off-peak cost of electricity, and the calculation of the fraction of air-conditioning load attributable to lighting.
- Be accurate and consistent. The integrity of a report is grounded in its accuracy. This does not just mean correctness of calculations. Clearly, inaccurate calculations will destroy a report's credibility, but other problems can also undermine the value of a report. Use the same terminology so that the reader is not confused. Make sure that the same values are used throughout the report. Do not use two different load factors for the same piece of equipment in different recommendations. This, for example, could happen if one calculated the loss of energy due to leaks from a compressor in one recommendation and the energy savings due to replacing the compressor motor with a high-efficiency motor in another recommendation.
- Proofread the report carefully: Typographical and spelling errors devalue an otherwise good product. With computer spell checkers, there is very little excuse for misspelled words. Nontechnical readers are likely to notice this type of error, and they will wonder if the technical calculations are similarly flawed.

14.2.2.1 Report Sections

The authors have found that the following report format meets their clients' needs and fits the authors' definition of a user-friendly report.

Executive Summary. The audit report should start with an executive summary that basically lists the recommended energy conservation measures and shows the implementation cost and dollar savings amount. This section is intended for the readers who want to see only the bottom line. Although the executive summary can be as simple as a short table, the authors add some brief text to explain the recommendations and sometimes include other special information needed to implement the recommendations. They also copy the executive summary on colored paper so that it stands out from the rest of the report.

Energy Management Plan. Following the executive summary, some information is provided to the decision makers on how to set up an energy management program in their facility. The authors view this section as one that encourages implementation of the report, so every attempt is made to try to make it as helpful as possible.

Energy Action Plan. In this subsection, the authors describe the steps that a company should consider in order to start implementing the report's recommendations.

Energy Financing Options. The authors also include a short discussion of the ways that a company can pay for the recommendations. This section covers the traditional use of company capital, loans for small businesses, utility incentive programs, and the shared savings approach of the energy service companies.

Maintenance Recommendations. The authors do not usually make formal maintenance recommendations in the technical supplement, because the savings are not often easy to quantify. However, in this section of the report, energy savings maintenance checklists are provided for lighting, heating/ventilation/air conditioning, and boilers.

The Technical Supplement. The technical supplement is the part of the report that contains the specific information about the facility and the audit recommendations. The authors' technical supplement has two main sections: one includes the report's assumptions and general calculations and the other describes the recommendations in detail, including the calculations and methodology. The authors sometimes include a third section that describes measures that were analyzed and determined not to be cost-effective, or that have payback times beyond the client's planning horizon.

Standard Calculations and Assumptions. This section was briefly described above when the importance of explaining assumptions was discussed. Here, the reader is provided with the basis for understanding many of the authors' calculations and assumptions. Included is a short description of the facility: square footage (both air-conditioned and unconditioned areas); materials of construction; type and level of insulation; etc. If the authors are breaking the facility down into subareas, those areas are described and each area is assigned a number which is then used throughout the recommendation section.

Standard values calculated in this section include operating hours, average cost of electricity, demand rate, off-peak cost of electricity, and the calculation of the fraction of air-conditioning load attributable to lighting. When a value is calculated in this section, the variable is labeled with an identifier that remains consistent throughout the rest of the report.

Audit Recommendations. This section contains a discussion of each of the energy management opportunities the authors have determined to be cost-effective. Each energy management recommendation (or EMR) which was capsulized in the executive summary is described in-depth here.

Again, the authors try to make the EMRs user-friendly. To do this, the narrative discussion is placed at the beginning of a recommendation and the technical calculations are left for the very end. In this manner, the authors allow the readers to decide for themselves whether they want to wade through the calculations.

Each EMR starts with a table that summarizes the energy, demand and cost savings, implementation cost, and simple payback period. Then follows a short narrative section that provides some brief background information about the recommended measure and explains how it should be implemented at the facility in question. If the authors are recommending installation of more than one item (lights, motors, air-conditioning units, etc.), a table is often used to break down the savings by unit or by area.

The final section of each EMR is the calculation section. Here the authors explain the methodology that was used to arrive at the report's savings estimates. The equations are provided and it is shown how the calculations are performed so that the clients can see what has been done. If they want to change the report's assumptions, they can. If some of the data the authors have used is incorrect, they can replace it with the correct data and recalculate the results. However, by placing the calculations away from the rest of the discussion rather than intermingling it, the authors do not scare off the readers who need to know the other information.

Appendix. The authors use an appendix for lengthy data tables. For example, there is a motor efficiencies table which is used in several of the authors' EMRs. Instead of repeating it in each EMR, it is printed in the appendix. The authors also include a table showing the facility's monthly energy-use history and a table listing the major energy-using equipment. Similar to the calculation section of the EMRs, the appendix allows the authors to provide backup information with out cluttering up the main body of the report.

14.3 Improving Industrial Energy Audits

14.3.1 Preventing Overestimation of Energy Savings in Audits

A frequent criticism of energy audits is that they overestimate the savings potential available to the facility. This section addresses several problem areas which can result in overly optimistic savings projections, and suggests ways to prevent mistakes.[13] This possibility of overestimation concerns many of the people and organizations that are involved in some part of this energy audit process. It concerns utilities who do not want to pay incentives for demand-side management programs if the facilities will not realize the expected results in energy or demand savings. Overestimates also make clients unhappy when their energy bills do not decrease as much as promised. The problem multiplies when a shared savings program is undertaken by the facility and an energy service company. Here, the difference between the audit projections and the actual metered and measured savings may be so significantly different that either there are no savings for the facility, or the energy service company makes no profit.

More problems are likely with the accuracy of the energy audits for industrial and manufacturing facilities than for smaller commercial facilities or even large buildings because the equipment and operation of industrial facilities is more complex. However, many of the same problems discussed here in terms of industrial and manufacturing facilities can occur in audits of large commercial facilities and office buildings. Based on the authors' auditing experience for industrial and manufacturing facilities over the last five years, it is possible to identify a number of areas where problems are likely to occur, and a number of these are presented and discussed. In addition, the authors have developed a few methods and approaches to dealing with these potential problems, and a few ways have been found to initiate energy audit analyses that lead the authors to improved results. One of these approaches is to collect data on the energy-using equipment in an industrial or manufacturing facility and then to perform both an energy and a demand balance to help insure that reasonable estimates of energy uses—and therefore, energy savings—are available for this equipment.

In addition, unfortunately, some analysts use the average cost of electricity to calculate energy savings. This can give a false picture of the actual savings and may result in overly optimistic savings predictions. This section also discusses how to calculate the correct values from the electricity bills, and when to use these values. Finally, this section discusses several common energysavings measures that are frequently recommended by energy auditors. Some of these may not actually save as much energy or demand as expected, except in limited circumstances. Others have good energy-saving potential but must be implemented carefully to avoid increasing energy use rather than decreasing it.

14.3.2 Calculating Energy and Demand Balances

The energy and demand balances for a facility are an accounting of the energy flows and power used in the facility. These balances allow the energy analyst to track the energy and power inputs and outputs (uses) and see whether they match. A careful energy analyst should perform an energy and demand balance on a facility before developing and analyzing any energy management recommendations.[14] In this way, the analyst can determine what the largest energy users are in a facility, can find out whether all—or almost all—energy uses have been identified, and can see whether more savings have been identified than are actually achievable. Making energy-use recommendations without utilizing the energy and demand balances is similar to making budget-cutting recommendations without knowing exactly where the money is currently being spent.

When the authors perform an energy survey (audit), all of the major energy-using equipment in the facility is inventoried. Then the authors list the equipment and estimate its energy consumption and demand using the data gathered at the facility, such as nameplate ratings of the equipment and operating hours. The energy balance is developed by major equipment category, such as lighting, motors, HVAC, air compressors, etc. There is also a category called *miscellaneous* to account for loads that were not individually surveyed, such as copiers, electric typewriters, computers, and other plug loads. The authors typically allocate 10% of the actual energy use and demand to the *miscellaneous* category in the demand and energy balances. (For an office building instead of a manufacturing facility, this miscellaneous load might be 15%–20%). Then, the energy and demand for each of the other categories is calculated.

14.3.2.1 Lighting

The first major category analyzed is lighting, because this is usually the category in which the authors have the most confidence for knowing the actual demand and hours of use. Thus, they believe that the energy and demand estimates for the lighting system are the most accurate, and can then be subtracted from the total actual use to let the authors continue to build up the energy and demand balance for the facility. The authors record the types of lamps and number of lamps used in each area of the facility and ask the maintenance person to show them the replacement lamps and ballasts used. With this lamp and ballast wattage data, together with a good estimate of the hours that the lights are on in the various areas, they can construct what they believe to be a fairly accurate description of the energy and demand for the lighting system.

14.3.2.2 Air-Conditioning

There is generally no other "easy" or "accurate" category to work on, so the authors proceed to either air-conditioning or motors. In most facilities, there will be some air-conditioning, even if it is just for the offices that are usually part of the industrial or manufacturing facility. Many facilities—particularly in the hot and humid southeast—are fully air-conditioned. Electronics, printing, medical plastics and devices, and many assembly plants are common ones seen to be fully air-conditioned. Boats, metal products, wood products, and plastic pipe-manufacturing facilities are most often not air conditioned. Air-conditioning system nameplate data is usually available and readable on many units, and efficiency ratings can be found from published ARI data,[15] or from the manufacturers of the equipment. The biggest problem with air-conditioning is to get runtime data that will allow the author(s) of the report to determine the number of full-load equivalent operating hours for the air-conditioning compressors or chillers. From the authors' experience in north and north-central Florida, about 2200–2400 h are used per year of compressor runtime for facilities that have air-conditioning that responds to outdoor temperature. Process cooling requirements are much different, and would typically have much larger numbers of full-load equivalent operating hours. With the equipment size, the efficiency data, and the full-load equivalent operating hours, it is possible to construct a description of the energy and demand for the air-conditioning system.

14.3.2.3 Motors

Turning next to motors, the authors begin looking at one of the most difficult categories to deal with in the absence of fully metered and measured load factors on each motor in the facility. In a one-day plant visit, it is usually impossible to get actual data on the load factors for more than a few motors. Even then, that data is only good for the one day that it was taken. Very few energy-auditing organizations can afford the time and effort to make long-term measurements of the load factor on each motor in an industrial or manufacturing facility. Thus, estimating motor load factors becomes a critical part of the energy and demand balance, and also a critical part of the accuracy of the actual energy audit analysis. Motor nameplate data shows the horsepower rating, the manufacturer, and sometimes the efficiency. If not, the efficiency can usually be obtained from the manufacturer, or from standard references such as the *Energy-Efficient Motor Systems Handbook*,[16] or from software databases such as MotorMaster produced by the Washington State Energy Office.[17] The authors inventory all motors over 1 hp, and sometimes try to look at the smaller ones if there is enough time.

Motor runtime is another parameter that is very difficult to obtain. When the motor is used in an application where it is constantly on is an easy case. Ventilating fans, circulating pumps, and some process-drive motors are often in this class because they run for a known, constant period of time each year. In other cases, facility operating personnel must help provide estimates of motor runtimes. With data on the horsepower, efficiency, load factor, and runtimes of motors, it is possible to construct a detailed table of motor energy and demands to use in the report's balances. Motor load factors will be discussed further in a later section of this paper.

14.3.2.4 Air Compressors

Air compressors are a special case of motor use with most of the same problems. Some help is available in this category because some air compressors have instruments showing the load factor, and some have runtime indicators for hours of use. Most industrial and manufacturing facilities will have several air compressors, and this may lead to some questions as to which air compressors are actually used and how many hours they are used. If the air compressors at a facility are priority-scheduled, it may turn out that one or more of the compressors are operated continuously, and one or two smaller compressors are cycled or unloaded to modulate the need for compressed air. In this case, the load factors on the larger compressors may be unity. Using this data on the horsepower, efficiency, load factor, and runtimes of the compressors, the authors develop a detailed table of compressor energy use and demand for the report's energy and demand balances.

14.3.2.5 Other Process Equipment

Specialized process equipment must be analyzed on an individual basis because it will vary tremendously depending on the type of industry or manufacturing facility involved. Much of this equipment will utilize electric motors and will be covered in the motor category. Other electrically-powered equipment, such as drying ovens, cooking ovens, welders, and laser and plasma cutters, are nonmotor electric uses and must be treated separately. Equipment nameplate ratings and hours of use are necessary to compute the energy and demand for these items. Process chillers are another special class that are somewhat different from the comfort air-conditioning equipment, because the operating hours and loads are driven by the process requirements and not the weather patterns and temperatures.

14.3.2.6 Checking the Results

After the complete energy and demand balances are constructed for the facility, the authors check to see if the cumulative energy/demand for these categories plus the miscellaneous category is substantially larger or smaller than the actual energy usage and demand over the year. If it is, and it is certain that all of the major energy uses have been identified, the authors know that a mistake was made somewhere in their assumptions. As mentioned above, one area that has typically been difficult to accurately survey is the energy use by motors. Measuring the actual load factors is difficult on a one-day walk-through audit visit, so the authors use the energy balance data to help estimate the likely load factors for the motors. This is done by adjusting the load factor estimates on a number of the motors to arrive at a satisfactory level of the energy and demand from the electric motors. Unless this is done, it is likely that the energy used by the motors will be overestimated, and thus overestimate the energy savings from replacing standard motors with high-efficiency motors.

As an example, the authors performed an energy audit for one large manufacturing facility with a lot of motors. It was first assumed that the load factors for the motors were approximately 80%, based on what the facility personnel explained. Using this load factor gave a total energy use for the motors of over 16 million kWh/year and a demand of over 2800 kW. Because the annual energy use for the entire facility was just over 11 million kWh/year and the demand never exceeded 2250 kW, this load factor was clearly wrong. The authors adjusted the average motor load factor to 40% for most of the motors, which reduced the energy use figure to 9 million kWh and the demand to just under 1600 kW. These values are much more reasonable with motors making up a large part of the electrical load of this facility.

After the energy/demand balances have been satisfactorily compiled, the authors use a graphics program to draw a pie chart showing the distribution of energy/demand between the various categories. This allows visual representation of which categories are responsible for the majority of the energy use. It also makes it possible to focus the energy savings analyses on the areas of largest energy use.

14.3.3 Problems with Energy Analysis Calculations

Over the course of performing 120 industrial energy audits, the authors have identified a number of problem areas. One lies with the method of calculating energy cost savings: whether to use the average cost of electricity or break the cost down into energy and demand cost components. Other problems include instances where the energy and demand savings associated with specific energy-efficiency measures may not be fully realized or where more research should go into determining the actual savings potential.

14.3.3.1 On-Peak and Off-Peak Uses: Overestimating Savings by Using the Average Cost of Electricity

One criticism of energy auditors is that they sometimes overestimate the dollar savings available from various energy-efficiency measures. One way overestimation can result is when the analyst uses only the average cost of electricity to compute the savings. Because the average cost of electricity includes a demand component, using this average cost to compute the savings for companies who operate on more than one shift can overstate the dollar savings. This is because the energy cost during the off-peak hours does not include a demand charge. A fairly obvious example of this type of problem occurs when the average cost of electricity is used to calculate savings from installing high-efficiency security lighting. In this instance, there is no on-peak electricity use, but the savings will be calculated as if all the electricity was used on-peak.

The same problem arises when an energy-efficiency measure does not result in an expected—or implicitly expected—demand reduction. Using a cost of electricity that includes demand in this instance will again overstate the dollar savings. Examples of energy-efficiency measures that fall into this category are occupancy sensors, photosensors, and adjustable-speed drives (ASDs). Although all of these measures can reduce the total amount of energy used by the equipment, there is no guarantee that the energy use will only occur during off-peak hours. While an occupancy sensor will save lighting kWh, it will not save any kW if the lights come on during the peak load period. Similarly, an ASD can save energy use for a motor, but if the motor needs its full load capability—as an air-conditioning fan motor or chilled-water pump motor might—during the peak load period, the demand savings may not be there. The reduced use of the device or piece of equipment on-peak load times may introduce a diversity factor that produces some demand savings. However, even this savings will be overestimated by using the average cost of electricity in most instances.

On the other hand, some measures can be expected to provide their full demand savings at the time of the facility's peak load. Replacing 40-W T12 fluorescent lamps with 32-W T8 lamps will provide a verifiable demand savings because the wattage reduction will be constant at all times, and will specifically show up during the period of peak demand. Shifting loads to off-peak times should also produce verifiable demand savings. For example, putting a timer or energy management system control on a constant-load electric drying oven to insure that it does not come on until the off-peak time will result in the full demand savings. Using high-efficiency motors also seems like it would also produce verifiable savings because of its reduced kW load, but in some instances, there are other factors that tend to negate these benefits. This topic is discussed later.

To help solve the problem of overestimating savings from using the average cost of electricity, the authors divide their energy savings calculations into a demand savings and an energy savings. In most instances, the energy savings for a particular piece of equipment is calculated by first determining the demand savings for that equipment and then multiplying by the total operating hours of the equipment. To calculate the annual cost savings (CS), the following formula is used:

$$\begin{aligned} CS = & [\text{Demand savings} \times \text{Average monthly demand rate} \times 12 \text{ months/year}] \\ & + [\text{Energy savings} \times \text{Average cost of electricity without demand}]. \end{aligned}$$

If a recommended measure has no demand savings, then the energy cost savings is simply the energy savings times the average cost of electricity without demand (or off-peak cost of electricity). This procedure forces us to think carefully about which equipment is used on-peak and which is used off-peak.

To demonstrate the difference in savings estimates, consider replacing a standard 30-hp motor with a high-efficiency motor. The efficiency of a standard 30-hp motor is 0.901 and a high-efficiency motor is 0.931. Assume the motor has a load factor of 40% and operates 8760 h/year (three shifts). Assume also that the average cost of electricity is \$0.068/kWh (including demand), the average demand cost is \$3.79/kW/mo, and the average cost of electricity without demand is \$0.053/kWh. The equation for calculating the demand of a motor is:

$$D = HP \times LF \times 0.746 \times 1/\text{Eff}.$$

The savings on demand (or demand reduction) from installing a high-efficiency motor is:

$$\begin{aligned} DR = & HP \times LF \times 0.746 \times (1/\text{Eff}_S - 1/\text{Eff}_H) = 30 \text{ hp} \times 0.40 \times 0.746 \text{ kW/hp} \\ & \times (1/0.901 - 1/0.931) = 0.32 \text{ kW}. \end{aligned}$$

The annual energy savings (ES) is:

$$ES = DR \times H = 0.32 \text{ kW} \times 8760 \text{ h/year} = 2803.2 \text{ kWh/year}.$$

Using the average cost of electricity above, the cost savings (CS_1) calculated as:

$$CS_1 = ES \times (\text{Average cost of electricity}) = 2803.2 \text{ kWh/year} \times 0.068/\text{kWh} = 190.62/\text{yr}.$$

Using the recommended formula above:

$$\begin{aligned} CS = & [\text{Demand savings} \times \text{Average monthly demand rate} \times 12\text{months/year}] \\ & + [\text{Energy savings} \times \text{Average cost of electricity without demand}] \\ & = (0.32 \text{ kW} \times 3.79/\text{month} \times 12\text{months/year}) + (2803.2 \text{ kWh/year} \times 0.053/\text{kWh}) \\ & = (14.55 + 148.57)/\text{year} = 163.12/\text{year}. \end{aligned}$$

In this example, using the average cost to calculate the energy cost savings overestimates the cost savings by \$27.50 per year, or 17%. Although the actual amount is small for one motor, if this error is repeated for all the motors for the entire facility as well as all other measures that reduce the demand component only during the on-peak hours, then the cumulative error in cost savings predictions can be substantial.

14.3.3.2 Motor Load Factors

Many in the energy-auditing business started off assuming that motors ran at full load or near full load, and based their energy consumption analysis and energy savings analysis on that premise. Most books and publications that give a formula for finding the electrical load of a motor do not even include a term for the motor load factor. However, since experience soon showed the authors that few motors actually run at full load or near full load, they were left in a quandary about what load factor to actually use in calculations, because good measurements on the actual motor load factor are rarely to be had. A classic paper by R. Hoshide shed some light on the distribution of motor load factors from his experience.[18] In this paper, Hoshide noted that only about one-fourth of all three-phase motors run with a load factor greater than 60%, with 50% of all motors running at load factors between 30 and 60%, and one-fourth running with load factors less than 30%. Thus, those auditors who had been assuming that a typical motor load factor was around 70 or 80% had been greatly overestimating the savings from high-efficiency motors, adjustable-speed drives, high-efficiency belts, and other motor-related improvements.

The energy and demand balances discussed earlier also confirm that overall motor loads in most facilities cannot be anywhere near 70%–80%. The authors' experience in manufacturing facilities has been that motor load factors are more correctly identified as being in the 30%–40% range. With these load factors, one obtains very different savings estimates and economic results than when one assumes that a motor is operating at a 70% or greater load factor, as shown in the example earlier.

One place where the motor load factor is critical—but often overlooked—is in the savings calculations for ASDs. Many motor and ASD manufacturers provide easy-to-use software that will determine savings with an ASD if you supply the load profile data. Usually a sample profile is included that shows calculations for a motor operating at full load for some period of time, and at a fairly high overall load factor—e.g., around 70%. If the motor has a load factor of only 50% or less to begin with, the savings estimates from a quick use of one of these programs may be greatly exaggerated. If the actual motor use profile with the load factor of 50% is use, one may find that the ASD will still save some energy and money, but often not as much as it looks like when the motor is assumed to run at the higher load factor. For example, a 20-hp motor may have been selected for use on a 15-hp load to insure that there is a "safety factor." Thus, the maximum load factor for the motor would be only 75%. A typical fan or pump in an air-conditioning system that is responding to outside weather conditions may operate at its maximum load only about 10% of the time. Because that maximum load here is only 15 hp, the average load factor for the motor might be more like 40%, and will not be even close to 75%.

14.3.3.3 High-Efficiency Motors

Another interesting problem area is associated with the use of high-efficiency motors. In Hoshide's paper mentioned earlier, he notes that, in general, high-efficiency motors run at a faster full-load speed than standard-efficiency motors. This means that when a standard motor is replaced by a high-efficiency motor, the new motor will run somewhat faster than the old motor in almost every instance. This is a problem for motors that drive centrifugal fans and pumps, because the higher operating speed means greater power use by the motor. Hoshide provides an example where he shows that a high-efficiency motor that should be saving about 5% energy and demand actually uses the same energy and demand as the old motor. This occurs because the increase in speed of the high-efficiency motor offsets the power savings by almost exactly the same 5% due to the cube law for centrifugal fans and pumps.

Few energy auditors ever monitor fans or pumps after replacing a standard motor with a high-efficiency motor; therefore, they have not realized that this effect has cancelled the expected energy and demand savings. Since Hoshide noted this feature of high-efficiency motors, the authors have been careful to make sure that their recommendations for replacing motors with centrifugal loads carry the notice that it will probably be necessary to adjust the drive pulleys or drive system so that the load is operated at the same speed to achieve the expected savings.

14.3.3.4 Motor Belts and Drives

The authors have developed some significant questions about the use of cogged and synchronous belts, and the associated estimates of energy savings. It seems fairly well accepted that cogged and synchronous belts do transmit more power from a motor to a load than if standard smooth V-belts are used. In some instances, this should certainly result in some energy savings. A constant-torque application like a conveyor drive may indeed save energy with a more efficient drive belt because the motor will be able to supply that torque with less effort. Consider also a feedback-controlled application, such as a thermostatically controlled ventilating fan or a level-controlled pump. In this case, the greater energy transmitted to the fan or pump should result in the task being accomplished faster than if less drive power were supplied, and some energy savings should exist. However, if a fan or a pump operates in a nonfeedback application—as is common for many motors—then there will not be any energy savings. For example, a large ventilating fan that operates at full load continuously without any temperature or other feedback may not use less energy with an efficient drive belt, because the fan may run faster as a result of the drive belt having less slip. Similarly, a pump that operates continuously to circulate water may not use less energy with an efficient drive belt. This is an area that needs some monitoring and metering studies to check the actual results.

Whether efficient drive belts result in any demand savings is another question. Because, in many cases, the motor is assumed to be supplying the same shaft horsepower with or without high-efficiency drive belts, a demand savings does not seem likely in these cases. It is possible that using an efficient belt on a motor with a constant-torque application which is controlled by an ASD might result in some demand savings. However, for the most common applications, the motor is still supplying the same load, and thus would have the same power demand. For feedback-controlled applications, there might be a diversity factor involved so that the reduced operation times could result in some demand savings—but not the full value otherwise expected. Thus, using average cost electricity to quantify the savings expected from high-efficiency drive belts could well overestimate the value of the savings. Verification of the cases where demand savings are to be expected is another area where more study and data are needed.

14.3.3.5 Adjustable-Speed Drives

The authors would like to close this discussion with a return to ASDs because these are devices that offer a great potential for savings, but have far greater complexities than are often understood or appreciated. Fans and pumps form the largest class of applications where great energy savings is possible from the use of ASDs. This is a result again of the cube law for centrifugal fans and pumps where the power required to drive a fan or pump is specified by the cube of the ratio of the flow rates involved. According to the cube law, a reduction in flow to one-half the original value could now be supplied by a motor using only one-eighth of the original horsepower. Thus, whenever an airflow or liquid flow can be reduced, such as in a variable-air-volume system or with a chilled-water pump, there is a dramatic savings possible with an ASD. In practice, there are two major problems with determining and achieving the expected savings.

The first problem is the one briefly mentioned earlier, and that is determining the actual profile of the load involved. Simply using the standard profile in a piece of vendor's software is not likely to produce very realistic results. There are so many different conditions involved in fan and pump applications that taking actual measurements is the only way to get a good idea of the savings that will occur with an ASD. Recent papers have discussed the problems with estimating the loads on fans and pumps, and have shown how the cube law itself does not always give a reasonable value.[19–21] The Industrial Energy Center at Virginia Polytechnic Institute and Virginia Power Company have developed an approach wherein they classify potential ASD applications into eight different groups, and then estimate the potential savings from analysis of each system and from measurements of that system's operation.[22] Using both an analytical approach and a few measurements allows them to get a reasonable estimate of the motor load profile, and thus a reasonable estimate of the energy and demand savings possible.

The second problem is achieving the savings predicted for a particular fan or pump application. It is not sufficient to just identify the savings potential and then install an ASD on the fan or pump motor.

In most applications, there is some kind of throttling or bypass action that results in almost the full horsepower still being required to drive the fan or pump most of the time. In these applications, the ASD will not save much, unless the system is altered to remove the throttling or bypass device and a feedback sensor is installed to tell the ASD what fraction of its speed to deliver. This means that in many airflow systems, the dampers or vanes must be removed so that the quantity of air can be controlled by the ASD changing the speed of the fan motor. In addition, some kind of feedback sensor must be installed to measure the temperature or pressure in the system to send a signal to the ASD or a PLC controller to alter the speed of the motor to meet the desired condition. The additional cost of the alterations to the system and the cost of the control system needed greatly change the economics of an ASD application compared to the case where only the purchase cost and installation cost of the actual ASD unit is considered.

For example, a dust collector system might originally be operated with a large 150-hp fan motor running continuously to pick up the dust from eight saws. However, because production follows existing orders for the product, sometimes only two, three, or four saws are in operation at a particular time. Thus, the load on the dust collector is much lower at these times than if all eight saws are in use. An ASD is a common recommendation in this case, but estimating the savings is not easy to begin with, and after the costs of altering the collection duct system and of adding a sophisticated control system to the ASD are considered, the bottom line result is much different than the cost of the basic ASD with installation. Manual or automatic dampers must be added to each duct at a saw so that it can be shut off when the saw is not running. In addition, a PLC for the ASD must be added to the new system, together with sensors added to each damper so that the PLC will know how many saws are in operation and therefore what speed to tell the ASD for the fan to run to meet the dust collection load of that number of saws. Without these system changes and control additions, the ASD itself will not save any great amount of energy or money. Adding them in might well double the cost of the basic ASD, and double the payback time that may have originally been envisioned.

Similarly, for a water or other liquid-flow application, the system piping or valving must be altered to remove any throttling or bypass valves, and a feedback sensor must be installed to allow the ASD to know what speed to operate the pump motor. If several sensors are involved in the application, then a PLC may also be needed to control the ASD. For example, putting an ASD on a chilled-water pump for a facility is much more involved, and much more costly, than simply cutting the electric supply lines to the pump motor and inserting an ASD for the motor. Without the system alterations and without the feedback control system, the ASD cannot provide the savings expected.

14.3.4 General Rules

New energy auditors often do not have the experience to have engineering judgment about the accuracy of their analyses. That is, they cannot look at the result and immediately know that it is not within the correct range of likely answers. Because the authors' IAC program has a fairly steady turnover of students, they find the same type of errors cropping up over and over as draft audit reports are reviewed. To help new team members develop the engineering judgment that they will eventually gain through experience, the authors are developing "rules of thumb" for energy analyses. The rules of thumb are intended to provide a ballpark estimate of the expected results. For example, if the rule of thumb for the percent for installing high-efficiency motors says that the savings range is 3%–5% of the energy use by the motors, then a student who comes up with a savings of 25% will immediately know that the calculations are wrong and will know to check the assumptions and data entry to see where the error lies. Without these rules of thumb, the burden for checking these results is shifted to the team leaders and program directors. Although this does not obviate the need for report review, it minimizes the likelihood that errors will occur. The authors suggest that other organizations who frequently utilize and train new energy auditors consider developing such rules of thumb for the major types of facilities or geographic areas that they audit.

Energy auditing is not an exact science, but a number of opportunities is available for improving the accuracy of the recommendations. Techniques which may be appropriate for small-scale energy audits can introduce significant errors into the analyses for large complex facilities. This chapter began by discussing how to perform an energy and demand balance for a company. This balance is an important step in doing an energy-use analysis, because it provides a check on the accuracy of some of the assumptions necessary to calculate savings potential. It also addressed several problem areas which can result in overly optimistic savings projections, and suggested ways to prevent mistakes. Finally, several areas where additional research, analysis, and data collection are needed were identified. After this additional information is obtained, everyone can produce better and more accurate energy audit results.

14.3.4.1 Decision Tools for Improving Industrial Energy Audits—OIT Software Tools

The Office of Industrial Technologies—a program operated by the U.S. Department of Energy, Division of Energy Efficiency and Renewable Energy—provides a series of computer software tools that can be obtained free from their Web site, or by ordering a CD at no cost from them. With the right know-how, these powerful tools can be used to help identify and analyze energy system savings opportunities in industrial and manufacturing plants. Although the tools are accessible at the U.S. DOE Web site for download, they also encourage users to attend a training workshop to enhance their knowledge and take full advantage of opportunities identified in the software programs. For some tools, advanced training is also available to help further increase expertise in their use.

14.3.4.2 Decision Tools for Industry—Order the Portfolio of Tools on CD

The Decision Tools for Industry CD contains the MotorMaster+(MM+), Pump System Assessment Tool, Steam System Tool Suite, 3E Plus, and the new AirMaster+software packages described here. In addition, it includes MM+training. The training walks the user through both the fundamentals and the advanced features of MM+and provides examples for using the software to make motor purchase decisions. The CD can be ordered via email from the EERE Information Center or by calling the EERE Information Center at 1-877-EERE-INF (877-337-3463).

DOE Industry Tools:

- AIRMaster+
- Chilled Water System Analysis Tool (CWSAT)
- Combined Heat and Power Application Tool (CHP)
- Fan System Assessment Tool (FSAT)
- MotorMaster+4.0
- MotorMaster+International
- NO_x and Energy Assessment Tool (NxEAT)
- Plant Energy Profiler for the Chemical Industry (ChemPEP Tool)
- Process Heating Assessment and Survey Tool (PHAST)
- Pumping System Assessment Tool 2004 (PSAT)
- Steam System Tool Suite

Other Industry Tools:

- ASDMaster: Adjustable Speed Drive Evaluation Methodology and Application

AIRMaster+. This tool provides comprehensive information on assessing compressed-air systems, including modeling, existing and future system upgrades, and evaluating savings and effectiveness of energy-efficiency measures.

Chilled Water System Analysis Tool Version 2.0. Use the CWSAT to determine energy requirements of chilled-water distribution systems, and to evaluate opportunities for energy and costs savings by applying improvement measures. Provide basic information about an existing configuration to calculate current energy consumption, and then select proposed equipment or operational changes for comparison. The results of this analysis will help the user quantify the potential benefits of chilled-water system improvements.

Combined Heat and Power Application Tool. The CHP Application Tool helps industrial users evaluate the feasibility of CHP for heating systems such as fuel-fired furnaces, boilers, ovens, heaters, and heat exchangers. It allows analysis of three typical system types: fluid heating, exhaust-gas heat recovery, and duct burner systems. Use the tool to estimate system costs and payback period, and to perform "what-if" analyses for various utility costs. The tool includes performance data and preliminary cost information for many commercially available gas turbines and default values that can be adapted to meet specific application requirements.

Fan System Assessment Tool. Use the FSAT to help quantify the potential benefits of optimizing fan system configurations that serve industrial processes. FSAT is simple and quick, and requires only basic information about the fans being surveyed and the motors that drive them. With FSAT, one can calculate the amount of energy used by one's fan system, determine system efficiency, and quantify the savings potential of an upgraded system.

MotorMaster+4.0. An energy-efficient motor selection and management tool, MotorMaster+4.0 software includes a catalog of over 20,000 AC motors. This tool features motor inventory management tools, maintenance log tracking, efficiency analysis, savings evaluation, energy accounting, and environmental reporting capabilities.

MotorMaster+International. MotorMaster+International includes many of the capabilities and features of MotorMaster+; however, now it can help evaluate repair/replacement options on a broader range of motors, including those tested under the Institute of Electrical and Electronic Engineers (IEEE) standard, and those tested using International Electrical Commission (IEC) methodology. With this tool, analyses can be conducted in different currencies, and it will calculate efficiency benefits for utility rate schedules with demand charges, edit and modify motor rewind efficiency loss defaults, and determine "best available" motors. The tool can be modified to operate in English, Spanish, and French.

NO_x and Energy Assessment Tool. The NxEAT helps plants in the petroleum refining and chemical industries to assess and analyze NO_x emissions and application of energy-efficiency improvements. Use the tool to inventory emissions from equipment that generates NO_x, and then compare how various technology applications and efficiency measure affect overall costs and reduction of NO_x. Perform "what-if" analyses to optimize and select the most cost-effective methods for reducing NO_x from systems such as fired heaters, boilers, gas turbines, and reciprocating engines.

Plant Energy Profiler for the Chemical Industry. The ChemPEP Tool provides chemical plant managers with the information they need to identify savings and efficiency opportunities. The ChemPEP Tool enables energy managers to see overall plant energy use, identify major energy-using equipment and operations, summarize energy cost distributions, and pinpoint areas for more detailed analysis. The ChemPEP Tool provides plant energy information in an easy-to-understand graphical manner that can be very useful to managers.

Process Heating Assessment and Survey Tool. The PHAST provides an introduction to process heating methods and tools to improve thermal efficiency of heating equipment. Use the tool to survey process-heating equipment that uses fuel, steam, or electricity, and identify the most energy-intensive equipment. It can also help perform an energy (heat) balance on selected equipment (furnaces) to identify and reduce nonproductive energy use. Compare performance of the furnace under various operating conditions and test "what-if" scenarios.

Pumping System Assessment Tool 2004. The Pumping System Assessment Tool helps industrial users assess the efficiency of pumping system operations. PSAT uses achievable pump performance data from Hydraulic Institute standards and motor performance data from the MotorMaster+database to calculate potential energy and associated cost savings.

Steam System Tool Suite. In many industrial facilities, steam system improvements can save 10%–20% in fuel costs. To help tap into potential savings in typical industrial facilities, DOE offers a suite of tools for evaluating and identifying steam system improvements.

- *Steam System Assessment Tool (SSAT) Version 2.0.0*. The SSAT allows steam analysts to develop approximate models of real steam systems. Using these models, SSAT can be applied to quantify the magnitude—energy, cost, and emission savings—of key potential steam improvement opportunities. SSAT contains the key features of typical steam systems. The enhanced and improved version includes features such as a steam demand savings project; a user-defined fuel model; a boiler stack loss worksheet for the SSAT fuels; a boiler flash steam recovery model; and improved steam-trap models.
- *3E Plus, Version 3.2*. The program calculates the most economical thickness of industrial insulation for user input operating conditions. Calculations can be made using the built-in thermal performance relationships of generic insulation materials or supply conductivity data for other materials.
- *Steam Tool Specialist Qualification Training*. Industry professionals can earn recognition as Qualified Specialists in the use of the BestPractices Steam Tools. DOE offers an in-depth two-and-a-half-day training session for steam system specialists, including two days of classroom instruction and a written exam. Participants who complete the workshop and pass the written exam are recognized by DOE as Qualified Steam Tool Specialists. Specialists can assist industrial customers in using the BestPractices Steam Tools to evaluate their steam systems.

ASDMaster: Adjustable Speed Drive Evaluation Methodology and Application. This Windows™ software program helps plant or operations professionals determine the economic feasibility of an ASD application, predict how much electrical energy may be saved by using an ASD, and search a database of standard drives. The package includes two 3.5-inch diskettes, a user's manual, and a user's guide. Please order from EPRI. For more information, see the ASDMaster Web Site.

14.3.4.3 Energy-Auditing Help From Industrial Assessment Centers

The Industrial Assessment Centers (IACs), sponsored by the U.S. Department of Energy, Energy Efficiency and Renewable Energy Division (EERE) Industrial Technologies Program (ITP), provide eligible small- and medium-sized manufacturers with no-cost energy assessments. Additionally, the IACs serve as a training ground for the next generation of energy savvy engineers.

Teams composed mainly of engineering faculty and students from the centers, located at 26 universities around the country, conduct energy audits or industrial assessments and provide recommendations to manufacturers to help them identify opportunities to improve productivity, reduce waste, and save energy. Recommendations from industrial assessments have averaged about $55,000 in potential annual savings for each manufacturer.

As a result of performing these assessments, upper-class and graduate engineering students receive unique hands on assessment training and gain knowledge of industrial process systems, plant systems, and energy systems, making them highly attractive to employers.

To be eligible for an IAC assessment, a manufacturing plant must meet the following criteria:

- Within Standard Industrial Codes (SIC) 20–39
- Generally be located within 150 miles of a host campus
- Gross annual sales below $100 million
- Fewer than 500 employees at the plant site
- Annual energy bills more than $100,000 and less than $2.5 million
- No professional in-house staff to perform the assessment

FIGURE 14.5 Industrial Assessment Center (IAC) locations and service areas. (From U.S. Department of Energy, Office of Industrial Technologies, 2005, http://www.oit.doe.gov/iac/schools.shtml.)

Presently (2007), there are 26 schools across the country participating in the IAC Program. For additional information or to apply for an assessment, go to http://www.oit.doe.gov/iac/schools.shtml and click on one of the school names for contact information. A map of the IAC centers and their service areas is shown in Figure 14.5.

14.4 Industrial Electricity End Uses and Electrical Energy Management

14.4.1 The Importance of Electricity in Industry

Electricity use in industry is primarily for electric drives, electrochemical processes, space heating, lighting, and refrigeration. Table 14.2 lists the relative importance of use of electricity in the industrial sector. Timely data on industrial and manufacturing energy use is difficult to obtain, since its frequency of collection is only every three years, and then it takes EIS another three years to process the results. In addition, it is only a sample survey, so the accuracy of the data is less that what would be optimal. However, it is better data than is available for most countries in the world. The most recent, detailed data on manufacturing energy end use (as of early 2005) is still the 1998 MECS (Manufacturing Energy Consumption Survey) data. Table 14.2 shows the electrical energy consumed for different end uses in 1998.

Because several of the categories above involve motor use, the total manufacturing energy use for motors is actually the sum of several categories:

Machine drives	53.8%
Facility HVAC	8.3%
Process cooling and refrigeration	6.0%
Onsite transportation	0.1%
Total motors	68.2%

This gives the same results shown as the pie chart shown in Figure 14.3.

TABLE 14.2 Manufacturing Electricity by End Use, 1998

	GWh	Percent Use
Machine drives (motors)	551,318	53.8
Electrochemical processes	103,615	10.1
Process heat	106,330	10.4
Facility HVAC (motors)	84,678	8.3
Facility lighting	66,630	6.5
Process cooling and refrigeration	61,263	6.0
Onsite transportation (motors)	1427	0.1
Other process	3882	0.4
Other	46,005	
Total	1,025,148	100

Source: From Manufacturing Energy Consumption Survey, 1998; U.S. Department of Energy; Energy Information Agency, 2002.

14.4.2 Electric Drives

Electric drives of one type or another use 68% of industrial electricity. Examples include electric motors, machine tools, compressors, refrigeration systems, fans, and pumps. Improvements in these applications would have a significant effect on reducing industrial electrical energy.

Motor efficiency can be improved in some cases by retrofit (modifications, better lubrication, improved cooling, heat recovery), but generally requires purchasing of more efficient units. For motors in the sizes 1–200 hp, manufacturers today supply a range of efficiency. Greater efficiency in a motor requires improved design, more costly materials, and generally greater first cost. Losses in electric-drive systems may be divided into four categories:

	Typical Efficiency (%)
Prime mover (motor)	10–95
Coupling (clutches)	80–99
Transmission	70–95
Mechanical load	1–90

Each category must be evaluated to determine energy management possibilities. In many applications the prime mover will be the most efficient element of the system. Table 14.3 shows typical induction motor data, illustrating the improvement in efficiency due to the 1992 Energy Policy Act. Note that both efficiency and power factor decrease with partial load operation, which means that motors should be sized to operate at or near full load ratings.

TABLE 14.3 Typical Electric Motor Data

	Full Load Efficiency (%)	
Size (hp)	1975	1993 (EPACT)
1	76	82.5–85.5
2	80	84.0–86.5
5	84	87.5–89.5
10	85	89.5–91.7
20	85	91.0–93.0
40	85	93.0–94.5
100	91	94.1–95.4
200	90	95.0–96.2

Manufacturers have introduced new high-efficiency electric motors in recent years. Many utilities offer rebates of $5–$20 per horsepower for customers installing these motors.

14.4.2.1 Electrochemical Processes

Industrial uses of electrochemical processes include electrowinning, electroplating, electrochemicals, electrochemical machining, fuel cells, welding, and batteries.

A major use of electrolytic energy is in electrowinning—the electrolytic smelting of primary metals such as aluminum and magnesium. Current methods require on the order of 13–15 kWh/kg; efforts are under way to improve electrode performance and reduce this to 10 kWh/kg. Recycling now accounts for one-third of aluminum production; this requires only about 5% of the energy required to produce aluminum from ore.

Electrowinning is also an important low-cost method of primary copper production. Another major use of electrochemical processes is in the production of chlorine and sodium hydroxide from salt brine. Electroplating and anodizing are two additional uses of electricity of great importance. Electroplating is basically the electrodeposition of an adherent coating upon a base metal. It is used with copper, brass, zinc, and other metals. Anodizing is roughly the reverse of electroplating, with the workpiece (aluminum) serving as the anode. The reaction progresses inward from the surface to form a protective film of aluminum oxide on the surface.

Fuel cells are devices for converting chemical energy to electrical energy directly through electrolytic action. Currently they represent a small use of energy, but research is directed at developing large systems suitable for use by electric utilities for small dispersed generation plants. Batteries are another major use of electrolytic energy, ranging in size from small units with energy storage in the joule or fractional joule capacity up to units proposed for electric utility use that will store 18×10^9 J (5 MWh). Electroforming, etching, and welding are forms of electrochemical used in manufacturing and material shaping. The range of applications for these techniques stretches from microcircuits to aircraft carriers. In some applications, energy for machining is reduced and reduction of scrap also saves energy. Welding has benefits in the repair and salvage of materials and equipment, reducing the need for energy to manufacture replacements.

14.4.2.2 Electric Process Heat

Electricity is widely used as a source of process heat due to ease of control, cleanliness, wide range in unit capacities (watts to megawatts), safety, and low initial cost. Typical heating applications include resistance heaters (metal sheath heaters, ovens, furnaces), electric salt bath furnaces, infrared heaters, induction and high-frequency resistance heating, dielectric heating, and direct arc electric furnaces.

Electric-arc furnaces in the primary metals industry are a major use of electricity. Typical energy use in a direct arc steel furnace is about 2.0 kWh/kg. Electric-arc furnaces are used primarily to refine recycled scrap steel. This method uses about 40% of the energy required to produce steel from iron ore using basic oxygen furnaces. Energy savings can be achieved by using waste heat to preheat scrap iron being charged to the furnace.

Glass making is another process that uses electric heat. An electric current flows between electrodes placed in the charge, causing it to melt. Electric motors constitute a small part of total glass production. Major opportunities for improved efficiency with electric process heat applications in general include improved heat transfer surfaces, better insulation, heat recovery, and improved controls.

14.4.2.3 HVAC

Heating, ventilating, and air-conditioning (HVAC) is an important use of energy in the industrial sector. The environmental needs in an industrial operation can be quite different from residential or commercial operations. In some cases, strict environmental standards must be met for a specific function or process. More often, the environmental requirements for the process itself are not limiting, but space conditioning is a prerequisite for the comfort of production personnel. Chapters 10 and 11 have a more complete discussion of energy management opportunities in HVAC systems.

14.4.2.4 Lighting

Industrial lighting needs range from low-level requirements for assembly and welding of large structures (such as shipyards) to the high levels needed for manufacture of precision mechanical and electronic components such as integrated circuits. Lighting uses about 20% of U.S. electrical energy and 7% of all energy. Of all lighting energy about 20% is industrial, with the balance being sizes of systems, energy management opportunities in industrial lighting systems are similar to those in residential/commercial systems (see Chapter 10).

14.4.2.5 Electric Load Analysis

The energy audit methodology is a general tool that can be used to analyze energy use in several forms and over a short or long period of time. Another useful technique, particularly for obtaining a short-term view of industrial electricity use, is an analysis based on evaluation of the daily load curve. Normally this analysis uses metering equipment installed by the utility and therefore available at the plant. However, special metering equipment can be installed if necessary to monitor specific process or building.

For small installations both power and energy use can be determined from the kilowatt hour meter installed numbered the equations. Please by the utility. Energy in kWh is determined by

$$E = (0.001)(K_h P_t C_t N)\text{kWh} \tag{14.1}$$

where E = electric energy used, kwh; K_h = meter constant, watt hours/revolution; P_t = potential transformer ratio; C_t = current transformer ratio; and N = number of revolutions of the meter disk. (The value of K_h is usually marked on the meter. P_t and C_t are usually 1.0 for small installations.) To determine energy use, the meter would be observed during an operation and the number of revolutions of the disk counted. Then the equation can be used to determine E.

To determine the average load over some period p (hours), determine E as earlier for time p and then use the relation that

$$L = \frac{E}{p}\ \text{kW} \tag{14.2}$$

where E is in kWh, p is in hours, and L is the load in kW. Larger installations will have meters with digital outputs or strip charts. Often these will provide a direct indication of kWh and kW as a function of time. Some also indicate the reactive load (kVARs) or the power factor.

The first step is to construct the daily load curve. This is done by obtaining kWh readings each hour using the meter. The readings are then plotted on a graph to show the variation of the load over a 24-hour period. Table 14.4 shows a set of readings obtained over a 24-hour period in the XYZ manufacturing plant located in Sacramento, California and operating one shift per day. These readings have been plotted in Figure 14.6.

Several interesting conclusions can be immediately drawn from this figure:

- The greatest demand for electricity occurs at 11:00.
- Through the lunch break the third highest demand occurs.
- The ratio of the greatest demand to the least demand is approximately 3:1.
- Only approximately 50% of the energy used actually goes into making a product (54% on-shift use, 46% off-shift use).

When presented to management, these facts were of sufficient interest that a further study of electricity use was requested. Additional insight into the operation of a plant (and into the cost of purchase of electricity) can be obtained from the load analysis. Following a brief discussion of electrical load parameters, a load analysis for the XYZ Company will be described.

Any industrial electrical load consists of lighting, motors, chillers, compressors, and other types of equipment. The sum of the capacities of this equipment, in kW, is the *connected* load. The actual load at any point in time is normally less that the connected load since every motor is not turned on at the same

TABLE 14.4 Kilowatt Hour Meter Readings for XYZ Manufacturing Company

Time Meter Read	Elapsed kWh	Notes Concerning Usage	Percentage of Total Usage (%)
1:00 (a.m.)	640		
2:00	610		
3:00	570		
4:00	570	7 h preshift use	
5:00	640		
6:00	770		
7:00	1120		
Subtotal	4920		17
8:00	1470		
9:00	1700		
10:00	1790		
11:00	1850		
	(Peak)	9 h on-shift use	
12:00 (noon)	1830		
13:00 (1 p.m.)	1790		
14:00 (2 p.m.)	1790		
15:00 (3 p.m.)	1760		
16:00 (4 p.m.)	1690		
Subtotal	15,670		54
17:00 (5 p.m.)	1470		
18:00 (6 p.m.)	1310		
19:00 (7 p.m.)	1210		
20:00 (8 p.m.)	1090	8 h postshift use	
21:00 (9 p.m.)	960		
22:00 (10 p.m.)	800		
23:00 (11 p.m.)	730		
24:00 (12 a.m.)	640		
Subtotal	8210		29
Grand totals	28,800		100

time; only part of the lights may be on at any one time, and so on. Thus the load is said to be diversified, and a measure of this can be found by calculating a *diversity factor*:

$$\mathrm{DV} = \frac{(D_{m1} + D_{m2} + D_{m3} + \cdots)}{(D_{\max})} \tag{14.3}$$

where D_{m1}, D_{m2}, and so on = sum of maximum demand of individual loads in kW and D_{max} = maximum demand of plant in kW. If the individual loads do not occur simultaneously (usually they do not), the diversity factor will be greater than unity. Typical values for industrial plants are 1.3–2.5.

If each individual load operated to its maximum extent simultaneously, the *maximum* demand for power would be equal to the connected load and the diversity factor would be 1.0. However, as pointed out earlier, this does not happen except for special cases. The demand for power varies over time as loads are added and removed from the system. It is usual practice for the supplying utility to specify a demand interval (usually 0.25, 0.5, or 1.0 h) over which it will calculate the

$$D = \frac{E}{p}\ \mathrm{kW} \tag{14.4}$$

where D = demand in kW, E = kilowatt-hours used during p, and p = demand interval in hours. The demand calculated in this manner is an average value, being greater than the lowest instantaneous demand during the demand interval but less than the maximum demand during the interval.

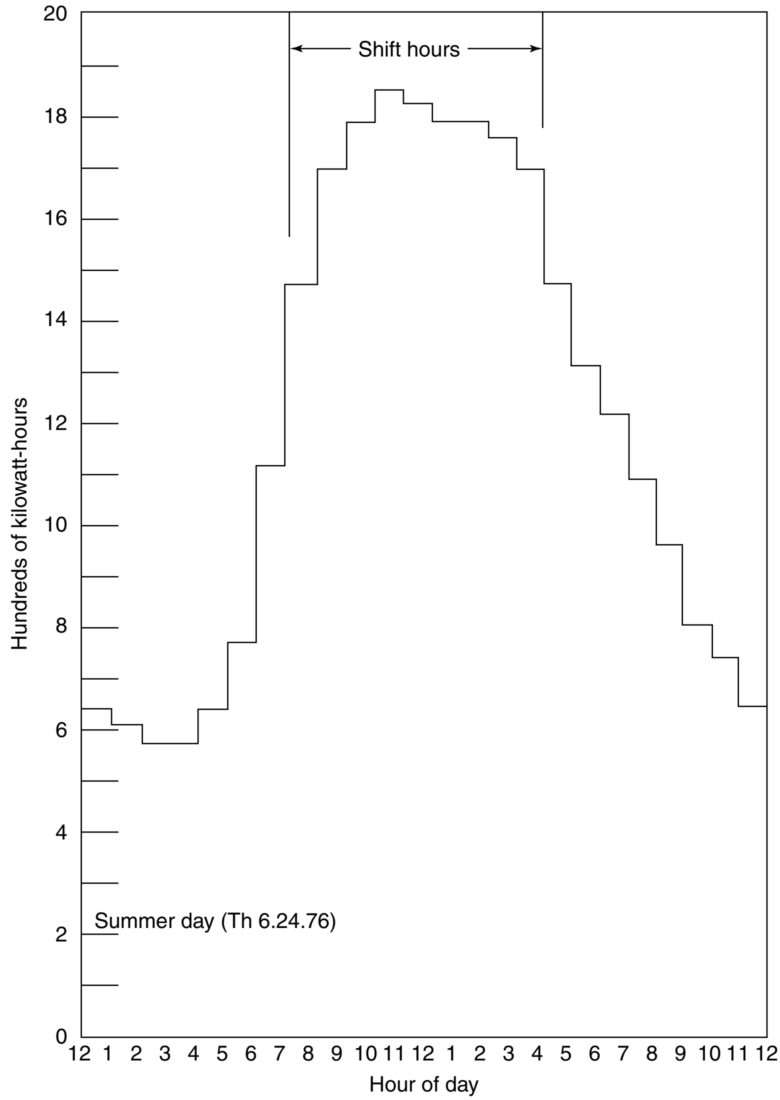


FIGURE 14.6 Daily load curve for XYZ company.

Utilities are interested in *peak demand*, since this determines the capacity of the equipment they must install to meet the customer's power requirements. This is measured by a demand factor, defined as

$$\mathrm{DF} = \frac{D_{\max}}{\mathrm{CL}} \tag{14.5}$$

where $D_{\max}$ = maximum demand in kW and CL = connected load in kW. The demand factor is normally less than unity; typical values range from 0.25 to 0.90.

Since the customer normally pays a premium for the maximum load placed on the utility system, it is of interest to determine how effectively the maximum load is used. The most effective use of the equipment would be to have the peak load occur at the start of the use period and continue unchanged throughout it. Normally, this does not occur, and a measure of the extent to which the maximum demand is sustained throughout the period (a day, month, or year) is given by the *hours use of demand*:

$$\mathrm{HUOD} = \frac{E}{D_{\max}} \text{ hours} \tag{14.6}$$

where HUOD = hours use of demand in hours; E = energy used in period p, in kWh; $D_{\max}$ = maximum demand during period p, in kW, and p = period over which HUOD is determined—for example, 1 day, 1 month, or 1 year (p is always expressed in *hours*).

The *load factor* is another parameter that measures the plant's ability to use electricity efficiently. In effect it measures the ratio of the average load for a given period of time to the maximum load which occurs during the same period. The most effective use results when the load factor is as high as possible once E or HUOD has been minimized (it is always less than one). The load factor is defined as

$$\mathrm{LF} = \frac{E}{(D_{\max})(p)} \tag{14.7}$$

where LF = load factor (dimensionless); E = energy used in period p in kWh; $D_{\max}$ = maximum demand during period p in kW; and p = period over which load factor is determined (e.g., 1 day, 1 month, or 1 year) in hours. Another way to determine LF is from the relation

$$\mathrm{LF} = \frac{\mathrm{HUOD}}{p} \tag{14.8}$$

Still another method is to determine the average load, L = kWh/p during p divided by p and then use the relation

$$\mathrm{LF} = \frac{L}{D_{\max}} \tag{14.9}$$

These relations are summarized for convenience in Table 14.5.

TABLE 14.5 Summary of Load Analysis Parameters

Formulas	Definitions
$E = K_h P_t C_t N/1,000$	E = Electric energy used in period p, kWh
$L = E/P$	$E_{\max}$ = Maximum energy used during period p, kWh
$DV = (D_{m1} + D_{m2} + D_{m3})/D_{\max}$	K_h = Meter constant, watt hours/revolution
$D = E/P$ $D_{\max} = E_{\max}/P$	P_t = Potential transformer ratio
$DF = D_{\max}/CL$	C_t = Current transformer ratio
$\mathrm{HUOD} = E/D_{\max}$	N = Number of revolutions of the meter disk
$LF = \frac{E}{(D_{\max})(p)} = \frac{\mathrm{HUOD}}{p} = \frac{L}{D_{\max}}$	L = Average load, kW
	p = Period of time used to determine load, demand, electricity use, etc., normally 1 hour, day, month, or year; measured in hours
	DV = Diversity factor, dimensionless
	$D_{\max}$ = Maximum demand in period p, kW
	D_{m1}, D_{m2}, etc. = Maximum demand of individual load, kW
	D = Demand during period p, kW
	DF = Demand factor for period p, dimensionless
	CL = Connected load, kW
	HUOD = Hours use of demand during period p, hours
	LF = Load factor during period p, dimensionless

TABLE 14.6 Data for Load Analysis of XYZ Plant

p	$=24$ h
E	$=28{,}800$ kWh/day
E_{max}	$=1850$ kWh
CL	$=2792$ kW
D_{m1}, D_{m2}, etc.	$=53$, 62, 144, 80, 700, 1420 kW

Returning to the XYZ plant, the various load parameters can now be calculated. Table 14.6 summarizes the needed data and the results of the calculations. The most striking thing shown by the calculations is the hours use of demand, equal to 15.6. This is a surprise, since the plant is only operating one shift. The other significant point brought out by the calculations is the low load factor.

An energy audit of the facility was conducted and the major loads were evaluated number of energy management opportunities whereby both loads (kW) and energy use (kWh) could be reduced. The audit indicated that inefficient lighting (on about 12 h per day) could be replaced in the parking lot. General office lighting was found to be uniformly at 100 fc; by selective reduction and task lighting the average level could be reduced to 75 fc or less. The air-conditioning load would also be reduced. Improved controls could be installed to automatically shut down lighting during off-shift and weekend hours (the practice had been to leave the lights on). Some walls and ceilings were selected for repainting to improve reflectance and reduced lighting energy. It was found that the air-conditioning chillers operated during weekends and off-hours; improved controls would prevent this. Also, the ventilation rates were found to be excessive and could be reduced. In the plant, compressed-air system leaks, heat losses from plating tanks, and on-peak operation of the heat treat furnace represented energy and load management opportunities.

The major energy management opportunities were evaluated to have the following potential savings, with a total payback of 5.3 months, as shown in Table 14.7 the average daily savings of electricity amounted to approximately 4400 kWh/day. This led to savings of $80,000 per year, with the cost of the modification being $36,000.

This can be compared to the original situation (Table 14.4). See also Figure 14.7, which shows the daily load curve after the changes have been made. The percentage of use on-shift is now higher. Note that D_{max} has been improved significantly (reduced by 13%); the HUOD has improved slightly (about 3% lower now); and the LF is slightly lower. Furthermore improvements are undoubtedly still possible in this facility; they should be directed first at reducing nonessential uses, thereby reducing HUOD.

So far the discussion has dealt entirely with power and has neglected the reactive component of the load. In the most general case the apparent power in kVA that must be supplied to the load is the sum of the active power in kW and the *reactive power* in KVAR (the reader who is unfamiliar with these terms should refer to a basic electrical engineering text):

$$|S| = \sqrt{P^2 + Q^2} \tag{14.10}$$

where $S=$apparent power in kVA; $P=$active power in kW; and Q$=$reactive power in KVAR. In this notation the apparent power is a vector of magnitude S and angle θ where θ is commonly referred to as the phase angle and given as

$$\theta = \tan^{-1}(\text{KVAR/kW}) \tag{14.11}$$

Another useful parameter is the *power factor*, given by

$$\text{pf} = \cos\theta \tag{14.12}$$

TABLE 14.7 Sample Calculations for XYZ Plant

1. $D_{\max} = \dfrac{E_{\max}}{p} = \dfrac{1,850\ \text{kWh}}{1\ \text{hr}} = 1,850\ \text{kW}$

2. $DV = \dfrac{D_{m1} + D_{m2} + D_{m3} + \dots}{D_{\max}} = \dfrac{2,459}{1,850} = 1.33$

3. $DF = \dfrac{D_{\max}}{CL} = \dfrac{1,850}{2,792} = 0.66$

4. $\text{HOUD}_{(\text{daily})} = \dfrac{E}{D_{\max}} = \dfrac{28,800\ \text{kWh/day}}{1,850\ \text{kW}} = 15.6\ \text{hr/day}$

5. $LF_{(\text{daily})} = \dfrac{\text{HUOD}}{p} = \dfrac{15.6}{24.0} = 0.65$

The load parameters after the changes were made can be found:

$D_{\max} = \dfrac{1,850 - 235}{1\ \text{hr}} = 1,615\ \text{kW}$

$\text{HOUD} = \dfrac{24,400\ \text{kWh/day}}{1,615\ \text{kW}} = 15.1\ \text{hr/day}$

$LF = \dfrac{15.1}{24} = 0.63$

	Savings	
Calculated Savings	kW	kWh/yr
More efficient parking lot lighting	16	67,000
Reduce office lighting	111	495,000
Office lighting controls to reduce off-shift use	—	425,000
Air-conditioning controls and smaller fan motor	71	425,000
Compressed-air system repairs and reduction of heat losses from plating tanks	—	200,000
Shift heat treat oven off-peak	37	—
Totals	235	1,612,000

The revised electricity use was found to be:

	kWh	%
Preshift	4400	18
On-shift	14,400	59
Postshift	5600	23
Totals	24,400	100

The power factor is also given by

$$\text{pf} = \frac{|P|}{|S|} \tag{14.13}$$

The power factor is always less than or equal to unity. A high value is desirable because it implies a small reactive component to the load. A low value means the reactive component is large.

The importance of the power factor is related to the reactive component of the load. Even though the reactive component does not dissipate power (it is stored in magnetic or electric fields), the switch gear and distribution system must be sized to handle the current required by the apparent power, or the vector sum of the active and reactive components. This results in a greater capital and operating expense. The operating expense is increased due to the standby losses that occur in supplying the reactive component of the load.

The power factor can be improved by adding capacitors to the load to compensate for part of the inductive reactance. The benefit of this approach depends on the economics of each specific case and generally requires a careful review or analysis.

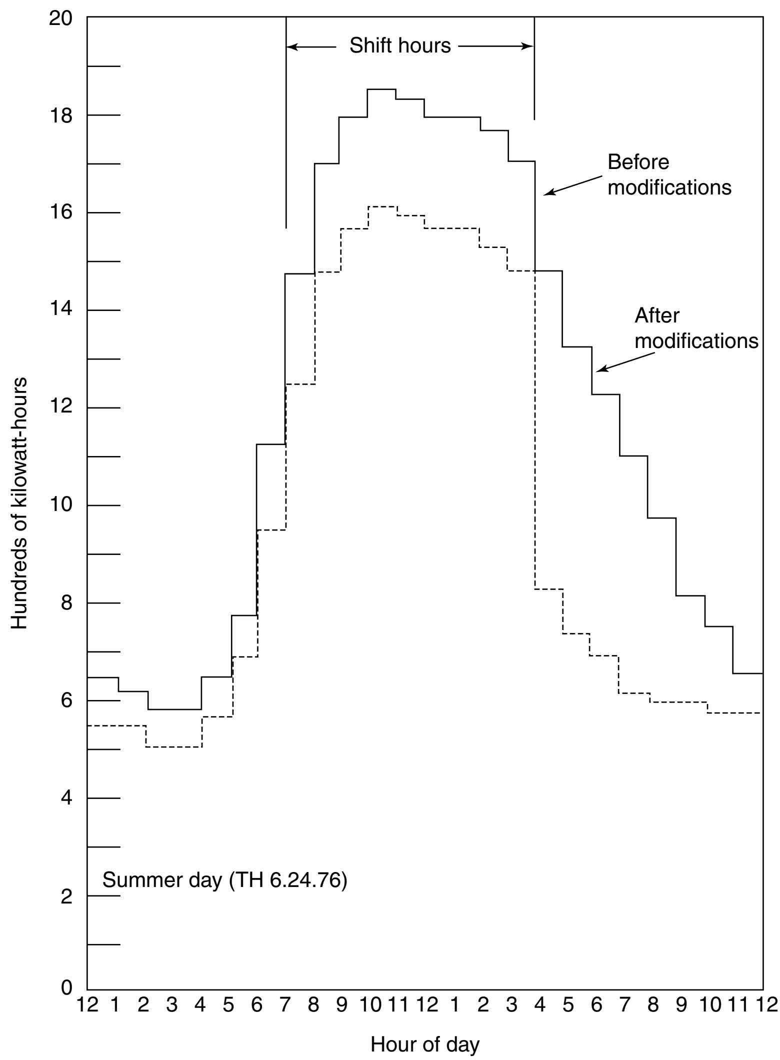


FIGURE 14.7 Daily load curve for XYZ company after modifications.

These points can be clarified with an example. Consider the distribution system shown in Figure 14.8. Four loads are supplied by a 600 A bus. Load A is a distant load that has a large reactive component and a low power factor (pf=0.6). To supply the active power requirement of 75 kW, an apparent power of 125 kVA must be provided and a current of 150 A is required.

The size of the wire to supply the load is dictated by the current to be carried and voltage drop considerations. In this case, #3/0 wire that weighs 508 lb per 1000 ft. and has a resistance of 0.062 Ω per 1000 ft. is used. Since the current in this conductor is 150 A, the power dissipated in the resistance of the conductor is

$$P = \sqrt{3}i^2 r = (1.732)(150^2)(0.124)\mathrm{W}$$

$$P = 4.8\ \mathrm{kW}$$

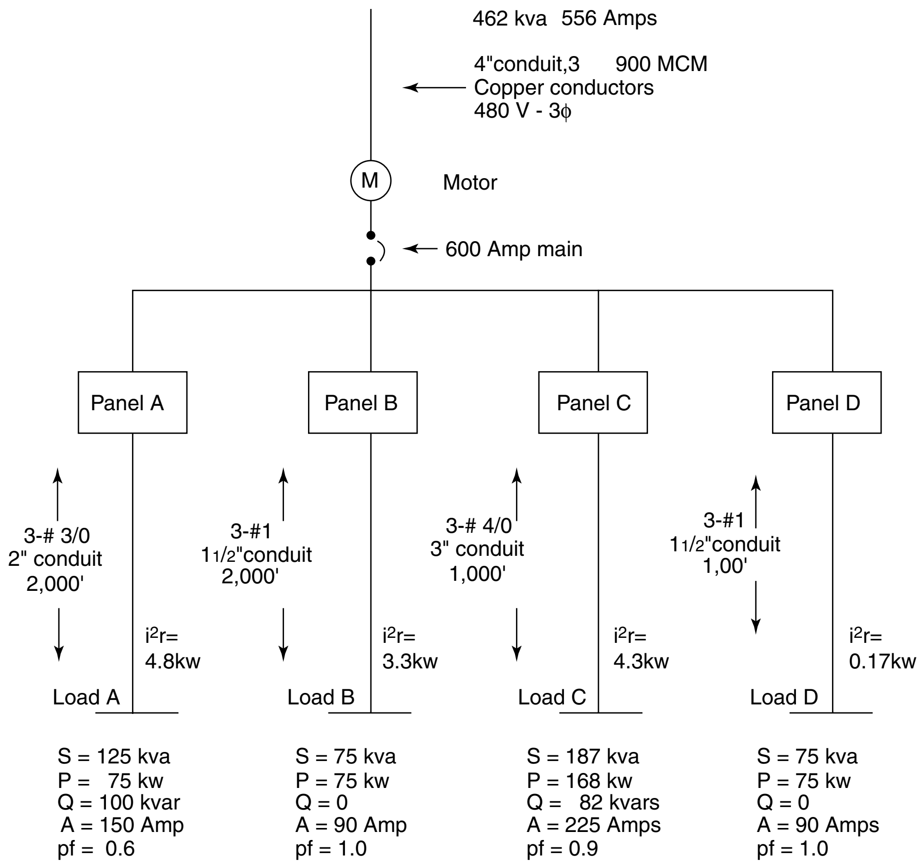


FIGURE 14.8 Electrical diagrams for building 201, XYZ company.

Similar calculations can be made for load B, which uses #1 wire, at 253 lb per 1000 ft. and 0.12 Ω/1000 ft.

Now the effect of the power factor is visible. Although the active power is the same for both load A and load B, load A requires 150 A vs. 90 A for load B. The i^2r standby losses are also higher for load A as opposed to load B. The installation cost to service load B is roughly half that of load A, due to the due to the long conduit run of load B compared to load D. For large loads that are served over long distances and operate continuously, consideration should be given to using larger wire sizes to reduce standby losses.

An estimate of the annual cost of power dissipated as heat in these conduit runs can be made if the typical operating hours of each load are known:

Load	Line Losses in kW	Operating Hours/yr	kWh/year
A	4.8	2000	9600
B	3.3	2000	6600
C	4.3	4000	17,200
D	0.17	2000	340
Total			33,740

At an average cost of 6¢/kWh (includes demand and energy costs), the losses in the distribution system alone are $2022 per year. Over the life of the facility, this is a major expense for a totally unproductive use of energy.

14.4.2.6 Data Acquisition and Control Systems for Energy Management

Data acquisition is essential in energy management for at least three reasons: (1) base-line operational data is an absolute requirement for understanding the size and timing of energy demands for each plant, division, system, and component, and design of the energy management strategy; (2) continuing data acquisition during the course of the energy conservation effort is necessary to calculate the gains made in energy-use efficiency and to measure the success of the program; and (3) effective automatic control depends upon the accurate measurement of the controlled variables and the system operational data. More information on control systems can be found in Chapter 11.

With small, manually controlled systems, data acquisition is possible using indicating instruments and manual recording of data. With large systems, or almost any size automatically controlled system, manual data collection is impractical. The easy availability and the modest price of personal digital computers (PCs) and their present growth in speed and power as their price declines makes them the preferred choice for data acquisition equipment. The only additions needed to the basic PC are the video monitor; a mass data storage device (MSD), analog to digital (A/D) interface cards; software for controlling the data sampling, data storage, and data presentation; and a suitable enclosure for protecting the equipment from the industrial environment.

That same computer is also suitable for use as the master controller of an automatic control system. Both functions can be carried on simultaneously, with the same equipment, with a few additions necessary for the control part of the system. These additional electronic components include multiplexers to increase the capacity of the A/D cards; direct memory access (an addition to the A/D boards), which speeds operation by allowing the data transfer to bypass the CPU and programmable logic controllers (PLCs); and digital input/output cards (I/Os), both used for controlling the equipment controllers. The most critical parts of either the data acquisition or the control system are the software. The hardware can be ordered off-the-shelf; but the software must be either written from scratch or purchased and modified for each particular system in order to achieve the reliable, high-performance operation that is desired. Although many proprietary program languages exist for the PCs and PLCs used for control purposes, BASIC is the most widely used. A schematic diagram of a typical data acquisition and control system is shown in Figure 14.9.

A major application has been in the control of mechanical and electrical systems in commercial and industrial buildings. These have been used to control lighting, electric demand, ventilating fans, thermostat setbacks, air-conditioning systems, and the like. These same computer systems can also be used in industrial buildings with or without modification for process control. Several manufactures of the computer systems will not only engineer and install the system but will maintain and operate it from a remote location. Such operations are ordinarily regional. For large systems in large plants, one may be able to have the same service provided in-plant. However, there are also many small standalone analog control systems that can be used advantageously for the control of simple processes. Examples of these are a temperature controller using a thermocouple output to control the temperature of a liquid storage tank; a level controller, which keeps the liquid level in a tank constant by controlling a solenoid valve; and an oxygen trim system for a small boiler, which translates the measurement of the oxygen concentration in the exhaust stream into the jack shaft position, which regulates the combustion airflow to the burner.

After installing a demand limiter on an electric-arc foundry cupola, the manager was able to reduce the power level from 7100 kW to 4900 kW with negligible effect on the production time and no effect on product quality. The savings in demand charges alone were $4400 per month with an additional savings in energy costs.

14.4.3 Web-Based Facility Automation Systems

Of all recent developments affecting computerized energy management systems, the most powerful new technology to come into use in the last several years has been information technology, or IT. The combination of cheap, high-performance microcomputers, together with the emergence of high-capacity communication lines, networks, and the Internet has produced explosive growth in IT and its

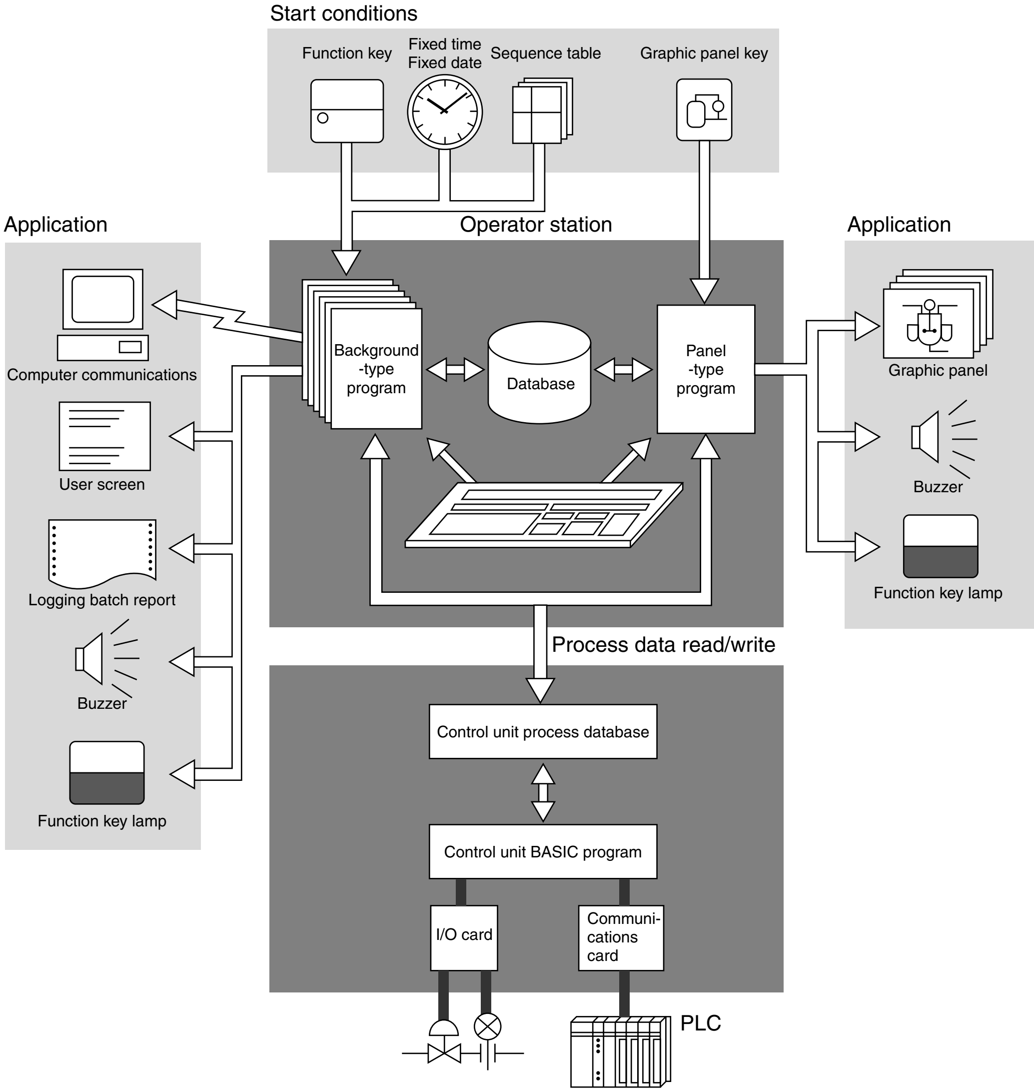


FIGURE 14.9 Typical data acquisition and control system.

application throughout the U.S. economy. Energy information and control systems have been no exception. IT and Internet-based systems are the wave of the future. Almost every piece of equipment and almost every activity will be connected and integrated into the overall facility operation in the next several years.[23]

In particular, the future of DDC in facility automation systems (FAS) can be found on the Web. Almost all FAS manufacturers see the need to move their products to the Internet. Tremendous economies of scale and synergies can be found there. Manufacturers no longer have to create the transport mechanisms for data to flow within a building or campus. They just need to make sure their equipment can utilize the network data paths already installed or designed for a facility. Likewise, with the software to display data to users, manufacturers that take advantage of presentation layer standards such as HTML and JAVA can provide the end user with a rich, graphical, and intuitive interface to their FAS using a standard Web browser.

Owners will reap the benefits of Internet standards through a richer user interface, more competition among FAS providers, and the ability to use their IT infrastructure to leverage the cost of transporting data within a facility. Another area where costs will continue to fall in using Internet standards is the hardware required to transport data within a building or a campus. Off-the-shelf products such as routers, switches, hubs, and server computers make the FAS just another node of the IT infrastructure. Standard IT tools can be used to diagnose the FAS network, generate reports of FAS bandwidth on the intranet, and back up the FAS database.

The FAS of old relied heavily on a collection of separate systems that operated independently, and often with proprietary communication protocols that made expansion, modification, updating, and integration with other building or plant information and control systems very cumbersome, if not impossible. Today, the FAS is not only expected to handle all of the energy- and equipment-related tasks, but also to provide operating information and control interfaces to other facility systems, including the total facility or enterprise management system.

Measuring, monitoring, and maximizing energy savings is a fundamental task of all FAS, and is the primary justification for many FAS installations. Improving facility operations in all areas, through enterprise information and control functions, is fast becoming an equally important function of the overall FAS or facility management system. The Web provides the means to share information easier, quicker, and cheaper than ever before. There is no doubt that the Web is having a huge impact on the FAS industry. The FAS of tomorrow will rely heavily on the Web, TCP/IP, high-speed data networks, and enterprise-level connectivity. If it has not already been done, it is a good time for energy managers to get to know their IT counterparts at their facility, along with those in the accounting and maintenance departments. The future FAS will be here sooner than you think.[24]

14.4.3.1 Energy Management Strategies for Industry

Energy management strategies for industry can be grouped into three categories:

- Operational and maintenance strategies
- Retrofit or modification strategies
- New design strategies

The order in which these are listed corresponds approximately to increased capital investment and increased implementation times. Immediate savings at little or no capital cost can generally be achieved by improved operations and better maintenance. Once these "easy" savings have been realized, additional improvements in efficiency will require capital investments.

14.4.3.2 Electric Drives and Electrically Driven Machinery

About 68% of industrial electricity use is for electrical-motor-driven equipment. Integral horsepower (<0.75 kW) motors are more numerous. Major industrial motor loads are, in order of importance, pumps, compressors, blowers, fans, miscellaneous integral motor applications including conveyors, DC drives, machine tools, and fractional horsepower applications.

Numerous examples of pumping in industry can be observed. These include process pumping in chemical plants, fluid movement in oil refineries, and cooling water circulation. An example of compressors is in the production of nitrogen and oxygen, two common chemicals. Large amounts of electricity are used to drive the compressors, which supply air to the process.

The typical industrial motor is a polyphase motor rated at 11.2 kW (15 hp) and having a life of about 40,000 h. The efficiencies of electric motors have increased recently as a result of higher energy prices, conservation efforts, and new government standards (the Energy Policy Act). High-efficiency motors cost roughly 20%–30% more than standard motors, but this expense is quickly repaid for motors that see continuous use.

Most efficient use of motors requires that attention be given to the following:

Optimum power—Motors operate most efficiently at rated voltage. Three-phase power supplies should be balanced; an unbalance of 3% can increase losses 25%.[1]

Good motor maintenance—Provide adequate cooling, keep heat transfer surfaces and vents clean, and provide adequate lubrication. Improved lubrication alone can increase efficiency a few percentage points.

Equipment scheduling—Turn equipment off when not in use; schedule large motor operation to minimize demand peaks.

Size equipment properly—Match the motor to the load and to the duty cycle. Motors operate most efficiently at rated load.

Evaluate continuous vs. batch processes—Sometimes a smaller motor operating continuously will be more economical.

Power factor—Correct if economics dictate savings. Motors have the best power factor at rated load.

Retrofit or new designs permit use of more efficient motors. For motors up to about 10–15 kW (15–20 hp) there are variations in efficiency. Select the most efficient motor for the job. Check to verify that the additional cost (if any) will be repaid by the savings that will accrue over the life of the installation.

In addition to reviewing the electric-drive system, consider the power train and the load. Friction results in energy dissipation in the form of heat. Bearings, gears, and belt drives all have certain losses, as do clutches. Proper operation and maintenance can reduce energy wastage in these systems and improve overall efficiency.

Material shaping and forming, such as is accomplished with machine tools, requires that electrical energy be transformed into various forms of mechanical energy. The energy expenditure related to the material and to the depth and speed of the cut. By experimenting with a specific process, it is possible to establish cutting rates that are optimum for the levels of production required and are most efficient in terms of energy use. Motors are not the only part of the electric-drive system that sustains losses. Other losses occur in the electric power systems that supply the motor. Electric power systems include substations, transformers, switching gear, distribution systems, feeders, power and lighting panels, and related equipment. Possibilities for energy management include the following:

Use highest voltages that are practical. For a given application, doubling the voltage cuts the required current in half and reduces the i^2r losses by a factor of four.

Eliminate unnecessary transformers. They waste energy. Proper selection of equipment and facility voltages can reduce the number of transformers required and cut transformer losses. Remember, the customer pays for losses when the transformers are on his side of the meter. For example, it is generally better to order equipment with motors of the correct voltage, even if this costs more, than to install special transformers.

Energy losses are an inherent part of electric power distribution systems. This is primarily due to i^2r losses and transformers. The end use conversion systems for electrical energy used in the process also contribute to energy waste. Proper design and operation of an electrical system can minimize energy losses and contribute to the reduction of electricity bills. Where long feeder runs are operated at near-maximum capacities, check to see if larger wire sizes would permit savings and be economically justifiable.

The overall power factor of electrical systems should be checked for low power factor. This could increase energy losses and the cost of electrical service, in addition to excessive voltage drops and increased penalty charges by the utility. Electrical systems studies should be made and consideration should be given to power factor correction capacitors. In certain applications as much as 10%–15% savings can be achieved in a poorly operating plant.

Check load factors. This is another parameter that measures the plant's ability to use electrical power efficiently. It is defined as the ratio of the actual kWh used to the maximum demand in kW times the total hours in the time period. A reduction in demand to bring this ratio closer to unity without decreasing plant output means more economical operation. For example, if the maximum demand for a given month (200 h) is 30,000 kWe and the actual kWh is 3.6×10^6 kWh, the load factor is 60%. Proper management of operations during high demand periods, which may extend only 15–20 min, can reduce the demand during that time without curtailing production. For example, if the 30,000 kWe could be reduced to 20,000 kWe, this would increase the load factor to about 90%. Such a reduction could amount to a \$20,000–\$50,000 reduction in the electricity bill.

Reduce peak loads wherever possible. Many nonessential loads can be shed during the demand peak without interrupting production. These loads would include such items as air compressors, heaters, coolers, and air conditioners. Manual monitoring and control is possible but is often impractical because of the short periods of time that are normally involved and the lack of centralized control systems. Automatic power demand control systems are available.

Provide improved monitoring or metering capability, submeters, or demand recorders. While it is true that meters alone will not save energy, plant managers need feedback to determine if their energy management programs are taking effect. Often the installation of meters on individual processes or buildings leads to immediate savings of 5%–10% by virtue of the ability to see how much energy is being used and to test the effectiveness of corrective measures.

14.4.3.3 Fans, Blowers, and Pumps

Simple control changes are the first thing to consider with these types of equipment. Switches, timeclocks, or other devices can insure that they do not operate except when needed by the process. Heat removal or process mass flow requirements will determine the size of fans and pumps. Often there is excess capacity, either as a result of design conservatism or because of process changes subsequent to the installation of equipment. The required capacity should be checked, since excess capacity leads to unnecessary demand charges and decreased efficiency.

For fans, the volume rate of airflow Q varies in proportion to the speed of the impeller:

$$Q = c_f N \quad \text{m}^3/\text{s} \tag{14.14}$$

where Q = airflow in m^3/s; c_f = a constant with units m^3/r; and N = fan speed, r/s. The pressure developed by the fan varies as the square of the impeller speed. The important rule, however, is that the power needed to drive the fan varies as the cube of the speed:

$$P = P_c N^3 \text{ W} \tag{14.15}$$

where P = input power in watts and P_c = a constant with units W s^3/r^3.

The cubic law of pumping power indicates that if the airflow is to be doubled, eight (2^3) times as much power must be supplied. Conversely, if the airflow is to be cut in half, only one eighth ($1/2^3$) as much power must be supplied. Airflow (and hence power) can be reduced by changing pulleys or installing smaller motors.

Pumps follow laws similar to fans, the key being the cubic relationship of power to the volume pumped through a given system. Small decreases in flow rate, such as might be obtained with a smaller pump or gotten by trimming the impeller, can save significant amounts of energy.

Variable-speed drives (VSDs) are another technique for reducing process energy use. VSDs permit fans, blowers, and pumps to vary speed depending on process requirements. This can lead to significant savings on noncontinuous processes. Recent improvements in solid-state electronics have caused the price of VSDs to drop substantially. This is another technology that is supported by utility rebates in many areas.

14.4.3.4 Air Compressors

Compressed air is a major energy use in many manufacturing operations. Electricity used to compress air is converted into heat and potential energy in the compressed-air stream. Efficient operation of compressed-air systems therefore requires the recovery of excess heat where possible, as well as the maximum recovery of the stored potential energy.

Efficient operation is achieved in these ways:

Select the appropriate type and size of equipment for the duty cycle required. Process requirements vary, depending on flow rates, pressure, and demand of the system. Energy savings can be achieved by selecting the most appropriate equipment for the job. The rotary compressor is more popular for industrial operations in the range of 20–200 kW, even though it is somewhat less efficient than the reciprocal compressor. This has been due to lower initial cost and reduced maintenance. When operated at partial load, reciprocating units can be as much as 25% more efficient than rotary units. However, newer rotary units incorporate a valve that alters displacement under partial load conditions and improves efficiency. Selection of an air-cooled vs. a water-cooled unit would be influenced by whether water or air was the preferred medium for heat recovery.

Proper operation of compressed-air systems can also lead to improved energy utilization. Obviously, air leaks in lines and valves should be eliminated. The pressure of the compressed air should be reduced to a minimum. The percentage saving in power required to drive the compressor at a reduced pressure can be estimated from the fan laws described previously. For example, suppose the pressure were reduced to one-half the initial value. Since pressure varies as the square of the speed, this implies the speed would be 70.7% of the initial value. Since power varies as the cube of the speed, the power would now be $0.707^3 = 35\%$ of the initial value. Of course, this is the theoretical limit; actual compressors would not do as well, and the reduction would depend on the type of compressor. Measurements indicate that actual savings would be about half the theoretical limit; reducing pressure 50% would reduce brake horsepower about 30%. To illustrate this point further, for a compressor operating at 6.89×10^5 N/m^2 (100 psi) and a reduction of the discharge pressure to 6.20×10^5 N/m^2 (90 psi), a 5% decrease in brake horsepower would result. For a 373 kW (500 hp) motor operating for 1 year, the 150,000 kWh savings per year would result in about $9000 per year in electric power costs.

The intake line for the air compressor should be at the lowest temperature available. This normally means outside air. The reduced temperature of air intake results in a smaller volume of air to be compressed. The percentage horsepower saving relative to a 21°C (70°F) intake air temperature is about 2% for each 10°F drop in temperature. Conversely, input power increases by about 2% for each 10°F increase in intake air temperature.

Leakage is the greatest efficiency offender in compressed-air systems. The amount of leakage should be determined and measures taken to reduce it. If air leakage in a plant is more than 10% of the plant demand, a poor condition exists. The amount of leakage can be determined by a simple test during off-production hours (when air-using equipment is shut down) by noting the time that the compressor operates under load compared with the total cycle. This indicates the percentage of the compressor's capacity that is used to supply the plant air leakage. Thus if the load cycle compared with the total cycle were 60 s compared with 180 s, the efficiency would be 33%, or 33% of the compressor capacity is the amount of air leaking in m^3/min (ft.^{3}/min).

Recover heat where feasible. There are sometimes situations where water-cooled or air-cooled compressors are a convenient source of heat for hot water, space heating, or process applications. As a rough rule of thumb, about 300 J/m^3 min of air compressed ($\sim$10 Btu/ft.^{3} min) can be recovered from an air-cooled rotary compressor.

Substitute electric motors for air motor (pneumatic) drives. Electric motors are far more efficient. Typical vaned air motors range in size from 0.15 to 6.0 kW (0.2–8 hp), cost $300–$1500, and produce 1.4–27 N m (1–20 ft. lb) of torque at 620 kN/m^2 (90 psi) air pressure. These are used in manufacturing operations where electric motors would be hazardous, or where light weight and high power are essential. Inefficiency results from air system leaks and the need (compared to electric motors) to generate compressed air as an intermediate step in converting electric to mechanical energy.

Review air usage in paint spray booths. In paint spray booths and exhaust hoods, air is circulated through the hoods to control dangerous vapors. Makeup air is constantly required for dilution purposes. This represents a point of energy rejection through the exhaust air.

Examination should be made of the volumes of air required in an attempt in reduce flow and unnecessary operation. Possible mechanisms for heat recovery from the exhaust gases should be explored using recovery systems.

14.4.3.5 Electrochemical Operations

Electrochemical processes are an industrial use of electricity, particularly in the primary metals industry, where it is used in the extraction process for several important metals. Energy management opportunities include

Improve design and materials for electrodes. Evaluate loss mechanisms for the purpose of improving efficiency.

Examine electrolysis and plating operations for savings. Review rectifier performance, heat loss from tanks, and the condition of conductors and connections.

Welding is another electrochemical process. Alternating current welders are generally preferable when they can be used, since they have a better power factor, better demand characteristics, and more economical operation.

Welding operations can also be made more efficient by the use of automated systems which require 50% less energy than manual welding. Manual welders deposit a bead only 15%–30% of the time the machine is running. Automated processes, however, reduce the no-load time to 40% or less. Different welding processes should be compared in order to determine the most efficient process. Electroslag welding is suited only for metals over 1 cm (0.5 in.) thick but is more efficient than other processes.

Two other significant applications of electrolysis of concern to industry are batteries and corrosion. Batteries are used for standby power, transportation, and other applications. Proper battery maintenance, and improved battery design contribute to efficient energy use.

Corrosion is responsible for a large loss of energy-intensive metals every year and thus indirectly contributes to energy wastage. Corrosion can be prevented and important economies realized, by use of protective films, cathodic protection, and electroplating or anodizing.

14.4.3.6 Steam Systems

In as much as approximately 40% of the energy utilized in industry goes toward the production of process steam, it presents a large potential for energy misuse and fuel waste from improper maintenance and operation. Even though electrically generated steam and hot water is a small percentage of total industrial steam and hot water, the electrical fraction is likely to increase as other fuels increase in price. This makes increased efficiency even more important For example:

Steam leaks from lines and faulty valves result in considerable losses. These losses depend on the size of the opening and the pressure of the steam, but can be very costly. A hole 0.1 ft. in diameter with steam at 200 psig can bleed \$1000–\$2000 worth of steam (500 GJ) in a year.

Steam traps are major contributors to energy losses when not functioning properly. A large process industry might have thousands of steam traps, which could result in large costs if they are not operating correctly. Steam traps are intended to remove condensate and noncondensable gases while trapping or preventing the loss of steam. If they stick open, orifices as large as 6 mm (0.25 in.) can allow steam to escape. Such a trap would allow 1894 GJ/year (2000 MBtu/year) of heat to be rejected to the atmosphere on a 6.89×10^5 N/m (100 psi) pressure steam line. Many steam traps are improperly sized, contributing to an inefficient operation. Routine inspection, testing, and a correction program for steam valves and traps are essential in any energy program and can contribute to cost savings.

Poor practice and design of steam distribution systems can be the source of heat waste up to 10% or more. It is not uncommon to find an efficient boiler or process plant joined to an inadequate steam distribution system. Modernization of plants results from modified steam requirements. The old distribution systems are still intact, however, and can be the source of major heat losses. Large steam lines intended to supply units no longer present in the plant are sometimes used for minor needs, such as space heating and cleaning operations, that would be better accomplished with other heat sources.

Steam distribution systems operating on an intermittent basis require a start-up warming time to bring the distribution system into proper operation. This can extend up to 2 or 3 h, which puts a demand on fuel needs. Not allowing for proper ventilating of air can also extend the start-up time. In addition, condensate return can be facilitated if it is allowed to drain by gravity into a tank or receiver and is then pumped into the boiler feed tank.

Proper management of condensate return. Proper management can lead to great savings. Lost feedwater must be made up and heated. For example, every 0.45 kg (1 lb) of steam that must be generated from 15°C feedwater instead of 70°C feedwater requires an additional 1.056×10^5 J (100 Btu) more than 1.12 MJ (1063 Btu) required or a 10% increase in fuel. A rule of thumb is that a 1% fuel saving results for every 5°C increase in feedwater temperature. Maximizing condensate recovery is an important fuel laying procedure.

Poorly insulated lines and valves due either to poor initial design or a deteriorated condition. Heat losses from a poorly insulated pipe can be costly. A poorly insulated line carrying steam at 400 psig can lose ~1000 GJ/year (10^9 Btu/year) or more per 30 m (100 ft.) of pipe. At steam costs of \$2.00/GJ, this translates to a \$2000 expense per year.

Improper operation and maintenance of tracing systems. Steam tracing is used to protect piping and equipment from cold weather freezing. The proper operation and maintenance of tracing systems will not only insure the protection of traced piping but also saves fuel. Occasionally these systems are operating when not required. Steam is often used in tracing systems and many of the deficiencies mentioned earlier apply (e.g., poorly operating valves, insulation, leaks).

Reduce losses in process hot water systems. Electrically heated hot water systems are used in many industrial processes for cleaning, pickling, coating, or etching components. Hot or cold water systems can dissipate energy. Leaks and poor insulation should be repaired.

14.4.3.7 Electrical Process Heat

Industrial process heat applications can be divided into four categories: direct-fired, indirect-fired, fuel, or electric. Here we shall consider electric direct-fired installations (ovens, furnaces) and indirect-fired (electric water heaters and boilers) applications. Electrical installations use metal sheath resistance heaters, resistance ovens or furnaces, electric salt bath furnaces, infrared heaters, induction and high-frequency resistance heaters, dielectric heaters, and direct arc furnaces. From the housekeeping and maintenance point of view, typical opportunities would include:

Repair or improve insulation. Operational and standby losses can be considerable, especially in larger units. Remember that insulation may degrade with time or may have been optimized to different economic criteria.

Provide finer controls. Excessive temperatures in process equipment waste energy. Run tests to determine the minimum temperatures that are acceptable, then test instrumentation to verify that it can provide accurate process control and regulation.

Practice heat recovery. This is an important method, applicable to many industrial processes as well as HVAC systems and so forth. It is described in more detail in the next section.

14.4.3.8 Heat Recovery

Exhaust gases from electric ovens and furnaces provide excellent opportunities for heat recovery. Depending on the exhaust-gas temperature, exhaust heat can be used to raise steam or to preheat air or feedstocks. Another potential source of waste-heat recovery is the exhaust air that must be rejected from industrial operations in order to maintain health and ventilation safety standards. If the reject air has been subjected to heating and cooling processes, it represents an energy loss inasmuch as the makeup air must be modified to meet the interior conditions. One way to reduce this waste is through the use of heat wheels or similar heat exchange systems.

Energy in the form of heat is available at a variety of sources in industrial operations, many of which are not normally derived from primary heat sources. Such sources include electric motors, crushing and grinding operations, air compressors, and drying processes. These units require cooling in order to maintain proper operation. The heat from these systems can be collected and transferred to some appropriate use such as space heating or water heating.

The heat pipe is gaining wider acceptance for specialized and demanding heat transfer applications. The transfer of energy between incoming and outgoing air can be accomplished by banks of these devices. A refrigerant and a capillary wick are permanently sealed inside a metal tube, setting up a liquid-to-vapor circulation path. Thermal energy applied to either end of the pipe causes the refrigerant to vaporize. The refrigerant vapor then travels to the other end of the pipe, where thermal energy is removed. This causes the vapor to condense into liquid again, and the condensed liquid then flows back to the opposite end through the capillary wick.

Industrial operations involving fluid flow systems that transport heat such as in chemical and refinery operations offer many opportunities for heat recovery. With proper design and sequencing of heat exchangers, the incoming product can be heated with various process steams. For example, proper heat exchanger sequence in preheating the feedstock to a distillation column can reduce the energy utilized in the process.

Many process and air-conditioning systems reject heat to the atmosphere by means of wet cooling towers. Poor operation can contribute to increased power requirements.

Water flow and airflow should be examined to see that they are not excessive. The cooling tower outlet temperature is fixed by atmospheric conditions if operating at design capacity. Increasing the water flow rate or the airflow will not lower the outlet temperature.

The possibility of utilizing heat that is rejected to the cooling tower for other purposes should be investigated. This includes preheating feedwater, heating hot water systems, space heating and other low-temperature applications. If there is a source of building exhaust air with a lower wet bulb temperature, it may be efficient to supply this to a cooling tower.

14.4.3.9 Power Recovery

Power recovery concepts are an extension of the heat recovery concept described earlier. Many industrial processes have pressurized liquid and gaseous streams at 150°C–375°C (300°F–700°F) that present excellent opportunities for power recovery. In many cases high-pressure process stream energy is lost by throttling across a control valve.

The extraction of work from high-pressure liquid streams can be accomplished by means of hydraulic turbines (essentially diffuser-type or volute-type pumps running backward). These pumps can be either single or multistage. Power recovery ranges from 170 to 1340 kW (230–1800 hp). The lower limit of power recovery approaches the minimum economically justified for capital expenditures at present power costs.

14.4.3.10 Heating, Ventilating, and Air-Conditioning Operation

The environmental needs in an industrial operation can be quite different from those in a residential or commercial structure. In some cases strict environmental standards must be met for a specific function or process. More often the environmental requirements for the process itself are not severe; however, conditioning of the space is necessary for the comfort of operating personnel, and thus large volumes of air must be processed. Quite often opportunities exist in the industrial operation where surplus energy can be utilized in environmental conditioning. A few suggestions follow:

Review HVAC controls. Building heating and cooling controls should be examined and preset.

Ventilation, air, and building exhaust requirements should be examined. A reduction of airflow will result in a savings of electrical energy delivered to motor drives and additionally reduce the energy requirements for space heating and cooling. Because pumping power varies as the cube of the airflow rate, substantial savings can be achieved by reducing airflows where possible.

Do not condition spaces needlessly. Review air-conditioning and heating operations, seal off sections of plant operations that do not require environmental conditioning, and use air-conditioning equipment only when needed. During nonworking hours the environmental control equipment should be shut down or reduced. Automatic timers can be effective.

Provide proper equipment maintenance. Insure that all equipment is operating efficiently. (Filters, fan belts, and bearings should be in good condition.)

Use only equipment capacity needed. When multiple units are available, examine the operating efficiency of each unit and put operations in sequence in order to maximize overall efficiency.

Recirculate conditioned (heated or cooled) air where feasible. If this cannot be done, perhaps exhaust air can be used as supply air to certain processes (e.g., a paint spraybooth) to reduce the volume of air that must be conditioned.

For additional energy management opportunities in HVAC systems, see Chapter 11.

14.4.3.11 Lighting

Industrial lighting needs range from low-level requirements for assembly and welding of large structures (such as shipyards) to the high levels needed for manufacture of precision mechanical and electronic components (e.g., integrated circuits). There are four basic housekeeping checks that should be made:

Is a more efficient lighting application possible? Remove excessive or unnecessary lamps.

Is relamping possible? Install lower-wattage lamps during routine maintenance.

Will cleaning improve light output? Fixtures, lamps, and lenses should be cleansed periodically.

Can better controls be devised? Eliminate turning on more lamps than necessary. For modification, retrofit, or new design, consideration should be given to the spectrum of high-efficiency lamps and luminaires that are available. For example, high-pressure sodium lamps are finding increasing acceptance for industrial use, with savings of nearly a factor of five compared to incandescent lamps. See Chapter 10 for additional details.

14.4.3.12 The New Electrotechnologies

Electricity has certain characteristics that make it uniquely suitable for industrial processes. These characteristics include electricity's suitability for timely and precise control; its ability to interact with materials at the molecular level; the ability to apply it selectively and specifically, and the ability to vary its frequency and wavelength so as to enhance or inhibit its interaction with materials. These aspects may be said to relate to the *quality* of electricity as an energy form. It is important to recognize that different forms of energy have different qualities in the sense of their ability to perform useful work. Thus, although the Btu content of two energy forms may be the same, their ability to transform materials may be quite different

New electrotechnologies based on the properties of electricity are now finding their way into modern manufacturing. In many cases the introduction of electricity reduces manufacturing costs, improves quality reduces pollution, or has other beneficial results. Some examples include

Microwave heating	Ion nitriding
Induction heating	Infrared drying
Plasma processing	UV drying and curing
Magnetic forming	Advanced finishes
RF drying and heating	Electron beam heating

Microwave heating is a familiar technology that exhibits the unique characteristics of electricity described earlier. First it is useful to review how conventional heating is preformed to dry paint, anneal a part, or remove water. A source of heat is required, along with a container (oven, furnace, pot, etc.) to which the heat is applied. Heat is transferred from the container to the work piece by conduction, radiation, convection, or a combination of these. There are certain irreversible losses associated with heat transfer in this process. Moreover, since the container must be heated, more energy is expended than is really required. Microwave heating avoids these losses due to the unique characteristics of electricity.

Timely control. There is no loss associated with the warmup or cooldown of ovens. The heat is applied directly when needed.

Molecular interaction. By interacting at the molecular level, heat is deposited directly in the material to be heated, without having to preheat an oven, saving the extra energy required for this purpose and avoiding the losses that result from heat leakage from the oven.

Selective application. By selectively applying heat only to the material to be heated, parasitic losses are avoided. In fact, the specificity of heat applied this way can improve quality by not heating other materials.

Selective wavelength and frequency. A microwave frequency is selected that permits the microwave energy to interact with the material to be heated, and not with other materials. Typically the frequency is greater than 2000 MHz.

Microwave heating was selected for this discussion, but similar comments could be made about infrared, ultraviolet (UV), dielectric, induction, or electron beam heating. In each case the frequency or other characteristics of the energy form are selected to provide the unique performance required.

Ultraviolet curing (now used for adhesive and finishes) is another example. The parts to be joined or coated can be prepared and the excessive adhesive removed without fear of prehardening. Then the UV energy is applied, causing the adhesive to harden.

Induction heating is another example. It is similar to microwave heating except that the energy is applied at a lower frequency. Induction heating operates on the principal of inducing electric currents to flow in materials, heating them by the power dissipated in the material. The method has several other advantages. In a conventional furnace, the work piece has to be in the furnace for a sufficient time to reach temperature. Because of this, some of the material is oxidized and lost as scale. In a typical high-temperature gas furnace this can be 2% of the throughput. Additional product is scrapped as a result of surface defects caused by uneven heating and cooling. This can amount to another 1% of throughput. Induction heating can reduce these losses by a factor of four.

The fact that electricity can be readily controlled and carries with it a high information content through digitization or frequency modulation also offers the potential for quantum improvements in efficiency. A slightly different example is the printing industry.

Today the old linotype technology has been replaced by electronic processes. The lead melting pots that used to operate continuously in every newspaper plant have been removed, eliminating a major energy use and an environmental hazard. Books, magazines, and newspapers can be written composed, and printed entirely by electronic means. Text is processed by computer techniques. Camera-ready art is prepared by computers directly or prepared photographically and then optically scanned to create digital images. The resulting electronic files can be used in web offset printing by an electronic photochemical process. The same information can be transmitted electronically; via satellite, to a receiving station at a remote location where a high-resolution fax machine reconstitutes the image. This method is being used to simultaneously and instantaneously distribute advertising copy to multiple newspapers, using a single original. Previously, to insert an advertisement in 25 newspapers, 25 sets of photographic originals would have to be prepared and delivered, by messenger, air express, or mail, to each newspaper.

Some of the other applications of the new electrotechnologies include RF drying of plywood veneers, textiles, and other materials; electric infrared drying for automobile paint and other finishes; electric resistance melting for high purity metals and scrap recovery; and laser cutting of wood, cloth, and other materials.

14.4.3.13 General Industrial Processes

The variety of industrial processes is so great that detailed specific recommendations are outside the scope of this chapter. Useful sources of information are found in trade journals, vendor technical bulletins, and manufacturers' association journals. These suggestions are intended to be representative, but by no means do they cover all possibilities.

In machining operations, eliminate unnecessary operations and reduce scrap. This is so fundamental from a purely economic point of view that it will not be possible to find significant improvements in many situations. The point is that each additional operation and each increment of scrap also represents a needless use of energy. Machining itself is not particularly energy-intensive. Even so, there are alternate technologies that can not only save energy but reduce material wastage as well. For example, powder metallurgy generates less scrap and is efficient if done in induction-type furnaces.

Use stretch forming. Forming operations are more efficient if stretch forming is used. In this process sheet metal or extrusions are stretched 2%–3% prior to forming, which makes the material more ductile so that less energy is required to form the product. The finished part

Use alternate heat treating methods. Conventional heat treating methods such as carburizing are energy-intensive. Alternate approaches are possible. For example, a hard surface can be produced by induction heating, which is a more efficient energy process. Plating, metallizing, flame spraying, or cladding can substitute for carburizing, although they do not duplicate the fatigue strengthening compressive skin of carburization or induction hardening.

Use alternative painting methods. Conventional techniques using solvent-based paints require drying and curing at elevated temperature. Powder coating is a substitute process in which no solvents are used. Powder particles are electrostatically charged and attracted to the part being painted so that only a small amount of paint leaving the spray gun misses the part and the overspray is recoverable. The parts can be cured rapidly in infrared ovens, which require less energy than standard hot air systems. Water-based paints and high-solid coatings are also being used and are less costly than solvent-based paints. They use essentially the same equipment as the conventional solvent paint spray systems so that the conversion can be made at minimum costs. New water-based emulsion paints contain only 5% organic solvent and require no afterburning. High-solids coatings are already in use commercially for shelves, household fixtures, furniture, and beverage cans, and require no afterburning. They can be as durable as conventional finishes and are cured by either conventional baking or UV exposure.

Substitute for energy-intensive processes such as hot forging. Hot forging may require a part to go through several heat treatments. Cold forging with easily wrought alloys may offer a replacement. Lowering the preheat temperatures may also be an opportunity for savings. Squeeze forging is a relatively new process in which molten metal is poured into the forging dye. The process is nearly scrap free, requires less press power, and promises to contribute to more efficient energy utilization.

Movement of materials through the plant creates opportunities for saving energy. Material transport energy can be reduced by

Combining processes or relocate machinery to reduce transport energy. Sometimes merely relocating equipment can reduce the need to haul materials.

Turning off conveyors and other transport equipment when not needed. Look for opportunities where controls can be modified to permit shutting down of equipment not in use.

Using gravity feeds wherever possible. Avoid unnecessary lifting and lowering of products.

14.4.3.14 Demand Management

The cost of electrical energy for medium to large industrial and commercial customers generally consists of two components. One component is the *energy charge*, which is based on the cost of fuel to the utility, the average cost of amortizing the utility generating plant, and on the operating and maintenance costs experienced by the utility. Energy costs for industrial users in the United States are typically in the range of 0.05–0.10 $/kWh.

The second component in the *demand charge*, which reflects the investment cost the utility must make to serve the customer. Besides the installed generating capacity needed, the utility also provides distribution lines, transformers, substations, and switches whose cost depends on the size of the load being served. This cost is recovered in a demand charge which typically is 2–10 $/kW month.

Demand charges typically account for 10%–50% of the bill, although wide variations are possible depending on the type of installation. Arc welders, for example, have relatively high demand charges, since the installed capacity is great (10–30 kW for a typical industrial machine) and the energy use is low.

From the utility's point of view it is advantageous to have its installed generating capacity operating at full load as much of the time as possible. To follow load variations conveniently, the utility operates its largest and most economical generating units continuously to meet its base load, and then brings smaller (and generally more expensive) generating units on line to meet peak load needs.

Today consideration is being given to time-of-day or *peak load pricing* as a means of assigning the cost of operating peak generating capacity to those loads that require it. See Chapter 5 for a discussion of utility pricing strategies. From the viewpoint of the utility, *load management* implies maintaining a high-capacity factor and minimizing peak load demands. From the customer's viewpoint, *demand management* means minimizing electrical demands (both on- and off-peaks) so as to minimize overall electricity costs.

Utilities are experimenting with several techniques for load management. Besides rate schedules that encourage the most effective use of power, some utilities have installed remotely operated switches that permit the utility to disconnect nonessential parts of the customer's load when demand is excessive. These switches are actuated by a radio signal, through the telephone lines, of over the power grid itself through a harmonic signal (ripple frequency) that is introduced into the grid.

Customers can control the demand of their loads by any of several methods:

- Manually switching off loads ("load shedding")
- Use of timers and interlocks to prevent several large loads from operating simultaneously
- Use of controllers and computers to control loads and minimize peak demand by scheduling equipment operation
- Energy storage (e.g., producing hot or chilled water during off-peak hours and storing it for use on-peak)

Demand can be monitored manually (by reading a meter) or automatically using utility-installed equipment or customer-owned equipment installed in parallel with the utility meter. For automatic monitoring, the basic approach involves pulse counting.

The demand meter produces electronic pulses, the number of which is proportional to demand n kW. Demand is usually averaged over some interval (e.g., 15 min) for calculating cost. By monitoring the pulse rate electronically, a computer can project what the demand will be during the next demand measurement interval, and can then follow a preestablished plan for shedding loads if the demand set point is likely to be exceeded.

Computer control can assist in the dispatching of power supply to the fluctuating demands of plant facilities. Large, electrically-based facilities are capable of forcing large power demands during peak times that exceed the limits contracted with the utility or cause penalties in increased costs. Computer control can even out the load by shaving peaks and filling in the valleys, thus minimizing power costs. In times of emergency or fuel curtailment, operation of the plant can be programmed to provide optimum production and operating performance under prevailing conditions. Furthermore, computer monitoring and control provide accurate and continuous records of plant performance.

It should be stressed here that many of these same functions can be carried out by manual controls, time clocks, microprocessors, or other inexpensive devices. Selection of a computer system must be justified economically on the basis of the number of parameters to be controlled and the level of sophistication required. Many of the benefits described here can be obtained in some types of operations without the expense of a computer.

14.5 Thermal Energy Management in Industry

14.5.1 The Importance of Fuel Use and Heat in Industry

The U.S. manufacturing sector depends heavily on fuels for the conversion of raw materials into usable products. Industry uses a wide range of fuels, including natural gas, petroleum, coal, and renewables. The petroleum forms used include distillate fuel oil, residual fuel oil, gasoline, LPG, and others. How efficiently energy is used, its cost, and its availability consequently have a substantial impact on the competitiveness and economic health of U.S. manufacturers. More efficient use of fuels lowers production costs, conserves limited energy resources, and increases productivity. Efficient use of energy also has positive impacts on the environment—reductions in fuel use translate directly into decreased emissions of pollutants such as sulfur oxides, nitrogen oxides, particulates, and greenhouse gases (e.g., carbon dioxide).

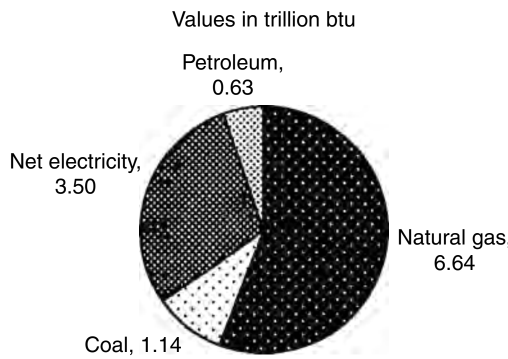


FIGURE 14.10 Manufacturing energy consumption by fuel (end use data). (From Manufacturing Energy Consumption Survey, 1998; U.S. Department of Energy; Energy Information Agency, 2002.)

From Figure 14.10, it can be seen that fuel use in manufacturing is just over 70% of the total energy used in manufacturing on an end use basis. Figure 14.11 shows the percentage of each fuel used for boiler fuel and process heat combined. Of this total, 53% is used for boiler fuel, and 47% for direct process heating. Thus, there is a huge potential for energy management and energy-efficiency improvement related to the use fuels and thermal energy in industry. Section 14.5 discusses the use of improved and new equipment and technology to accomplish some of these reductions in energy use and cost.

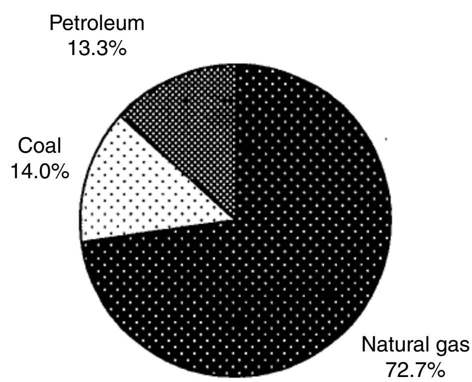


FIGURE 14.11 Manufacturing energy use for boiler fuel and process heating end use basis. (From Manufacturing Energy Consumption Survey, 1998; U.S. Department of Energy; Energy Information Agency, 2002.)

Energy efficiency can be defined as the effectiveness with which energy resources are converted into usable work. Thermal efficiency is commonly used to measure the efficiency of energy conversion systems such as process heaters, steam systems, engines, and power generators. Thermal efficiency is essentially the measure of the efficiency and completeness of fuel combustion, or, in more technical terms, the ratio of the net work supplied to the heat supplied by the combusted fuel. In a gas-fired heater, for example, thermal efficiency is equal to the total heat absorbed divided by the total heat supplied; in an automotive engine, thermal efficiency is the work done by the gases in the cylinder divided by the heat energy of the fuel supplied.[1]

Energy efficiency varies dramatically across industries and manufacturing processes, and even between plants manufacturing the same products. Efficiency can be limited by mechanical, chemical, or other physical parameters, or by the age and design of equipment. In some cases, operating and maintenance practices contribute to lower-than-optimum efficiency. Regardless of the reason, less-than-optimum energy efficiency implies that not all of the energy input is being converted to useful work—some is released as lost energy. In the manufacturing sector, these energy losses amount to several quadrillion BTUs (quadrillion British thermal units, or quads) and billions of dollars in lost revenues every year.

Typical thermal efficiencies of selected energy systems and industrial equipment[25] is provided in the following table:

Power generation	25%–44%
Steam boilers (natural gas)	80%
Steam boilers (coal and oil)	84%–85%
Waste-heat boilers	60%–70%
Thermal cracking (refineries)	58%–61%
EAF steelmaking	56%
Paper drying	48%
Kraft pulping	60%–69%
Distillation column	25%–40%
Cement calciner	30%–70%
Compressors	10%–20%
Pumps and fans	55%–65%
Motors	90%–95%

Boiler losses represent energy lost due to boiler inefficiency. In practice, boiler efficiency can be as low as 55%–60%, or as high as 90%. The age of the boiler, maintenance practices, and fuel type are contributing factors to boiler efficiency. It is assumed that the greater losses are in steam pipes (20%), with small losses incurred in other fuel transmission lines (3%) and electricity transmission lines (3%). Losses in steam pipes and traps have been reported to be as high as from 20%–40%. A conservative value of 20% was used for steam distribution losses in this study.[1]

14.5.2 Boiler Combustion Efficiency Improvement

Boilers and other fuel-fired equipment, such as ovens and kilns, combust fuel with air for the purpose of releasing chemical energy as heat. For an industrial boiler, the purpose is to generate high-temperature and high-pressure steam to use directly in a manufacturing process, or to operate other equipment such as steam turbines to produce shaft power. As shown earlier in Figure 14.11, the predominant boiler fuel is natural gas. The efficiency of any combustion process is dependent on the amount of air that is used in relation to the amount of fuel and how they are mixed. Air is about 20% oxygen, so approximately five units of air must be brought in to the boiler for every one unit of oxygen that is needed. Controlling this air–fuel mixture, and minimizing the amount of excess air while still obtaining safe mixing of the air and fuel, is key to insuring a high combustion efficiency in the boiler.[26]

14.5.2.1 Combustion Control

The stoichiometric equation for the combustion of methane, the principal constituent of natural gas, with air is

$$CH_4 + 2O_2 + 7.52N_2 \rightarrow CO_2 + 2H_2O + 7.52N_2 \tag{14.16}$$

The stoichiometric equation is the one representing the exact amount of air necessary to oxidize the carbon and hydrogen in the fuel to carbon dioxide and water vapor. However, it is necessary to provide more than the stoichiometric amount of air since the mixing of fuel and air is imperfect in the real combustion chamber. Thus the combustion equation for hydrocarbon fuels becomes

$$\begin{aligned} C_xH_y + \phi(x + y/4)O_2 + 3.76\phi(x + y/4)N_2 &\rightarrow xCO_2 + y/2H_2O \\ &+ (\phi - 1)(x + y/4)O_2 + 3.76\phi(x + y/4)N_2 \end{aligned} \tag{14.17}$$

Note that for a given fuel nothing in the equation changes except the parameter ϕ, the equivalence ratio, as the fuel–air ratio changes.

As ϕ is increased beyond the optimal value for good combustion, the stack losses increase and the heat available for the process decreases. As the equivalence ratio increases for a given flue temperature and a given fuel, more fuel must be consumed to supply a given amount of heat to the process.

The control problem for the furnace or boiler is to provide the minimum amount of air for good combustion over a wide range of firing conditions and a wide range of ambient temperatures. The most common combustion controller uses the ratio of the pressure drops across orifices, nozzles, or Venturis in the air and fuel lines. Since these meters measure volume flow, a change in temperature of combustion air with respect to fuel, or vice versa, will affect the equivalence ratio of the burner. Furthermore, since the pressure drops across the flow meters are exponentially related to the volume flow rates control dampers must have very complicated actuator motions. All the problems of ratio controllers are eliminated if the air is controlled from an oxygen meter. These are now coming into more general use as reasonably priced, high-temperature oxygen sensors become available. It is possible to control to any set value of percentage oxygen in the products; that is,

$$\%\ O_2 = \frac{\phi - 1}{\frac{x+y/2}{x+y/4} + 4.76\ \phi - 1} \tag{14.18}$$

Figure 14.12 is a nomograph from the Bailey Meter Company (Dukelow, 1974) that gives estimates of the annual dollar savings resulting from the reduction of excess air to 15% for gas-, oil-, or coal-fired boilers with stack temperatures from 300°F to 700°F. The fuel savings are predicted on the basis that as excess air is reduced, the resulting reduction in mass flow of combustion gases results in reduced gas velocity and thus a longer gas residence time in the boiler. The increased residence time increases the heat transfer from the gases to the water. The combined effect of lower exhaust-gas flows and increased heat exchange effectiveness is estimated to be 1.5 times greater than that due to the reduced mass flow alone.

As an example assume the following data pertaining to an oil-fired boiler. Entering the graph at the top abscissa with 6.2% O_2, we drop to the oil fuel line and then horizontally to the 327°C (620°F) flue gas temperature line. Continuing to the left ordinate we can see that 6.2% O_2 corresponds to 37.5% excess air. Dropping vertically from the intersection of the flue gas temperature line and the excess air line we note a 3.4% total fuel savings. Fuel costs are

Burner capacity	63 GJ/hr (60×10^6 Btu/hr)
Annual operating hours	6,200
Fuel cost	\$0.38/L (\$1.44/gal)
Heating value fuel	42.36 MJ/L (152,000 Btu/gal)
Percent O_2 in exhaust gases	6.2%
Stack temperature	327°C (620°F)

$$10^9\text{J/GJ}(10^6\text{Btu/million Btu}) \times \frac{\$0.38/\text{L}}{42,365,00\text{J/L}}\left(\frac{\$1.44/\text{gal}}{152,000\ \text{Btu/gal}}\right) = \$8.96/\text{GJ}(\$9.48/\text{million Btu})$$

Continuing the vertical line to intersect the \$2.50/million Btu and then moving to left ordinate shows a savings of \$140,000 per year for 8000 h of operation, 100×10^6 Btu/hr input and \$5.00/million Btu fuel

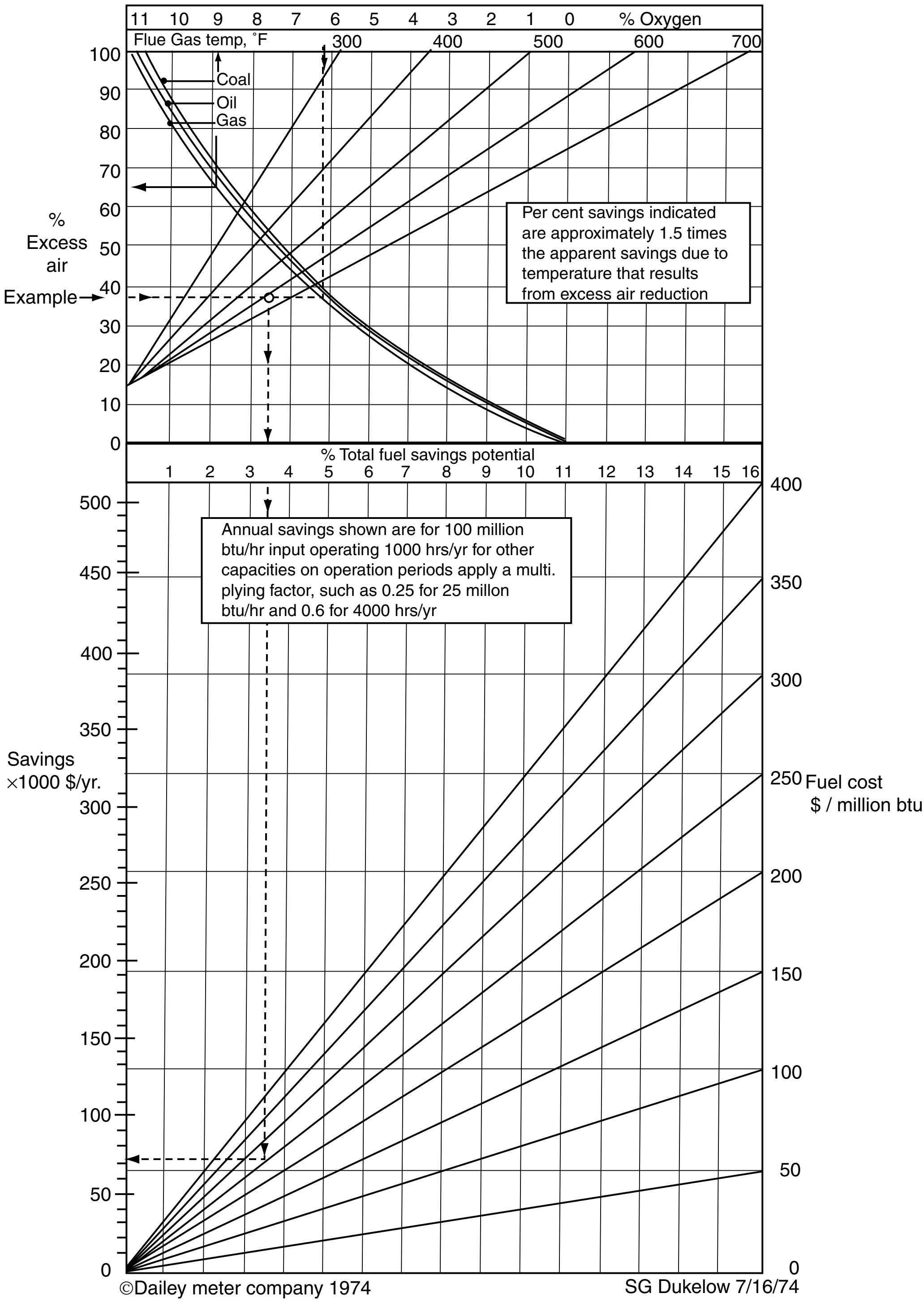


FIGURE 14.12 Nomograph for estimating savings from adjustment of burners. (From Dukelow, S. G., Bailey Meter Company, 29801 Euclid Avenue, Wickliffe, Ohio, 1974. With permission.)

cost. Adjusting that result for the assumed operating data:

$$\text{Annual savings} = \frac{6,200}{8,000} \times \frac{60 \times 10^6}{100 \times 10^6} \times \frac{9.48}{5.00} \times \ \$140,000 = \ \$123,430 \text{ per year}$$

This savings could be obtained by installing a modern oxygen controller, an investment with approximately a 1-year payoff, or from heightened operator attention with frequent flue gas testing and manual adjustments. Valuable sources of information concerning fuel conservation in boilers and furnaces are given by the DOE/OIT website.

14.5.2.2 Waste-Heat Management

Waste heat as generally understood in industry is the energy rejected from any process to the atmospheric air or to a body of water. It may be transmitted by radiation, conduction, or convection, but often it is contained in gaseous or liquid streams emitted to the environment. Almost 50% of all fuel energy used in the United States is transferred as waste heat to the environment, causing thermal pollution as well as chemical pollution of one sort or another. It has been estimated that half of that total may be economically recoverable for useful heating functions.

What must be known about waste-heat streams in order to decide whether they can become useful? Here is a list along with a parallel list of characteristics of the heat load that should be matched by the waste-heat supply.

Waste-Heat Supply
Quantity
Quality
Temporal availability of supply
Characteristics of fluid
Heat Load
Quantity required
Quality required
Temporal availability of load
Special fluid requirements

Let us examine the particular case of a plant producing ice-cream cones. All energy quantities are given in terms of 15.5°C reference temperature. Sources of waste heat include

- Products of combustion from 120 natural gas jets used to heat the molds in the carousel-type baking machines. The stack gases are collected under an insulated hood and released to the atmosphere through a short stack. Each of six machines emits 236.2 m^3/hr of stack gas at 160°C. Total source rate is 161,400 kJ/hr or 3874 MJ/day for a three-shift day.
- Cooling water from the jackets, intercoolers, and aftercoolers of two air compressors used to supply air to the pneumatic actuators of the cone machines; 11.36 L/min of water at 48.9°C is available. This represents a source rate of 96 MJ/hr. The compressors run an average of 21 h per production day. Thus this source rate is 2015 MJ/day.
- The water chillers used to refrigerate the cone batter make available—at 130°F—264 MJ/hr of water heat. This source is available to heat water to 48.9°C using desuperheaters following the water chiller compressors. The source rate is 6330 MJ/day.
- 226 m^3/min of ventilating air is discharged to the atmosphere at 21.2°C. This is a source of rate of less than 22.2 MJ/hr or 525 MJ/day.

Uses for waste heat include

- 681 L/hr of hot water at 82.2°C is needed for cleanup operations during 3 h of every shift or during 9 of every 24 h. Total daily heat load is 4518 MJ.

- Heating degree-days total in excess of 3333 annually. Thus any heat available at temperatures above 21.1°C can be utilized with the aid of runaround systems during the 5(1/2)-month heating season. Estimated heating load per year is 4010 GJ.
- Total daily waste heat available—12.74 GJ/day
- Total annual waste heat available—3.19 TJ/year
- Total annual worth of waste heat (at $5.00/GJ for gas)—$15,893
- Total daily heat load—this varies from a maximum of 59.45 GJ/day at the height of the heating season to the hot-water load of 4.52 GJ/day in the summer months.

Although the amount of waste heat from the water chillers is 40% greater than the load needed for hot-water heating, the quality is insufficient to allow its full use, since the hot water must be heated to 82°C and the compressor discharge is at a temperature of 54°C.

However the chiller waste heat can be used to preheat the hot water. Assuming 13°C supply water and a 10° heat exchange temperature approach, the load that can be supplied by the chiller is

$$\frac{49-13}{82-13}\times 4.52 = 2.36\ \text{GJ/day}$$

Since the cone machines have an exhaust-gas discharge of 3.87 GJ/day at 160°C, the remainder of the hot-water heating load of 2.17 GJ/day is available. Thus a total saving of 1129 GJ/year in fuel is possible with a cost saving of $5645 annually based on $5.00/GJ gas. The investment costs will involve the construction of a common exhaust heater for the cone machines, a desuperheater for each of the three water chiller compressors, a gas-to-liquid heat exchanger following the cone-machine exhaust heater, and possibly an induced draft fan following the heat exchanger, since the drop in exhaust-gas temperature will decrease the natural draft due to hot-gas buoyancy.

It is necessary to almost match four of four characteristics. Not exactly, of course, but the closer means thermodynamic availability of the waste heat. Unless the energy of the waste stream is sufficiently hot, it will be impossible to even transfer it to the heat load, since spontaneous heat transfer occurs only from higher to lower temperature.

The quantity and quality of energy available from a waste-heat source or for a heat load are studied with the aid of a heat balance. Figure 14.13 shows the heat balance for a steam boiler. The rates of enthalpy entering or leaving the system fluid streams must balance with the radiation loss rate from the boiler's external surfaces. Writing the First-Law equation for a steady-flow-steady-state process

$$\dot{q}_L = \dot{m}_f h_f + \dot{m}_a h_a + \dot{m}_c h_c - \dot{m}_s h_s - \dot{m}_g h_g \tag{14.19}$$

and referring to the heat-balance diagram, one sees that the enthalpy flux $\dot{m}_g h_g$, leaving the boiler in the exhaust-gas stream, is a possible source of waste heat. A fraction of that energy can be transferred in a heat exchanger to the combustion air, thus increasing the enthalpy flux $\dot{m}_g h_a$ and reducing the amount of fuel required. The fraction of fuel that can be saved is given in the equation

$$\frac{\dot{m}_f - \dot{m}_{f'}}{\dot{m}_f} = 1 - \left[\frac{K_1 - (1+\phi)\bar{C}_p T_g}{K_1 - (1+\phi')\bar{C}'_p T'_g}\right] \tag{14.20}$$

where the primed values are those obtained with waste-heat recovery. K_1 represents the specific enthalpy of the fuel–air mixture, $h_f + \phi h_a$, which is presumed to be the same with or without waste-heat recovery, ϕ is the molar ratio of air to fuel, and $\bar{C}_p$ is the specific heat averaged over the exhaust-gas components. Figure 14.14, which is derived from Equation 14.20, gives possible fuel savings from using high-temperature flue gas to heat the combustion air in industrial furnaces.

It should be pointed out that the use of recovered waste heat to preheat combustion air, boiler feedwater, and product to be heat treated, confers special benefits not necessarily accruing when the heat

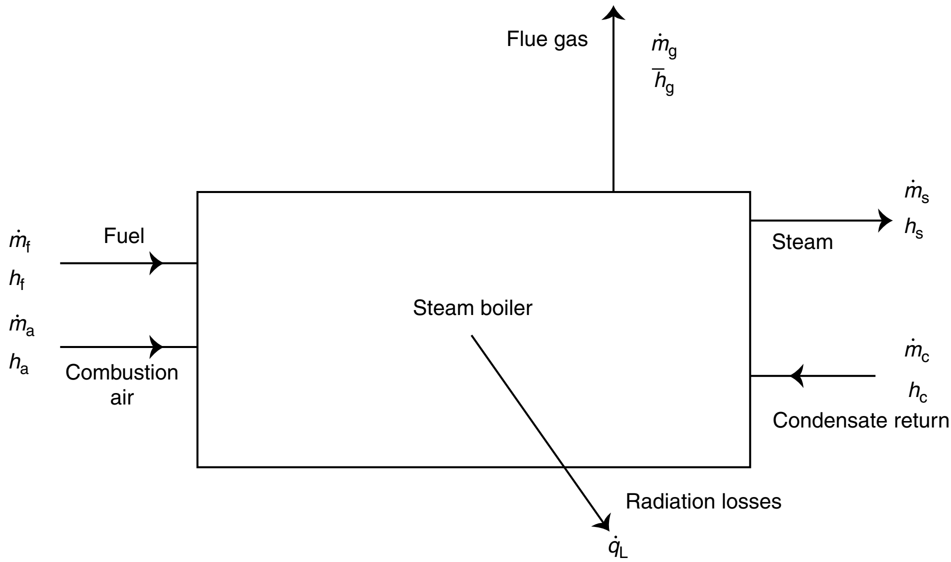


FIGURE 14.13 Heat balance on steam boiler.

is recovered to be used in another system. The preheating operation results in less fuel being consumed in the furnace, and the corresponding smaller air consumption means even smaller waste heat being ejected from the stacks.

Table 14.8 shows heat balances for a boiler with no flue gas–heat recovery, with a feedwater preheater (economizer) installed and with an air preheater, respectively.

It is seen that air preheater alone saves 6% of the fuel and the economizer saves 9.2%. Since the economizer is cheaper to install than the air preheater, the choice is easy to make for an industrial boiler. For a utility boiler, both units are invariably used in series in the exit gas stream.

Table 14.9 is an economic study using 2005 prices for fuel, labor, and equipment. At that time it was estimated that a radiation recuperator fitted to a fiberglass furnace would cost $820,200 and effect a

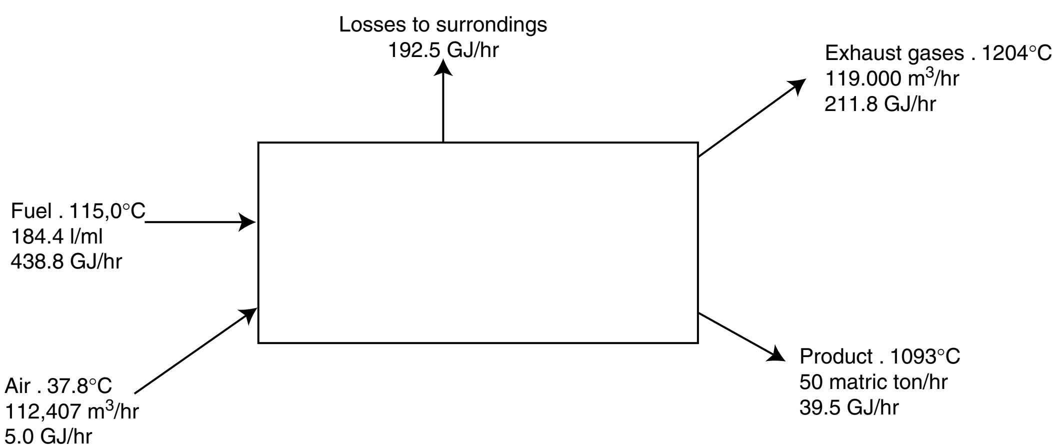


FIGURE 14.14 Heat balance for a simple continuous steel tube furnace.

TABLE 14.8 Heat Balances for a Steam Generator

Case	Input Streams				Output Streams			
	Name	Temperature (°C)	Flow Rate	Energy	Name	Temperature (°C)	Flow Rate	Energy
Without economizer or air preheater	Natural gas	26.7	3611 m^3/hr	134.78 GJ/hr	Steam	185.6	45,349 kg/hr	126.23
	Air	26.7	35,574 m^3/hr	560.32 MJ/hr	Flue gas	372.2	41,185 m^3/hr	19.49
	Makeup water	10.0	7076 kg/hr	297.25 MJ/hr	Surface losses	—	—	1.21
	Condensate return	82.2	41,277 kg/hr	14.21 GJ/hr	Blow down	185.6	2994 kg/hr	2.36
With air preheater	Natural gas	26.7	3395 m^3/hr	126.72 GJ/hr	Steam	185.6	45,359 kg/hr	126.23
	Air	232.2	32,114 m^3/hr	478.90 GJ/hr	Flue gas	260.0	35,509 m^3/hr	11.42
	Makeup water	10	70,767 kg/hr	297.25 MJ/hr	Surface losses	—	—	1.21
	Condensate return	82.2	41,277 kg/hr	14.21 GJ/hr	Blow down	185.6	2994 kg/hr	2.36
With economizer	Natural gas	26.7	3278 m^3/hr	122.37 GJ/hr	Steam	185.6	45,359 kg/hr	126.23
	Air	26.7	31,013 m^3/hr	462.48 MJ/hr	Flue gas	176.7	34,281 m^3/hr	7.08
	Makeup water	101	7076 kg/hr	297.25 GJ/hr	Surface loss	—	—	1.21
	Condensate water	82.2	41,277 kg/hr	12.21 GJ/hr	Blow down	185.6	2994 kg/hr	2.36

TABLE 14.9 Cost-Fuel Savings Analysis of a Fiberglass Furnace Recuperator

Operation	Continuous
Fuel input	19.42 GJ/hr
Fuel	No. 3 fuel oil
Furnace temperature	1482°C
Flue gas temperature entering recuperator	1204°C
Air preheat (at burner)	552°C
Fuel savings = 37.4%	
Q = 7.26 GJ/hr or 173.6 L/hr of oil	
Fuel cost savings estimation	
per GJ = $5.00	
per hour = $36.30	
per year (8,000 h) = $290,435	
Cost of recuperator	$421,400
Cost of installation, related to recuperator	$398,800
Total cost of recuperator installation	$820,200
Approximate payback time	2.82 yrs

savings of $90,435/year, making for a payoff period of approximately 2.8 years. This assumes of course that the original burners, combustion-control system, and so on, could be used without modification.

At this point, we can take some time to relate waste-heat recovery to the combustion process itself. We can first state categorically that the use of preheated combustion air improves combustion conditions and efficiency at all loads, and in a newly designed installation permits a reduction in size of the boiler or furnace. It is true that the increased mixture temperature that accompanies air preheat results in some narrowing of the mixture stability limits, but in all practical furnaces this is of small importance.

In many cases low-temperature burners may be used for preheated air, particularly if the air preheat temperature is not excessive and if the fuel is gaseous. However, the higher volume flow of preheated air in the air piping may cause large enough increases in pressure drop to require a larger combustion air fan. Of course larger diameter air piping can prevent the increased pressure drop, since air preheating results in reduced quantities of fuel and combustion air. For high preheat temperatures alloy piping, high-temperature insulation, and water-cooled burners may be required. Since many automatic combustion-control systems sense volume flows of air and/or fuel, the correct control settings will change when preheated air is used. Furthermore, if air preheat temperature varies with furnace load, then the control system must compensate for this variation with an auxiliary temperature-sensing control. On the other hand, if control is based on the oxygen content of the flue gases, the control complications arising from gas volume variation with temperature is obviated. This is the preferred control system for all furnaces, and only cost prevents its wide use in small installations. Burner operation and maintenance for gas burners is not affected by preheating, but oil burners may experience accelerated fuel coking and resulting plugging from the additional heat being introduced into the liquid fuel from the preheated air. Careful burner design, which may call for water cooling or for shielding the fuel tip from furnace radiation, will always solve the problems. Coal-fired furnaces may use preheated air up to temperatures that endanger the grates or burners. Again, any problems can be solved by special designs involving water cooling if higher air temperatures can be obtained and/or desired.

The economics of waste-heat recovery today range from poor to excellent depending upon the technical aspects of the application as detailed earlier, but the general statement can be made that, at least for most small industrial boilers and furnaces, standard designs and/or off-the-shelf heat exchangers prove to be the most economic. For large systems one can often afford to pay for special designs, construction, and installations. Furthermore, the applications are often technically constrained by material properties and space limitations, and as shall be seen later, always by economic considerations. For further information on heat exchangers see Chapter 13.

14.5.2.3 Heating, Ventilating, and Air-Conditioning

Heating, ventilating, and air-conditioning, while not usually important in the energy-intensive industries, may be responsible for the major share of energy consumption in the light manufacturing field, particularly in high-technology companies and those engaged primarily in assembly.

Because of air pollution from industrial processes, many HVAC systems require 100% outside ventilating air. Furthermore, ventilating air requirements are often much in excess of those in residential and commercial practice (Hayashi et al. 1985). An approximate method for calculating the total heat required for ventilating air in kJ per heating season is given by

$$\begin{aligned} E_{\mathrm{v}}(\mathrm{kJ}) &= 60 \times 24\left(\frac{\mathrm{min}}{\mathrm{day}}\right) \times (1.2 \times 0.519)\left(\frac{\mathrm{kJ}}{\mathrm{m}^3 - \mathrm{K}^\circ}\right) \times \mathrm{SCMM} \times \mathrm{DD} \\ &= 896.8 \times \mathrm{SCMM} \times \mathrm{DD} \end{aligned} \tag{14.21}$$

where SCMM = standard cubic meter per minute of total air entering plant including unwanted infiltration; DD = heating degree-days (C).

This underestimates the energy requirement, because degree-days are based on 18.33°C reference temperature and indoor temperatures are ordinarily held 1.6–3.9° higher. For a location with 3333 degree-days each year the heating energy given by Equation 14.13 is about 17% low.

Savings can be effected by reducing the ventilating air rate to the actual rate necessary for health and safety and by ducting outside air into direct-fired heating equipment such as furnaces, boilers, ovens, and dryers. Air infiltration should be prevented through a program of building maintenance to replace broken windows, doors, roofs, and siding, and by campaigns to prevent unnecessary opening of windows and doors.

Additional roof insulation is often economic, particularly because thermal stratification makes roof temperatures much higher than average wall temperatures. Properly installed vertical air circulators can prevent the vertical stratification and save heat loss through the roof. Windows can be double glazed, storm windows can be installed, or windows can be covered with insulation. Although the benefits of natural lighting are eliminated by this measure, it can be very effective in reducing infiltration and heat transfer losses.

Waste heat from ventilating air itself, from boiler and furnace exhaust stacks, and from air-conditioning refrigeration compressors can be recovered and used to preheat makeup air. Consideration should also be given to providing spot ventilation in hazardous locations instead of increasing general ventilation air requirements.

As an example of the savings possible in ventilation air control, a plant requiring 424.5 CMM outside airflow is selected. A gas-fired boiler with an energy input of 0.0165 GJ is used for heating and is supplied with room air.

$$\text{Combustion air} = \frac{16.5 \times 10^6}{37,281} \times \frac{12}{60} = 88.52 \text{ CMM}$$

for a fuel with 37,281 kJ/m^3 heating value and an air fuel ratio of 12 m^3 air per m^3 fuel. The number of annual degree-days was 3175.

A study showed that the actual air supplied through infiltration and air handlers was 809 m^3/min. An outside air duct was installed to supply combustion air for the boiler and the actual ventilating air supply was reduced to the required 424 m^3/min. The fuel saving that resulted using Equation 14.21 was

$$896.8(809 - 424)\ 3,175 = 1,096.2 \text{ GJ}$$

worth $5482 in fuel at $5.00/GJ for natural gas.

14.5.2.4 Modifications of Unit Processes

The particular process used for the production of any item affects not only the cost of production but also the quality of the product. Since the quality of the product is critical in customer acceptance and therefore in sales, the unit process itself cannot be considered a prime target for the energy conservation program. That does not say that one should ignore the serendipitous discovery of a better and cheaper way of producing something. Indeed, one should take instant advantage of such a situation, but that clearly is the kind of decision that management could make without considering energy conservation at all.

14.5.2.5 Optimizing Process Scheduling

Industrial thermal processing equipment tends to be quite massive compared to the product treated. Therefore, the heat required to bring the equipment to steady-state production conditions may be large enough to make start-ups fuel intensive. This calls for scheduling this equipment so that it is in use for as long periods as can be practically scheduled. It also may call for idling the equipment (feeding fuel to keep the temperature close to production temperature) when it is temporarily out of use. The fuel rate for idling may be between 10 and 40% of the full production rate for direct-fired equipment. Furthermore, the stack losses tend to increase as a percentage of fuel energy released. It is clear that overfrequent start-ups and long idling times are wasteful of energy and add to production costs. The hazards of eliminating some of that waste through rescheduling must not be taken lightly. For instance, a holdup in an intermediate heating process can slow up all subsequent operations and make for inefficiency down the line. The breakdown of a unit that has a very large production backlog is much more serious than that of one having a smaller backlog. Scheduling processes in a complex product line is a very difficult exercise and perhaps better suited to a computerized PERT program than to an energy conversation engineer. That does not mean that the possibilities for saving through better process scheduling should be ignored. It is only a warning to move slowly and take pains to find the difficulties that can arise thereby.

A manufacturer of precision instruments changed the specifications for the finishes of over half of his products, thereby eliminating the baking required for the enamel that had been used. He also rescheduled the baking process for the remaining products so that the oven was lighted only twice a week instead of every production day. A study is now proceeding to determine if electric infrared baking will not be more economic than using the gas-fired oven.

14.5.2.6 Cogeneration of Process Steam and Electricity

In-plant (or on-site) electrical energy cogeneration is nothing new. It has been used in industries with large process steam loads for many years, both in the U.S. and Europe. It consists of producing steam at a higher pressure than required for process use, expanding the high-pressure steam through a back-pressure turbine to generate electrical energy, and then using the exhaust steam as process steam. Alternatively, the power system may be a diesel engine that drives an electrical generator. The diesel engine exhaust is then stripped of its heat content as it flows through a waste-heat boiler where steam is generated for plant processes. A third possibility is operation of a gas-turbine generator to supply electric power and hot exhaust gases, which produce process steam in a waste-heat boiler. As will be seen later, the ratio of electric power to steam-heat rate varies markedly from one of these types of systems to the next. In medium to large industrial plants the cogeneration of electric power and process steam is economically feasible provided certain plant energy characteristics are present. In small plants or in larger plants with small process steam loads, cogeneration is not economic because of the large capital expenditure involved. Under few circumstances is the in-plant generation of electric power economic without a large process steam requirement. A small industrial electric plant cannot compete with an electric utility unless the generation required in-plant exceeds the capacity of the utility. In remote areas where no electric utility exists, or where its reliability is inferior to that of the on-site plant, the exception can be made.

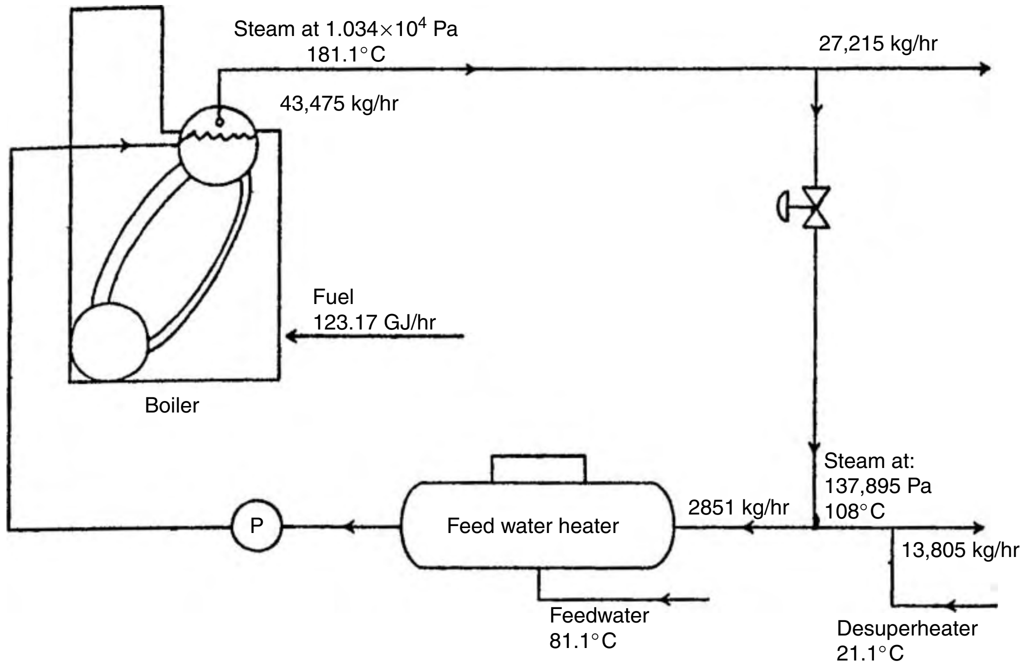


FIGURE 14.15 Steam plant schematic before adding electrical generation.

Cogeneration if applied correctly is not only cost-effective, it is fuel conserving. That is, the fuel for the on-site plant is less than that used jointly by the utility to supply the plant's electric energy and that used on-site to supply process steam. Figure 14.15 and Figure 14.16 illustrate the reasons for and the magnitude of the savings possible. However, several conditions must be met in order that an effective application be possible. First, the ratio of process steam heat rate to electric power must fall close to these given in the table below:

Heat Engine Type	$E_{\text{stream}}/E_{\text{elect}}$
Steam turbine	2.3
Gas turbine	4.0
Diesel engine	1.5

The table is based upon overall electric plant efficiencies to 30, 20, and 40% respectively, for steam turbine, gas turbine, and diesel engine. Second, it is required that the availability of the steam load coincide closely with the availability of the electric load. If these temporal availabilities are out of phase, heat-storage systems will be necessary and the economy of the system altered. Third, it is necessary to have local electric utility support. Unless backup service is available from your utility, the cost of building in redundancy is too great. This may be the crucial factor in some cases. The subject of cogeneration and the available technology is covered in Chapter 16.

14.5.2.7 Commercial Options in Waste-Heat Recovery Equipment

The equipment that is useful in recovering waste heat can be categorized as heat exchangers, heat-storage systems, combination heat storage-heat exchanger systems, and heat pumps.

Heat exchangers certainly constitute the largest sales volume in this group. They consist of two enclosed flow paths and a separating surface that prevents mixing, upports any pressure difference

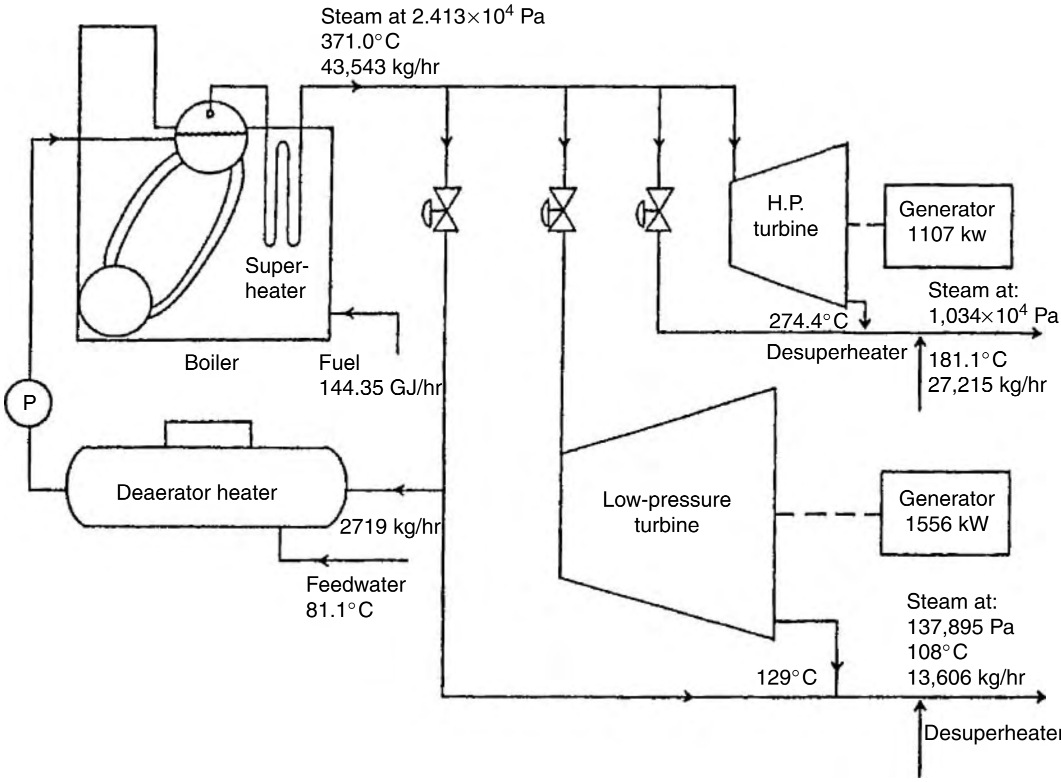


FIGURE 14.16 Steam plant schematic after installing electrical generation.

between the fluids of the two fluids, and provides the means through which heat is transferred from the hotter to the cooler fluid. These are ordinarily operated at steady-state-steady-flow condition. The fluids may be gases, liquids, condensing vapors, or evaporating liquids, and occasionally fluidized solids. For more information, see Chapter 13.

Radiation recuperators are high-temperature combustion-air preheaters used for transferring heat from furnace exhaust gases to combustion air. As seen in Figure 14.17 they consist of two concentric cylinders, the inner one as a stack for the furnace and the concentric space between the inner and outer cylinders as the path for the combustion air, which ordinarily moves upward and therefore parallel to the flow of the exhaust gases. With special construction materials these can handle 1355°C furnace gases and save as much as 30% of the fuel otherwise required. The main problem in their use is damage due to overheating for reduced airflow or temperature excursions in the exhaust-gas flow.

Convective air preheaters are corrugated metal or tubular devices that are used to preheat combustion air in the moderate temperature range (121°C–649°C) for ovens, furnaces, boilers, and gas turbines, or to preheat ventilating air from sources as low in temperature as 21°C. Figure 14.18 and Figure 14.19 illustrate typical construction. These are often available in modular design so that almost any capacity and any degree of effectiveness can be obtained by multiple arrangements. The biggest problem is keeping them clean.

Economizer is the name traditionally used to describe the gas-to-liquid heat exchanger used to preheat the feedwater in boilers from waste heat in the exhaust-gas stream. These often take the form of loops, spiral or parallel arrays of finned tubing through which the feedwater flows and over which the exhaust gases pass. They are available in modular form to be introduced into the exhaust stack or into

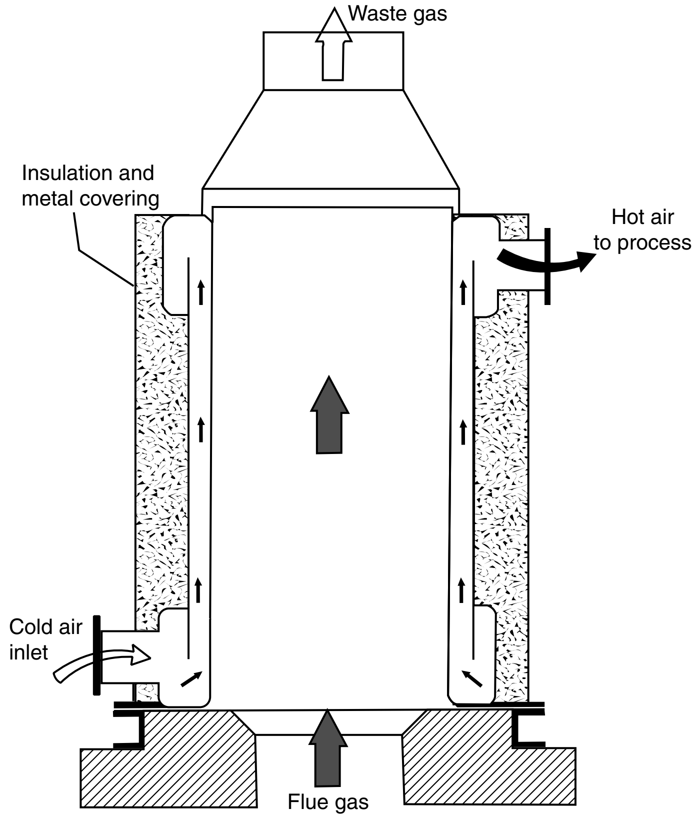


FIGURE 14.17 Metallic radiation recuperator.

the breeching. They can also be used in reverse to heat air or other gases with waste heat from liquid streams.

A more recent development is the use of condensing economizers which are placed in the exhaust stream following high-temperature economizers. They are capable of extracting an additional 6%–8% of the fuel input energy from the boiler exhaust gases. However, they are only used under certain restricted conditions. Obviously, the cooling fluid must be at a temperature below the dew point of the exhaust stream. This condition is often satisfied when boilers are operated with 100% make-up water. A second, less restrictive condition is that the flue gases be free of sulfur oxides. This is normally the case for natural gas-fired boilers. Otherwise the economizer tubes will be attacked by sulfurous and/or sulfuric acid. Acid corrosion can be slowed down markedly by the use of all-stainless steel construction, but the cost of the equipment is increased significantly.

Heat-pipe arrays are often used for air-to-air heat exchangers because of their compact size. Heat-transfer rates per unit area are quite high. A disadvantage is that a given heat pipe (that is, a given internal working substance) has a limited temperature range for efficient operation. The heat pipe transfers heat from the hot end by evaporative heating and at the cold end by condensing the vapor. Figure 14.20 is a sketch of an air preheater using an array of heat pipes.

Waste-heat boilers are water-tube boilers, usually prefabricated in all but the largest sizes, used to produce saturated steam from high-temperature waste heat in gas streams. The boiler tubes are often finned to keep the dimensions of the boiler smaller. They are often used to strip waste heat from diesel-engine exhausts, gas-turbine exhausts, and pollution-control incinerators or afterburners. Figure 14.21 is

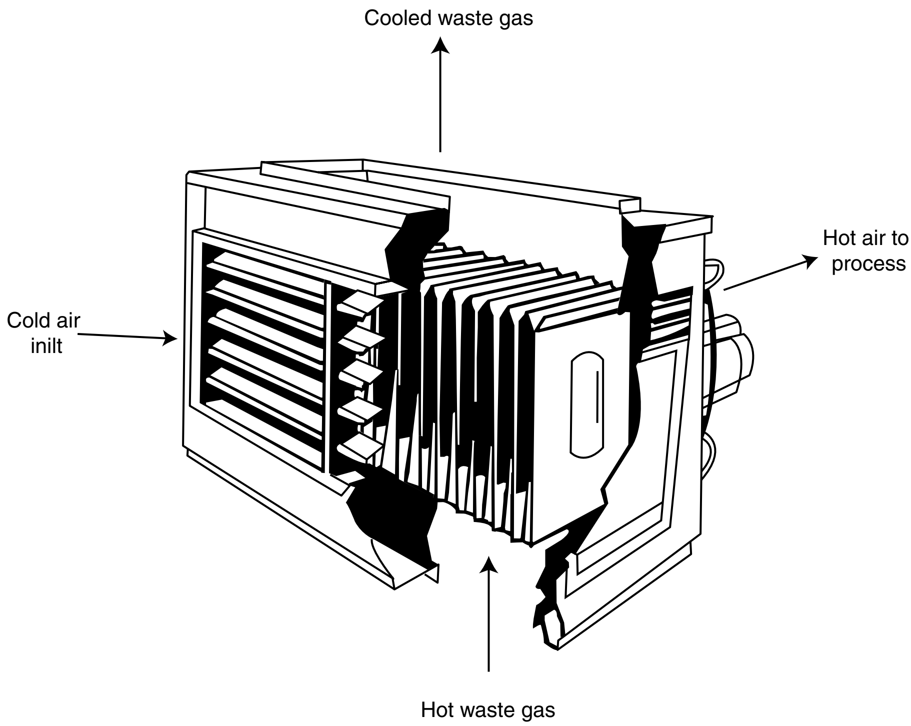


FIGURE 14.18 Air preheater.

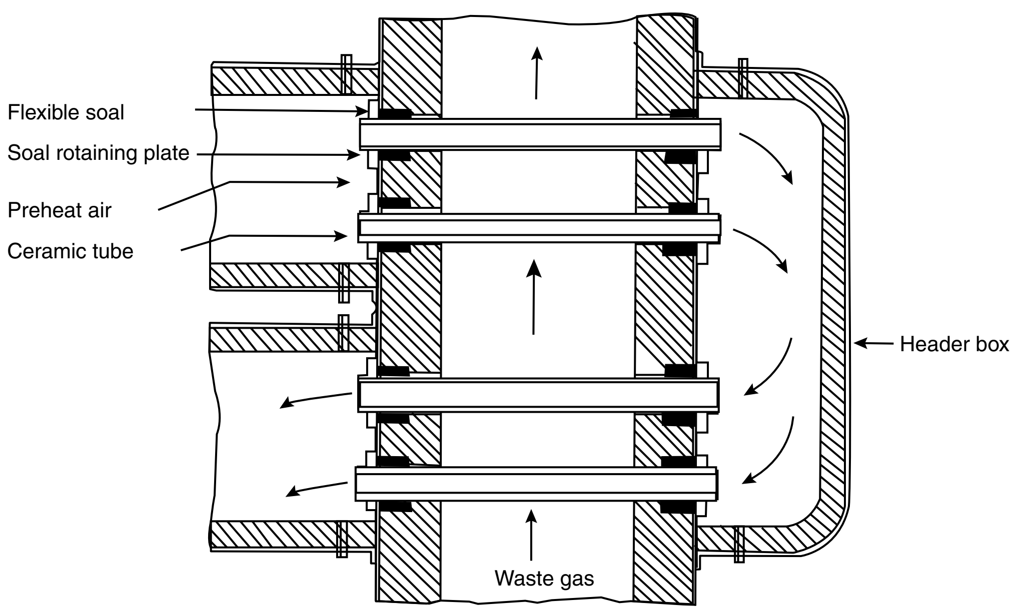


FIGURE 14.19 Ceramic tube recuperator.

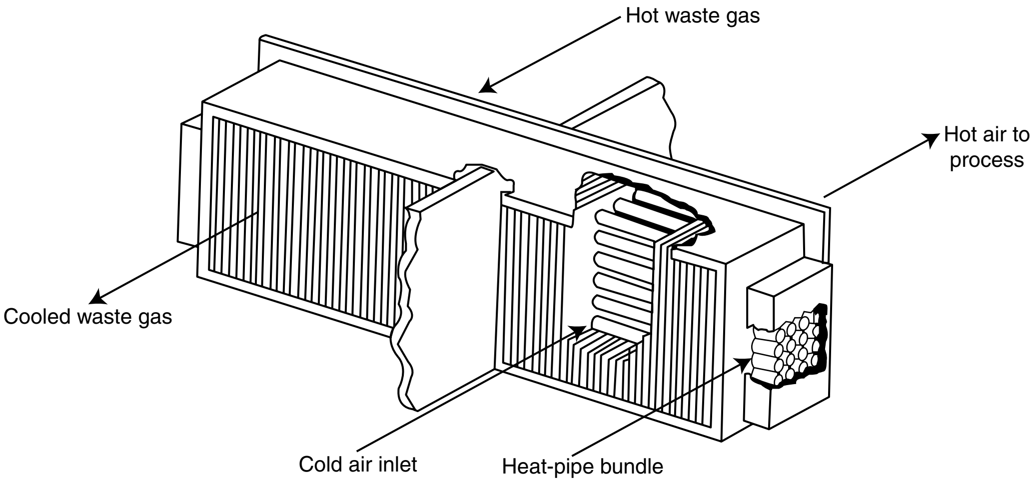


FIGURE 14.20 Heat-pipe recuperator.

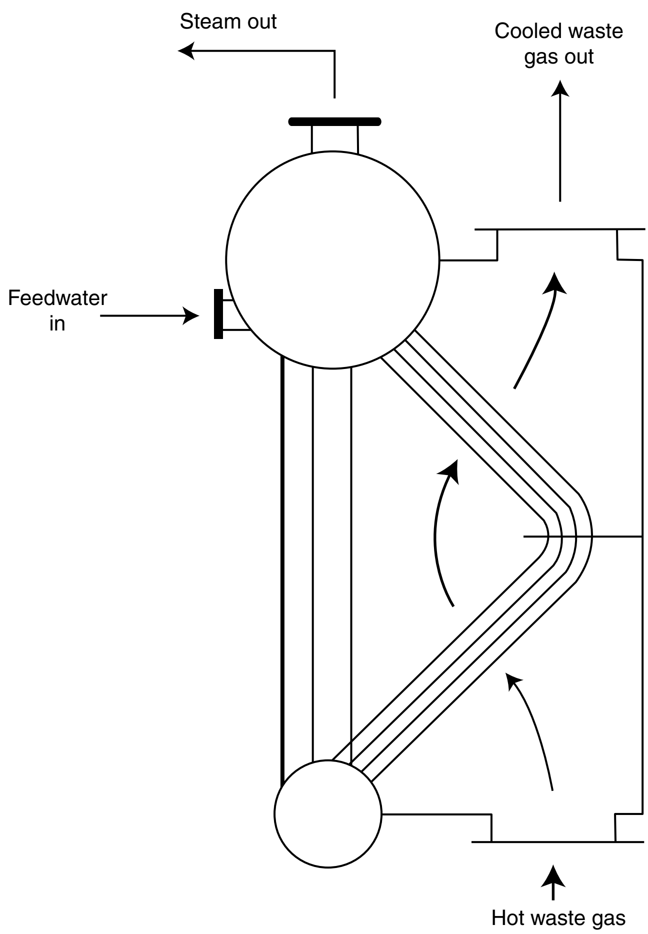


FIGURE 14.21 Waste-heat boiler.

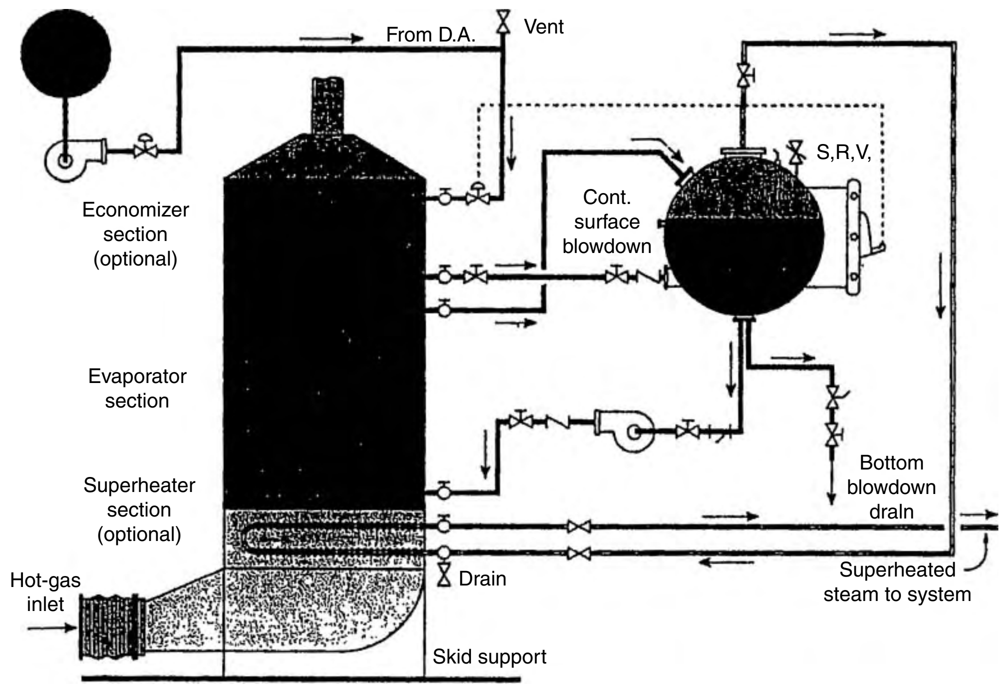


FIGURE 14.22 Schematic diagram of a finned-tube waste-heat boiler. (Courtesy Cannon Technology.)

a diagram of the internals of a typical waste-heat boiler. Figure 14.22 is a schematic diagram showing a waste-heat boiler for which the evaporator is in the form of a finned-tube economizer. Forced water circulation is used giving some flexibility in placing the steam drum and allowing the use of smaller tubes. It also allows the orientation of the evaporator to be either vertical or horizontal. Other advantages of this design are the attainment of high boiler efficiencies, a more compact boiler, less cost to repair or retube, the ability to make superheated steam using the first one or more rows of downstream tubes as the superheater, and the elimination of thermal shock, since the evaporator is not directly connected to the steam drum.

Heat-storage systems, or regenerators, once very popular for high-temperature applications, have been largely replaced by radiation recuperators because of the relative simplicity of the latter. Regenerators consist of twin flues filled with open ceramic checkerwork. The high-temperature exhaust of a furnace flowing through one leg of a swing valve to one of the flues heated the checkerwork while the combustion air for the furnace flowed through the second flue in order to preheat it. When the temperatures of the two masses of checkerwork were at proper levels, the swing valve was thrown and the procedure was continued, but with reversed flow in both flues. Regenerators are still used in some glass- and metal-melt furnaces, where they are able to operate in the temperature range 1093°C–1649°C. It should be noted that the original application of the regenerators was to achieve the high-melt temperatures required with low-heating-value fuel.

A number of ceramic materials in a range of sizes and geometric forms are available for incorporation into heat-storage units. These can be used to store waste heat in order to remedy time discrepancies between source and load. A good example is the storage of solar energy in a rock pile so that it becomes available for use at night and on cloudy days. Heat storage, other than for regenerators in high-temperature applications, has not yet been used a great deal for waste-heat recovery but will probably become more popular as more experience with it accumulates.

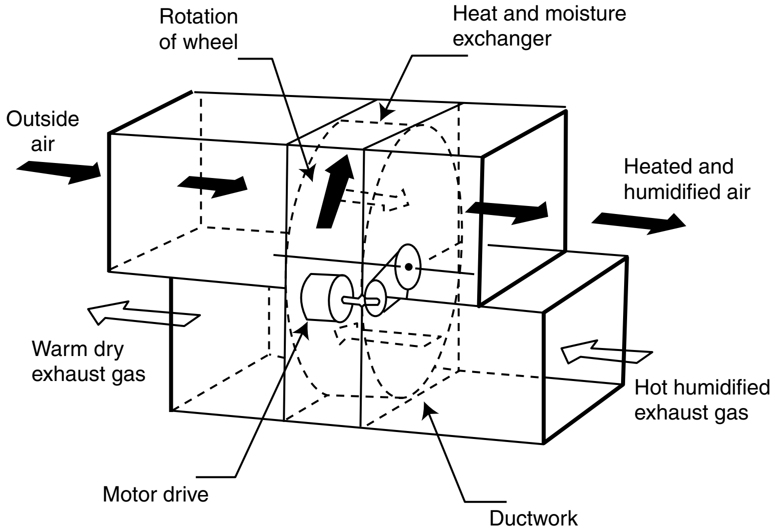


FIGURE 14.23 Heat Wheel.

Combination heat-storage unit-heat exchangers called heat wheels are available for waste-heat recovery in the temperature range 0°C–982°C. The heat wheel is a porous flat cylinder that rotates within a pair of parallel ducts, as can be observed in Figure 23. As the hot gases flow through the matrix of the wheel they heat one side of it, which then gives up that heat to the cold gases as it passes through the second duct. Heat-recovery efficiencies range to 80%. In low- and moderate-wire matrix of the same material. In the high-temperature range the material is ceramic. In order to prevent cross-contamination of the fluid streams, a purge section, cleared by fresh air, can be provided. If the matrix of the wheel is covered with a hygroscopic material, latent heat as well as sensible heat can be recovered. Problems encountered with heat wheels include freeze damage in winter, seal wear, and bearing maintenance in high-temperature applications.

The heat pump is a device operating on a refrigeration cycle that is used to transfer energy from a low-temperature source to a higher temperature load. It has been highly developed as a domestic heating plant using energy from the air or from well but has not been used a great deal for industrial applications. The COP (or the ratio of heat delivered to work input) for an ideal Carnot refrigeration cycle equals $T_H/(T_H - T_L)$, where T_H is the load temperature and T_L is the source temperature. It is obvious that when the temperature difference $T_H - T_L$ becomes of the order of T_H, the heat could be derived almost as cheaply from electric resistance heating. However for efficient refrigeration machines and a small temperature potential to overcome, the actual COP is favorable to moderate-cost waste energy. The heat pump can be used to transfer waste heat from and to any combination of liquid, gas, and vapor.

14.6 The Role of New Equipment and Technology in Industrial Energy Efficiency

14.6.1 Industrial Energy Savings Potential

The last part of Section 14.4 pointed to several studies of industrial energy efficiency that showed that a 20% reduction could be accomplished in a relatively easy and cost-effective manner. The purpose of this section is to provide the equipment, technology, and operational changes that could lead to industrial energy savings on the order of 20% or more.

14.6.2 The U.S. DOE Energy-Loss Study and the NAM Efficiency and Innovation Study

The most recent major study on the potential for improving energy efficiency in industry was conducted for the U.S. Department of Energy, Office of Industrial Technologies, for their ITP in December 2004. This study was called the "Energy Use, Loss and Opportunities Analysis: U.S. Manufacturing and Mining."[1] This study was then used by the NAM together with the Alliance to Save Energy to produce a report on the "Efficiency and Innovation in U.S. Manufacturing Energy Use"[4] Both the U.S. DOE/OIT and NAM conclude that there is a significant opportunity for reducing industrial energy use by 20%.

As stated in the NAM report, "[i]ndustry's best R&D options for reducing energy costs were summarized in a study sponsored by the U.S. DOE. This study identifies energy-efficiency opportunities that yield energy, economic and environmental benefits, primarily for large volume, commodity/process industries. Opportunities were prioritized to reflect the magnitude of potential savings, broadness of suitability across industries, and feasibility to implement. In total, these energy-saving opportunities represent 5.2 quadrillion BTU—21% of primary energy consumed by the manufacturing sector. These savings equate to almost $19 billion for manufacturers, based on 2004 energy prices and consumption volumes." Table 14.10 summarizes these leading opportunities. An expanded version of this information appears in Table 14.11.

14.6.3 The ACEEE Fan and Pump Study

In April 2003, the ACEEE released a report on their study "Realizing Energy Efficiency Opportunities in Industrial Fan and Pump Systems." They concluded that fans and pumps account for more than a quarter of industrial electricity consumption, and that optimization of the operation of these fans and

TABLE 14.10 Industries' Best Opportunities for Future Energy Savings

Top R&D Opportunities for Energy Savings in Commodity/Process Manufacturing; Initiatives that Provide the Largest Energy and Dollar Savings

Type of Opportunity[a]	Total Energy Savings		Total Cost Savings	
	(trillion BTU)	Percent of Total %	($mill.)	Percent of Total %
Waste heat and energy recovery	1831	35	$6408	34
Improvements to boilers, fired systems, process heaters, and cooling opportunities	907	17	$3077	16
Energy system integration and best practices opportunities	1438	28	$5655	30
Energy source flexibility and combined heat and power	828	16	$3100	16
Improved sensors, controls, automation and robotics for energy systems	191	4	$630	3
Totals	5195		$18,870	

[a] See Appendix A for an expanded version of this table.

Source: From U.S. Department of Energy, Industrial Technologies Program. See References section (DOE-TP, 2004); *NAM, Efficiency and Innovation in U.S. Manufacturing Energy Use*, National Association of Manufacturers, Washington, DC, 2005.

TABLE 14.11 Top R&D Opportunities for Industrial Energy Savings

Top R&D Opportunities for Energy Initiatives that Provide the Largest Energy	Savings in Commodity/Process Manufacturing and Dollar Savings	Total Energy Savings		Total Cost Savings	
Type of Opportunity	Leading Industry Recipients	(trillion Btu)	Percent of Total %	($mil.)	Percent of Total %[a]
Waste Heat and Energy Recovery		1831	35	$6408	34
…from gases and liquids, including hot gas cleanup and dehydration of liquid waste streams	Chemicals, petroleum, forest products	851	16	$2271	12
…from drying processes	Chemicals, forest products, food processing	377	7	$1240	7
…from gases in metals and nonmetallic minerals manufacture (excluding calcining), including hot gas cleanup	Iron and steel, cement	235	5	$1133	6
…from by-product gases	Petroleum, iron and steel	132	3	$750	4
…using energy export and co-location (fuels from pulp mills, forest bio-refineries, co-location of energy sources/sinks)	Forest products	105	2	$580	3
…from calcining (not flue gases)	Cement, forest products	74	1	$159	1
…from metal quenching/cooling processes	Iron and steel, cement	57	1	$275	1
Improvements to Boilers, Fired Systems, Process Heaters and Cooling Opportunities		907	17	$3077	16
Advanced industrial boilers	Chemicals, forest products, petroleum, steel, food processing	400	8	$1090	6
Improved heating/heat transfer systems (heat exchangers, new materials, improved heat transport)	Petroleum, chemicals	260	5	$860	5
Improved heating/heat transfer for metals, melting, heating, annealing (cascade heating, batch to continuous process, improved heat channeling, modular systems)	Iron and steel, metal casting, aluminum	190	4	$915	5
Advanced process cooling and refrigeration	Food processing, chemicals, petroleum and forest products	57[b]	1	$212	1
Energy System Integration and Best Practices Opportunities		1438	28	$5655	30
* Steam best practices (improved generation, distribution and recovery), not including advanced boilers	All manufacturing	310	6	$850	5
* Pump system optimization	All manufacturing	302[b]	6	$1370	7
* Energy system integration	Chemicals, petroleum, forest products, iron and steel, food, aluminum	260	5	$860	5
* Energy-efficient motors and rewind practices	All manufacturing	258[b]	5	$1175	6
* Compressed-air system optimization	All manufacturing	163[b]	3	$740	4
* Optimized materials processing	All manufacturing	145[b]	3	$660	3
Energy Source Flexibility and Combined Heat and Power		828	16	$3100	16
* Combined heat and power onsite in manufacturers' central plants, producing both thermal and electricity needs	Forest products, chemicals, food processing, metals, machinery	634	12	$2000	11

Energy source flexibility (heat-activated power generation, waste steam for mechanical drives, indirect vs. direct heat vs. steam)	Chemicals, petroleum, forest products, iron and steel	194	4	$1100	6
Improved Sensors, Controls, Automation and Robotics for Energy Systems	Chemicals, petroleum, forest products, iron and steel, food, cement, aluminum	191	4	$630	3
Totals		5195		$18,870	

Note: All are R&D opportunities except for items denoted by an asterisk (*), which are near-term best practices, applicable to current assets.

[a] Totals may not add up due to rounding.

[b] Energy savings figures include the corresponding recapture of losses inherent in electricity generation, transmission, and distribution.

Source: From U.S. DOE, Office of Industrial Technologies, Industrial Technology Program, *Annual Report: Technology, Delivery, Industry of the Future*, U.S. Department of Energy, Washington, DC, 2004.

pumps could achieve electricity savings ranging from 20% to well over 50% of this category of use.[11] This report says that most optimization projects involve greater engineering costs than equipment costs, but the average payback for a good optimization project is about 1.2 years, with the cost of saved energy on the order of $0.012/kWh. In addition, these estimates do not account for productivity gains known to exist at many of the plant sites, which are sometimes as much as two to five times the energy savings.

This ACEEE report contains some excellent data on motor systems end use that is very difficult to find in general. Their data shows that 40% of industrial motor use is for fans and pumps. Because the MECS data from EIA shows that about 68% of electric use in industry is motors, this leads to the fraction of industrial electric use from fans and pumps to be $(0.68)^*(0.4)=0.32$, or 32%. The end use data from ACEEE is reproduced below in Table 14.12.

To see the final impact of this estimate of industrial energy savings, now apply the ACEEE estimates of 20% to 50% savings on fan and pump energy to the 32% fraction of fan and pump contribution to the total industrial electricity use. Thus, the overall savings are in the range $(0.2)^*(0.32)=6.4\%$ to $(0.5)^*(0.32)=16\%$. Just this opportunity alone could result in achieving over half of the 20% savings contained in the U.S. DOE/OIT study.

Table 14.13 presents the ACEEE estimates of the relative magnitude of electricity consumed by fans and pump systems for the important industries on a national basis.

14.6.4 The LBL/ACEEE Study of Emerging Energy-Efficient Industrial Technologies

In October 2000, ACEEE released a report of a study they did in conjunction with staff from Lawrence Berkeley Laboratories (LBL), where they identified 175 emerging energy-efficient technologies, and honed this list down to 32 technologies that had a high likelihood of success and a high energy savings.[27] An interesting aspect of this study is that it shows that the U.S. is not running out of technologies to improve energy efficiency and economic and environmental performance, and will not run out in the future. The study shows that many of the technologies have important nonenergy benefits, ranging from reduced environmental impact to improved productivity. Several technologies have reduced capital costs compared to the current technology used by those industries. Nonenergy benefits such as these are frequently a motivating factor in bringing this kind of technology to market.

The LBL/ACEEE list of 32 most beneficial technologies is shown in Table 14.14.

TABLE 14.12 National Industrial Motor Systems Energy End Use

Pumps	25%
Materials processing	22%
Compressed air	16%
Fans	14%
Material handling	12%
Refrigeration	7%
Other	4%

Source: From Elliott, R. N., and Nadel, S., *Realizing Energy Efficiency Opportunities in Industrial Fan and Pump Systems*, Report A034, American Council for an Energy Efficient Economy, Washington, DC, 2003.

TABLE 14.13 Characterization of Industrial Fan and Pump Load in the United States

NAICS	Industry	Electricity Demand 1997	Pumps (%)	Fans and Blowers (%)	Total Motors (%)	Motor Electricity	Fans/Pumps Share of Electricity (%)	Fans/Pumps Electricity Use
11	Agriculture	16,325	25	20	75	12,244	45	7346
22	Mining	85,394	7	21	90	76,854	29	24,363
311	Food mfg.	66,166	11	5	81	53,756	16	10,809
314	Textile product mills	5135	14	15	82	4221	30	1523
321	Wood product mfg.	21,884	4	10	80	17,464	14	3064
322	Paper mfg.	119,627	28	16	84	101,078	44	52,636
324	Petroleum and coal products mfg.	69,601	51	13	85	59,369	63	44,061
325	Chemical mfg.	212,709	18	8	73	154,693	26	54,797
326	Plastics & rubber mfg.	52,556	9	4	66	34,847	13	6729
327	Nonmetallic minerals product mfg.	37,416	4	4	65	24,328	8	3037
331	Primary metal mfg.	172,518	2	4	26	44,855	6	10,351
332	Fabricated metal product mfg.	49,590	7	5	65	32,462	12	6149
333	Machinery mfg.	27,295	8	4	67	18,391	12	3330
334	Computer & electronic product mfg.	40,099	2	3	54	21,783	4	1801
336	Transportation equipment mfg.	54,282	4	6	64	34,629	11	5753
	Total	1,030,598				690,974		235,750
				Fraction of total elec.		67%		23%

TABLE 14.14 Technologies with High Energy Savings and a High Likelihood of Success

Technology	Code	Total Energy Savings	Likelihood of Success	Recommended Next Steps
Efficient cell retrofit designs	Alum-2	High	High	Demo
Advanced lighting technologies	Lighting-1	High	High	Dissem., demo
Advance ASD designs	Motorsys-1	High	High	R&D
Membrane technology wastewater	Other-3	High	High	Dissem., R&D
Sensors and controls	Other-5	High	High	R&D, demo, dissem.
Black liquor gasification	Paper-1	High	High	Demo
Near net shape casting/strip casting	Steel-2	High	High	R&D
New EAF furnace processes	Steel-3	High	High	Field test
Oxy-fuel combustion in reheat furnace	Steel-4	High	High	Field test
Advanced CHP turbine systems	Utilities-1	High	High	Policies
Autothermal reforming-ammonia	Chem-7	High	Medium	Dissemination
Membrane technology—food	Food-3	High	Medium	Dissem., R&D
Advanced lighting design	Lighting-2	High	Medium	Dissem., demo
Compressed-air system management	Motorsys-3	High	Medium	Dissem.
Motor system optimization	Motorsys-5	High	Medium	Dissem., training
Pump efficiency improvement	Motorsys-6	High	Medium	Dissem., training
High-efficiency/low NOX burners	Other-2	High	Medium	Dissem., demo
Process integration (pinch analysis)	Other-4	High	Medium	Dissemination
Heat recovery—paper	Paper-5	High	Medium	Demo
Impulse drying	Paper-7	High	Medium	Demo
Smelting reduction processes	Steel-5	High	Medium	Demo
Advanced reciprocating engines	Utilities-2	High	Medium	R&D, demo
Fuel cells	Utilities-3	High	Medium	Demo
Microturbines	Utilities-4	High	Medium	R&D, demo
Inert anodes/wetted cathodes	Alum-4	High	Medium	R&D
Advanced forming	Alum-1	Medium	High	R&D
Plastics recovery	Chem-8	Medium	High	Demo
Continuous melt silicon crystal growth	Electron-1	Medium	High	R&D
100% recycled glass cullet	Glass-1	Medium	High	Demo
Anaerobic wastewater treatment	Other-1	Medium	High	Dissem., demo
Dry sheet forming	Paper-4	Medium	High	R&D, demo
Biodesulfurization	Refin-1	Medium	High	R&D, demo

Technologies in this table are listed in alphabetical order based on industry sector.

14.7 Conclusion

Energy is the lifeblood of industry; it is used to convert fuels to thermal, electric, or motive energy to manufacture all the products of daily life. Using this energy efficiently is a necessity to keep industries competitive, clean, and at their peak of productivity. Energy management programs that improve the operational efficiency and the technological efficiency of industry are critical to the long-term success of industry and manufacturing in the U.S. One important result in this area has been a recognition that the U.S. is not running out of technologies to improve industrial energy efficiency, productivity, and environmental performance, and it is not going to run out in the foreseeable future. A substantial opportunity to the country's industrial energy use by over 20% is currently available using better operational procedures and using improved equipment in industrial plants. These savings to industry are worth almost $19 billion at 2004 energy prices. With crude oil prices edging toward $70 in late summer 2005, this dollar savings amount should be substantially higher. It is time to capture the benefits of this opportunity.

References

1. U.S. Department of Energy, 2004. *Energy Use, Loss and Opportunities Analysis: U.S. Manufacturing and Mining*, U.S. Department of Energy, Office of Industrial Technologies, Industrial Technology Program, Washington, DC.
2. U.S. DOE Energy Information Agency, 2005. *Industrial Sector Energy Price and Expenditure Estimates for 2001*. U.S. Department of Energy, Office of Industrial Technologies, Industrial Technology Program, Washington, DC.
3. U.S. DOE Energy Information Administration, 1995. *Manufacturing Energy Consumption Survey: Changes in Energy Efficiency 1985–1991*, U.S. Department of Energy, Office of Industrial Technologies, Industrial Technology Program, Washington, DC.
4. NAM, 2005. *Efficiency and Innovation in U.S. Manufacturing Energy Use,* National Association of Manufacturers, Washington, DC.
5. U.S. DOE, Office of Industrial Technologies, Industrial Technology Program, 2004. *Annual Report: Technology, Delivery, Industry of the Future*, U.S. Department of Energy, Washington, DC.
6. U.S. DOE, 1990. *The National Energy Strategy, Chapter on Industrial Energy Use*, U.S. Department of Energy, Washington, DC.
7. Oak Ridge National Laboratory, 2000. *Scenarios for a Clean Energy Future, Interlaboratory Working Group,* 2000. Oak Ridge National Laboratory, Oak Ridge, TN.
8. Price, L. and Worrell. E. 2004. Improving industrial energy efficiency in the U.S.: Technologies and policies for 2010–2050. *Proceedings of The 10–50 Solution: Technologies and Policies for a Low Carbon Future*, Lawrence Berkeley National Laboratory, Berkeley, CA.
9. Elliott, R. N. and Nadel, S. 2003. Realizing Energy Efficiency Opportunities in Industrial Fan and Pump Systems, American Council for an Energy Efficient Economy, Washington, DC.
10. Aspenson, R. L. 1989. Testimony to the U.S. Department of Energy on the National Energy Strategy, hearings on energy and productivity, Providence, RI.
11. 3M Company. Improving energy efficiency. http://www.MMM.com (accessed June 15, 2005).
12. Capehart, L. C. and Capehart. B. L. 1994. Writing user-friendly energy audit reports, *Strategic planning for energy and the environment*, 14, 2, pp. 17–26. Published by the Association of Energy Engineers, Atlanta, GA.
13. Capehart, B. L. and Capehart. L. C. 1994. Improving industrial energy audit analyses. *Proceedings of the ACEEE Summer Study of Industrial Energy Use*, ACEEE, Washington, DC.
14. Pawlik, K.-D., Capehart, L. C., and Capehart, B. L. 2001. Analyzing facility energy use: A balancing act. *Strategic Planning for Energy and the Environment*, 21, 2, 8–23.
15. Air-Conditioning & Refrigeration Institute. *ARI Unitary Directory*. http://www.ari.org (accessed June 15, 2005.)

16. Nadel, S., Elliott, R. N., Shepard, M., Greenberg, S., and Katz, G. 2002. *Energy-Efficient Motor Systems: A Handbook on Technology, Program and Policy Opportunities, 2nd Ed*, American Council for an Energy-Efficient Economy, Washington, DC.
17. Washington State Energy Office, *MotorMaster Electric Motor Selection Software and Database*. http://www.oit.doe.gov/bestpractices/software_tools.shtml (accessed June 15, 2005).
18. Hoshide, R. K. 1994. Electric motor 'do's and 'don't's. *Energy Engineering*, 91, 1, pp. 6–24. Published by the Association of Energy Engineers, Atlanta, CA.
19. Stebbins, W. L. 1994. Are you certain you understand the economics for applying ASD systems to centrifugal loads? *Energy Engineering*, 91, 1, pp. 25–44. Published by the Association of Energy Engineers, Atlanta, CA.
20. Vaillencourt, R. R. 1994. Simple solutions to VSD pumping measures. *Energy Engineering*, 91, 1, pp. 45–59. Published by the Association of Energy Engineers, Atlanta, CA.
21. Kempers, G. 1995. DSM pitfalls for centrifugal pumps and fans. *Energy Engineering*, 92, 2, pp. 15–23. Published by the Association of Energy Engineers, Atlanta, CA.
22. Personal communication with Mr. Mark Webb, Senior Engineer, Virginia Power Company, Roanoke, VA, January 1995.
23. Capehart, B. L. ed. 2004. *Information Technology for Energy Managers: Understanding Web Based Energy Information and Control Systems*, Fairmont Press, Atlanta, GA.
24. Capehart, B. L. and Capehart, L. C. eds. 2005. *Web Based Energy Information and Control Systems: Case Studies and Applications*, Fairmont Press, Atlanta, GA.
25. Capehart, B. L. and Capehart, L. C. eds. 2005. *Web Based Energy Information and Control Systems: Case Studies and Applications*, Fairmont Press, Atlanta, GA.
26. Turner, W. C. ed. 2005. *Energy Management Handbook*, Fairmont Press, Atlanta, GA.
27. Martin, N. and Elliott, R. N. 2000. *Emerging Energy Efficient Technologies*, American Council for an Energy Efficient Economy, Washington, DC.
28. Manufacturing Energy Consumption Survey, 1998; U.S. Department of Energy; Energy Information Agency, 2002.

15

Process Energy Efficiency: Pinch Technology

Kirtan K. Trivedi
Parsons Energy and Chemicals Group, Inc.

Ed Fouche and
Kelly E. Parmenter
Global Energy Partners, LLC

15.1 Pinch Technology in Theory

Kirtan K. Trivedi

Many industrial processes require that heat be added to and removed from streams within the process. This is notably the case in the chemical process industry. Raw materials are heated or cooled to the appropriate reaction temperature. Heat may be added or removed to carry out the reaction at the specified condition of temperature and pressure. The product separation is achieved by heat addition (to a reboiler in a distillation column) and removal (from a condenser of the distillation column). The product may be heated or cooled to the correct temperature for storage and transportation. Thus, at all times heat is either added or removed to a variety of process streams by utilities or by heat exchange between process streams. In an integrated plant this is normally achieved with a *heat exchanger network* (HEN). A large fraction of the capital cost of many chemical process plants is attributed to heat recovery networks.

The aim of a designer is to synthesize a near optimal configuration for a process. this indicates that a proper trade-off between the capital invested and the operating cost of the plant should be achieved. The capital cost depends on the type, number, and size of units utilized to satisfy the design objectives. A substantial part of the operating cost usually depends upon the utilities consumed. To reduce these costs, the process designer should aim for an economic combination having near the theoretical minimum number of heat exchanger units and aim to recover the maximum possible heat with them. An obvious way to recover heat is by exchanging it between hot process streams that need to be cooled and cold process streams that need to be heated, in addition to the heating and cooling utilities. Furthermore, the designer should also investigate the operability of the final design.

These objectives can be achieved by synthesizing a good heat exchanger network. However, a very large number of alternatives exist. Because of this, it will be highly rewarding to synthesize quickly and systematically the best possible alternatives. It is now possible to do this with the aid of "pinch technology."

Pinch technology refers to a large and growing set of methods for analyzing process energy requirements in order to find economically optimal and controllable designs. Considerable development has taken place in pinch technology during the last two decades, mainly due to the efforts of Linnhoff and co-workers. Pinch technology has proved to be effective and is successfully applied to process integration[1] that encompasses overall plant integration and includes heat exchanger networks, heat and power integration or cogeneration, and thermal integration of distillation columns. To date, there are 65 concepts used by this technology. However, due to the limited scope of this chapter, only the most important concepts are discussed here.

Industrial applications of this technology include capital cost reduction, energy cost reduction, emissions reduction, operability improvement, and yield improvement for both new process design and revamp process design. Imperial Chemical Industries (ICI) where this technology was first developed reported an averaged energy saving of about MM$11/year, (about 30%) in processes previously thought optimized.[2] The payback time was typically in the order of 12 months. Union Carbide showed even better results.[3] Studies conducted by Union Carbide in none projects showed an average savings of 50% with an average payback period of 6 months. BASF reports energy savings of 25% obtained by application of pinch technology to their Ludwigshafen site in Germany.[4] Over a period of 3 years about 150 projects were undertaken by BASF. Fraser and Gillespie[5] reported energy savings of 10% with a payback period of 2 years by applying pinch technology to the Caltex Refinery situated in Milnerton, Cape, South Africa. The energy consumption before the study was 100 MW for the whole refinery. A newsfront article in *Chemical Engineering*[6] gives more details about the experience of various companies in using pinch technology and the benefits obtained.

15.1.1 Fundamental Principles and Basic Concepts

Pinch technology is based on thermodynamic principles. Hence, in this section we will review the important thermodynamic principles and some basic concepts.

15.1.1.1 Temperature–Enthalpy Diagram

Whenever there is a temperature change occurring in a system, the enthalpy of the system will change. If a stream is heated or cooled, then the amount of heat absorbed or liberated can be measured by the amount of the enthalpy change. Thus,

$$Q = \Delta H \tag{15.1}$$

For sensible heating or cooling at constant pressure where $\mathrm{CP} = m_\mathrm{p}$,

$$\Delta H = \mathrm{CP}\Delta T \tag{15.2}$$

For latent heating or cooling,

$$\Delta H = m\lambda \tag{15.3}$$

If we assume that the temperature change for latent heating or cooling is 1°C, then

$$CP = m\lambda \tag{15.4}$$

Equation 15.2 enables us to represent a heating or cooling process on a temperature–enthalpy diagram. The abscissa is the enthalpy and the ordinate is the temperature. The slope of the line is (1/CP). Figure 15.1a shows a cold stream being heated from 20°C to 80°C with CP = 2.0 kW/°C.

As enthalpy is a relative function, the stream can be drawn anywhere on the enthalpy scale as long as it is between its starting and target temperatures and has the same enthalpy change. Figure 15.1b shows such a case. Thus, one of the important advantages of representing a stream on the temperature–enthalpy plot is that the stream can be moved horizontally between the same temperature intervals. Figure 15.1c shows two different streams in the same temperature interval.

Two streams can be easily added on the temperature–enthalpy plot to represent a composite of the two streams. Figure 15.2 shows how to obtain the composite of two streams on this plot. This feature will be used in later sections to predict the minimum utility requirement for multistream problems.

A heat exchanger is represented by two heat curves on the temperature–enthalpy diagram as shown in Figure 15.3. This figure also shows how we will represent heat exchangers in a grid diagram form and the conventional form. In the grid diagram form, the hot stream goes from left to right and the cold stream goes from right to left. If the flow is counter to the current in the exchanger, then the temperature will decrease from left to right. The exchanger is represented by two circles connected by a vertical line. The advantages of the grid diagram will become apparent when we discuss the design of heat exchanger networks. The exchanger has a temperature difference at the hot end and another at the cold end of the heat exchanger. The smaller of these two temperature differences is called Δt_{min}.

15.1.1.2 Second Law of Thermodynamics and Exergy Analysis

According to the Second Law of Thermodynamics, heat can be transferred from a higher temperature to a lower temperature only. Design of low-temperature processes uses exergy analysis that is based on this law.

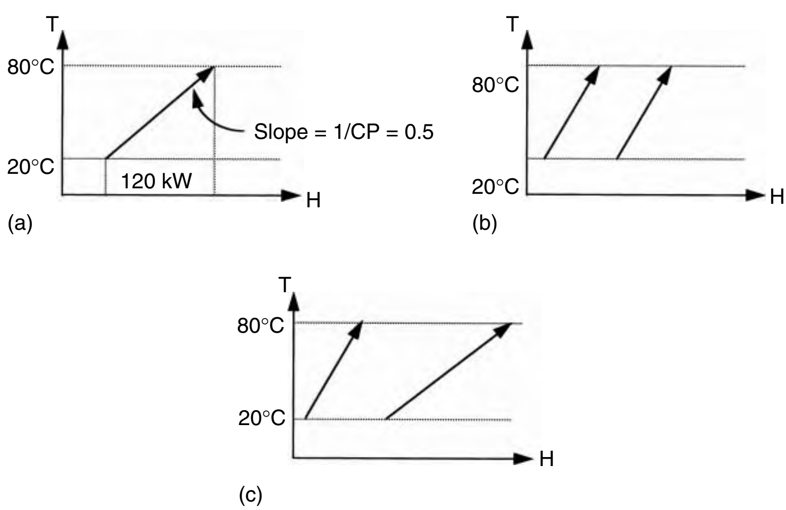


FIGURE 15.1 The temperature–enthalpy diagram. (a) Representation of a cold stream on the *T–H* diagram. (b) A stream can be moved horizontally on the *T–H* diagram in a given temperature interval. (c) Two different streams in the same temperature interval.

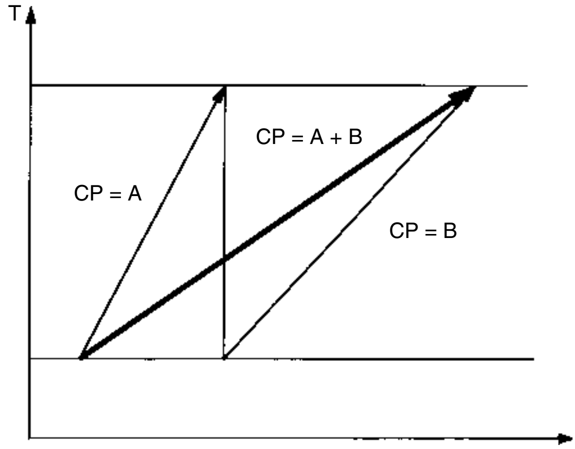


FIGURE 15.2 Composite stream obtained from two different streams in the same temperature interval.

15.1.1.3 Some Definitions

A *match* between a hot and a cold stream indicates that heat transfer is taking place between the two streams. A match between two streams is physically achieved via a heat exchanger *unit*. The number of heat exchanger units impacts the plot plan and determines the piping and the foundation cost.

For reasons of fouling, mechanical expansion, size limitation, cleaning, improved heat transfer coefficients, and so on, most process heat exchangers are shell-and-tube type with 1 shell pass and 2 tube passes. Often what appears as a single match between two streams in a heat exchanger network representation is actually installed as several 1–2 exchangers in series or parallel. The term *shell* will be

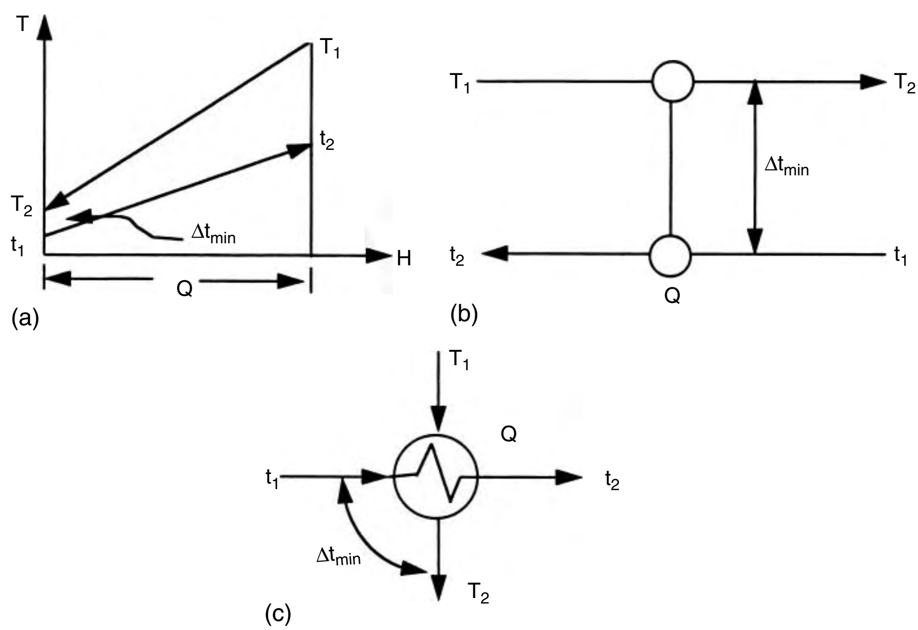


FIGURE 15.3 Heat exchanger representation. (a) Representation of an exchanger on the *T*–*H* diagram. (b) Grid diagram representation of an exchanger. (c) Conventional representation of an exchanger.

TABLE 15.1 Commercially Available Computer Programs for Pinch Technology

Software Name	Marketed By
Supertarget	Linnhoff-March Inc., 9800 Richmond Av. Suite 560, Houston, TX 77042 Phone: (713) 787 6861 Fax: (713) 787 6865
Advent	Aspen Technology Inc., Ten Canal Park, Cambridge, MA 02141 Phone: (617) 577 0100 Fax: (617) 577 0303
Heatnet	Britain's National Engineering Laboratory East Kilbride Scotland
Hextran	Simulation Sciences Inc. 601 S. Valencia Ave., Brea, CA 92621 Phone: (714) 579 0354 Fax: (714) 579 0412

used to represent a single 1–2 shell-and-tube heat exchanger. See Chapter 13 for more discussion on 1–2 shell-and-tube heat exchangers.

15.1.2 Software

Pinch technology is a mature state-of-the-art technology for process integration and design. While various programs have been written by different researchers at various universities, only a handful of commercial programs are available. Table 15.1 shows the commercially available programs for process integration.

15.1.3 Heat Exchanger Network Design Philosophy

This section discusses the design philosophy of pinch technology.

15.1.3.1 Optimization Variables

The objective of the heat exchanger network synthesis problem is to design a network that meets an economic criterion such as minimum total annualized cost. The total annualized cost is the sum of the annual operating cost (which consists mainly of energy costs) and annualized capital cost. The capital cost of a network primarily depends on the total surface area, the number of shells, and the number of units that will be installed. The capital cost will also depend on the individual type of heat exchangers and their design temperature, pressure, and material of construction.

If we include the pressure drop incurred in a heat exchanger, then the capital and operating cost of the pumps will also have to be taken into account. We will limit the scope of this chapter by not including stream pressure drop constraints.

Figure 15.4 shows the various variables that affect the optimization of a heat exchanger network. To make the network operable or flexible, extra cost may be incurred. This cost may be in the form of added equipment, use of extra utilities, or use of additional area on some of the exchangers in the network.

From this discussion, it is clear that the synthesis of heat exchanger networks is a multivariable optimization problem. The design can never violate the laws of thermodynamics. Hence, the philosophy adopted in pinch technology is to establish targets for the various optimization variables based on thermodynamic principles. These targets set the boundaries and constraints for the design problem. Further, targets help us identify the various trade-offs between the optimization parameters. They help us in obtaining a bird's-eye view of the solution space and identify the optimal values of the optimization parameters. Once the optimal values are identified, design the network at these values. This approach will always lead to an optimal design.

To understand the interaction between the different optimization variables consider the heat curves with a small $\Delta t_{\min}$ for a simple two-stream system shown in Figure 15.5a. Three different sections can be easily identified on this diagram. Section 1 represents the cold utility requirement, Section 2 represents the process-process heat exchange, and Section 3 represents the hot utility requirement. From this set of heat curves we can calculate the utility requirements for the system, and for each section we can calculate the area required as we know the duty for each section and the terminal temperatures. Further, we can

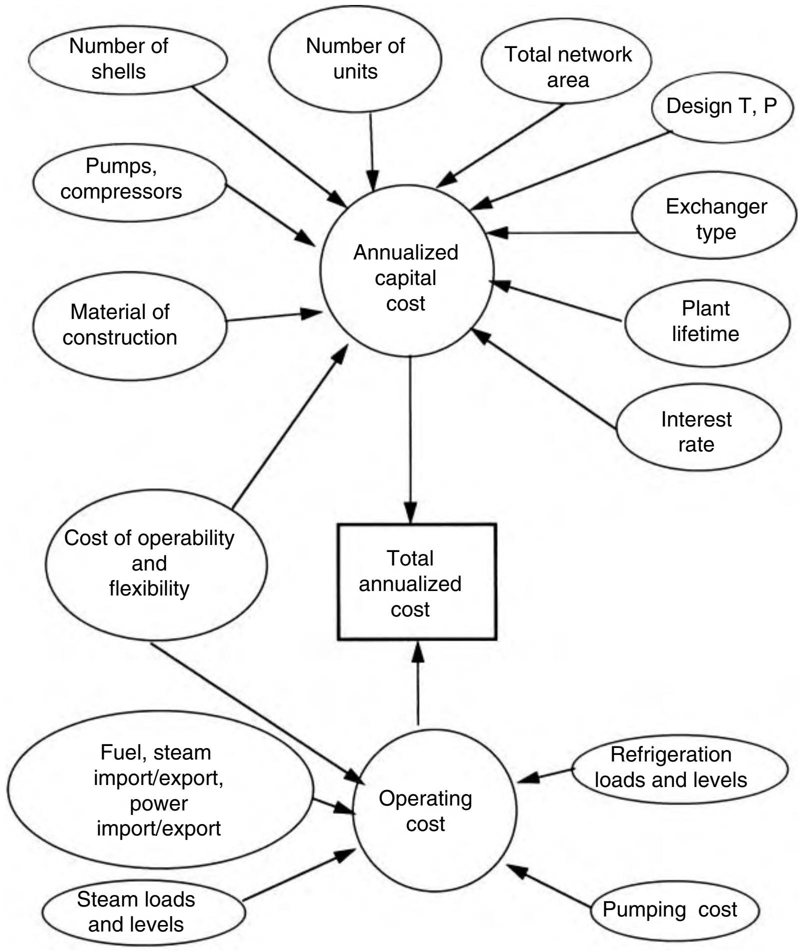


FIGURE 15.4 Optimization variables in the design of heat exchanger networks.

deduce that each section will require one unit. We can use Bell's method[7] to estimate the number of shells required for each unit that represents the different sections. In Section 2, two shells will be required. The heat curves for the hot and cold streams establish targets for energy, area, units, and shells. These are the major components that contribute towards the annualized cost of the network.

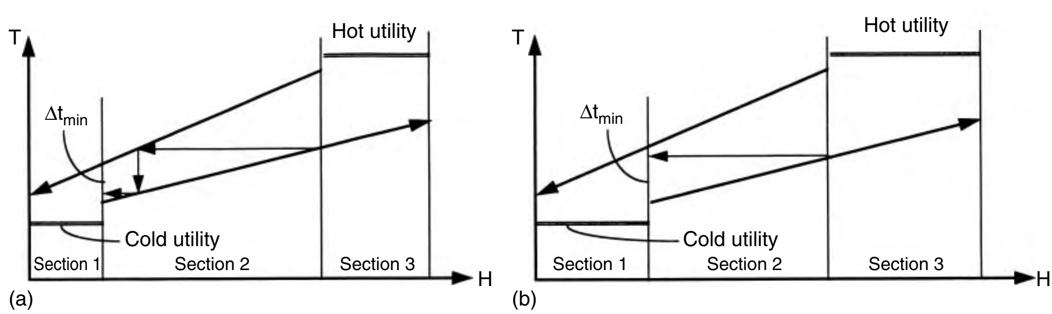


FIGURE 15.5 Effects of Δt_{min} on the energy, area, units, and shells required for a heat exchange system.

Now for the same system, let us increase the Δt_{min}. The new set of heat curves is shown in Figure 15.5b. The preceding exercise can now be repeated to obtain targets for energy, area, units, and shells. When we increase the value of Δt_{min}, the utility requirement will increase and the total network area required will be decreased. The number of units for this system remains the same, but the number of shells required will decrease. Section 2 now only requires one shell. Thus, as the value of Δt_{min} increases, the utility cost increases and the capital cost decreases. This indicates that Δt_{min} is the single variable that fixes the major optimization variables shown in Figure 15.4. Hence, we can reduce the multivariable optimization problem to a single-variable optimization problem. This single variable is Δt_{min}. At the optimum value of Δt_{min}, the other optimization variables—that is, number of shells, number of units, total network area requirement, and the hot and cold utility requirements—will also be optimal. For a multistream problem the same conclusion can be easily derived. Hence, for the multistream problem optimization discussed in the next sections, we will develop methods to optimize the value of Δt_{min}.

15.1.4 Design Problem

Consider an illustrative plant flowsheet shown in Figure 15.6. The feed is heated to the reaction temperature. The reactor effluent is further heated, and the products are separated in a distillation column. The reboiler and condenser use external utilities for control purposes. The overhead and bottoms products are cooled and sent for further processing. We shall use this problem to illustrate the concepts of pinch technology.

Table 15.2 lists the starting and target temperatures of all the streams involved in the flowsheet. It also shows the CP values. While pressure drop constraints on the individual streams determine the film heat transfer coefficients, we shall not take into account this constraint. Instead, we use (unrealistically) the same heat transfer coefficient value of 1 kW/m^2 K for all streams. Our objective is to design an optimum heat exchanger network for this process using the economics outlined in Table 15.3. Further, let Δt_{min} denote the minimum temperature difference between any hot stream and any cold stream in any exchanger in the network. It is desired that the final network should be able to control the target temperature of all the streams at the design temperature.

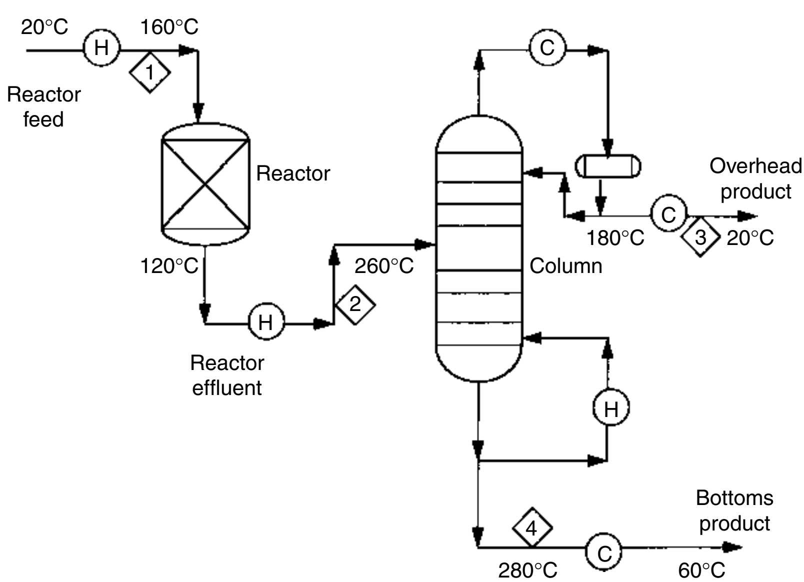


FIGURE 15.6 A typical plant flowsheet.

TABLE 15.2 Stream Data for the Flowsheet in Figure 15.6

Stream Number	Stream Name	Stream No.	T_s (°C)	T_t (°C)	CP (kW/°C)
1	Reactor Feed	1	20	160	40
2	Reactor Effluent	2	120	260	60
3	Overhead Product	3	180	20	45
4	Bottoms Product	4	280	60	30

15.1.5 Targets for Optimization Parameters

15.1.5.1 Energy Targets

15.1.5.1.1 Composite Curves

Let us plot the heat curves for all the hot streams on the *T–H* diagram (see Figure 15.7). We can divide the diagram into a number of temperature intervals, defined by the starting and target temperatures for all the streams. Between two adjacent temperatures we can calculate the total heat content of all the streams that are present in this temperature interval. For example, between 180°C and 60°C, the sum total of the heat available is calculated as

$$\begin{aligned} \Delta H &= (T_3 - T_2) \textstyle\sum_j \mathrm{CP}_j \\ \Delta H_2 &= (180 - 60)(40 + 30) = 8,400 \text{ kW} \end{aligned} \tag{15.5}$$

A composite curve that represent the heat content of all the hot streams is obtained by summing the heat available in each of these temperature intervals as shown in Figure 15.7. Similarly, a composite curve for all the cold streams can be obtained. These two composite curves can be used as the heat curves for the whole process.

To obtain the energy target, fix the hot composite curve and move the cold composite curve horizontally till the shortest vertical distance between the two curves is equal to the value of $\Delta t_{\min}$. The overshoot of the cold composite curve is the minimum hot utility requirement and the overshoot of the hot composite curve is the minimum cold utility requirement.[8] For $\Delta t_{\min} = 30$°C, the minimum hot utility requirement is 4,750 kW and the cold utility requirement is 4,550 kW (see Figure 15.8).

15.1.5.1.2 The Problem Table

The preceding procedure for obtaining the energy targets using composite curves is time-consuming and clumsy. An alternative method based on thermodynamic principles is developed. Hohmann[8] called it the feasibility table. Linnhoff and Flower[9] independently developed the problem table. The problem table algorithm is easy and involves no trial and error.

The algorithm consist of the following steps:

TABLE 15.3 Economics for the Flowsheet in Figure 15.6

	Heat Exchangers	
Installed cost per shell ($)		$= 10{,}000\ A^{0.6}$ (m^2)
	Utility Data	
Cost of using hot oil		= 68 ($/kW yr), ($2.36/MMBtu)
Cooling water		= 2.5 ($/kW yr), ($0.09/MMBtu)
Temperature range of hot oil		= 320°–310°C
Temperature range of cooling water		= 10°–20°C
	Plant Data	
Interest rate		= 10%/year
Lifetime		= 5 years
Operation time		= 8000 hr/yr
Calculated capital recovery factor, CRF (see Chapter 3 for details)		= 0.2638/yr

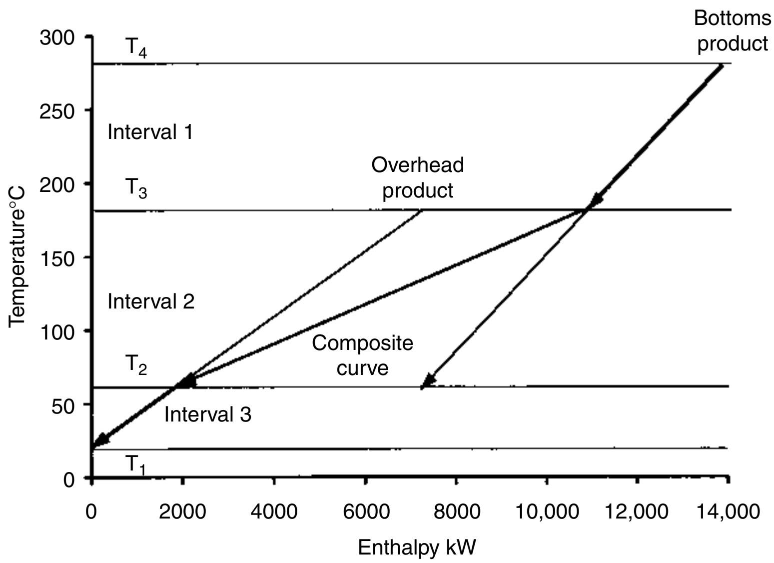


FIGURE 15.7 Construction of the hot composite curve.

- Select a value of Δt_{min}. Since we have already established the targets using composite curves for $\Delta t_{min} = 30°C$, we shall use that value here.
- Convert the actual terminal temperatures into interval temperatures as follows:

for the hot streams: $T_{int} = T_{act} - \Delta t_{min}/2$

for the cold streams: $T_{int} = T_{act} + \Delta t_{min}/2$

where T_{int} = interval temperature

T_{act} = actual stream temperature

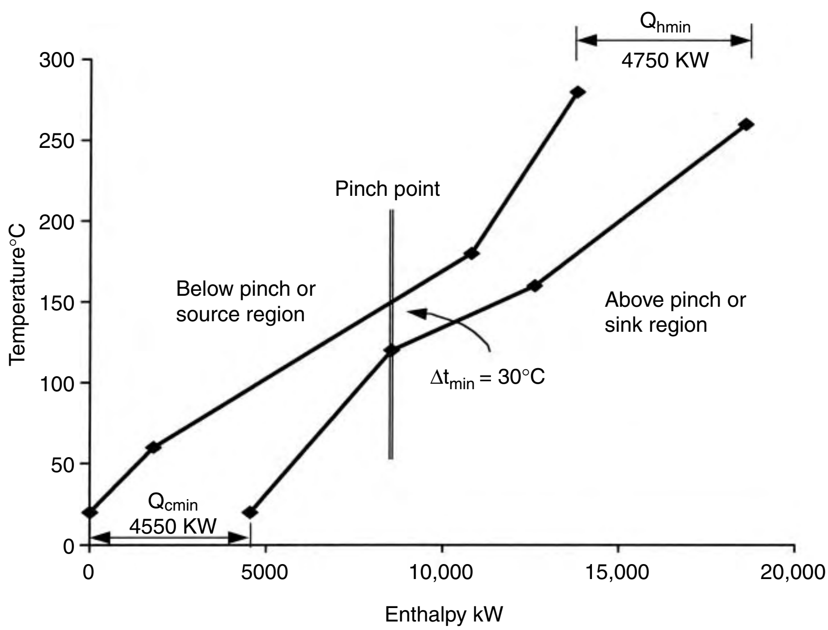


FIGURE 15.8 Composite curves for the flowsheet in Figure 15.9.

TABLE 15.4 Generation of Temperature Intervals for $\Delta t_{min}=30°C$

Stream No.	Actual Temperature (°C)		Interval Temperature (°C)		Interval Number	Ordered Interval Temperatures (°C)
	T_s	T_t	T_s	T_t		
1	20	160	35	175	1	275
2	120	260	135	275	2	265
3	180	20	165	5	3	175
4	280	60	265	45	4	165
					5	135
					6	45
					7	35
					8	5

The interval temperatures now have the allowance for Δt_{min}. This modification guarantees that for a given T_{int}, the actual temperature difference between the hot and cold stream will always be greater than or equal to Δt_{min}.

- All the interval temperatures for both the hot and cold streams are sorted in descending order and the duplicate intervals are removed (see Table 15.4).
- For each interval, an enthalpy balance is made. The enthalpy balance for interval *i* is calculated using the following equation:

$$\Delta H_i = (T_{\text{int},i} - T_{\text{int},i+1})\left(\sum_{j}^{N_{\text{streams}}} CP_{cj} - \sum_{j}^{N_{\text{streams}}} CP_{hj}\right) \tag{15.6}$$

where ΔH_i = net heat surplus or deficit in interval *i*
CP_c = mass specific heat of a cold stream
CP_h = mass specific heat of a hot stream

For the example problem this calculation is shown in Table 15.5. With each interval the enthalpy balance will indicate that there is either heat deficit or surplus or the interval is in heat balance. A heat surplus is negative and a heat deficit is positive.

TABLE 15.5 The Problem Table

Stream No.		1	2	3	4					
CP		40	60	45	30					
Temperature Interval	Interval Number					ΔT_{int}	$\Sigma CP_{cj} - \Sigma CP_{hj}$	ΔH_{int}	Heat Cascade	Corrected Heat Cascade
Hot utility–275	0								0	4,750
275–265	1					10	60	600	−600	4,150
265–175	2					90	30	2,700	−3,300	1,450
175–165	3					10	70	700	−4,000	750
165–135	4					30	25	750	−4,750	0
135–45	5					90	−35	−3,150	−1,600	3,150
45–35	6					10	−5	−50	−1,550	3,200
35–5	7					30	−45	−1,350	−200	4,550
5–Cold utility	8									4,550

The Second Law of Thermodynamics only allows us to cascade heat from a higher temperature to a lower temperature. Thus, a heat deficit from any interval can be satisfied in two possible ways—by using an external utility or by cascading the surplus heat from a higher temperature interval. If no external utility is used, the heat cascade column of Table 15.5 can be constructed. All intervals in this heat cascade have negative heat flows. This is thermodynamically impossible in any interval (it would mean heat flowing from lower to higher temperature). To correct this situation, we take the largest negative flow in the cascade and supply that amount of external hot utility at the highest temperature interval. This modification will make the heat cascade feasible (i.e., none of the intervals will have negative heat flows). The amount of external heat supplied to the first interval is minimum amount of hot utility required. The surplus heat in the last interval is the minimum amount of cold utility required.

For the example under consideration, the minimum hot utility required is 4,750 kW and the minimum cold utility required is 4,550 kW. These are the same targets obtained from the composite curves. It is clear, however, that the problem table algorithm is an easier, quicker, and more exact method for setting the energy targets. One can easily set up a spreadsheet to implement the problem table algorithm. Further, once the energy targets are obtained, the composite curves can be easily drawn to visualize the heat flows in the system. It should be noted that the absolute minimum utility targets for a fixed flowsheet are determined by the case of $\Delta t_{\min} = 0$.

15.1.5.2 Capital Cost Targets

To find the optimal value of $\Delta t_{\min}$ ahead of design, we need to set the targets for the capital and energy costs. As seen before, the capital cost target for a given $\Delta t_{\min}$ depends on the total area of the network, the number of shells required in the network, and the number of units required in the network. We will now establish these targets.

15.1.5.2.1 Target for Minimum Total Area for the Network

The composite curves can be divided into different enthalpy intervals at the discontinuities in the hot and cold composite curves. Between any two adjacent enthalpy intervals, if the hot streams in that interval transfer heat only to the cold streams in that interval and vice versa, then we say that vertical heat transfer takes place along the composite curves.[10] This mode of heat transfer models the pure countercurrent heat exchange.

The equation for establishing the minimum area target is based on a complex network called the "spaghetti" network. This network models the vertical heat transfer on the composite curves. In this network, for any enthalpy interval defined by the discontinuities in the composite curves, each hot stream is split into the number of cold streams in that enthalpy interval. Each cold stream in the enthalpy interval is also split into the number of hot streams in that interval. A match is made between each hot stream and each cold stream. Figure 15.9 shows this network. The area target is obtained by the following equation:

$$A_{\min} = \sum_{i}^{N_{\text{intervals}}} \frac{1}{\Delta T_{\text{LMTD}i}} \left(\sum_{j}^{N_{\text{streams}}} \frac{q_j}{h_j} \right)_{\text{i}} \tag{15.7}$$

where

$A_{\min}$ = total minimum area for the network
$\Delta T_{\text{LMTDi}} = \Delta T_{\text{LMTD}}$ for enthalpy interval i
q_{j} = heat content of stream j in enthalpy interval i
h_{f} = film plus fouling heat transfer coefficient of stream j in enthalpy interval i

The equation gives a minimum total surface area for any system where the process streams have uniform heat transfer coefficients. Townsend and Linnhoff[10] claim that for nonuniform heat transfer coefficients the equation gives an useful approximation of the minimum area, with errors being typically within 10%.

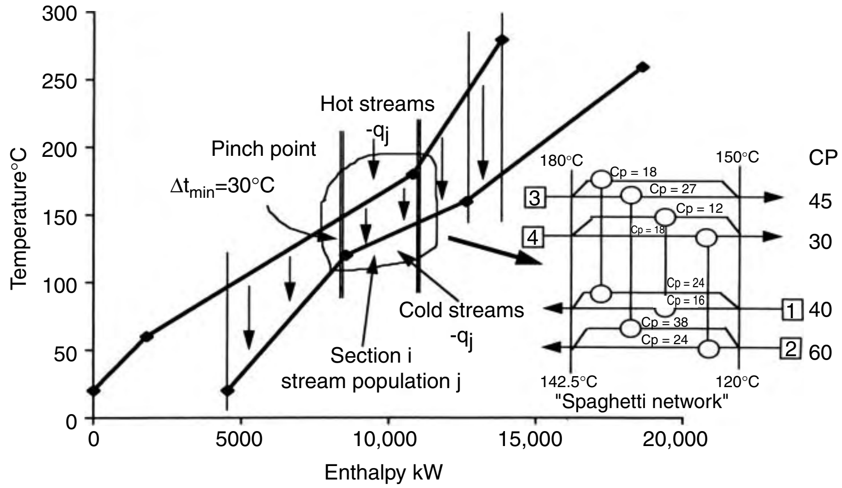


FIGURE 15.9 Enthalpy intervals for area targeting and the spaghetti network that mimics vertical heat transfer.

15.1.5.2.2 Minimum Number of Units Target

The minimum number of units is given by:[11]

$$u = N + L - s \tag{15.8}$$

where

u = number of units including heaters and coolers
N = number of streams including utilities
L = number of loops
s = number of separate networks

A loop is a closed path through the network, and its effect is to increase the number of units. For a network to have the minimum number of units, the number of loops should be zero. Figure 15.10 shows a loop on the grid diagram. Normally for a given stream system, only a single network exists and if the number of loops is assumed to be zero, then the equation can be reduced to[8]

$$u = N - 1 \tag{15.9}$$

Figure 15.10 illustrates the occurrence of independent networks. The total network consists of two independent subnetworks. Thus, for the system shown in this figure, N = 5, $L = 1$, $s = 2$, and $u = 4$.

15.1.5.2.3 Target for Minimum Number of Shells

Trivedi et al.[12] have developed a method for estimating the total number of shells required for the heat exchanger network. The method starts by setting the energy targets for the selected value of Δt_{min}. The hot and cold utilities are added to the process stream system and the composite curves for the process, and the utility streams are constructed. No external utility requirement will be needed for these set of composite curves. Hence, these are called the balanced composite curves.[13] The method estimates the number of shells based on the balanced composite curves. The method involves two steps:

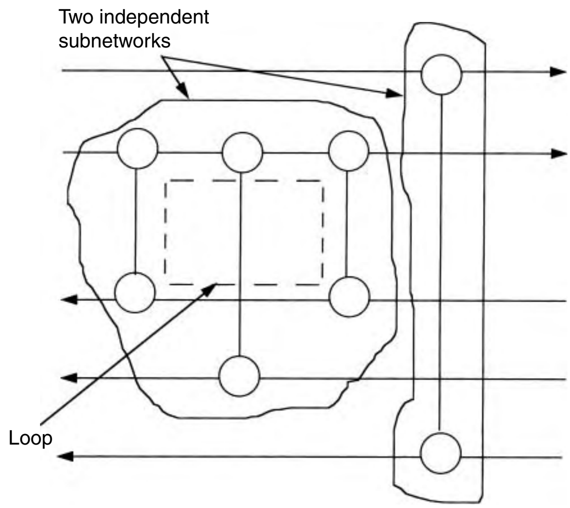


FIGURE 15.10 Loops and independent subnetworks in a heat exchanger network.

Step 1 Estimate the total number of shells required by the cold process streams and the cold utility streams. See Figure 15.11a.

- Commencing with a cold stream target temperature, a horizontal line is drawn until it intercepts the hot composite curve. From that point a vertical line is dropped to the cold composite curve. This section, defined by the horizontal line, represents a single exchanger shell in which the cold stream under consideration gets heated without the possibility of a temperature cross. In this section, the cold stream will have at least one match with a hot stream. Thus, this section implies that the cold stream will require at least one match with a hot stream. Further, it ensures that log mean temperature (LMTD) correction factor, $F_t \geq 0.8$.
- Repeat the procedure until a vertical line intercepts the cold composite curve at or below the starting temperature of that particular stream.
- The number of horizontal lines will be the number of shells the cold stream is likely to require to reach its target temperature.
- Repeat the procedure for all the cold streams including the cold utility streams.

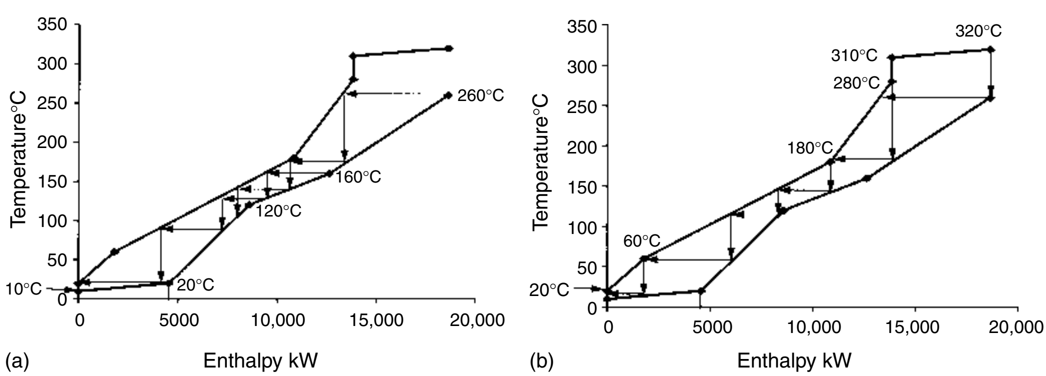


FIGURE 15.11 (a) Shell targeting for cold streams on the balanced composite curves. Cold streams will require 8 shells. (b) Shell targeting for hot streams on the balanced composite curves. Hot streams will require 9 shells.

- The sum of the number of shells for all the cold streams is the total number of shells required by the cold streams to reach their respective target temperatures.

Step 2 Estimate the total number of shells required by the hot process streams and the hot utility streams. See Figure 15.11b.

- Starting from the hot stream initial temperature, drop a vertical line on the balanced composite curve until it intercepts the cold composite curve. From this point construct the horizontal and vertical lines until a horizontal line intercepts the hot composite curve at or below the hot stream target temperature.
- The number of horizontal lines will be the number of shells required by the hot streams for heat exchange in the network.
- Repeat the procedure for all the hot streams including the hot utility streams.
- The sum of the number of shells required by the hot streams will be the total number of shells required by the hot streams to reach their respective target temperatures.

The quasi-minimum number of shells required in the network would be the larger of either the total number of shells required by the hot streams or the total number of shells required by the cold streams.

Ahmad and Smith[14] have pointed out that this procedure works for the problems whose composite curves are wide. They suggest that the procedure should be applied independently to the above and below pinch subnetworks. (Above and below pinch subnetworks are discussed in "Pinch Point and Network Design.") They have also proposed another method for finding the number of shells in a network. Their method is more sophisticated and is not within the scope of this chapter.

15.1.5.3 Optimum $\Delta t_{\min}$ Value

The procedures presented so far establish, for a selected $\Delta t_{\min}$ value, targets for minimum energy, minimum area, minimum number of units, and minimum number of shells. These targets along with the cost data can be translated into capital and energy cost for the network. The targets can be evaluated at different values of $\Delta t_{\min}$ to obtain a bird's-eye view of the solution space and the optimal value of $\Delta t_{\min}$ ahead of design. This philosophy was first proposed by Hohmann[8] and later developed by Linnhoff and Ahmad[15] as "supertargeting."

For the process flowsheet illustrated in Figure 15.6, the different targeting curves are shown in Figure 15.12(a–f). It is seen from the total annualized cost curve, Figure 15.12f, that the optimal region is very flat. While selecting a value of $\Delta t_{\min}$ from the optimal region, a couple of points should be kept in mind. Different values of $\Delta t_{\min}$ lead to different topologies. Hence, we should take into account different factors that can affect the final cost. They are the wideness of the composite curves and the problem constraints that have significance in the network synthesis and refinement. The optimal value of $\Delta t_{\min}$ selected for this problem is 40°C and the target values for energy, area, units, shells, and annualized total cost are given in Table 15.6.

15.1.6 The Pinch Point

On the composite curves, are one or more enthalpy values for which the two composite curves are $\Delta t_{\min}$ apart. For the example under consideration (see Figure 15.8), this occurs at a hot stream temperature interval in the problem table. In this interval, the heat cascade has zero heat flow (i.e., no heat is transferred across this interval when minimum hot utility is used). This interval, identified by a zero in the corrected heat cascade column, is referred to as the pinch point.[11,16] The significance of the pinch point is now clear—for the minimum external utility requirement do not transfer heat across the pinch point. Any extra amount of external heat that is put into the system above the minimum will be transferred across the pinch point and will be removed by the cold utility.[17]

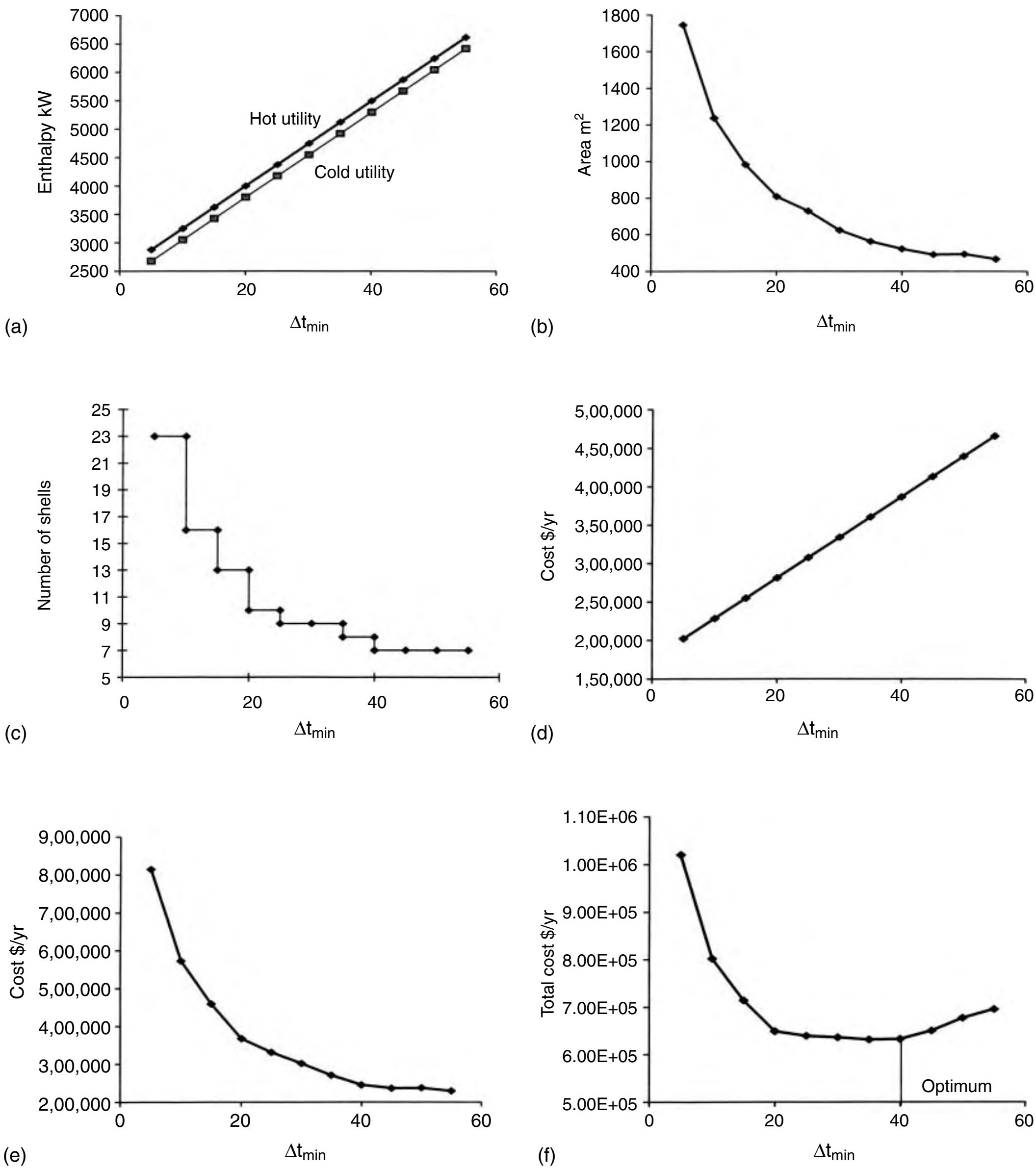


FIGURE 15.12 (a) Minimum energy targeting plot. (b) Minimum network area targeting plot. (c) Targeting plot for minimum number of shells. (d) Minimum energy cost targeting plot. (e) Minimum annualized capital cost targeting plot. (f) Targeting plot for total annualized cost.

15.1.6.1 Cross Pinch Principle

For a given value of Δt_{min}, if the network is using Q_{h} units of hot utility and if Q_{hmin} is the minimum energy target,[17] then

$$Q_{\mathrm{h}} = Q_{\mathrm{hmin}} + \alpha \tag{15.10}$$

TABLE 15.6 Comparison of Target and Design Values for Different Network Optimization Variables at $\Delta t_{\min} = 40°C$

	Target	Figure 15.24 MER Design		Figure 15.25 Design	
			% Deviation from Target		% Deviation from Target
Total network area, m^2	687	779	13	708	3
Total hot utility consumption, kW	5,500	5,500		6,400	16
Total cold utility consumption, kW	5,300	5,300		6,200	17
Number of units	7	7		6	
Number of shells	8	9		7	
Energy cost, $/yr	387,250	387,250		450,700	16
Annualized capital cost $/yr	305,407	333,576	9	282,996	−7
Total annualized cost, $/yr	692,657	720,826	4	733,696	6

If the network uses Q_c units of cold utility and if Q_{cmin} is the minimum energy target, then

$$Q_c = Q_{\text{cmin}} + \alpha \tag{15.11}$$

where

α = the amount of cross-pinch heat transfer.

15.1.6.2 The Pinch Point and Its Significance

The pinch point divides the stream system into two independent subsystems. The subsystem above the pinch point is a net heat sink, and the subsystem below the pinch point is a net heat source. Thus, for a system of hot and cold streams, to design a network that meets the minimum utility targets the following rules set by the pinch principle should be follows:[17]

- Do not transfer heat across the pinch point.
- Do not use hot utility below the pinch point.
- Do not use cold utility above the pinch point.

We shall use these principles to design the network for the optimal value of $\Delta t_{\min}$.

15.1.7 Network Design

15.1.7.1 Network Representation on the Grid Diagram

If we attempt to design the network using the conventional flowsheet format, any changes in the design may lead to redrawing the flowsheet. Hence, we shall use the grid diagram to represent and design the network. Figure 15.13 shows the grid diagram for the flowsheet shown in Figure 15.6. On this grid diagram we can place a heat exchanger match between two streams without redrawing the whole stream system. The grid representation reflects the countercurrent nature of heat transfer that makes it easier to check temperature feasibility of the match that is placed. Furthermore, we can easily represent the pinch point on the grid diagram as shown in Figure 15.13.

15.1.7.2 Pinch Point and Network Design

We now continue with the design of the example problem. Figure 15.8 shows the composite curves with $\Delta t_{\min} = 30°C$ for the process flowsheet. It also shows the pinch point. The pinch point is the most constrained region of the composite curves. From optimization principles, we know that for a

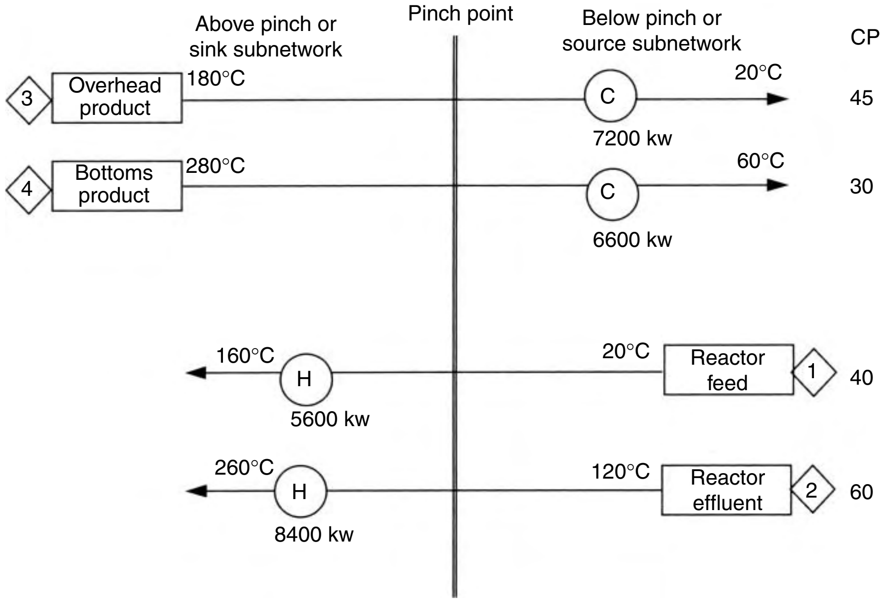


FIGURE 15.13 Grid diagram representation of the network in the flowsheet of Figure 15.9.

constrained problem, the optimal solution is located at the point formed by the intersection of multiple constraints. Hence, if we can satisfy the constraints at that point then we are guaranteed an optimal design. Thus, we should start the design where the problem is most constrained, that is, the pinch point.[17] The pinch point divides the problem into two independent subnetworks—an above pinch subnetwork that requires only hot utility and a below pinch subnetwork that requires only cold utility.

15.1.7.2.1 Pinch Design Rules and Maximum Energy Recovery

Let us make some observations at the pinch point. Immediately above the pinch point,[17]

$$\sum CP_h \leq \sum CP_c \tag{15.12}$$

and immediately below the pinch point,

$$\sum CP_h \geq \sum CP_c \tag{15.13}$$

This condition defines the pinch point, that is, the constraint which the designer has to satisfy. Thus, each and every match in the sink subnetwork at the pinch point should be placed such that $CP_h \leq CP_c$. Similarly, each and every match in the source subnetwork at the pinch point should be placed such that $CP_c \leq CP_h$. Figure 15.14 shows these conditions on the *T–H* diagram.

Two situations arise where the above condition may not be satisfied. The first condition is shown in Figure 15.15. The solution is to split the hot stream. Another situation is shown in Figure 15.16. Here we will have to split the cold stream.

Consider the situation shown in Figure 15.17, which shows the stream population at the pinch point in a sink subnetwork. Here, the number of hot streams is greater than the number of cold streams. Two matches can be easily placed, but for the third stream there is no cold stream to match unless we split a cold stream. Thus, immediately at the pinch for a sink subnetwork, the stream population constraint tells us that the number of hot streams should be less than or equal the number of cold stream. Applying the same logic to the source subnetwork immediately at the pinch point, the number of cold streams should

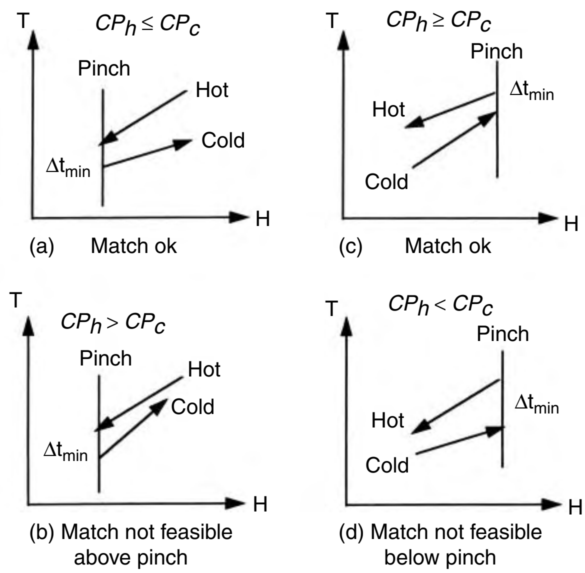


FIGURE 15.14 CP matching rules at the pinch point.

be less than or equal to the number of hot streams. If this condition is not satisfied, then split a hot stream. Figure 15.18 shows the algorithms that are developed for placing matches immediately at the pinch point for both the above and below pinch subnetworks.

The matches and stream splitting can be easily identified with the CP table.[17] Consider the stream system shown in Figure 15.19. The CP table is shown in the same figure. The first row in this table contains the conditions that need to be satisfied for the subnetwork under consideration. The CP values of the hot and cold streams are arranged in a descending order in the columns. The stream numbers are shown in brackets adjacent to the CP values. For the sink subnetwork under consideration, the CP value for the hot streams is listed in the left column and the CP value for the cold stream is listed in the right column. There is no feasible match for hot stream 1. Hence, it will have to be split. Hot stream 2 can match with either stream 3 or 4. Once we split hot stream 1 we violate the stream population constraint. To satisfy this constraint we will have to split a cold stream. Either stream 3 or 4 can be split. The designer can use his or her judgment and decide which stream should be split depending on controllability and other physical constraints. For example, it is not advisable to split a stream that may have a two-phase flow.

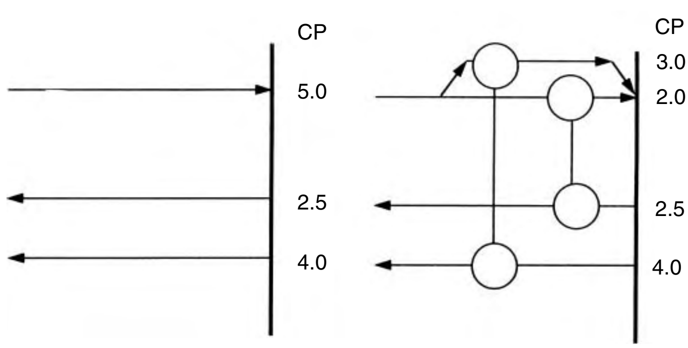


FIGURE 15.15 Stream splitting when CP rule is not satisfied.

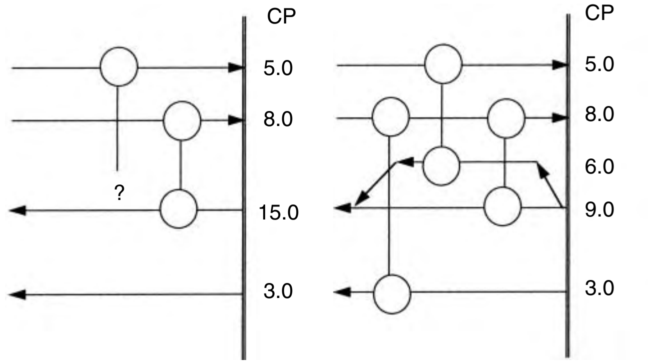


FIGURE 15.16 Stream splitting at the pinch point when CP rule is not satisfied for all hot streams above the pinch point.

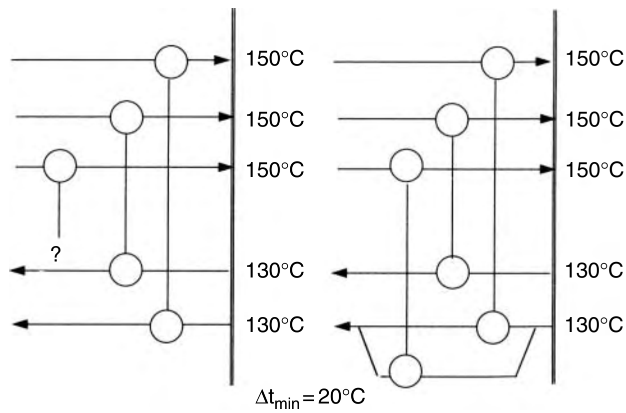


FIGURE 15.17 Stream population at the pinch point may force a stream to split.

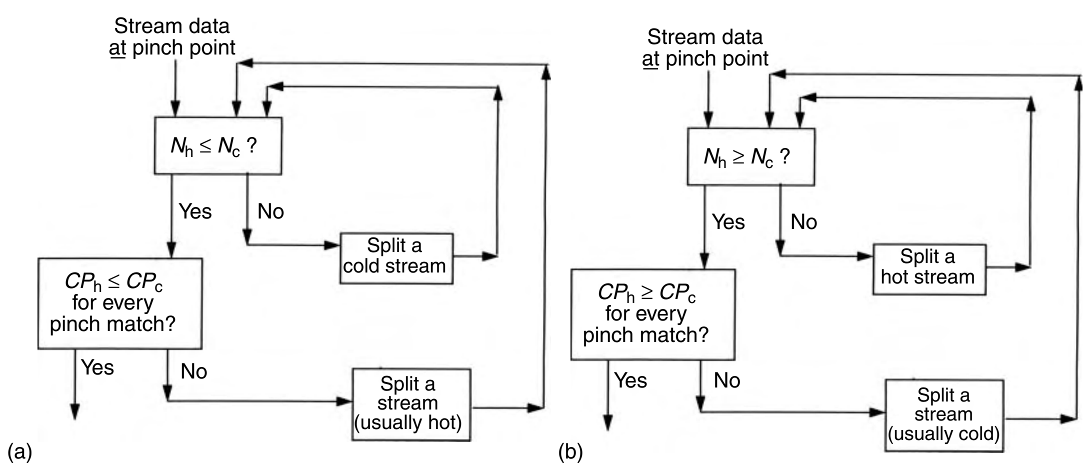


FIGURE 15.18 (a) Algorithm for sink subnetwork design at the pinch point. (b) Algorithm for source subnetwork design at the pinch point.

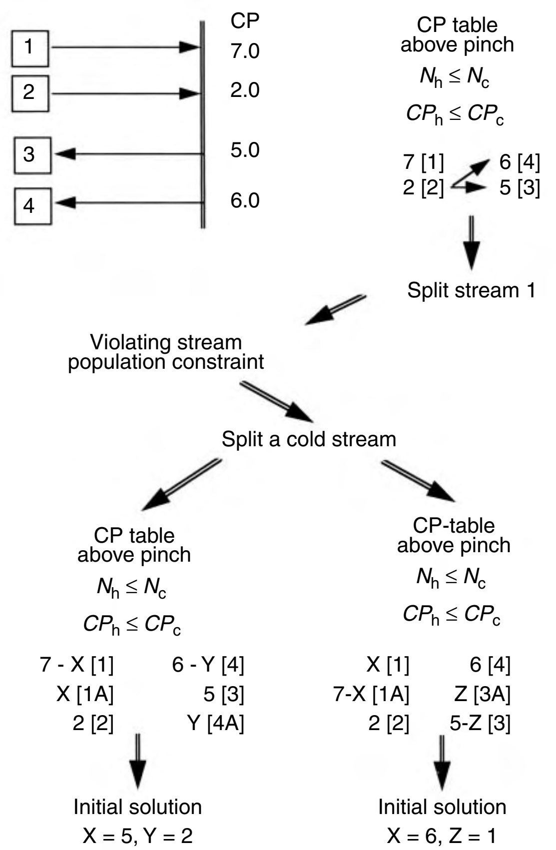


FIGURE 15.19 Identifying matches and stream splitting using the CP table.

Two different designs will be obtained depending on which stream is split. The hot stream is split into two streams having CP values of X and $7-X$. In the first alternative, we split stream 4 into streams having CP values of Y and $6-Y$. In the second option we split stream 3 into streams having CP values of Z and $5-Z$. To find the initial values of X, Y, and Z it is recommended that all matches except for one are set for CP equality. Thus, for the first option set $X=5$ and $Y=2$. For the second option set $X=6$ and $Z=1$. This is an initial solution. The values of X, Y, and Z can be adjusted to obtain an optimal design. The two different topologies obtained for this example are shown in Figure 15.20.

Once the matches are identified at the pinch, the heat loads for these matches are fixed to maximize the heat exchange to the limit of heat load of either the hot stream or the cold stream. This will eliminate that stream from the analysis. Fixing the heat loads for match with this heuristic will also help us to minimize the number of units and thus the installation cost.[17]

Matches away from the pinch point do not have to satisfy the condition just outlined, as the system is not constrained in this region. Matches are easily identified away from the pinch point. The network obtained will satisfy the energy targets set. Such a network is called maximum energy recovery (MER) network.[17]

15.1.7.3 $\Delta T-T_c$ Plot

To appropriately place matches in the network that utilize the correct driving forces, the driving force plot is very useful design aid.[19] A plot of the temperature difference ΔT between the hot composite curve

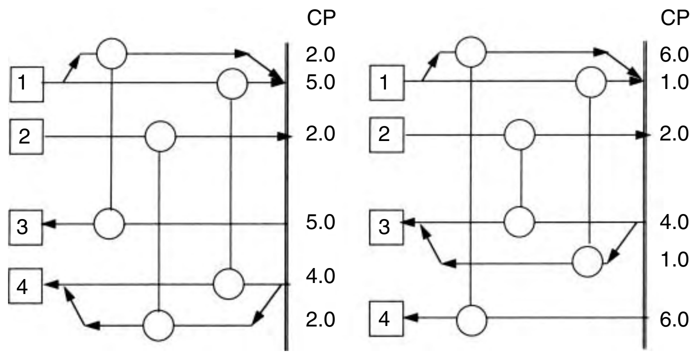


FIGURE 15.20 Two different topologies for the problem in Figure 15.23.

and the cold composite curve vs. the cold composite temperature is derived from the composite curves as shown in Figure 15.21. Different versions of this plot can also be used such as the $\Delta T - T_h$ plot or $T_h - T_c$ plot. To meet the area targets established one must place the matches such that the $\Delta T - T_c$ plot of the exchanger must fall in line with the $\Delta T - T_c$ plot obtained from the composite curves. This is a good guiding principle when the heat transfer coefficients of all the streams are in the same order of magnitude.

Figure 15.22 shows two matches placed in the network. The match in Figure 15.22a is a good match, as it makes the correct utilization of the available driving forces. The match in Figure 15.22b is a bad match, as it uses more ΔT than available. If a match uses more driving force than is available, it will force some other match in the network to use less driving force than available. This results in an overall increase in the total surface area of the network. Badly placed matches do not waste energy but waste surface area.[3]

15.1.7.4 Remaining Problem Analysis

As we have noted, the rules that are established for design guidelines can identify multiple options. The option that satisfies the targets established ahead of design will be the optimal design. To help screen different options, Remaining Problem Analysis (RPA) technique[19] should be used. RPA evaluates the impact of the design decision made on the remaining problem. Once a matching decision is made, the remaining stream data excluding the match is retargeted. The targets for the remaining problem should be in line with the original targets. In case they are not, look for other options such as decreasing the heat duty on the match or finding a new match.

Figure 15.23 shows how RPA is undertaken. For the original problem the targets are

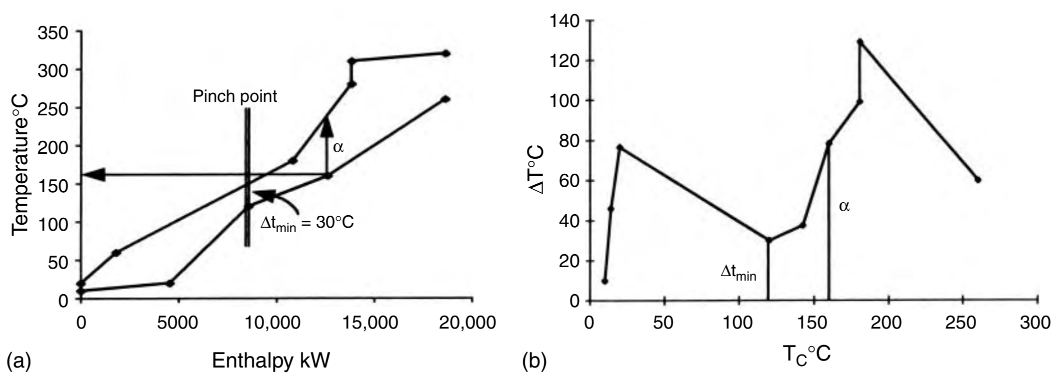


FIGURE 15.21 Construction of the $\Delta T - T_c$ plot.

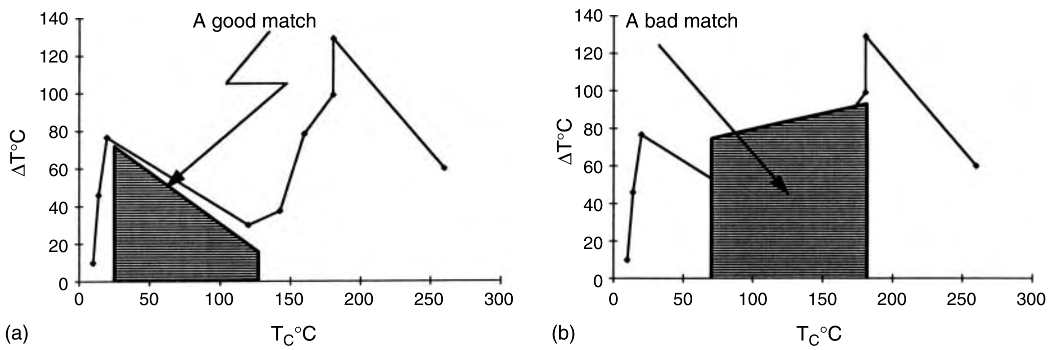


FIGURE 15.22 Identifying good and bad matches using the $\Delta T - T_c$ plot.

Q_{hmin}	= 30 MW
Q_{cmin}	= 20 MW
Minimum area	= 1,000 m^2
Minimum number of shells	= 15

A match is placed that has a duty of 40 MW, has a 200 m^2 area, and needs 3 shells (see Figure 15.23b). The targets for the remaining stream data that excluded the match are

Q_{hmin}	= 40 MW
Q_{cmin}	= 30 MW
Minimum area	= 900 m^2
Minimum number of shells	= 14

Thus, by placing the selected match, 10 MW of extra energy will be required, 100 m^2 more area will be required, and two extra shells will be needed. If we place a new match as shown in Figure 15.23c the remaining problem targets are in line with the original targets (i.e., only a small penalty is associated with this match). Hence this match should be selected.

The procedure is repeated after placing each and every match. This way the designer gets continuous feedback on the impact of the matching decisions and the heat load assigned to the match. The final design will meet all the targets set before we start the design.

15.1.7.5 Local Optimization—Energy Relaxation

The optimal region is generally very flat, and the optimal value of Δt_{min} is not a single point but a region. The targeting exercise helps identify this region. Once a value of Δt_{min} is selected, and the network is designed using the methodology outlined earlier, there is still some scope for further optimization. This is achieved by a process called energy relaxation.[17]

The number of units in the design obtained using this procedure will generally be greater than the target value. This is due to the presence of loops. If we can eliminate the loops in a network, then the heat transfer area is concentrated on fewer matches and hence will decrease the piping and foundation requirements. This will tend to decrease the capital cost of the network. However, the same energy penalty may be incurred.

Identifying loops in a large network cannot be done visually as we had done in the example problem. Various algorithms exist to identify loops in a given network.[20,21] Trivedi et al.[22] have proposed LAPIT (Loop and Path Identification Tree) for identifying loops and paths present in the network. A path is the connection between a heater and a cooler via process-process matches.

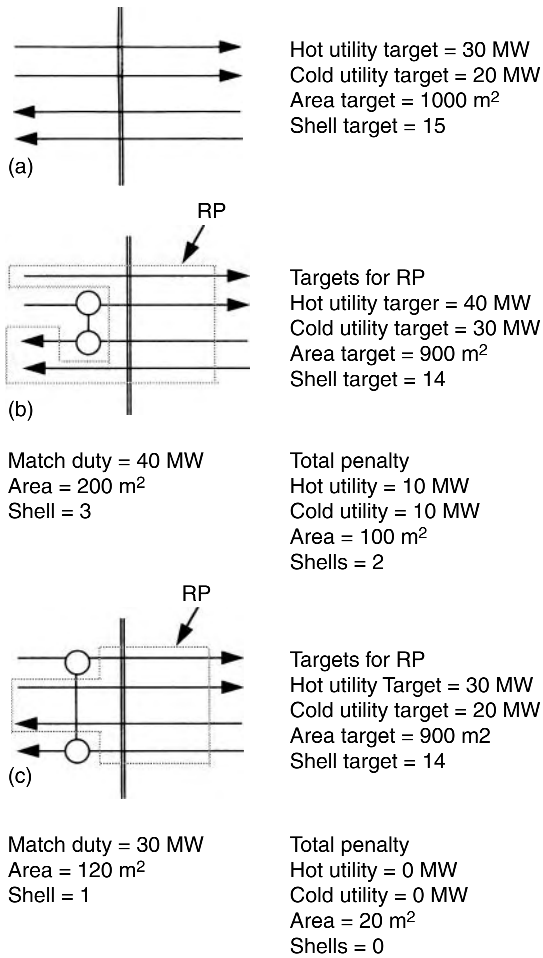


FIGURE 15.23 Remaining problem analysis.

Loops can be broken by removal of unit in a loop and redistributing the load of the unit among the remaining units of the loop. Some exchangers may result having a very small $\Delta t_{\min}$ when a loop is broken. The $\Delta t_{\min}$ across such exchangers can be increased by increasing the utility consumption along a path that consists of a heater, the unit having a small value of $\Delta t_{\min}$, and a cooler.

Loop breaking is a very complex optimization process. Trivedi et al.[22,23] have proposed a detail method that systematically breaks loop and identifies options available at each step of the process. The method is based on loop network interaction and load transfer analysis (LONITA)[23] and on a best-first-search procedure.[22]

15.1.7.6 Summary of the Design Procedure

The pinch design procedure can be summarized now. Establish the value of the optimum $\Delta t_{\min}$ using the targeting techniques. This identifies the region in which the design should be initialized. Using the pinch design procedure, design the network. Reduce the number of units using the loop breaking and energy relaxation techniques outlined earlier.

15.1.7.7 Example

The principles and procedures just discussed are used for the design of the network for the flowsheet of Figure 15.6. The value of the optimum Δt_{min} is 40°C. The final network MER design is shown in Figure 15.24. Note that the annualized capital cost of this network is only about 4% higher than the target value of $692,700. This design is in the neighborhood of the predicted optimum.

The design after energy relaxation is shown in Figure 15.25. This design has one less unit than the MER design in Figure 15.24. Hence, this will decrease the piping, installation, and maintenance cost. The cost of the new design is about 2% higher than the MER design. Since we are using simplified models for costing, the cost of the design in Figure 15.25 is within the errors of cost estimating. However, this design will be cheaper when detail cost is evaluated, as it has less number of units.

Table 15.6 compares the target values of the various optimization parameters with actual design values for Figure 15.24 and 15.25.

15.1.8 The Dual Temperature Approach Design Method

15.1.8.1 Concept

The pinch design procedure fixes the Δt_{min}, designs a network, and then undertakes local optimization. During the process of local optimization either the constraint of Δt_{min} across exchangers may be relaxed or the energy consumption may be increased. Thus, the final network will have a Δt_{min} less than the Δt_{min} on the composite curves corresponding to the final energy consumption.

The final network for the flowsheet in Figure 15.6 using pinch design is shown in Figure 15.25. In this network the energy consumption correspond to a Δt_{min} of 52°C on the composite curves. However, the minimum approach temperature is 40°C across exchanger 2. The dual temperature approach design method for the heat exchanger network synthesis is based on this observation.

The network is characterized by two approach temperatures. One approach temperature is called the network minimum approach temperature Δt_{Nmin}. This is the minimum approach temperature on the composite curves and thus fixes the energy consumption for the network. The other approach temperature is called the exchanger minimum approach temperature Δt_{Emin}. This is the minimum approach temperature that any exchanger will have in the network. In the network of Figure 15.25, $\Delta t_{Nmin} = 52$°C and $\Delta t_{Emin} = 40$°C. Generally, $\Delta t_{Emin} \leq \Delta t_{Nmin}$. When $\Delta t_{Emin} = Dt_{Nmin}$ the network has a

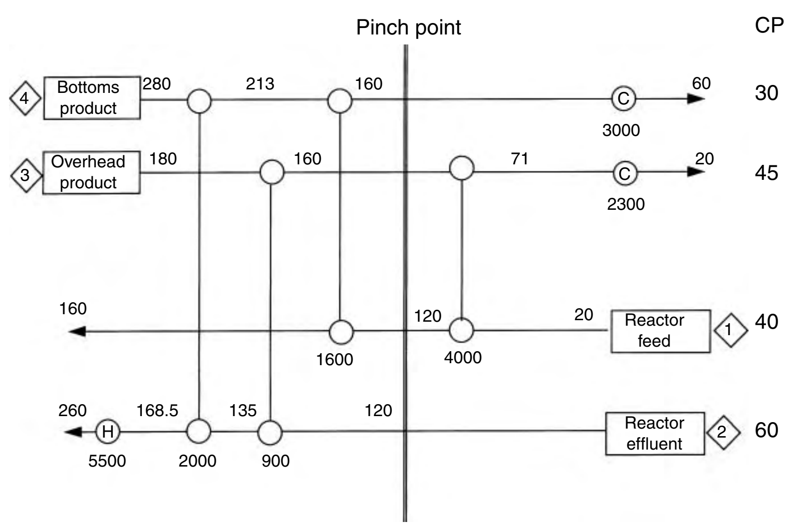


FIGURE 15.24 Pinch design for the flowsheet in Figure 15.9. Total annualized cost is $720,800.

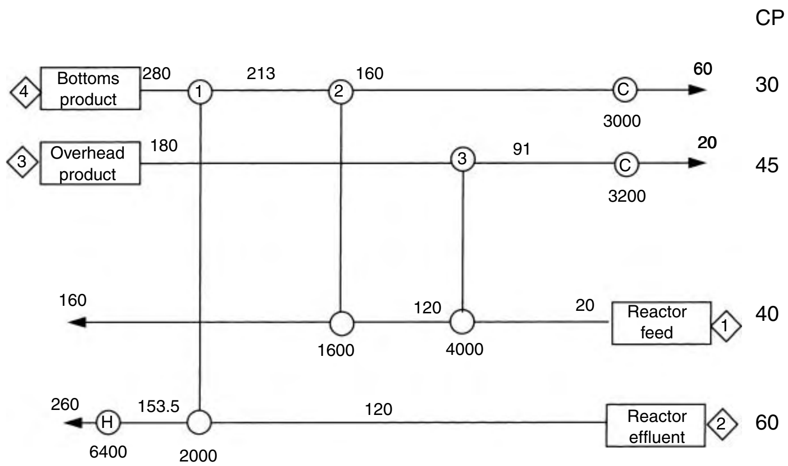


FIGURE 15.25 Energy of the pinch design in Figure 15.30. Total annualized cost is \$773,700. $\Delta t_{\text{Nmin}} = 52$ K and $\Delta t_{\text{Emin}} = 40$ K.

global minimum approach temperature just like the $\Delta t_{\min}$ parameter of the pinch design procedure. This is in contrast to the pinch design method, which utilizes a global value of $\Delta t_{\min}$ (i.e., $\Delta T_{\text{Nmin}} = \Delta t_{\text{Emin}}$) and then relaxes the energy consumption of the network (with the obvious implication of increase in $\Delta t_{\min}$).

The concept of using dual approach temperatures for the design of the network is very powerful as the designer retains complete control over the energy consumption of the network. This is in contrast with the pinch design method where energy consumption increases following loop breaking and is determined by the new topology and not the designer. Another advantage in using the dual temperature method is the reduction in the number of shells and units. As well, unnecessary stream splits may be avoided, making a more practical network with better operability characteristics. This may result in substantial decrease (~20%) in the capital cost for a fixed operating cost, representing millions of dollars savings in industrial applications.[24]

The design procedure proposed by Trivedi et al.[25] incorporates the best features of both the pinch design method and the dual temperature approach design procedure. Their design procedure is based on the concept of a pseudo-pinch point.

15.1.8.2 Location of the Pseudo-Pinch Point

Initially two sets of composite curves are generated using Δt_{Nmin} and Δt_{Emin}, with Δt_{Nmin} selected such that $\Delta t_{\text{Nmin}} \geq \Delta t_{\text{Emin}}$. Both sets will have different energy consumption denoted as EC. As $\Delta t_{\text{Emin}} \leq \Delta t_{\text{Nmin}}$, $\text{EC}(\Delta t_{\text{Nmin}}) \geq \text{EC}(\Delta t_{\text{Emin}})$.

Let us define the energy difference α (see Figure 15.26a and b) as

$$\alpha = \text{EC}(\Delta t_{\text{Nmin}}) - \text{EC}(\Delta t_{\text{Emin}}) \tag{15.14}$$

As the network is designed employing two approach temperatures, a new interpretation of the problem could be this: As Δt_{Emin} is less than Δt_{Nmin}, an amount of energy α traverses the Δt_{Emin} conventional pinch point. This can be compensated by energy carried upward across the pinch because heat exchange is allowed at temperature differences as low as Δt_{Emin} between streams. The maximum amount of this upward carriage that can be achieved is the value of α defined earlier. This represents additional flexibility in stream matching by providing for upward and downward carriage, which is permitted by the reduced Δt_{Emin} between streams, while the total energy requirement remains fixed at $\text{EC}(\Delta t_{\text{Nmin}})$, which in turn is defined by the network approach temperature Δt_{Nmin}.

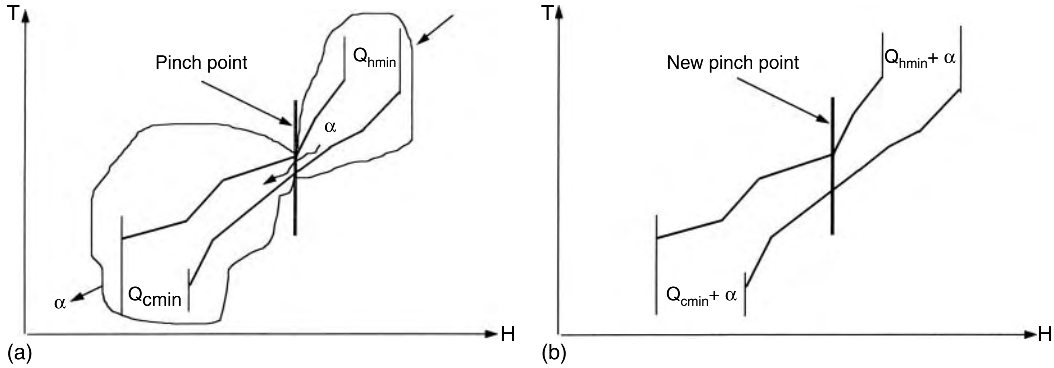


FIGURE 15.26 (a) Energy relaxation with heat transfer across the pinch point. (b) New pinch point with minimum energy consumption plus (α).

A pseudo-pinch point (actually a set of stream temperatures) is defined so that the stream temperatures at this point allow the problem to be partitioned as shown in Figure 15.27. Thus, the two parts of the network are in enthalpy balance with the utility consumption determined by Δt_{Nmin}.

The pseudo-pinch point is defined based on the observation that the changes in slope of the composite curve that occur at the conventional pinch point require (at least) one hot or cold stream entering at a real pinch point temperature.[26] This stream's starting temperature is chosen to determine the pseudo-pinch point. Let the subscripts pp denote the pseudo-pinch temperatures. To determine the pseudo-pinch point temperatures with α units of heat transferred across the conventional pinch point at Δt_{Emin}, various strategies for allocating the α units are proposed by Trivedi et al.[25] Generally the following heuristics are used:

1. Pseudo-pinch point temperatures having $\Delta T = \Delta t_{\mathrm{Emin}}$: If a hot stream with a starting temperature T_{hs} determines the conventional pinch point for Δt_{Nmin} or Δt_{Emin}, then this stream is also assumed to be at the pseudo-pinch point with all the cold stream temperatures given by $T_{\mathrm{hs}} - \Delta t_{\mathrm{Emin}}$. Likewise, if a cold stream entrance determines the conventional pinch point, then the pseudo-pinch temperatures of all the hot streams are given by $T_{\mathrm{cs}} + \Delta t_{\mathrm{Emin}}$. It is possible that different

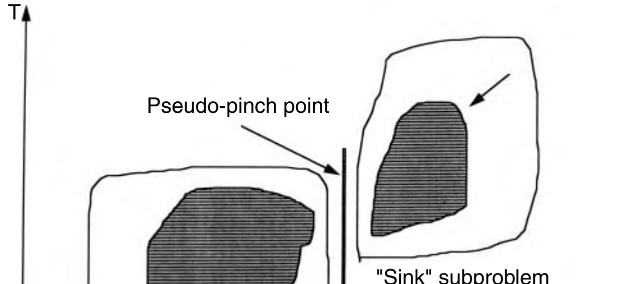


FIGURE 15.27 Pseudo-pinch division of the system.

streams may determine the conventional pinch point for both Δt_{Emin} and Δt_{Nmin}. In such a situation different topologies will result and all possible configurations should be investigated.

2. Pseudo-pinch point temperatures with $\Delta T > \Delta t_{\text{Emin}}$: The α units of heat carried across the conventional pinch point will increase the pseudo-pinch point ΔT's for some stream combinations. Assuming that a hot stream determines the conventional pinch point, various strategies are used. Generally, if there are N other hot streams with $T_{\text{hsj}} = T_{\text{hspp}}$, then $\alpha_j = \alpha / N$ is allocated to each of these streams so that their pseudo-pinch point temperatures are increased by α/NCP_j.

15.1.8.3 Design of the Sink Subproblem

15.1.8.3.1 Feasibility Criteria at the Pseudo-Pinch

As the problem constraints are relaxed at the conventional pinch point by passing heat across it, a wide variety of network topologies can be generated. The method commences the design at the pseudo-pinch point. As there are only a few essential matches to be made at the pseudo-pinch point, all the options available to the designer are readily identified. The designer, utilizing the process knowledge and experience, can now select the necessary matches. These could include imposed and constrained matches required for safe, controllable and practical network or any other preferences the designer may have. The feasibility criteria for stream matching and splitting follow.

Number of Process Streams and Branches. In a conventional pinch situation, the stream population is compatible with a minimum utility design only if a pinch match is found for each hot stream above the conventional pinch point. The situation is illustrated in Figure 15.17, where stream splitting is unavoidable. However, for the pseudo-pinch situation, splitting may not be necessary as demonstrated in Figure 15.28. This results from the relaxation of the conventional pinch temperatures resulting in an increase in the available driving forces for the pseudo-pinch matches.

It is exceedingly difficult to determine in advance if stream splitting is necessary. Hence, two approaches are suggested. The first approach is identical to the pinch design method. Stream splits may be removed once the initial design is generated. This will be discussed later. In the second approach, we place the pseudo-pinch matches and see if the unmatched streams at the pseudo-pinch can be

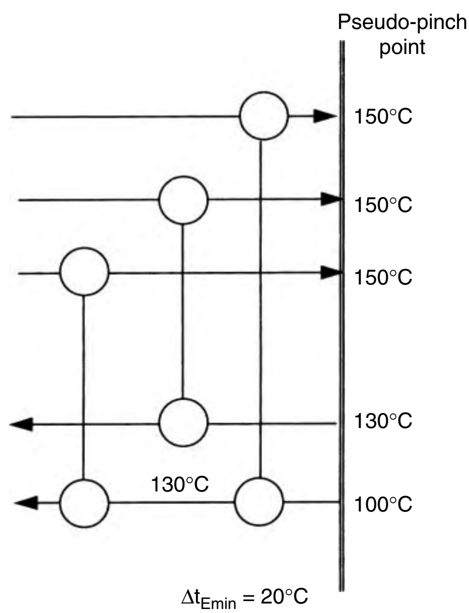


FIGURE 15.28 Stream splitting may not be necessary in pseudo-pinch problems.

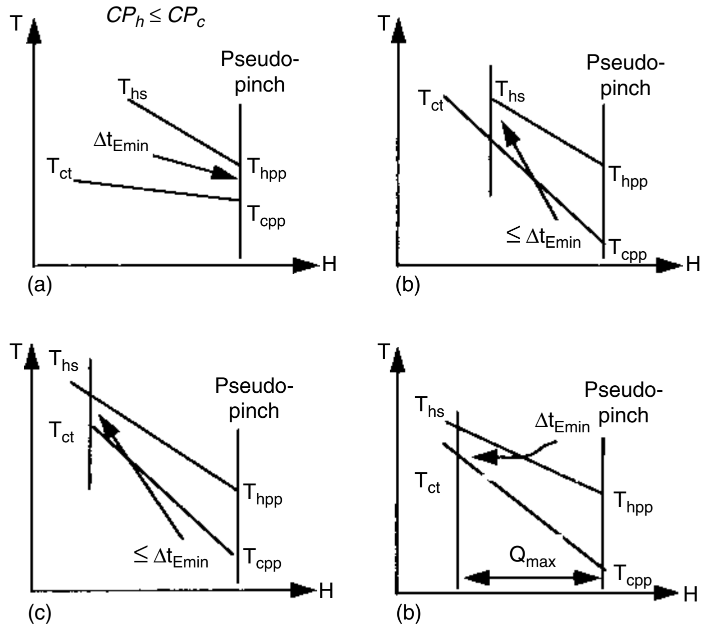


FIGURE 15.29 Temperature feasibility for stream splitting immediately adjacent to the pseudo-pinch point for the sink subproblem.

satisfied. If the streams are matching violating Δt_{Emin}, then a stream is split and the pseudo-pinch matching is repeated.

Temperature Feasibility Criteria. Four temperature profiles are possible for a match at the pseudo-pinch point as shown in Figure 15.29. These profiles are for the sink subproblem. The match illustrated in Figure 15.29a possesses the same characteristics as that of an above conventional pinch match. However, for a pseudo-pinch, other matches of the type illustrated in Figure 15.29b, c, and d are also possible. Trivedi et al.[25] have outlined the various criteria to be used at the pseudo-pinch that determine the maximum heat load Q_{max} for a match between a hot and a cold stream.

15.1.8.4 Design of the Source Subproblem

For the source subproblem analogous feasibility criteria can readily be developed. These depend both on the stream populations and temperature feasibility shown in Figure 15.30. The complete design procedure is summarized in Figure 15.31a and b.

15.1.8.5 Design of the Remaining Network

Once the pseudo-pinch matches are placed, the remaining problem can be designed employing the rules mentioned earlier depending on the subproblem. It should be noted that the remaining problem is still in enthalpy balance and only requires either a hot or cold utility, depending on the subproblem. The utilities can be placed on either end or at an appropriate temperature level. (The rules outlined by Trivedi et al.[25] can also be employed in selecting the nonpinch matches when the pinch design method is employed.)

Once an initial topology is synthesized, its capital cost can be further reduced by decreasing the number of units and removing stream splits keeping the energy consumption constant. A linear program can be used to remove stream splits.[25] To reduce the number of units, LONITA can be used.

15.1.8.6 Example

We shall illustrate the dual temperature design method for the flowsheet shown in Figure 15.6. A final topology obtained by employing the pinch design method is illustrated in Figure 15.25. In this design

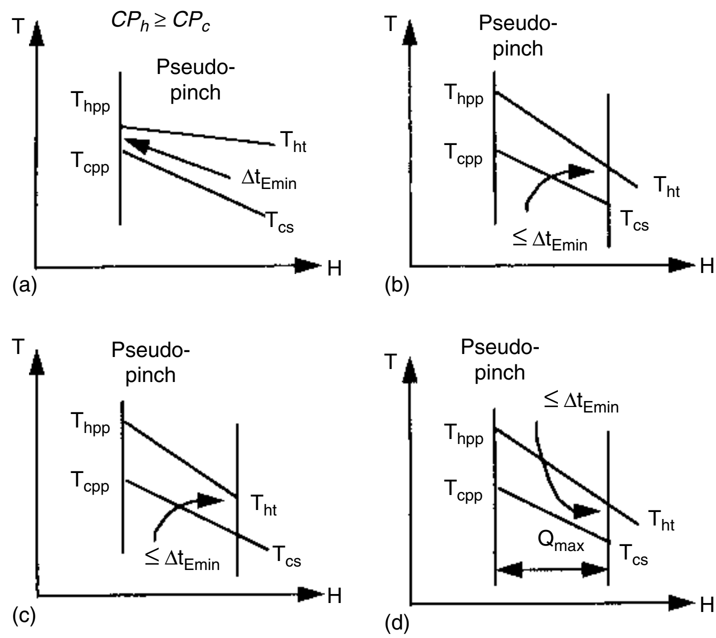


FIGURE 15.30 Temperature feasibility for stream splitting immediately adjacent to the pseudo-pinch point for the source subproblem.

$\Delta t_{Emin} = 40°C$ and $\Delta t_{Nmin} = 52°C$. For $\Delta t_{Emin} = 40°C$, the pinch point is determined by the entrance of cold stream 2. Hence, the pseudo-pinch temperature for all the hot streams is 160°C and for the cold stream 2 is 120°C. As the energy consumption of the network corresponds to $\Delta t_{Nmin} = 52°C$, $\alpha = 900$ kW. This additional heat load is carried by stream 1, and hence the pseudo-pinch temperature of cold stream 1 is 97.5°C.

15.1.8.6.1 Above Pseudo-Pinch Design

The two possible designs are illustrated in Figure 15.32a and b. It is interesting to note that the topology of Figure 15.32b will not be identified if the pinch design method is employed.

15.1.8.6.2 Below Pseudo-Pinch Design

Again two designs are possible as illustrated in Figure 15.32c and d. The matches suggested in Figure 15.32c will not be identified by the use of pinch design rules.

The above and below pseudo-pinch designs can be combined to create four different fixed-energy designs. Combining Figure 15.32a and c gives the same design as shown in Figure 15.25. The other designs are illustrated in Figure 15.33(a–c). The capital cost of the designs obtained so far are within 3% of the one shown in Figure 15.24. As the cost of all these networks is very close, we will choose the final design based on other factors such as operability, controllability, and flexibility. A detailed analysis may have to be under taken on all these designs.

The design shown in Figure 15.33a has utility matches for all the streams. This design will probably be the most controllable design. Hence it is selected. The final process flow diagram for the flowsheet in Figure 15.6 is shown in Figure 15.33d. Observe that there is a heater on the reactor feed before it enters the reactor. This heater controls the reactor feed temperature. Similarly, the heater on the reactor effluent stream to the distillation column controls the distillation column feed enthalpy. The coolers on the products will control the final product temperatures.

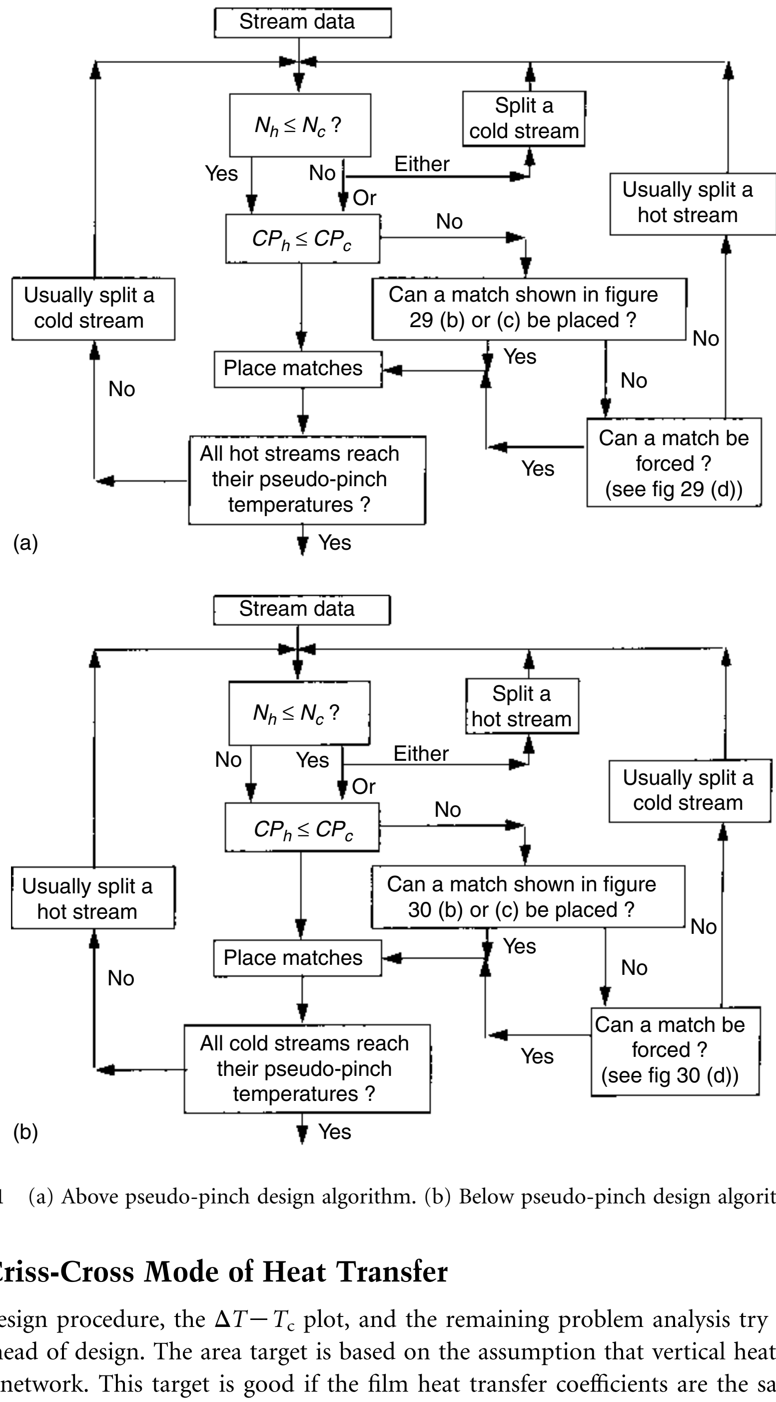


FIGURE 15.31 (a) Above pseudo-pinch design algorithm. (b) Below pseudo-pinch design algorithm.

15.1.9 Criss-Cross Mode of Heat Transfer

The pinch design procedure, the $\Delta T - T_c$ plot, and the remaining problem analysis try to achieve the targets set ahead of design. The area target is based on the assumption that vertical heat transfer takes place in the network. This target is good if the film heat transfer coefficients are the same for all the streams. In the majority of the processes phase changes will be occurring, and the streams will be in different phases. Thus, a process may consist of streams that have, at times, film heat transfer coefficients that differ by an order of magnitude. Such as situation is shown in Figure 15.34, where there are two gas streams and two liquid streams. A network that criss-crosses on the composite curves may have a smaller area than predicted by vertical heat transfer mode.[27]

A network will generally have different types of exchangers. Further, each exchanger will have different material of construction and design temperature and pressure. These factors contribute to the capital cost

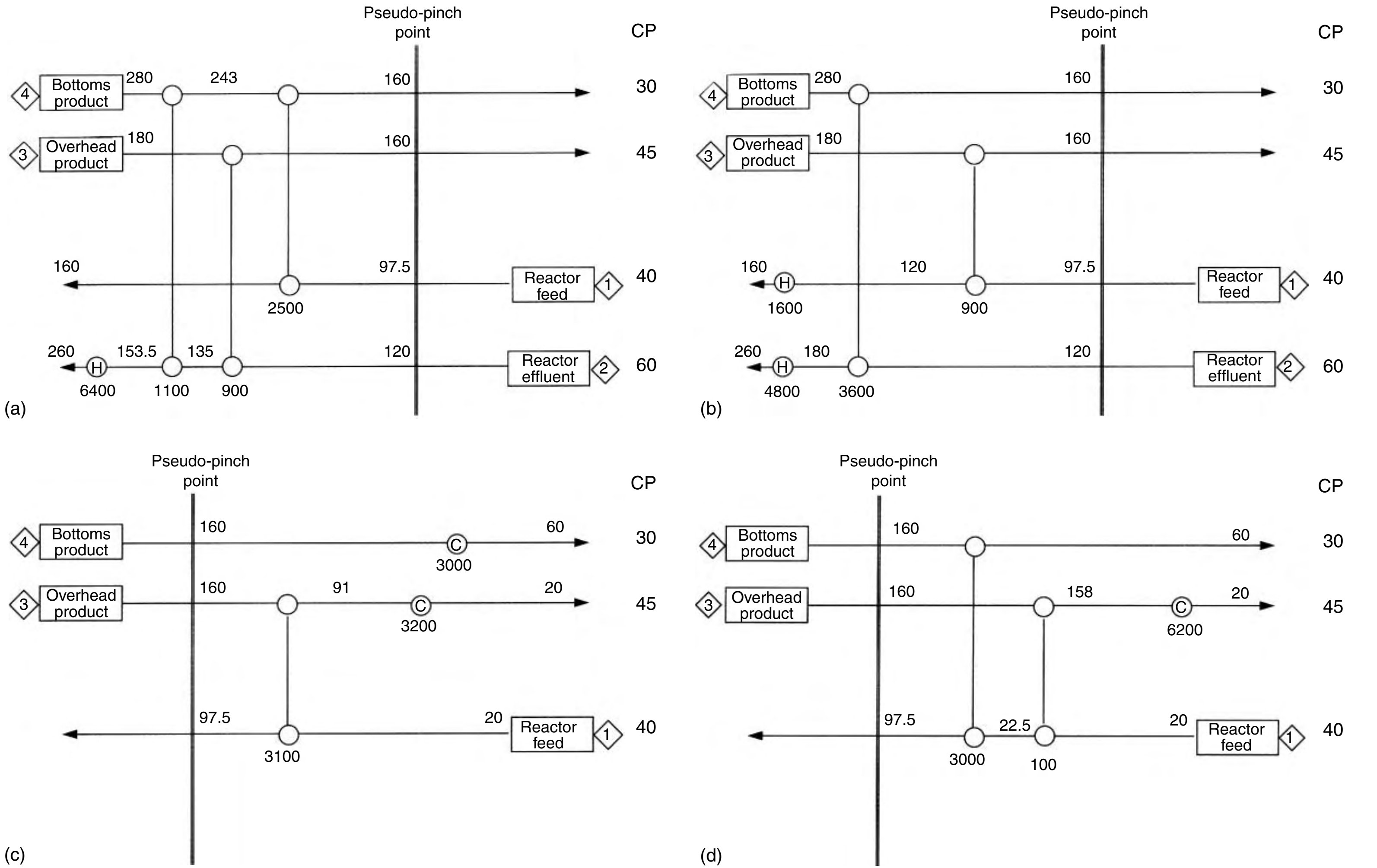


FIGURE 15.32 (a) Above pseudo-pinch design—1. (b) Above pseudo-pinch design—2. (c) Below pseudo-pinch design—1. (d) Below pseudo-pinch design—2.

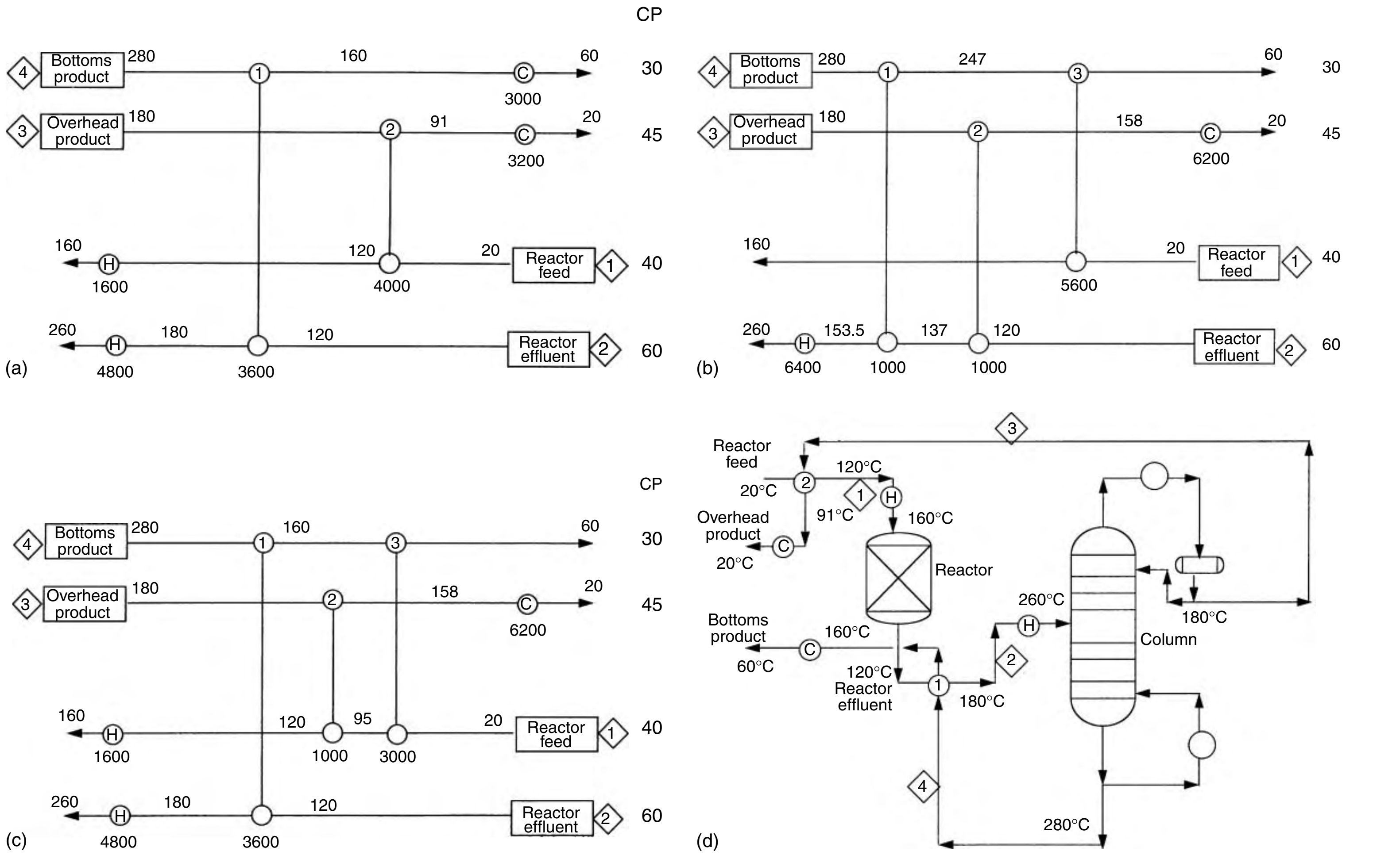


FIGURE 15.33 (a) Pseudo-pinch design—2. Total annualized cost is $743,400. (b) Pseudo-pinch design—3. Total annualized cost is $727,600. (c) Pseudo-pinch design—4. Total annualized cost is $740,100. (d) Final flow diagram for the plant in Figure 15.9.

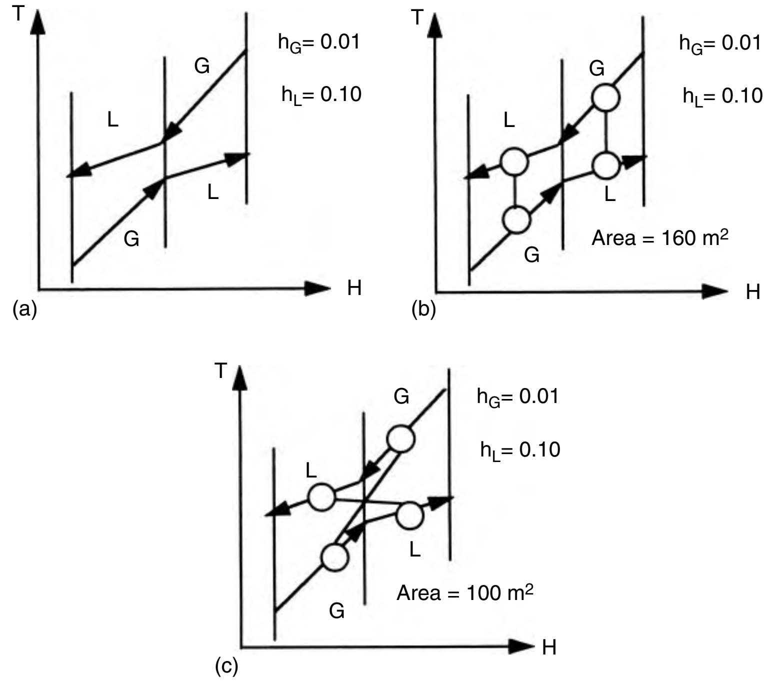


FIGURE 15.34 Criss-cross heat transfer will reduce the network area when heat transfer coefficients are very different. (b) Vertical heat transfer. (c) Criss-cross heat transfer.

of the network. Thus each stream match will have different coefficients in the cost equation. To optimize the network one will have to undertake criss-crossing on the composite curves.

Stream matching constraints may also require criss-crossing on the composite curve. For example, a match between two streams may be prohibited to maintain process safety or to avoid incurring excessive piping costs associated with a particular match. When constraints are imposed, extra utilities may be required. As a result, one may have to use cold utilities above the pinch point, use hot utilities below the pinch point, or transfer heat across the pinch point. This is totally in contrast to the pinch design rules.

All these scenarios force the designer to criss-cross on the composite curves. Various design methods are proposed in the literature to obtain optimal designs using the criss-cross mode of heat transfer in the network and is not with the scope of this work.

15.1.10 Selection of Utility Loads and Levels

The annual operating cost depends on the amount and type of utilities used. In a complex process such as an ethylene plant, there would be about four or five steam levels and about seven or eight refrigeration levels. High-pressure steam generated within the process is let down to other steam levels via steam turbines that generate power. Steam from these levels is used for heating the process. The power generated from the turbines is used by the compressors and pumps and to generate various levels of refrigeration needed. The question faced is what pressure levels of steam to use and what the load is on each level. Also, what are the best temperature levels of refrigeration and what will be their respective loads?

Pinch technology helps us answer these questions in a very simple manner using the Grand Composite Curves (GRCC).[28] The GRCC is the curve that shows the heat demand and supply within each temperature interval. This curve is derived from the problem table (refer to Table 15.5). In the problem table, we had modified the stream starting and target temperatures depending on the value

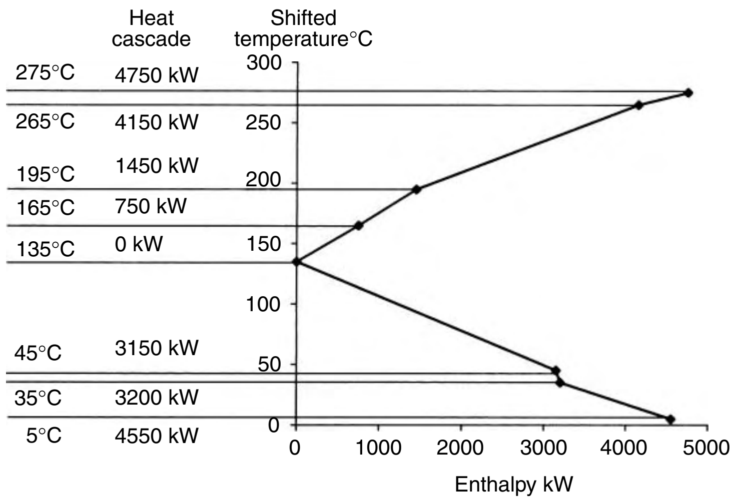


FIGURE 15.35 Grand composite curve for the flowsheet of Figure 15.9.

of Δt_{min}. From now on we shall refer these modified temperatures as shifted temperature. The heat flows between two adjacent shifted temperature intervals can be plotted on the shifted temperature–enthalpy plot. Figure 15.35 shows the heat cascade and how the grand composite curve is developed for the flowsheet shown in Figure 15.6. The grand composite curve gives us a graphical representation of heat flows taking place in the system. At the pinch point the heat flow is zero.

The grand composite curve is piecewise linear. The slope of this curve also changes from interval to interval. A line with a positive slope indicates that the system in that region needs external heat. A line with a negative slope indicate that there is surplus heat available within that temperature interval that can be cascade down within the system and used at a lower temperature interval. It is very clear from the grand composite curve that the above pinch region is a heat sink and the below pinch region is a heat source.

To further explain the importance of the grand composite curve and the kind of information that can be extracted from it, consider the curve shown in Figure 15.36. Consider the section AB that is between

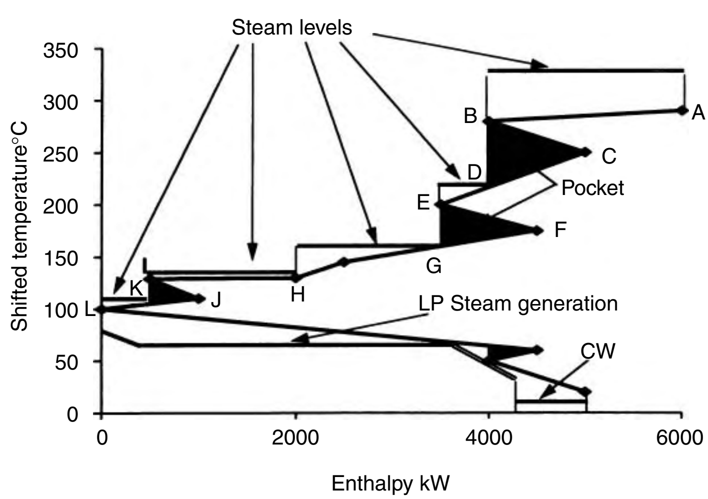


FIGURE 15.36 Selection of utility loads and levels.

the shifted temperature interval of 295°C and 280°C. This section demands external heat of 2,000 kW. Hence we should place a hot utility such as high-pressure steam or hot oil to supply it.

Next, consider section BC of the GRCC between the shifted temperature intervals of 280°C and 250°C. This section has a heat surplus, so we can use it else where in the system. We can drop a vertical line from *B* to meet the GRCC at point *D*. Section CD of the GRCC, which is a heat deficit region of the system, can now be satisfied by the heat surplus from the BC section. This section of the GRCC is called a pocket, and the process is self-sufficient with respect to energy in this region. The section DE now needs external heat. This heat can be supplied at any temperature ranging from the highest available temperature to a minimum temperature corresponding to point *D*.

Following the same logic, EFG will be a pocket. In section GHI one hot utility level can be used at a shifted temperature level of point *G* with a total duty of 3,000 kW, or two levels can be used—one at 160°C with a duty of 1,500 kW, and other at the shifted temperature level of point *H* (i.e., 140°C with a duty of 1,500 kW). The choice is dictated by the trade-off between the power requirement, capital investment, and complexity of the design. Using only one level will make the design of the utility system simpler and the capital cost of the heat exchanger smaller due to higher-temperature approaches. On the other hand, if there is demand for power, then using two levels will produce more power.

A similar economic trade-off will be required for supplying the external heating requirement in section KL of the composite curve. Point L is the pinch point. Below point L heat needs to be rejected into a cooling utility such as an air cooler or cooling water. Also, in the below pinch section of the process we can address the question, Is it possible to raise steam at some temperature? If so, how much?

For example, for the process grand composite curve shown in Figure 15.36, we want to find out how much low-pressure superheated steam can be generated. The saturation temperature of the low-pressure steam is 70°C and boiler feed water is available at 30°C. The superheat is 10°C. Using a simple trial-and-error procedure we can find out how much steam will be generated. Assume the amount of steam that is generated. Develop a heat curve for the low-pressure steam generation on the shifted temperature scale. As generation of low-pressure steam will be a cold stream, the temperature of the stream will be increased by $\Delta t_{\min/2}$. Keep on increasing the amount of steam generated till the steam generation heat curve touches the process grand composite curve at any point. The selection of refrigeration levels uses the same technique.

Once the utility levels are decided, introduce them into the stream data and obtain the balanced composite curves. The number of pinch points will increase. In addition to the original process pinch point, each utility level will introduce at least one pinch point. A balanced grid diagram that includes all the utility streams and all the pinch points identified on the balanced composite curve can now be used along with the network design algorithms to develop a network that achieves the target set. Advances have made the task of selecting and optimizing multiple utilities easier.[29,30]

15.1.11 Process Integration

Until now we have discussed the design and optimization of heat exchanger networks. But heat exchanger networks are a part of a whole process. Processes consist of reactors, distillation columns, utilities, and so on. For the process to be optimally designed, all the unit operations should be properly integrated. Generally each unit operation is individually optimized. It is a misconception that if each individual unit is individually optimized then the resulting process is also optimized. Each unit operation interacts with the other in the process. Hence, for the process to be optimized, each unit operation should be properly integrated.

Figure 15.37 shows the onion diagram proposed by Linnhoff et al.[18] that represents the hierarchy of process design. Generally, the reactors are designed first that fix the heat and material balance for the whole process. The products, by-products, and reactants are then separated using the most common unit operation—distillation. A number of columns may be used if a large number of components need to be separated. To achieve the separation and to carry out the reaction in the reactor, the heat exchanger

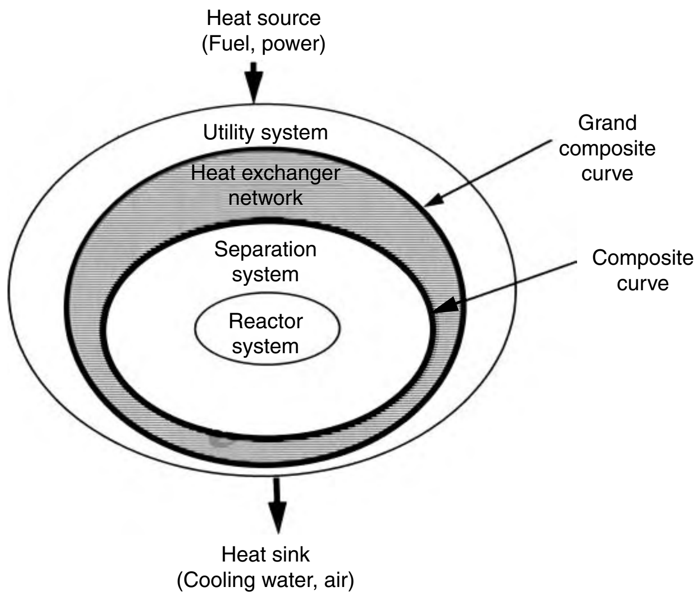


FIGURE 15.37 Onion diagram.

network, using heat exchangers, heats and cools the process streams. External heating and cooling is supplied by the utility system. The utility system may consist of a simple boiler system or a cogeneration unit that produces the necessary steam for the process and the power requirement. All these layers interact. An optimal process design must take into account these interactions and must be properly integrated.

Process integration uses principles that we have already established. Process integration starts by optimizing stand-alone distillation columns with respect to reflux ratio, feed conditioning, side-reboilers, and side-condensers.[31] To properly integrate the distillation column with the background process, it should be placed above or below but not across the background process pinch point.[32] The same principle applies for the appropriate integration of heat engines. Heat pumps should be placed across the pinch point.[28] Linnhoff and Leur[33] have developed procedures for integrating furnaces with the rest of the process.

15.1.12 Process Modification

In the above pinch region, hot streams should be modified such that they can transfer more heat to the cold streams. The cold streams should be modified such that they require less heat. Both of these modifications will decrease the amount of external hot utility requirement. Similarly, in the below pinch region, the cold streams should be modified such that they require more heat and the hot streams should be modified such that they transfer less heat. These modifications will reduce the amount of cold utility requirement. This principle is called the plus/minus principle[3] and is illustrated in Figure 15.38.

15.1.13 Shaftwork Targets

Recently, Linnhoff and Dhole[34] have proposed shaftwork targets based on exergy analysis of the stream data for a plant. The shaftwork targeting procedure calculates the change in the total shaftwork requirement of the system due to any changes in the base case. They use exergy composite and grand composite curves that are obtained from the composite and grand composite curves by changing the temperature axis to the carnot factor. Shaftwork targets are very important in the design of low-temperature processes.

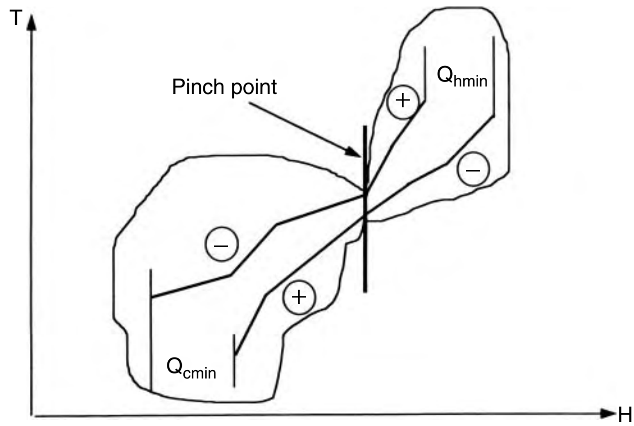


FIGURE 15.38 The plus/minus principle.

15.1.14 Sitewide Integration

Generally, individual process are located in a site. The site consist of multiple units. For example a refinery may consist of a crude unit, vacuum unit, naphtha hydrotreater, diesel hydrotreater, fluidized catalytic cracking (FCC) unit, visbreaker unit coker unit, and so on. A petrochemical complex may consist of an ethylene unit, polyethylene unit, and so on. All these units will have one utility system that provides hot and cold utilities as well as power to the whole site. Each unit will have steam demands at different temperatures. At the same time, each unit may be producing steam at different levels. It is possible to directly integrate different units, but that may cause political problems or piping problems. Hence, different units are generally integrated indirectly through the utility system. The problem that sitewide integration addresses is what is the correct level of steam for a site and what is the trade-off between heat and power. Should power be imported or cogenerated?[35]

The most important benefit of the sitewide analysis is that correct pricing for different levels of steam are no longer required. The impact of any modification in terms of fuel or power can be easily evaluated. There is no need for "cost of steam," as energy pricing is only done with respect to either fuel or power at the battery limit.

Sitewide analysis has increased the understanding of global emissions associated with any processing industry. To minimize the emissions associated with a process, sharper separations are required. Since distillation is the workhorse of the chemical industry, to obtain sharp separation, the reboiler and condenser duties will increase or additional processes may be added that require additional heating and cooling. External energy is obtained by burning fuel, and when fuel is burned emissions are generated. If extra power is used, then the emissions in the utility company will increase. Thus, to decrease the emissions in the process we might end up increasing the emissions associated with burning the fuel. The net effect might be that we increase the global emissions.[1] Recently, Smith and Delaby[36] have established techniques to target for CO_2 emissions associated with energy. Combining these CO_2 targeting techniques with sitewide analysis will help in trading off different emissions.[37]

15.1.15 Data Extraction

Process integration studies start from a base-case flowsheet. This flowsheet may be existing or may be developed from designer's experience. To conduct pinch analysis properly, it is important to extract the flow rate, temperature, and heat duty data correctly.

Stream target and starting temperatures should be chosen so that we do not generate the original flowsheet.[18] To illustrate this, consider the flowsheet shown in Figure 15.39. If we extract the data as two streams, then we might end up with the original flowsheet. If the drum temperature is not important

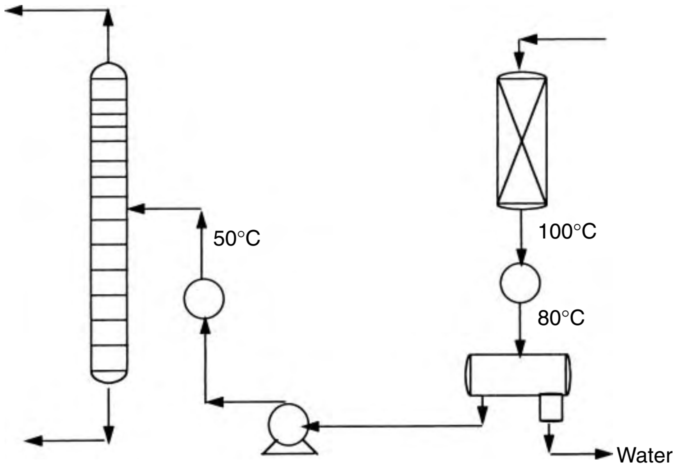


FIGURE 15.39 Flowsheet for data extraction example.

then we can consider it to be one stream and we stand a chance for finding new matches. The drum and the pump can then be kept as a natural break point in the system.

While extracting data, extra care should be taken when streams are mixing nonisothermally.[18] Consider the system shown in Figure 15.40. Stream A is being cooled to the mixed temperature, and stream B is being heated to the mixed temperature. This happens due to mixing the streams. The mixed stream is then heated to a higher temperature. If we extract the data as shown in Figure 15.40a and if the pinch temperature is 70°C then we are inherently transferring heat across the pinch point by mixing the two streams. In the process of mixing we are cooling stream A and then subsequently heating it up again. The correct way to extract the data is shown in Figure 15.40b.

When a stream is split and the split streams have two different target temperatures then each stream is considered as two separate streams. However during the design phase we can use the stream splitting and mixing technique to eliminate one unit.[18,38] See Figure 15.41.

15.1.16 Procedure for Optimization of an Existing Design

Process design generally starts with an existing flowsheet. The flowsheet is modified to meet the new requirements defined by the project. Figure 15.42 shows the procedure proposed by Trivedi et al.[39] that

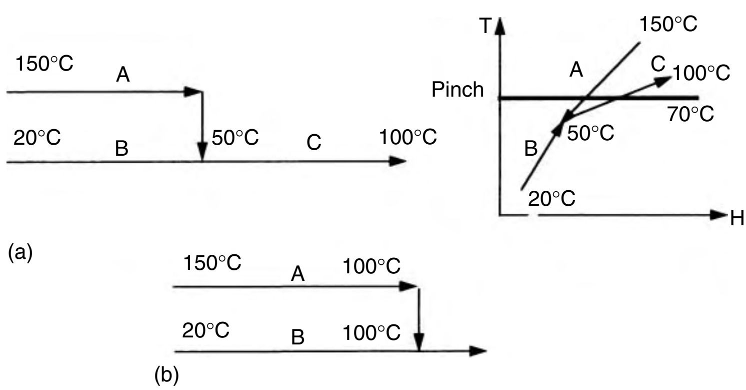


FIGURE 15.40 Data extraction for streams that are mixing.

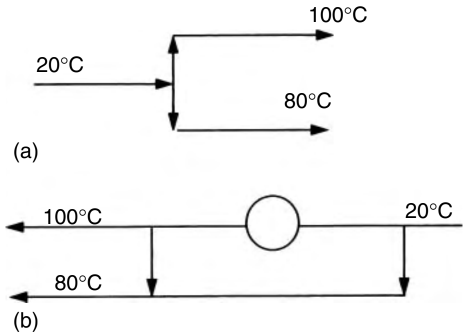


FIGURE 15.41 Stream bypassing and mixing can save a unit.

utilizes all the concepts of pinch technology for optimizing an existing design. The following sections discuss various steps of this procedure.

15.1.16.1 Problem Definition

From the existing process flow diagram and the project requirement develop a problem definition and scope of the project. At this early stage in the design process, identify all the key project constraints. These constraints may be utility temperature levels (if the construction of the plant is going to be in an existing site), environmental constraints, budget constraints, schedule constraints, and so on.

15.1.16.2 Conceptual Flow Design

On the basis of the problem definition and project constraints, develop a new conceptual flow diagram. This flow diagram need not be detailed. However, it should at least contain the reactors, the distillation columns, and the associated conceptual utility system. Furthermore, it should also contain key operating parameters such as feed conditions to distillation columns, operating temperature, and pressure of distillation columns. This flow diagram becomes the basis for optimization.

15.1.16.3 Marginal Cost of Utilities

Develop a simulation model of the conceptual utility system to evaluate the marginal cost of various utility levels. With this model it is possible to evaluate the interactions between the various utility levels. Use a reference utility level as the basis for calculating the savings associated with the other levels. The reference utility level may be the fuel fired to generate steam or to export/import steam at a certain level.

15.1.16.4 Simulate and Optimize Distillation Columns

In this step, simulate the distillation columns using the conditions set by the problem definition and the conceptual flow diagram. Optimize these columns on a stand-alone basis using the procedures outlined by Dhole and Linnhoff.[31]

15.1.16.5 Stream Data Extraction

The conceptual flow diagram along with the simulation of the distillation columns establish the heat and material balance for the entire process. Simulation will also generate the heat curves for all the hot and cold streams. Extract stream data from the process flow diagram using the guidelines established earlier.

15.1.16.6 Targeting

Assume a value of $\Delta t_{\min}$. Set overall energy targets using the composite curves. Divide the problem into above ambient temperature and below ambient temperature subproblems. The temperature of cooling water will generally decide this division point. Construct the grand composite curve (GCC) for the above

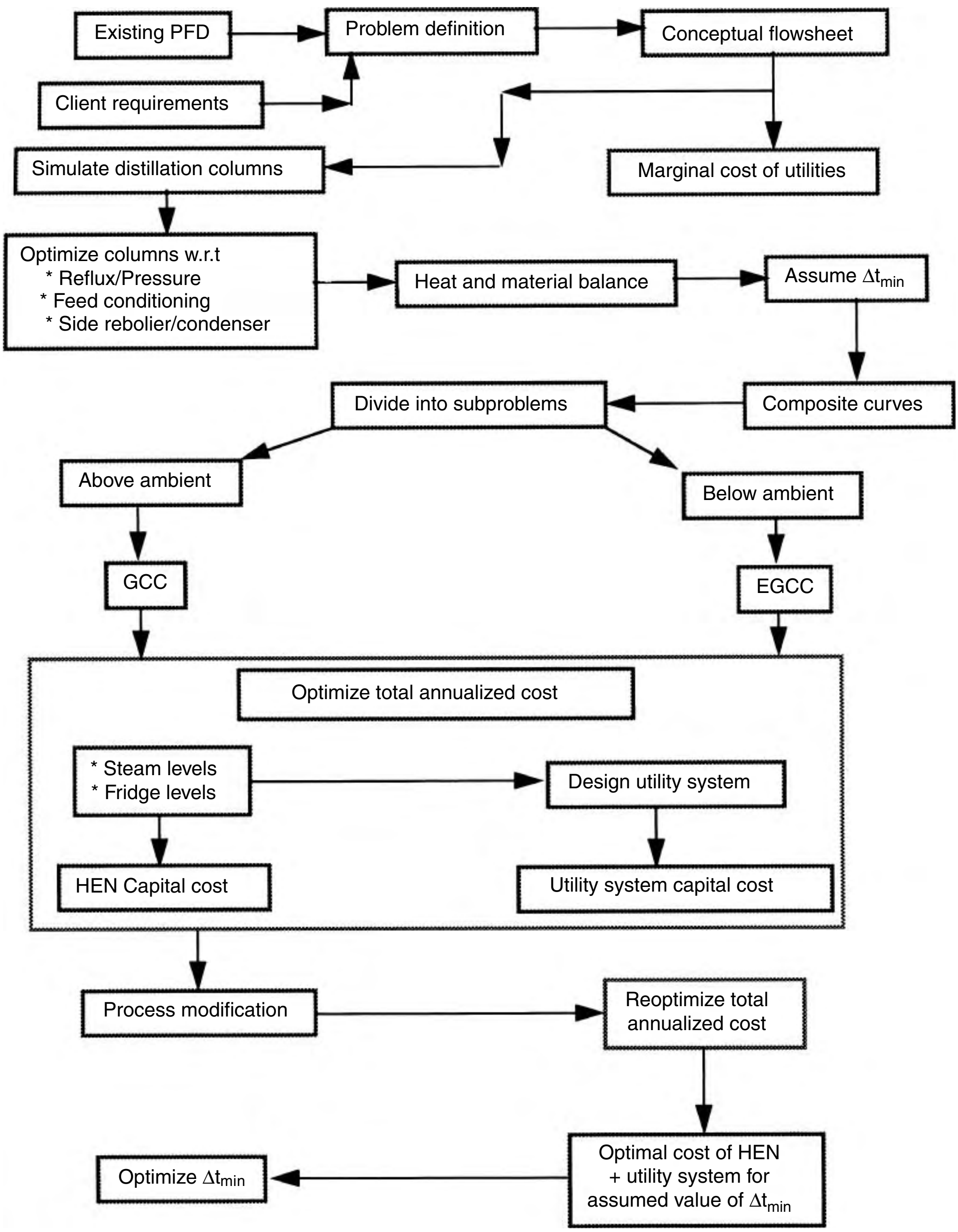


FIGURE 15.42 Proposed new procedure.

ambient temperature subproblem. For the below ambient temperature subproblem, construct the process exergy grand composite curve. Fine-tune the column optimization undertaken previously.

15.1.16.7 Optimize Total Annualized Cost

From the grand composite curve, optimize the steam levels and their corresponding loads. From these utility loads and levels, design the utility system and estimate its capital cost. Now construct the composite curves that include the utilities. These set of composite curves are the balanced composite curves. Estimate the capital cost of the heat exchanger network from these curves. Calculate the utility cost from the utility loads and levels. Hence, estimate the total annualized cost for the heat exchanger network and the utility system. Optimize this cost by changing the utility loads and levels.

15.1.16.8 Process Modification

Use the balanced composite curves to identify process modifications that can potentially reduce the total annualized cost. Changing the operating pressure of a distillation column or modifying the flow rate of a stream are some of the potential process modifications that can be made. The total annualized cost needs to be reoptimized if the utility system design changes significantly.

15.1.16.9 Optimize the Value of $\Delta t_{\min}$

At this stage we have the optimal total annualized cost targets for the heat exchanger network and the utility system for the assumed value of $\Delta t_{\min}$. Assuming another value of $\Delta t_{\min}$ will give another optimized total annualized capital cost. Repeat the exercise over a range of $\Delta t_{\min}$ values to find the optimal value. Design the heat exchanger network at the optimal value of $\Delta t_{\min}$.

Trivedi et al.[39] have reported a case study that discusses the application of the above outlined approach to optimize Brown & Root's state-of-the-art Ethylene process. They achieved 12% savings in energy consumption. The capital cost increased by 2% with a simple payback period of 2 years. (Typical specific energy consumption of an ethylene plant for cracking ethane through naphtha ranges from 3,000 to 5,000 Kcal/kg. The total world production capacity of ethylene in 1989 was 61 million MTY[40]).

15.1.17 Recent Developments

Recent developments have been made to solve different process design problems using pinch technology principles and concepts. We shall discuss these developments in brief.

Normally plants have many variables that change during the course of time. For example, depending on market conditions the through put may change or the feed stock can change. Process conditions can change from summer to winter operations. Catalytic reactors lose their effectiveness and heat exchangers get fouled over time. The final process design should be able to handle such changes. Sensitivity analysis and multiple base-case design procedures are developed that address the issue of flexibility.[8,41]

The pressure drop in a heat exchanger network is also important. Polley et al.[42] have developed methods that account for stream pressure drops during the targeting and detail design phase. Stream pressure drop constraints are very important during revamp projects. Pinch technology concepts are also applied to retrofit projects, heat integrate batch processes, minimize waste water, integrate evaporator systems, and design power cycles.[1]

15.2 Pinch Technology in Practice

Ed Fouche and Kelly E. Parmenter

15.2.1 Introduction

Pinch technology is a practical technique for maximizing energy efficiency, reducing costs, and improving environmental performance across multiple processes in an industrial setting. Its implementation can lead to higher profitability and increased competitiveness for the plant.

In industrial processes, some of the energy that goes into a process must be released at a lower temperature. Thus, the process "consumes" higher-temperature (and therefore higher-quality) energy. In principle, a low-temperature process can potentially use the waste heat from other processes that operate at higher temperatures. Pinch analysis was pioneered by Linnhoff to identify opportunities for utilizing heat across different industrial processes. It was first applied to industrial energy systems in response to

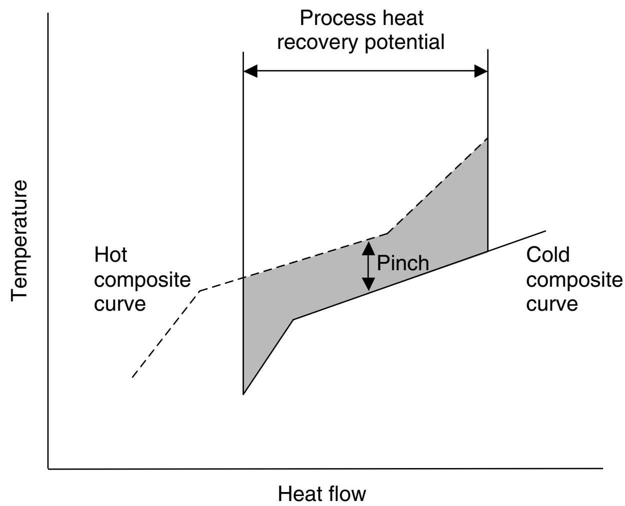


FIGURE 15.43 Composite curves used in pinch analysis.

the energy crisis of the late 1970s. The theory behind pinch technology is discussed in great detail in Section 15.1.

In short, to perform a pinch analysis, engineers integrate the temperature characteristics for all of the streams that need to be heated (cold streams) into a single cold composite curve and then combine this curve with an analogous hot composite curve that represents all of the streams that need to be cooled (hot streams). The result is a diagram showing two temperature vs. heat flow plots—or composite curves—as shown in Figure 15.43. The overlap region between the two curves represents the potential for plant-wide heat recovery. The curves also reveal the minimum net heating and cooling requirements, or energy targets. The minimum vertical distance between the hot and cold composite curves is called the *process pinch*. The pinch represents the bottleneck for heat transfer. This leads to three important principles for efficient energy use:

1. Heat should not be transferred across the pinch.
2. There should be no external cooling above the pinch.
3. There should be no external heating below the pinch.

Therefore, the identification of the bottleneck, or pinch, enables heat transfer to be optimized. The benefit of this technique is that it illustrates the big picture and provides a comprehensive overview. Prior to development of pinch technology, the energy requirements of a unit operation in a large processing plant were evaluated without consideration of what energy sources or sinks might be available elsewhere in the plant. Pinch analysis provides this overview and forces discipline; in many cases, the installation of a boiler or chiller has been avoided by taking advantage of temperature surpluses and demands elsewhere in the plant.

Key features of pinch analysis are that:

- The minimum heating and cooling needed to operate the process are explicitly quantified.
- The maximum possible process heat recovery is defined.
- Unnecessary capital costs are avoided.
- Possible uses for low-grade waste heat are identified.

The remaining sections of this article discuss the application of pinch technology in a variety of real-life settings. First, an overview of the general methods employed by pinch experts is provided; this is then followed by summaries of several specific case studies.

15.2.2 Pinch Technology in Practice

Pinch technology can be employed to identify practical and cost-effective ways to improve substantially the energy efficiency of a processing facility. It can be used to optimize the energy use in individual processes or across a complete site. In general, the objectives of a pinch analysis are to quantify the minimum heating and cooling required for process operation, to pinpoint the maximum heat recovery potential, and to develop technically and economically viable projects to achieve energy saving targets.

The approach typically taken in a pinch study begins with the formation of a team consisting of pinch experts, key plant personnel, and possibly utility representatives. The team then visits the site to gather data pertaining to the processes, as well as economic and utility information. With data in hand, the pinch experts characterize the heating and cooling requirements of the plant's processing units using pinch technology. They next quantify the scope for potential improvement in each unit and the overall site. Then, the team identifies specific projects based on process change, process heat recovery, etc. that can be implemented to achieve target energy savings. The target energy savings is the summation of savings from practical, cost-effective projects identified to move the energy use at the site as close as possible to the minimum total process heating and cooling energy defined by the pinch analysis.

Since its inception, a variety of experts have implemented pinch technology with a high degree of success to improve energy utilization in processes. Early application focused on designing optimum heat-exchanger networks; recent efforts have extended the analysis to include boilers, turbines, heat pumps, and refrigeration systems, and techniques have been developed for designing effective combined heat and power systems.

The Electric Power Research Institute (EPRI) and its utility members in the U.S. and overseas have sponsored a variety of energy-related projects in the process industries through EPRI's Industrial Technology Application Service. The Industrial Technology Application Service is now owned and operated by Global Energy Partners. Pinch studies have historically been a key element of EPRI's process industries programs, particularly in the Pulp, Paper, and Forest Products program (see Section 15.2.3) and the Chemicals, Pharmaceuticals, Petroleum Refining, and Natural Gas program (see Section 15.2.4). Electric Power Research Institute has also carried out numerous studies for the food and beverage industry and several studies for the textile and fiber industry. In all, EPRI has conducted more than fifty pinch studies in the last 15 years that, on average, have identified energy cost savings of more than 20 percent with one-to-three year payback periods.

15.2.3 Pinch Technology in Pulp and Paper Industries

Pulping and papermaking are energy-intensive. Rising energy costs have triggered a new awareness of energy usage and have spawned a number of energy efficiency studies across the industry. The U.S. Department of Energy cofunds several programs for improving energy efficiency. Its short-term initiatives include the popular plant-wide energy assessments where energy audits identify cost-reduction opportunities. Longer term, its Agenda 2020 program focuses on developing major technology jumps meant to reduce the energy costs of papermaking and to provide environmental benefits. Energy and environmental issues are intertwined; lower energy use reduces emissions from both the mill and from the power plant that generates the energy.

EPRI and associates have conducted over 30 pinch studies for the pulp, paper, and forest products industry alone. Collectively, the studies have pinpointed potential energy savings of over \$80 million/year, with many projects having a simple payback period of about one year. Table 15.7 summarizes results from some of these studies. The table shows the target, actual, and scope energy values, as well as practical energy and cost savings by mill type. Target values represent the minimum energy required to run a process; actual values represent the current, actual energy use; scope values represent the maximum possible heat recovery; and practical values represent what is realistically feasible

TABLE 15.7 Summary of Results from Pinch Studies in the Pulp, Paper, and Forest Products Industry

Mill Type	Target 10^6 Btu/h (MW)	Actual 10^6 Btu/h (MW)	Scope 10^6 Btu/h (MW)	Practical Savings	
				10^6 Btu/h (MW)	$\$10^6$/year
Semisulfite/OCC linerboard	297 (87)	417 (122)	120 (35)	45–80 (13–23)	1.4–2.8
Kraft/NSSC/OCC linerboard	939 (275)	1250 (366)	311 (91)	177 (52)	1.37
Continuous bleached kraft	NA	NA	177 (52)	127 (37)	3.2
Bleached kraft	428 (125)	553 (162)	125 (37)	85 (25)	3.2
Bleached kraft	732 (215)	866 (254)	134 (39)	80 (23)	1
Newsprint kraft	516 (151)	921 (270)	405 (119)	160 (47)	1.6
Bleached kraft	292 (86)	442 (130)	150 (44)	100 (29)	3
Continuous bleached kraft HWD	NA	NA	140 (41)	118 (35)	2.5
Bleached kraft/TMP/groundwood	792 (232)	1154 (338)	362 (106)	300 (88)	4.0
Bleached kraft	826 (242)	1096 (321)	270 (79)	160 (47)	2.0
Bleached kraft	1265 (371)	1904 (558)	639 (187)	235 (69)	2.8
Bleached semisulfite	235 (69)	314 (92)	79 (23)	37 (11)	1.4
Bleached kraft/TMP	625 (183)	965 (283)	340 (100)	294 (86)	2.1
Bleached kraft & NSSC	499 (146)	688 (202)	189 (55)	100 (29)	2.1
TMP/recycle mill greenfield design	Capital cost avoided ~ $5 million				
Bleached kraft	531 (156)	693 (203)	162 (47)	79 (23)	2
TMP	NA	NA	137 (40)	70 (21)	1.5
Bleached kraft	1212 (355)	1387 (407)	175 (51)	100 (29)	1.8
TMP/Deink Mill	19 (6)	156 (46)	137 (40)	67 (20)	1.3
Bleached kraft/TMP/groundwood	Multiple case scenario			350 (103)	5.8
Bleached kraft	542 (159)	864 (253)	322 (94)	175 (51)	2.3
Bleached kraft	NA	NA		80 (23)	3.5
Continuous bleached kraft northern	NA	NA	190 (56)	110 (32)	3.1
Groundwood/coaters	255 (75)	360 (106)	105 (31)	53 (16)	0.7
Kraft/NSSC/OCC linerboard	933 (273)	1178 (345)	245 (72)	225 (66)	4.8

TABLE 15.8 Practical Steam Cost Savings Identified by Pinch Technology in the Pulp, Paper, and Forest Products Industry

Mill Type	$/BDST ($/metric ton)
Bleached kraft/NSSC	2.9 (3.2)
Bleached kraft/TMP & other	5.0 (5.5)
Bleached market pulp	5.5 (6.1)
Nonintegrated papermaking	1.3 (1.4)
Kraft/NSSC/OCC	2.9 (3.2)
Sulfite or semisulfite	4.4 (4.9)

considering constraints. As the table shows, practical energy savings per mill range from about 50 to 350 million Btu/h [~15 to 100 MW], and practical cost savings range from about $1 million/year to $6 million/year.

Table 15.8 lists the average cost savings per ton of product for several types of mills. Savings range from $1.3 per bone dry short ton (BDST) ($1.4/metric ton) for nonintegrated papermaking to $5.5/BDST ($6.1/metric ton) for bleached market pulp.

Two specific examples of pinch technology being applied to the pulp, paper, and forest products industry are described in the following case studies.

15.2.3.1 Case Study 1: Minimizing Process Energy Use in a Large Thermomechanical Pulp (TMP) Mill[a]

EPRI and American Process, Inc. conducted an energy targeting scoping study using pinch analysis at a large TMP mill in Canada. A local utility company cofunded the project. The mill produces bleached, unbleached, and semibleached kraft market pulp and standard and offset newsprint. Fiber is prepared in kraft continuous and batch digesters, groundwood, and TMP departments. The kraft digesters separate wood chips into fiber by use of chemicals that act on the wood's binding agent (lignin). In the groundwood process, wood chips are mechanically pressed against a grindstone and, with the addition of water, are separated into fibers. The TMP process uses both heat and mechanical forces to convert wood chips into fibers.

Spent chemicals and lignin, which are referred to as *black liquor*, are recovered from the kraft process in a recovery cycle that includes three evaporators and three recovery boilers, a recausticizing line, and two kilns. First, the black liquor is removed from the pulp in washers; next, it is thickened in the evaporators; then, it is burned in the recovery boilers. During the burning process, the chemicals and lignin are recovered and the steam raised in the boilers is used to support operations. Two lime kilns provide calcium oxide for causticizing. Chlorine dioxide is produced onsite and used in a six-stage bleach plant.

The mill also has power boilers that use oil and hog fuels. The steam from the power boilers is used in steam turbines to generate a portion of the mill's power requirement, as well as to support various operations. The balance of the mill's power is purchased.

The intent of the energy targeting scoping study using pinch analysis was to identify ways for the mill to reduce energy costs as part of a long-term improvement in competitiveness. In particular, operators of the mill wanted to reduce process steam consumption so that they could decrease steam production from the power boilers.

Pinch analysis yielded the hot and cold composite curves shown in Figure 15.44 and the grand composite curve shown in Figure 15.45. The figures show that the pinch interval temperature is 246°F (119°C) and the thermal energy target (minimum thermal energy required to run the extracted process

[a]Data from EPRI, Minimizing Process Energy Use for a Large TMP Mill with Pinch Technology, EPRI, Palo Alto, CA, 2000.

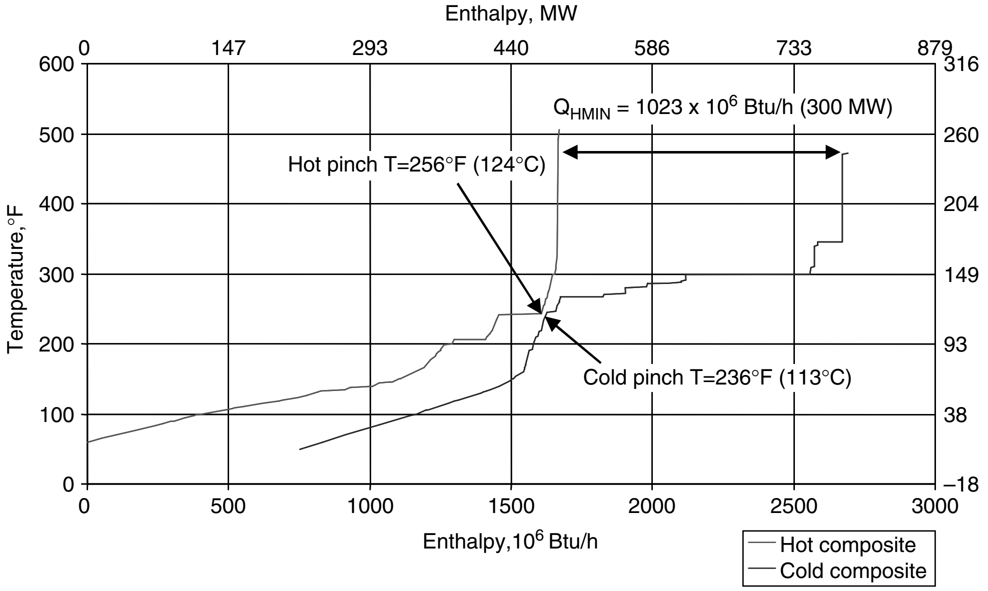


FIGURE 15.44 Hot and cold composite curves for thermomechanical pulping (TMP) mill.

steam) is 1023 million Btu/h (300 MW). Because the existing process thermal consumption is 1336 million Btu/h (392 MW), the maximum potential process steam savings (scope) is about 313 million Btu/h (∼310,000 lb/h, 92 MW, 140,000 kg/h).

To accomplish the objective of reducing steam consumption, two options were developed for the recovery of TMP dirty steam. Option 1 uses TMP steam in the mill's pulp *and* paper companies, and option 2 restricts the use of TMP steam to the paper company. For option 1, implementation of TMP heat recovery in conjunction with several other recommended projects was found to potentially save

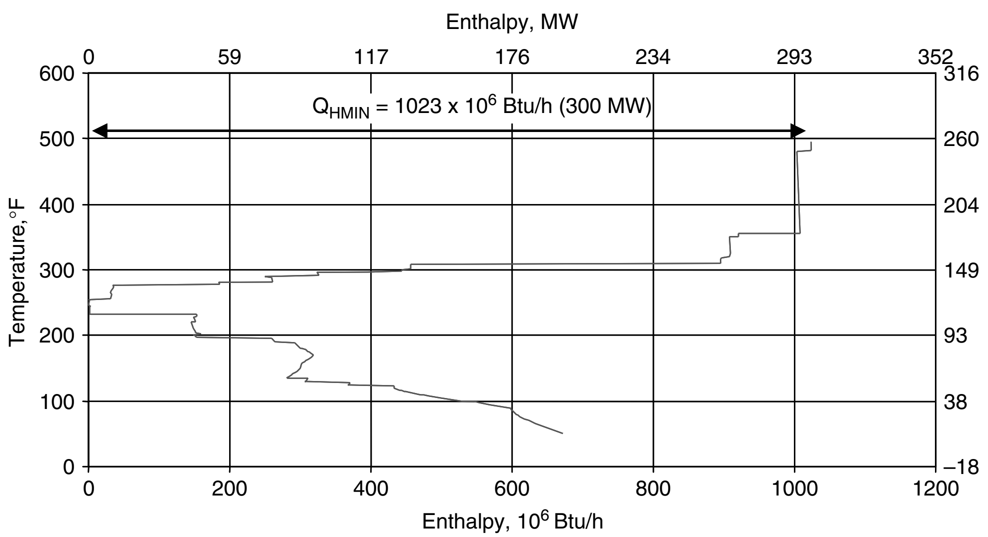


FIGURE 15.45 Grand composite curve for thermomechanical pulping (TMP) mill.

292,400 lb/h (132,600 kg/h) of process steam; for option 2, the potential savings were found to be 232,800 lb/h (105,600 kg/h). These steam savings equate to cost savings of about Can$9.5 million/year (Canadian dollars per year) for option 1 and Can$8.7 million/year for option 2 as a result of decreased steam production in the power boilers. The simple payback periods were estimated to be 2.5 years for option 1 and 2.1 years for option 2.

Additionally, a target cogeneration analysis showed that with the existing steam consumption, there was an opportunity to increase onsite power generation by about 21 MW (which equates to a savings of about Can$6.3 million/year) by correcting turbine inefficiency, pressure-reducing valve operation, and incorrect use of steam levels. After implementation of the proposed projects and the resulting new steam consumption, this additional power generation opportunity decreases to about 9.1 MW and Can$2.7 million/year.

Furthermore, a comparison of the steam cycle with a combined cogeneration cycle showed that a combined cycle would potentially generate 49 MW as opposed to only 8.6 MW in the steam cycle. This generation capacity equates to a monetary savings of Can$5.3 million/year.

15.2.3.2 Case Study 2: Minimizing Process Energy Use in a Newsprint Mill[b]

With support from a local utility company, EPRI and American Process, Inc. carried out an energy-targeting scoping study using pinch analysis at a newsprint mill in Mississippi. The mill generates 780 tons/d (707 metric tons/d) of newsprint from a combination of TMP and purchased kraft pulp. Steam is produced in a reboiler using TMP-generated steam and in a bark boiler using mill-generated and purchased hog fuel. To take advantage of off-peak power rates, the mill swings the refiner operation.

Prior to the pinch analysis, the mill had already completed a cost assessment study and established several possible energy savings projects. The objective of the pinch analysis was to augment previous efforts and identify additional possible projects, as well as to prioritize all projects and ensure that the savings were additive.

For the steam duties evaluated, the hot and cold composite curves and the grand composite curve yielded a pinch interval temperature of 244°F (118°C) for the minimum temperature driving force of the $\Delta T_{\min} = 20$°F (11°C), and a thermal energy target (minimum thermal energy required to run the extracted process streams) of 42 million Btu/h (12 MW). Because the existing process thermal consumption of the streams is 96 million Btu/h (28 MW), the scope (or maximum potential to reduce process heat demand) is 54 million Btu/h (16 MW).

The significance of the pinch results is that, ideally:

- No utility steam should be used for heating cold streams below 224°F (107°C).
- No cold water should be used for cooling streams above 264°F (129°C) and no heat should be rejected to the environment above 264°F (129°C).
- No heat exchange should occur between above-the-pinch streams and below-the-pinch streams.

The instances where these rules are violated are referred to as *cross-pinch heat exchanges* or "XPs," and are the reasons why the actual steam consumption by the mill is larger than the minimum or "target" quantity. These XP heat exchanges need to be corrected for the mill to get closer to its target energy consumption.

The existing configuration was placed on a heat recovery grid to identify where the 54 million Btu/h (16 MW) of XP heat transfer was being used inappropriately. Table 15.9 lists the XP occurrences.

The team developed a variety of projects for the mill to consider to achieve energy savings. For the energy savings to correspond to realistic monetary savings, the energy savings by steam reduction must

[b]Data from EPRI, Energy Targeting Scoping Study Using Pinch Analysis: Bowater Newsprint, Grenada, MS, EPRI, Palo Alto, CA, 2004.

TABLE 15.9 Existing Cross-Pinch Heat Transfer Occurrences in a Newsprint Mill

Stream	Process-Process XP 10^6 Btu/h (MW)	Stream XP 10^6 Btu/h (MW)	Waste XP 10^6 Btu/h (MW)	Total XP 10^6 Btu/h (MW)
Bark boiler air heating		3.78 (1.11)		
Demin water heating		7.88 (2.31)		
Condensate return heating		2.90 (0.85)		
White water heating		3.70 (1.08)		
Off machine silo heating		5.64 (1.65)		
Warm water heating		1.30 (0.38)		
Paper machine PV air heating		14.12 (4.14)		
Paper machine building air heating		0.86 (0.25)		
Primary & secondary refiners steam			12.20 (3.58)	
Primary & secondary refiners steam condensate	1.36 (0.40)		0.63 (0.18)	
Subtotals	1.36 (0.40)	40.18 (11.78)	12.83 (3.76)	54.37 (15.93)

lead to a decrease in purchased hog fuel. Some of the projects initially developed would have had good payback periods if the steam were produced with fossil fuel; however, they yielded poor returns based on the cost of marginal steam from hog fuel. Six additional low-cost projects were proposed that would result in steam savings of 9000 lb/h (4080 kg/h) in the summer and 29,000 lb/h (13,150 kg/h) in the winter with payback periods of only 5–20 months. An additional project was recommended to temper incoming mill water with effluent during the winter months to reduce the seasonal variation in steam demand, yielding a potential savings of 15,000 lb/h (6800 kg/h) during the winter. Table 15.10 summarizes the energy costs savings, capital costs, and payback periods for all of the recommended projects. If implemented, the potential savings would be about \$273,100/year with a capital cost of \$462,100 and a simple payback period of 1.69 years.

15.2.4 Pinch Technology in Chemical and Petroleum Industries

The chemical and petroleum industries offer tremendous opportunities for energy savings from pinch analysis. EPRI and associates have carried out several dozen pinch studies in these industries since 1988. The calculated payback periods for most of the projects have been below two years and energy cost

TABLE 15.10 Summary of Costs and Benefits of Projects Developed for a Newsprint Mill

No.	Description	Savings (\$/year)	Capital Cost (\$)	Payback (years)
1	Scrubber-condensate storage tank vent	30,000	25,000	0.83
2	Sewer pressate purge after pressate coolers	25,000	10,300	0.41
3	Turpentine condenser—heated water generation	30,000	50,100	1.67
4	Optimize reject refinery vent heat recovery	39,900	55,500	1.39
5	Ejector for vent steam to chip bin	58,400	103,200	1.77
6	Series water flow through pressate coolers	15,000	10,700	0.71
7	Tempering mill water with process effluent	74,800	207,300	2.77
1–7	Totals	273,100	462,100	1.69

TABLE 15.11 Representative Results from Pinch Studies in the Chemical and Petroleum Industries

Type of Facility	Study Highlights
Petroleum refinery	\$840,000/year savings at 1.5-year payback, plus capital savings of \$390,000 for planned retrofit
Petroleum refinery	Scope for 32% reduction in net fuel consumption
Petroleum refinery	Revised heat exchanger network in crude distillation unit to reduce costs by \$1 million/year at less than 1.5-year payback
Petroleum refinery	Revised heat exchanger network in hydrocracker to save \$2.7 million/year in steam/fuel costs at 1.6-year payback
Petroleum refinery	Steam cost savings of \$5 million/year with 1–4-year payback
Pharmaceutical company	Increased heat recovery to reduce fuel costs by 13% at less than 1-year payback
Pharmaceutical company	Identified measures to reduce thermal energy requirements by 30%; annual savings of \$70,000/year at 1.3–2.5-year payback
Specialty chemical plant: kelp harvesting and extraction	Steam savings of \$2.8 million/year at 1–2-year payback
Wax extraction plant	Cost savings of \$930,000/year at 2.3-year payback

savings have typically been over \$1 million. Table 15.11 summarizes a few of the studies conducted. Two more detailed examples are provided in the following case studies.

15.2.4.1 Case Study 1: Minimizing Process Energy Use in a Petroleum Refinery[c]

With cofunding from local utility company, EPRI and Aspen Technology, Inc. conducted a sitewide pinch study at an oil refinery in Illinois. The objective of the study was to recommend ways to improve the refinery processes and utility system to achieve plant-wide energy savings. The project was conducted in two phases. The first phase consisted of a targeting and scoping study of the overall site. The second phase involved revising the basic design of selected process units.

The following processing units were included in the study:

- Atmospheric crude and vacuum distillation unit (VDU)
- Two delayed coker units
- Reformer and naphtha hydrotreater
- CCR reformer and naphtha hydrotreater
- Isom and LSR hydrotreater
- HF alkylation
- FCCU and light ends recovery
- Diesel hydrotreater
- MTBE
- Two saturated gas plants
- Two sulfur recovery trains and tail gas units
- Hydrocracker
- Utility system

During the first phase, the team performed an overall site analysis that included energy targeting, utility placement, and total site targeting. Table 15.12 summarizes the results of the study for each of the individual processes and for the total site. The table compares the total existing hot and cold utility usage for each of the processes to the total hot and cold utility targets. The final rows repeat this comparison for the total site.

[c]Data from EPRI, Pinch Technology/Process Optimization: Volume 9: Case Study—Marathon Oil Company, EPRI, Palo Alto, CA, 1997.

TABLE 15.12 Sitewide Energy Targets for a Petroleum Refinery

Unit	Rating	Total Hot Existing	Total Hot Target	Hot Delta	Total Cold Existing	Total Cold Target	Cold Delta	Sum Delta
Crude/Vacuum	1	199.9	181.5	18.4	106.9	88.0	18.9	−0.5
FCC	1	20.8	0.0	20.8	177.9	157.1	20.8	0.0
Isom	1	15.0	11.0	4.0	17.8	13.8	4.0	0.0
CCR reformer	1	196.4	167.5	28.9	115.1	86.2	28.9	0.0
Coker A	1	34.1	31.6	2.5	52.2	49.7	2.5	0.0
Coker B	1	39.1	38.1	1.0	53.7	52.7	1.0	0.0
Reformer	1	159.9	139.1	20.8	87.0	66.2	20.8	0.0
Hydrocracker	1	65.4	37.1	28.3	140.0	111.7	28.3	0.0
HF alkylation	2	62.3	59.8	2.5	81.2	78.7	2.5	0.0
DHT	2	31.9	21.0	10.9	71.7	60.8	10.9	0.0
MTBE	2	30.6	29.3	1.3	31.6	30.3	1.3	0.0
Amine	3	22.1	22.1	0.0	21.3	21.3	0.0	0.0
Naphtha splitter	3	12.2	12.2	0.0	12.0	12.0	0.0	0.0
BT/PFS towers	3	26.2	26.2	0.0	22.1	22.1	0.0	0.0
Sat gas 23	3	22.1	21.9	0.2	23.0	22.7	0.3	−0.1
Sat gas 8	3	31.0	31.1	−0.1	32.0	32.1	−0.1	0.0
SWS	3	24.2	24.2	0.0	23.1	23.1	0.0	0.0
Sulfur plants	3	6.9	2.2	4.7	51.4	46.7	4.7	0.0
Total		1000.0	856.0	144.0	1119.9	975.3	144.6	−0.6
Total site		0.0	853.6		0.0	968.2		

The estimated existing process hot utility usage for the refinery is 1000 energy units and the energy target for the complete refinery is 856 energy units. Therefore, the scope for energy reduction is 144 energy units, or 14.4%.

The table indicates that of the 18 units included in the study, there are eight that are key to the overall refinery efficiency improvement, three others have an impact but are less important, and the remaining seven need not be considered further. The key units are the crude/vacuum unit, FCC, Isom, CCR reformer, two cokers, reformer, and hydrocracker. The eight key units account for 73% of the existing hot utility usages, or 730 energy units, and 71% of the energy target, or 606 energy units. Therefore, concentrating on the eight key units reduces the scope for energy reduction to 124 energy units, or 12% of the site.

Phase 2 of the project involved using pinch analysis to study the key units in detail. Six of the eight key units were chosen:

1. Coker A
2. Coker B
3. Crude/vacuum unit
4. CCR reformer
5. Hydrocracker
6. Reformer

For reasons mainly associated with future investment plans and plant turnaround schedules, the FCC and Isom were not studied further.

Phase 2 analysis resulted in the identification of several projects with energy cost savings ranging from \$200,000/year to \$1,410,000/year and simple payback periods of as low as 0.5 years up through 16 years. The projects with the biggest energy savings were associated with the longest payback periods; however, these projects also offer significant, cost-effective debottlenecking potential for the units. Implementation of all mutually exclusive projects would result in energy savings of about \$4.6 million/year.

A plant-side steam balance was also conducted. The results from the steam balance showed an overall energy savings potential in the steam system of about \$5 million/year with a payback of 1–4 years. However, it is important to note that projects identified within the units and within the steam system sometimes compete. Therefore, energy cost savings for the projects and steam system are not directly additive.

15.2.4.2 Case Study 2: Minimizing Process Energy Use in a Wax Extraction Plant[d]

Together with support from a local utility company, EPRI and Linnhoff March, Inc. conducted a comprehensive energy study using pinch analysis at a wax extraction plant in Pennsylvania. The plant produces high-grade waxes for use in a variety of commercial and consumer markets. The objective of the study was to identify ways for the plant to reduce processing costs per barrel of throughput. The plant's first priority was to implement quick payback projects such as insulation of storage tanks. The plant was also interested in increasing manufacturing flexibility and achieving compliance with a nitrogen oxide (NO_x) abatement order at a minimum cost.

The waxes produced at the plant are petroleum based. Waxy crudes and distillates are first fractionated in a VDU to separate the oil into several distinct cuts. The VDU bottoms stream is treated with propane in the propane deasphalting (PDA) unit to separate out the asphaltenes (or heavy resin fraction), and the main oil fraction is recovered for further processing. The side cuts from VDU and recovered oil from PDA are then subjected to solvent extraction using methyl ethyl ketone (MEK) to separate wax from oil. The waxes are purified by bauxite filtration. The residual oils are sold as by-products to lube oil processors and heating oil dealers, and asphalt is used onsite as fuel.

[d]Data from EPRI, Pinch Technology/Process Optimization: Volume 7: Case Study—A Wax Extraction Plant, EPRI, Palo Alto, CA, 1995.

TABLE 15.13 Economics of Developed Projects for a Wax Extraction Plant

Project	Savings ($/year)	Capital Cost ($)	Payback (years)
VDU retrofit	300,000	300,000	1.0
MEK retrofit	350,000	1,500,000	4.2
Tank insulation	280,000	300,000	1.0
Total	930,000	2,100,000	2.3

The project team was charged with improving operating economics, debottlenecking processing capacity, evaluating the feasibility of hydrotreating vs. bauxite filtration for the wax finishing operation, and determining the optimum combined heat and power strategy for the site, paying particular heed to compliance with rather stringent NO_x abatement requirements.

The results of the study yielded a comprehensive process improvement strategy that improved wax yield by 8%, debottlenecked refrigeration capacity by 10%, and reduced total steam demand by 22%. The steam savings would allow shut down of two of the plant's four boilers, automatically reducing NO_x emissions, enabling compliance with the NO_x abatement order without installing expensive end-of-pipe controls. Implementation of suggested retrofits in the VDU and MEK process areas along with tank insulation would potentially yield operating savings of $930,000 million/year with a cost of $2.1 million and a simple payback period of 2.3 years. Evaluations related to cogeneration and to replacing the existing bauxite process with a hydrotreater showed that the economics did not favor either measure. Table 15.13 summarizes the economics of the proposed retrofits. Note that although the stand-alone economics of the MEK revamp are not particularly attractive (i.e., the payback period is 4.2 years), this project is essential to enable shutdown of two of the boilers, thereby achieving a relatively low-cost NO_x reduction.

References

1. Linnhoff, B. 1994. "Use pinch analysis to knock down capital costs and emissions," *Chem. Eng. Prog.*, 33, Aug.
2. Linnhoff, B. and Turner, J. A. 1981. "Heat recovery networks: new insights yield big savings," *Chem. Eng.*, 56, Nov. 2.
3. Linnhoff, B. and Vredeveld, D. R. 1984. "Pinch technology has come of age," *Chem. Eng. Prog.*, 80 (7), 33.
4. Korner, H. 1988. "Optimal use of energy in the chemical industry," *Chem. Eng. Tech.*, 60 (7), 511.
5. Fraser, D. M. and Gillespie, N. E. 1992. "The application of pinch technology to retrofit energy integration of an entire oil refinery," *Trans. IChemE*, 70, Part A, 395.
6. Samdani, G. and Moore, S. 1993. "Pinch technology: doing more with less," *Chem. Eng.*, 43.
7. Bell, K. J. 1978. "Estimate S & T exchanger design fast," *Oil Gas J.*, 59, 4.
8. Hohmann, E. C. 1971. "Optimum networks for heat exchange," PhD. thesis, University of Southern California, Los Angeles.
9. Linnhoff, B. and Flower, J. R. 1978. "Synthesis of heat exchanger networks," *AIChE J*, 633.
10. Townsend, D. W. and Linnhoff, B. 1984. "Surface area targets for heat exchanger networks," *IChemE 11th Annual Research Meeting*, Bath University, U.K., April.
11. Linnhoff, B., Mason, D. R., and Wardle, I. 1979, "Understanding heat exchanger networks," *Comp. & Chem. Eng.*, 3, 295.
12. Trivedi, K. K., Roach, J. R., and O'Neill, B. K. 1987. "Shell targeting in heat exchanger networks," *AIChE J.*, 33 (12), 2087.
13. Linnhoff, B. 1986. "The process/utility interface," paper presented at the second international meeting, "*National Use of Energy,*" Liege, Belgium, March.

14. Ahmad, S. and Smith, R. 1989. "Targets and design for minimum number of shells in heat exchanger networks," *Chem. Eng. Res. Des. Devel.*, 67, 481.
15. Linnhoff, B. and Ahmad, S. 1989. "Supertargeting: optimum synthesis of energy management systems," *J. Energ. Resour. Technol.*, 111, 121.
16. Umeda, T., Itoh, J., and Shiroko, K. 1978. "Heat exchange system synthesis," *Chem. Eng. Prog.*, 70.
17. Linnhoff, B. and Hindmarsh, E. 1983. "The pinch design method for heat exchanger networks," *Chem. Eng. Sci.*, 38 (5), 745.
18. Linnhoff, B., Townsend, D. W., Boland, D., Hewitt, G. F., Thomas, B. E. A., Guy, A. R., and Marshland, R. H. 1982. "Users guide on process integration for the efficient use of energy," *IChemE*, Rugby, U.K.
19. Linnhoff, B. and Ahmad, S. 1990. "Cost optimum heat exchanger networks—part i: minimum energy and capital using simple models for capital cost," *Comp. & Chem. Eng.*, 14 (7), 729.
20. Su, L.J. 1979. "A loop breaking evolutionary method for the synthesis of heat exchanger networks," MS thesis, Sever Institute of Washington University.
21. Forder, G. J. and Hutchison, H. P. 1969. The analysis of chemical flowsheets. *Chem. Eng. Sci.*, 24, 771.
22. Trivedi, K. K., O'Neill, B. K., Roach, J. R., and Wood, R. M. 1990. A best-first search method for energy relaxation. *Eng. Optim.*, 16, 291.
23. Trivedi, K. K., O'Neill, B. K., Roach, J. R., and Wood, R. M. 1990. Systematic energy relaxation in mer heat exchanger networks. *Comp. & Chem. Eng.*, 14 (6), 601.
24. O'Reilly, M. 1985. Personal view. *Chem. Eng.*, 410, 46.
25. Trivedi, K. K., O'Neill, B. K., Roach, J. R., and Wood, R. M. 1989. A new dual-temperature design method for the synthesis of heat exchanger networks. *Comp. & Chem. Eng.*, 13 (6), 667.
26. Grimes, L. E., Rychener, M. D., and Westerberg, A. W. 1982. The synthesis and evolution of networks of heat exchanges that feature the minimum number of units. *Chem. Eng. Commu.*, 14, 339.
27. Townsend, D. W., Surface area and capital cost targets for process energy systems. Ph.D. thesis, UMIST, Manchester, U.K.
28. Townsend, D. W. and Linnhoff, B. 1983. Heat and power networks in process design: part I: criteria for placement of heat engines and heat pumps in process networks; part II: design procedure for equipment selection and process matching. *AIChE J*, 29 (5), 742.
29. Hall, S. G., 1989. Targeting for multiple utilities in pinch technology. Ph.D. thesis, UMIST, Manchester, U.K.
30. Parker, S. J. 1989. Supertargeting for multiple utilities. Ph.D. thesis, UMIST, Manchester, U.K.
31. Dhole, V. R. and Linnhoff, B. 1993. Distillation column targets. *Comp. & Chem. Eng.*, 17 (5/6), 549.
32. Linnhoff, B., Dunford, H., and Smith, R. 1983. Heat integration of distillation columns into overall processes. *Chem. Eng. Sci.*, 38 (8), 1175.
33. Linnhoff, B. and de Leur, J. 1988. Appropriate placement of furnaces in the integrated process, Paper presented at IChemE Symposium—Understanding Process Integration—II, UMIST, Manchester, U.K.
34. Linnhoff, B. and Dhole, V. R. 2081. Shaftwork targets for low temperature process design. *Chem. Eng. Sci.*, 47 (8), 2081.
35. Raissi, K. 1994. Total site integration. Ph.D. thesis, UMIST, Manchester, U.K.
36. Smith, R. and Delaby, O. 1992. Targeting flue gas emissions. *Trans., IChemE*, 69, Part A, 492.
37. Linnhoff, B. and Dhole, V. R. 1993. Targeting for CO_2 emissions for total sites. *Chem. Ing. Tech.*, 16, 256.
38. Wood, R. M., Wilcox, R. J., and Grossmann, I. E. 1985. A note of the minimum number of units for heat exchanger network synthesis. *Chem. Eng. Commun.*, 39, 371.
39. Trivedi, K. K., Pang, K. H., Klavers, H. R., O'Young, D. L., and Linnhoff, B. 1994. Integrated Ethylene Process Design using Pinch Technology," presented at AIChE Meeting, Atlanta, April.

40. Ma, J. L. James. 1991. Ethylene Supplement E, Report 29E, SRI International, Menlo Park, CA, Oct.
41. Kotjabasakis, E. and Linnhoff, B. 1986. Sensitivity tables for the design of flexible processes, part I: how much contigency in heat exchange networks is cost-effective?. *Chem. Eng. Res. & Des.*, 64, 197.
42. Polley, G. T., Panjeh Shahi, M. H., Jegede, F. O. 1990. Pressure drop consideration in the retrofit of heat exchanger networks, *Trans. IChemE*, 68, Part A, 211.

16

Energy Audits for Buildings

Moncef Krarti

University of Colorado

16.1 Introduction

This chapter describes energy audit procedures commonly used to improve the energy efficiency of residential and commercial buildings as well as industrial facilities. Moreover, the chapter summarizes proven energy-efficient technologies in the building sectors with some examples to highlight the cost-effectiveness of some of these technologies. A brief overview is also provided for currently available energy management programs where energy audit is crucial for their proper and successful implementation.

16.2 Background

To reduce the operating costs and the environmental impact associated with utilizing conventional energy resources, energy conservation and energy efficiency offer attractive solutions. Moreover, energy efficiency can avoid the need to build new power plants—that use conventional energy sources—at little cost and with no adverse environmental impact. In addition, energy efficiency and energy conservation have other beneficial impacts:

- Increases economic competitiveness. As stated by the International Energy Agency (IEA), investment in energy conservation provides a better return than investment in energy supply.

- Stretches the availability of limited nonrenewable energy resources and gains time for possible development of renewable and reliable energy resources such as solar energy.
- Decreases air and water pollution and thus improves health conditions.

Around the world, there is a vast potential for energy efficiency that has begun to be tapped in only a few countries. This potential exists for all energy end use sectors including buildings, industries, and transportation. One of the main challenges in this new millennium will be to increase the efficiency of production, distribution, and consumption of energy that will reduce costs and lower the environmental impacts. Therefore, energy efficiency can have beneficial impacts on economic competitiveness, the environment, and health.

In several industrialized countries, energy consumption has fluctuated in response to significant changes in oil prices, economic growth rates, and environmental concerns, especially since the oil crisis of the early 1970s. For instance, the U.S. energy consumption increased from 66 quadrillion British thermal units (Btu) in 1970–1994 quadrillion Btu in 1998 (EIA 1998). The energy costs in the U.S. economy represent about 8% of the gross domestic product (GDP), which is one of the highest among industrialized countries. One of the reasons for the high energy costs is that the U.S. consumes a significant fraction of the total world energy. Thus, the U.S. has the highest per capita energy-use rate in the world with an average of 350 million Btu per year, or the equivalent of 7 gallons of oil per person per day.

Figure 16.1 illustrates the rate of growth of the per capita energy-use and the population relative to 1973. It is interesting to note that the per capita energy-use rate remains almost constant—with relatively small fluctuations—since 1973 even though the population growth rate has clearly increased throughout the years. The higher oil prices in the 1970s (oil embargo in 1973 and the Iranian revolution in 1979) have mandated energy conservation and increased energy efficiency. The trend toward energy conservation, although relaxed during the 1980s, had continued in the 1990s due to the 1992 National Energy Policy Act (EPACT) which promotes more efficient use of energy in the U.S. In particular, the EPACT revises energy efficiency standards for buildings, promotes use of alternative fuels, and reduces the monopolistic structure of electric and gas utilities.

Figure 16.2 presents the total U.S. energy consumption distribution by major sectors for 1996. As indicated, buildings and industrial facilities are responsible, respectively, for 36 and 38% of the total U.S. energy consumption. The transportation sector, which accounts for the remaining 26% of the total U.S. energy consumption, uses mostly fuel products. However, buildings and industries consume predominantly electricity and natural gas. Because of its low price, coal is primarily used as an energy source for electricity generation.

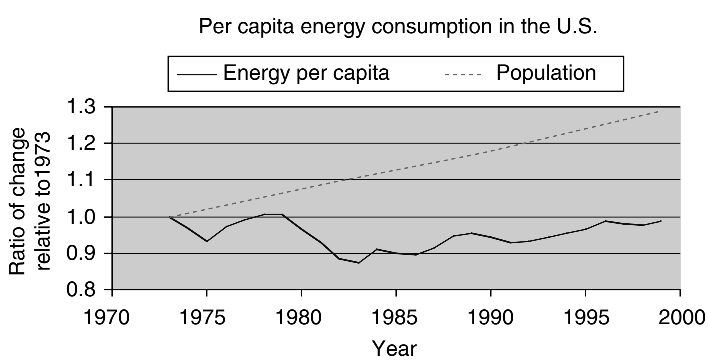


FIGURE 16.1 Per capita energy-use and population growth since 1973. (From EIA, *Annual Energy Review*, Department of Energy, Energy Information Administration, 1998. http://www.doe.eia.gov).

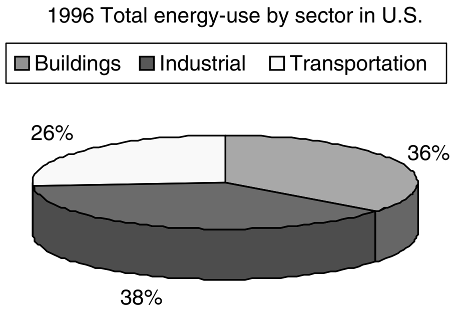


FIGURE 16.2 Distribution of U.S. energy consumption by end use sector. (From EIA, *Annual Energy Review*, Department of Energy, Energy Information Administration, 1998. http://www.doe.eia.gov).

Despite some improvements in energy efficiency over the last 25 years, the U.S. remains the most energy-intensive in the world. If it wants to maintain its lead in a global and competitive world economy, it is imperative that the U.S. continues to improve its energy efficiency.

In most countries, residential and commercial buildings account for a significant portion of the total national energy consumption (almost 40% in the U.S. and in France). Typically, buildings use electricity and a primary energy source such as natural gas or fuel oil. Electricity is used for lighting, appliances, and HVAC equipment. Typical energy density for selected types of commercial and institutional buildings are summarized in Table 16.1 for both the U.S. and France.

The industrial sector consumes more than 35% of the total U.S. energy-use as indicated in Figure 16.2. Fossil fuels constitute the main source for the U.S. industry. Electricity accounts for about 15% of the total U.S. industrial energy-use. In some energy-intensive manufacturing facilities, cogeneration systems are used to produce electricity from fossil fuels. A significant potential for energy savings exist in industrial facilities due to the vast amounts of energy wasted in the industrial processes. Using improved housekeeping measures and recovering some of the waste heat, the U.S. could save up to 35% of the total energy-used in the industry (Ross and Williams 1977).

The potential for energy conservation for both buildings and industrial sector remains large is the U.S. and other countries despite the improvements in the energy efficiency since the 1970s. Energy management programs using proven and systematic energy audit procedures suitable for both buildings and industrial facilities are provided in the following sections. In addition, some proven and cost-effective energy efficiency technologies are summarized.

TABLE 16.1 Energy Intensity by Principal Building Activity in kWh/m^2

Major Building Activity	France	US
Office	395	300
Education	185	250
Health care	360	750
Lodging	305	395
Food service	590	770
Mercantile and service	365	240
Sports	405	NA[a]
Public assembly	NA	375
Warehouse and storage	NA	125

[a] Not Available.

Source: From CEREN, *La Consommation d"Energie Dans les Regions Francaises*. Report from Centre d'Etudes et de Recherches Economiques sur l Energie 1997; Energy Information Administration (EIA), Annual Energy review. Department of Energy. 1998. http://www.doe.eia.gov (accessed on 2005).

16.3 Energy Audit Procedures

16.3.1 Energy Audit Types

Energy audits are the first step to improve the energy efficiency of buildings and industrial facilities. Generally, four types of energy audits can be distinguished as briefly described below (Krarti 2000):

- A walk-through audit consists typically of a short on-site visit of the facility to identify areas where simple and inexpensive actions (typically operating and maintenance measures) can provide immediate energy-use and/or operating cost savings.
- A utility cost analysis includes a careful evaluation of metered energy-uses and operating costs of the facility. Typically, the utility data over several years are evaluated to identify the patterns of energy-use, peak demand, weather effects, and potential for energy savings.
- A standard energy audit consists of a comprehensive energy analysis for the energy systems of the facility. In particular, the standard energy audit includes the development of a baseline for the energy-use of the facility, the evaluation of the energy savings, and the cost-effectiveness of appropriately selected energy conservation measures.
- A detailed energy audit is the most comprehensive but also time-consuming energy audit type. Specifically, the detailed energy audit includes the use of instruments to measure energy-use for the whole building and/or for some energy systems within the building (for instance by end uses such as lighting systems, office equipment, fans, chillers, etc.). In addition, sophisticated computer simulation programs are typically considered for detailed energy audits to evaluate and recommend energy retrofits for the facility.

16.3.2 General Procedure for a Detailed Energy Audit

To perform an energy audit, several tasks are typically carried out depending on the type of the audit and the size and function of the audited building. Some of the tasks may have to be repeated, reduced in scope, or even eliminated based on the findings of other tasks. Therefore, the execution of an energy audit is often not a linear process and is rather iterative. However, a general procedure can be outlined for most facilities.

Step 1: Facility and Utility Data Analysis. The main purpose of this step is to evaluate the characteristics of the energy systems and the patterns of energy-use for the building or the facility. The building/facility characteristics can be collected from the architectural/mechanical/electrical drawings and/or from discussions with building operators. The energy-use patterns can be obtained from a compilation of utility bills over several years. Analysis of the historical variation of the utility bills allows the energy auditor to determine if there are any seasonal and weather effects on the building energy-use. Some of the tasks that can be performed in this step are presented below with the key results expected from each task noted:

- Collect at least three years of utility data (to identify a historical energy-use pattern).
- Identify the fuel types used such as electricity, natural gas, oil, etc. (to determine the fuel type that accounts for the largest energy-use).
- Determine the patterns of fuel use by fuel type (to identify the peak demand for energy-use by fuel type).
- Understand utility rate structure (energy and demand rates) (to evaluate if the building is penalized for peak demand and if cheaper fuel can be purchased).
- Analyze the effect of weather on fuel consumption (to pinpoint any variations of energy-use related to extreme weather conditions).

- Perform utility energy-use analysis by building type and size; building signature can be determined including energy-use per unit area (to compare against typical indices).

Step 2: *Walk-Through Survey*. From this step, potential energy savings measures should be identified. The results of this step are important because they determine if the building warrants any further energy auditing work. Some of the tasks involved in this step are:

- Identify the customer concerns and needs.
- Check the current operating and maintenance procedures.
- Determine the existing operating conditions of major energy-use equipment (lighting, HVAC systems, motors, etc.).
- Estimate the occupancy, equipment, and lighting (energy-use density and hours of operation).

Step 3: *Baseline for Building Energy-Use*. The main purpose of this step is to develop a base-case model that represents the existing energy-use and operating conditions for the building. This model is to be used as a reference to estimate the energy savings incurred from appropriately selected energy conservation measures. The major tasks to be performed during this step are as follows:

- Obtain and review architectural, mechanical, electrical, and control drawings.
- Inspect, test, and evaluate building equipment for efficiency, performance, and reliability.
- Obtain all occupancy and operating schedules for equipment (including lighting and HVAC systems).
- Develop a baseline model for building energy-use.
- Calibrate the baseline model using the utility data and/or metered data.

Step 4: *Evaluation of Energy Savings Measures*. In this step, a list of cost-effective energy conservation measures is determined using both energy savings and economic analysis. To achieve this goal, the following tasks are recommended:

- Prepare a comprehensive list of energy conservation measures (using the information collected in the walk-through survey).
- Determine the energy savings due to the various energy conservation measures pertinent to the building using the baseline energy-use simulation model developed in phase 3.
- Estimate the initial costs required to implement the energy conservation measures.
- Evaluate the cost-effectiveness of each energy conservation measure using an economical analysis method (simple payback or life cycle cost analysis).

Table 16.2 and Table 16.3 provide summaries of the energy audit procedure recommended respectively for commercial buildings and for industrial facilities (Krarti 2000). Energy audits for thermal and electrical systems are separated because they are typically subject to different utility rates.

16.4 Energy Management Programs

This section describes energy conservation and energy efficiency programs that require energy auditing as a tool for proper and successful implementation and completion. These programs have been found to be effective in improving the energy efficiency of commercial buildings and industrial facilities.

16.4.1 Performance Contracting

Over the last decade, a new mechanism for funding energy projects has been proposed to improve energy efficiency of existing buildings. This mechanism, often called *performance contracting*, can be structured

TABLE 16.2 Energy Audit Summary for Residential and Commercial Buildings

Phase	Thermal Systems	Electric Systems
Utility analysis	Thermal energy-use profile (building signature)	Electrical energy-use profile (building signature)
	Thermal energy-use per unit area (or per student for schools or per bed for hospitals)	Electrical energy-use per unit area (or per student for schools or per bed for hospitals)
	Thermal energy-use distribution (heating, DHW, process, etc.)	Electrical energy-use distribution (cooling, lighting, equipment, fans, etc.)
	Fuel types used	Weather effect on electrical energy-use
	Weather effect on thermal energy-use	Utility rate structure (energy charges, demand charges, power factor penalty, etc.)
	Utility rate structure	
On-site survey	Construction Materials (thermal resistance type and thickness)	HVAC system type
	HVAC system type	Lighting type and density
	DHW system	Equipment type and density
	Hot water/steam use for heating	Energy-use for heating
	Hot water/steam for cooling	Energy-use for cooling
	Hot water/steam for DHW	Energy-use for lighting
	Hot water/steam for specific applications (hospitals, swimming pools, etc.)	Energy-use for equipment
		Energy-use for air handling
		Energy-use for water distribution
Energy-use baseline	Review architectural, mechanical, and control drawings	Review architectural, mechanical, electrical, and control drawings
	Develop a base-case model (using any baselining method ranging from very simple to more detailed tools)	Develop a base-case model (using any baselining method ranging from very simple to more detailed tools)
	Calibrate the base-case model (using utility data or metered data)	Calibrate the base-case model (using utility data or metered data)
Energy conservation measures	Heat recovery system (heat exchangers)	Energy-efficient lighting
	Efficient heating system (boilers)	Energy-efficient equipment (computers)
	Temperature setback	Energy-efficient motors
	EMCS	HVAC system retrofit
	HVAC system retrofit	EMCS
	DHW use reduction	Temperature setup
	Cogeneration	Energy-efficient cooling system (chiller)
		Peak demand shaving
		Thermal energy storage system
		Cogeneration
		Power factor improvement
		Reduction of harmonics

using various approaches. The most common approach for performance contacting consists of the following steps:

- A vendor or contractor proposes an energy project to a facility owner or manager after conducting an energy audit. This energy project would save energy-use and energy cost and thus would reduce the facility operating costs.
- The vendor/contractor funds the energy project using typically borrowed moneys from a lending institution.
- The vendor/contractor and facility owner/manager agree on a procedure to repay the borrowed funds from energy cost savings that may result from the implementation of the energy project.

An important feature of performance contracting is the need for a proven protocol for measuring and verifying energy cost savings. This measurement and verification protocol has to be accepted by all the parties involved in the performance contracting project: the vendor/contractor, the facility owner/manager, and the lending institution. For different reasons, all the parties have to insure that cost savings

TABLE 16.3 Energy Audit Summary for Industrial Facilities

Phase	Thermal Systems	Electric Systems
Utility analysis	Thermal energy-use profile (building signature)	Electrical energy-use profile (building signature)
	Thermal energy-use per unit of a product	Electrical energy-use per unit of a product
	Thermal energy-use distribution (heating, process, etc.)	Electrical energy-use distribution (cooling, lighting, equipment, process, etc.)
	Fuel types used	Analysis of the electrical energy input for specific processes used in the production line (such as drying)
	Analysis of the thermal energy input for specific processes used in the production line (such as drying)	
	Utility rate structure	Utility rate structure (energy charges, demand charges, power factor penalty, etc.)
On-site survey	List of equipment that use thermal energy	List of equipment thqt use electrical energy
	Perform heat balance of the thermal energy	Perform heat balance of the electrical energy
	Monitor thermal energy-use of all or part of the equipment	Monitor electrical energy-use of all or part of the equipment
	Determine the by-products of thermal energy-use (such emissions and solid waste)	Determine the by-products of electrical energy-use (such pollutants)
Energy-use baseline	Review mechanical drawings and production flow charts	Review electrical drawings and production flow charts
	Develop a base-case model using (any baselining method)	Develop a base-case model (using any baselining method)
	Calibrate the base-case model (using utility data or metered data)	Calibrate the base-case model (using utility data or metered data)
Energy conservation measures	Heat recovery system	Energy-efficient motors
	Efficient Heating and drying system	Variable speed drives
	EMCS	Air compressors
	HVAC system retrofit	Energy-efficient lighting
	Hot water and steam use reduction	HVAC system retrofit
	Cogeneration (possibly with solid waste from the production line)	EMCS
		Cogeneration (possibly with solid waste from the production line)
		Peak demand shaving
		Power factor improvement
		Reduction of harmonics

have indeed incurred from the implementation of the energy project and are properly estimated. Over the last decade, several methods and protocols for measuring and verifying actual energy savings from energy efficiency projects in existing buildings have been developed (Krarti 2000). Among the methods proposed for the measurement of energy savings are those proposed by the National Association of Energy Service Companies (NAESCO 1993), the Federal Energy Management Program (FEMP 1992), the American Society of Heating Refrigeration and Air Conditioning Engineers (ASHRAE 1997), the Texas LoanSTAR program (Reddy, Kissock and Katipamula 1994), and the North American Energy Measurement and Verification Protocol (NEMVP) sponsored by DOE and later updated and renamed the International Performance Measurement and Verification Protocol (IPMVP 1997).

16.4.2 Commissioning of Building Energy Systems

Before final occupancy of a newly constructed building, it is recommended to perform commissioning of its various systems including structural elements, building envelope, electrical systems, security systems, and HVAC systems. The commissioning is a quality assurance process to verify and document the performance of building systems as specified by the design intent. During the commissioning process, operation and maintenance personnel are trained to properly follow procedures in order that all building systems are fully functional and are properly operated and maintained.

16.4.3 Energy Rating of Buildings

In the U.S., a new building rating system has been recently developed and implemented by the U.S. Green Building Council. This rating system is called the *Leadership in Energy and Environmental Design* (LEED), and it considers the energy and the environmental performance of all the systems in a building over its life cycle. Currently, the LEED rating system evaluates new and existing commercial, institutional, and high-rise residential buildings. The rating is based on credits that can be earned if the building satisfies a list of criteria based on existing proven technology. Different levels of green building certification are awarded based on the total credit earned.

Other countries have similar rating systems. In fact, England was the first country to develop and implement a national green building rating system, the Building Research Establishment's Environmental Assessment Method (BREEAM). The Building Research Establishment estimates that up to 30% of office buildings constructed in the last 7 years have been assessed using BREEAM rating system. Currently, BREEAM rating system can be applied to new and existing office buildings, industrial facilities, residential homes, and superstores.

16.5 Energy Conservation Measures

In this section energy conservation measures commonly implemented for commercial and industrial facilities are briefly discussed. The potential energy savings and the cost-effectiveness of some of the energy efficiency measures are discussed through illustrative examples. The calculation details of the energy savings incurred for common energy conservation measures can be found in Krarti (2000).

16.5.1 Building Envelope

For some buildings, the envelope (i.e., walls, roofs, floors, windows, and doors) has an important impact on the energy-used to condition the facility. The energy efficiency of the building envelope can be characterized by its building load coefficient (BLC). The BLC can be estimated either by a regression analysis of the utility data or by a direct calculation using the thermal resistance of the construction materials used in the building-envelope assemblies (i.e., walls, roofs, windows, doors, etc.). Figure 16.3 illustrates how the BLC for a given building can be estimated using utility data (Krarti 2000).

Some of the commonly recommended energy conservation measures to improve the thermal performance of the building envelope include the following:

1. *Addition of Thermal Insulation*. For building surfaces without any thermal insulation, this measure can be cost-effective.
2. *Replacement of Windows*. When windows represent a significant portion of the exposed building surfaces, using more energy-efficient windows (high R-value, low-emissivity glazing, air tight, etc.) can be beneficial in both reducing the energy-use and improving the indoor comfort level.
3. *Reduction of Air Leakage*. When infiltration load is significant, leakage area of the building envelope can be reduced by simple and inexpensive weather-stripping techniques. In residential buildings, the infiltration rate can be estimated using a blower door test setup as shown in Figure 16.4. The blower test door setup can be used to estimate the infiltration or exfiltration rates under both pressurization and depressurization conditions.

The energy audit of the envelope is especially important for residential buildings. Indeed, the energy-use from residential buildings are dominated by weather because heat gain and/or loss from direct conduction of heat or from air infiltration/exfiltration through building surfaces accounts for a major portion (50%–80%) of the energy consumption. For commercial buildings, improvements to the building envelope are often not cost-effective because modifications to the building envelope (replacing windows, adding thermal insulation in walls) are typically considerably expensive. However, it is recommended to

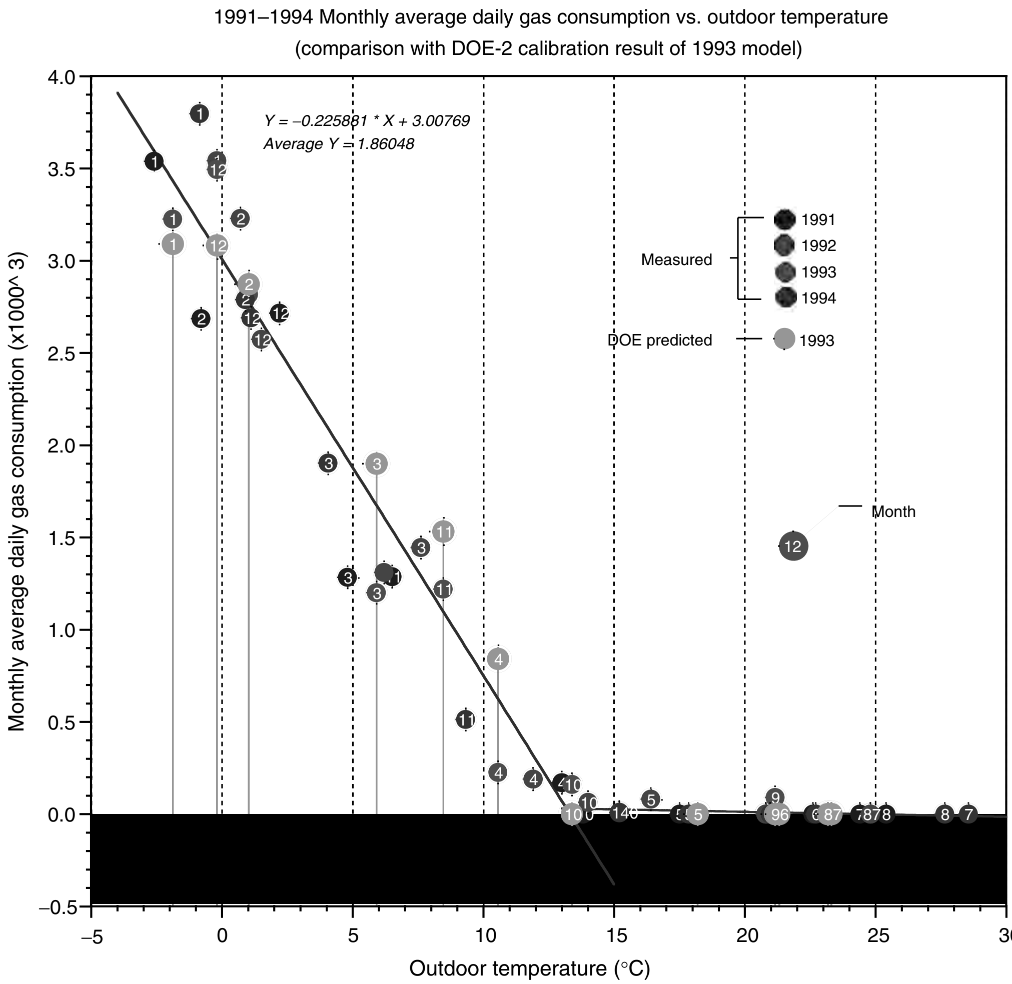


FIGURE 16.3 Estimation of the BLC based on a regression analysis of the monthly gas consumption. (From Krarti, M., *Energy Audit of Building Systems: An Engineering Approach*, CRC Press, Boca Raton, FL, 2000.)

systematically audit the envelope components not only to determine the potential for energy savings but also to insure the integrity of its overall condition. For instance, thermal bridges, if present, can lead to heat transfer increase and to moisture condensation. The moisture condensation is often more damaging and costly than the increase in heat transfer because it can affect the structural integrity of the building envelope.

16.5.2 Ventilation and Indoor Air Quality

16.5.2.1 Ventilation in Commercial/Institutional Buildings

The energy required to condition ventilation air can be significant in both commercial buildings and industrial facilities, especially in locations with extreme weather conditions. Whereas the ventilation is used to provide fresh air to occupants in commercial buildings, it is used to control the level of dust, gases, fumes, or vapors in several industrial applications. The auditor should estimate the existing volume of fresh air and compare this estimated amount of the ventilation air with that required by the

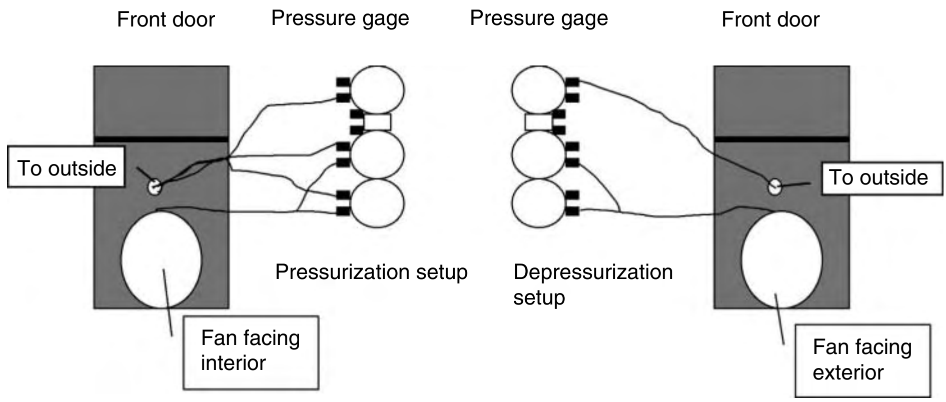


FIGURE 16.4 A blower door test setup for both pressurization and depressurization. (From Krarti, M., *Energy Audit of Building Systems: An Engineering Approach*, CRC Press, Boca Raton, FL, 2000.)

appropriate standards and codes. Excess in air ventilation should be reduced if it can lead to increases in heating and/or cooling loads. However, in some climates and periods of the year or the day, providing more air ventilation can be beneficial and may actually reduce cooling and heating loads through the use of air-side economizer cycles.

Table 16.4 summarizes some of the minimum outdoor air requirements for selected spaces in commercial buildings.

If excess ventilation air is found, the outside air damper setting can be adjusted to supply the ventilation that meets the minimum outside requirements as listed in Table 16.5. Further reductions in outdoor air can be obtained by using demand ventilation controls by supplying outside air only during periods when there is need for fresh air. A popular approach for demand ventilation is the monitoring of CO_2 concentration level within the spaces. CO_2 is considered as a good indicator of pollutants generated by occupants and other construction materials. The outside air damper position is controlled to maintain a CO_2 set-point within the space. CO_2-based demand-controlled ventilation has been implemented in various buildings with intermittent occupancy patterns including cinemas, theaters, classrooms, meeting rooms, and retail establishments. However, the ventilation for several office buildings has been controlled using CO_2 measurements (Emmerich and Persily 1997). Based on field studies, it has been found that significant energy savings can be obtained with a proper implementation of CO_2-based demand-controlled ventilation. Typically, the following building features are required for an effective performance of demand ventilation controls (Davidge 1991):

- Unpredictable variations in the occupancy patterns
- Requirement of either heating or cooling for most of the year
- Low pollutant emissions from non-occupant sources (i.e. furniture, equipment, etc.)

TABLE 16.4 Minimum Ventilation Rate Requirements for Selected Spaces in Commercial Buildings

Space and or Application	Minimum Outside Air Requirements
Office space	9.5 L/s (20 cfm) per person
Corridor	0.25 L/s per m^2 (0.05 cfm/ft.^{2})
Restroom	24 L/s (50 cfm) per toilet
Smoking lounge	28.5 L/s (60 cfm) per person
Parking garage	7.5 L/s (1.5 cfm/ft.^{2})

Source: From ASHRAE, *Ventilation for Acceptable Indoor Air Quality, Standard, 62*, American Society of Heating. Refrigerating and Air-Conditioning Engineers, Inc., Atlanta, GA, 1989.

TABLE 16.5 Typical Efficiencies of Motors

Motor Size (HP)	Standard Efficiency	High Efficiency(%)
1	72%	81%
2	76%	84%
3	77%	89%
5	80%	89%
7.5	82%	89%
10	84%	89%
15	86%	90%
20	87%	90%
30	88%	91%
40	89%	92%
50	90%	93%

It should be noted that although CO_2 can be used to control occupant-generated contaminants, it may not reliably control pollutants generated from non-occupant sources such as building materials. As a solution, a base ventilation rate can be maintained at all times to ensure that non-occupant contaminants are controlled (Emmerich and Persily 1997).

16.5.2.2 Ventilation of Parking Garages

Automobile parking garages can be partially open or fully enclosed. Partially open garages are typically above-grade with open sides and do not generally need mechanical ventilation. However, fully enclosed parking garages are usually underground and require mechanical ventilation. Indeed, in absence of ventilation, enclosed parking facilities present several indoor air quality problems. The most serious is the emission of high levels of carbon monoxide (CO) by cars within the parking garages. Other concerns related to enclosed garages are the presence of oil and gasoline fumes, and other contaminants such as oxides of nitrogen (NO_x) and smoke haze from diesel engines.

To determine the adequate ventilation rate for garages, two factors are typically considered: the number of cars in operation and the emission quantities. The number of cars in operation depends on the type of the facility served by the parking garage and may vary from 3% (in shopping areas) up to 20% (in sports stadium) of the total vehicle capacity (ASHRAE 1999). The emission of carbon monoxide depends on individual cars including such factors as the age of the car, the engine power, and the level of car maintenance.

For enclosed parking facilities, ASHRAE standard 62-1989 specifies fixed ventilation rate of below 7.62 L/sm^2 (1.5 cfm/ft.^{2}) of gross floor area (ASHRAE 1989). Therefore, a ventilation flow of about 11.25 air changes per hour is required for garages with 2.5-m ceiling height. However, some of the model code authorities specify an air change rate of four to six air changes per hour. Some of the model code authorities allow ventilation rate to vary and be reduced to save fan energy if CO demand-controlled ventilation is implemented, that is, a continuous monitoring of CO concentrations is conducted, with the monitoring system being interlocked with the mechanical exhaust equipment. The acceptable level of contaminant concentrations varies significantly from code to code. A consensus on acceptable contaminant levels for enclosed parking garages is needed. Unfortunately, ASHRAE standard 62-1989 does not address the issue of ventilation control through contaminant monitoring for enclosed garages. Thus, ASHRAE commissioned a research project 945-RP (Krarti, Ayari, Grot 1999) to evaluate current ventilation standards and recommend rates appropriate to current vehicle emissions/usage. Based on this project, a general methodology has been developed to determine the ventilation requirements for parking garages.

Figure 16.5 indicates also the fan energy savings achieved by the on–off and VAV systems (relative to the fan energy use by the CV system). As illustrated in Figure 16.5, significant fan energy savings can be

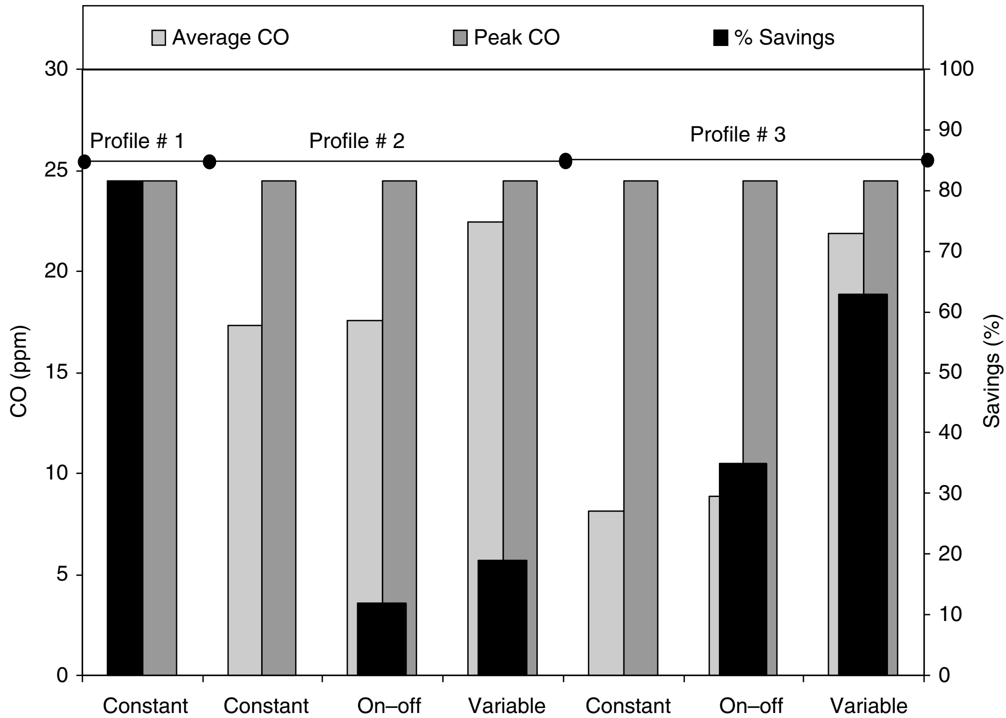


FIGURE 16.5 Typical energy savings and maximum CO level obtained for demand CO ventilation controls. (From Krarti, M., Ayari, A., and Grot, D., *Ventilation Requirements for Enclosed Vehicular Parking Garages*, Final Report for ASHRAE RP-945, American Society of Heating, Refrigerating, and Air Conditioning Engineering, Atlanta, GA, 1999.)

obtained when demand CO-ventilation control strategy is used to operate the ventilation system while also maintaining acceptable CO levels within the enclosed parking facility. These energy savings depend on the pattern of car movement within the parking facility. Figure 16.6 indicates three types of car movement profiles considered in the analysis considered by Krarti, Ayari and Grot (1999).

16.5.3 Electrical Systems

For most commercial buildings and a large number of industrial facilities, the electrical energy cost constitutes the dominant part of the utility bill. Lighting, office equipment, and motors are the electrical systems that consume the major part of energy in commercial and industrial buildings.

16.5.3.1 Lighting

For a typical office building lighting represents on average 40% of the total electrical energy-use. There are a variety of simple and inexpensive measures to improve the efficiency of lighting systems. These measures include the use of energy-efficient lighting lamps and ballasts, the addition of reflective devices, delamping (when the luminance levels are above the recommended levels by the standards), and the use of daylighting controls. Most lighting measures are especially cost-effective for office buildings for which payback periods are less than one year.

16.5.3.2 Daylighting

Several studies indicated that daylighting can offer a cost-effective alternative to electrical lighting for commercial and institutional buildings. Through sensors and controllers, daylighting can reduce and even eliminate the use of electrical lighting required to provide sufficient illuminance levels inside office

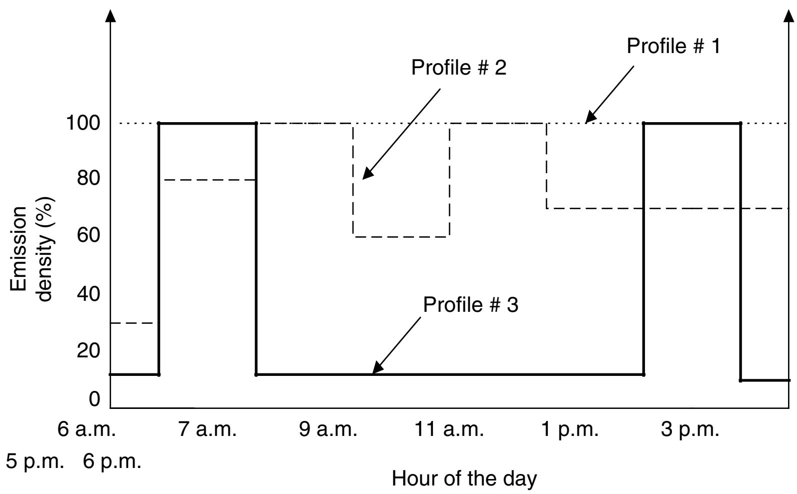


FIGURE 16.6 Car movement profiles used in the analysis conducted by Krarti et al. (From Krarti, M., Ayari, A., and Grot, D., *Ventilation Requirements for Enclosed Vehicular Parking Garages*, Final Report for ASHRAE RP-945, American Society of Heating, Refrigerating, and Air Conditioning Engineering, Atlanta, GA, 1999.)

spaces. Recently, a simplified calculation method has been developed by Krarti, Erickson and Hillman (2005) to estimate the reduction in the total lighting energy-use due to daylighting with dimming controls for office buildings. The method has been shown to apply for office buildings in the U.S. as well as in Egypt (Al-Moheimen, Hanna and Krarti 2005). The simplified calculation method is easy to use and can be applied as a predesign tool to assess the potential of daylighting in saving electricity use associated with artificial lighting for office buildings.

To determine the percent savings, f_d, in annual use of artificial lighting due to implementing daylighting using daylighting controls in office buildings, Krarti, Erickson and Hillman (2005) found that the following equation can be used:

$$f_d = b[1 - \exp(-a\tau_w A_w/A_p)]\frac{A_p}{A_f}, \tag{16.1}$$

where A_w/A_p is the window-to-perimeter floor area; this parameter provides a good indicator of the window size relative to the daylit floor area. A_p/A_f is the perimeter-to-total floor area; this parameter indicates the extent of the daylit area relative to the total building floor area. Thus, when $A_p/A_f = 1$, the whole building can benefit from daylighting. Parameters a and b in Equation 16.1 are coefficients that depends only on the building location and are given by Table 16.6 for various sites throughout the world.

16.5.3.3 Office Equipment

Office equipment constitutes the fastest growing part of the electrical loads, especially in commercial buildings. Office equipment includes computers, fax machines, printers, and copiers. Today, there are several manufacturers that provide energy-efficient office equipment (such those that comply with the U.S. EPA Energy Star specifications). For instance, energy-efficient computers automatically switch to a low-power "sleep" mode or off-mode when not in use.

16.5.3.4 Motors

The energy cost to operate electric motors can be a significant part of the operating budget of any commercial and industrial building. Measures to reduce the energy cost of using motors include reducing operating time (turning off unnecessary equipment), optimizing motor systems, using controls to match

TABLE 16.6 Coefficients *a* and *b* of Equation 16.1 for Various Locations throughout the World

Location	*a*	*b*	Location	*a*	*b*
Atlanta	19.63	74.34	Casper	19.24	72.66
Chicago	18.39	71.66	Portland	17.79	70.93
Denver	19.36	72.86	Montreal	18.79	69.83
Phoenix	22.31	74.75	Quebec	19.07	70.61
New York City	18.73	66.96	Vancouver	16.93	68.69
Washington DC	18.69	70.75	Regina	20.00	70.54
Boston	18.69	67.14	Toronto	19.30	70.48
Miami	25.13	74.82	Winnipeg	19.56	70.85
San Francisco	20.58	73.95	Shanghai	19.40	67.29
Seattle	16.60	69.23	K-Lumpur	20.15	72.37
Los Angels	21.96	74.15	Singapore	23.27	73.68
Madison	18.79	70.03	Cairo	26.98	74.23
Houston	21.64	74.68	Alexandria	36.88	74.74
Fort Worth	19.70	72.91	Tunis	25.17	74.08
Bangor	17.86	70.73	Sao Paulo	29.36	71.19
Dodge City	18.77	72.62	Mexico91	28.62	73.63
Nashville	20.02	70.35	Melbourne	19.96	67.72
Oklahoma City	20.20	74.43	Roma	16.03	72.44
Columbus	18.60	72.28	Frankfurt	16.22	69.69
Bismarck	17.91	71.50	Kuwait	21.98	65.31
Minneapolis	18.16	71.98	Riyadh	21.17	72.69
Omaha	18.94	72.30			

Source: From Krarti, M., Erickson, P., and Hillman. T., *Daylighting Building and Environment*, 40, 747–754, 2005.

motor output with demand, using variable speed drives for air and water distribution, and installing energy-efficient motors. Table 16.5 provides typical efficiencies for several motor sizes.

In addition to the reduction in the total facility electrical energy-use, retrofits of the electrical systems decrease space cooling loads and therefore further reduce the electrical energy-use in the building. These cooling energy reductions as well as possible increases in thermal energy-use (for space heating) should be accounted for when evaluating the cost-effectiveness of improvements in lighting and office equipment.

16.5.4 HVAC Systems

The energy-use due to HVAC systems can represent 40% of the total energy consumed by a typical commercial building. A large number of measures can be considered to improve the energy performance of both primary and secondary HVAC systems. Some of these measures are listed below:

- *Setting Up/Back Thermostat Temperatures*. When appropriate, setback of heating temperatures can be recommended during unoccupied periods. Similarly, setup of cooling temperatures can be considered.
- *Retrofit of Constant-Air-Volume Systems*. For commercial buildings, variable-air-volume (VAV) systems should be considered when the existing HVAC systems rely on constant volume fans to condition part or the entire building.
- *Retrofit of Central Heating Plants*. The efficiency of a boiler can be drastically improved by adjusting the fuel air ratio for proper combustion. In addition, installation of new energy-efficient boilers can be economically justified when old boilers are to be replaced.
- *Retrofit of Central Cooling Plants*. Currently, there are several chillers that are energy-efficient, easy to control and operate, and are suitable for retrofit projects. In general, it is cost-effective to recommend energy-efficient chillers such as those using scroll compressors for replacement of existing chillers.

- *Installation of Heat Recovery Systems*. Heat can be recovered from some HVAC equipment. For instance, heat exchangers can be installed to recover heat from air handling unit (AHU) exhaust air streams and from boiler stacks.

It should be noted that there is a strong interaction between various components of heating and cooling system. Therefore, a whole-system analysis approach should be followed when retrofitting a building HVAC system. Optimizing the energy-use of a central cooling plant (which may include chillers, pumps, and cooling towers) is one example of using a whole-system approach to reduce the energy-use for heating and cooling buildings.

16.5.5 Compressed-Air Systems

Compressed air has become an indispensable tool for most manufacturing facilities. Its uses span a range of instruments from air-powered hand tools and actuators to sophisticated pneumatic robotics. Unfortunately, staggering amounts of compressed air are currently wasted in a large number of facilities. It is estimated that only 20%–25% of input electrical energy is delivered as useful compressed-air energy. Leaks are reported to account for 10%–50% of the waste and misapplication accounts for 5–40% of loss in compressed air (Howe and Scales 1998).

The compressor can be selected from several types such as centrifugal, reciprocating, or rotary screw with one or multiple stages. For small and medium sized units, screw compressors are currently the most commonly used in the industrial applications. Table 16.7 provides typical pressure, airflow rate, and mechanical power requirement ranges for different types of compressors.

Some of the energy conservation measures that are suitable for compressed-air systems are listed below:

- Repair of air leaks in the distribution lines. Several methods do exist to detect these leaks ranging from the use of water and soap to the use of sophisticated equipment such as ultrasound leak detectors.
- Reduction of inlet air temperature and/or the increase of inlet air pressure.
- Reduction of the compressed-air usage and air pressure requirements by making some modifications to the processes.
- Installation of heat recovery systems to use the compression heat within the facility for either water heating or building space heating.
- Installation of automatic controls to optimize the operation of several compressors by reducing part load operations.
- Use of booster compressors to provide higher discharge pressures. Booster compressors can be more economical if the air with the highest pressure represents a small fraction of the total compressed air used in the facility. Without booster compressors, the primary compressor will have to compress the entire amount of air to the maximum desired pressure.

TABLE 16.7 Typical Ranges of Application for Various Types of Air Compressors

Compressor Type	Airflow Rate (m^3/s)	Absolute Pressure (MPa)	Mechanical power requirement (kW/L/s)
Reciprocating	0.0–5.0	0.340–275.9	0.35–0.39
Centrifugal	0.5–70.5	3.5–1034.3	0.46
Rotary screw	0.5–16.5	0.1–1.8	0.33–0.41

Source: From Herron, D. J., *Energy Engineering*, 96(2), 19, 1999.

16.5.6 Energy Management Controls

With the constant decrease in the cost of computer technology, automated control of a wide range of energy systems within commercial and industrial buildings is becoming increasingly popular and cost-effective. An energy management and control system (EMCS) can be designed to control and reduce the building energy consumption within a facility by continuously monitoring the energy-use of various equipments and making appropriate adjustments. For instance, an EMCS can automatically monitor and adjust indoor ambient temperatures, set fan speeds, open and close air handling unit dampers, and control lighting systems.

If an EMCS is already installed in the building, it is important to recommend a system tune-up to insure that the controls are properly operating. For instance, the sensors should be calibrated regularly in accordance with manufacturer specifications. Poorly calibrated sensors may cause increase in heating and cooling loads and may reduce occupant comfort.

Precooling building thermal mass is an example of the application of the EMCS to reduce operating costs. Precooling of the building thermal mass can be effective at lowering building operating costs. This strategy can have a large impact when chillers have high loads during periods of high occupancy and high outdoor temperatures (which typically coincide with on-peak periods in rate structures). By reducing the on-peak cooling load, it is possible to reduce chiller energy-use during these critical periods, thereby reducing energy costs.

Based on long-term simulation analysis, the annual energy cost savings associated with precooling has been estimated for various time-of-use utility rates (Morgan and Krarti 2005). For time-of-use rates, on-peak to off-peak ratios for energy and demand charges were defined as follows:

For the ratio of on-peak to off-peak energy charges, R_e:

$$R_e = \frac{(\text{PeakEnergyRate/kWh})}{\text{Off} - \text{Peak Energy Rate(/kWH)}.} \tag{16.2}$$

For the ratio of on-peak to off-peak demand charges, R_d:

$$R_d = \frac{(\text{PeakDemandRate/kW})}{\text{Off} - \text{Peak Demand Rate(/kW)}.} \tag{16.3}$$

Figure 16.7 and Figure 16.8 show the variation of the annual energy cost savings for a typical office building located in four U.S. locations due to a 4-h precooling period as a function of R_d and R_e, respectively. The office building has a heavy thermal mass of 105 lbm/ft.^{2} (513.7 kg/m^2) and the time-of-use rate has an 8-h on-peak period (Morgan and Krarti 2005).

16.5.7 Indoor Water Management

Water and energy savings can be achieved in buildings by using water-saving fixtures instead of the conventional fixtures for toilets, faucets, showerheads, dishwashers, and clothes washers. Savings can also be achieved by eliminating leaks in pipes and fixtures.

Table 16.8 provides typical water use of conventional and water-efficient fixtures for various end uses. In addition, Table 16.8 indicates the hot water use by each fixture as a fraction of the total water. With water-efficient fixtures, savings of 50% of water use can be achieved for toilets, showers, and faucets.

16.5.8 New Technologies

The energy auditor may consider the potential of implementing and integrating new technologies within the facility. It is therefore important that the energy auditor understands these new technologies and knows how to apply them. The following listing includes new technologies that can be considered for commercial and industrial buildings:

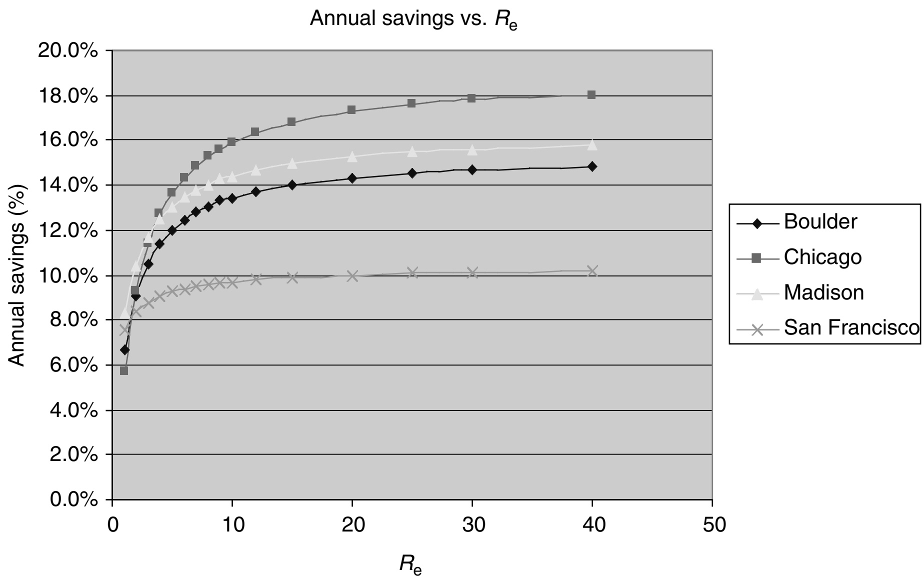


FIGURE 16.7 Annual energy cost savings due to precooling relative to conventional controls as a function of R_e. (From Morgan, S. and Krarti, M., *Impact of Electricity Rate Structures on the Energy Cost Savings of Precooling Controls for Office Buildings*, *Building and Environment* submitted, 2005.)

1. *Building-Envelope Technologies*. Recently several materials and systems have been proposed to improve the energy efficiency of building envelope and especially windows including:
   - Spectrally selective glasses that can optimize solar gains and shading effects
   - Chromogenic glazings that change their properties automatically depending on temperature and/or light level conditions (similar to sunglasses that become dark in sunlight)

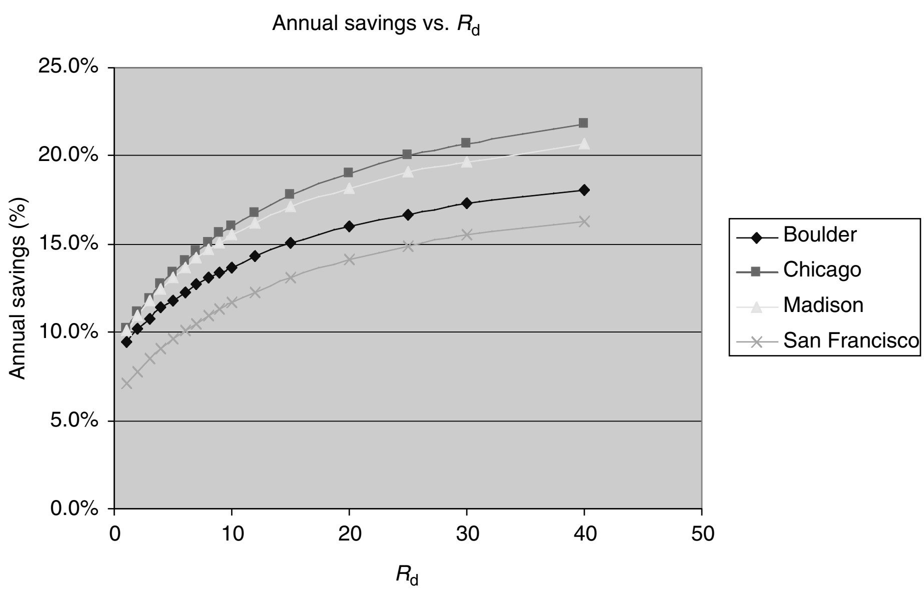


FIGURE 16.8 Annual energy cost savings due to precooling relative to conventional controls as a function of R_d. (From Morgan, S. and Krarti, M., *Impact of Electricity Rate Structures on the Energy Cost Savings of Precooling Controls for Office Buildings*, *Building and Environment* submitted, 2005.)

TABLE 16.8 Usage Characteristics of Water-Using Fixtures

End-Use	Conventional Fixtures	Water-Efficient Fixtures	Usage Pattern	% Hot Water
Toilets	3.5 gal/flush	1.6 gal/flush	4 flushes/pers/day	0%
Showers	5.0 gal/min	2.5 gal/min	5 min./shower	60%
Faucets	4.0 gal/min	2.0 gal/min	2.5 min/pers/day	50%
Dishwashers	14.0 gal/load	8.5 gal/load	0.17 loads/pers/day	100%
Clothes washers	55.0 gal/load	42.0 gal/load	0.3 loads/pers/day	25%
Leaks	10% of total use	2% of total use	N/A	50%

Source: From Krarti, M., *Energy Audit of Building Systems: An Engineering Approach*, CRC Press, Boca Raton, 2000.

- Building integrated photovoltaic panels that can generate electricity while absorbing solar radiation and reducing heat gain through building envelope (typically roofs)

2. *Light-Pipe Technologies*. Although the use of daylighting is straightforward for perimeter zones that are near windows, it is not usually feasible for interior spaces, particularly those without any skylights. Recent but still emerging technologies allow the "piping" of light from roof or wall-mounted collectors to interior spaces that are not close to windows or skylights.
3. *HVAC Systems and Controls*. Several strategies can be considered for energy retrofits including:
    - Thermal comfort controls can reduce energy consumption for heating or cooling buildings. Some HVAC control manufacturers have recognized the potential benefits from thermal comfort controls—rather than controls relying on only dry-bulb temperature—and have already developing and producing thermal comfort sensors. These sensors can be used to generate comfort indicators such as predicted mean vote (PMV) and/or predicted percent dissatisfied (PPD).
    - Heat recovery technologies, such rotary heat wheels and heat pipes, can recover 50–80% of the energy-used to heat or cool ventilation air supplied to the building.
    - Desiccant-based cooling systems are now available and can be used in buildings with large dehumidification loads during long periods (such as hospitals, swimming pools, and supermarket fresh produce areas).
    - Geothermal heat pumps can provide an opportunity to take advantage of the heat stored underground to condition building spaces.
    - Thermal energy storage (TES) systems offer a means of using less expensive off-peak power to produce cooling or heating to condition the building during on-peak periods. Several optimal control strategies have been developed in recent years to maximize the cost savings of using TES systems.
4. *Cogeneration*. This is not really a new technology. However, recent improvements in its combined thermal and electrical efficiency made cogeneration cost-effective in several applications, including institutional buildings such hospitals and universities.

16.6 Summary

In this chapter, simple yet proven analysis, procedures, and technologies have been described to improve energy efficiency for buildings and industrial facilities. If the energy management procedures are followed properly and if some cost-effective energy conservation measures—briefly described in this chapter—are implemented, it is expected that significant savings in energy-use and cost can be achieved. The efficient use of energy will continue to be vital to improve the environment and to increase the economic competitiveness.

References

Al-Mohimen, G. Hanna and Krarti, M. 2005. Analysis of daylighting benefits for office buildings in Egypt. *ASME Solar Energy Engineering*, 127, 366–370.

American Society of Heating Refrigeration and Air Conditioning Engineers (ASHRAE). 1989. *Ventilation for Acceptable Indoor Air Quality, Standard 62-1989*, ASHRAE, Atlanta, GA.

American Society of Heating Refrigeration and Air Conditioning Engineers (ASHRAE). 1997. *Proposed Guideline 14P, Measurement of Energy and Demand Savings*, ASHRAE, Atlanta, GA.

American Society of Heating Refrigeration and Air Conditioning Engineers (ASHRAE). 1999. *Handbook of HVAC Applications*, ASHRAE, Atlanta, GA.

Centre d'Etudes et de Recherches Economiques sur l, Energie (CEREN). 1997. *La Consommation d Energie Dans les Regions Francaises*. Report by the Centre d'Etudes et de la Recherche Economiques sur l'Energie, Paris, France.

Davidge, B. 1991. Demand-controlled ventilation systems in office buildings. *Proceedings of the 12th AIVC Conference Air Movement and Ventilation Control within Buildings*, pp. 157–171. Air Infiltration and Ventilation Center, Belgium.

Energy Information Administration (EIA). 1998. Annual energy review. Department of Energy. http://www.doe.eia.gov (accessed on 2005).

Emmerich, S. J. and Persily, A. K. 1997. Literature review on CO_2-based demand-controlled ventilation. *ASHRAE Transactions*, 103, (2), 139–148.

Federal Energy Management Program (FEMP). 1992. Energy policy act of 1992 becomes law. *FEMP Focus Special Edition*, Vol. 2, pp. 2–3.

Herron, D. J. 1999. Understanding the basics of compressed air systems. *Energy Engineering*, 96 (2), 14–17.

Howe, B. and Scales, B. 1998. Beyond leaks: deman-side Strategies for improving compressed air efficiency. *Energy Engineering*, 95, 31–35.

International Performance Monitoring and Verification Protocol (IPMVP), 1997. U.S. Department of Energy DOE/EE-0157, US Government Printing Office, Washington, DC.

Krarti, M. 2000. *Energy Audit of Building Systems: An Engineering Approach*. CRC Press, Boca Raton, FL.

Krarti, M., Ayari, A., and Grot, D. 1999. *Ventilation Requirements for Enclosed Vehicular Parking Garages*. Final Report for ASHRAE RP-945, Atlanta, GA.

Krarti, M., Erickson, P., and Hillman, T. 2005. A simplified method to estimate energy savings of artificial lighting use. *Daylighting Building and Environment*, 40, 747–754.

Morgan, S., and Krarti, M. 2006. Impact of electricity rate structures on the energy cost savings of pre-cooling controls for office buildings, Building and Environment In-Press.

National Association of Energy Services Companies (NAESCO). 1993. *Standard for Measurement of Energy Savings for Electric Utility Demand Side Management (DSM) Projects.* National Association for Energy Service Companies, Washington, DC.

Reddy, T. A., Kissock, K., Katipamula, S., Ruch, D. and Claridge., D. 1994. *An Overview of Measured Energy Retrofit Saving Methodologies Developed in the Texas LoanSTAR Program*. Energy Systems Laboratory Technical Report ESL-TR-94/03-04. Texas A&M University, College Station, TX.

Ross, M. H. and Williams, R. H. 1997. The Potential for fuel conservation. *Technological Review*, 79 (4), 49–53.

17

Cogeneration

W. Dan Turner

Texas A&M University

17.1 Introduction

17.1.1 Background

The term "cogeneration" as used in this chapter is defined as the combined production of electrical power[1] and useful thermal energy by the sequential use of a fuel or fuels. Each term in this definition is important. Combined means that the production processes of the electric power and thermal energy are linked, and often are accomplished in a series or parallel fashion. Electrical power is the electricity produced by an electrical generator, which is most often powered by a prime mover such as a steam turbine, gas turbine, or reciprocating engine. Thermal energy is that product of the process which provides heating or cooling. Forms of this thermal energy include hot exhaust gases, hot water, steam, and chilled water. Useful means that the energy is directed at fulfilling an existing need for heating or cooling. Simply exhausting hot exhaust gases does not meet the definition of useful thermal energy. In other words, cogeneration is the production of electrical power and the capture of coexisting thermal energy for useful purposes.

The importance of cogeneration is monetary and energy savings. Any facility that uses electrical power and needs thermal energy is a candidate for cogeneration. Although many considerations are involved in determining if cogeneration is feasible for a particular facility, the basic consideration is if the saving on thermal energy costs is sufficient to justify the capital expenditures for a cogeneration system. Facilities that may be considered for cogeneration include those in the industrial, commercial, and institutional sectors.

The technology for cogeneration is for the most part available and exists over a range of sizes: from less than 100 kW to over 100 MW. The major equipment requirements include a prime mover, electrical generator, electrical controls, heat recovery systems, and other typical power plant equipment. These components are well developed, and the procedures to integrate these components into cogeneration systems are well established.

In addition to the economic and technical considerations, the application of cogeneration systems involves an understanding of the governmental regulations and legislation on electrical power production and on environmental impacts. With respect to electrical power production, certain governmental regulations were passed during the late 1970s that remove barriers and provide incentives to encourage cogeneration development. Finally, no cogeneration assessment would be complete without an understanding of the financial arrangements that are possible.

17.1.2 Overall Scope of the Chapter

The objective of this chapter is to provide a complete overview of the important aspects necessary to understand cogeneration. Specifically, this chapter includes information on the technical equipment and components, technical design issues, regulatory considerations, economic evaluations, financial aspects, computer models and simulations, and future technologies. As briefly described earlier, a thorough discussion of cogeneration includes many aspects of engineering, economics, law, finance, and other topics. The emphasis of this chapter is on the technical and engineering considerations. The descriptions of the economic, legal, and governmental aspects are provided for completeness, but should not be considered a substitute for consulting with appropriate specialists such as attorneys, accountants, and bankers.

This chapter is divided into sections on basic cogeneration systems and terminology, technical components, design issues, regulatory considerations, economic evaluations, and financial aspects. Also, two case studies are presented that illustrate the application of cogeneration to an industrial and an institutional facility. Finally, the chapter concludes with some comments on the future

[1]Although the power output of a cogeneration system could be mechanical power as well as electrical power, the majority of systems produce electrical power. This chapter will consider only cogeneration systems that produce electrical power.

technologies of cogeneration. This introductory section ends with a brief review of the history of cogeneration.

17.1.3 History of Cogeneration

At the beginning of the twentieth century, electrical power generation was in its infancy. Most industrial facilities generated all their own electrical power and often supplied power to nearby communities. They used the thermal energy that was available during the electrical power production to provide or supplement process or building heat. These industrial facilities, therefore, were the first "cogenerators." The dominant prime mover at this time was the reciprocating steam engine, and the low-pressure exhaust steam was used for heating applications.

Between the early 1920s and through the 1970s, the public electric utility industry grew rapidly because of increasing electrical power demands. Coincident with this rapid growth was a general reduction in the costs to produce electrical power, mainly due to the economies of scale, more efficient technologies, and decreasing fuel costs. During this period, industry often abandoned their own electrical power generation because of (1) the decreasing electrical rates charged by public utilities, (2) income tax regulations that favored expenses instead of capital investments, (3) increasing costs of labor, and (4) the desire of industry to focus on their product rather than the side issue of electrical power generation. Estimates are available that suggest that industrial cogenerated electrical power decreased from about 25 to 9% of the total electrical power generated in the country between the years of 1954 and 1976. Since about the mid-1980s, this percentage has been fairly constant at about 5%. For example, at the end of 1992, 5.1% of the total U.S. electrical capacity was due to cogeneration systems.

During the 1960s and 1970s, the natural gas industry promoted a "total energy" concept of cogeneration. This effort was not very successful due to the relatively poor economics (e.g., relatively inexpensive electricity and expensive fuels), and the lack of governmental regulations to ease the interface with public utilities.

In late 1973 and again in 1979, America experienced major "energy crises" that were largely a result of reduced petroleum imports. Between 1973 and 1983, the prices of fuels and electrical power increased by a factor of approximately five. Any facility purchasing electrical power began to consider (or reconsider) the economic savings associated with cogeneration. These considerations were facilitated by Federal regulations that were enacted to ease or remove barriers to cogeneration.

In 1978, the government passed the National Energy Act (NEA) that included several important pieces of legislation. The NEA included the Fuel Use Act, the Natural Gas Policy Act, and the Public Utility Regulatory Policies Act (PURPA). Each of these acts had a direct impact on cogeneration, but PURPA was the most significant. In particular, PURPA defined cogeneration systems to include those power plants that supplied a specified fraction of their input energy as useful thermal output in addition to a mechanical or electrical output. This legislation is discussed in more detail in later sections of this chapter.

In addition, other regulatory legislation passed beginning in the late 1950s and continuing through the early 1990s impacts the installation of cogeneration systems. In particular, Federal legislation directed at managing air and water quality significantly affects the installations of cogeneration systems. For managing air pollution, the original legislation is the Air Quality Act of 1967 that was amended with the Clean Air Act Amendments in 1970, 1977, and 1990. The main legislative basis for managing water pollution is the Federal Water Pollution Control Act of 1956, as amended by the Water Quality Act of 1965, the Federal Water Pollution Control Act amendments of 1972, and Clean Water Act of 1977. These acts and other legislation, including the Energy Policy Act of 2005, and their impacts on the development of cogeneration projects, are discussed in a subsequent section of this chapter.

The term *cogeneration* lost favor in the U.S. in the 1980s and 1990s, particularly with the electric utilities. Instead, a new term, "combined heat and power," or CHP, became popular. A CHP Association was formed, and the U.S. Department of Energy began funding demonstration projects. A typical system might include a small combustion turbine producing electrical power used in conjunction with an absorption chiller that is direct-fired with the gas turbine exhaust gases. These smaller CHP systems are also being touted for use in buildings, called *building CHP* or *BCHP*. There has been an increased interest in distributed generation since parts of the U.S. and Canada suffered a massive electrical blackout in August 2003 that left eight states in the northeast and parts of Canada without power for several days.

The number of smaller CHP systems increased during the early part of the twenty-first century; however, relatively cheap electricity prices and moderate natural gas prices still hindered the widespread applications of CHP. Two natural disasters in the summer of 2005, Hurricanes Katrina and Rita, caused widespread damage to offshore natural gas wells in the Gulf of Mexico. The price of natural gas virtually doubled over a two-month period. This could have a profound impact on new cogeneration systems because, despite the increased system efficiency, the economics of most CHP systems depend on moderate gas prices (relative to electricity). The increased demand for natural gas, both in the U.S. and worldwide, and the price volatility make the near-term future of CHP highly uncertain. However, new technologies, transmission-grid problems forcing a greater reliance on distributed generation, and favorable electricity prices will make CHP systems more prominent in the twenty-first century. The overall CHP system efficiency will eventually drive the market to more widespread acceptance.

17.2 Basic Cogeneration Systems

17.2.1 Advantages of Cogeneration

To illustrate the fuel and monetary savings of a cogeneration system, the following simple comparison is presented. Figure 17.1 shows the details of this comparison. The power and thermal needs to be satisfied are 30 energy units[2] of electrical power and 34 energy units of heat. These power and thermal needs may be satisfied by either a cogeneration system or by conventional systems. Conventional systems for this example would be an electric utility and a standard boiler.

At the left of Figure 17.1 is the cogeneration system, and at the right of Figure 17.1 is the conventional system. All the components of the cogeneration system are represented by the box. Assuming an electrical generation efficiency of 30% for the cogeneration system, then 30 energy units are produced with 100 energy units of fuel energy. This cogeneration system then supplies 34 energy units of heat. These are conservative numbers for the electrical and thermal outputs of the cogeneration system, and many systems may have higher efficiencies. The overall thermodynamic efficiency[3] of the cogeneration system is defined as

$$\eta_0 = (P + T)/F \tag{17.1}$$

where P represents the power, T represents the thermal or heat energy rate, and F represents the fuel input rate (all in consistent units). This overall efficiency also may be found as the sum of the individual efficiencies of the electrical and thermal production. For this example, the cogeneration system has an overall efficiency of

$$\eta_0 = (30 + 34)/100 = 64\% \tag{17.2}$$

The conventional systems, diagrammed at the right of Figure 17.1, satisfy the power and thermal needs by the use of an electric utility and a boiler. In this case, the utility is assumed to be able to deliver the 30 energy units of electrical power with a plant efficiency of 35% (which is on the high side for most

[2]For this example, an energy unit is defined as any consistent unit of energy such as kW, MW, Btu, MBtu, or Btu/h. The use of arbitrary units avoids confusion and the need for unit conversions.

[3]This overall thermodynamic efficiency should not be confused with the "PURPA" efficiency, which is a legislated definition and is described in a later subsection.

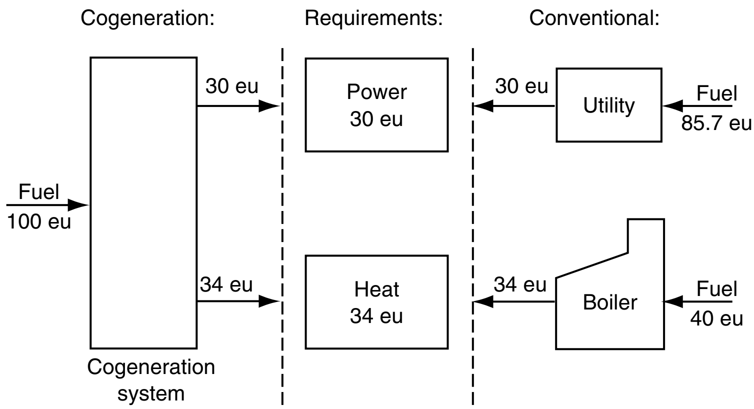


FIGURE 17.1 A schematic illustration of the advantages of cogeneration relative to conventional systems.

utilities). This results in a fuel input of 85.7 energy units. The boiler supplies the 34 energy units of heat with an 85% efficiency, which requires a fuel input of 40 energy units. The overall efficiency is, therefore,

$$\eta_0 = (P + T)/F = (30 + 34)/(85.7 + 40) = 51\% \tag{17.3}$$

This simple example demonstrates the thermodynamic advantage of a cogeneration system relative to conventional systems for accomplishing the same objectives. In this example, the cogeneration system had an overall efficiency of 64% compared to 51% for the conventional systems. Compared to the conventional systems, this is an absolute increase of 13% and a relative improvement of 25% (based on the 51% efficiency).

To determine the fuel and monetary savings of the cogeneration system relative to the conventional systems, the output power will be assumed to be 50 MW. This is a typical power level for a large university or hospital. The associated thermal output rate for this example would be

$$T = (34/30)50\ \text{MW} = 56.7\ \text{MW}. \tag{17.4}$$

This may be converted to units of MBtu/h (i.e., million Btu/h)

$$T = 56.7\ \text{MW}(3.41\ \text{MBtu/h/MW}) = 193.4\ \text{MBtu/h}. \tag{17.5}$$

To determine the fuel inputs, the power must be converted to consistent units

$$P = 50\ \text{MW}(3.41\ \text{MBtu/h/MW}) = 170.7\ \text{MBtu/h}. \tag{17.6}$$

Now, the fuel inputs are

$$(\text{Fuel})_{\text{cogen}} = \frac{(P + T)}{(\eta_0)_{\text{cogen}}} = \frac{(170.7 + 193.4)}{0.64} = 569\ \text{MBtu/h}. \tag{17.7}$$

$$(\text{Fuel})_{\text{conv}} = \frac{(P + T)}{(\eta_0)_{\text{conv}}} = \frac{(170.7 + 193.4)}{0.51} = 714\ \text{MBtu/h}. \tag{17.8}$$

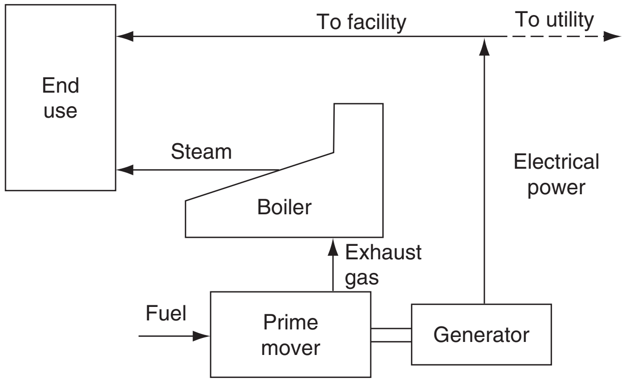


FIGURE 17.2 A schematic illustration of a cogeneration topping-cycle system.

The fuel savings is the difference between these two numbers:

$$(\text{Fuel savings}) = (\text{Fuel})_{\text{conv}} - (\text{Fuel})_{\text{cogen}} = 714 - 569 = 145 \text{ MBtu/h}. \tag{17.9}$$

If the plant operates 6,000 h/year (this is about 68% of the time on average—a conservative estimate), then the fuel energy savings per year is 870,000 MBtu. If the fuel price is \$8.00/MBtu, then the monetary saving is \$$6.96 \times 10^6$ per year. Although this is a simple economic analysis, the general trends are relevant. A later section of this chapter will describe more comprehensive economic analyses used in actual assessments of cogeneration applications.

In summary, the use of cogeneration is an effective way to more efficiently use fuels. The energy released from the fuels is used to both produce electrical power as well as provide useful thermal energy. The savings in energy and money can be substantial, and such systems have been shown to be technically and economically feasible in a wide range of applications.

17.2.2 Topping and Bottoming Cycles

A cogeneration system may be classified as either a topping-cycle system or a bottoming-cycle system. Figure 17.2 is a schematic illustration of a topping-cycle system. As shown, a prime mover uses fuel to power an electrical generator to produce electricity. This electricity may be used completely on-site or may be tied into an electrical distribution network for sale to the local utility or other customers. The hot exhaust gases are directed to a heat recovery boiler (HRB)[4] to produce steam or hot water. This steam or hot water is used on-site for process or building heat.

This cogeneration system is classified as a topping cycle because the electrical power is first generated at the higher (top) temperatures associated with the fuel combustion process and then the rejected or exhausted energy is used to produce useful thermal energy (such as the steam or hot water in this example). Figure 17.3 shows the thermodynamic states of the exhaust gases for this process on a temperature-entropy diagram. For this process, as energy is removed from the combustion gases, the temperature and entropy decrease (since the process involves energy removal). As shown, the top or high-temperature gases are used first to produce the electrical power and then the lower-temperature gases (exhaust) are used to produce useful thermal energy. The majority of cogeneration systems are based on topping cycles.

[4]Many other terms for this boiler are common: for example, waste heat boiler (WHB) and heat recovery steam generator (HRSG).

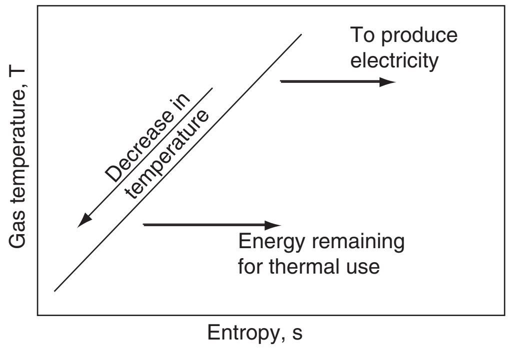


FIGURE 17.3 The gas temperature as a function of entropy for a topping-cycle system.

The other classification of cogeneration systems is bottoming cycle systems. Figure 17.4 is a schematic illustration of a bottoming cycle system. As shown, the high-temperature combustion gases are used first in a high-temperature thermal process (such as high-temperature metal treatment) and then the lower-temperature gases are used in a special low-temperature cycle to produce electrical power. Figure 17.5 shows the thermodynamic states of the exhaust gases for this process on a temperature-entropy diagram. After the energy is removed at the high temperatures, the energy available at the bottom or lower temperatures is then used to produce electrical power.[5]

Bottoming-cycle cogeneration systems have fewer applications than topping-cycle systems and must compete with waste heat recovery systems such as feedwater heaters, recuperators, and process heat exchangers. One of the difficulties with bottoming cycle systems is the low-temperature electrical power producing cycle. One example, depicted in Figure 17.4, is a low-temperature Rankine cycle. The low-temperature Rankine cycle is a power cycle similar to the conventional steam Rankine cycle, but a special fluid such as an organic substance (like a refrigerant) is used in place of water. This fluid vaporizes at a lower temperature than water, so this cycle is able to utilize the low-temperature energy. These cycles are generally much less efficient than conventional power cycles, often involve special equipment, and use more expensive working fluids. The majority of this chapter on cogeneration is based on topping cycle systems.

17.2.3 Combined Cycles

One power plant configuration that is based on a form of a topping cycle and is widely used in industry and by electrical utilities is known as a combined cycle. Typically in this configuration, a gas turbine is used to generate electricity and the exhaust gas is ducted to a heat recovery steam generator. The steam then is ducted to a steam turbine, which produces additional electricity. Such a combined cycle gas turbine power plant is often denoted as CCGT. In a cogeneration application, some steam would then need to be used to satisfy a thermal requirement. As might be expected, combined cycles have high power-to-heat ratios and high electrical efficiencies. Current designs have electrical efficiencies of up to 55% depending on the equipment, location, and details of the specific application. These current designs for combined cycle plants result in gas turbine power of between 1.5 and 3.5 times the power obtained from the steam turbine. These plants are most often base load systems operating more than 6000 hours per year. More details on gas turbines and steam turbines are provided in the following sections on the prime movers.

[5]Other definitions of bottoming cycles are common. These other definitions often include any second use of the energy. For example, some authors refer to the steam turbine in a combined cycle as using a bottoming cycle. The more precise thermodynamic definition employed here is preferred, although fewer applications meet this definition.

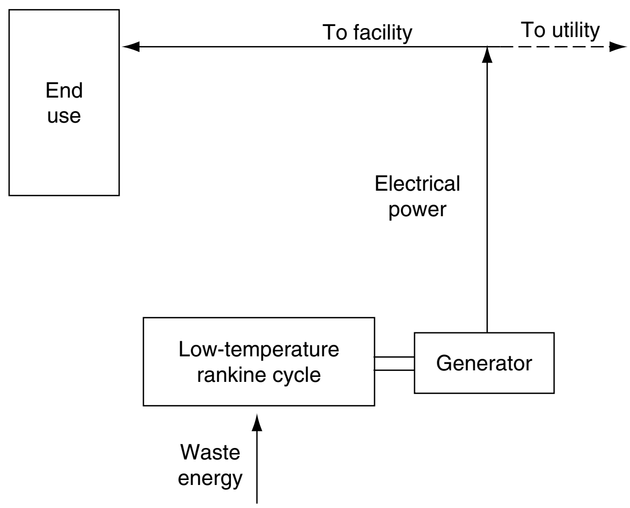


FIGURE 17.4 A schematic illustration of a cogeneration bottoming-cycle system.

17.2.4 Applications of Cogeneration Systems

17.2.4.1 General

Cogeneration systems may involve different types of equipment and may be designed to satisfy specific needs at individual sites. On the other hand, many sites have similar needs and packaged (pre-engineered) cogeneration systems may satisfy these needs and are more economical than custom-engineered systems. The following are examples of cogeneration systems in three different economic sectors.

Cogeneration systems are found in all economic sectors of the world. For convenience, cogeneration systems are often grouped into one of three sectors: (1) industrial, (2) institutional, and (3) commercial. The types and sizes of the cogeneration systems in these sectors overlap to varying degrees, but these sectors are nonetheless convenient for describing various applications of cogeneration. This section will provide examples of the variety of applications that exist. Obviously, these examples are not inclusive, but are intended to illustrate the breadth of possibilities.

17.2.4.2 Industrial Sector

Compared to the other economic sectors, the industrial sector includes the oldest, largest, and greatest number of cogeneration systems. As mentioned in the preceding history, cogeneration was first used by

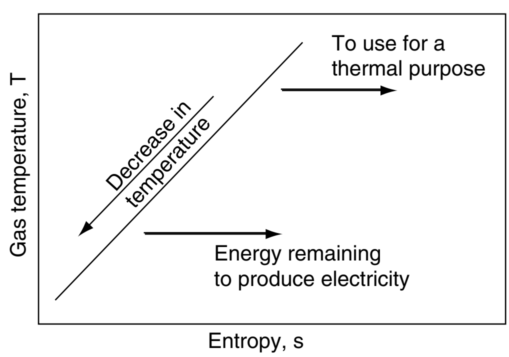


FIGURE 17.5 The gas temperature as a function of entropy for a bottoming-cycle system.

industry in the early 1900s to supply both electrical and thermal needs in an efficient manner. Many industries have a rich and continuous history of cogeneration applications. The industrial sector is dominant in cogeneration for several reasons. Industrial facilities often operate continuously, have simultaneous electrical and thermal requirements, and already have a power plant and operating staff. Those industries that are particularly energy intensive, such as the petrochemical and paper and pulp industries, are significant cogenerators. Many of these industries are cogenerating hundreds of megawatts of electrical power at one site. In addition, of course, mid-sized and smaller industries also use cogeneration.

17.2.4.3 Institutional Sector

The institutional sector includes a wide range of largely not-for-profit enterprises including universities, colleges, other schools, government building complexes, hospitals, military bases, and not-for-profit institutes. Many of these enterprises operate for a majority of the day, if not continuously. Some, such as hospitals, may already have provisions for emergency backup power and have the staff to operate a cogeneration system. Although not as large as the largest industrial cogenerators, a large hospital complex or university may require 50 or more megawatts of cogenerated electrical power.

Figure 17.6 is a schematic of a cogeneration system installed in 1986 at Rice University in Houston, Texas. As shown, this is a topping-cycle cogeneration system that uses a 3.2-MW gas turbine, a heat recovery steam generator, and two 1000-ton absorption chillers. Because this system was so successful, a second cogeneration system using a combined cycle was installed in 1989. It included a 3.7-MW gas turbine, a 400-kW steam turbine, and a 1500-ton absorption chiller. These two cogeneration systems supply the university with 90% of its electrical power, heating, and cooling requirements.

17.2.4.4 Commercial Sector

The commercial sector includes a wide range of for-profit enterprises including businesses, hotels, motels, apartment and housing complexes, restaurants, shopping centers, laundries, and laboratories. Generally, this sector includes the smallest cogenerators, and unless the electrical rates are unusually high, the economics are often less favorable than for the other sectors. There has been a greater interest in building cogeneration systems with the push for distributed generation and

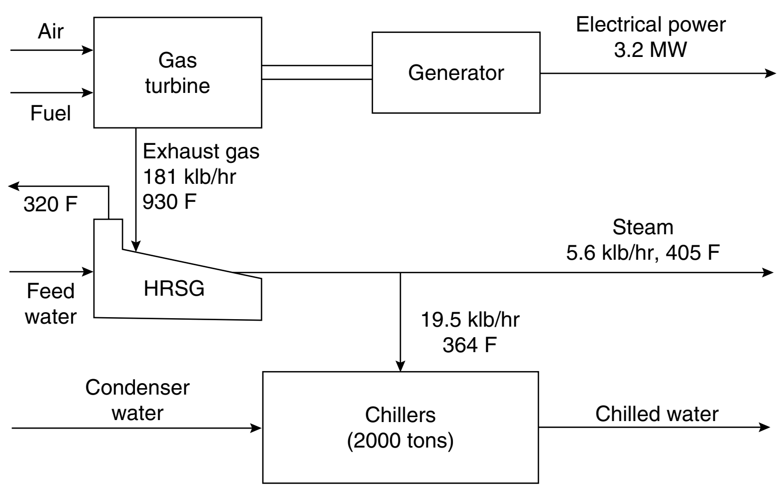


FIGURE 17.6 A schematic diagram of the topping-cycle cogeneration system installed in 1986 at Rice University in Houston, Texas.

increased efficiency of electrical generation. Fuel cells are also being used for cogeneration applications in buildings.

17.3 Equipment and Components

Cogeneration systems consist of several major pieces of equipment and many smaller components. This section will describe and provide guidance on the selection of this equipment and components. The following discussion is grouped into four subsections: (1) prime movers, (2) electrical equipment, (3) heat recovery devices, and (4) absorption chillers.

17.3.1 Prime Movers

Prime movers include those devices that convert fuel energy into rotating-shaft power to drive electrical generators. The prime movers that are used most often in cogeneration systems are steam turbines, gas turbines, and reciprocating engines. Each of these prime movers is described here. Important distinctions between the prime movers are the fuels that they may use, their combustion processes, their overall thermal efficiency, and the type, amount, and temperature of their rejected thermal energy. In cogeneration applications, a significant parameter for each type of prime mover is the ratio of the rate of supplied thermal energy and the output power. This ratio is called the heat-to-power ratio, and is unitless (i.e., kW/kW or Btu hr/Btu hr). Knowing the value of the heat-to-power ratio assists in matching a particular prime mover to a particular application. This matching is discussed in a later section.

17.3.1.1 Steam Turbines

Steam turbines are widely used in power plants throughout industry and electric utilities. Steam turbines use high-pressure, high-temperature steam from a boiler. The steam flows through the turbine, forcing the turbine to rotate. The steam exits the turbine at a lower pressure and temperature. A major difference of the steam turbine relative to the reciprocating engines and gas turbines is that the combustion occurs externally in a separate device (boiler). This allows a wide range of fuels to be used including solid fuels such as coal or solid waste materials. The exit steam, of course, can be used for thermal heating or to supply the energy to an absorption chiller.

Steam turbines are available in a multitude of configurations and sizes. This description will highlight only the major possibilities, but more complete information is available elsewhere (Wood 1982). A major distinction is whether the machine is a condensing or noncondensing (back-pressure) steam turbine. Condensing steam turbines are steam turbines designed so that the steam exits at a low pressure (less than atmospheric) such that the steam may be condensed in a condenser at near ambient temperatures. Condensing steam turbines provide the maximum electrical output, and hence, are most often used by central plants and electric utilities. Since the exiting steam possesses little available energy, the application of condensing steam turbines for cogeneration is negligible.

Noncondensing steam turbines are steam turbines designed such that the exiting steam is at a pressure above atmospheric. These steam turbines are also referred to as back-pressure steam turbines. The exiting steam possesses sufficient energy to provide process or building heat. Either type of steam turbine may be equipped with one or more extraction ports so that a portion of the steam may be extracted from the steam turbine at pressures between the inlet and exit pressures. This extracted steam may be used for higher temperature heating or process requirements.

Noncondensing steam turbines are available in a wide range of outputs beginning at about 50 kW and increasing to over 100 MW. Inlet steam pressures range from 150 to 2000 psig, and inlet temperatures range from 500 to 1,050°F. Depending on the specific design and application, the heat-to-power ratio for steam turbines could range from 4 to over 10. The thermal efficiency increases with size (or power level) from typically 8 to 20%. Although the major source of thermal energy is the exit or extracted steam, the boiler exhaust may be a possible secondary source of thermal energy in some cases.

17.3.1.2 Gas Turbines

As with steam turbines, stationary gas turbines are major components in many power plants. Stationary gas turbines share many of the same components with the familiar aircraft gas turbines. In fact, both stationary (or industrial) and aircraft gas turbines are used in cogeneration systems. This brief description will highlight the important characteristics of gas turbines as applied to cogeneration.

Many configurations, designs, and sizes of gas turbines are available. The simple-cycle gas turbine uses no external techniques such as regeneration to improve its efficiency. The thermal efficiency of simple-cycle gas turbines may therefore be increased by the use of several external techniques, but the designs and configurations become more complex. Many of these modifications to the simple-cycle gas turbine are directed at using the energy in the exhaust gases to increase the electrical output and efficiency. Of course, such modifications will decrease the available energy in the exhaust. For cogeneration applications, therefore, the most efficient gas turbine may not always be the appropriate choice.

Gas turbines are available in a wide range of outputs beginning at about 100 kW and increasing to over 100 MW. Depending on the specific design, the heat-to-power ratio for gas turbines could range from about 1 to 3. The design point thermal efficiency increases with size (or power level) and complexity from typically 20 to 45%. The higher thermal efficiencies, compared to steam turbines, is the reason the heat-to-power ratio is lower than for steam turbines. These high efficiencies are for full-load (design point) operation. At part load, a gas turbine's efficiency decreases rapidly. As mentioned earlier, the use of regeneration, intercooling, reheating, and other modifications are used to improve the overall performance of the simple-cycle gas turbine.

Due to the large amount of excess air (the total air mass may be on the order of 100 times the fuel mass) used in the combustion process of gas turbines, the exiting exhaust gas contains a relatively high concentration of nitrogen and oxygen. Hence, the gas turbine exhaust may be characterized as mostly heated air and is nearly ideal for process or heating purposes. Gas turbines may use liquid fuels such as jet fuel or kerosene, or they may use gaseous fuels such as natural gas or propane. The highest performance is possible with liquid fuels, but the lowest emissions have been reported for natural gas operation.

Power ratings for gas turbines are provided for continuous and intermittent duty cycles. The continuous ratings are slightly lower than the intermittent ratings to provide long life and good durability. In both cases, the power ratings are provided for a set of standard operating conditions known as ISO conditions. ISO conditions include 1 atm, 59°F (15°C), sea level, and no inlet or exhaust pressure losses. For a specific application, the local conditions must be used to adjust the rated power to reflect the actual operating conditions. Most manufacturers will provide potential users with their recommended adjustments, which depend on the difference between the local and standard conditions. In particular, the local ambient inlet air temperature affects the gas turbine output in a significant manner. As the ambient air temperature increases, the performance of a gas turbine decreases due to the lower air density.

Gas turbines require more frequent and specialized maintenance than steam turbines. Major overhauls are required at 20,000–75,000 hours of operation, depending on the service duty and the manufacturer. Aircraft gas turbines are designed to have their "hot section" returned to the manufacturer and replaced with a conditioned unit to minimize downtime, although the overhaul is more expensive. Industrial gas turbines are designed to be overhauled in the field, and this is generally a less expensive procedure.

17.3.1.3 Reciprocating Engines

A third category of prime movers for cogeneration systems is internal combustion (IC), reciprocating engines. Although rotary engines could be used in cogeneration systems, no significant applications are known. The remaining discussion will focus on reciprocating engines. These engines are available in several forms. Probably the most common form of the reciprocating engine is the typical spark-ignited, gasoline engine used in automobiles. For cogeneration applications, the spark-ignited gasoline engine has been converted to operate in a stationary, continuous mode with fuels such as natural gas.

The majority of reciprocating engines for mid- to large-sized cogeneration systems are stationary diesel engines operating either with diesel fuel or in a dual-fuel mode with natural gas. These engines share some common characteristics for cogeneration applications and have some distinctive features as well.

Power ratings for reciprocating engines are similar to those for gas turbines in that both continuous and intermittent duty cycle ratings are provided. As with the gas turbines, these power ratings are provided for a set of standard conditions for ambient temperature, pressure, and elevation. The standard power ratings need to be adjusted for the local conditions at the site of the installation. For cogeneration applications, reciprocating engines are available in many power levels and designs. These power levels range from less than 50 kW to over 200 MW. Some manufacturers even offer "mini" cogeneration systems with outputs as low as 6 kW.

For those reciprocating, internal combustion engines that are liquid cooled (the majority of the engines considered here), the cooling liquid is a secondary source of thermal energy. Although not at the high temperatures of exhaust gas, this energy can be used to produce hot water or low-pressure steam. Several designs are available for recovering the energy in the cooling liquid. These designs use one or more direct or indirect heat exchangers to generate the hot water or low-pressure steam. Liquid-to-liquid heat exchangers can have high efficiencies, and most of this energy is recoverable (but at relatively low temperatures). Other energy sources from reciprocating engines include oil-coolers and turbocharger after-coolers. This energy is usually at temperatures below 160°F, and would only be practical to recover for low-temperature requirements. Another benefit of the reciprocating engine is that the maintenance and repair is less specialized than for gas turbines. On the other hand, the maintenance may be more frequent and more costly.

17.3.1.4 Heat-to-Power Characteristics

As mentioned earlier, an important feature in the selection of a prime mover for a cogeneration system is the heat-to-power ratio. This ratio may be a fairly constant characteristic of a family of a particular type of prime mover. Figure 17.7 and Figure17.8 show the electrical power output as a function of the potential available thermal energy rate output for a specific family of gas turbines and a specific family of reciprocating engines, respectively. The symbols represent the power and thermal output for individual gas turbines and engines in the two "families." The solid lines are the linear best fits to this data and represent the average heat-to-power (H/P) ratio for each family. As shown, the data does scatter some

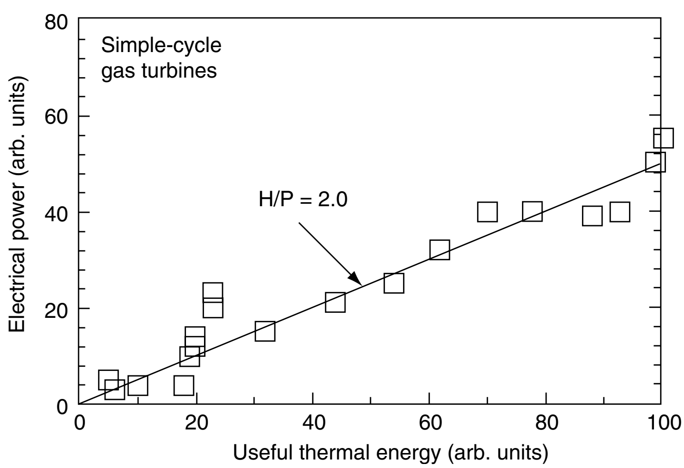


FIGURE 17.7 The electrical power as a function of the useful thermal energy (in arbitrary units) for a number of simple-cycle gas turbines. The average heat-to-power ratio for this family of gas turbines is 2.0, and is shown as the solid line.

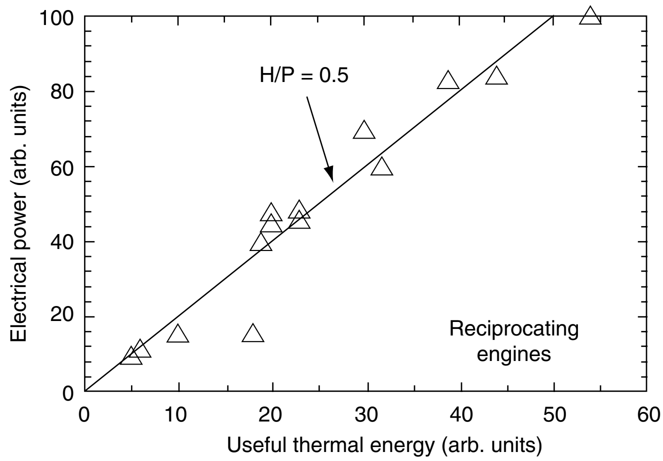


FIGURE 17.8 The electrical power as a function of the useful thermal energy (in arbitrary units) for similar diesel engines. The average heat-to-power ratio for this family of reciprocating engines is 0.5, and is shown as the solid line.

about the linear line, but the characteristic for each prime mover family is a good average. Also as shown, the average heat-to-power ratio is 2.0 for the gas turbine and 0.5 for the reciprocating engine. These are representative values for these two types of prime movers. These characteristic heat-to-power ratios will be used here in matching the prime mover to a specific application.

17.3.2 Electrical Equipment

The electrical equipment for cogeneration systems includes electrical generators, transformers, switching components, circuit breakers, relays, electric meters, controls, transmission lines, and related equipment. In addition to the equipment that supports electrical production, cogeneration systems may need equipment to interconnect with an electric utility to operate in parallel for the use of backup (emergency) power or for electrical sales to the utility. This section will briefly highlight some of the aspects regarding electrical generators and interconnections, but will not be able to review the other important electrical considerations.

The electric generator is a device for converting the rotating mechanical energy of a prime mover to electrical energy (i.e., electricity). The basic principle for this process, known as the Faraday effect, is that when an electrically conductive material such as a wire moves across a magnetic field, an electric current is produced in the wire. This can be accomplished in a variety of ways, so there are several types of electric generators. The frequency of the generator's output depends on the rotational speed of the assembly.

An important feature of generators is that they require a magnetic field to operate. The source of the energy for this magnetic field serves to distinguish the two major types of generators. If the generator is connected to an electric source and uses that source for the magnetic field, it is called an induction generator. In this case, the generator operates above the synchronous speed, and cannot operate if the external current (usually from the electric utility) is not available. On the other hand, if the magnetic field is generated internally using a small alternator, the generator is called a synchronous generator and operates at the synchronous speed. Synchronous generators can operate independent of the external electric grid.

Most often the manufacturer of the prime mover will provide the prime mover and generator as an integrated, packaged assembly (called a genset). Performance characteristics of generators include power rating, efficiency, voltage, power factor, and current ratings. Each of these performance characteristics must be considered when selecting the proper generator for a given application. Electric generators may have conversion efficiencies of between about 50 to over 98%. The efficiency increases in this range as the generator increases in size and power rating. Only the largest electric generators (say, on the order of 100 MW) can attain efficiencies of 98%.

In many cases, the cogeneration system will need to be interconnected to the utility. Although a cogeneration system could be isolated from the electric grid, there are several reasons why the system may need to be interconnected. An interconnection is necessary for receiving supplementary or backup (emergency) electric power. Also, if the cogenerator elects to sell excess power, such an interconnection would be needed. The interconnection equipment includes relays, circuit breakers, fuses, switches, transformers, synchronizers, and meters. The specific equipment and design for the interconnection largely is dictated by the utility for safety and compatibility reasons. The utility has specific responsibilities for the integrity of the electric grid and for maintaining the electrical power quality.

17.3.3 Heat Recovery Equipment

The primary heat recovery equipment used in cogeneration systems includes several types of steam and hot-water production facilities. In addition, absorption chillers could be considered in this section, but for organizational reasons, chillers will be discussed in the following section. Several configurations of heat recovery devices are available. As mentioned earlier, these devices may be referred to as "heat recovery steam generators" or HRSGs. HRSGs are often divided into the following categories: (1) unfired, (2) partially fired, and (3) fully fired. An unfired HRSG is essentially a convective heat exchanger. A partially fired HRSG may include a "duct burner," which often uses a natural gas burner upstream of the HRSG to increase the exhaust gas temperature. A fully fired HRSG is basically a boiler that simply uses the exhaust gas as preheated air.

For most of these heat recovery devices, exhaust gas flows up through the device and exits at the top. Energy from the exhaust gas is used to heat and vaporize the water, and to superheat the steam. Figure 17.9 shows this process on a temperature diagram. The top line shows the exhaust gas temperature decreasing from left to right as energy is removed from the gas to heat the water. The lower line represents the water heating up from right to left in the diagram. The lower-temperature exhaust is used to preheat the water to saturation conditions in the economizer. The intermediate-temperature exhaust is used to vaporize (or boil) the water to form saturated steam. Finally, the highest-temperature exhaust is used to superheat the steam.

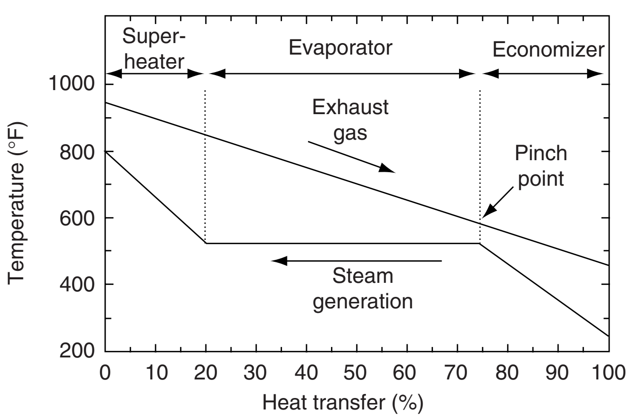


FIGURE 17.9 The temperatures of the exhaust gas and water/steam as a function of the heat transfer coordinate for the steam generation process of an HRSG.

The temperature difference between the exhaust gas and the water where the water first starts to vaporize is referred to as the pinch point temperature difference. This is the smallest temperature difference in the HRSG and may limit the overall performance of the heat recovery device. Since the rate of heat transfer is proportional to the temperature difference, the greater this difference the greater the heat transfer rate. On the other hand, as this temperature difference increases the steam flow rate must decrease and less of the exhaust gas energy will be utilized. To use smaller temperature differences and maintain higher heat transfer rates, larger heat exchanger surfaces are required. Larger heat transfer surface areas result in higher capital costs. These, then, are the types of trade-offs that must be decided when incorporating a heat recovery device into a cogeneration system design.

The proper selection of HRSG equipment depends on the prime mover, the required steam conditions, and other interdependent factors. Some of these considerations are described next. Most of these considerations involve a trade-off between increased performance and higher initial capital cost.

The "back-pressure" of the HRSG unit affects the overall system performance. As the back-pressure decreases, the HRSG efficiency increases, but the cost of the HRSG unit increases. Successful units often have 10–15 in. water pressure of back-pressure. As mentioned earlier, the selection of the "pinch point" temperature difference affects the performance and cost. Typical pinch point temperature differences are between 30 and 80°F. High-efficiency, high-cost units may use a temperature difference as low as 25°F.

The final outlet stack gas temperature must be selected so as to avoid significant acid formation. This requires the outlet stack temperature to be above the water condensation temperature. This is typically above 300°F. The amount of sulfur in the fuel affects this decision, since the sulfur forms sulfuric acid. As the amount of sulfur in the fuel increases, the recommended stack gas outlet temperature increases.

The final consideration is the steam temperature and pressure. This is a complex decision that depends on many factors such as the application for the steam, the source of the exhaust gas, the exhaust gas temperature and flow rate, and inlet water condition and temperature.

17.3.4 Absorption Chillers

Absorption chillers may use the thermal energy from cogeneration systems to provide cooling for a facility. This section will briefly review the operation of absorption chillers and their application to cogeneration systems. Absorption chillers use special fluids and a unique thermodynamic cycle that produces low temperatures without the requirement of a vapor compressor, which is used in mechanical chillers. Instead of the vapor compressor, an absorption chiller uses liquid pumps and energy from low-temperature sources such as hot water, steam, or exhaust gas.

Absorption chillers utilize fluids that are solutions of two components. The basic principle of the operation of absorption chillers is that once the solution is pumped to a high pressure, low-temperature energy is used to vaporize one component from the solution. This component serves as the "refrigerant" for this cycle. Examples of these solutions are (1) water and ammonia, (2) lithium bromide and water, and (3) lithium chloride and water. In the first case the ammonia serves as the refrigerant, and in the latter two cases the water serves as the refrigerant.

For cogeneration applications, the important feature of absorption chillers is that they use relatively low-temperature energy available directly or indirectly from the prime mover and produce chilled water for cooling. The use of absorption chillers is particularly advantageous for locations where space and water heating loads are minimal during a good part of the year. For these situations, the thermal output of a cogeneration system can be used for heating during the colder part of the year and, using an absorption chiller, for cooling during the warmer part of the year. Furthermore, by not using electric chillers, the electric loads are more constant throughout the year. In warm climates, absorption chillers are often an important, if not an essential, aspect of technically and economically successful cogeneration systems.

Some machines are designed as indirect-fired units using hot water or steam. As examples of typical numbers, a single-stage unit could use steam at 250°F to produce a ton of cooling for every 18 pounds of steam flow per hour. A dual-stage unit would need 365°F steam to produce a ton of cooling for every

10 pounds of steam flow per hour. If hot water is available, a ton of cooling could be produced for every 220 pounds of 190°F hot water per hour.

Other machines use the exhaust gas directly and are called direct-fired units. In these cases, the exhaust gas temperature needs to be 550 to 1000°F. The higher the exhaust temperature, the less energy (or exhaust gas flow) is needed per ton of cooling. For example, for 1000°F exhaust gas, a ton of cooling requires 77 pounds per hour of flow whereas for 550°F exhaust gas a ton of cooling requires 313 pounds per hour of flow.

17.4 Technical Design Issues

17.4.1 Selecting and Sizing the Prime Mover

The selection of a prime mover for a cogeneration system involves the consideration of a variety of technical and nontechnical issues. Technical issues, which often dominate the selection process, include the operating mode or modes of the facility, the required heat-to-power ratio of the facility, the overall power level, and any special site considerations (e.g., low noise). Other issues, which may play a role in the selection process, include matching existing equipment and utilizing the skills of existing plant personnel. Of course, the final decision is often dominated by the economics.

Steam turbines and boilers usually are selected for a cogeneration system if the fuel of choice is coal or another solid fuel. Occasionally, for very large ($>$50 MW) systems operating at base load, a steam turbine system may be selected even for a liquid or gaseous fuel. Also, steam turbines and boilers would be selected if a high heat-to-power ratio is needed. Steam turbines also may be selected for a cogeneration system in certain specialized cases. For example, a large pressure reduction valve in an existing steam system could be replaced with a steam turbine and thereby provide electrical power and thermal energy. In many applications, however, steam turbines are selected to be used in conjunction with a gas turbine in a combined cycle power plant to increase the power output. Combined cycle gas turbine power plants were described in an earlier section.

Gas turbines are selected for many cogeneration systems where the required heat-to-power ratio and the electrical power need are high. Also, gas turbines are the prime mover of choice where minimal vibration or a low weight-to-power ratio (such as for a roof installation) is required. Reciprocating engines are selected where the heat-to-power ratio is modest, the temperature level of the thermal energy is low, and the highest electrical efficiency is necessary for the economics. Additionally, reciprocating engines may be selected if the plant personnel are more suited to the operation and maintenance of these engines.

Selecting the appropriate size prime mover involves identifying the most economic cogeneration operating mode. This is accomplished by first obtaining the electrical and thermal energy requirements of the facility. Next, various operating modes are considered to satisfy these loads. By conducting a comprehensive economic analysis, the most economic operating mode and prime mover size can be identified. The process of matching the prime mover and the loads is described next.

17.4.2 Matching Electrical and Thermal Loads

To properly select the size and operating mode of the prime mover, the electric and thermal loads of the facility need to be obtained. For the most thorough "matching," these loads are needed on an hourly, daily, monthly, and yearly basis. Figure 17.10 is an example of the hourly electrical and thermal loads of a hypothetical facility for a typical work day in the winter. As shown, the heating and power demands begin to increase at 6:00 a.m. as the day's activities begin. The heating load is shown to peak shortly after 8:00 a.m. and then decrease for the remainder of the day. The electrical load remains nearly constant during the working part of the day, and is somewhat lower during the evening and nighttime hours.

Figure 17.11 shows the month totals for the electrical and thermal loads for the same hypothetical facility. For the summer months, the heating and electrical loads are minimum. In addition, this figure

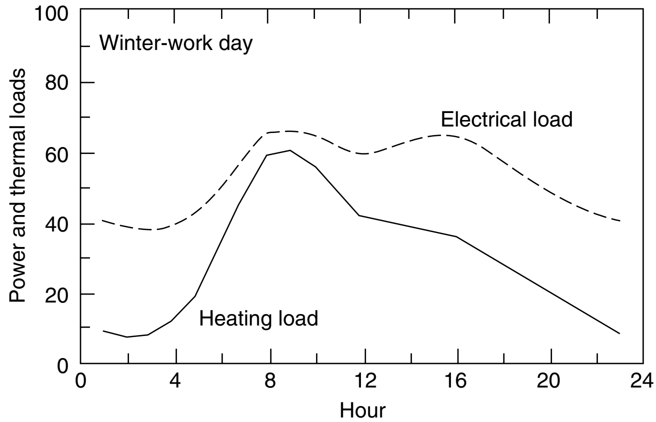


FIGURE 17.10 The power and thermal loads (in arbitrary units) as a function of the hour of the day for a hypothetical facility for one work day in the winter.

shows dotted and dashed lines, which represent the "base loads" for the electrical and heating loads, respectively. The base loads are the minimum loads during the year and form a floor or base for the total loads. Often a cogeneration system may be sized so as to provide only the base loads. In this case, auxiliary boilers would provide the additional heating needed during the days where the heating needs exceeded the base amount. Similarly, electrical power would need to be purchased to supplement the base power provided by the cogeneration system.

Several options exist in matching the facility's electrical power and thermal needs to a prime mover for a cogeneration system. To illustrate this matching procedure, consider Figure 17.12, which shows electrical power as a function of thermal energy. The three dashed lines are the heat-to-power characteristics of three families of prime movers. In this example, the heat-to-power ratio increases from A to B to C. The facility's needed electrical power and thermal energy are shown as the horizontal and vertical dotted lines, respectively. This needed power and thermal energy is often the base load of the facility as explained earlier.

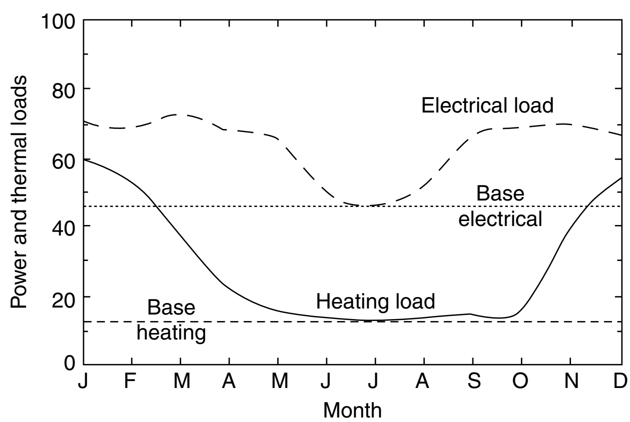


FIGURE 17.11 The total power and thermal loads (in arbitrary units) as a function of the month of the year for a hypothetical facility.

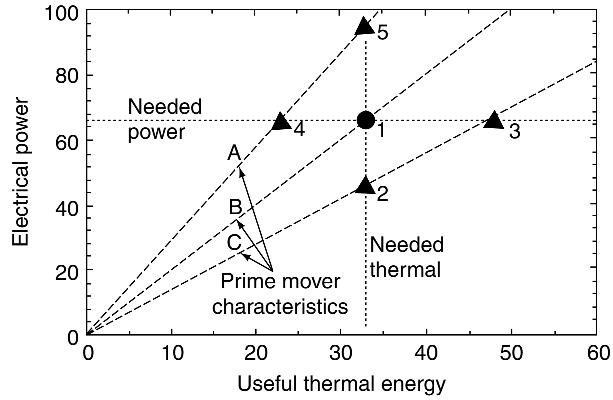


FIGURE 17.12 The electrical power as a function of the useful thermal energy (in arbitrary units).

Five operating points are identified:

1. Point 1 represents an exact match between the facility's needed power and thermal energy, and the power and thermal energy available from a prime mover with the heat-to-power ratio characteristic B. Such an exact match is not likely or desired, so the other operating points must be considered.
2. Point 2 represents a case where a prime mover is selected with a characteristic C such that the thermal needs are matched but the electrical power supplied is too low. In this case, the facility would need to purchase supplementary electrical power usually from the local electrical utility. This is a common case that is often the best economic and operational choice.
3. Point 3 represents a case where a prime mover is selected with a characteristic C such that the electrical power needs are matched, but excess thermal energy is available. Generally, this is not an economical choice because the overall system efficiency is low. For certain situations, however, this case might be selected—for example, if future thermal loads are expected to increase or if a customer can be identified to purchase the excess thermal energy.
4. Point 4 represents a case where a prime mover is selected with a characteristic A such that the electrical power needs are matched but the available thermal energy does not satisfy the needs of the facility. In this case, supplementary boilers are often used to supply the additional thermal energy. This may be the best economic choice, particularly if existing boilers are available and the desire is to size the cogeneration system such that all the thermal energy is used.
5. Point 5 represents a case where a prime mover is selected with a characteristic A such that the thermal needs are matched but the electrical power supplied is too high. This would only be economical if an external customer was available to purchase the excess electrical power. This is often the situation with third-party cogenerators, who own and operate a cogeneration system and sell the thermal energy and electrical power to one or more customers.

These cases represent distinct operating mode selections. In practice, other technical and nontechnical aspects must be considered in the final selection. These cases are presented to illustrate the global decision process.

The possible overall operating modes for a cogeneration power plant are often categorized into one of three classes. (1) The plant may operate as a base load system with little or no variation in power output. Base load plants operate in excess of 6,000 hours per year. Power needs above the base load are typically provided by interconnections to a local utility or by an auxiliary power plant. (2) The plant may operate as an intermediate system for 3,000–4,000 hours per year. These systems are less likely than base load systems, but if the economics are positive, may have application for facilities that are not continuously operated such as some commercial enterprises. (3) Finally, a third class of plant is a peaking system, which operates only for 1,000 hours or less per year. Utility plants often use peaking systems to provide peaking power during periods of high electrical use. For cogeneration applications, peaking units may be economical where the cost of the electricity above a certain level is unusually high. These units are sometimes referred to as peak shaving systems.

17.4.3 Packaged Systems

In general, facilities with low electrical power needs cannot utilize customized cogeneration systems because of the relatively high initial costs that are associated with any size system. These initial costs include at least a portion of the costs related to the initial design, engineering, and related development and installation matters. Also, smaller facilities often do not have the specialized staff available to develop and operate complex power plants.

To solve some of these above problems, pre-engineered, factory-assembled, "packaged" cogeneration systems have been developed. Packaged cogeneration systems range from less than 50 kW to over 1 MW. Some larger units (say, 20 MW or larger) using gas turbines or reciprocating engines are sometimes described as packaged cogeneration systems, and are, at least to some degree, "packaged." For example, a gas turbine and generator may be factory assembled, tested, and skid mounted for shipping. These units, although packaged, do not completely eliminate the problems of project development, installation, and operation. The remainder of this section will consider only the smaller packaged cogeneration systems.

The major advantage of packaged cogeneration systems is that the initial engineering, design, and development costs can be spread over many units, which reduces the capital cost (per kW) for these systems. Other advantages of packaged cogeneration systems include factory assembly and testing of the complete system. If there are any problems, they can be fixed while the system is still at the manufacturer's plant. The standard design and reduced installation time result in short overall implementation times. In some cases, a packaged cogeneration system could be operational within a few months after the order is received. This short implementation time reduces the project's uncertainty, which eases making decisions and securing financing.

Another advantage of packaged cogeneration systems is that the customer only interacts with one manufacturer. In some cases, the packaged cogeneration system manufacturer will serve as project engineer and take the project from initial design to installation to operation. The customer often may decide to purchase a Turnkey system. This provides the customer with little uncertainty and places the burden of successful project completion on the manufacturer. Also, the manufacturer of a packaged cogeneration system will have experience interacting with regulating boards, financing concerns, and utilities, and may assist the customer in these interactions.

The major disadvantage of packaged cogeneration systems is that the system is not customized for a specific facility. This may mean some compromise and lack of complete optimization. Specialized configurations may not be available. Also, beyond a certain size, packaged cogeneration systems are simply not offered and a customized unit is the only alternative.

Although the initial capital costs of packaged cogeneration systems are low (on a per kW basis), the other site-specific costs could be relatively high and result in unsatisfactory economics. These other costs depend on the selected fuel, the available space at the site, and the compatibility of existing electrical and mechanical facilities. In particular, the interconnection costs associated with the electric utility and the mechanical systems can be prohibitively high. The electric interconnection costs depend on the electric utility's requirements. The fuel system costs depend on whether the selected fuel is already available on site at the appropriate conditions (such as gas pressure for gaseous fuels). Other miscellaneous costs include those for services and assistance with engineering, permitting, zoning, financing, legal review, and construction oversight.

In summary, packaged cogeneration systems offer a cost-effective solution for facilities with low power and thermal requirements such as small to medium hospitals, schools, hotels, and restaurants. The majority of current packaged cogeneration systems use reciprocating engines and standard components. They are factory assembled and tested, and are often sold as Turnkey installations. Although the packaged cogeneration systems' costs can be attractive, each facility must be assessed for total costs to determine the economic feasibility. Future developments of packaged cogeneration systems will include even more efficient and cost-effective engines, more advanced microprocessor control systems, and more effective integration of the various components.

17.5 Regulatory Considerations

This section includes brief overviews of the relevant Federal regulations on electric power generation and the related environmental constraints. It should be pointed out that some sections of PURPA were repealed by the Energy Policy Act of 2005; however, its importance to cogeneration is so significant that it deserves the recognition and discussion as noted below.

17.5.1 Federal Regulations Related to Cogeneration—Early History

As mentioned earlier, the passage of PURPA, the Public Utility Regulatory Policies Act, in 1978, helped make cogeneration a much more attractive option for electric power generation for a variety of facilities. PURPA is one of the most controversial bills relating to power generation that has been passed by Congress. PURPA was one of five separate acts comprising the National Energy Act passed by Congress in 1978:

1. National Energy Conservation Policy Act (NECPA)
2. Natural Gas Policy Act (NGPA)
3. Powerplant and Industrial Fuel Use Act (FUA)
4. Energy Tax Act (ETA)
5. Public Utility Regulatory Policies Act (PURPA)

The Federal government, and for that matter the entire nation, was concerned about the oil crisis in 1973, the rapidly increasing price of energy, the increased dependence on foreign oil, and our lack of concern for energy efficiency and the use of renewable energy. The National Energy Act addressed most of those concerns, and PURPA, in particular, dealt with independent power generation, energy efficiency, and the use of renewables for power generation. Although all five acts had an impact on our nation's energy consumption and energy use patterns, PURPA had the greatest overall impact and led to a resurgence of interest in cogeneration. Prior to the passage of PURPA, the three most common barriers to cogeneration were (Laderoute 1980) (1) no general requirement by electric utilities to purchase electric power from cogenerators, (2) discriminatory backup power for cogenerators, and (3) fear on the part of the cogenerator that they might become subject to the same state and Federal regulations as electric utilities. The passage of PURPA helped to remove these obstacles to development of cogeneration facilities (Spiewak 1980; Spiewak 1987).

Although PURPA was passed in 1978, it was not until 1980 that the Federal Energy Regulatory Commission (FERC) issued its final rulemakings and orders on PURPA. In the National Energy Act, FERC was designated as the regulatory agency for implementation of PURPA. The regulations dealing with PURPA are contained in Part 292 of the FERC regulations, and Sections 201 and 210 are the two primary sections relevant to small power production and cogeneration (FERC 1978).

TABLE 17.1 Required Efficiency Standards for Qualified Facilities (QFs)

If the Useful Thermal Energy Fraction is	The Required η_{PURPA} Must be
$\geq 5.0\%$	$\geq 45.0\%$
$\geq 15.0\%$	$\geq 42.5\%$

Section 201 contains definitions of cogeneration and sets annual efficiency standards for new topping cycle[6] cogeneration facilities that use oil or natural gas.[7] For a cogenerating facility to qualify for the privileges and exclusions specified in PURPA, the facility must meet these legislated standards. These standards define a legislated or artificial “efficiency,” which facilities must equal or exceed to be considered a “qualified facility” (QF). A qualified facility is eligible to use the provisions outlined in PURPA regarding nonutility electric power generation. This legislated “efficiency” is defined as

$$\eta_{\mathrm{PURPA}} = \left(P + \frac{1}{2}T\right)/F \qquad (17.10)$$

where P is the electrical energy output, T is the useful thermal energy, and F is the fuel energy used. Table 17.1 lists the standards for the PURPA efficiencies. These standards state that a facility must produce at least 5% of the site energy in the form of useful thermal energy, and must meet the efficiency standards in Table 17.1. Values for the thermal fraction and the PURPA efficiency are based upon projected or estimated annual operations.

The purpose of introducing the “artificial” standards was to insure that useful thermal energy was produced on-site in sufficient quantities to make the cogenerator more efficient than the electric utility. Any facility that meets or exceeds the required efficiencies will be more efficient than any combination of techniques producing electrical power and thermal energy separately. Section 201 also put limitations on cogenerator ownership; that is, electric utilities could not own a majority share of a cogeneration facility, nor could any utility holding company, nor a combination thereof. This section also defined the procedures for obtaining QF status.

Section 210 of the PURPA regulations specifically addressed major obstacles to developing cogeneration facilities. In fact, the regulations in Section 210 not only leveled the playing field, they tilted it in favor of the cogenerator. The principal issues in Section 210 include the following legal obligations of the electric utility toward the cogenerator:

- Obligation to purchase cogenerated energy and capacity from QFs.
- Obligation to sell energy and capacity to QFs.
- Obligation to interconnect.
- Obligation to provide access to transmission grid to “wheel” to another electric utility.
- Obligation to operate in parallel with QFs.
- Obligation to provide supplementary power, Backup power, maintenance power, and interruptible power.

Section 210 also exempted QFs from utility status and established a cost basis for purchase of the power from QFs. FERC specified that the price paid to the QF must be determined both on the basis of the utility’s avoided cost for producing that energy and, if applicable, on the capacity deferred as a result

[6]Since bottoming cycle cogeneration facilities do not use fuel for the primary production of electrical power, these facilities are only regulated when they use oil or natural gas for supplemental firing. The standards states that during any calendar year the useful power output of the bottoming cycle cogeneration facility must equal or exceed 45% of the energy input if natural gas or oil is used in the supplementary firing. The fuels are used first in the thermal process prior to the bottoming cycle cogeneration facility are not taken into account for satisfying PURPA requirements.

[7]For topping cycle cogeneration facilities using energy sources other than oil or natural gas (or facilities installed before 13 March 1980), no minimum has been set for efficiency.

of the QF power (e.g., cost savings from not having to build a new power plant). Other factors, such as QF power dispatchability, reliability, and cooperation in scheduling planned outages, could also be figured into the price paid to the QFs by the electric utilities. The state public utility commissions were responsible for determining the value of these avoided cost rates.

PURPA also gave special consideration to power produced from small (less than 80-MW) power production facilities (SPPFs). These must be fueled by biomass, geothermal, wastes, renewable resources, or any combination thereof. H.R. 4808, passed by the 101st Congress, lifted the 80-MW size restriction on wind, solar, geothermal cogeneration, and some waste facilities.[8] These SPPFs could become QFs and have the same status as a large gas-fired industrial cogeneration facility. Facilities between 1 kW and 30 MW may receive exemptions from the Federal Power Act (FPA) and the Public Utility Holding Company Act (PUHCA), and SPPFs using biomass are exempt from PUHCA and certain state restrictions, but not from the FPA.

17.6 Regulatory Developments of the 1990s and Early Twenty-First Century

New and restored terms in the power generation area were introduced in the early 1990s: nonutility generators (NUGs), independent power producers (IPPs), and exempt wholesale generators (EWGs). An EWG is an entity that wishes to build a new nonrate-based power plant under PUHCA and is only in the business of selling wholesale power. A revised PUHCA and a 1992 FERC ruling gave more open access to transmission lines and limited what a utility could charge third-party transmission users for that access.

The Energy Policy Act of 1992 enacted into law many of FERC's rules on EWGs and defined their legal status. In addition to opening up transmission-line access, the Energy Act also opened the possibility of retail wheeling, an area strongly favored by industry and heavily opposed by the electric utilities. An example of retail wheeling might be an industrial cogenerator in the Texas Gulf Coast area selling power to an industry in the Dallas area. Two or more electric utilities would be affected on their transmission lines, and the Dallas-area utility would lose the revenue from the customer in their service territory. From industry's viewpoint, they would like the ability to buy power wherever they could purchase it cheaper. The Energy Policy Act of 1992 was, in part, responsible for the deregulation of the electric utility industry.

The California Public Utility Commission held hearings in August 1994 on a proposal to have individual consumers choose their own electric company. California led the way "for the now-inevitable downfall of regulated electric utility monopolies…."[9] The California experience with deregulation was not pleasant. Electric rates remained high for consumers, brownouts occurred, and price fixing was prevalent among some of the state's power generators. The California experiences in 2000 and 2001 discouraged many states from deregulating their electric utilities. However, in 2005 approximately half of the states in the lower 48 have some form of deregulation already in place or a time schedule for deregulation. States in the northeastern U.S., anchored by Pennsylvania, and the state of Texas accomplished deregulation without experiencing many of the problems that occurred in California.

The Energy Policy Act of 2005 modified a number of provisions in PURPA. Because so many states have already deregulated their electrical utilities, some of the PURPA provisions discussed above were no longer needed. Section 1251 of the EPACT of 2005 amended PURPA by adding a "Net Metering" requirement for cogenerating facilities; Sec. 1252 added a requirement for "Smart Meters"; Section 1253 terminated the mandatory purchase of cogenerated power and power produced by small power cogenerators. Section 1254 of the 2005 act requires electric utilities to interconnect with customers who are also cogenerators. Section 1263 of the EPACT 2005 repeals PUHCA in its entirety and shifts all PUHCA responsibilities to FERC.

[8]Serial 101–160, Hearing of the Committee on Energy and Commerce, U.S. House of Representatives, June 14, 1990.
[9]USA Today, August 5, 1994, "Electricity Consumers May Get Power of Choice." p. 10A.

The Energy Policy Act of 2005 is a very broad energy act, placing more emphasis on energy efficiency, distributed generation, improved transmission systems, and fuel cells. Overall, it is favorable to cogeneration.

17.7 Environmental Considerations, Permitting, Water Quality

As described in the introduction, air and water regulations have been legislated since the 1950s. The Clean Air Act Amendments of 1977 established a new source of performance standards and placed environmental enforcement in the hands of the Environmental Protection Agency (EPA). The National Energy Policy Act of 1992 and the revised Clean Air Act Amendments (1990) placed even more stringent regulations on environmental pollutants. Concerns over acid rain, global warming, and the depletion of the ozone layer have prompted these tougher regulations.

The Clean Air Act Amendment of 1990 also produced new challenges and opportunities for power generation—trading and selling of emissions. Certain areas in the U.S. were designated as nonattainment zones where no new emissions sources could be located. Industries could, however, trade or even sell emissions (i.e., so much per ton of CO_2 or SO_2). This could become very important if a cogenerator wanted to locate a plant in a nonattainment zone and could buy emissions from a company that reduced production or somehow cut emissions from their plant.

A number of certifications are required to get a cogeneration plant approved. These include not only a FERC certificate but also various state permits. Since each state will set its own permitting requirements, it is not possible to generalize what is required on each application, but the following information will likely be required:

1. Nature of pollutant source.
2. Type of pollutant.
3. Level of emissions.
4. Description of process technology.
5. Process flow diagram.
6. Certification that the "Best Available Control Technology" (BACT) is being used.

Different requirements must be met if the proposed site is in an attainment or nonattainment zone. An attainment area will require sufficient modeling of ambient air conditions to insure that no significant deterioration of existing air quality occurs. This is PSD modeling or "prevention of significant deterioration." For a nonattainment area, no new emissions can be added unless they are offset by the removal of existing emissions. This has led to the selling or trading of emissions by industrial facilities and utilities.

17.7.1 Water Quality and Solid Waste Disposal

Water quality is usually not a problem with natural gas-fired cogeneration plants but will probably be monitored and will require both state and Federal permits if the wastewater is discharged into a public waterway. In extreme conditions it may be necessary to control both wastewater temperature and pH.

Solid waste disposal is not a problem with either natural gas or oil-fired cogeneration plants but could be a major problem for a coal-fired, coal gasification, or waste-to-energy cogeneration system. In some states the bottom ash from waste-to-energy plants was considered hazardous waste and the ash disposal cost per ton was more expensive than the refuse disposal cost. In 1994, bottom ash was declared by the EPA to be nonhazardous and therefore was allowed for land filling in standard landfills. Again, all states have different standards for both the quality of the water discharged and the requirements for solid waste disposal, and all project planners should check with the appropriate regulatory agencies in the state where the project is planned to be sure what air, water, and solid waste (if applicable) permits are required.

17.8 Economic Evaluations

This section will present a brief summary of the important aspects of economic evaluations of cogeneration systems. This section is not intended to be a complete presentation of engineering economics, which would be beyond the scope of this chapter. Rather, those aspects of economic evaluations that are specific or especially important to cogeneration systems will be described. These aspects include the baseline (or "business as usual") financial considerations, cogeneration economics including capital and operating costs, and the available economic and engineering computer programs. The two case studies presented at the end of this chapter will further illustrate some of the economic principles discussed in this section.

17.8.1 Baseline Considerations

Because a cogeneration project is capital intensive, it is extremely important to have accurate data to determine current and estimated (projected) future energy consumption and costs. In a large industrial plant with known process steam needs and 15-minute electric data available, it is usually fairly straightforward to establish the baseline case (i.e., the no-cogeneration case). For universities, hospitals, small manufacturing plants, and so on, the data are not always available. Typically whole campus/facility electrical data are available on a monthly basis only (from utility bills), and hourly electrical profiles may have to be "constructed" from monthly energy and demand data. Hourly thermal data also may not exist. Boiler operators may have daily logs, but these are typically not in electronic form. Constructing accurate hourly thermal energy profiles can involve hours of tedious work pouring over graphs and boiler operator logs. In the worst-case scenario, only monthly gas bills may be available, and it may be necessary to construct hourly thermal profiles from monthly bills, boiler efficiencies, and Btu content of the fuel. Since the monthly dollars have to be matched, there is often a great deal of trial and error involved before there is a match between assumed hourly energy profiles and monthly energy consumption and costs from the utility bills.

While a first-cut energy and cost analysis may be done on a monthly basis for an industrial facility that operates 7 days a week, 24 hours a day, that is not the case for a university, a commercial building complex, or a one- or two-shift manufacturing facility. Energy loads are highly variable, and hourly analyses are required. The possible exception may be hospitals, where 7-day-a-week operation is normally the case. However, there will often be a significant difference between weekday and weekend loads, which would require some sort of "day typing" in the analysis.

Assuming a good energy profile (hourly preferred) can be constructed, the "business as usual" scenario costs are determined. This is the annual cost of doing business without the cogeneration system, including annual purchased electrical energy sales, electrical demand charges, boiler fuel costs, and boiler maintenance. That total number is the baseline to which the cogeneration economics are compared.

17.8.2 Cogeneration Economics

If a simple payback approach is used, the cogeneration system costs divided by the savings over the baseline costs, would be the simple payback:

$$\text{Annual savings} = \text{Baseline costs} - \text{Cost with cogeneration system}$$

$$\text{Simple payback} = \frac{\text{Cogeneration system capital costs}}{\text{Annual savings}} \tag{17.11}$$

For many analyses, a simple payback approach is often enough. Once an acceptable payback period is defined, the cogeneration decision can be made. If the simple payback period is fairly short (2–3 years), then small variations in energy prices will have little effect on the decision. As the payback period lengthens (e.g., 4–8 years), then other factors should be examined. The time value of money has to be considered, as well as projected energy rates, and projected changes in energy needs for the facility. See Chapter 3 for more information.

A typical scenario for a project financed over 15 to 20 years might require an economic analysis that would include the following steps:

1. Determine current energy costs and current thermal and electrical energy requirements.
2. Project energy costs into the 15- to 20-year future using forecasts from the local utility or from the state public utility commission.
3. Project thermal and electrical needs into the future, considering possible growth (or declines) in energy requirements.

Note: Any projections of future thermal and electrical loads should carefully consider the impact of energy conservation on the total energy needs. Energy conservation is extremely important in the cost analysis because it represents dollars that do not have to be spent. The term “negawatt” has been applied to demand-side management programs as a means of expressing a decrease in electrical demand at a facility.

4. Project increasing maintenance costs associated with an expected increase in thermal loads (if appropriate).

Summing these costs will provide a more realistic “business as usual” case over the expected financing period of the cogeneration system.

Once the base case is determined, various sizing and operating scenarios for the cogeneration system should then be evaluated. These scenarios include (1) sizing a base-load cogeneration system and, if necessary, purchasing additional required electrical energy and producing supplemental required thermal energy, (2) following the thermal load profile and purchasing electrical energy needs, (3) following the electrical load profile and producing supplemental thermal energy, and (4) following the thermal profile with excess electricity produced. Note that with any potential scenario considered, all the thermal energy produced by the cogeneration system is used. It is the use of all the thermal energy that generally makes a cogeneration system feasible. If electricity production were the primary output, the cogenerator could not compete with the local electric utility. It is the simultaneous production of and need for the thermal energy that makes cogeneration attractive.

The next step is to determine the costs associated with a cogeneration system. This will depend on the type and size of system selected—diesel engine/generator, dual fuel diesel engine/generator, natural gas-fired engine/generator, or gas turbine/generator. As described in a previous section, a variety of possibilities exist in sizing the cogeneration system for a particular application (see “Matching Electrical and Thermal Loads”). Ideally, the cogeneration system could be sized to match the electrical and thermal loads exactly; however, seldom is there an exact match. Table 17.2 summarizes four possible outcomes.

17.8.3 Operating and Capital Costs

In any scenario, the savings and income from the cogeneration system must be determined as well as the additional costs such as the additional fuel, and the maintenance. Steps for determining the savings and income of the cogeneration system include

a. Savings in electricity generated (avoided purchase) = kWh × ¢/kWh.
b. Savings in demand charges (decreased demand) = kW × $/kW.
c. Savings in HRSG-produced steam.
d. Income from excess electrical power sales (if applicable).
e. Income from excess thermal energy sales (if applicable).

TABLE 17.2 Possible Scenarios to Consider for Cogeneration System Sizing

Scenario	Result	Application
Baseline cogeneration system that meets a portion of the facility's electrical and thermal requirements	Requires production of supplemental thermal energy and purchase of additional electrical energy	Universities/commercial buildings/thermal energy plants that have variable loads and do not want to get into the power sales business
Cogeneration system where the thermal load exceeds the equivalent electrical output of the generator-thermal load sizing	Electrical energy will have to be purchased from the local utility	Industrial facility or manufacturing plant with relatively constant thermal load
Cogeneration system designed to match electrical loads	Thermal energy to be produced by a boiler	Industrial facility with fairly high and constant electrical loads and lower, but variable, thermal loads
Cogeneration system sized to meet high thermal loads, but with lower equivalent electrical requirements	Excess electrical power sales to the local utility	Industrial or university facility that has lower or variable electrical needs, which require electricity sales to make the project economical

There will be additional costs for the cogeneration system. These include:

- Turbine/generator annual maintenance—use manufacturer's recommended fees, typically expressed in terms of ¢/kWh of electricity produced. The range may be from 1.0 ¢/kWh for gas turbines to as much as 3.0 ¢/kWh for small engines.
- Fuel cost for the prime mover (fuel oil and/or natural gas) at cost per MBtu
- Maintenance cost for HRSG—use manufacturer's recommended fee schedule, typically around $1/klb of steam.
- Any electrical utility charges for standby power, maintenance energy or power, etc., expressed in $/kW or ¢/kWh

As an example of the savings and costs of a cogeneration system, see Table 17.4 in the following case study of the Austin State Hospital. This table provides detailed cost and savings data for that application.

Finally, the system or capital costs must be determined. A general "rule of thumb" for estimating these system costs is as follows:

Major equipment packages (combustion turbine and HRSG)	40%
Balance of plant equipment	25%
Engineering and construction	15%
Other (interest during construction, permitting, contingency, etc.)	20%
	100%

Any cogeneration analysis is very complicated, and the results will vary greatly depending upon interest rates, cost of fuel, permitting requirements, cost of electricity sales, and other factors.

17.8.4 Final Comments on Economic Evaluations

Economic evaluations for assessing the feasibility of cogeneration projects is an important aspect of any development. Although many of the considerations are similar to other engineering economic evaluations, several aspects are unique to cogeneration projects. Obtaining detailed electric and thermal energy data is a prerequisite to completing accurate economic assessments. After the data are available, the savings and income of a cogeneration project must be compared to the additional costs of operating the system. If the net savings will "pay back" the capital expense of the cogeneration system in a reasonable time, then the project may make economic sense. A variety of sizes and operating modes of the cogeneration system should be examined to determine the most economical scenario. Computer programs are available to help in this process. Other detailed considerations that are often necessary in a thorough economic evaluation include tax and financing arrangements, but these considerations are beyond the scope of this chapter.

17.9 Financial Aspects

17.9.1 Overall Considerations

Cogeneration facilities can be financed by a variety of options. The traditional approach to financing is owner financing; however, cogeneration facilities are expensive to construct, and a company may choose not to use its own capital. When deciding on the most favorable financial arrangement, companies often consider the following factors:

- Shortage of capital.
- Effect on credit rating (if borrowing the money to finance the plant).
- Impact on balance sheet.
- Desire to fix savings (minimize risk).
- Inability to utilize available tax benefits.
- Lack of interest in operating or owning the plant (not main line of business).

Financing is critical to the success of a cogeneration facility, and it is best to determine, early on, the financing method to be used. Various financial structures are described below. There may be other, more creative ways, to finance projects, including combinations of these, but most financing will fall under one of the following:

1. Conventional ownership and operation.
2. Joint venture partnership.
3. Lease.
4. Third-party ownership.
5. Guaranteed savings.

17.9.2 Conventional Ownership and Operation (100% Ownership)

The heat load owner has two basic financing options in a conventional owner/operate structure: (1) to fund the project internally from profits in other areas of the business, or (2) to fund part of the project from internal sources and borrow the remainder from a conventional lending institution. Figure 17.13 shows one such financial structure.

Most businesses have a minimum internal rate of return on equity that they require for any investment. They may not be willing to fund any project that does not meet the internal hurdle rate using 100% equity (internal) financing. With 100% internal financing the company avoids the problems of arranging external financing (perhaps having to add partners). If there is a marginal return on equity, however, the company will not finance the project, especially if the money could be used to expand a product line or create a new product that could provide a greater return on equity.

Because the cost of borrowed money is typically lower than a business's own return on equity requirements, the combination of partial funding internally and conventional borrowing is often used. By borrowing most of the funds, the internal funds can be leveraged for other projects, thus magnifying the overall return on equity.

External contract issues are simpler in conventional ownership. Contracts will be required for the gas supply, for excess power sales to the local utility, possible O&M agreements, plus any agreements with possible lenders, but no contracts may be needed for the steam and electricity if all is used internally.

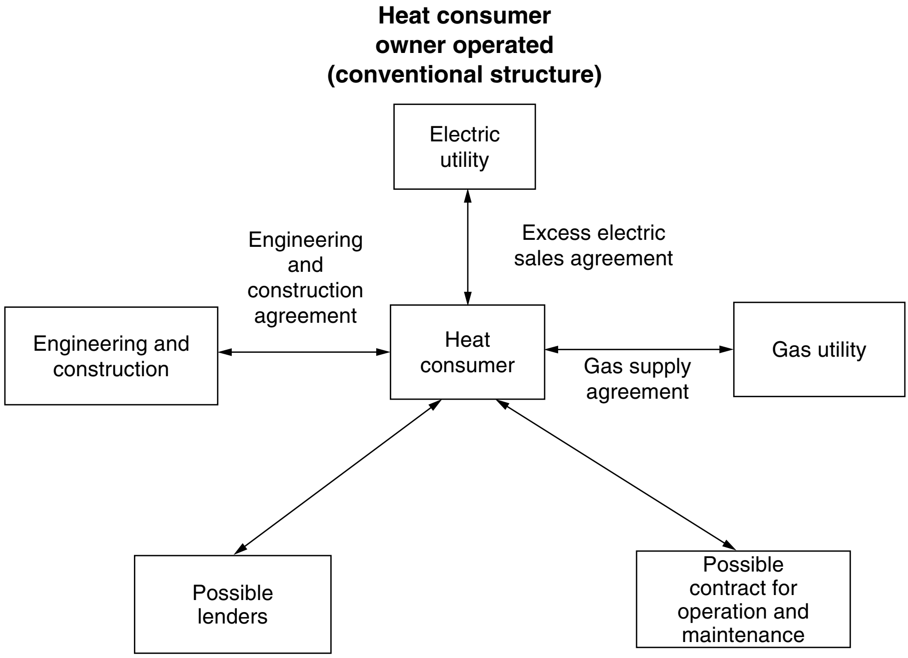


FIGURE 17.13 A schematic illustration of one financial structure for conventional ownership and operation of a cogeneration facility.

17.9.3 Joint Venture

One alternative to 100% ownership is to share ownership. Figure 17.14 shows the outline of one possible arrangement of a joint venture project. PURPA regulations allow electric utilities to own up to a 50% share in a cogeneration facility. Profits shared with a utility in a partial ownership role are unregulated. Other partners might include a gas utility or a major equipment vendor, such as a major gas turbine producer.

The major advantages of a joint venture are the sharing of risks and credit. The disadvantages are that profits are also shared and that contract complexity increases. Since the joint venture company is the owner, steam and power sales agreements, in addition to the other contracts required under conventional ownership, have to be signed with the heat consumer.

17.9.4 Leasing

In this financing arrangement, a company will build the cogeneration plant with the agreement that the steam host or heat consumer will lease the plant from the builder. In this model, the heat consumer is still heavily involved with the project construction but keeps at arm's length from the financing, since the lessor will have put that package together. These deals are often more difficult to put together and could take longer to develop than either joint venture or conventional ownership financing. Figure 17.15 is a schematic of one such lease agreement approach.

17.9.5 Third-Party Ownership

In third-party ownership, the heat consumer is distanced from both the financing and construction of the cogeneration facility. A third party arranges the finances, develops the project, and arranges for gas supply, power sales for the excess power produced, steam sales to the heat consumer, and O&M

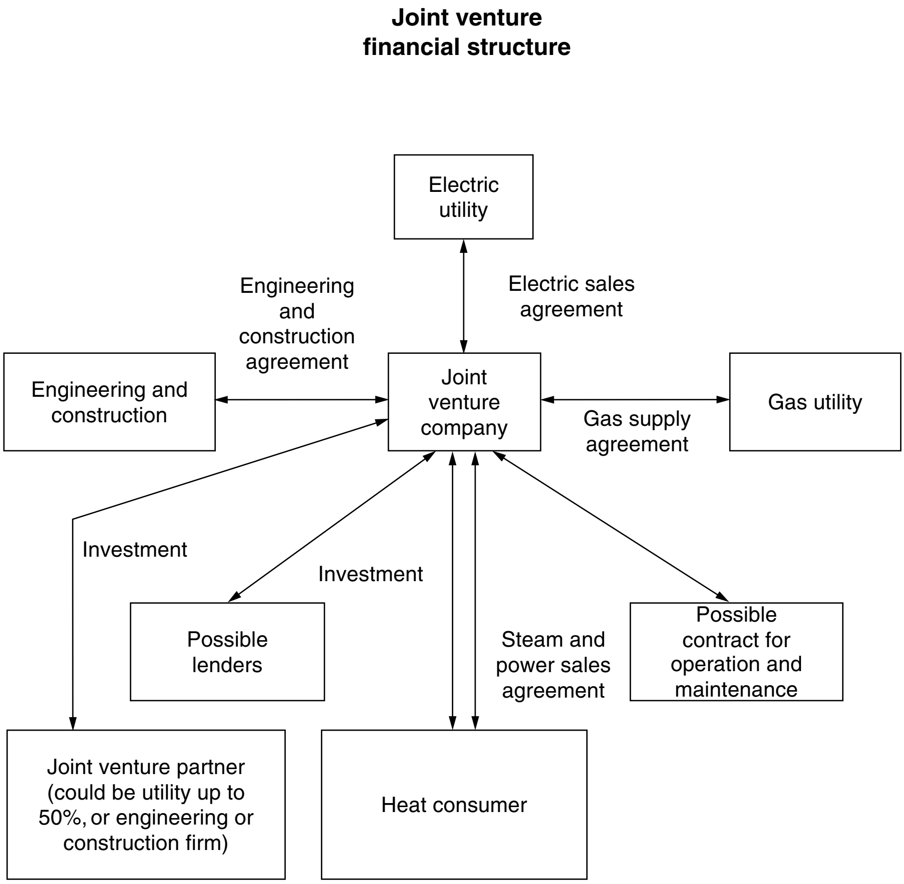


FIGURE 17.14 A schematic illustration of one financial structure for a joint venture ownership of a cogeneration facility.

agreements. Under the 1992 National Energy Policy Act, the third party may also be able to enter into a power sales contract with the heat consumer as well. Figure 17.16 is a schematic of one type of third-party arrangement. The lessee, shown in Figure 17.16, can be a minimally capitalized entity that is primarily in the business of plant operation and handling energy sales once the plant is built. The contracts for firm energy sales are key to this type of financing operation. This type of financing option can be complex, requiring more development time, so development costs could be much higher.

17.9.6 Guaranteed Savings Contracts

These types of contracts are more common in smaller sized cogeneration systems, and typically may use packaged facilities. Figure 17.17 is a schematic of one possible guaranteed savings arrangement. A developer may pay all costs for construction of a cogeneration facility and then operate and maintain the system. Further, a guarantee of savings would be made to the owner over a specified length of time. Typically, a guaranteed savings contract will run from 5 to 10 years, with an average of 8 years, and will guarantee a fixed savings per year. The "split" between the guaranteed savings contractor usually ranges from a 75/25 division to a 95/5 division, with a "typical" split of 85/15; that is, 85% of the savings to the guaranteed savings contractor and 15% to the owner. Any additional savings above the guarantee are often split between the guaranteed savings contractor and the owner.

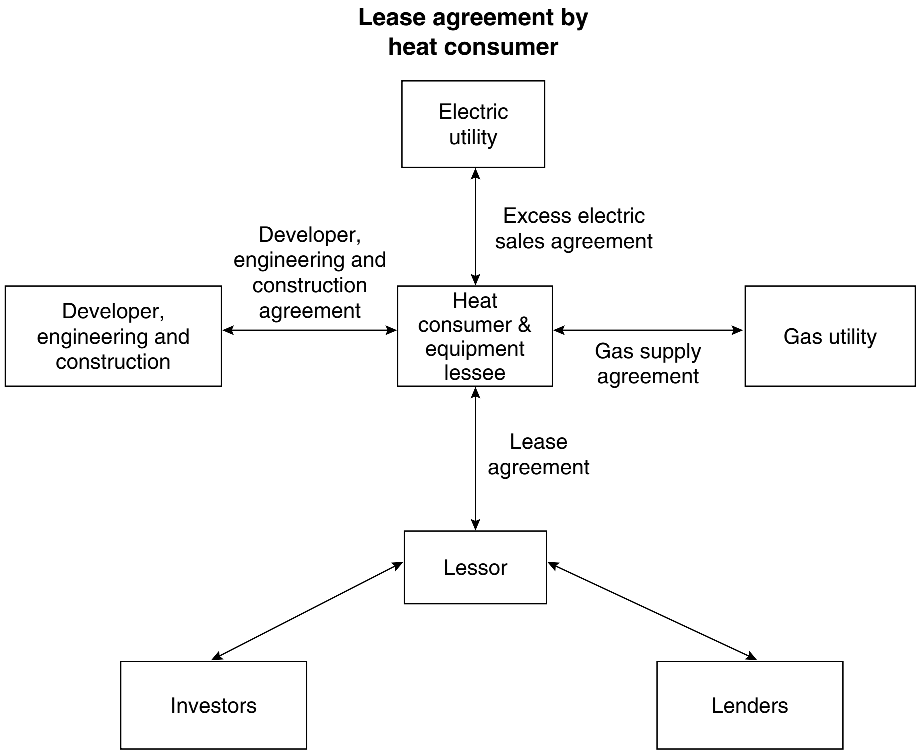


FIGURE 17.15 A schematic illustration of one financial structure for a lease arrangement for a cogeneration facility.

Guaranteed savings contracts have a distinct advantage to the owner in that all the capital and most of the risks are taken by the guaranteed savings contractor. The owner is guaranteed a percentage savings each year. These types of contracts should be considered when (1) the owner does not have capital to invest in a cogeneration system, or (2) when the owner does not have the necessary experience in-house to operate/manage such a project. If the capital is available, and if the in-house expertise exists, the company would be ahead financially to do the project itself, as the net return would be much higher.

17.9.7 Final Comments on Financial Aspects

Financing arrangements are a crucial aspect of most cogeneration developments. These arrangements may range from simple to highly complex. They are affected by internal factors such as ownership arrangements, credit ratings, and risk tolerance. In addition, these financial aspects are affected by external factors such as the financial and credit markets, tax laws, and cogeneration regulations. A variety of initial financial arrangements have been outlined in this section to illustrate the nature of these arrangements. Much more detailed arrangements are possible and often necessary, but these are beyond the scope of this chapter.

17.10 Case Studies

Case studies will be presented. One is a state hospital where a small gas turbine was installed, and the other is a utility application that generates power and sells steam to a local paper mill.

17.10.1 The Austin State Hospital Case Study

The Austin State Hospital was a facility operated by the Texas Mental Health and Mental Retardation (MHMR) state agency. It was one of 15 sites selected for a potential cogeneration site and was analyzed by

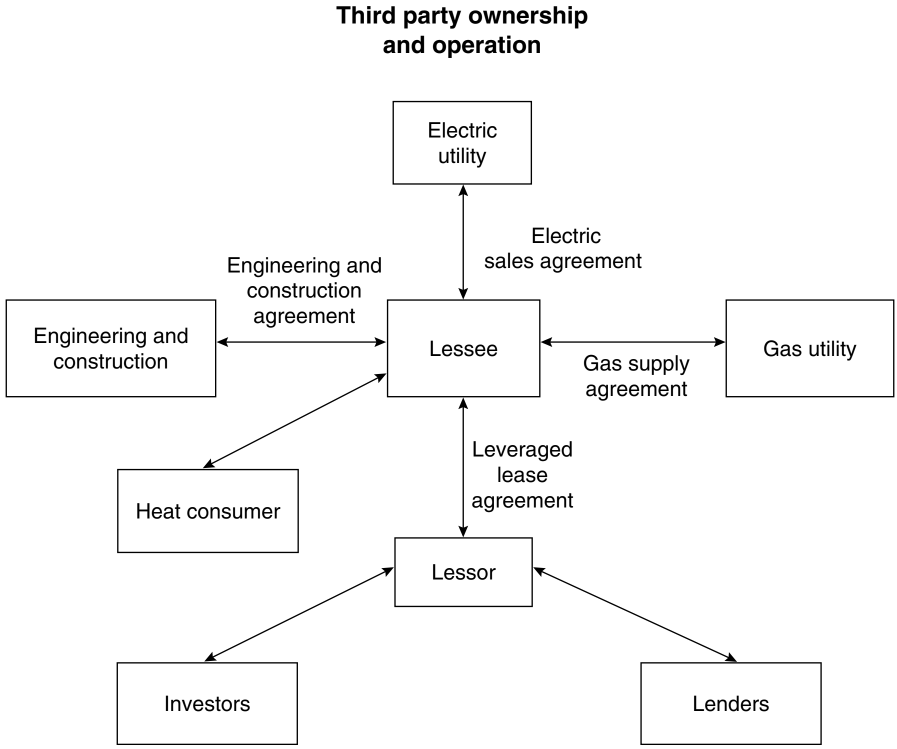


FIGURE 17.16 A schematic illustration of one financial structure for third-party ownership and operation of a cogeneration facility.

Texas A&M's Energy Systems Laboratory (ESL) in 1985. A large laundry and cafeteria provided most of the thermal loads needed for the project, but there were also significant heat loads during the winter. In 1985, the ESL study recommended that gas turbines totaling 1.5 MW be installed, at an estimated cost of $1.5 million. An engineering study was then initiated in 1986–1987, and steam data was taken to determine the steam loads more accurately. The consulting engineering study recommended a 1-MW system at an estimated cost of $1.3 million. A $1.5 million grant application was submitted to and approved by Department of Energy (DOE). Specifications were prepared, the project was bid, and all bids came in too high. The project was then stalled for nearly 2 years until additional funding was secured and a long-term gas contract could be signed with the Texas General Land Office (GLO). The project was rebid in 1990, and a contractor was selected (Caton, Muraya, and Turner 1991).

The 1-MW gas turbine, heat-recovery steam generator, and controls were completed in May 1992, at a cost of over $1.8 million, or slightly more than $1,800/kW. The system supplies steam to the Austin State Hospital complex and provides base-load electricity for the facility. The cogeneration system was projected to save $225,000 per year, for a simple payback of 8 years. A long-term gas contract with the GLO assured the long-term economics of the system. The GLO established a maximum gas cost escalation rate of 5% per year for the duration of the payback period. Because DOE was providing a $1.5 million grant for the system, the benefit to the state of Texas was large. The state contribution of $300,000 would be paid back in less than 2 years. Table 17.3 summarizes 2 years of operation of the cogeneration system. The system was monitored by the ESL at Texas A&M University.

Shortly after the installation of the cogeneration system was completed (in May 1992), lightning struck the transformer and power line into the Austin State Hospital and burned out many of the controls of the cogeneration system. This caused several weeks of downtime during that summer. Additional start-up

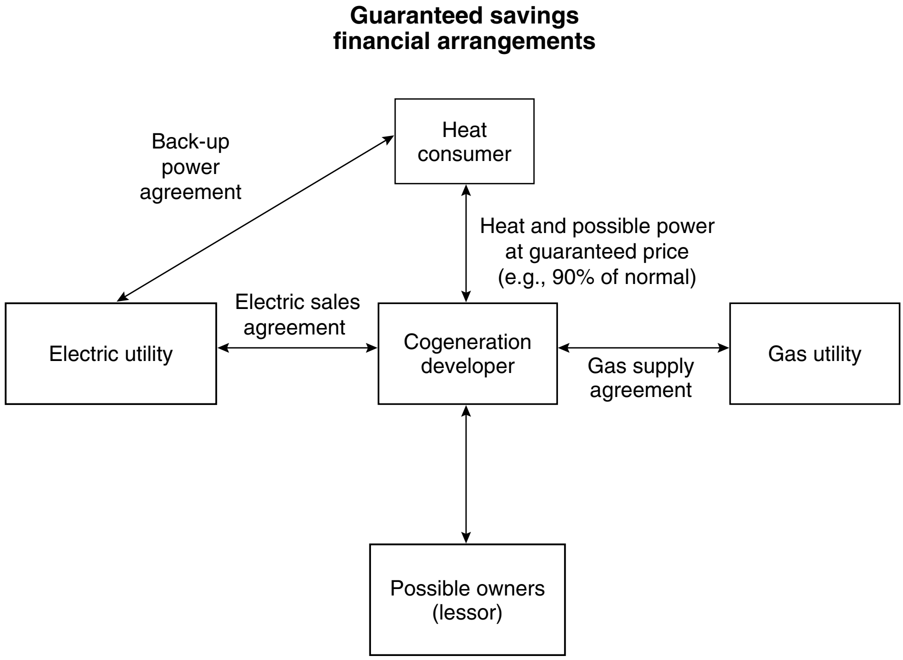


FIGURE 17.17 A schematic illustration of one financial structure for a guaranteed savings arrangement of a cogeneration facility.

TABLE 17.3 Energy Usage for the Austin State Hospital Cogeneration System

Date	Purchased Electricity (kWh/mo)	Cogenerated Electricity (kWh/mo)	Total Electricity (kWh/mo)	HRSG Heating (MMBtu/mo)
August 1992	994,909	629,714	1,624,623	3851
September 1992	1,159,110	395,537	1,554,647	1182
October 1992	694,243	655,740	1,349,983	4030
November 1992	475,120	550,046	1,025,166	3449
December 1992	303,951	721,222	1,025,173	5168
January 1993	251,741	704,480	956,221	4862
February 1993	268,584	663,623	932,207	4755
March 1993	346,492	698,975	1,045,467	4910
April 1993	541,877	557,241	1,099,118	3947
May 1993	601,993	649,361	1,251,354	5066
June 1993	795,057	601,731	1,396,788	4923
July 1993	880,342	629,613	1,509,955	4913
August 1993	887,767	641,948	1,529,715	5221
September 1993	756,969	610,222	1,367,191	5016
October 1993	605,129	588,025	1,193,154	4812
November 1993	284,797	684,472	969,269	5077
December 1993	305,742	703,319	1,009,061	5535
January 1994	320,072	676,443	996,515	4708
February 1994	246,905	656,497	903,402	4643
March 1994	326,711	661,360	988,071	4841
April 1994	409,269	641,899	1,051,168	5009
May 1994	515,273	655,417	1,170,690	5437
June 1994	866,959	582,881	1,449,840	5220

TABLE 17.4 Summary of Savings Calculations for the Austin State Hospital (1992 dollars)

Electrical energy cost		$0.02609/kWh
Average demand cost		$11.69/kW
Natural gas fuel cost		$3.19/MCF
(August 1992–June 1994)		
	Energy usage (per year)	Savings
a. Electricity	14,599,766 kWh	$198,209
b. Demand (avoided)	23,190 kW	$141,183
c. Heating (HRSG)	106,477 Mbtu	$229,565
d. Maintenance of conventional boiler at $1.10/klb of steam (not required)		$57,496
	Total savings	$626,452
Energy costs for the cogeneration System	Expenses (per year)	
Natural gas for the gas turbine	$332,394	
Standby charge	30,600	
Maintenance of the cogeneration system ($0.004/kWh)	30,365	
Maintenance of HRSG ($1.00/klb)	52,261	
Total costs	$445,620	
Net savings	$181,000/year	

problems were attributed to the waste heat system. Overall, the cogeneration system performed exceptionally well, with a turbine/generator availability of 96.8% for a 2-year period (July 1992 through June 1994). This 96.8% availability includes downtime for scheduled maintenance. Excluding scheduled maintenance downtime, the turbine/generator availability was nearly 100%.

Problems with the waste-heat boiler reduced its availability to approximately 92% over the same period. This reduced the expected steam recovery below the assumed availability. This is not the reason, however, for the lower-than-projected savings. The city of Austin changed the type of electrical service and increased the demand charges in 1992. An extra standby charge added approximately $30,000 annually to the bill. Without this charge, the cogeneration system would be saving nearly $210,000 annually, almost exactly the annual amount predicted by the savings analysis. Table 17.4 is a summary based on actual metered data for the ASH cogeneration system for 23 months of operation (August 1992 through June 1994). The annual dollar savings are approximately $181,000 per year, which gives a simple payback of approximately 10 years. Because the bulk of the money was provided by a Federal grant, the state portion has already been repaid. Therefore, the cogeneration project saved the state over $180,000 annually based on 1992 electricity and gas prices. The Austin State Hospital cogeneration system was shut down after 10 years of operation. The gas turbine remained reliable throughout most of the period of operation. The waste-heat recovery system had many problems during its operating period. However, Table 17.3 and Table 17.4 give actual data and measured performance of a small-scale system.

17.10.2 Klamath Cogeneration Project—An Industrial/Utility Case Study

The Klamath cogeneration project (KCP) is a 484-MW natural-gas fired, combined cycle plant that was completed in 2001. This plant is an example of community and state involvement that is prevalent in twenty-first century plants. State involvement came from the state of Oregon's Energy Facility Siting Council (EFSC), which approved the project based on reduced CO_2 emissions from the more efficient cogeneration system. Community involvement was also important because a nearby industry (wood processing facility) was to be a steam host for the cogeneration plant. The facility's boilers

TABLE 17.5 KCP Facility's Provision of Steam to Collins

Year	Months of Operation	lbs. Steam Sent to Collins	Hours of Operation (approximately)	Flow Rate (lb/h)	CO_2 Offsets (Short tons)
2001	July–December	181,288,984	3884	46,675	15,301
2002	January–December	533,655,766	8551	62,408	50,077
2003	January–December	623,281,664	8758	71,159	53,324
2004	January–December	623,759,744	8713	71,214	53,154

Source: From TRC Global Management Solutions, *Klamath Cogeneration Project, 2005. Annual Report to the Energy Facility Siting Council*, Revision 1.

were to be shut down, thus reducing CO_2 emissions. Although there are other aspects of the overall project, the bottom line is a significant offset of CO_2 emissions, largely from the cogeneration project.

17.10.3 Project Overview

The original intent of the project was to generate 484 MW of electricity and supply 200,000 pounds/h of steam to the wood processing plant. The Oregon EFSC approved the permitting based on shutting down the plant's boilers, which would guarantee an offset of 4,464,395 short tons of CO_2 over 30 years. Unfortunately, after the permitting was approved, the wood processing plant shut down one of their process lines, which reduced their steam needs to less than half on the original design.

Table 17.5 gives the steam supplied to the wood processing plant since 2001. Although the steam usage is increasing, the steam supplied to the plant is far short of the 200,000 pounds/h required in the permitting. Thus, the Klamath cogeneration project will have to find another large steam host or purchase additional CO_2 offsets in order to meet the state of Oregon environmental requirements.

The city of Klamath Falls also had to guarantee the project's heat rate. If it fell short, then additional carbon offsets had to be provided to cover any shortfalls.

This case study is being provided to show the links between cogeneration projects and the environment. Even in a very strict environmental state such as Oregon, a highly efficient cogeneration project can be permitted and licensed to operate. It also points to a problem area discussed earlier, i.e., the need to have a thermal load for the excess steam. When the project was first initiated, there was a large steam host capable of accepting huge quantities of steam. The needs of the wood processing plant changed, however, requiring only one-third of the original steam, and this has created a problem for the cogeneration system. The economics could change dramatically if the success of the project hinged on selling 200,000 pounds/h of steam, and the environmental impact by reducing the CO_2 offset could place severe environmental constraints on the city. At the end of 2004, the city was over 3 million tons of CO_2 short of achieving its offset goal.

17.11 Small-Scale Cogeneration Applications in Buildings

Large-scale cogeneration systems have been shown to be very attractive in increasing thermal efficiency and reducing fuel costs and emissions. This technology also offers opportunities for cogeneration systems to be applied to commercial and residential buildings. Recent development of equipment such as microturbines, small-sized reciprocating engines, Stirling engines, and fuel cells makes cogeneration suitable for building use. Building-oriented, small-scale CHP is called *BCHP*. BCHP uses the waste heat for heating, cooling, and humidity control of buildings. The overall thermal efficiency of small-scale CHP can be as much as 85%–90%.

17.11.1 Technology Status of Small-Scale CHP

Microscale CHP started in western Europe and Japan, and is gaining popularity in these countries, Driven by the potential market and advanced technology, the U.S. DOE and the U.S. EPA initiated the

"CHP Challenge" in December 1998 to double the installed capacity of CHP systems in the U.S. by 2010. In 1998, the installed CHP capacity was 46 GW, and the goal set for 2010 is 92 GW. As of the end of December 2003, the installed capacity of CHP in the U.S. was approximately 71 GW. About 0.04% of this installed capacity is coming from systems with less than 100 kW of power output, and much less coming from systems with less than 10 kW of power output (Bernstein and Knoke 2004). However, with technology advances, the small-scale cogeneration systems are expected to experience greater growth in the near future.

Table 17.6 compares some of the important performance characteristics of the various types of small-scale systems.

Natural gas is the preferred fuel for the spark-ignition engines, but they can also run on propane or gasoline. Compression-ignition engines can operate on diesel fuel or heavy oil, or they can be set up in a dual-fuel configuration that burns primary natural gas with a small amount of diesel fuel. The engines used for micro-CHP units are normally designed as packaged units. The size of micro-CHP units, which use internal combustion engine as a prime mover, varies from 10 to 200 kW. The unit can be setup easily on-site. However, some drawbacks exist in the reciprocating-engine package units because they require regular maintenance, are noisy without proper noise abatement, and have potentially high emissions. They are not as attractive for residential applications.

Microturbines are a newly developed small-sized gas turbine with power generation from 25 to 250 kW. Like the lager gas turbines, the microturbine generator consists of a compressor, a combustion chamber, a one-stage turbine, and a generator. The rotating speed of the generator can be up to 10,000 rpm. The high-frequency electricity output is first rectified and then converted to 60 Hz. A rectifier and a transformer are invaluable devices in the generation set. Several companies are developing packaged microturbine CHP systems. The 30-kW and 60-kW units are already in the marketplace. The electricity efficiency of microturbines is about 25%–30%. The advantages of microturbine CHP are low noise and relatively low NO_x emission. The disadvantages are low electricity efficiency and high cost. Microturbine CHP are suitable for taking a base-load of electricity, heating, and cooling because of their inflexibility in handling load changes.

TABLE 17.6 Technical Features of Small-Scale CHP Devices

	Reciprocating Engines	Microturbines	Stirling Engines	PEM Fuel Cells
Electrical power (kW)	10–200	25–250	2–50	2–200
Electrical efficiency, full load (%)	24–45	25–30	15–35	40
Electrical efficiency, half load (%)	23–40	20–25	35	40
Total efficiency (%)	75–85	75–85	75–85	75–85
Heat/electrical power ratio	0.9–2	1.6–2	1.4–3.3	0.9–1.1
Output temperature level (°C)	85–100	85–100	60–80	60–80
Fuel	Natural or biogas, diesel fuel oil	Natural or biogas, diesel, gasoline, alcohols	Natural or biogas, LPG, several liquid or solid fuels	Hydrogen, gases, including hydrogen, methanol
Interval between maintenance (h)	5000–20,000	20,000–30,000	5,000	N/A
Investment cost ($/kW)	800–1500	900–1500	1300–2000	2500–3500
Maintenance costs (¢/kW)	1.2–2.0	0.5–1.5	1.5–2.5	1.0–3.0

Source: From Alanne, K. and Saari, A. 2004. *Renewable and Sustainable Energy Reviews*, 8, 401–431.

The Stirling engines are small-scale engines ranging from 2 to 50 kW that are targeting the future residential CHP needs. The Stirling engine uses the Stirling cycle and is a reciprocating engine. However, unlike the internal-combustion engine, the Stirling engine is an external-combustion engine. The working gases inside the engine cylinder are typically helium or hydrogen and never leave the engine. The combustion takes place outside the cylinder. The piston is driven by compression or expansion of working gases due to the alternating heating and cooling of the cylinder by external heat sources. The engine converts the temperature difference into electricity. The Stirling engine can run on various fuels, both gas fuels and solid fuels, due to its external combustion. The advantages of Stirling engines are its quiet operation, little maintenance, and low NO_x emission. The disadvantages are relative lower electricity efficiency, typically about 25%–30%. When the Stirling engine is used as prime mover in micro-CHP systems, the total thermal efficiency is about the same level of other micro-CHP applications. Stirling-engine CHP packaged units are being developed. Several demonstration projects have shown the system efficiency can be over 90%.

A fuel cell is an electrochemical device that produces electricity by an oxidation-reduction reaction between a fuel, typically hydrogen, and atmospheric oxygen. A fuel cell consists of an anode, a cathode, and an electrolyte. The fuel (hydrogen) is fed to the anode and releases electrons on the anode that are conducted to the cathode via an external circuit. The electron current can be utilized to generate electricity. The hydrogen protons diffuse through the electrolyte material to the cathode. Oxygen or air passes through the cathode in the fuel cell, where oxygen is reacted with the hydrogen proton to form water and release heat. A typical fuel cell system consists of three parts: a fuel reformer section, power generation section, and power conditioning section. The heat rejected from the fuel reformer and power generation section can be recovered and used to drive absorption chillers and heat exchanges to supply hot water and chilled water for building air conditioning systems. The fuel cell technologies can be classified as polymer electrolyte membrane (PEM), phosphoric acid fuel cell (PAFC), solid oxide fuel cell (SOFC), and molten carbonate fuel cell (MCFC), according to their types of electrolyte. The PEM and PAFC are lower-temperature fuel cells, with operating temperatures of 200 and 400°F, respectively. The SOFC and MCFC are high-temperature fuel cells with operating temperatures of 1750 and 1200°F, respectively. The advantages of fuel cell technology are its high electrical efficiency, which can be as high as 45%–55%, and low emissions. Fuel cells are applicable for various purposes, based on the operating temperature levels of the cells. The major disadvantage of a fuel cell is the high price. In 2005, the 200-kW PAFC became available in the market.

17.11.2 Building Combined Heat and Power Systems

Commercial and residential buildings are responsible for about 37% of total U.S. primary energy consumption. Ninety-five percent of the commercial buildings are powered by electricity. Ninety-seven percent of these buildings use electricity for air conditioning. Approximately 73.9% of the electricity is generated from combustion of coal, petroleum, and natural gas in the U.S. Because many commercial buildings have a simultaneous need for electricity and thermal energy, there is a potential application for BCHP. The energy consumed by the BCHP system will be utilized at a higher efficiency than conventional means.

Modular CHP units or packaged systems have been developed in recent years to cater to the needs of buildings. Modular cogeneration systems are compact, and can be manufactured economically. These systems, ranging in size from 20 to 650 kW, produce electricity and hot water from engine waste heat. Figure 17.18 shows an IC-engine-driven BCHP system. The engine exhaust gases pass through the heat recovery steam generator, which generates superheated steam to drive the steam-driven absorption chiller and hot-water heat exchanger. The condensed water first passes through the engine to take away the engine heat, then passes through HRSG. This system makes use of engine exhaust heat and engine cooling heat to supply hot water and chilled water for the building air conditioning system. It is usually better to size the systems to meet the basic electricity load and use an auxiliary boiler or duct burner to

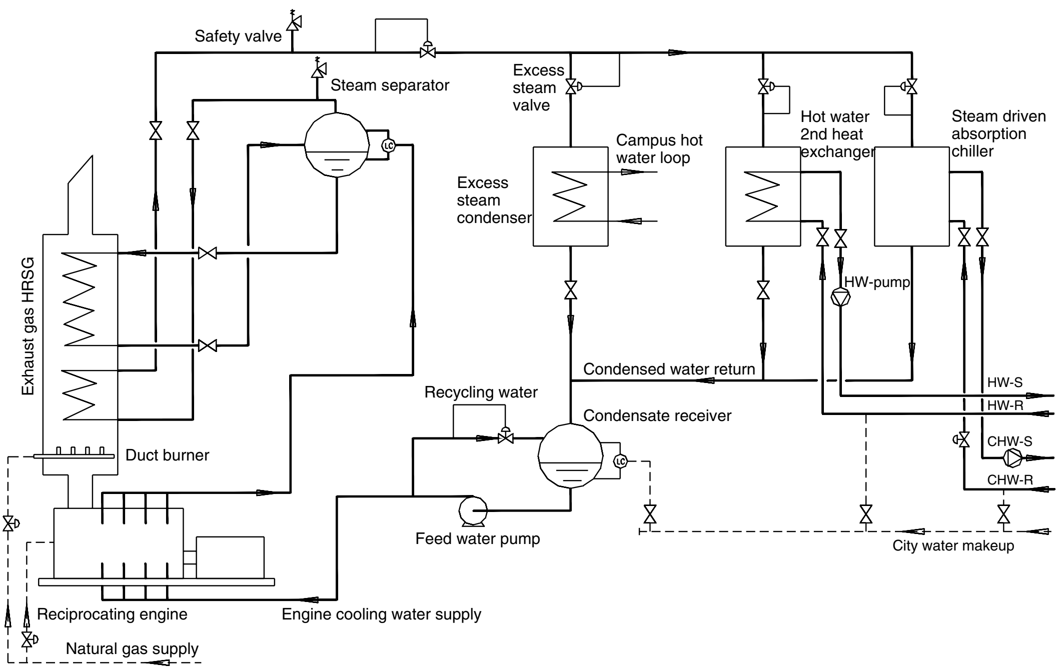


FIGURE 17.18 IC-engine-driven microcogeneration system.

adjust CHP thermal output. Therefore, this type of BCHP can run continuously. The best CHP applications for buildings are hospitals or restaurants that have a year-round need for hot water or steam.

Microscale CHP systems using different prime movers are also being developed by several companies to satisfy the electricity and thermal need of residential buildings. These micro-CHP package have a capacity of up to 16 kW, and are capable of providing most of the heating and electrical needs for a home. Currently, there are several micro-CHP demonstration units running, in which the prime movers vary as IC engine, microturbine, and fuel cell. The major barriers in commercializing these systems in the U.S. market are cost, reliability, and regulations.

17.11.3 Market Barriers and Drivers for Building CHP

Electricity deregulation has opened up the traditional power generation market for competition. On the scale of BCHP, deregulation takes the form of net metering in many states, which requires utilities to pay the customer retail price for the electricity exported to the grid. Deregulation, environmental concerns, and incentive programs of the government are the major market drivers for the installation of BCHP. However, several factors also affect the growth of installation of BCHP. These factors include the initial cost of buying and installing a micro-CHP system, maintenance costs, and environmental control requirements. Some electric utilities charge stand-by fees for the installation of micro-CHP so that the utilities may recover the costs of the stranded assets when customers turn to on-site generation. Customers using micro-CHP typically require a backup source of power to meet load requirements during outages or scheduled maintenance. The stand-by charge sometimes can kill a micro-CHP project.

Currently, the high cost is a major barrier preventing the installation of BCHP. High initial cost relative to the central station electric power generation can be a deterrent to the investment in micro-CHP. Although micro-CHP products normally have a higher efficiency compared to the traditional power generation, the cost of the electricity and heat generated by micro-CHP products highly depends on the cost of the micro-CHP products and the cost of fuel. Technology development will certainly reduce the price of the micro-CHP products in the future.

17.12 Future of Cogeneration

The future of cogeneration will include refined and better versions of the technology already described. In addition, several advanced cogeneration cycles are currently in various stages of development. These include (1) the Cheng cycle, (2) the Kalina cycle, and (3) coal gasification combined cycles.

17.12.1 Cheng Cycle

The Cheng cycle for optimizing the performance of gas-turbine cogeneration systems was first identified by Dr. Dan Yu Cheng in 1974. It uses the steam produced in the HRSG for reinjection into the combustion chamber to provide additional power from the gas turbine. Because the gas turbine is a mass-flow device, combining the steam with the combustion gases will increase the total electrical power produced. This cycle has an advantage over the traditional combined cycle (Rankine cycle) because of its simplicity. There is no need for steam turbine condensers and pumps. Part of the steam produced in the HRSG can also be used for process steam. Following electrical loads or following process steam loads is possible using steam injection in the Cheng cycle. Substantial increases in both power output and generating efficiency are possible.

An added benefit from the steam injection is reduced NO_x emissions from the system. Strasser (1991) summarizes the results of several Cheng cycle installations in the U.S. and abroad, the oldest of which was an installation at San Jose State University in San Jose, California. That particular site had accumulated over 46,000 operating hours through November 1991.

17.12.2 Kalina Cycle

The Kalina cycle has been proposed by Dr. Alex Kalina as a means of improving the overall efficiency of combined-cycle plants. Its application to cogeneration systems would be to replace the conventional Rankine cycle power generation with a more efficient Kalina cycle. Dr. Kalina proposed the use of an ammonia-water mixture as the working fluid instead of water. The advantages include a 10%–20% increased thermal efficiency over a Rankine cycle with the same boundary conditions.

Ibrahim and Kovach (1993) describe a power generation application using the Kalina cycle, and Zervos, Leibowitz, and Robinson (1992) describe a 3-MW demonstration plant that uses the Kalina cycle. The complexity of the plant increases with the use of the NH_3–H_2O mixture. The conventional boiler is replaced by a vapor generator, the steam turbine by a vapor turbine, and the condenser by a distillation/condensation system. In the DOE demonstration project (Zervos, Leibowitz, and Robinson 1992), the liquid ammonia concentration is 70% by weight as it leaves the distillation/condensation system and enters the simulated waste-heat boiler. The superheated vapor then enters a specially designed vapor turbine, which drives a generator and produces electrical power. The Kalina cycle demonstration project operates between 2,000 and 20 psia with ammonia. Mixture concentrations vary from 30 to 70%. The vapor turbine inlet conditions include a pressure of 1,600 psia, a temperature of 960°F, an enthalpy of 1,221 Btu/lb, an ammonia concentration of 70%, and a flow rate of 31,000 lb/h. The distillation/condensation system consists of a low-pressure condenser, vaporizer, preheater, and pump; an intermediate pressure shell-and-tube heat exchanger, falling film evaporator, flash drum, and pump; and a high-pressure condenser, preheater, pump, and solids removal system. Obviously, the management of the distillation/condensation system is crucial to the operation of a Kalina cycle system.

The vapor turbine has to withstand the high-temperature corrosion of the ammonia-water mixture and is made of type-316 stainless steel. Pressurized steam is injected into the labyrinth seals to contain the ammonia-water mixture and prevent its leakage during the vapor expansion phase in the turbine. This demonstration plant using the Kalina cycle began operation in December 1991.

17.12.3 Coal Gasification Combined Cycle Applications

The United States Clean Coal Technology (CCT) program was established in 1984 with the set-aside of $750 million for research and development of commercial processes to accelerate the use of clean coal technologies for power production. The Clean Air Act amendments of 1990 prescribed reductions in acid rain effects and emissions, aimed primarily at the electric utility industry. Coal gasification projects, as well as fluidized-bed combustion, have been the focus of the DOE CCT program (Hemenway, Williams, and Huber 1991.

The purpose of the coal gasification project is to produce a syngas that can be burned in a combustion turbine, hopefully at a price that is competitive with natural gas. The obstacles are coal cleanup, including desulfurization and particulate removal, development of components capable of operating at higher temperatures, disposal of solids and particulates from the process, high-temperature cleaning of the gas stream, and optimization of combustor design for the lower-temperature gas.

Sulfur removal is a must for these units to meet the standards of the Clean Air Act for electric power production. There are various approaches used to remove sulfur, including the addition of limestone or dolomite to the gasifier. Particulate cleanup is usually accomplished by a series of cyclone separators and particle filters. The tricky problem with the gas cleanup is that it should be accomplished in the heated gas stream that may at a temperature of 1000°F or more. Waste heat from the combustion turbine exhaust is routed through the HRSG for steam production. The steam may then be piped to a steam turbine to produce additional electrical power or used partly for process heating. If some stage in the gasification process requires a fixed temperature to be maintained, it is possible that additional steam could be produced in the gasifier that could also be used for process heating or for additional power production.

Ruth and Bedick (1992) summarize five major integrated gasification combined cycle (IGCC) applications sponsored by the DOE. All of the projects are scheduled for start-up/demonstration in the 1995–1996 calendar years. The five activities are (1) Combustion Engineering IGCC Repowering, (2) Tampa Electric IGCC, (3) Pinion Pine IGCC, (4) Toms Creek IGCC Demonstration, and (5) Wabash River Coal Gasification Repowering Project.

Of the five planned projects discussed by Ruth and Bedick, three have been built and are in operation. The Wabash (Indiana) plant is a repowering plant, another application of the IGCC technology—to replace aging coal-fired plants with more efficient and more environmentally friendly IGCC plants. The first "greenfield" (or new plant) is Tampa's 313 MW facility, dedicated in 1997, received many awards for its gas cleaning technology (DOE Fossil Energy website 2006). In February 1997 there were nine IGCC plants operating worldwide, including three in the United States (The third U.S. plant is outside Reno, Nevada, the Pinion Pine IGCC project.). By 2000 nearly 4 Gigawatts (GW) were in use worldwide (Coal 21 website [http://www. Coal21.com.au/IGCC.php]). Most of the commercial IGCC plants are in the 250–300 MW capacity, currently restricted by the size of the pressurized gasifiers. Since the gasifiers are pressure vessels, they cannot be fabricated on site, and transport restrictions currently limit size (IEA Clean Coal Centre [http://www.iea-coal.org.uk]).

Other authors have examined the economics of obtaining and marketing the by-products produced by the gasification process, including methanol, acetic acid, nitric acid, and formaldehyde, as a means of enhancing revenues (Bahmann, Epstein, and Kern 1992). Several scenarios were analyzed for cost effectiveness depending on the type of coproduct produced. The IGCC research at the U.S. DOE and throughout the world is focused on high temperature gas cleanup and improved turbines, both of which will improve IGCC efficiency. There is still a huge emphasis on improved methods to remove harmful pollutants that are emitted from the IGCC process. A good overview of the current IGCC technology is a summary by Maurstad (2005).

17.13 Summary and Conclusions

The importance of cogeneration is monetary, environmental, and energy savings. Any facility that uses electrical power and has thermal energy needs is a candidate for cogeneration. Facilities that may be considered for cogeneration include those in the industrial, commercial, and institutional sectors. The technology for conventional cogeneration systems is for the most part available and exists over a range of sizes: from less than 100 kW to over 100 MW. The major equipment requirements include a prime mover, electrical generator, electrical controls, heat recovery systems, and other typical power plant equipment. These components are well developed, and the procedures to integrate these components into cogeneration systems are well established.

In addition to the economic and technical considerations, the application of cogeneration systems involves an understanding of the governmental regulations and legislation on electrical power production and on environmental impacts. With respect to electrical power production, certain governmental regulations were passed during the late 1970s, again in the 1990s, and in 2005 that remove barriers and provide incentives to encourage cogeneration development. Finally, no cogeneration assessment would be complete without an understanding of the financial arrangements that are possible.

This chapter has provided a brief survey of the important aspects necessary to understand cogeneration. Specifically, this chapter has included information on the technical equipment and components, technical design issues, regulatory considerations, economic evaluations, financial aspects, and future cogeneration technologies. The chapter included two case studies to illustrate the range of technologies available for cogeneration systems, the types of considerations involved in selecting a cogeneration system, and the overall economic and environmental evaluations that may be completed.

Cogeneration will continue to experience progressive growth because of its inherent efficiency over conventional power generation and thermal energy production. As distributed generation becomes more widespread, we will see advances in smaller systems, sized for single buildings, including residential applications.

Acknowledgments

The author wishes to acknowledge the assistance of Gary Gong, PhD candidate at Texas A&M University, for his assistance with this chapter. Also, Dr. Jerald Caton of the Mechanical Engineering Department at Texas A&M University was a contributing author to the earlier chapter.

References

Alanne, K. and Saari, A. 2004. Sustainable small-scale CHP technologies for building: The basis for multi-perspective decision-making. *Renewable and Sustainable Energy Reviews*, 8 (5), 401–431.

Bahmann, P. D., Epstein, M., and Kern, E. E. 1992. Coal gasification-based integrated coproduction energy facilities. *1992 ASME Cogen-Turbo Conference, IGTI*, 7, 69–74.

Bernstein, S. and Knoke, S. 2004. Micro-CHP: U.S. market potential and complex challenges. *Refocus*, 5 (2), 36–39.

Caton, J. A., Muraya, N., and Turner, W. D. 1991. Engineering and economic evaluations of a cogeneration system for the Austin State Hospital, In *Proceedings of the 26th Intersociety Energy Conversion Engineering Conference*, Vol. 5, pp. 438–443.

Coal 21 website, http://www.coal21.com.au/IGCC.php

DOE Fossil Energy website. 2006. http://www.fe.doe.gov/programs/powersystems/gasification.

FERC Regulations. 1978. Part 292—Regulations under sections 201 and 210 of the Public Utility Regulatory Policies Act of 1978 with regard to small power production and gogeneration.

Hemenway, A., Williams, W. A., and Huber, D. A. 1991. Effects of the Clean Air Act Amendments of 1990 on the commercialization of fluidized bed technology. *ASME Fluidized Bed Combustion*219–224.

Ibrahim, M. and Kovach, R. M. 1993. A Kalina cycle application for power generation. *Energy*, 18 (9), 961–969.

IEA Clean Coal Centre, http://www.iea-coal.org.uk

Laderoute, C. D. 1980. The FERC's PURPA cogeneration rules: Economics, rate design, and policy aspects, *Annual UMM DNR 7th Conference Energy Proceedings*, pp. 114–122.

Maurstad, O. 2005. *An Overview of Coal Based Integrated Gasification Combined Cycle (IGCC) Technology*, MIT LFEE 2005-002 WP, September 2005, http://lfee.mit.edu, click on publications, then click on working papers.

Paffenbarger, J. A. 1991. A GCC power plant with methanol storage for intermediate-load duty. *Journal of Engineering for Gas Turbines and Power*, 113, 151–157.

Ruth, L. K. and Bedick, R. C. 1992. Research and development efforts at the Department of Energy (DOE) supporting integrated gasification combined cycle (IGCC) demonstrations. *1992 ASME Cogen-Turbo Conference, IGTI*, 7, 87–94.

Shipley, A. M. and Elliot, R. N. 2003. Phantom power: The status of fuel cell technology markets. *Industrial Energy Technology Conference Procedings*. 119–127.

Spiewak, S. 1980. Regulation of cogeneration. *McGraw-Hill Conference on Industrial Power—Electrical Rates, Reliability, and Energy Management*, October 7–8.

Spiewak, S. A. 1987. *Cogeneration and Small Power Production Manual*. Fairmont Press, Atlanta, GA.

Strasser, A. 1991. The Cheng cycle cogeneration system: Technology and typical applications. *ASME Cogen-Turbo Conference, JGTI*, 6, 419–428.

Swenson, A. 1998. *A Look at Commercial Buildings in 1995: Characteristics, Energy Consumption, and Energy Expenditures*. U.S. Department of Energy, Washington, DC.

The Energy Information Administration (EIA) at the U.S. Department of Energy (DOE). 2005. Building energy data. http://www.eia.doe.gov.

TRC Global Management Solutions. 2005. *Klamath Cogeneration Project. 2005 Annual Report to the energy facility siting Council*. Revision 1.

Turbine-based cogeneration systems. In *Planning Cogeneration Systems*, D. R. Limaye, ed., pp. 119–143. Fairmont Press, Atlanta, GA.

United States Combined Heat and Power Association 2001. *National CHP Roadmap*.

Wood, B. D. 1982. *Applications of Thermodynamics. 2nd Ed.* Addison-Wesley, Reading, MA.

Zervos, N. G., Leibowitz, H. M., and Robinson, K. S. 1992. Startup and demonstration experience of the Kalina cycle demonstration plant. *ASME Cogen-Turbo Conference, IGTI*, 7, 187–191.

18

Energy Storage, Transmission, and Distribution

Roel Hammerschlag
Institute for Lifecycle Environmental Assessment

Robert Pratt
Pacific Northwest National Laboratory

Christopher P. Schaber
Institute for Lifecycle Environmental Assessment

Steve Widergren
Pacific Northwest National Laboratory

18.1 Energy Storage Technologies

Roel Hammerschlag and Christopher P. Schaber

18.1.1 Overview of Storage Technologies

Energy storage will play a critical role in an efficient and renewable energy future; much more so than it does in today's fossil-based energy economy. There are two principal reasons that energy storage will grow in importance with increased development of renewable energy:

- Many important renewable energy sources are intermittent, and generate when weather dictates, rather than when energy demand dictates.
- Many transportation systems require energy to be carried with the vehicle.[1]

Energy can be stored in many forms: as mechanical energy in rotating, compressed, or elevated substances; as thermal or electrical energy waiting to be released from chemical bonds; or as electrical charge ready to travel from positive to negative poles on demand.

[1]This is almost always true for private transportation systems, and usually untrue for public transportation systems, which can rely on rails or overhead wires to transmit electric energy. However, some public transportation systems such as buses do not have fixed routes and also require portable energy storage.

TABLE 18.1 Overview of Energy Storage Technologies and Their Applications

	Utility Shaping	Power Quality	Distributed Grid	Automotive
		Direct electric		
Ultracapacitors		✓		✓
SMES		✓		
		Electrochemical		
Batteries				
Lead–acid	✓	✓	✓	
Lithium-ion	✓	✓	✓	✓
Nickel–cadmium	✓	✓		
Nickel–metal hydride				✓
Zebra				✓
Sodium–sulfur	✓	✓		
Flow Batteries				
Vanadium redox	✓			
Polysulfide bromide	✓			
Zinc bromide	✓			
Electrolytic hydrogen				✓
		Mechanical		
Pumped hydro	✓			
Compressed air	✓			
Flywheels		✓		✓
		Direct Thermal		
Sensible Heat				
Liquids			✓	
Solids			✓	
Latent Heat				
Phase change	✓		✓	
Hydration–dehydration	✓			
Chemical reaction	✓		✓	
		Thermochemical		
Biomass solids	✓		✓	
Ethanol	✓			✓
Biodiesel				✓
Syngas	✓			✓

All technologies are discussed in this chapter except hydrogen-based solutions, which are treated in chapter 27 and chapter 28.

Storage media that can take and release energy in the form of electricity have the most universal value, because electricity can efficiently be converted either to mechanical or heat energy, whereas other energy conversion processes are less efficient. Electricity is also the output of three of the most promising renewable energy technologies: wind turbines, solar thermal, and photovoltaics. Storing this electricity in a medium that naturally accepts electricity is favored, because converting the energy to another type usually has a substantial efficiency penalty.

Still, some applications can benefit from mechanical or thermal technologies. Examples are when the application already includes mechanical devices or heat engines that can take advantage of the compatible energy form; lower environmental impacts that are associated with mechanical and thermal technologies; or low cost resulting from simpler technologies or efficiencies of scale.

In this chapter, the technologies are grouped into five categories: direct electric, electrochemical, mechanical, direct thermal, and thermochemical. Table 18.1 is a summary of all of the technologies covered. Each is listed with indicators of appropriate applications that are further explained in Section 18.1.3.

18.1.2 Principal Forms of Stored Energy

The storage media discussed in this chapter can accept and deliver energy in three fundamental forms: electrical, mechanical, and thermal. Electrical and mechanical energy are both considered high-quality energy because they can be converted to either of the other two forms with fairly little energy loss (e.g., electricity can drive a motor with only about 5% energy loss, or a resistive heater with no energy loss).

The quality of thermal energy storage depends on its temperature. Usually, thermal energy is considered low quality because it cannot be easily converted to the other two forms. The theoretical maximum quantity of useful work $W_{\max}$ (mechanical energy) extractable from a given quantity of heat Q is

$$W_{\max} = \frac{T_1 - T_2}{T_1} \times Q,$$

where T_1 is the absolute temperature of the heat and T_2 is the surrounding, ambient absolute temperature.

Any energy storage facility must be carefully chosen to accept and produce a form of energy consistent with either the energy source or the final application. Storage technologies that accept and/or produce heat should, as a rule, only be used with heat energy sources or with heat applications. Mechanical and electric technologies are more versatile, but in most cases electric technologies are favored over mechanical because electricity is more easily transmitted, because there is a larger array of useful applications, and because the construction cost is typically lower.

18.1.3 Applications of Energy Storage

In Table 18.1 above, each technology is classified by its relevance in one to four different, principal applications:

- *Utility shaping* is the use of very large capacity storage devices to answer electric demand, when a renewable resource is not producing sufficient generation. An example would be nighttime delivery of energy generated by a solar thermal plant during the prior day.
- *Power quality* is the use of very responsive storage devices (capable of large changes in output over very short timescales) to smooth power delivery during switching events, short outages, or plant run-up. Power-quality applications can be implemented at central generators, at switchgear locations, and at commercial and industrial customers' facilities. Uninterruptible power supplies (UPS) are an example of this category.
- *Distributed grid technologies* enable energy generation and storage at customer locations, rather than at a central (utility) facility. The distributed grid is an important, enabling concept for photovoltaic technologies that are effective at a small scale and can be installed on private homes and commercial buildings. When considered in the context of photovoltaics, the energy storage for the distributed grid is similar to the utility shaping application in that both are solutions to an intermittent, renewable resource, but distributed photovoltaic generation requires small capacities in the neighborhood of a few tens of MJ, while utility shaping requires capacities in the TJ range.[2] Renewable thermal resources (solar, geothermal) can also be implemented on a distributed scale, and require household-scale thermal storage tanks. For the purposes of this chapter, district-heating systems are also considered a distributed technology.
- *Automotive applications* include battery-electric vehicles (EVs), hybrid gasoline–electric vehicles, plug-in hybrid electric vehicles (PHEVs), and other applications that require mobile batteries larger than those used in today's internal combustion engine cars. A deep penetration of automotive batteries also could become important in a distributed grid. Large fleets of EVs or PHEVs that are grid connected when parked would help enable renewable technologies, fulfilling utility shaping and distributed grid functions as well as their basic automotive function.

---

[2]Storage capacities in this chapter are given in units of MJ, GJ, and TJ: 1 MJ = 0.28 kWh, 1 GJ = 280 kWh, and 1 TJ = 280 MWh.

Additional energy storage applications exist, most notably portable electronics and industrial applications. However, the four applications described here make up the principal components that will interact in a significant way with the global energy grid.

18.1.4 Specifying Energy Storage Devices

Every energy storage technology, regardless of category, can be roughly characterized by a fairly small number of parameters. Self-discharge time, unit size, and efficiency serve to differentiate the various categories. Within a category, finer selections of storage technology can be made by paying attention to cycle life, specific energy, specific power, energy density, and power density.

Self-discharge time is the time required for a fully charged, noninterconnected storage device to reach a certain depth of discharge (DOD). DOD is typically described as a percentage of the storage device's useful capacity, so that, for instance, 90% DOD means 10% of the device's energy capacity remains. The relationship between self-discharge time and DOD is rarely linear, so self-discharge times must be measured and compared at a uniform DOD. Acceptable self-discharge times vary greatly, from a few minutes for some power-quality applications, to years for devices designed to shape annual power production.

Unit size describes the intrinsic scale of the technology, and is the least well-defined of the parameters listed here. If the unit size is small compared to the total required capacity of a project, complexity and supply shortages can increase the cost relative to technologies with a larger unit size. Some technologies have a fairly large unit size that prohibits small-scale energy storage.

Figure 18.1 maps all of the technologies discussed in this chapter, according to their unit size and 10% self-discharge time. The gamut of technologies available covers many orders of magnitude on each axis, illustrating the broad choice available. Utility shaping applications require a moderate self-discharge time

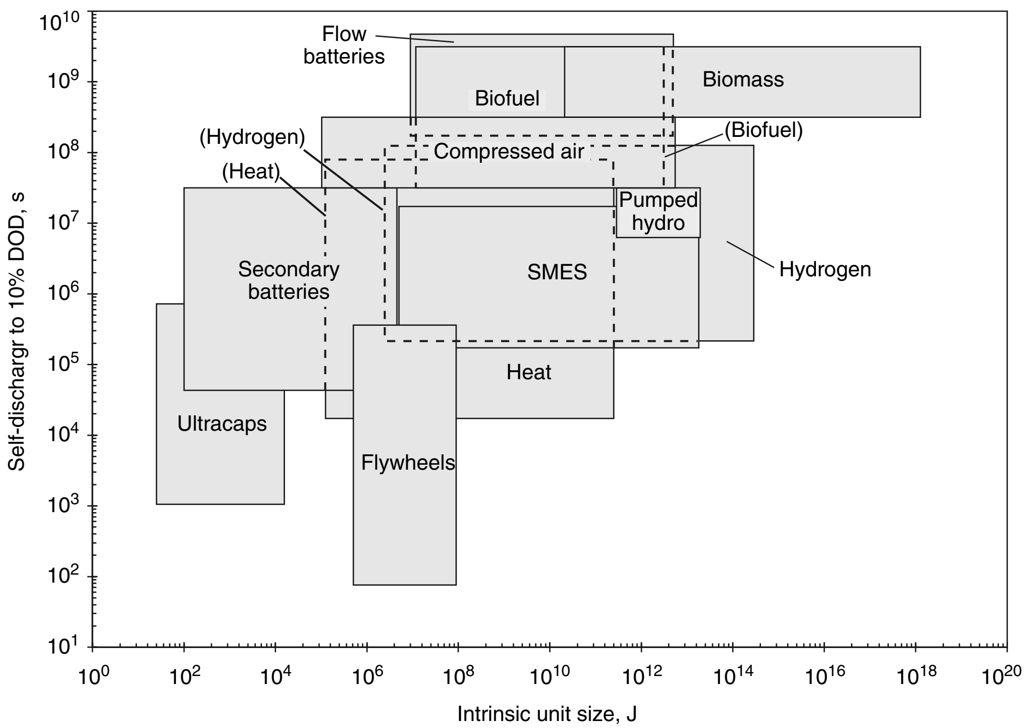


FIGURE 18.1 All storage technologies, mapped by self-discharge time and unit size. Not all hidden lines are shown. Larger self-discharge times are always more desirable, but more or less important depending on the application. Intrinsic unit size does not have a desirability proportional to its value, but rather must be matched to the application.

and a large unit size; power-quality applications are much less sensitive to self-discharge time but require a moderate unit size. Distributed grid and automotive applications both require a moderate self-discharge time and a moderate unit size.

Efficiency is the ratio of energy output from the device, to the energy input. Like energy density and specific energy, the system boundary must be carefully considered when measuring efficiency. It is particularly important to pay attention to the form of energy required at the input and output interconnections, and to include the entire system necessary to attach to those interconnections. For instance, if the system is to be used for shaping a constant-velocity, utility wind farm, then presumably both the input and output will be AC electricity. When comparing a battery with a fuel cell in this scenario, it is necessary to include the efficiencies of an AC-to-DC rectifier for the battery, an AC-powered hydrogen generation system for the fuel cell system, and DC-to-AC converters associated with both systems.

Efficiency is related to self-discharge time. Technologies with a short self-discharge time will require constant charging to maintain a full charge; if discharge occurs much later than charge in a certain application, the apparent efficiency will be lower because a significant amount of energy is lost in maintaining the initial, full charge.

Cycle life is the number of consecutive charge–discharge cycles a storage installation can undergo while maintaining the installation's other specifications within certain, limited ranges. Cycle-life specifications are made against a chosen DOD depending on the application of the storage device. In some cases, for example pressurized hydrogen storage in automobiles, each cycle will significantly discharge the hydrogen canister and the appropriate DOD reference might be 80% or 90%. In other cases, for example a battery used in a hybrid electric vehicle, most discharge cycles may consume only 10% or 20% of the energy stored in the battery. For most storage technologies, cycle life is significantly larger for shallow discharges than deep discharges, and it is critical that cycle-life data be compared across a uniform DOD assumption.

Specific energy is a measure of how heavy the technology is. It is measured in units of energy per mass, and in this chapter this quantity will always be reported in MJ/kg. The higher the specific energy, the lighter the device. Automotive applications require high specific energies; for utility applications, specific energy is relatively unimportant, except where it impacts construction costs.

Energy density is a measure of how much space the technology occupies. It is measured in units of energy per volume, and in this chapter we will always report this quantity in MJ/L. The higher the energy density, the smaller the device. Again, this is most important for automotive applications, and rarely important in utility applications. Typical values for energy density associated with a few automotive-scale energy technologies are listed in Table 18.2, together with cycle-life and efficiency data.

Energy-density and specific-energy estimates are dependent on the system definition. For example, it might be tempting to calculate the specific energy of a flow battery technology by dividing its capacity by the mass of the two electrolytes. But it is important to also include the mass of the electrolyte storage

TABLE 18.2 Nominal Energy Density, Cycle Life and Efficiency of Automotive Storage Technologies

	Energy Density MJ/L	Cycle Life at 80% DOD[a]	Electric Efficiency %
Ultracapacitors	0.2	50,000	95
Li-ion batteries	1.8	2,000	85
NiMH batteries	0.6	1,000	80
H_2 at 350 bar	3.0	n/a[b]	47
H_2 at 700 bar	5.0	n/a	45
Air at 300 bar	<0.1	n/a	37
Flywheels	<0.1	20,000	80
Ethanol	23.4	n/a	n/a

Electric efficiencies are calculated for electric-to-electric conversion and momentary storage.
[a]Depth of discharge.
[b]Not applicable.

containers, and of the battery cell for a fair and comparable estimate of its specific energy. Therefore, the energy density and specific energy are dependent on the size of the specific device; large devices benefit from efficiency of scale with a higher energy density and specific energy. *Specific power* and *power density* are the power correlates to specific energy and energy density.

18.1.5 Specifying Fuels

A fuel is any (relatively) homogenous substance that can be combusted to produce heat. Though the energy contained in a fuel can always be extracted through combustion, other processes may be used to extract the energy (e.g., reaction in a fuel cell). A fuel may be gaseous, liquid, or solid. All energy storage technologies in the thermochemical category store energy in a fuel. In the electrochemical category, electrolytic hydrogen is a fuel.

A fuel's lower heating value (LHV) is the total quantity of sensible heat released during combustion of a designated quantify of fuel. For example, in the simplest combustion process, that of hydrogen,

$$2\,H_2 + O_2 \rightarrow 2\,H_2O(\text{vapor}) + \text{LHV},$$

or for the slightly more complex combustion of methane,

$$CH_4 + 2\,O_2 \rightarrow CO_2 + 2\,H_2O(\text{vapor}) + \text{LHV}.$$

In this chapter, the quantity of fuel is always expressed as a mass, so that LHV is a special case of specific energy. Like specific energy, LHV is expressed in units of MJ/kg in this chapter.

Higher heating value (HHV) is the LHV, plus the latent heat contained in the water vapor resulting from combustion.[3] For the examples of hydrogen and methane, this means

$$2\,H_2 + O_2 \rightarrow 2\,H_2O(\text{liquid}) + \text{HHV},$$

and

$$CH_4 + 2\,O_2 \rightarrow CO_2 + 2\,H_2O(\text{liquid}) + \text{HHV}.$$

The latent heat in the water vapor can be substantial, especially for the hydrogen-rich fuels typical in renewable energy applications. Table 18.3 lists LHVs and HHVs of fuels discussed in this chapter; in the most extreme case of molecular hydrogen, the HHV is some 18% higher than the LHV. Recovery of the latent heat requires controlled condensation of the water vapor; technologies for doing so are described in Chapter 13.

In this chapter, all heating values are reported as HHV rather than LHV. HHV is favored for two reasons: (1) its values allow easier checking of energy calculations with the principle of energy conservation, and (2) when examining technologies for future implementation, it is wise to keep an intention of developing methods for extracting as much of each energy source's value as possible.

18.1.6 Direct Electric Storage

18.1.6.1 Ultracapacitors

A capacitor stores energy in the electric field between two oppositely charged conductors. Typically, thin conducting plates are rolled or stacked into a compact configuration with a dielectric between them. The dielectric prevents arcing between the plates and allows the plates to hold more charge, increasing the maximum energy storage. The ultracapacitor—also known as supercapacitor, electrochemical capacitor, or electric double layer capacitor (EDLC)—differs from a traditional capacitor in that it employs a thin

[3]The concepts of sensible and latent heat are explained further in Section 18.1.9.

TABLE 18.3 Properties of Fuels

	Chemical Formula	Density g/L	LHV MJ/kg	HHV MJ/kg
Methanol	CH_3OH	794	19.9	22.7
Ethanol	C_2H_5OH	792	26.7	29.7
Methane	CH_4	0.68	49.5	54.8
Hydrogen	H_2	0.085	120	142
Dry syngas, airless process[a]	$40H_2 + 21CO + 10CH_4 + 29CO_2$	0.89	11.2	12.6
Dry syngas, air process[a]	$25H_2 + 16CO + 5CH_4 + 15CO_2 + 39N_2$	0.99	6.23	7.01

[a]Chemical formulae and associated properties of syngas are representative; actual composition of syngas will vary widely according to manufacturing process.

Source: From All except syngas from U.S. Department of Energy, *Properties of Fuels*, Alternative Fuels Data Center 2004.

electrolyte, on the order of only a few angstroms, instead of a dielectric. This increases the energy density of the device. The electrolyte can be made of either an organic or an aqueous material. The aqueous design operates over a larger temperature range, but has a smaller energy density than the organic design. The electrodes are made of a porous carbon that increases the surface area of the electrodes and further increases energy density over a traditional capacitor.

Ultracapacitors' ability to effectively equalize voltage variations with quick discharges make them useful for power-quality management and for regulating voltage in automotive systems during regular driving conditions. Ultracapacitors can also work in tandem with batteries and fuel cells to relieve peak power needs (e.g., hard acceleration) for which batteries and fuel cells are not ideal. This could help extend the overall life and reduce lifetime cost of the batteries and fuel cells used in hybrid and electric vehicles. This storage technology also has the advantage of very high cycle life of greater than 500,000 cycles and a 10- to 12-year life span.[1] The limitations lie in the inability of ultracapacitors to maintain charge voltage over any significant time, losing up to 10% of their charge per day.

18.1.6.2 Superconducting Magnetic Energy Storage

An superconducting magnetic energy storage (SMES) system is well suited to storing and discharging energy at high rates (high power.) It stores energy in the magnetic field created by direct current in a coil of cryogenically cooled, superconducting material. If the coil were wound using a conventional wire such as copper, the magnetic energy would be dissipated as heat due to the wire's resistance to the flow of current. The advantage of a cryogenically cooled, superconducting material is that it reduces electrical resistance to almost zero. The SMES recharges quickly and can repeat the charge/discharge sequence thousands of times without any degradation of the magnet. A SMES system can achieve full power within 100 ms.[2] Theoretically, a coil of around 150–500 m radius would be able to support a load of 18,000 GJ at 1000 MW, depending on the peak field and ratio of the coil's height and diameter.[3] Recharge time can be accelerated to meet specific requirements, depending on system capacity.

Because no conversion of energy to other forms is involved (e.g., mechanical or chemical), the energy is stored directly and round-trip efficiency can be very high.[2] SMES systems can store energy with a loss of only 0.1%; this loss is due principally to energy required by the cooling system.[3] Mature, commercialized SMES is likely to operate at 97%–98% round-trip efficiency and is an excellent technology for providing reactive power on demand.

18.1.7 Electrochemical Energy Storage

18.1.7.1 Secondary Batteries

A secondary battery allows electrical energy to be converted into chemical energy, stored, and converted back to electrical energy. Batteries are made up of three basic parts: a negative electrode, positive

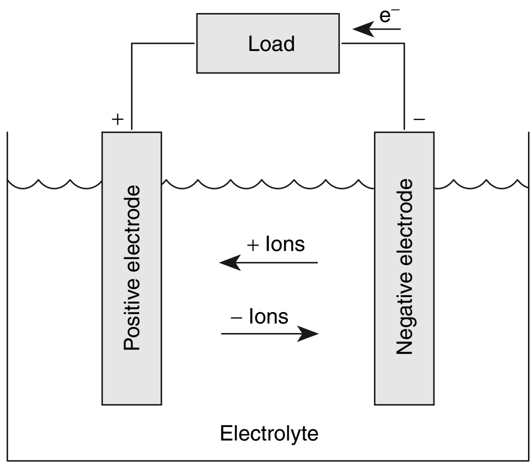


FIGURE 18.2 Schematic of a generalized secondary battery. Directions of electron and ion migration shown are for discharge, so that the positive electrode is the cathode and the negative electrode is the anode. During charge, electrons and ions move in the opposite directions and the positive electrode becomes the anode while the negative electrode becomes the cathode.

electrode, and an electrolyte (Figure 18.2). The negative electrode gives up electrons to an external load, and the positive electrode accepts electrons from the load. The electrolyte provides the pathway for charge to transfer between the two electrodes. Chemical reactions between each electrode and the electrolyte remove electrons from the positive electrode and deposit them on the negative electrode. This can be written as an overall chemical reaction that represents the states of charging and discharging of a battery. The speed at which this chemical reaction takes place is related to the internal resistance that dictates the maximum power at which the batteries can be charged and discharged.

Some batteries suffer from the "memory effect" in which a battery exhibits a lower discharge voltage under a given load than is expected. This gives the appearance of lowered capacity but is actually a voltage depression. Such a voltage depression occurs when a battery is repeatedly discharged to a partial depth and recharged again. This builds an increased internal resistance at this partial depth of discharge and the battery appears as a result to only be dischargeable to the partial depth. The problem, if and when it occurs, can be remedied by deep discharging the cell a few times. Most batteries considered for modern renewable applications are free from this effect, however.

18.1.7.2 Lead–Acid

Lead–acid is one of the oldest and most mature battery technologies. In its basic form, the lead–acid battery consists of a lead (Pb) negative electrode, a lead dioxide (PbO_2) positive electrode and a separator to electrically isolate them. The electrolyte is dilute sulfuric acid (H_2SO_4), which provides the sulfate ions for the discharge reactions. The chemistry is represented by:

$$PbO_2 + Pb + 2\,H_2SO_4 \leftrightarrow 2\,PbSO_4 + 2\,H_2O.$$

(In all battery chemistries listed in this chapter, left-to-right indicates battery discharge and right-to-left indicates charging.)

There are three main types of lead–acid batteries: the flooded cell, the sealed gel cell, and the sealed absorbed glass mat (AGM) lead–acid battery. The wet cell has a liquid electrolyte that must be replaced occasionally to replenish the hydrogen and oxygen that escape during the charge cycle. The sealed gel cell has a silica component added to the electrolyte to stiffen it. The AGM design uses a fiberglass-like separator to hold electrolyte in close proximity to the electrodes, thereby increasing efficiency. For both the gel and AGM configurations, there is a greatly reduced risk of hydrogen explosion and corrosion from disuse. These two types do require a lower charging rate, however. Both the gel cells and the AGM batteries are sealed and pressurized so that oxygen and hydrogen produced during the charge cycle are recombined into water.

The lead–acid battery is a low-cost and popular storage choice for power-quality applications. Its application for utility shaping, however, has been very limited due to its short cycle life. A typical installation survives a maximum of 1500 deep cycles.[4] Yet, lead–acid batteries have been used in a few commercial and large-scale energy management applications. The largest one is a 140-GJ system in Chino, California, built in 1988. Lead–acid batteries have a specific energy of only 0.18 MJ/kg and would therefore not be a viable automobile option apart from providing the small amount of energy needed to start an engine. It also has a poor energy density at around 0.25 MJ/L. The advantages of the lead–acid battery technology are low cost and high power density.

18.1.7.3 Lithium-Ion

Lithium-ion and lithium polymer batteries, although primarily used in the portable electronics market, are likely to have future use in many other applications. The cathode in these batteries is a lithiated metal oxide ($LiCoO_2$, $LiMO_2$, etc.) and the anode is made of graphitic carbon with a layer structure. The electrolyte consists of lithium salts (such as $LiPF_6$) dissolved in organic carbonates; an example of Li-ion battery chemistry is

$$\mathrm{Li}_x\mathrm{C} + \mathrm{Li}_{1-x}\mathrm{CoO}_2 \leftrightarrow \mathrm{LiCoO}_2 + \mathrm{C}.$$

When the battery is charged, lithium atoms in the cathode become ions and migrate through the electrolyte toward the carbon anode where they combine with external electrons and are deposited between carbon layers as lithium atoms. This process is reversed during discharge. The lithium polymer variation replaces the electrolyte with a plastic film that does not conduct electricity but allows ions to pass through it. The 60°C operating temperature requires a heater, reducing overall efficiency slightly.

Lithium-ion batteries have a high energy density of about 0.72 MJ/L and have low internal resistance; they will achieve efficiencies in the 90% range and above. They have an energy density of around 0.72 MJ/kg. Their high energy efficiency and energy density make lithium-ion batteries excellent candidates for storage in all four applications considered here: utility shaping, power quality, distributed generation, and automotive.

18.1.7.4 Nickel–Cadmium

Nickel–cadmium (NiCd) batteries operate according to the chemistry:

$$2\,\mathrm{NiOOH} + 2\,\mathrm{H_2O} + \mathrm{Cd} \leftrightarrow 2\,\mathrm{Ni(OH)_2} + \mathrm{Cd(OH)_2}.$$

NiCd batteries are not common for large stationary applications. They have a specific energy of about 0.27 MJ/kg, an energy density of 0.41 MJ/L and an efficiency of about 75%. Alaska's Golden Valley Electric Association commissioned a 40-MW/290-GJ nickel–cadmium battery in 2003 to improve reliability and to supply power for essentials during outages.[5] Resistance to cold and relatively low cost were among the deciding factors for choosing the NiCd chemistry.

Cadmium is a toxic heavy metal and there are concerns relating to the possible environmental hazards associated with the disposal of NiCd batteries. In November 2003, the European Commission adopted a proposal for a new battery directive that includes recycling targets of 75% for NiCd batteries. However, the possibility of a ban on rechargeable batteries made from nickel–cadmium still remains and hence the long-term viability and availability of NiCd batteries continues to be uncertain. NiCd batteries can also suffer from "memory effect," where the batteries will only take full charge after a series of full discharges. Proper battery management procedures can help to mitigate this effect.

18.1.7.5 Nickel–Metal Hydride

The nickel–metal hydride (Ni*M*H) battery operates according to the chemistry:

$$M\text{H} + \text{NiOOH} \leftrightarrow M + \text{Ni(OH)}_2,$$

where *M* represents one of a large variety of metal alloys that serve to take up and release hydrogen. Ni*M*H batteries were introduced as a higher energy density and more environmentally friendly version of the nickel–cadmium cell. Modern nickel–metal hydride batteries offer up to 40% higher energy density than nickel–cadmium. There is potential for yet higher energy density, but other battery technologies (lithium-ion, in particular) may fill the same market sooner.

Nickel–metal hydride is less durable than nickel–cadmium. Cycling under heavy load and storage at high temperature reduces the service life. Nickel–metal hydride suffers from a higher self-discharge rate than the nickel–cadmium chemistry. Nickel–metal hydride batteries have a specific energy of 0.29 MJ/kg, an energy density of about 0.54 MJ/L and an energy efficiency of about 70%. These batteries have been an important bridging technology in the portable electronics and hybrid automobile markets. Their future is uncertain because other battery chemistries promise higher energy storage potential and cycle life.

18.1.7.6 Sodium–Sulfur

A sodium–sulfur (NaS) battery consists of a liquid (molten) sulfur positive electrode and liquid (molten) sodium negative electrode, separated by a solid beta-alumina ceramic electrolyte (Figure 18.3). The chemistry is as follows:

$$2\,\text{Na} + x\,\text{S} \leftrightarrow \text{Na}_2\text{S}_x.$$

When discharging, positive sodium ions pass through the electrolyte and combine with the sulfur to form sodium polysulfides. The variable x in the equation is equal to 5 during early discharging, but after free sulfur has been exhausted a more sodium-rich mixture of polysulfides with lower average values of x develops. This process is reversible as charging causes sodium polysulfides in the positive electrode to

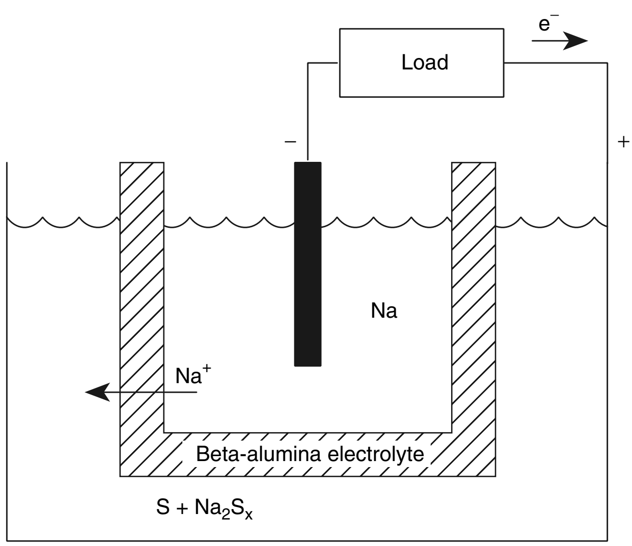


FIGURE 18.3 Sodium–sulfur battery showing discharge chemistry. The sodium (Na) and sulfur (S) electrodes are both in a liquid state and are separated by a solid, beta-alumina ceramic electrolyte that allows only sodium ions to pass. Charge is extracted from the electrolytes with metal contacts; the positive contact is the battery wall.

release sodium ions that migrate back through the electrolyte and recombine as elemental sodium. The battery operates at about 300°C. NaS batteries have a high energy density of around 0.65 MJ/L and a specific energy of up to 0.86 MJ/kg. These numbers would indicate an application in the automotive sector, but warm-up time and heat-related accident risk make its use there unlikely. The efficiency of this battery chemistry can be as high as 90% and would be suitable for bulk storage applications while simultaneously allowing effective power smoothing operations.[6]

18.1.7.7 Zebra

Zebra is the popular name for the sodium–nickel-chloride battery chemistry:

$$NiCl_2 + 2\,Na \leftrightarrow Ni + 2\,NaCl.$$

Zebra batteries are configured similarly to sodium–sulfur batteries (see Figure 18.3), and also operate at about 300°C. Zebra batteries boast a greater than 90% energy efficiency, a specific energy of up to 0.32 MJ/kg and an energy density of 0.49 MJ/L.[7] Its tolerance for a wide range of operating temperature and high efficiency, coupled with a good energy density and specific energy, make its most probable application the automobile sector, and as of 2003 Switzerland's MES-DEA is pursuing this application aggressively.[8] Its high energy efficiency also makes it a good candidate for the utility sector.

18.1.7.8 Flow Batteries

Most secondary batteries use electrodes both as an interface for gathering or depositing electrons, and as a storage site for the products or reactants associated with the battery's chemistry. Consequently, both energy and power density are tied to the size and shape of the electrodes. Flow batteries store and release electrical energy by means of reversible electrochemical reactions in two liquid electrolytes. An electrochemical cell has two compartments—one for each electrolyte—physically separated by an ion exchange membrane. Electrolytes flow into and out of the cell through separate manifolds and undergo chemical reaction inside the cell, with ion or proton exchange through the membrane and electron exchange through the external electric circuit. The chemical energy in the electrolytes is turned into electrical energy and vice versa for charging. They all work in the same general way but vary in chemistry of electrolytes.[9]

There are some advantages to using the flow battery over a conventional secondary battery. The capacity of the system is scaleable by simply increasing the amount of solution. This leads to cheaper installation costs as the systems get larger. The battery can be fully discharged with no ill effects and has little loss of electrolyte over time. Because the electrolytes are stored separately and in large containers (with a low surface area to volume ratio), flow batteries show promise to have some of the lowest self-discharge rates of any energy storage technology available.

Poor energy densities and specific energies remand these battery types to utility-scale power shaping and smoothing, although they might be adaptable for distributed-generation use. There are three types of flow batteries that are closing in on commercialization: vanadium redox, polysulfide bromide, and zinc bromide.

18.1.7.8.1 Vanadium Redox

The vanadium redox flow battery (VRB) was pioneered at the University of New South Wales, Australia, and has shown potentials for long cycle life and energy efficiencies of over 80% in large installations.[10] The VRB uses compounds of the element vanadium in both electrolyte tanks. The reaction chemistry at the positive electrode is:

$$V^{5+} + e^- \leftrightarrow V^{4+},$$

and at the negative electrode,

$$V^{2+} \leftrightarrow V^{3+} + e^{-}.$$

Using vanadium compounds on both sides of the ion-exchange membrane eliminates the possible problem of cross-contamination of the electrolytes and makes recycling easier.[11] As of 2005, two small, utility-scale VRB installations are operating, one 2.9-GJ unit on King Island, Australia and one 7.2-GJ unit in Castle Valley, Utah.

18.1.7.8.2 Polysulfide Bromide

The polysulfide bromide battery (PSB) utilizes two salt solution electrolytes, sodium bromide (NaBr) and sodium polysulfide (Na_2S_x). PSB electrolytes are separated in the battery cell by a polymer membrane that only passes positive sodium ions. The chemistry at the positive electrode is

$$\mathrm{NaBr_3} + 2\,\mathrm{Na^+} + 2\,e^{-} \leftrightarrow 3\,\mathrm{NaBr},$$

and at the negative electrode,

$$2\,\mathrm{Na_2S_2} \leftrightarrow \mathrm{Na_2S_4} + 2\,\mathrm{Na^+} + 2\,e^{-}.$$

The PSB battery is being developed by Canada's VRB Power Systems, Inc.[12] This technology is expected to attain energy efficiencies of approximately 75%.[13] Although the salt solutions themselves are only mildly toxic, a catastrophic failure by one of the tanks could release highly toxic bromine gas. Nevertheless, the Tennessee Valley Authority released a finding of no significant impact for a proposed 430-GJ facility and deemed it safe.[14]

18.1.7.8.3 Zinc Bromide

In each cell of a zinc bromide (ZnBr) battery, two different electrolytes flow past carbon-plastic composite electrodes in two compartments separated by a microporous membrane. Chemistry at the positive electrode follows the equation:

$$\mathrm{Br_2}(aq) + 2\,e^{-} \leftrightarrow 2\,\mathrm{Br},$$

and at the negative electrode:

$$\mathrm{Zn} \leftrightarrow \mathrm{Zn^{2+}} + 2e^{-}.$$

During discharge, Zn and Br combine into zinc bromide. During charge, metallic zinc is deposited as a thin-film on the negative electrode. Meanwhile, bromine evolves as a dilute solution on the other side of the membrane, reacting with other agents to make thick bromine oil that sinks to the bottom of the electrolytic tank. During discharge, a pump mixes the bromine oil with the rest of the electrolyte. The zinc bromide battery has an energy efficiency of nearly 80%.[15]

Exxon developed the ZnBr battery in the early 1970 s. Over the years, many GJ-scale ZnBr batteries have been built and tested. Meidisha demonstrated a 1-MW/14-GJ ZnBr battery in 1991 at Kyushu Electric Power Company. Some GJ-scale units are now available preassembled, complete with plumbing and power electronics.

18.1.7.9 Electrolytic Hydrogen

Diatomic, gaseous hydrogen (H_2) can be manufactured with the process of electrolysis; an electric current applied to water separates it into components O_2 and H_2. The oxygen has no inherent energy value, but the HHV of the resulting hydrogen can contain up to 90% of the applied electric energy, depending on the technology.[16] This hydrogen can then be stored and later combusted to provide heat or work, or to power a fuel cell (see Chapter 26).

The gaseous hydrogen is low density and must be compressed to provide useful storage. Compression to a storage pressure of 350 bar, the value usually assumed for automotive technologies, consumes up to

12% of the hydrogen's HHV if performed adiabatically, although the loss approaches a lower limit of 5% as the compression approaches an isothermal ideal.[17] Alternatively, the hydrogen can be stored in liquid form, a process that costs about 40% of HHV using current technology, and that at best would consume about 25%. Liquid storage is not possible for automotive applications, because mandatory boil-off from the storage container cannot be safely released in closed spaces (i.e., garages).

Hydrogen can also be bonded into metal hydrides using an absorption process. The energy penalty of storage may be lower for this process, which requires pressurization to only 30 bar. However, the density of the metal hydride can be between 20 and 100 times the density of the hydrogen stored. Carbon nanotubes have also received attention as a potential hydrogen storage medium.[18] Hydrogen storage technologies are covered in more detail in Chapter 25.

18.1.8 Mechanical Energy Storage

18.1.8.1 Pumped Hydro

Pumped hydro is the oldest and largest of all of the commercially available energy storage technologies, with existing facilities up to 1000 MW in size. Conventional pumped hydro uses two water reservoirs, separated vertically. Energy is stored by moving water from the lower to the higher reservoir, and extracted by allowing the water to flow back to the lower reservoir. Energy is stored according to the fundamental physical principle of potential energy. To calculate the stored energy, E_s, in joules, use the formula:

$$E_s = Vdgh,$$

where V is the volume of water raised (m^3), d is the density of water (1000 kg/m^3), g is the acceleration of gravity (9.8 m/s^2), and h is the elevation difference between the reservoirs (m) often referred to as the *head*.

Though pumped hydro is by nature a mechanical energy storage technology, it is most commonly used for electric utility shaping. During off-peak hours electric pumps move water from the lower reservoir to the upper reservoir. When required, the water flow is reversed to generate electricity. Some high dam hydro plants have a storage capability and can be dispatched as pumped hydro storage. Underground pumped storage, using flooded mine shafts or other cavities, is also technically possible but probably prohibitively expensive. The open sea can also be used as the lower reservoir if a suitable upper reservoir can be built at close proximity. A 30-MW seawater pumped hydro plant was first built in Yanbaru, Japan in 1999.

Pumped hydro is most practical at a large scale with discharge times ranging from several hours to a few days. There is over 90 GW of pumped storage in operation worldwide, which is about 3% of global electric generation capacity.[19] Pumped storage plants are characterized by long construction times and high capital expenditure. Its main application is for utility shaping. Pumped hydro storage has the limitation of needing to be a very large capacity to be cost-effective, but can also be used as storage for a number of different generation sites.

Efficiency of these plants has greatly increased in the last 40 years. Pumped storage in the 1960s had efficiencies of 60% compared with 80% for new facilities. Innovations in variable speed motors have helped these plants operate at partial capacity, and greatly reduced equipment vibrations, increasing plant life.

18.1.8.2 Compressed Air

A relatively new energy storage concept that is implemented with otherwise mature technologies is compressed air energy storage (CAES). CAES facilities must be coupled with a combustion turbine, so are actually a hybrid storage/generation technology.

A conventional gas turbine consists of three basic components: a compressor, combustion chamber, and an expander. Power is generated when compressed air and fuel burned in the combustion chamber drive turbine blades in the expander. Approximately 60% of the mechanical power generated by the expander is consumed by the compressor supplying air to the combustion chamber.

A CAES facility performs the work of the compressor separately, stores the compressed air, and at a later time injects it into a simplified combustion turbine. The simplified turbine includes only the combustion chamber and the expansion turbine. Such a simplified turbine produces far more energy than a conventional turbine from the same fuel, because there is potential energy stored in the compressed air. The fraction of output energy beyond what would have been produced in a conventional turbine is attributable to the energy stored in compression.

The net efficiency of storage for a CAES plant is limited by the heat energy loss occurring at compression. The overall efficiency of energy storage is about 75%.[20]

CAES compressors operate on grid electricity during off-peak times, and use the expansion turbine to supply peak electricity when needed. CAES facilities cannot operate without combustion because the exhaust air would exit at extremely low temperatures causing trouble with brittle materials and icing. If 100% renewable energy generation is sought, biofuel could be used to fuel the gas turbines. There might still be other emissions issues but the system could be fully carbon neutral.

The compressed air is stored in appropriate underground mines, caverns created inside salt rocks or possibly in aquifers. The first commercial CAES facility was a 290-MW unit built in Hundorf, Germany in 1978. The second commercial installation was a 110-MW unit built in McIntosh, Alabama in 1991. The third commercial CAES is a 2,700-MW plant under construction in Norton, Ohio. This nine-unit plant will compress air to about 100 bar in an existing limestone mine 2200 ft. (766 m) underground.[21] The natural synergy with geological caverns and turbine prime movers dictate that these be on the utility scale.

18.1.8.3 Flywheels

Most modern flywheel energy storage systems consist of a massive rotating cylinder (comprised of a rim attached to a shaft) that is supported on a stator by magnetically levitated bearings that eliminate bearing wear and increase system life. To maintain efficiency, the flywheel system is operated in a low vacuum environment to reduce drag. The flywheel is connected to a motor/generator mounted onto the stator that, through some power electronics, interact with the utility grid.

The energy stored in a rotating flywheel, in joules, is given by

$$E = \frac{1}{2} I \omega^2$$

where I is the flywheel's moment of inertia (kg m^2), and ω is its angular velocity (s^{-2}). I is proportional to the flywheel's mass, so energy is proportional to mass and the square of speed. In order to maximize energy capacity, flywheel designers gravitate toward increasing the flywheel's maximum speed rather than increasing its moment of inertia. This approach also produces flywheels with the higher specific energy.

Some of the key features of flywheels are low maintenance, a cycle life of better than 10,000 cycles, a 20-year lifetime and environmentally friendly materials. Low-speed, high-mass flywheels (relying on I for energy storage) are typically made from steel, aluminum, or titanium; high-speed, low-mass flywheels (relying on ω for energy storage) are constructed from composites such as carbon fiber.

Flywheels can serve as a short-term ride-through before long-term storage comes online. Their low energy density and specific energy limit them to voltage regulation and UPS capabilities. Flywheels can have energy efficiencies in the upper 90% range depending on frictional losses.

18.1.9 Direct Thermal Storage

Direct thermal technologies, although they are storing a lower grade of energy (heat, rather than electrical or mechanical energy) can be useful for storing energy from systems that provide heat as a native output (e.g., solar thermal, geothermal), or for applications where the energy's commodity value is heat (e.g., space heating, drying).

Although thermal storage technologies can be characterized by specific energy and energy density like any other storage technology, they can also be characterized by an important, additional parameter: the delivery temperature range. Different end uses have more or less allowance for wide swings of the delivery temperature. Also, some applications require a high operating temperature that only some thermal storage media are capable of storing.

Thermal storage can be classified into two fundamental categories: sensible heat storage and latent heat storage. Applications that have less tolerance for temperature swings should utilize a latent heat technology.

Input to and output from heat energy storage is accomplished with heat exchangers. The discussion below focuses on the choice of heat storage materials; the methods of heat exchange will vary widely depending on properties of the storage material, especially its thermal conductivity. Materials with higher thermal conductivity will require a smaller surface area for heat exchange. For liquids, convection or pumping can reduce the need for a large heat exchanger. In some applications, the heat exchanger is simply the physical interface of the storage material with the application space (e.g., phase-change drywall, see below).

18.1.9.1 Sensible Heat

Sensible heat is the heat that is customarily and intuitively associated with a change in temperature of a massive substance. The heat energy, E_s, stored in such a substance is given by:

$$E_s = (T_2 - T_1)cM,$$

where c is the specific heat of the substance (J/kg °C) and M is the mass of the substance (kg); T_1 and T_2 are the initial and final temperatures, respectively (°C). The specific heat c is a physical parameter measured in units of heat per temperature per mass: substances with the ability to absorb heat energy with a relatively small increase in temperature (e.g., water) have a high specific heat, whereas those that get hot with only a little heat input (e.g., lead) have a low specific heat. Sensible heat storage is best accomplished with materials having a high specific heat.

18.1.9.1.1 Liquids

Sensible heat storage in a liquid is, with very few exceptions, accomplished with water. Water is unique among chemicals in having an abnormally high specific heat of 4,186 J/kg K, and furthermore has a reasonably high density. Water is also cheap and safe. It is the preferred choice for most nonconcentrating solar thermal collectors.

Liquids other than water may need to be chosen if the delivery temperature must be higher than 100°C, or if the system temperature can fall below 0°C. Water can be raised to temperatures higher than 100°C, but the costs of storage systems capable of containing the associated high pressures are usually prohibitive. Water can be mixed with ethylene glycol or propylene glycol to increase the useful temperature range and prevent freezing.

When a larger temperature range than that afforded by water is required, mineral, synthetic, or silicone oils can be used instead. The tradeoffs for the increased temperature range are higher cost, lower specific heat, higher viscosity (making pumping more difficult), flammability, and, in some cases, toxicity.

For very high temperature ranges, salts are usually preferred that balance a low specific heat with a high density and relatively low cost. Sodium nitrate has received the most prominent testing for this purpose in the U.S. Department of Energy's Solar Two Project located in Barstow, California.

Liquid sensible heat storage systems are strongly characterized not just by the choice of heat-transfer fluid, but also by the system architecture. Two-tank systems store the cold and hot liquids in separate tanks (Figure 18.4). Thermocline systems use a single tank with cold fluid entering or leaving the bottom of the tank and hot fluid entering or leaving the top (Figure 18.5). Thermocline systems can be particularly low cost because they minimize the required tank volume, but require careful design to prevent mixing of the hot and cold fluid.

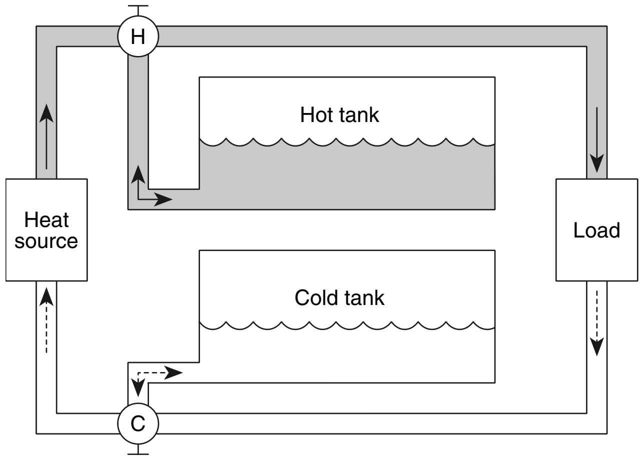


FIGURE 18.4 Two-tank thermal storage system; hot water is shown in gray and cold water is shown in white. When the heat source is producing more output than required for the load, valve H is turned to deposit hot liquid in the tank. When it is producing less than required for the load, the valve is turned to provide supplemental heat from the storage tank. Note that each tank must be large enough to hold the entire fluid capacity of the system.

One particularly interesting application of the thermocline concept is nonconvecting, salinity-gradient solar ponds that employ the concept in reverse. Solar ponds are both an energy collection and energy storage technology. Salts are dissolved in the water to introduce a density gradient, with the densest (saltiest) water on the bottom and lightest (freshest) on top. Solar radiation striking the dark bottom of the pond heats the densest water, but convection of the heated water to the top cannot occur because the density gradient prevents it. Salinity-gradient ponds can generate and store hot water at temperatures approaching 95°C.[22]

18.1.9.1.2 Solids

Storage of sensible heat in solids is usually most effective when the solid is in the form of a bed of small units, rather than a single mass. The reason is that the surface-to-volume ratio increases with the number

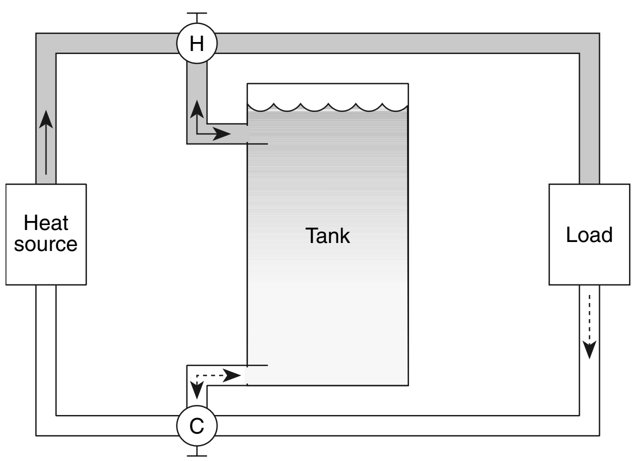


FIGURE 18.5 Thermocline storage tank. Thermocline storage tanks are tall and narrow to encourage the gravity-assisted separation of hot and cold fluid, and include design features (especially at the input/output connectors) to prevent mixing in the stored fluid.

of units, so that heat transfer to and from the storage device is faster for a greater number of units. Energy can be stored or extracted from a thermal storage bed by passing a gas (such as air) through the bed. Thermal storage beds can be used to extract and store the latent heat of vaporization from water contained in flue gases.

Although less effective for heat transfer, monolithic solid storage has been successfully used in architectural applications and solar cookers.

18.1.9.2 Latent Heat

Latent heat is absorbed or liberated by a phase change or a chemical reaction and occurs at a constant temperature. A phase change means the conversion of a homogenous substance among its various solid, liquid, or gaseous phases. One very common example is boiling water on the stovetop: though a substantial amount of heat is absorbed by the water in the pot, the boiling water maintains a constant temperature of 100°C. The latent heat, E_s, stored through a phase change is:

$$E_s = lM,$$

where M is the mass of material undergoing a phase change (kg), and l is the latent heat of vaporization (for liquid–gas phase changes) or the latent heat of fusion (for solid–liquid phase changes), in J/kg; l is measured in units of energy per mass. Conservation of energy dictates that the amount of heat absorbed in a given phase change is equal to the amount of heat liberated in the reverse phase change.

Although the term *phase change* is used here to refer only to straightforward freezing and melting, many sources use the term *phase-change materials* or *PCMs* to refer to any substance storing latent heat (including those described in Section 18.1.9.6 and Section 18.1.9.7, as well.)

18.1.9.2.1 Phase Change

Practical energy storage systems based on a material phase change are limited to solid–solid and solid–liquid phase changes. Changes involving gaseous phases are of little interest due to the expense associated with containing a pressurized gas, and difficulty of transferring heat to and from a gas.

Solid–solid phase changes occur when a solid material reorganizes into a different molecular structure in response to temperature. One particularly interesting example is lithium sulfate (Li_2SO_4) which undergoes a change from a monoclinic structure to a face-centered cubic structure at 578°C, absorbing 214 J/g in the process, more than most solid–liquid phase changes.[23]

Some common chemicals, their melting points and heats of fusion are listed in Table 18.4. Fatty acids and paraffins received particular attention in the 1990s as candidate materials for the heat storage component of phase-change drywall, a building material designed to absorb and release heat energy near room temperature for the purpose of indoor temperature stabilization.[24] In this application, solids in the drywall maintain the material's structural integrity even though the phase-change materials are transitioning between solid and liquid states.

18.1.9.2.2 Hydration–Dehydration

In this process, a salt or similar compound forms a crystalline lattice with water below a "melting-point" temperature, and at the melting point the crystal dissolves in its own water of hydration. Sodium sulfate (Na_2SO_4) is a good example, forming a lattice with ten molecules of water per molecule of sulfate ($Na_2SO_4 \cdot 10H_2O$) and absorbing 241 J/g at 32°C.[25]

Hydration–dehydration reactions have not found significant application in renewable energy systems, although they, too, have been a candidate for phase-change drywall.

18.1.9.2.3 Chemical Reaction

A wide variety of reversible chemical reactions are available that release and absorb heat (see, for example, Hanneman, Vakil, and Wentorf[26]). The principal feature of this category of latent heat storage technologies is the ability to operate at extremely high temperatures, in some cases over 900°C.

TABLE 18.4 Melting Points and Heats of Fusion for Solid–Liquid Phase Changes

	Melting Point°C	Heat of Fusion J/g
Aluminum bromide	97	42
Aluminum iodide	191	81
Ammonium bisulfate	144	125
Ammonium nitrate	169	77
Ammonium thiocyanate	146	260
Anthracine	96	105
Arsenic tribromide	32	37
Beeswax	62	177
Boron hydride	99	267
Metaphosphoric acid	43	107
Naphthalene	80	149
Naphthol	95	163
Paraffin	74	230
Phosphoric acid	70	156
Potassium	63	63
Potassium thiocyanate	179	98
Sodium	98	114
Sodium hydroxide	318	167
Sulfur	110	56
Tallow	76	198
Water	0	335

Source: From Kreith, F. and Kreider J.F., *Principles of Solar Engineering*, Taylor & Francis, 1978. With permission

Extremely high temperature applications have focused primarily on fossil and advanced nuclear applications; to date, none of these chemical methods of heat storage have been deployed in commercial renewable energy applications.

18.1.10 Thermochemical Energy Storage

This section provides an overview of biomass storage technologies from an energetic perspective only. Additional details on biomass fuels are presented in Chapter 25.

18.1.10.1 Biomass Solids

Plant matter is a storage medium for solar energy. The input mechanism is photosynthesis conversion of solar radiation into biomass. The output mechanism is combustion of the biomass to generate heat energy.

Biologists measure the efficiency of photosynthetic energy capture with the metric net primary productivity (NPP), which is usually reported as a yield in units similar to dry Mg/ha-yr (dry metric tons per hectare per year). However, to enable comparisons of biomass with other solar energy storage technologies, it is instructive to estimate a solar efficiency by multiplying the NPP by the biomass heating value (e.g., MJ/dry Mg) and then dividing the result by the average insolation at the crop's location (e.g., MJ/ha-yr). The solar efficiency is a unitless value describing the fraction of incident solar energy ultimately available as biomass heating value. Most energy crops capture between 0.2 and 2% of the incident solar energy in heating value of the biomass; Table 18.5 shows examples of solar efficiencies estimated for a number of test crops.

The principal method for extracting useful work or electricity from biomass solids is combustion. Therefore, the solar efficiencies listed in Table 18.5 need to be multiplied by the efficiency of any associated combustion process to yield a net solar efficiency. For example, if a boiler-based electric generator extracts 35% of the feedstock energy as electricity, and the generator is sited at a switchgrass plantation achieving 0.30% solar capture efficiency on a mass basis, the electric plant has a net solar

TABLE 18.5 Primary Productivity and Solar Efficiency of Biomass Crops

Location	Crop	Yield (dry Mg/ha-yr)	Average Insolation (W/m^2)	Solar Efficiency(%)
Alabama	Johnsongrass	5.9	186	0.19
Alabama	Switchgrass	8.2	186	0.26
Minnesota	Willow and hybrid poplar	8–11	159	0.30–0.41
Denmark	Phytoplankton	8.6	133	0.36
Sweden	Enthropic lake angiosperm	7.2	106	0.38
Texas	Switchgrass	8–20	212	0.22–0.56
California	*Euphorbia lathyris*	16.3–19.3	212	0.45–0.54
Mississippi	Water hyacinth	11.0–33.0	194	0.31–0.94
Texas	Sweet sorghum	22.2–40.0	239	0.55–0.99
Minnesota	Maize	24.0	169	0.79
West Indies	Tropical marine angiosperm	30.3	212	0.79
Israel	Maize	34.1	239	0.79
Georgia	Subtropical saltmarsh	32.1	194	0.92
Congo	Tree plantation	36.1	212	0.95
New Zealand	Temperate grassland	29.1	159	1.02
Marshall Islands	Green algea	39.0	212	1.02
New South Wales	Rice	35.0	186	1.04
Puerto Rico	*Panicum maximum*	48.9	212	1.28
Nova Scotia	Sublittoral seaweed	32.1	133	1.34
Colombia	Pangola grass	50.2	186	1.50
West Indies	Tropical forest, mixed ages	59.0	212	1.55
California	Algae, sewage pond	49.3–74.2	218	1.26–1.89
England	Coniferous forest, 0–21 years	34.1	106	1.79
Germany	Temperate reedswamp	46.0	133	1.92
Holland	Maize, rye, two harvests	37.0	106	1.94
Puerto Rico	*Pennisetum purpurcum*	84.5	212	2.21
Hawaii	Sugarcane	74.9	186	**2.24**
Java	Sugarcane	86.8	186	2.59
Puerto Rico	Napier grass	106	212	2.78
Thailand	Green algae	164	186	4.90

Source: From Klass, D. L., Biomass for Renewable Energy, Fuels, and Chemicals, Academic Press, San Diego, CA, 1998. With permission.

efficiency of $0.30\% \times 35\% = 0.11\%$. Because biomass is a low-efficiency collector of solar energy, it is very land intensive compared to photovoltaic or solar thermal collectors that deliver energy at solar efficiencies over 20% (see Chapter 19 and Chapter 21 for a full discussion). However, the capacity of land to store standing biomass over time is extremely high, with densities up to several hundred Mg/ha (and therefore several thousand GJ/ha), depending on the forest type. Standing biomass can serve as long-term storage, although multiple stores need to be used to accommodate fire risk. For short-term storage, woody biomass may be dried, and is frequently chipped or otherwise mechanically treated to create a fine and homogenous fuel suitable for burning in a wider variety of combustors.

18.1.10.2 Ethanol

Biomass is a more practical solar energy storage medium if it can be converted to liquid form. Liquids allow for more convenient transportation and combustion, and enable extraction on demand (through reciprocating engines) rather than through a less dispatchable, boiler- or turbine-based process. This latter property also enables its use in automobiles.

Biomass grown in crops or collected as residue from agricultural processes consists principally of cellulose, hemicellulose, and lignin. The sugary or starchy by-products of some crops such as sugarcane, sugar beet, sorghum, molasses, corn, and potatoes can be converted to ethanol through fermentation processes, and these processes are the principal source of ethanol today. Starch-based ethanol production is low efficiency, but does succeed in transferring about 16% of the biomass heating value to the ethanol fuel.[27]

When viewed as a developing energy storage technology, ethanol derived from cellulose shows much more promise than the currently prevalent starch-based ethanol.[28] Cellulosic ethanol can be manufactured with two fundamentally different methods: either the biomass is broken down to sugars using a hydrolysis process, and then the sugars are subjected to fermentation; or the biomass is gasified (see below), and the ethanol is subsequently synthesized from this gas with a thermochemical process. Both processes show promise to be far cheaper than traditional ethanol manufacture via fermentation of starch crops, and will also improve energy balances. For example, it is estimated that dry sawdust can yield up to 224 L/Mg of ethanol, thus recovering about 26% of the higher heating value of the sawdust.[29] Because the ethanol will still need to be combusted in a heat engine, the gross, biomass-to-useful-work efficiency will be well below this. In comparison, direct combustion of the biomass to generate electricity makes much more effective use of the biomass as an energy storage medium. Therefore, the value of ethanol as an energy storage medium lies mostly in the convenience of its liquid (rather than solid) state.

18.1.10.3 Biodiesel

As starch-based ethanol is made from starchy by-products, most biodiesel is generated from oily by-products. Some of the most common sources are rapeseed oil, sunflower oil, and soybean oil. Biodiesel yields from crops like these range from about 300 to 1000 kg/ha-yr, but the crop as a whole produces about 20 Mg/ha-yr, meaning that the gross solar capture efficiency for biodiesel from crops ranges between 1/20 and 1/60 the solar capture efficiency of the crop itself. Because of this low solar-capture efficiency, biomass cannot be the principal energy storage medium for transportation needs.[30]

Biodiesel can also be manufactured from waste vegetable or animal oils; however, in this case, the biodiesel is not functioning per se as a solar energy storage medium, so is not further treated in this work.

18.1.10.4 Syngas

Biomass can be converted to a gaseous state for storage, transportation, and combustion (or other chemical conversion).[31] Gasification processes are grouped into three different classes: *pyrolysis* is the application of heat in anoxic conditions; *partial oxidation* is combustion occurring in an oxygen-starved environment; *reforming* is the application of heat in the presence of a catalyst. All three processes form syngas, a combination of methane, carbon monoxide, carbon dioxide and hydrogen. The relative abundances of the gaseous products can be controlled by adjusting heat, pressure, and feed rates. The HHV of the resulting gas can contain up to 78% of the original HHV of the feedstock, if the feedstock is dry.[29] Compositions and heating values of two example syngases are listed in Table 18.3.

The equivalent of up to 10% of the gas HHV will be lost when the gas is pressurized for transportation and storage. Even with this loss, gasification is a considerably more efficient method than ethanol manufacture for transferring stored solar energy to a nonsolid medium.

18.2 Advanced Concepts in Transmission and Distribution

Robert Pratt, Christopher P. Schaber, and Steve Widergren

18.2.1 Introduction

Global prosperity depends upon efficient and affordable energy. Without a major shift in the way energy systems are planned, built, and operated, the world will invest trillions of dollars in conventional electric infrastructure over the next 20 years to meet expected growth.[32,33]

The future electric system will incorporate information technology in revolutionary ways, much as information technology has transformed other aspects of business because, fundamentally, "bits are cheaper than iron." The impact of the information age on the electric system will allow nations to realize the benefits already achieved by leading-edge industries that use real-time information, distributed e-business systems and market efficiencies to minimize the need for inventory and infrastructure and to maximize productivity, efficiency, and reliability.

With the help of information technologies, the creation of a distributed, yet integrated, system will empower consumers to participate in energy markets—the key to stabilizing prices. Market participants from utilities to new third parties to consumers will create value by developing and deploying solutions that cross enterprise and regulatory boundaries. At the same time, this transformation of the energy system addresses the urgent need to enhance national security. A distributed, network-based electric system will reduce single-point vulnerabilities and allow the grid to become "self-healing" by incorporating autonomic system reconfiguration in response to human-caused or natural disruptions.

Deploying information technology in a highly connected world will maximize the use of existing assets and minimize the need for new assets. To do this, the future electric system will:

- Provide the incentive for customer and third-party assets to collaborate with existing grid assets to control costs and improve reliability by revealing the true time- and location-dependent value of electricity
- Provide the basis for collaboration by allowing the revealed values to be shared in real time by leveraging broadband communications that are rapidly becoming ubiquitous
- Provide the means to take advantage of the opportunities for collaboration and capture value in return through rapid advances in distributed controls and e-business applications

Much of the challenge and opportunity presented by these solutions is that their value proposition often involves crossing enterprise and regulatory boundaries. This means that nontechnical challenges must be addressed to support policy changes and business models that allow new technology and operational practices to flourish.

Information technology will profoundly transform the planning and operation of the power grid, just as it has changed business, education, and entertainment. It will form the "nervous system" that integrates new distributed technologies—demand response, distributed generation, and storage—with traditional grid generation, transmission, and distribution assets to share responsibility for managing the grid as a collaborative "society" of devices.[35–38]

18.2.2 Technology Advancements for Electricity Delivery

The following sections describe new technologies, or concepts that are becoming more economically feasible, for application to improved operation of the electric system. These are summarized in three general categories: distributed resources, transmission and distribution resources, information and communications technology, and control, applications, and operational paradigms.

18.2.2.1 Distributed Resources

Distributed resources are physical, capital assets that generate, store, or consume electric power at the fringes of the electric delivery system. Traditionally, these resources have not been broadly considered or deployed as transmission and distribution assets; however, significant efforts are underway to take advantage of such existing resources, and to provide motivation for future installations that can further benefit the system.[34] This section describes several of the resources that fall into this category:

End Use Control. Manual or automated control of end use systems to reduce or defer electric power use is an important element in active measures for enhancing delivery operations. Automatic load shedding (under-frequency, under-voltage), operator-initiated interruptible load, demand-side management programs, voltage reduction, and other load curtailment strategies have long been an integral part of coping with unforeseen contingencies as a last resort, and/or as a means of assisting the system during high-stress, overloaded conditions. Advances in load control technology will allow end use systems to play a more active role in the day-to-day operations of the electric system as well as create a more graceful response to emergency conditions.

Intelligent Building Systems. Intelligent building systems are emerging in factories, commercial buildings, and now residential facilities. These systems save energy by increasing efficient operations. Coordinated utilization of cooling, heating, and electricity in these establishments can significantly reduce energy consumption. Operated in a system that supports price-responsive demand programs, intelligent building systems can locally optimize operations to the benefit of larger system concerns.

System-Sensitive Appliances. System-sensitive appliances are emerging with the ability to measure and respond to frequency or voltage conditions or to react to simple signals, such as energy prices or emergency conditions. Washers, dryers, HVAC units, and other devices now come equipped with onboard intelligent systems that have the potential to respond to energy signals (see Control, Applications, and Operational Paradigms section below).

Distributed Generation (DG). Fuel cells, microturbines, reciprocating engines, and small renewable generation such as wind and solar power, are increasingly being deployed across the system. DG can supply local load or sell into the system and offers owners self-determination. As these distributed resources increase in number, they can become a significant resource for reliable system operations. Their vast numbers teamed with end use systems adds a further dimension to the advantages of controllable load discussed above.

Combined-Heating Power (CHP). Combined-heating-power and buildings-cooling-heating-power (BCHP) are systems that combine distributed fuel-fired generation with equipment that recovers waste heat for heating or cooling at the for a building, facility, or district, maximizing the utilization of fuel.

Renewable Generation. Renewable generation includes small wind-power, photovoltaic solar, and other renewable generation systems connected at the distribution level. Note that large wind farms are similar to central generation in that they do not displace the need for transmission and distribution assets. Unlike distributed fuel-fired generation, renewable generation is difficult to dispatch coincident with the needs of the grid, unless a storage system is involved.

Local Energy Storage. Systems that store energy locally for the benefit of end use systems can also be aggregated to the benefit of larger system operation concerns.

Electric Storage. Systems that store electricity for later discharge include traditional batteries, flow batteries, supercapacitors, flywheels, and extend to uninterruptible power supplies modified to support local system operations.

Thermal Storage. These systems store energy used for heat or cooling, usually in the form of hot or chilled water or ice, for later use in a building or an industrial process. The ability to "charge" to or "draw" from these resources provides flexibility to the timing of electricity use.

Direct DC End Use. Returning to the days of Edison, equipment within customer premises are emerging that distribute and use DC power directly (instead of converting it from AC with attendant cost and inefficiencies) such as computers and electronic equipment, fans, pumps and other equipment with variable speed drives, battery chargers, and uninterruptible power supplies. DC distribution is especially useful if onsite generation inherently involves DC power directly or at an intermediate stage, such as from photovoltaics, microturbines, and fuel cells.

Local Voltage Regulation. Transformers, capacitors, or power electronics equipment that maintains voltage at the customer meter at a specified nominal level, can reduce wasted consumption from over-voltage while allowing higher-than-nominal distribution voltages to increase system throughput during times of peak demand.

Advanced Energy Efficiency. Investments in improved equipment, building structure, or industrial processes can reduce the quantity of electricity consumed yet provide equivalent service. Efficiency benefits occur with sensitivity to the end use being served, but are indifferent to the needs of the grid.

18.2.2.2 Transmission and Distribution Resources

The transmission and distribution (T & D) system of the future must not only have increased capacity to support the market demand for energy transactions, it must also be flexible to adapt to alterations in energy delivery patterns. These patterns change at various timescales: hourly, daily, weekly, and seasonally. The delivery system must also adapt to operation patterns dictated by the evolving geographical distribution of load and generation. As generation siting and dispatch decision making become less predictable and demand-side resources play a larger role in system operations, new technologies that afford transmission and distribution planners a wider range of alternatives for deployment of power become more attractive. This section discusses some of the newer hardware technologies that are being researched and deployed to reinforce grid operations. The list presented is not exhaustive.[39–41]

Conductors. The siting of the new transmission will continue to be a major challenge. Getting the most out of existing rights of way minimizes the need for new lines and rights of way and can minimize the societal concerns associated with visual pollution and high-energy EMFs.

Advanced Composite Conductors. Usually, transmission lines contain steel-core cables that support strands of aluminum wires that are the primary conductors of electricity. New cores developed from composite materials are proposed to replace the steel core. A new core consisting of composite glass-fiber materials shows promise as stronger than steel-core aluminum conductors while being 50% lighter in weight with up to 2.5 times less sag. The reduced weight and higher strength equate to greater current-carrying capability. This technology can be integrated in the field by most existing reconductoring equipment.

High-Temperature Superconducting (HTSC) Technology. The conductors in HTSC devices operate at extremely low resistances. They require refrigeration (generally liquid nitrogen) to supercool ceramic superconducting materials. The technology is applicable to transmission lines, transformers, reactors, capacitors, and current limiters. HTSC cable occupies less space (AC transmission lines bundle three phase together; transformers and other equipment occupy smaller footprint for same level of capacity). Exposure to EMFs can be reduced with buried cables and this also counteracts visual pollution issues. Transformers can reduce or eliminate cooling oils that, if spilled, can damage the environment. For now, the maintenance costs remain high.

Below-Surface Cables. The underground cable advancements are employing fluid-filled polypropylene paper laminate (PPL) and extruded dielectric polyethylene (XLPE) cable technologies. Other approaches, such as gas-insulated transmission lines (GIL), are being researched and hold promise for future applications. Although there have been significant improvements in the technology, manufacturing costs remain high.

T & D Configurations. Advances are being made in the configuration of transmission lines. New design processes coupled with powerful computer programs can optimize the height, strength, and positioning of transmission towers, insulators, and associated equipment in order to meet engineering standards appropriate for the conductor (e.g., distance from ground and tension for a given set of weather parameters).

Tower and Pole Design Tools. A set of tools is emerging to analyze upgrades to existing transmission facilities or the installation of new facilities to increase their power-transfer capacity and reduce maintenance. Unused potential can be discovered in existing facilities to enhance upgrades, whereas new facilities can be engineered to closer design tolerances. Visualization techniques are adding the regulatory and public review cycles.

Modular Equipment. One way to gain flexibility for changing market and operational situations is to develop standards for the manufacture and integration of modular equipment. This can reduce overall the time and expense for transmission systems to adapt to the changing economic and reliability landscape.

T & F System Devices. Implemented throughout the system, these devices include capacitors, phase shifters, static VAR compensators (SVCs), thyristor-controlled series capacitors (TCSC), thyristor-controlled dynamic brakes, and other similar devices. Used to adjust system impedance, these devices can increase the delivery system's transfer capacity, support bus voltages by providing reactive power, or enhance dynamic or transient stability.

HVDC. With active control of real and reactive power transfer, HVDC can be modulated to damp oscillations or provide power-flow dispatch independent of voltage magnitudes or angles (unlike conventional AC transmission). HVDC runs independent of system frequency and can control the amount of power sent through the line. This latter benefit is the same as for FACTS devices discussed below. The high cost of converter equipment and its maintenance limits the application; however, as these costs continue to drop, the number of implementations will rise. Proposals have been made for DC application at the distribution level. The military continues to advance the technology in shipboard systems and mobile land-based facilities.

Flexible AC Transmission System (FACTS). FACTS devices use power electronics to adjust the apparent impedance of the system. Capacitor banks are applied at loads and substations to provide capacitive reactive power to offset the inductive reactive power typical of most power system loads and transmission lines. With long inter-tie transmission lines, series capacitors are used to reduce the effective impedance of the line. By adding thyristors to both of these types of capacitors, actively controlled reactive power is available using static VAR compensators and thyristor-controlled series compensators, which are shunt- and series-controlled capacitors, respectively. The thyristors are used to adjust the total impedance of the device by switching individual modules. Unified power-flow controllers (UPFCs) also fall into this category. Phase shifters are transformers configured to change the phase angle between buses; they are particularly useful for controlling the power flow on the transmission network. Adding thyristor control to the various tap settings of the phase-shifting transformer, permits continuous control of the effective phase angle (and thus control of power flow). As with HVDC, the power electronics used in these devices are expensive in price and maintenance, but costs are dropping.

T & D Energy Storage. The traditional function of an energy storage device is to save production costs by holding cheaply generated off-peak energy that can be dispatched during peak consumption periods. By virtue of its attributes, energy storage can also provide effective power system control with modest incremental investment. Different dispatch modes can be superimposed on the daily cycle of energy storage, with additional capacity reserved for the express purpose of providing these control functions. Storage at the bulk system level also benefits the integration of intermittent renewable resources such as wind and solar power. The loss of efficiency between converting electricity into and out of storage is an important consideration in deploying these resources.

Batteries. Large battery systems use converters to transform the DC in the storage device to the AC of the power grid. Converters also operate in the opposite direction to recharge the batteries. Battery converters use power electronics that, by the virtue of their ability to change the power exchange rapidly, can be utilized for a variety of real-time control applications ranging from enhancing transient stability to preconditioning the area control error for automatic generator control enhancement. The expense of manufacturing and maintaining batteries has limited their impact in the industry.

Superconducting Magnetic Energy Storage (SMES). SMES uses cryogenic technology to store energy by circulating current in a superconducting coil. SMES devices are efficient and compact because of their superconductive properties. As with the superconducting equipment mentioned above, SMES entails costs for the cooling system, special protection, and the specialized skills required to maintain the device.

Pumped Hydro and Compressed Air Storage. Pumped hydro consists of large ponds with turbines that can be run in either pump or generation modes. During periods of light load (e.g., night) excess, inexpensive capacity drives the pumps to fill the upper pond. During heavy-load periods, the water generates electricity into the grid. Compressed air storage uses the same principle except that large, natural underground vaults are used to store air under pressure during light-load periods. Pumped hydro, like any hydro generation project, requires significant space and has corresponding ecological impact. Compressed air storage systems require special geological formations.

Flywheels. Flywheels spin at high velocity to store energy. As with pumped hydro or compressed air storage, the flywheel is connected to a motor that either accelerates the flywheel to store energy or draws energy to generate electricity. Superconductivity technology can also be deployed to increase efficiency.

18.2.2.3 Information and Communications Technology

Automation and the ability to coordinate intelligent systems across wide areas are transforming all areas of our economy. The electric system is leveraging advances in generally available information and communication resources to enhance its operational effectiveness and incorporate new participants at the fringes of the system in operations.

Communication Media. Wireless, power line carrier, cable, fiber, and other forms of broadband communication to delivery system components and end use premises is beginning to carry the information needed to implement new applications such as, consumption data, sensor data on grid conditions, and e-commerce information including service offerings, prices, contract terms, and incentives.

Information Security. Technology that provides secure and reliable communications is essential to prevent intrusion by unauthorized parties or those without a need or right to know specific information.

Privacy and Authentication. Grid operators, consumers and businesses all demand technology that ensures the privacy of information is maintained and the identity of participants is unambiguous.

E-commerce Transactions. Technology that manages a multitude of small financial transactions in an auditable and traceable way will provide the financial incentive for day-to-day, moment-by-moment collaboration of end users with the grid (e.g., micropayments).

Real-Time Monitoring. The capability of the electricity grid is restricted through a combination of the limitations on individual devices and the composite capacity of the system. Improving monitoring to determine these limits in real time and to measure the system state directly can increase grid capability.

Power-System Device Sensors. Advancement in sensor technology is enabling dynamic ratings on the use of T&D resources. This includes the measurement of conductor sag, transformer coil temperature, and underground cable monitoring and diagnostics.

System-State Sensors. Technology advancements are improving the operational view of large regions of the power network. Power-system monitors collect essential signals (key power flows, bus voltages, alarms, etc.) from local monitors and make them available to site operation functions. This provides regional surveillance over important parts of the control system to verify system performance in real time. Sample rates and data quality is increasing as costs decrease. For example, phasor measurement units (PMUs) are synchronized digital transducers that stream data, in real time, to phasor data concentrators (PDC). The general functions and topology for this network resemble those for dynamic monitor networks. Data quality for phasor technology is trustworthy, and secondary processing of the acquired phasor information can provide a broad range of signal types.

Advanced Meters. Electric usage meters at customer premises are being replaced with technology that not only measures the energy usage, but also provides flexible interval energy monitoring and real-time power measurement to address needs such as peak usage, and power quality.

18.2.3 Control, Applications, and Operational Paradigms

The technologies summarized in the previous sections can be integrated into the electric system in a cooperative manner to address many applications: existing, proposed, and yet to be invented. There are many potential applications; some emerging ones are listed below and organized by end use, distribution, transmission, electricity service provider, and market operations.

18.2.3.1 End Use Operations and Planning

The management of end use resources in factories, commercial buildings, and residential facilities offer arguably the greatest area for advancement and contribution to system operations. Some important applications include the following:

- Customer information gateways: information portals that support two-way communications, including real-time price signals and long-term contracts, across the customer enterprise boundary to suppliers, grid operators, and other third parties.
- Demand response and energy management systems: controls that optimize the scheduling of energy use by appliances, equipment, and processes to minimize overall costs for electricity and respond to incentives from the service providers to curtail loads at times of peak demand or grid distress.
- System-sensitive appliances, equipment, and processes: controls that autonomously sense

disturbances in grid frequency and voltage, directing end use devices to immediately curtail their demand for periods of up to a few minutes (or, conversely, turning things like heating elements on to soak up momentary fluctuations of excess power) to prevent or arrest cascading blackouts. They also delay device restart after a voltage collapse or a blackout, easing service restoration by preventing the surge in demand from devices that have been without power for a while.

- Autonomous agents for power purchases: software serving as the trading agent on behalf of customers, searching market opportunities and arranging for the lowest cost suppliers, including subsequently auditing power bills to ensure fair play.

18.2.3.2 Distribution Operations and Planning

- Automated meter reading: a smart meter allows the customer and their suppliers to access the electric meter to enable automated meter reading for billing and other purposes as well as access to consumption data by end use energy management systems.
- Distribution automation: advanced distribution grid management systems that optimize supply voltage, manage peak demands by reconfiguring feeders to switch customers from one supply point to another, minimize the number of customers affected by outages, and accelerate outage restoration. This includes advanced protection schemes that adapt protection device settings based on operating conditions.
- Fault location and isolation: use advanced sensing and communications to better identify the location of faults, then invoke distribution automation capabilities to reconfigure feeders, safely reenergize as much of the affected region as possible, and inform work crews to repair the problem. Diagnostics and fault location tools for underground cables represent an important aspect of this area.
- Distribution capacity marketplaces: software that projects the need for new distribution capacity, posting the cost of required upgrades and allowing customers and third parties to offer distributed resource projects (such as distributed generation, efficiency, and demand response) to defer or avoid construction; then operating those resources through local market signals or incentives to manage net demand at the capacity constraint.
- Distributed generation for reactive power support and other ancillary services: software that manages customer-owned distributed generators to support the power grid by supplying reactive power in addition to real power and other ancillary services like spinning and nonspinning reserves, via direct dispatch based on prearranged contracts or indirect dispatch through markets or incentives.

18.2.3.3 Transmission Operations and Planning

- Transactive control for transmission grids: systems for trading real-time transmission rights, analogous to automated stock market trading systems, that allow market signals to cause suppliers and load serving entities to shift or curtail wheeling of power from one region to another, effectively reconfiguring the power grid in real time in response to a potential or pending crisis.
- Substation automation: device-level intelligence combines with high-speed, reliable communications to improve operations effectiveness. For example, transformer tap changes will coordinate with capacitor banks to perform voltage or reactive power control functions. More and better quality information is used locally for adaptive relay protection schemes that change their settings based on present system configuration information. Information exchanged between neighboring substations can be used to further coordinate system protection schemes and economic operation.

- Regional control: on a regional level, remedial action schemes can adaptively arm themselves based on updated information from several points in the system. Regional control centers can also coordinate the transmission operating configuration based on gathering better information from the field. For example, the results from system studies, such as contingency analysis and optimal power flow can make use of dynamic line ratings that are regularly updated based on weather forecasts and sag measurements from the field.

18.2.3.4 Electricity Service Provider Operations and Planning

- Differentiating customer service levels: systems that integrate demand response, distribution automation, and local distributed generation to supply premium reliability service to customers that require it, providing digital quality power to those willing to pay a premium without "gold-plating" the system for other customers satisfied with current service.
- Emergency end use curtailment: communications and control software that broadcasts emergency status information and leverages demand response capabilities to ration power to customers on a prearranged basis, providing power for critical customers and end uses while curtailing less critical demand to match the available supply; the intent is for everyone to have some power in a crisis rather than using rolling blackouts, for instance.

18.2.3.5 Market Operations and Forecasting

- Market operations: software that operates markets and incentives that engage distributed resources and demand response to manage peak demands and ancillary services in supply, transmission, and local distribution systems.
- Load forecasters: algorithms that provide continually updated, adaptive predictions of the demand for electricity at various levels in the system ranging from end use to distribution and transmission, as it responds to changes in conditions such as weather and price; based on techniques ranging from statistical methods and neural networks to engineering model parameter estimation.

18.2.4 System-Sensitive Appliances

The abundance of information, including price signals and grid conditions, will allow demand-response technologies to play a significant role in a virtual energy infrastructure. Consider that regardless of time of day or even time of year, about 20% of the load on the electric system is from consumer appliances that cycle on and off, such as heating, air conditioning, water heaters, and refrigerators. At the same time, generators also maintain steady operating reserves on "hot standby" that are equivalent to about 13% of the total load on the system in case problems suddenly arise . Why not find a way to reduce the demand of these appliances in times when operators would typically be dipping into the expensive cushion?

The (Grid Friendly™ controller) low-cost sensors embedded in appliances can sense grid conditions by monitoring the frequency of the system and provides automatic demand response in times of disruption. Within each interconnected operating region, a disturbance reflected in the system frequency is a universal indicator of serious imbalance between supply and demand that, if not relieved, can lead to a blackout. A simple computer chip installed in end use appliances can reduce the appliance's electricity demand for a few minutes or even a few seconds to allow the grid to stabilize. The controllers can be programmed to autonomously react in fractions of a second when a disturbance is detected whereas power plants can take minutes to come up to speed.[42]

As these controllers penetrate the appliance marketplace, important control issues must be addressed. For example, the control approach must deactivate appliances in a graduated manner so as not to "shock" the grid by dropping more load than necessary to rebalance supply and demand.[43] They must also be reactivated after the crisis in a similarly smooth fashion. It may be highly desirable to organize this response as a hierarchy ordered from the least to the highest priority end uses—from the least critical functions, such as air conditioning, to the most critical functions, such as communications, traffic control, water, sewage, and fuel pumps. Then, if a crisis persists, grid operators can allocate power to end uses simply by maintaining grid frequency at small increments below the normal range. Managing demand in the power grid in an "all-or-nothing" fashion will no longer be necessary.

By integrating the controllers with appliances at the factory, costs can be reduced to a few dollars per appliance. Done properly, consumers will not even notice the short interruption (by turning off the compressor in a refrigerator, but leaving the light on, for example). In the process, consumers become an integral part of power-grid operations and could even be rewarded for their participation in helping prevent a widespread outage. Therefore, without the need for any formal communications capability, appliances are transformed from being part of the problem to part of the solution. They now act as assets that form a much quicker and better safety net under the power grid, freeing up power plants from standby duty to increase competition, lower prices and meet future load growth. Moreover, because they act autonomously, no communication system is required beyond the power grid itself. However, when a communication system becomes available, the "smarts" are already on board the appliances to do much more sophisticated negotiation and control, such as reducing peak loads.

18.2.5 Markets + Control = The Transactive Network

Transactive (e.g., contract) networks and agent-based systems present an opportunity to implement process controls in which highly optimized control (both local and global) is an inherent attribute of the strategy rather than an explicitly programmed feature. The premise of transaction-based control is that interactions between various components in a complex energy system can be controlled by negotiating immediate and contingent contracts on a regular basis in lieu of standard on/off command and control. Each device is given the ability to negotiate deals with its peers, suppliers and customers in order to maximize revenues while minimizing costs. This is best illustrated by an example.

18.2.5.1 The End Use Facilities Marketplace

A typical building might have several chillers that supply a number of air handlers with chilled water on demand. If several air handlers require the full output of one chiller, and another air handler suddenly also requires cooling, traditional building control algorithms simply startup a second chiller to meet the demand and the building's electrical load ratchets upward accordingly.

A transaction-based building control system behaves differently. Instead of honoring an absolute demand for more chilled water, the air handler requests such service in the form of a bid (expressed in dollars), increasing its bid in proportion to its "need" (divergence of the zone temperature from its setpoint). The chiller controls, with knowledge of the electric rate structure, can easily express the cost of service as the cost of the kWh to run the additional chiller plus the incremental kW demand charge. If the zone served by this air handler just began to require cooling, its "need" is not very great at first, and so it places a low value on its bid for service and the additional chiller stays off until the level of need increases. Meanwhile, if another air handler satisfies its need for cooling, the cost of chilled water immediately drops below the bid price because a second chiller is no longer required, and the air handler awaiting service receives chilled water.

This is analogous to air traffic control where a limited resource is managed by scheduling demand into time slots. Alternatively, a peer-to-peer transaction can take place in which an air handler with greater need for service displaces (literally outbids) another whose thermostat is nearly satisfied.

In this way, the contract-based control system accomplishes several things. First, it inherently limits demand by providing the most "cost-effective" service. In doing this, it inherently prioritizes service to the most important needs before serving less important ones. Second, it decreases energy consumption by preventing the operation of an entire chiller to meet a small load, where it operates inefficiently.

Third, contract-based controls inherently propagate cost impacts up and down through successive hierarchical levels of the system being controlled (in this example, a heating/cooling system). The impacts on the utility bill, which are easily estimated for the chiller operation, are used as the basis for expressing the costs of air handler and zone services. Using cost as a common denominator for control makes expression of what is effectively a multilevel optimization much simpler to express than an engineered solution would be. It allows controls to be expressed in local, modular terms while accounting for their global impact on the entire system.

In effect, the engineering decision-making process is subsumed by a market value-based decision-making process that injects global information conveyed by market activity (i.e., asks, bids, immediate and contingent contracts, closings, and defaults) into the local engineering parameters that govern the behavior of individual systems over multiple timescales.

18.2.5.2 The Distribution Marketplace

One of the most critical, yet difficult, parts of the value chain to reveal is the value of new or expanded distribution capacity. When distribution utilities determine a need to add or increase the capacity of a substation, it is typical for the costs to range from $100/kW to $200/kW, or even higher in less common cases such as urban centers.[44] Although this is generally not enough value to fund a distributed-generation project that may cost $800/kW, for example, it nevertheless is a substantial fraction of the needed investment. When added to other values seen by the customer, such as backup power and reduced demand and energy charges, plus values for grid support and displaced central generation at the transmission and wholesale levels, the project may be attractive to the customer, the utility, or a third party if all these values can be accumulated or shared.

This suggests a role for a new distribution-level marketplace. The distribution utility that traditionally deals with operations and planning now has an additional function—operating a local marketplace or incentive structure where these values are revealed. By posting long-term upgrade costs as an opportunity notice to customers and third parties to bid alternative investments, such opportunities will reveal themselves.

For example, if a single distributed generator is located at a substation or feeder to displace an upgrade, then sooner or later it will go down just when the load is peaking on a hot summer day. To maintain the initial level of reliability, multiple, smaller units must be installed, or load management contracts must be signed with enough customers to back up the resource occasionally. Direct load curtailment contracts are one way to do this, using technology such as transactive control to make it seamless, automatic, and minimally disruptive. A market-like economic dispatch based on a local congestion surcharge is another way to let these resources compete for the right to serve peak loads at minimal cost. In either case, the distribution marketplace becomes a focal point for the transformation.

18.2.5.3 The Transmission Marketplace

The transmission power grid and wholesale markets can potentially be operated in an analogous, transactive fashion that provides a mechanism for the “self-healing” features that are the key to increasing the reliability of the grid while minimizing the infrastructure cost involved. A scenario best illustrates this concept. In today’s wholesale market, a large customer shopping for a long-term contract for power finds an energy service provider offering the minimal cost contract guaranteeing generation and transmission of power to the local distribution utility on its behalf. When a crisis occurs that suddenly disrupts the grid and severely constricts the transmission capacity of the corridor being used, the frequency disruption immediately triggers frequency-sensitive appliances throughout the entire region to temporarily rebalance supply and demand. This allows time for an automated power-trading system, analogous to an automated stock market trading scheme, to reconfigure the system in response to the crisis.

To do this, the energy management system immediately posts an emergency congestion surcharge on the constrained pathway. This surcharge triggers the energy service provider to issue a short-term subcontract to a generator in another location to use a different transmission pathway to deliver power on behalf of the customer. In effect, this is what grid operators do today, reconfiguring the grid and voiding power wheeling contracts with a telephone and paperwork. However, the fact that the August 14, 2003 blackout in the Eastern Interconnection, once begun,[45] rolled from Ohio to New York in nine seconds serves as a reminder that information technology is essential for reacting with the speed necessary to arrest such events.

The key is tapping into the remarkable property of markets to reorganize themselves efficiently and the speed of high technology to effect that reorganization. Although engineers speak of "control" and economists speak of "markets," in the future power grid, these concepts will merge and blend to form a transactive network that effectively forms the central nervous system of an adaptive, evolving system.

18.2.6 Renewable Energy Sources Integration

Broadly defined, renewable energy sources include wind, solar (solar photovoltaic (PV) and solar thermal), biomass, hydroelectric, ocean and river currents, waves, and tides. However, only wind, solar, and biomass sources are largely unconstrained by resource location. Thus, these renewable energy sources are ideally suited as distributed sources within the future electric system. Biomass-derived fuels can power conventional and emerging external and internal combustion engines. Thus, this renewable source brings the additional value of regulated/controllable power output and benefits in the form of voltage and reactive power support, spinning reserve, and improved power quality. In short, biomass-derived fuels can simply replace conventional fossil-based fuels in powering conventional power generator units such as diesel generator sets. However, both solar and wind energy sources can also bring many of these additional values.

Both solar and wind energy have, to date, been considered to be intermittent sources (when energy storage is absent). However, as these power sources become accepted as conventional components of the power generation mix and are used more effectively, there has come the recognition that these resources are best described as "variable" and that furthermore the outputs are relatively highly predictable. In addition, the parallel development of reliable, intelligent, low-cost, electronic-based voltage-source power converters has brought the ability of solar and wind generators to supply both high-quality power and reactive power support. Therefore, wind and solar power generators should now be included in the distributed power generator resource category.

The integration of ubiquitous communications and integrated distributed control now allows distributed generation, and particularly renewable power sources to utilize loads (especially thermostatic loads) to "firm up" intermittent sources making them collectively a more reliable resource. Additionally, opportunities for distributed renewables to participate in real-time emission management and provide grid support such as reactive power, cold load pickup during restoration, and backup for nearby critical loads provide benefits that can be realized economically.

Renewable power generators can be assigned effective load-carrying capability (ELCC) (or capacity credit) values. These values can be established when the deployment configuration, integration, and dispatch of these resources, along with other distributed resources, is optimized within a specific distribution or transmission system. For example, it has been shown that had a dependable capacity of 5000 MW of PV been installed in the California ISO region in 2000, the peak load on June 15, 2000 could have been reduced by 3000 MW. This would have cut by one-half the number of equivalently sized natural-gas-driven combustion gas turbines needed to ensure system reserve capacity.[46] Although PV power is only available during daylight hours, the annual average ELCC estimated for PV power plants distributed throughout California has nonetheless been calculated to be 64% of the PV plant rating. Furthermore, when the California system electric load is driven by the sun, the ELCC exceeds 80%. Thus a 5-MW PV system would be considered a summer peaking power source of 4 MW.

Wind-power output in many locations also has a definite diurnal pattern. However, analyses of real operating data currently underway in both the United States and Europe show that the very favorable geographic diversity effect (due to many wind turbines scattered throughout a relatively large geographical area), along with the significant advances in short-term wind forecasting, can permit electric power system planners and operators to assign high values of ELCC and capacity factor to wind-power systems.[47]

In short, the intelligent deployment, integration, and operation of wind and solar energy sources within the future electric system enable renewable power resources to be categorized as distributed power generation assets without qualification, bringing with them many of the values of the traditional distributed power generation sources (such as diesel generators and gas turbine generators) that can be exploited within the system, renewable power to provide ancillary benefits such as grid reactive support, and additional societal benefits to be realized such as real-time emissions management.

References

1. Linden, D. and Reddy, T. B. 2002. *Handbook of Batteries, 3rd Ed.*, pp. 40–163. McGraw Hill, New York.
2. Luongo, C. A. 1996. Superconducting storage systems: An overview. *IEEE Transactions on Magnetics, Vol. 32*, 2214–2223.
3. Cheung, K. Y. C., Cheung, S. T. H. et al. 2003. *Large Scale Energy Storage Systems*, pp. 40–163. Imperial College London, London.
4. EA Technology, *Review of Electrical Energy Storage Technologies and Systems and of their Potential for the UK*, pp. 40–163. Department of Trade and Industry, London, UK.
5. DeVries, T. 2002. World's biggest battery helps stabilize Alaska. *Modern Power Systems*, 22, 40.
6. Nourai, A. 2003. NaS battery demonstration in the USA. *Electricity Storage Association Spring Meeting*, Electricity Storage Association, Morgan Hill, California.
7. Sudworth, J. L. 2001. Sodium/nickel chloride (Zebra) battery. *Journal of Power Sources*, 100, 149–163.
8. Fleets and Fuels. 2003. *The Zebra*, (February 17), 8.
9. Price, A. 2000. Technologies for energy storage, present and future: Flow batteries. *2000 Power Engineering Society Meeting*. IEEE Power Engineering Society, Piscataway, New Jersey.
10. Skyllas-Kazacos, M. 2000. *Recent Progress with the Vanadium Redox Battery*, pp. 27–148. University of New South Wales, Sydney.
11. Menictas, C. 1994. Status of the vanadium battery development program. *Proceedings of the Electrical Engineering Congress*, Sydney.
12. VRB Power Systems, Inc. http://www.vrbpower.com (accessed August 15, 2005).
13. Wilks, N. 2000. Solving current problems. *Professional Engineering*, 13, 27.
14. Scheffler, P. 2001. *Environmental Assessment for the Regenesys Energy Storage System*, pp. 142–148. Tennessee Valley Authority.
15. Lex, P. and Jonshagen, B. 1999. The zinc/bromide battery system for utility and remote applications. *Power Engineering Journal*, 13, 142–148.
16. Kruger, P. 2001. Electric power requirement for large-scale production of hydrogen fuel for the world vehicle fleet. *International Journal of Hydrogen Energy*, 26, 1137–1147.
17. Bossel, U., Eliasson, B., and Taylor, G. 2003. *The Future of the Hydrogen Economy: Bright or Bleak? Oberrohrdorf*, pp. 1863–1874. European Fuel Cell Forum, Switzerland.
18. Dillon, A. et al. Storage of hydrogen in single-walled carbon nanotubes. *Nature*, 386, 377–379.
19. Donalek, P. 2003. Advances in pumped storage. *Electricity Association Spring Meeting*, Chicago, IL.
20. Kondoh, J. et al. 2000. Electrical energy storage systems for energy networks. *Energy Conservation and Management*, 41, 1863–1874.
21. van der Linden, S. 2002. The case for compressed air energy system. *Modern Power Systems*, 22 19–21.
22. Hull, J. et al. 1988. *Salinity Gradient Solar Ponds*, pp. 393–403. CRC Press, Boca Raton, FL.
23. Sørensen, B. 2004. *Renewable Energy: Its Physics, Engineering, Environmental Impacts, Economics and Planning, 3rd Ed.*, pp. 393–403. Elsevier Academic Press, Burlington, MA.

24. Neeper, D. A. 2000. Thermal dynamics of wallboard with latent heat storage. *Solar Energy*, 68, 393–403.
25. Goswami, D. Y., Kreith, F., and Kreider, J. F. 2000. *Principles of Solar Engineering*, *2nd Ed*., pp. 631–653. Taylor & Francis, New York.
26. Hanneman, R., Vakil, H., and Wentorf, R. Jr. 1974. Closed loop chemical systems for energy transmission, conversion and storage, *Proceedings of the 9th Intersociety Energy Conversion Engineering Conference*, pp. 631–653. American Society of Mechanical Engineers, New York.
27. Shapouri, H., Duffield, J. A., Wang, M. 2002. *The Energy Balance of Corn Ethanol: An Update. Report 814*. USDA Office of Energy Policy and New Uses, Agricultural Economics, Washington, DC.
28. Hammerschlag, R. 2006. Ethanol's energy return on investment: A survey of the literature 1990–Present. *Environmental Science and Technology*, 40, 1744–1750.
29. Klass, D. L. 1998. *Biomass for Renewable Energy, Fuels, and Chemicals*, pp. 631–653. Academic Press, San Diego.
30. Bockey, D. and Körbitz, W. 2003. *Situation and Development Potential for the Production of Biodiesel—An International Study*, pp. 631–653. Union zur Forderung von Oel- und Proteinpflanzen, Berlin.
31. Bridgwater, A. V. 1995. The technical and economic feasibility of biomass gasification for power generation. *Fuel*, 74, 631–653.
32. Kannberg, L., Chassin, D., Desteese, J., Hauser, S., Kintner-Meyer, M., Pratt, R., Schienbein, L., and Warwick, W. 2003. *GridWise™: The Benefits of a Transformed Energy System*. p. 10. Pacific, Northwest National Laboratory, Richland, WA.
33. Baer, W., Fulton, B., Mahnovski, S. 2004. *Estimating the Benefits of the GridWise Initiative*. pp. 17–20. Rand Corporation, Santa Monica, CA.
34. Chapel, S., Coles, L., Iannucci, J., and Pupp, R. 1993. *Distributed Utility Valuation Project Monograph*, p. 6. EPRI, NREL, and PG&E, San Francisco, CA.
35. Pratt, R. 2004. Transforming the electric energy system, *Proceedings of 2004 IEEE PES Power Systems Conference and Exhibition*, Vol. 3, pp. 1651-1654. New York, NY.
36. Intelligrid Architecture, 2005. http://intelligrid.info/IntelliGrid_Architecture/Overview_Guidelines/index.htm (February 3, 2006).
37. The Modern Grid Initiative, 2006. http://www.themoderngrid.org (February 3, 2006).
38. Widergren, S. et. al. 2005. GridWise architecture council constitutional convention, *Proceedings*, pp. 1–45. Pacific Northwest National Laboratory, Richland, WA.
39. Hauer, J., Overbye, T., Dagle, J., and Widergren, S. 2002. *National Transmission Grid Study Issue Papers: Advanced Transmission Technologies*, pp. F1–F44. US Department of Energy, Washington, DC.
40. Office of Electric Transmission and Distribution, 2004. *DOE Electric Distribution Multi-Year Research, Development, Demonstration, and Deployment Technology Roadmap Plan:* 2005–2009. pp. 1–82. US Department of Energy.
41. Office of Electric Transmission and Distribution, 2003. Grid 2030: *National Vision for Electricity's Second 100 Years*, pp. 1–44. US Department of Energy, Washington, DC.
42. Dagle, J., Winiarski, D., and Donnelly, M. 1997. *End-Use Load Control for Power System Dynamic Stability Enhancement*. pp. 1–57. Pacific Northwest National Laboratory, Richland, WA.
43. Trudknowski, D., Donnelly, M., and Lightner, E. 2005. *Power System Frequency and Stability Control Using Decentralized Intelligent Loads*, pp. 1–7. Pacific Northwest National Laboratory, Richland, WA.
44. Balducci, P. J., Schienbein, L. A., Nguyen, T. B., Brown, D. R., and Fathelrahman, E. M. 2004. *An Examination of the Costs and Critical Characteristics of Electricity Distribution System Capacity Enhancement Projects for Three Western Utilities IEEE Power Systems Conference and Exposition October* 10–13, 2004, Vol. 1, p. 517.

45. Final Report on the August 14, 2003 Blackout in the United States and Canada: Causes and Recommendations, U.S.-Canada Power System Outage Task Force, April 2004, https://reports.energy.gov/BlackoutFinal-Web.pdf. (February 3, 2006).
46. Herig, C. Using Photovoltaics to Preserve California's Electricity Capacity Reserves. http://www.nrel.gov/docs/fy01osti/31179.pdf (February 3, 2006).
47. Soder, L. and Ackerman, T. 2005. Wind power in power systems: an introduction. In *Wind Power in Power Systems*, T. Ackerman, ed., pp. 36–38. John Wiley and Sons, Ltd., West Sussex, England.

19

Availability of Renewable Resources

David S. Renné,
Stephen Wilcox, and
William Marion,
National Renewable Energy Laboratory

Gene L. Maxwell
Retired Solar Expert

Martin Rymes
National Renewable Energy Laboratory

Julie Phillips
Julie Phillips and Associates

Dale E. Berg
*Sandia National Laboratories**

Marjorie A. Franklin
Franklin Associates

Robert C. Brown
Iowa State University

19.1 Solar Energy

David S. Renné, Stephen Wilcox, William Marion, Gene L. Maxwell, Martin Rymes, and Julie Phillips

19.1.1 Introduction

Solar radiation is the electromagnetic energy emitted by the sun. The amount of solar radiation available for terrestrial solar energy projects is that portion of the total solar radiation that reached the earth's surface. In this chapter, the terms *insolation* and *irradiance* are used interchangeably to define solar radiation incident on the earth per unit of area per unit of time as measured in watts per square meter (W/m^2) or watt-hours per square meter per 24-h day ($Wh/m^2/day$).

*Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

The amount of solar radiation that reached the earth's surface at any give location and time depends on many factors. They include time of day, season, latitude, surface albedo, the translucence of atmosphere, and the weather.

This chapter describes practical methods for estimating the amount of solar radiation at specific locations. Section 19.1.2, "Extraterrestrial Solar Radiation," discusses the measurement and modeling of solar energy outside the earth's atmosphere. Section 19.1.3, "Terrestrial Solar Radiation," describes the factions affecting solar radiation at the earth's surface and provides a general description of considerations important in estimating the available resource. Section 19.1.4, "Solar Radiation Data for the United States," describe the National Solar Radiation Database and presents example data tables for selected stations in the U.S. Section 19.1.5, "Solar Design Tools," describes publications available from the National Renewable Energy Laboratory that provide estimates of the U.S. solar resource and also gives equations needed to estimate insolation on both horizontal and tilted surfaces. Section 19.1.6, "International Solar Radiation Data," describes databases that may be useful in estimating solar resources in other countries.

19.1.2 Extraterrestrial Solar Radiation

Extraterrestrial radiation is the amount of solar energy above the earth's atmosphere. Also known as *top-of-the-atmosphere* radiation, this parameter indicates the amount of solar energy that would fall on the earth in the absence of an atmosphere.

The sun's energy is nearly constant, varying from year to year by less than 1%. The small interannual variations are associated primarily with the 22-year sunspot cycle. However, extraterrestrial solar radiation also varies due to the earth's elliptical orbit around the sun. Extraterrestrial radiation increases by about 7% from July 4 to January 3, at which time the earth reaches the point in its orbit closest to the sun. Figure 19.1 shows a diagram of the earth's orbit around the sun.

Calculations of extraterrestrial solar radiation typically include an eccentricity correction factor, E_0, to account for this variation. For many engineering applications, the following hand calculation is

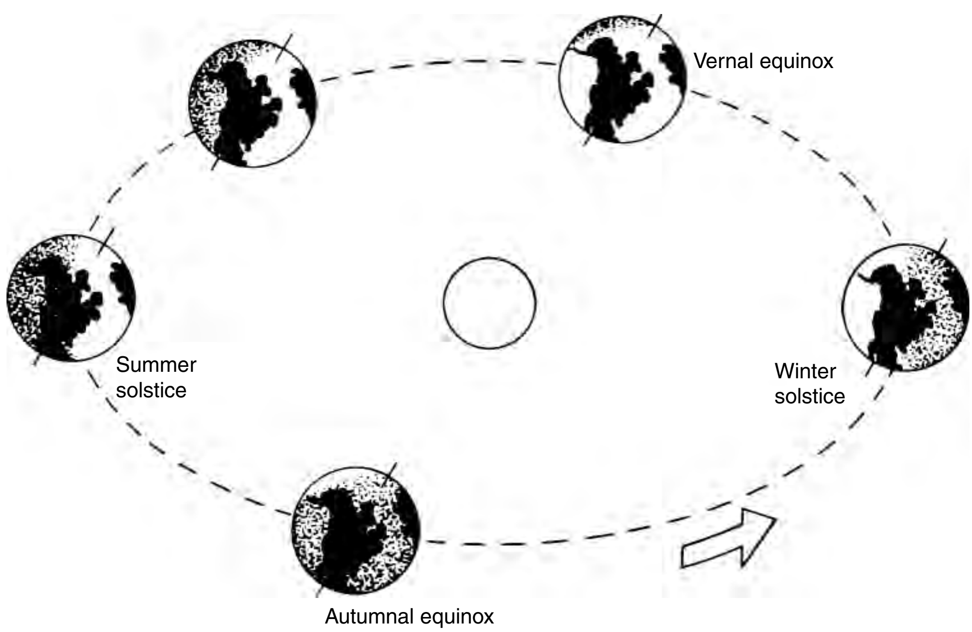


FIGURE 19.1 The earth's elliptical orbit around the sun.

adequate (Goswami, Kreith, and Kreider 2000):

$$E_0 = 1 + 0.033\cos\left(\frac{360n}{365}\right), \tag{19.1}$$

where n is the day of the year counted from January 1, such that $1 \leq n \leq 365$.

The earth's axis is tilted 23.5° with respect to the plan of its orbit around the sun. Tilting results in longer days in the northern hemisphere from the spring equinox (approximately March 23) to the autumnal equinox (approximately September 22) and longer days in the southern hemisphere during the other six months. On the March and September equinoxes, the sun is directly over the equator, both poles are equidistant from the sun, and the entire earth experiences 12 h of daylight and 12 h of darkness.

In the temperate latitudes ranging from 23.45° to 66.5° north and south, variations in insolation are pronounced. At a latitude of 40° N, for example, the average daily total extraterrestrial solar radiation (on a horizontal surfaces) varies from 3.94 kWh/m^2 in December to 11.68 kWh/m^2 in June. Therefore, it is necessary to know the location of the sun in the sky to determine the extraterrestrial radiation at a specific location and time.

The sun's position relative to a location can be determined if a location's latitude, L, a location's hour angle, W, and the sun's declination are known. These three parameters are shown graphically in Figure 19.2. Latitude is the angular distance north or south of the earth's equator, measured in degrees along a meridian.

The hour angle is measured in the earth's equatorial plane. It is the angle between the projection of a line drawn from the location to the earth's center and the projection of a line drawn from the center of the earth to the sun's center. Thus, at solar noon, the hour angle is zero. At a specific location, the hour angle expresses the time of day with respect to solar noon, with one hour of time equal to 15 degrees angle. By convention, the westward direction from solar noon is positive.

The sun's declination is the angle between a projection of the line connecting the center of the earth with the center of the sun and the earth's equatorial plane. It varies from −23.45° on the winter solstice (December 21), to +23.45° on the summer solstice (June 22). The approximate declination of any given

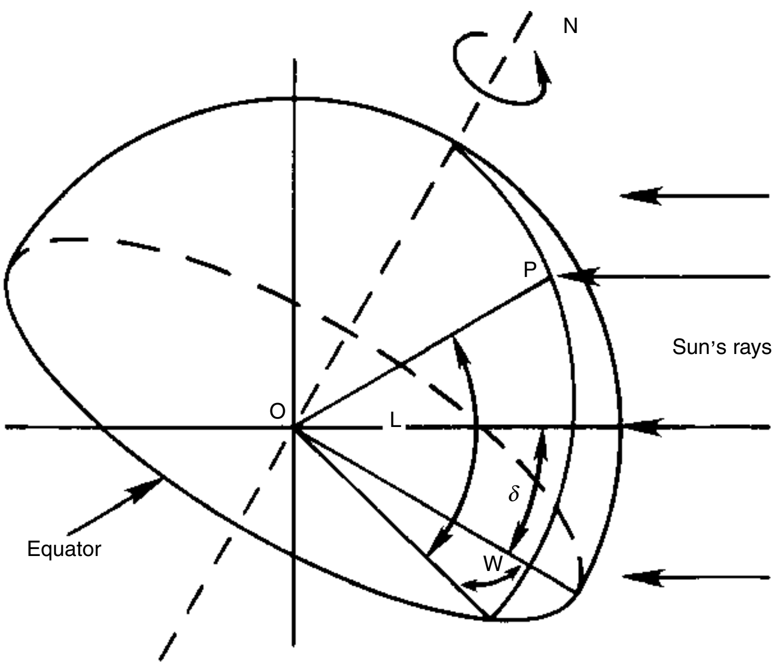


FIGURE 19.2 Solar geometry: latitude, hour angle, and the sun's declination.

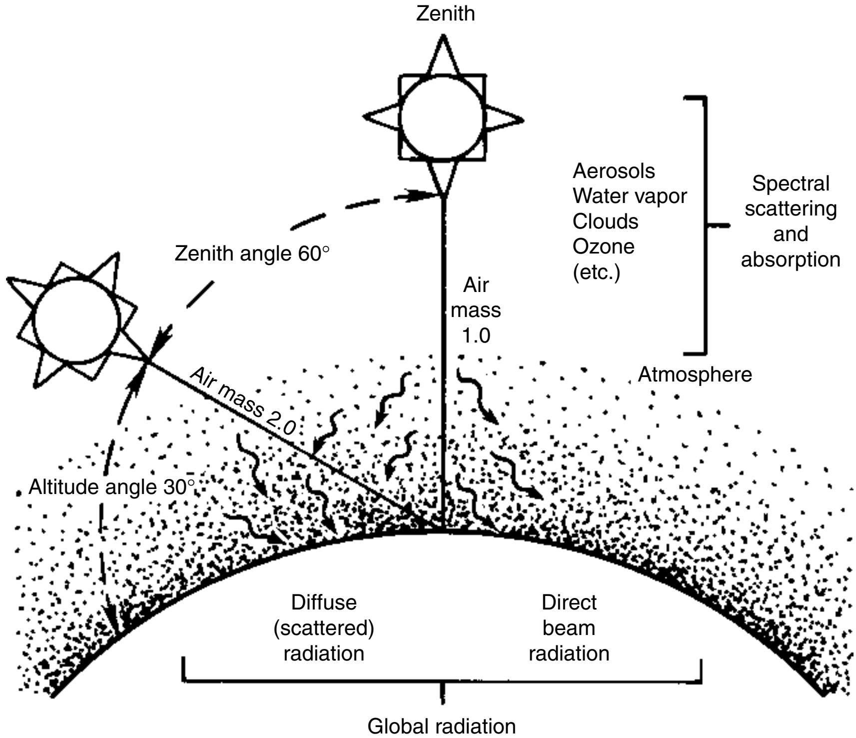


FIGURE 19.3 Solar geometry: solar altitude and solar zenith angles.

day is given by (Cooper 1969):

$$\delta = 23.45 \sin\left[360\left(\frac{284 + n}{365}\right)\right]. \tag{19.2}$$

Several other angles are used to identify the location of the sun in the sky. These angles include the solar azimuth angle (γ), the solar zenith angle (θ_z), the solar altitude angle (α), and the sunset hour angle, W_s. The solar azimuth angle is the compass direction. The solar zenith angle, as shown in Figure 19.3, is the angle of the sun from the vertical, or zenith. The solar altitude angle is the angle of the sun with the horizon. Therefore, $\alpha + \theta_z = 90°$.

If the solar zenith angle is known, the extraterrestrial radiation (H_0) can be calculated as follows:

$$H_0 = I_{SC} \cdot E_0 \cdot \cos(\theta_z), \tag{19.3}$$

where I_{SC} is the solar constant and E_0 is the correction factor for the eccentricity of earth's orbit.

The value of the solar constant (I_{SC}) endorsed by the World Radiation Center in Davos Dorf, Switzerland, is 1367 W/m^2. The eccentricity correction factor (E_0), calculated using Equation 19.1, allows for the daily adjustment of the solar constant.

The relationship among the solar zenith angle, the solar altitude angle, the sun's declination, the hour angle, and latitude is given as follows:

$$\cos \theta_z = \sin \alpha = \cos \delta \cos W \cos L + \sin \delta \sin L. \tag{19.4}$$

This equation can be used to compute the sunset hour angle. At sunset, the sine of α is negative. Therefore,

$$\cos W_S = -\tan \delta \tan L. \tag{19.5}$$

Because day length can be easily derived from Equation 19.5, W_S is also a measure of the position of the earth in its orbit around the sun. This relationship is given by

$$\text{day length} = 2 \cos^{-1}(-\tan\delta \quad \tan L). \tag{19.6}$$

19.1.3 Terrestrial Solar Radiation

Measuring or calculating the total global solar radiation striking the earth's surface is far more complicated than determining extraterrestrial radiation. Changing atmospheric conditions, topography, and the changing position of the sun interact to moderate the solar resource at a given time and location. The problem is further complicated for the solar engineer because the vast majority of solar radiation measurements (and extrapolations based on them) have been made on horizontal surfaces rather than on the tilted surfaces typically employed by efficient solar energy collection. Solar designers must use reference tables that translate horizontal data into directional data for tilted surfaces or make those calculations themselves. Section 19.1.5, "Solar Design Tools," describes such calculations and provides information on obtaining data manuals that describe the available solar resource for tilted surfaces in specific locations.

The total solar radiation on earth consists of (1) direct beam radiation that comes to the surface on a direct line from the sun, and (2) diffuse radiation from the sky, created when part of the direct beam radiation is scattered by atmospheric constituents (e.g., clouds and aerosols). Concentrating solar collectors rely almost entirely on direct beam radiation, whereas other collectors, including passive solar buildings, capture both direct beam and diffuse radiation. Radiation reflected from the surface in front of a collector may also contribute to total solar radiation.

The atmosphere through which solar radiation passes is variable and acts as a dynamic filter, absorbing and scattering solar radiation. Low- and mid-level cumulus and stratus clouds are generally opaque, blocking the direct beam of the sun. In contrast, high, thin cirrus clouds are usually translucent, scattering the direct beam, but not totally blocking it.

On sunny days with clear skies, most of the solar radiation is direct beam radiation and diffuse radiation only accounts for about 5–20% of the total. Under overcast skies, diffuse radiation, which is scattered out of the direct beam by gases, aerosols, and clouds, accounts for 100% of the solar radiation reaching the earth's surface. Under clear skies, the total instantaneous solar radiation at the planet's surface at midday can exceed 1000 W/m^2. In contrast, instantaneous midday radiation on a dark, overcast day can be less than 100 W/m^2.

The impact of clouds on the solar resource can be seen in Figure 19.4, which shows the inverse relationship between average monthly opaque cloud cover and average daily solar energy for Fresno, California. Fresno experiences relatively little seasonal variation in atmospheric turbidity and only small variations in atmospheric water vapor. Therefore, the city's seasonal variations in average daily solar radiation are due almost entirely to variations in cloud cover (Maxwell, Marion, and Myers 1995).

Aerosols, including dust, smoke, pollen, and suspended water droplets reduce the transmittance or amount of solar radiation reaching the planet's surface. These factors are influenced by climate and seasonal changes, as demonstrated in Figure 19.5, which shows seasonal variations in aerosol optical depth for six cities with different climates. If the annual mean aerosol optical depth and the amplitude of seasonal variations are known for a particular location, the aerosol optical depth for any day of the year can be estimated as follows (Maxwell and Myers 1992):

$$\tau_A = A \sin [(360n/365) - \varphi] + C, \tag{19.7}$$

where τ_A is the aerosol optical depth on day n, A is the amplitude of seasonal variations in τ_A, n is the day of the year (1–365), ϕ is the phase factor that determines the days when the maximum and minimum values occur (for most locations in the northern hemisphere, $\phi = 90°$), and C is the annual mean aerosol optical depth.

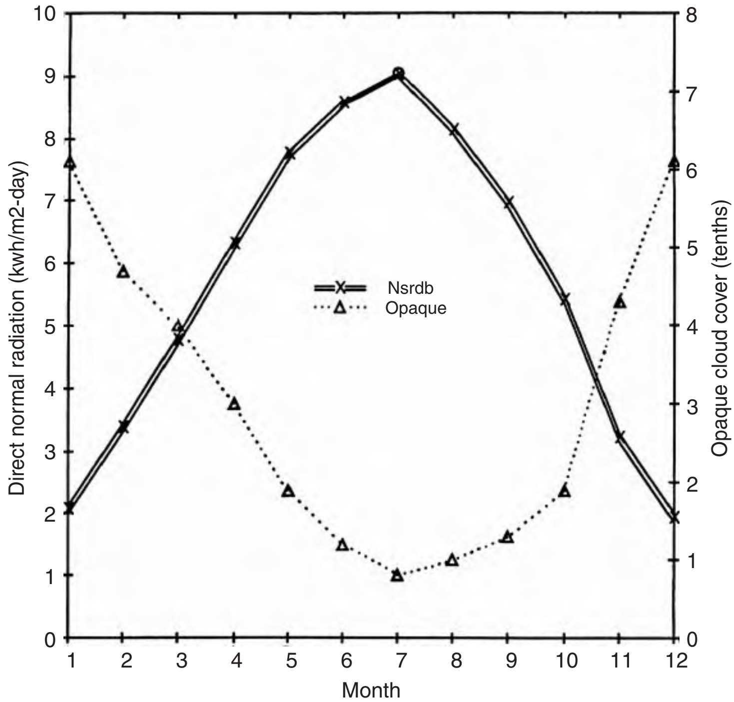


FIGURE 19.4 Comparison of direct normal solar radiation and opaque cloud cover for Fresno, CA. (From Maxwell, E. L., Marion, W., and Myers, D. R. 1995. *National Solar Radiation Data Base*, *Vol. 2*, (1961–1990), National Renewable Energy Laboratory, Golden, CO, 216–260.)

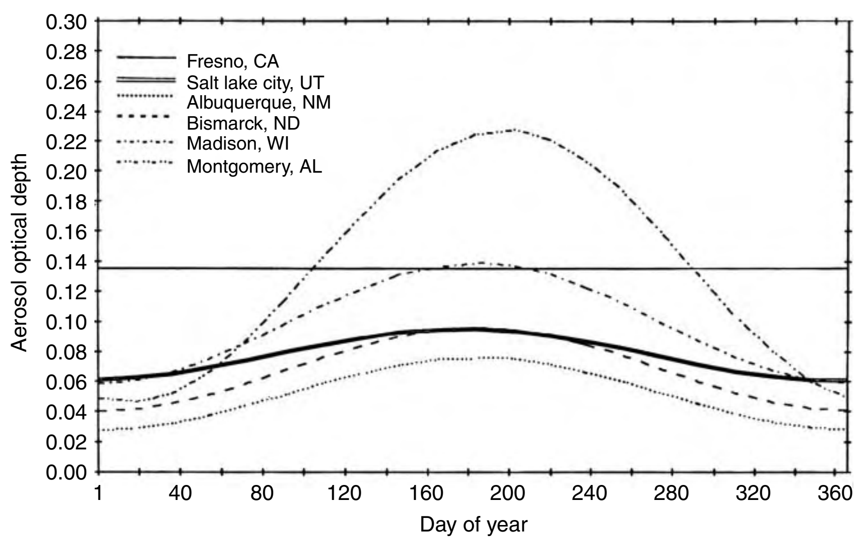


FIGURE 19.5 Seasonal variation of aerosol optical depth for six locations in the United States. (From Maxwell, E. L. and Myers D. R. 1992. *Proceedings of the 1992 Annual Conference of the American Solar Energy Society*, 323–327.)

The monthly average clearness index (K) is often used to quantify the relative transmittance of the atmosphere. The clearness index is (Iqbal 1983:249)

$$K = \frac{\bar{H}}{\bar{H}_0}, \tag{19.8}$$

where $\bar{H}$ is the monthly average daily total of the global horizontal radiation (see Section 19.1.3), and $\bar{H}_0$ is the average daily total amount of extraterrestrial radiation incident upon a horizontal surface during that month.

The albedo of a surface affects the total solar radiation available to a tilted collector. For example, with snow on the ground and scattered clouds overhead, surface-to-cloud reflections can greatly increase diffuse solar radiation. Under such circumstances, the total global radiation on the collector can increase. Any highly reflective surface material such as Utah's salt flats or light-colored sand will have similar effects. Engineers planning a solar project in highly reflective areas should measure the average albedo in the region and work with a model that incorporates these surface reflection effects.

Seasonal and diurnal changes in insolation impact resource availability. Insolation is normally greater at midday, when the path of the sun's rays through the atmosphere is shortest, than in the early morning or late afternoon, when sunlight must pass through more of the atmosphere to reach the plant's surface, as shown in Figure 19.3. The longer the path through the atmosphere, the greater will be the absorption and rescattering of solar radiation. In the temperate and polar latitudes, daily total solar radiation is much higher during summer months because there are more daylight hours.

Some solar energy applications may require spectral information in addition to the parameters discussed earlier. The spectral distribution of extraterrestrial solar radiation, with wavelengths ranging from 0.3 m to 3 μm, is shown in Figure 19.6. Figure 19.6 also shows an example of a similar spectral distribution after solar radiation has passed through the atmosphere and is absorbed by particulates, ozone, water vapor, and other gases. Because many photovoltaic devices are optimized for specific wavelengths in the solar spectrum, spectral information is necessary for evaluating their performance. Such information is available from the American Society for Testing and Materials' "Standard for Terrestrial Solar Spectral Irradiance Tables at Air Mass 1.5 for a 37° Tilted Surface, ASTM Standard E892-82." The standard is found in the *Annual Book of ASTM Standards*, Volume 12.02, Section 12, 1984,

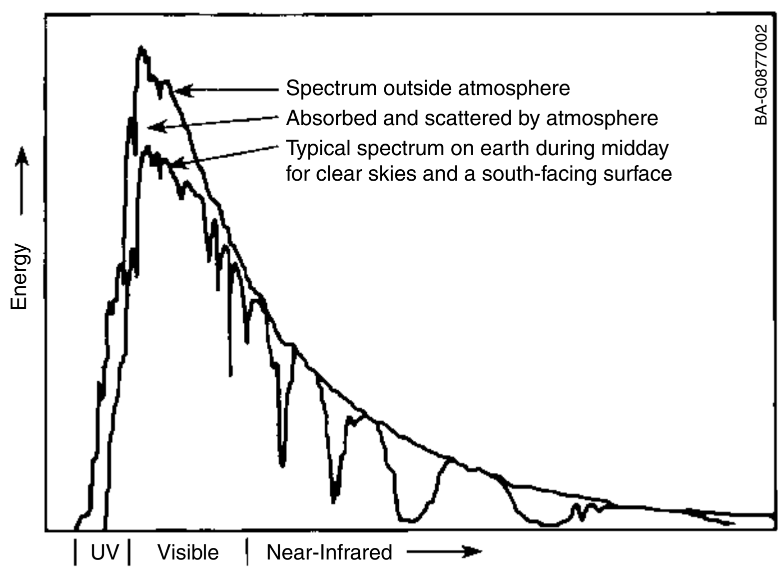


FIGURE 19.6 Comparison of solar spectral distributions in space and on earth.

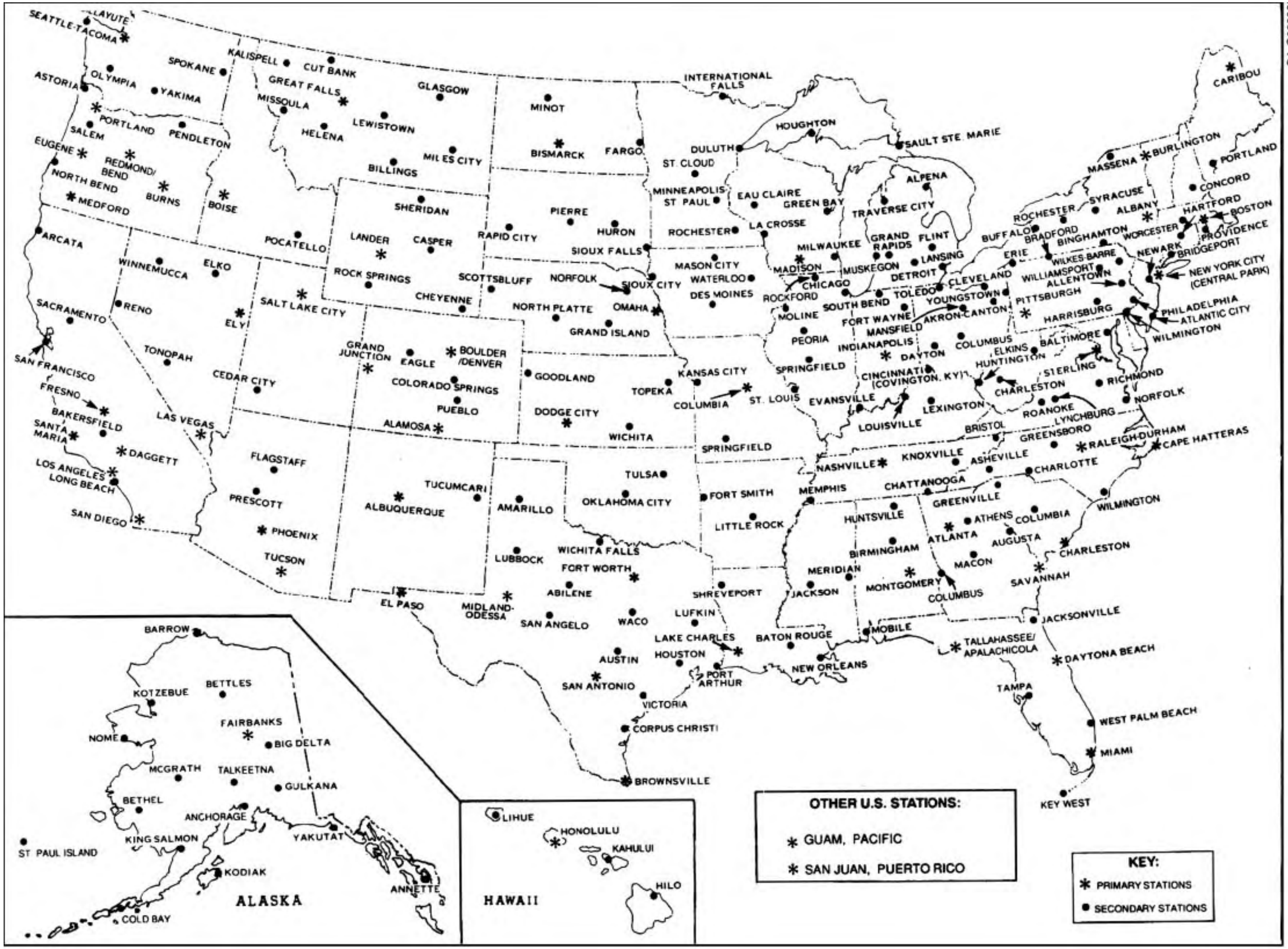


FIGURE 19.7 Map showing the location of 239 stations in the National Solar Radiation Database.

pp. 712–719 and is available from the American Society of Testing and Materials in Philadelphia, Pennsylvania.

Even with the best instruments and calibration methods available today, solar radiation data are no more certain than $\pm 5\%$ and the uncertainty is usually much higher. Solar designers should take this uncertainty into account in predicting the performance of solar energy systems.

19.1.4 Solar Radiation Data for the United States

Solar designers and engineers can obtain basic information about the nation's solar resource for the National Solar Radiation Database (NSRDB 1992) that was developed as part of a national resource assessment project conducted by the National Renewable Energy Laboratory for the U.S. Department of Energy. A series of solar resource maps of the United States has been produced from this database. The database and products are described in this section.

19.1.4.1 National Solar Radiation Database

The National Renewable Energy Laboratory completed the National Solar Radiation Database in 1992 to assist designers, planners, and engineers with solar resource assessments. The database contains 30 years of hourly values of measured or model solar radiation together with meteorological information from 239 stations across the United States, covering the period from 1961 through 1990. Figure 19.7 shows the locations of these stations and indicates whether each location was a primary station that measured solar radiation data for at least 1 year, or a secondary station. Solar radiation data for secondary stations were derived from computer models. In most cases, primary and secondary stations were located at National Weather Service stations that collected hourly or three-hourly meteorological data throughout the period from 1961 to 1990. Table 19.1 shows the solar radiation and meteorological elements included in the database. An update to the NSRDB is currently underway; data for the period 1991–2005 will be available in the near future.

All data are referenced to local standard time. The solar radiation elements represent the total energy received during the hour preceding the designated times. The position of the sun in the sky and the earth's position in orbit were calculated at the midpoint of the hour preceding the designated time. Meteorological elements are the values observed at the designated time.

The most comprehensive products available from the database are serial hourly data in either a synoptic or TD-3282 format. The synoptic format presents all solar radiation and meteorological data for 1 h in a line of data. The next line contains data for the next hour.

In the TD-3282 format, each line of data contains a day's worth of data (24 h) for one element. The next line contains a day of data for the next element. This format is more flexible in that it allows the user to tailor data files to a specific project's requirements. Both synoptic and TD-3282 formats present the

TABLE 19.1 National Solar Radiation Database

Hourly Solar Radiation Elements	Meteorological Elements
Global horizontal radiation (in Wh/m^2)	Total sky cover (in tenths)
Direct normal radiation (in Wh/m^2)	Opaque sky cover (in tenths)
Diffuse horizontal radiation (in Wh/m^2)	Dry bulb temperature (in °C)
Extraterrestrial radiation (in Wh/m^2)	Dew point temperature (in °C)
Direct normal ETR (in Wh/m^2)	Wind direction (in degrees)
	Wind speed (in m/s)
	Horizontal visibility (in km)
	Ceiling height (in km)
	Present weather
	Total precipitable water (in cm)
	Aerosol optical depth
	Snow depth (in cm)
	Days since last snowfall

degree of uncertainty for global horizontal, direct normal, and diffuse horizontal radiation data. They also indicate whether these elements were measured or modeled.

Other database products include (1) hourly, daily, and quality statistics for the solar radiation elements, (2) daily statistics for the meteorological elements, and (3) persistence statistics for daily total solar radiation. The statistical summaries include hourly, monthly, annual, and 30-year averages and their standard deviations. The persistence statistics show the number of times the daily solar energy persisted above or below set thresholds for periods from 1 to 15 days during the 30-year period. Summary characteristics of specific measurement sites are also available. The statistical products and the synoptic and TD-3282 data formats are completely described in a user's manual (NSRDB 1992).

The National Solar Radiation Database is available in the synoptic format on a three-disk set of CD-ROMs from:

User Services
National Climatic Data Center
Federal Building
Asheville, NC 28801-2696
Telephone: 704-259-0682
Facsimile: 704-259-0876

The user's manual is included on the CD-ROMs, but a hard copy is also available. Other products developed from the database are also available from the National Climatic Data Center. These products include:

1. Hourly statistics files on floppy disks: statistics include means, standard deviations, and cumulative frequency distributions of hourly solar radiation (Wh/m^2) for global, direct, and diffuse elements for each station-year-month, each station-year, and for the 30-year period
2. Daily statistics files on floppy disks: Statistics include means and standard deviations of daily total radiation (Wh/m^2) for global, direct, and diffuse elements for each station-year-month, each station-year, and for the 30-year period. Statistics for meteorological elements include:
   a. Mean daily temperature in °C
   b. Mean daily temperature during daylight hours in °C
   c. Maximum daily temperature in °C
   d. Minimum daily temperature in °C
   e. Mean wind speed
   f. Mean total and opaque sky cover
   g. Mean total precipitable water in cm
   h. Mean aerosol optical depth
   i. Mean heating and cooling degree days
3. Persistence files of daily total solar radiation on floppy disks

19.1.4.2 Solar Radiation Data Manual

The *Solar Radiation Data Manual for Flat-Plate and Concentrating Collectors* (Marion and Wilcox 1994), published by the National Renewable Energy Laboratory, provides information on the solar resource for 239 stations in the United States and its territories. Researchers developed the manual's data from the hourly values of direct beam and diffuse horizontal solar radiation values in the National Solar Radiation Database.

The manual converts raw data regarding insolation into a form directly useful to solar system designers and engineers. It contains estimates of the resource available to flat-plate collectors mounted at fixed tilt angles of 0°, latitude minus 15°, latitude, latitude plus 15°, and 90°. The manual also shows solar irradiation for flat-plate and concentrating collectors, including parabolic trough designs, with one-axis tracking and two-axis tracking. Figure 19.8 shows drawings of five collector types for which resource estimates were made.

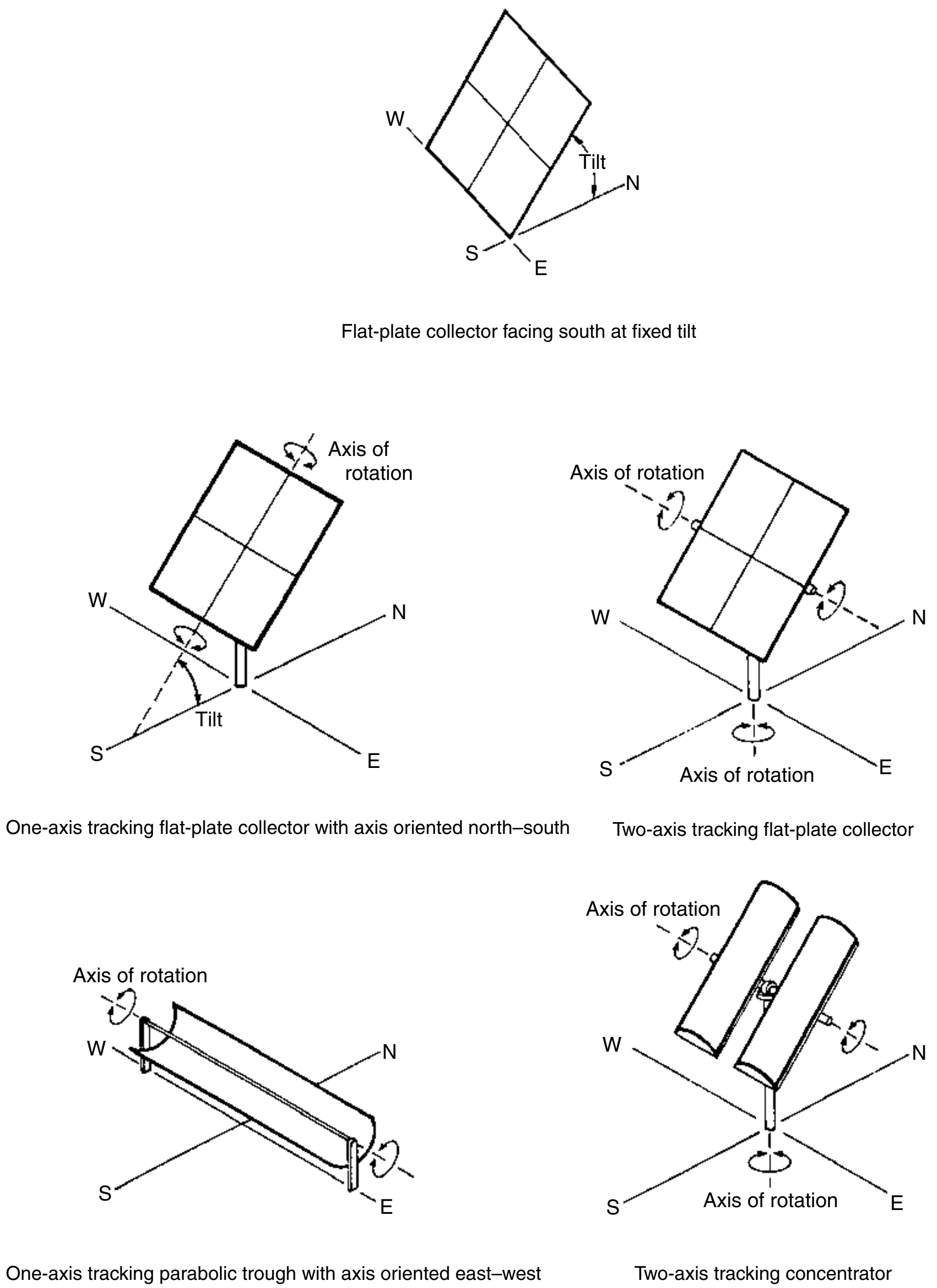


FIGURE 19.8 Flat-plate and concentrating solar collector configurations.

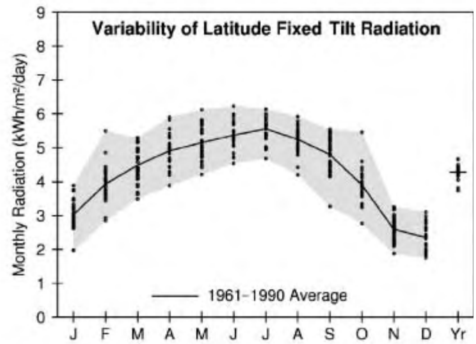
Data include monthly and hourly averages for 14 solar collectors for the period of 1961–1990. Data tables for each station are presented on a single page, which also contains a graph highlighting the variability of the solar resource at that particular station. Figure 19.9 through Figure 19.22 show sample data pages for 14 cities throughout the United States and Puerto Rico. Each data page also contains a table listing climatic conditions such as average temperatures, average daily minimum and maximum

Albany, NY

WBAN NO. 14735

LATITUDE: 42.75° N
LONGITUDE: 73.80° W
ELEVATION: 89 m
MEAN PRESSURE: 1006 millibars

STATION TYPE: Primary



Solar Radiation for Flat-Plate Collectors Facing South at a Fixed Tilt ($kWh/m^2/day$), Uncertainty ±9%

Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	1.8	2.6	3.6	4.7	5.5	6.0	6.1	5.2	4.1	2.8	1.7	1.4	3.8
	Min/Max	1.2/2.1	2.1/3.3	3.1/4.1	3.9/5.4	4.5/6.4	5.1/6.9	5.4/6.6	4.6/5.8	3.2/4.6	2.2/3.6	1.4/2.0	1.2/1.6	3.6/4.1
Latitude -15	Average	2.7	3.6	4.4	5.0	5.5	5.8	6.0	5.5	4.8	3.7	2.4	2.1	4.3
	Min/Max	1.8/3.4	2.6/4.9	3.4/5.1	4.0/6.0	4.5/6.5	4.9/6.8	5.0/6.6	4.4/6.2	3.3/5.5	2.6/5.1	1.8/2.9	1.6/2.7	3.7/4.7
Latitude	Average	3.0	3.9	4.5	4.9	5.1	5.4	5.5	5.2	4.8	3.9	2.6	2.4	4.3
	Min/Max	2.0/3.9	2.8/5.5	3.5/5.3	3.9/5.9	4.2/6.1	4.5/6.2	4.7/6.1	4.2/5.9	3.3/5.6	2.8/5.5	1.9/3.3	1.8/3.1	3.7/4.7
Latitude +15	Average	3.2	4.1	4.4	4.5	4.6	4.6	4.8	4.8	4.6	3.9	2.7	2.5	4.1
	Min/Max	2.1/4.2	2.9/5.8	3.4/5.2	3.6/5.5	3.8/5.4	4.0/5.4	4.1/5.3	3.8/5.4	3.1/5.3	2.7/5.5	1.9/3.4	1.9/3.3	3.5/4.4
90	Average	3.1	3.7	3.5	3.1	2.7	2.6	2.8	3.0	3.3	3.2	2.4	2.4	3.0
	Min/Max	2.0/4.1	2.5/5.4	2.6/4.2	2.4/3.8	2.3/3.1	2.2/2.9	2.3/3.0	2.4/3.3	2.3/3.9	2.3/4.6	1.6/3.0	1.8/3.2	2.6/3.3

Solar Radiation for 1-Axis Tracking Flat-Plate Collectors with a North–South Axis ($kWh/m^2/day$), Uncertainty ±9%

Axis Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	2.5	3.6	4.8	6.0	6.9	7.5	7.7	6.7	5.4	3.8	2.2	1.9	4.9
	Min/Max	1.6/3.2	2.7/5.2	3.8/5.7	4.6/7.4	5.2/8.5	5.8/8.9	6.5/8.7	5.3/8.0	3.6/6.3	2.6/5.3	1.6/2.7	1.4/2.4	4.4/5.4
Latitude -15	Average	3.2	4.4	5.4	6.4	7.0	7.5	7.7	7.0	6.0	4.5	2.7	2.4	5.4
	Min/Max	2.1/4.1	3.2/6.4	4.3/6.4	4.7/7.9	5.3/8.7	5.8/8.9	6.6/8.8	5.6/8.3	3.9/7.0	3.1/6.4	1.9/3.5	1.8/3.2	4.8/5.9
Latitude	Average	3.5	4.6	5.5	6.3	6.8	7.2	7.5	6.9	6.0	4.6	2.9	2.6	5.4
	Min/Max	2.2/4.5	3.3/6.8	4.3/6.6	4.6/7.8	5.1/8.4	5.5/8.6	6.3/8.5	5.4/8.2	3.9/7.0	3.2/6.8	2.0/3.7	1.9/3.6	4.8/5.9
Latitude +15	Average	3.6	4.7	5.4	6.0	6.4	6.7	7.0	6.5	5.9	4.6	3.0	2.7	5.2
	Min/Max	2.3/4.8	3.4/7.0	4.2/6.6	4.4/7.6	4.7/7.9	5.1/8.0	5.9/7.9	5.2/7.8	3.8/6.9	3.2/6.8	2.0/3.8	2.0/3.8	4.7/5.7

Solar Radiation for 2-Axis Tracking Flat-Plate Collectors ($kWh/m^2/day$), Uncertainty ±9%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
2-Axis	Average	3.7	4.7	5.5	6.4	7.1	7.7	7.9	7.1	6.1	4.7	3.0	2.8	5.6
	Min/Max	2.4/4.8	3.4/7.0	4.3/6.6	4.7/7.9	5.3/8.8	5.9/9.1	6.7/9.0	5.6/8.4	4.0/7.1	3.2/6.8	2.1/3.8	2.0/3.8	5.0/6.1

Direct Beam Solar Radiation for Concentrating Collectors ($kWh/m^2/day$), Uncertainty ±8%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
1-Axis, E-W	Average	1.9	2.2	2.4	2.7	3.0	3.3	3.4	3.0	2.8	2.3	1.5	1.4	2.5
Horiz Axis	Min/Max	1.1/2.6	1.5/3.8	1.6/3.3	1.5/3.7	1.7/4.3	2.0/4.4	2.4/4.3	2.3/3.9	1.7/3.6	1.5/4.0	0.8/2.1	0.7/2.3	2.2/2.8
1-Axis, N-S	Average	1.3	1.9	2.7	3.5	3.9	4.4	4.5	3.9	3.3	2.2	1.1	0.9	2.8
Horiz Axis	Min/Max	0.8/1.8	1.3/3.4	1.8/3.6	2.0/4.8	2.1/5.6	2.5/5.8	3.1/5.6	3.0/5.2	2.0/4.2	1.4/3.8	0.6/1.6	0.4/1.5	2.5/3.2
1-Axis, N-S	Average	2.0	2.7	3.2	3.7	3.9	4.1	4.3	4.1	3.7	2.9	1.6	1.4	3.2
Tilt=Latitude	Min/Max	1.2/2.8	1.8/4.6	2.2/4.3	2.1/5.1	2.1/5.5	2.4/5.5	3.0/5.4	3.1/5.3	2.3/4.8	1.9/5.0	0.9/2.4	0.8/2.4	2.8/3.6
2-Axis	Average	2.2	2.7	3.2	3.8	4.1	4.5	4.7	4.2	3.8	3.0	1.7	1.6	3.3
	Min/Max	1.3/3.0	1.8/4.7	2.2/4.3	2.1/5.2	2.2/5.8	2.6/5.9	3.2/5.8	3.2/5.5	2.3/4.8	1.9/5.1	0.9/2.5	0.8/2.6	2.9/3.7

Average Climatic Conditions

Element	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
Temperature (°C)	-6.3	-4.7	1.3	8.0	14.2	19.4	22.1	20.9	16.3	10.1	4.3	-3.1	8.6
Daily Minimum Temp	-11.7	-10.1	-4.2	1.7	7.4	12.6	15.3	14.3	9.7	3.7	-0.7	-7.7	2.6
Daily Maximum Temp	-1.0	0.7	6.7	14.2	20.9	26.1	28.9	27.4	22.9	16.6	9.3	1.6	14.5
Record Minimum Temp	-33.3	-29.4	-29.4	-12.2	-3.3	2.2	4.4	1.1	-4.4	-8.9	-15.0	-30.0	-33.3
Record Maximum Temp	16.7	19.4	30.0	33.3	34.4	37.2	37.8	37.2	37.8	31.7	27.8	21.7	37.8
HDD, Base 18.3°C	764	646	529	310	137	19	0	7	78	255	422	663	3830
CDD, Base 18.3°C	0	0	0	0	10	51	118	86	17	0	0	0	282
Relative Humidity (%)	71	68	65	61	66	70	71	74	76	72	73	74	70
Wind Speed (m/s)	4.2	4.4	4.7	4.6	3.9	3.6	3.3	3.1	3.3	3.4	4.1	4.2	3.9

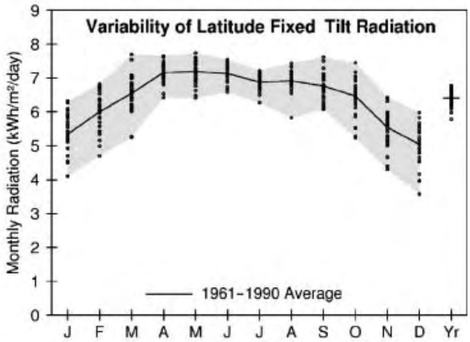
FIGURE 19.9 Solar radiation data manual table for Albany, NY.

Albuquerque, NM

WBAN NO. 23050

LATITUDE: 35.05° N
LONGITUDE: 106.62° W
ELEVATION: 1619 m
MEAN PRESSURE: 838 millibars

STATION TYPE: Primary



Solar Radiation for Flat-Plate Collectors Facing South at a Fixed Tilt (kWh/m²/day), Uncertainty ±9%

Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	3.2	4.2	5.4	6.8	7.7	8.1	7.5	6.9	5.9	4.7	3.5	2.9	5.6
	Min/Max	2.6/3.7	3.4/4.7	4.4/6.2	6.2/7.3	6.9/8.3	7.4/8.7	6.8/7.9	5.9/7.4	5.4/6.6	4.0/5.3	2.8/3.9	2.2/3.3	5.1/5.9
Latitude -15	Average	4.6	5.4	6.3	7.3	7.7	7.8	7.4	7.2	6.6	5.9	4.8	4.3	6.3
	Min/Max	3.6/5.4	4.3/6.1	5.1/7.3	6.5/7.8	6.8/8.3	7.2/8.3	6.7/7.8	6.1/7.7	6.0/7.4	4.9/6.8	3.8/5.5	3.1/5.0	5.7/6.6
Latitude	Average	5.3	6.0	6.5	7.2	7.2	7.1	6.9	6.9	6.8	6.5	5.5	5.0	6.4
	Min/Max	4.1/6.3	4.7/6.8	5.3/7.7	6.4/7.7	6.4/7.7	6.6/7.5	6.3/7.2	5.8/7.4	6.1/7.6	5.2/7.4	4.3/6.4	3.6/6.0	5.8/6.8
Latitude +15	Average	5.8	6.2	6.5	6.6	6.3	6.1	6.0	6.3	6.5	6.6	5.9	5.5	6.2
	Min/Max	4.4/6.9	4.8/7.1	5.1/7.6	5.9/7.1	5.7/6.8	5.6/6.4	5.5/6.3	5.3/6.7	5.8/7.3	5.3/7.7	4.6/6.9	3.8/6.6	5.5/6.5
90	Average	5.2	5.1	4.5	3.7	2.8	2.4	2.5	3.2	4.2	5.1	5.2	5.1	4.1
	Min/Max	3.9/6.4	3.9/5.8	3.5/5.4	3.4/4.0	2.5/3.0	2.2/2.5	2.3/2.7	2.8/3.4	3.7/4.6	4.0/6.0	3.9/6.2	3.5/6.2	3.5/4.4

Solar Radiation for 1-Axis Tracking Flat-Plate Collectors with a North–South Axis (kWh/m²/day), Uncertainty ±9%

Axis Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	4.9	6.2	7.6	9.6	10.6	10.9	9.9	9.2	8.3	7.0	5.3	4.5	7.8
	Min/Max	3.7/5.9	4.7/7.1	5.7/9.3	8.4/10.6	9.0/11.8	9.5/12.1	8.5/10.8	7.4/10.3	7.0/9.6	5.5/8.4	4.1/6.3	3.2/5.4	6.9/8.3
Latitude -15	Average	5.9	7.1	8.3	10.0	10.6	10.8	9.9	9.5	8.8	7.9	6.3	5.5	8.4
	Min/Max	4.4/7.2	5.4/8.1	6.2/10.2	8.7/11.0	9.1/11.9	9.5/12.0	8.5/10.8	7.6/10.6	7.5/10.4	6.1/9.5	4.8/7.5	3.8/6.7	7.4/8.9
Latitude	Average	6.5	7.5	8.6	9.9	10.3	10.4	9.5	9.3	9.0	8.3	6.8	6.1	8.5
	Min/Max	4.8/7.9	5.7/8.7	6.4/10.5	8.7/10.9	8.8/11.6	9.1/11.5	8.1/10.4	7.5/10.4	7.6/10.5	6.3/10.0	5.2/8.1	4.2/7.4	7.5/9.1
Latitude +15	Average	6.9	7.7	8.5	9.5	9.7	9.7	8.9	8.9	8.8	8.4	7.1	6.5	8.4
	Min/Max	5.1/8.4	5.8/8.9	6.3/10.4	8.3/10.6	8.3/10.9	8.4/10.7	7.6/9.8	7.1/9.9	7.4/10.3	6.4/10.1	5.3/8.5	4.4/7.9	7.3/8.9

Solar Radiation for 2-Axis Tracking Flat-Plate Collectors (kWh/m²/day), Uncertainty ±9%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
2-Axis	Average	6.9	7.7	8.6	10.0	10.8	11.1	10.0	9.5	9.0	8.4	7.2	6.6	8.8
	Min/Max	5.1/8.5	5.8/8.9	6.4/10.5	8.8/11.1	9.2/12.1	9.7/12.3	8.6/11.0	7.7/10.6	7.6/10.5	6.4/10.1	5.3/8.6	4.4/8.1	7.7/9.4

Direct Beam Solar Radiation for Concentrating Collectors (kWh/m²/day), Uncertainty ±8%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
1-Axis, E-W	Average	4.5	4.6	4.6	5.3	5.9	6.3	5.5	5.2	4.9	5.1	4.6	4.4	5.1
Horiz Axis	Min/Max	3.1/5.8	3.0/5.7	3.0/6.2	4.3/6.1	4.7/7.1	5.3/7.4	4.3/6.4	3.7/6.0	3.9/6.2	3.3/6.5	2.8/5.9	2.6/5.7	4.3/5.5
1-Axis, N-S	Average	3.7	4.6	5.6	7.2	8.0	8.4	7.2	6.7	6.2	5.4	4.0	3.4	5.9
Horiz Axis	Min/Max	2.5/4.8	3.0/5.7	3.4/7.6	5.7/8.4	6.1/9.7	6.8/10.0	5.4/8.4	4.7/7.9	4.7/7.9	3.4/7.2	2.4/5.2	2.0/4.5	4.8/6.4
1-Axis, N-S	Average	5.1	5.7	6.3	7.4	7.7	7.8	6.8	6.8	6.7	6.5	5.4	4.8	6.4
Tilt=Latitude	Min/Max	3.4/6.6	3.7/7.2	3.9/8.6	5.8/8.6	5.9/9.3	6.4/9.4	5.1/7.9	4.7/8.0	5.1/8.6	4.1/8.6	3.2/6.9	2.8/6.3	5.3/7.0
2-Axis	Average	5.4	5.9	6.3	7.5	8.1	8.5	7.3	7.0	6.8	6.6	5.7	5.2	6.7
	Min/Max	3.7/7.1	3.8/7.3	3.9/8.6	5.9/8.8	6.2/9.9	6.9/10.2	5.5/8.5	4.9/8.2	5.2/8.7	4.2/8.6	3.4/7.3	3.1/6.8	5.6/7.3

Average Climatic Conditions

Element	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
Temperature (°C)	1.2	4.4	8.3	12.9	17.9	23.4	25.8	24.4	20.3	13.9	6.8	1.8	13.4
Daily Minimum Temp	-5.7	-3.1	0.1	4.2	9.2	14.6	18.0	17.0	12.9	6.1	-0.4	-4.9	5.7
Daily Maximum Temp	8.2	11.9	16.3	21.6	26.5	32.2	33.6	31.7	27.7	21.7	14.1	8.6	21.2
Record Minimum Temp	-27.2	-20.6	-13.3	-7.2	-2.2	4.4	11.1	11.1	2.8	-6.1	-21.7	-21.7	-27.2
Record Maximum Temp	20.6	24.4	29.4	31.7	36.7	40.6	40.6	38.3	37.8	32.8	25.0	22.2	40.6
HDD, Base 18.3°C	531	389	312	167	49	0	0	0	10	144	345	512	2458
CDD, Base 18.3°C	0	0	0	4	36	155	233	188	70	6	0	0	691
Relative Humidity (%)	56	50	40	33	31	30	42	47	48	45	50	57	44
Wind Speed (m/s)	3.7	3.9	4.5	4.9	4.8	4.5	4.0	3.8	3.8	3.6	3.6	3.5	4.1

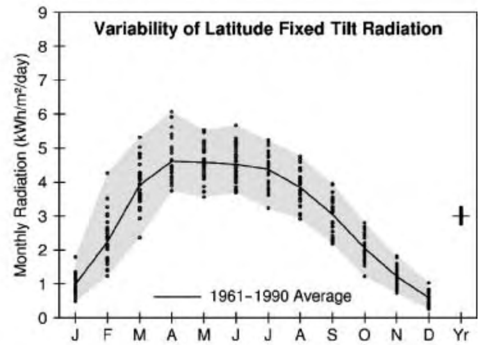
FIGURE 19.10 Solar radiation data manual table for Albuquerque, NM.

Anchorage, AK

WBAN NO. 26451

LATITUDE: 61.17° N
LONGITUDE: 150.02° W
ELEVATION: 35 m
MEAN PRESSURE: 1004 millibars

STATION TYPE: Secondary



Solar Radiation for Flat-Plate Collectors Facing South at a Fixed Tilt (kWh/m²/day), Uncertainty ±9%

Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	0.3	1.0	2.3	3.6	4.6	4.9	4.6	3.5	2.2	1.1	0.4	0.2	2.4
	Min/Max	0.2/0.4	0.7/1.3	1.7/2.6	3.2/4.5	3.8/5.3	4.1/6.0	3.7/5.4	2.9/4.2	1.9/2.6	0.9/1.4	0.3/0.5	0.1/0.2	2.2/2.5
Latitude -15	Average	0.9	2.1	3.8	4.7	4.9	5.0	4.8	4.1	3.1	2.0	1.1	0.5	3.1
	Min/Max	0.4/1.6	1.2/3.8	2.3/5.0	3.9/6.1	3.9/5.9	4.1/6.2	3.5/5.7	3.1/5.0	2.3/3.9	1.2/2.7	0.7/1.6	0.3/0.9	2.8/3.3
Latitude	Average	1.0	2.2	3.9	4.6	4.6	4.5	4.4	3.8	3.1	2.1	1.2	0.6	3.0
	Min/Max	0.5/1.8	1.2/4.3	2.4/5.3	3.7/6.1	3.6/5.5	3.7/5.7	3.2/5.2	2.9/4.8	2.2/4.0	1.2/2.8	0.7/1.8	0.3/1.0	2.8/3.2
Latitude +15	Average	1.0	2.3	3.9	4.3	4.0	3.9	3.8	3.4	2.9	2.0	1.3	0.6	2.8
	Min/Max	0.5/1.9	1.2/4.5	2.3/5.3	3.4/5.8	3.1/4.8	3.2/4.9	2.8/4.5	2.6/4.3	2.0/3.8	1.2/2.8	0.7/1.9	0.3/1.1	2.6/3.0
90	Average	1.0	2.3	3.7	3.8	3.3	3.2	3.1	2.9	2.6	1.9	1.3	0.6	2.5
	Min/Max	0.5/1.9	1.2/4.4	2.2/5.1	3.0/5.3	2.6/4.0	2.6/4.0	2.3/3.7	2.2/3.6	1.8/3.4	1.1/2.6	0.7/1.9	0.3/1.1	2.3/2.8

Solar Radiation for 1-Axis Tracking Flat-Plate Collectors with a North–South Axis (kWh/m²/day), Uncertainty ±9%

Axis Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	0.5	1.5	3.5	5.2	6.1	6.3	6.0	4.7	3.1	1.6	0.7	0.3	3.3
	Min/Max	0.3/0.8	0.9/2.8	2.1/4.7	4.1/6.9	4.6/7.6	4.7/8.4	4.3/7.5	3.4/5.9	2.2/3.9	1.0/2.2	0.5/0.9	0.1/0.4	2.9/3.6
Latitude -15	Average	0.9	2.4	4.6	6.1	6.6	6.6	6.3	5.2	3.8	2.2	1.2	0.5	3.9
	Min/Max	0.5/1.7	1.3/4.6	2.6/6.5	4.8/8.3	4.8/8.3	4.8/8.9	4.3/8.0	3.6/6.7	2.5/4.9	1.3/3.2	0.7/1.8	0.3/0.9	3.4/4.2
Latitude	Average	1.0	2.5	4.8	6.1	6.4	6.3	6.1	5.1	3.8	2.3	1.3	0.6	3.9
	Min/Max	0.5/1.9	1.3/5.0	2.7/6.7	4.7/8.3	4.7/8.1	4.6/8.6	4.1/7.7	3.5/6.6	2.5/5.0	1.3/3.3	0.8/2.0	0.3/1.1	3.4/4.2
Latitude +15	Average	1.1	2.6	4.7	5.8	6.0	5.9	5.7	4.8	3.6	2.3	1.4	0.6	3.7
	Min/Max	0.5/2.0	1.3/5.1	2.6/6.7	4.5/8.1	4.3/7.6	4.2/8.1	3.8/7.3	3.3/6.2	2.4/4.8	1.3/3.3	0.8/2.1	0.3/1.2	3.3/4.1

Solar Radiation for 2-Axis Tracking Flat-Plate Collectors (kWh/m²/day), Uncertainty ±9%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
2-Axis	Average	1.1	2.6	4.8	6.1	6.7	6.8	6.5	5.2	3.8	2.4	1.4	0.7	4.0
	Min/Max	0.5/2.1	1.3/5.1	2.7/6.7	4.8/8.3	5.0/8.4	5.0/9.2	4.5/8.2	3.7/6.8	2.6/5.0	1.3/3.3	0.8/2.1	0.3/1.2	3.6/4.4

Direct Beam Solar Radiation for Concentrating Collectors (kWh/m²/day), Uncertainty ±8%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
1-Axis, E-W	Average	0.6	1.3	2.2	2.6	2.8	2.8	2.6	2.1	1.7	1.2	0.8	0.4	1.8
Horiz Axis	Min/Max	0.2/1.5	0.4/3.4	0.8/3.7	1.5/4.3	1.6/4.0	1.6/4.5	1.2/3.8	1.2/3.2	0.9/2.6	0.5/1.9	0.3/1.4	0.1/0.8	1.5/2.0
1-Axis, N-S	Average	0.2	0.8	1.9	3.0	3.5	3.4	3.3	2.5	1.7	0.8	0.3	0.1	1.8
Horiz Axis	Min/Max	0.1/0.5	0.2/2.0	0.6/3.3	1.7/4.9	2.0/5.0	1.8/5.5	1.5/4.8	1.3/3.7	0.8/2.5	0.3/1.4	0.1/0.6	0.0/0.2	1.4/2.1
1-Axis, N-S	Average	0.6	1.5	2.8	3.6	3.7	3.5	3.4	2.9	2.3	1.4	0.8	0.4	2.3
Tilt=Latitude	Min/Max	0.2/1.5	0.4/3.9	1.0/4.8	2.1/6.0	2.2/5.4	1.9/5.7	1.6/5.0	1.5/4.3	1.1/3.5	0.6/2.3	0.3/1.5	0.1/0.8	1.8/2.6
2-Axis	Average	0.7	1.5	2.8	3.7	4.0	3.8	3.7	3.0	2.3	1.4	0.9	0.4	2.3
	Min/Max	0.2/1.6	0.4/4.0	1.0/4.9	2.1/6.1	2.3/5.7	2.1/6.2	1.7/5.4	1.6/4.4	1.1/3.5	0.6/2.3	0.3/1.5	0.1/0.9	1.9/2.7

Average Climatic Conditions

Element	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
Temperature (°C)	-9.5	-7.4	-3.5	2.1	8.1	12.4	14.7	13.5	9.1	1.4	-6.0	-8.7	2.2
Daily Minimum Temp	-13.1	-11.4	-7.7	-1.9	3.8	8.4	10.9	9.7	5.3	-1.8	-9.4	-12.2	-1.6
Daily Maximum Temp	-5.9	-3.4	0.6	6.0	12.4	16.4	18.4	17.2	12.9	4.7	-2.7	-5.3	5.9
Record Minimum Temp	-36.7	-32.2	-31.1	-20.0	-8.3	0.6	3.3	-0.6	-6.7	-20.6	-29.4	-34.4	-36.7
Record Maximum Temp	10.0	8.9	10.6	18.3	25.0	29.4	27.8	27.8	22.8	16.1	11.7	8.3	29.4
HDD, Base 18.3°C	863	720	677	487	317	177	114	150	277	523	730	839	5872
CDD, Base 18.3°C	0	0	0	0	0	0	0	0	0	0	0	0	0
Relative Humidity (%)	74	71	66	64	62	66	71	75	76	75	77	77	71
Wind Speed (m/s)	2.9	3.1	3.0	3.3	3.7	3.8	3.2	3.1	3.0	3.0	2.9	2.8	3.1

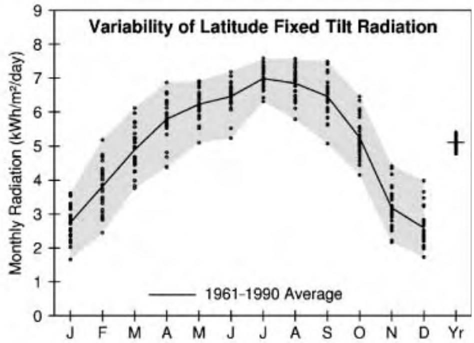
FIGURE 19.11 Solar radiation data manual table for Anchorage, AK.

Boise, ID

WBAN NO. 24131

LATITUDE: 43.57° N
LONGITUDE: 116.22° W
ELEVATION: 874 m
MEAN PRESSURE: 916 millibars

STATION TYPE: Primary



Solar Radiation for Flat-Plate Collectors Facing South at a Fixed Tilt (kWh/m²/day), Uncertainty ±9%

Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	1.6	2.5	3.8	5.3	6.5	7.2	7.6	6.6	5.1	3.4	1.9	1.4	4.4
	Min/Max	1.3/2.0	1.9/3.0	3.1/4.4	4.2/6.1	5.5/7.2	5.9/8.0	6.9/8.3	5.6/7.2	4.3/5.8	2.9/3.9	1.4/2.3	1.2/1.8	4.2/4.6
Latitude -15	Average	2.5	3.5	4.7	5.9	6.7	7.1	7.6	7.1	6.3	4.9	2.9	2.3	5.1
	Min/Max	1.6/3.2	2.3/4.7	3.7/5.8	4.5/7.0	5.5/7.4	5.7/7.9	6.8/8.2	6.0/7.9	5.0/7.3	3.9/5.9	2.0/3.9	1.6/3.4	4.8/5.4
Latitude	Average	2.8	3.8	4.9	5.8	6.2	6.5	7.0	6.8	6.5	5.2	3.2	2.6	5.1
	Min/Max	1.7/3.6	2.5/5.2	3.8/6.1	4.4/6.9	5.1/6.9	5.2/7.2	6.3/7.6	5.8/7.6	5.1/7.5	4.2/6.5	2.2/4.4	1.7/4.0	4.8/5.4
Latitude +15	Average	2.9	3.9	4.8	5.4	5.5	5.5	6.0	6.2	6.2	5.3	3.3	2.8	4.8
	Min/Max	1.7/3.9	2.5/5.4	3.6/6.1	4.0/6.4	4.5/6.1	4.5/6.2	5.5/6.5	5.3/6.8	4.8/7.2	4.2/6.6	2.2/4.7	1.8/4.3	4.5/5.1
90	Average	2.7	3.4	3.7	3.6	3.2	3.0	3.3	3.8	4.6	4.5	3.0	2.6	3.5
	Min/Max	1.5/3.8	2.1/4.9	2.8/4.8	2.8/4.3	2.7/3.5	2.6/3.2	3.0/3.5	3.3/4.2	3.5/5.3	3.5/5.6	1.9/4.3	1.7/4.2	3.2/3.8

Solar Radiation for 1-Axis Tracking Flat-Plate Collectors with a North–South Axis (kWh/m²/day), Uncertainty ±9%

Axis Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	2.2	3.5	5.4	7.4	9.0	10.0	10.9	9.5	7.8	5.2	2.7	2.0	6.3
	Min/Max	1.5/2.9	2.3/4.8	3.9/6.7	5.3/8.8	7.1/10.3	7.6/11.4	9.5/12.2	7.6/10.9	6.0/9.3	4.1/6.3	1.8/3.6	1.4/2.9	5.8/6.7
Latitude -15	Average	2.9	4.3	6.1	7.9	9.2	10.0	11.0	10.1	8.7	6.3	3.4	2.7	6.9
	Min/Max	1.7/3.8	2.7/6.0	4.4/7.7	5.5/9.6	7.3/10.6	7.6/11.4	9.6/12.4	8.0/11.6	6.6/10.5	4.8/7.8	2.3/4.8	1.7/4.1	6.3/7.3
Latitude	Average	3.1	4.5	6.2	7.8	9.0	9.6	10.7	9.9	8.8	6.6	3.7	2.9	6.9
	Min/Max	1.8/4.2	2.8/6.4	4.4/8.0	5.4/9.5	7.0/10.3	7.3/11.0	9.3/12.0	7.9/11.4	6.7/10.6	5.0/8.3	2.4/5.2	1.8/4.6	6.3/7.4
Latitude +15	Average	3.2	4.6	6.1	7.5	8.5	9.0	10.0	9.5	8.6	6.7	3.8	3.1	6.7
	Min/Max	1.8/4.4	2.8/6.6	4.3/7.9	5.2/9.2	6.6/9.8	6.8/10.3	8.7/11.3	7.5/10.9	6.5/10.4	5.0/8.4	2.4/5.5	1.9/4.9	6.1/7.2

Solar Radiation for 2-Axis Tracking Flat-Plate Collectors (kWh/m²/day), Uncertainty ±9%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
2-Axis	Average	3.3	4.6	6.2	7.9	9.4	10.3	11.3	10.1	8.8	6.7	3.8	3.1	7.1
	Min/Max	1.8/4.4	2.8/6.6	4.4/8.0	5.6/9.6	7.4/10.8	7.8/11.8	9.8/12.7	8.1/11.6	6.7/10.6	5.1/8.4	2.4/5.5	1.9/5.0	6.5/7.6

Direct Beam Solar Radiation for Concentrating Collectors (kWh/m²/day), Uncertainty ±8%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
1-Axis, E-W	Average	1.7	2.4	3.0	3.8	4.7	5.5	6.4	5.6	5.0	4.0	2.2	1.7	3.8
Horiz Axis	Min/Max	0.4/2.8	1.0/3.9	1.8/4.5	1.8/5.2	3.1/5.9	3.4/6.9	5.1/7.6	3.9/7.0	3.3/6.4	2.6/5.7	0.9/3.6	0.6/3.4	3.3/4.3
1-Axis, N-S	Average	1.1	2.1	3.4	4.9	6.3	7.3	8.7	7.5	6.1	3.9	1.6	1.1	4.5
Horiz Axis	Min/Max	0.3/1.9	0.9/3.4	1.9/5.1	2.4/6.7	4.2/8.0	4.5/9.2	6.8/10.4	5.1/9.4	4.0/7.9	2.5/5.4	0.7/2.6	0.4/2.1	3.8/5.0
1-Axis, N-S	Average	1.8	2.9	4.1	5.3	6.3	7.0	8.4	7.7	6.9	5.0	2.5	1.8	5.0
Tilt=Latitude	Min/Max	0.5/3.0	1.2/4.7	2.4/6.1	2.6/7.2	4.1/7.9	4.3/8.8	6.6/10.0	5.3/9.7	4.6/9.1	3.3/7.2	1.1/4.0	0.6/3.6	4.2/5.6
2-Axis	Average	1.9	3.0	4.1	5.4	6.6	7.6	9.0	8.0	7.0	5.1	2.6	2.0	5.2
	Min/Max	0.5/3.2	1.3/4.8	2.4/6.1	2.6/7.3	4.4/8.3	4.7/9.5	7.1/10.7	5.5/10.0	4.6/9.1	3.3/7.3	1.1/4.3	0.7/3.9	4.4/5.8

Average Climatic Conditions

Element	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
Temperature (°C)	-1.7	2.2	5.8	9.5	14.2	19.2	23.3	22.5	17.0	11.0	4.4	-1.1	10.5
Daily Minimum Temp	-5.8	-2.5	-0.1	2.6	6.6	11.2	14.3	13.8	9.0	3.9	-0.5	-5.3	3.9
Daily Maximum Temp	2.4	6.8	11.6	16.3	21.7	27.2	32.3	31.2	25.0	18.1	9.3	3.2	17.1
Record Minimum Temp	-27.2	-26.1	-14.4	-7.2	-5.6	-0.6	1.7	2.8	-5.0	-11.7	-19.4	-31.7	-31.7
Record Maximum Temp	17.2	21.1	27.2	33.3	36.7	42.8	43.9	43.3	38.9	32.8	23.3	18.3	43.9
HDD, Base 18.3°C	620	453	389	265	134	42	3	11	89	230	418	601	3256
CDD, Base 18.3°C	0	0	0	0	5	67	158	140	49	0	0	0	419
Relative Humidity (%)	75	70	59	52	49	45	36	37	45	54	69	75	55
Wind Speed (m/s)	3.2	3.8	4.2	4.3	4.2	3.9	3.6	3.5	3.5	3.4	3.6	3.3	3.7

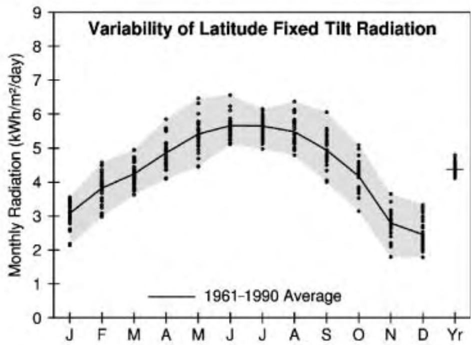
FIGURE 19.12 Solar radiation data manual table for Boise, ID.

Chicago, IL

WBAN NO. 94846

LATITUDE: 41.78° N
LONGITUDE: 87.75° W
ELEVATION: 190 m
MEAN PRESSURE: 992 millibars

STATION TYPE: Secondary



Solar Radiation for Flat-Plate Collectors Facing South at a Fixed Tilt (kWh/m²/day), Uncertainty ±9%

Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	1.8	2.6	3.5	4.6	5.7	6.3	6.1	5.4	4.2	3.0	1.8	1.5	3.9
	Min/Max	1.5/2.1	2.2/3.0	3.1/3.9	4.0/5.3	4.8/6.8	5.7/7.2	5.4/6.6	4.8/6.1	3.6/4.9	2.5/3.5	1.4/2.1	1.2/1.7	3.7/4.2
Latitude -15	Average	2.7	3.5	4.1	5.0	5.8	6.1	6.1	5.7	4.9	3.9	2.5	2.2	4.4
	Min/Max	2.0/3.1	2.8/4.2	3.6/4.8	4.2/5.9	4.8/6.9	5.5/7.1	5.4/6.6	5.0/6.6	4.0/5.9	3.0/4.7	1.7/3.3	1.6/2.9	4.2/4.8
Latitude	Average	3.1	3.8	4.2	4.9	5.4	5.7	5.6	5.5	4.9	4.2	2.8	2.4	4.4
	Min/Max	2.1/3.5	3.0/4.6	3.6/4.9	4.1/5.8	4.4/6.5	5.1/6.6	5.0/6.2	4.8/6.4	4.0/6.1	3.1/5.1	1.8/3.7	1.8/3.3	4.1/4.8
Latitude +15	Average	3.3	3.9	4.1	4.5	4.8	4.9	4.9	5.0	4.7	4.2	2.9	2.6	4.1
	Min/Max	2.2/3.8	3.0/4.7	3.5/4.9	3.8/5.4	4.0/5.7	4.4/5.7	4.3/5.4	4.3/5.8	3.8/5.8	3.1/5.2	1.8/3.8	1.8/3.6	3.8/4.5
90	Average	3.1	3.5	3.2	3.0	2.8	2.6	2.7	3.1	3.4	3.4	2.5	2.4	3.0
	Min/Max	2.0/3.7	2.5/4.3	2.7/3.8	2.5/3.6	2.4/3.2	2.5/3.0	2.5/2.9	2.8/3.5	2.7/4.2	2.5/4.3	1.5/3.5	1.7/3.5	2.7/3.3

Solar Radiation for 1-Axis Tracking Flat-Plate Collectors with a North–South Axis (kWh/m²/day), Uncertainty ±9%

Axis Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	2.5	3.5	4.5	5.9	7.3	8.0	7.9	7.0	5.6	4.1	2.4	1.9	5.1
	Min/Max	1.7/2.9	2.7/4.2	3.9/5.4	4.7/7.3	5.8/9.1	7.1/9.6	6.9/8.9	5.9/8.5	4.4/6.9	3.0/5.0	1.6/3.0	1.5/2.6	4.7/5.7
Latitude -15	Average	3.2	4.2	5.0	6.2	7.4	8.0	7.9	7.3	6.2	4.8	2.9	2.5	5.5
	Min/Max	2.2/3.7	3.2/5.2	4.3/6.1	5.0/7.8	5.9/9.4	7.1/9.6	6.9/9.0	6.1/9.0	4.8/7.7	3.5/6.0	1.8/3.9	1.8/3.4	5.1/6.2
Latitude	Average	3.5	4.5	5.1	6.1	7.2	7.7	7.7	7.2	6.2	5.0	3.1	2.7	5.5
	Min/Max	2.3/4.1	3.3/5.5	4.3/6.2	4.9/7.8	5.7/9.1	6.8/9.2	6.7/8.7	6.0/8.8	4.8/7.8	3.6/6.3	1.9/4.2	1.9/3.8	5.1/6.2
Latitude +15	Average	3.7	4.5	5.0	5.9	6.8	7.2	7.2	6.8	6.0	5.0	3.2	2.9	5.3
	Min/Max	2.4/4.3	3.3/5.6	4.2/6.1	4.6/7.4	5.4/8.6	6.3/8.6	6.2/8.1	5.7/8.4	4.6/7.6	3.5/6.3	1.9/4.4	2.0/4.1	4.9/6.0

Solar Radiation for 2-Axis Tracking Flat-Plate Collectors (kWh/m²/day), Uncertainty ±9%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
2-Axis	Average	3.7	4.6	5.1	6.2	7.5	8.2	8.1	7.4	6.2	5.0	3.2	2.9	5.7
	Min/Max	2.4/4.3	3.4/5.6	4.3/6.2	5.0/7.8	6.0/9.5	7.2/9.8	7.1/9.2	6.2/9.0	4.8/7.8	3.6/6.3	1.9/4.4	2.0/4.1	5.2/6.4

Direct Beam Solar Radiation for Concentrating Collectors (kWh/m²/day), Uncertainty ±8%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
1-Axis, E-W	Average	2.0	2.2	2.2	2.6	3.3	3.7	3.6	3.3	2.9	2.6	1.7	1.5	2.6
Horiz Axis	Min/Max	1.1/2.5	1.3/3.3	1.6/3.1	1.8/3.8	2.1/4.9	3.0/4.9	2.8/4.4	2.3/4.4	1.9/4.1	1.5/3.6	0.7/2.6	0.9/2.5	2.3/3.2
1-Axis, N-S	Average	1.3	1.9	2.4	3.3	4.3	4.8	4.8	4.2	3.4	2.5	1.3	1.0	2.9
Horiz Axis	Min/Max	0.7/1.7	1.1/2.8	1.6/3.4	2.2/4.8	2.8/6.4	4.0/6.4	3.8/5.9	3.0/5.9	2.2/4.7	1.4/3.5	0.5/1.9	0.6/1.6	2.6/3.6
1-Axis, N-S	Average	2.1	2.6	2.9	3.6	4.3	4.6	4.6	4.3	3.9	3.2	1.9	1.6	3.3
Tilt=Latitude	Min/Max	1.2/2.7	1.5/3.9	2.0/4.1	2.4/5.2	2.8/6.3	3.7/6.1	3.7/5.6	3.1/6.1	2.5/5.4	1.9/4.5	0.8/2.9	0.9/2.6	2.9/4.0
2-Axis	Average	2.3	2.7	2.9	3.6	4.5	5.0	4.9	4.5	3.9	3.3	2.0	1.7	3.4
	Min/Max	1.3/2.9	1.6/4.0	2.0/4.1	2.4/5.2	3.0/6.7	4.1/6.6	3.9/6.1	3.2/6.2	2.5/5.4	1.9/4.5	0.8/3.0	1.0/2.8	3.0/4.2

Average Climatic Conditions

Element	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
Temperature (°C)	-6.1	-3.7	2.9	9.2	14.9	20.3	22.9	22.1	18.0	11.6	4.4	-3.0	9.4
Daily Minimum Temp	-10.6	-8.2	-1.9	3.7	8.7	14.2	17.0	16.4	12.2	5.7	-0.2	-7.2	4.2
Daily Maximum Temp	-1.7	0.8	7.7	14.8	21.2	26.4	28.7	27.7	23.8	17.4	9.1	1.1	14.8
Record Minimum Temp	-32.8	-27.2	-22.2	-13.9	-4.4	2.2	4.4	5.0	-2.2	-8.3	-17.2	-31.7	-32.8
Record Maximum Temp	18.3	21.7	31.1	32.8	33.9	40.0	38.9	38.3	37.2	32.8	25.6	21.7	40.0
HDD, Base 18.3°C	758	616	479	273	131	19	3	11	47	217	417	661	3631
CDD, Base 18.3°C	0	0	0	0	26	79	144	126	37	7	0	0	418
Relative Humidity (%)	72	72	70	65	64	66	68	71	71	69	73	76	70
Wind Speed (m/s)	5.2	5.1	5.4	5.3	4.6	4.2	3.7	3.6	4.0	4.5	4.8	4.9	4.6

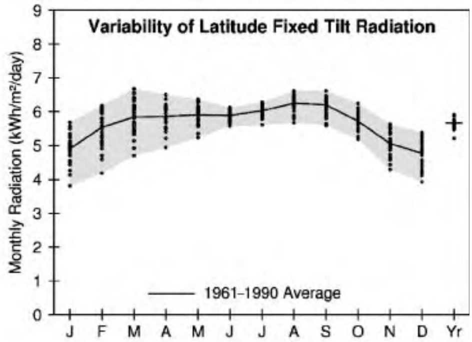
FIGURE 19.13 Solar radiation data manual table for Chicago, IL.

Honolulu, HI

WBAN NO. 22521

LATITUDE: 21.33° N
LONGITUDE: 157.92° W
ELEVATION: 5 m
MEAN PRESSURE: 1016 millibars

STATION TYPE: Primary



Solar Radiation for Flat-Plate Collectors Facing South at a Fixed Tilt (kWh/m²/day), Uncertainty ±9%

Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	3.9	4.7	5.4	5.9	6.4	6.5	6.6	6.5	5.9	5.0	4.1	3.7	5.4
	Min/Max	3.2/4.4	3.7/5.2	4.4/6.1	5.0/6.5	5.6/6.9	6.2/6.9	6.1/7.0	5.8/6.9	5.4/6.3	4.6/5.4	3.7/4.5	3.2/4.1	5.0/5.6
Latitude -15	Average	4.3	5.0	5.6	5.9	6.3	6.4	6.5	6.5	6.1	5.3	4.5	4.1	5.5
	Min/Max	3.4/4.8	3.9/5.5	4.6/6.3	5.1/6.6	5.6/6.8	6.1/6.7	6.0/6.8	5.9/6.9	5.5/6.5	4.9/5.7	3.9/4.9	3.5/4.6	5.1/5.8
Latitude	Average	4.9	5.5	5.8	5.9	5.9	5.9	6.0	6.2	6.2	5.7	5.1	4.8	5.7
	Min/Max	3.8/5.7	4.2/6.2	4.7/6.7	4.9/6.5	5.2/6.4	5.6/6.1	5.6/6.3	5.7/6.6	5.6/6.6	5.2/6.2	4.3/5.6	3.9/5.4	5.2/5.9
Latitude +15	Average	5.3	5.8	5.8	5.5	5.3	5.1	5.3	5.7	6.0	5.8	5.4	5.2	5.5
	Min/Max	4.0/6.2	4.3/6.5	4.6/6.6	4.6/6.1	4.7/5.6	4.8/5.3	4.9/5.4	5.2/6.0	5.4/6.4	5.3/6.4	4.5/6.0	4.2/5.9	5.0/5.7
90	Average	4.2	4.0	3.2	2.2	1.5	1.4	1.4	1.8	2.8	3.7	4.1	4.3	2.9
	Min/Max	3.1/5.0	2.9/4.6	2.6/3.6	1.9/2.3	1.5/1.6	1.4/1.5	1.4/1.5	1.8/1.9	2.6/3.0	3.3/4.1	3.3/4.7	3.3/4.9	2.6/3.1

Solar Radiation for 1-Axis Tracking Flat-Plate Collectors with a North–South Axis (kWh/m²/day), Uncertainty ±9%

Axis Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	5.3	6.4	7.2	7.7	8.3	8.5	8.7	8.7	8.0	6.8	5.6	5.1	7.2
	Min/Max	4.1/6.3	4.7/7.3	5.4/8.4	6.2/8.9	6.9/9.4	7.9/9.3	7.9/9.3	7.4/9.4	7.0/8.7	6.0/7.6	4.5/6.2	4.0/5.9	6.4/7.6
Latitude -15	Average	5.6	6.6	7.3	7.7	8.3	8.5	8.6	8.7	8.1	7.0	5.8	5.3	7.3
	Min/Max	4.2/6.6	4.8/7.6	5.5/8.6	6.2/9.0	6.9/9.3	7.8/9.2	7.9/9.3	7.4/9.4	7.1/8.9	6.2/7.8	4.7/6.5	4.2/6.2	6.5/7.7
Latitude	Average	6.1	7.0	7.5	7.7	8.0	8.1	8.3	8.5	8.2	7.3	6.2	5.9	7.4
	Min/Max	4.5/7.3	5.0/8.1	5.7/8.9	6.1/8.9	6.7/9.0	7.5/8.8	7.6/8.9	7.3/9.2	7.2/9.0	6.4/8.2	5.0/7.0	4.6/6.8	6.6/7.8
Latitude +15	Average	6.3	7.2	7.5	7.4	7.5	7.5	7.7	8.1	8.0	7.4	6.5	6.2	7.3
	Min/Max	4.7/7.6	5.1/8.3	5.6/8.8	5.9/8.6	6.3/8.5	7.0/8.2	7.1/8.3	6.9/8.8	7.0/8.8	6.5/8.3	5.2/7.3	4.8/7.2	6.4/7.7

Solar Radiation for 2-Axis Tracking Flat-Plate Collectors (kWh/m²/day), Uncertainty ±9%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
2-Axis	Average	6.4	7.2	7.5	7.8	8.3	8.6	8.8	8.7	8.2	7.4	6.5	6.3	7.7
	Min/Max	4.7/7.8	5.1/8.3	5.7/8.9	6.2/9.0	7.0/9.4	8.0/9.4	8.0/9.4	7.4/9.4	7.2/9.0	6.5/8.3	5.2/7.4	4.8/7.3	6.8/8.1

Direct Beam Solar Radiation for Concentrating Collectors (kWh/m²/day), Uncertainty ±8%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
1-Axis, E-W	Average	3.5	3.8	3.7	3.6	4.1	4.4	4.6	4.4	4.1	3.8	3.5	3.5	3.9
Horiz Axis	Min/Max	2.2/4.8	2.1/4.8	2.3/4.8	2.5/4.7	3.0/5.1	3.9/5.2	3.8/5.3	3.3/5.3	3.2/4.7	2.9/4.5	2.3/4.4	2.2/4.4	3.2/4.3
1-Axis, N-S	Average	3.5	4.3	4.7	4.9	5.5	5.7	6.0	6.0	5.5	4.5	3.6	3.3	4.8
Horiz Axis	Min/Max	2.1/4.7	2.4/5.5	2.9/6.2	3.3/6.4	3.9/6.9	5.0/6.8	5.0/7.0	4.4/7.1	4.3/6.3	3.5/5.5	2.3/4.4	2.0/4.2	3.9/5.3
1-Axis, N-S	Average	4.1	4.8	5.0	4.9	5.2	5.4	5.7	5.9	5.6	4.9	4.1	4.0	4.9
Tilt=Latitude	Min/Max	2.5/5.5	2.7/6.1	3.1/6.5	3.3/6.4	3.7/6.6	4.7/6.4	4.7/6.6	4.3/6.9	4.4/6.5	3.8/5.9	2.7/5.1	2.4/5.0	4.0/5.4
2-Axis	Average	4.3	4.9	5.0	4.9	5.5	5.8	6.1	6.0	5.6	5.0	4.3	4.3	5.2
	Min/Max	2.7/5.9	2.8/6.3	3.1/6.6	3.4/6.5	3.9/7.0	5.1/6.9	5.1/7.1	4.4/7.1	4.4/6.5	3.9/6.0	2.8/5.4	2.6/5.4	4.2/5.6

Average Climatic Conditions

Element	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
Temperature (°C)	22.7	22.8	23.6	24.3	25.3	26.3	26.9	27.4	27.2	26.4	25.1	23.4	25.1
Daily Minimum Temp	18.7	18.6	19.6	20.4	21.3	22.3	23.1	23.4	23.1	22.4	21.3	19.4	21.1
Daily Maximum Temp	26.7	26.9	27.6	28.2	29.3	30.3	30.8	31.5	31.4	30.5	28.9	27.3	29.1
Record Minimum Temp	11.7	11.7	12.8	13.9	15.6	18.3	18.9	19.4	18.9	17.8	13.9	12.2	11.7
Record Maximum Temp	30.6	31.1	31.1	31.7	33.9	33.3	33.3	33.9	34.4	34.4	33.9	31.7	34.4
HDD, Base 18.3°C	0	0	0	0	0	0	0	0	0	0	0	0	0
CDD, Base 18.3°C	136	124	162	180	216	240	267	282	267	252	203	157	2486
Relative Humidity (%)	73	71	69	67	66	64	65	64	66	68	70	72	68
Wind Speed (m/s)	4.3	4.3	5.0	5.1	5.1	5.5	5.7	5.4	4.9	4.6	4.6	4.3	4.9

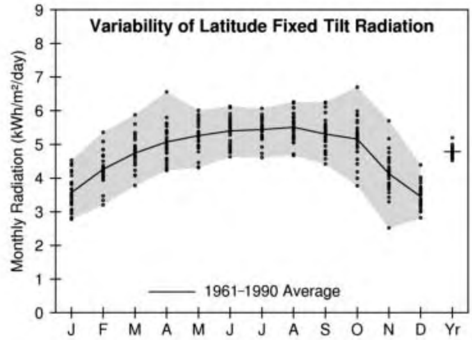
FIGURE 19.14 Solar radiation data manual table for Honolulu, HI.

Houston, TX

WBAN NO. 12960

LATITUDE: 29.98° N
LONGITUDE: 95.37° W
ELEVATION: 33 m
MEAN PRESSURE: 1014 millibars

STATION TYPE: Secondary



Solar Radiation for Flat-Plate Collectors Facing South at a Fixed Tilt (kWh/m²/day), Uncertainty ±9%

Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	2.7	3.4	4.2	5.0	5.6	6.0	5.9	5.6	4.9	4.2	3.1	2.5	4.4
	Min/Max	2.2/3.2	2.8/4.1	3.6/5.1	4.3/6.4	4.6/6.4	5.1/6.9	5.0/6.7	4.8/6.3	4.2/5.7	3.3/5.2	2.2/3.9	2.2/2.9	4.2/4.8
Latitude -15	Average	3.2	3.9	4.6	5.2	5.6	5.9	5.8	5.7	5.2	4.8	3.7	3.0	4.7
	Min/Max	2.6/3.9	3.1/4.8	3.8/5.6	4.3/6.7	4.6/6.4	5.0/6.7	4.9/6.6	4.8/6.5	4.4/6.1	3.6/6.1	2.4/4.9	2.6/3.8	4.5/5.1
Latitude	Average	3.6	4.3	4.7	5.1	5.3	5.4	5.4	5.5	5.3	5.2	4.1	3.5	4.8
	Min/Max	2.8/4.5	3.2/5.4	3.8/5.9	4.2/6.6	4.3/6.0	4.6/6.1	4.6/6.1	4.7/6.3	4.4/6.3	3.8/6.7	2.5/5.7	2.8/4.4	4.5/5.2
Latitude +15	Average	3.8	4.4	4.6	4.7	4.7	4.7	4.8	5.0	5.1	5.2	4.3	3.7	4.6
	Min/Max	2.9/4.9	3.2/5.6	3.6/5.8	3.9/6.1	3.8/5.3	4.1/5.3	4.1/5.3	4.3/5.7	4.2/6.0	3.8/6.9	2.5/6.1	2.9/4.8	4.3/5.0
90	Average	3.1	3.3	3.0	2.5	2.0	1.8	1.9	2.4	3.0	3.8	3.5	3.2	2.8
	Min/Max	2.3/4.2	2.3/4.3	2.3/3.8	2.2/3.0	1.8/2.2	1.7/1.9	1.8/2.0	2.1/2.6	2.5/3.6	2.6/5.0	1.9/5.1	2.4/4.2	2.6/3.1

Solar Radiation for 1-Axis Tracking Flat-Plate Collectors with a North–South Axis (kWh/m²/day), Uncertainty ±9%

Axis Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	3.5	4.5	5.4	6.3	7.0	7.6	7.5	7.2	6.4	5.7	4.1	3.3	5.7
	Min/Max	2.7/4.5	3.3/5.8	4.1/7.0	4.9/8.8	5.3/8.3	6.1/9.0	6.0/8.9	5.9/8.4	5.3/7.7	3.9/7.4	2.4/5.8	2.7/4.2	5.3/6.3
Latitude -15	Average	3.9	4.9	5.7	6.4	7.0	7.5	7.4	7.3	6.7	6.1	4.6	3.7	5.9
	Min/Max	3.0/5.1	3.5/6.3	4.3/7.4	5.0/9.0	5.3/8.3	6.0/8.9	6.0/8.8	6.0/8.5	5.5/8.0	4.2/8.1	2.6/6.6	3.0/4.8	5.6/6.5
Latitude	Average	4.2	5.1	5.8	6.4	6.8	7.2	7.2	7.2	6.7	6.4	4.9	4.0	6.0
	Min/Max	3.1/5.5	3.7/6.7	4.3/7.6	4.9/8.9	5.1/8.1	5.8/8.5	5.8/8.5	5.8/8.4	5.5/8.1	4.3/8.5	2.7/7.2	3.2/5.3	5.6/6.6
Latitude +15	Average	4.3	5.2	5.7	6.1	6.4	6.7	6.7	6.8	6.6	6.4	5.1	4.2	5.9
	Min/Max	3.2/5.8	3.7/6.9	4.2/7.6	4.7/8.6	4.8/7.6	5.4/7.9	5.4/7.9	5.6/8.0	5.3/8.0	4.3/8.6	2.7/7.5	3.3/5.6	5.5/6.5

Solar Radiation for 2-Axis Tracking Flat-Plate Collectors (kWh/m²/day), Uncertainty ±9%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
2-Axis	Average	4.4	5.2	5.8	6.5	7.1	7.7	7.6	7.3	6.7	6.5	5.1	4.2	6.2
	Min/Max	3.2/5.8	3.7/6.9	4.3/7.6	5.0/9.0	5.3/8.4	6.1/9.1	6.1/8.9	6.0/8.6	5.5/8.1	4.3/8.6	2.7/7.5	3.3/5.7	5.8/6.8

Direct Beam Solar Radiation for Concentrating Collectors (kWh/m²/day), Uncertainty ±8%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
1-Axis, E-W	Average	2.3	2.7	2.7	2.8	3.1	3.4	3.3	3.2	3.1	3.4	2.8	2.3	2.9
Horiz Axis	Min/Max	1.4/3.6	1.4/4.1	1.4/3.9	1.6/4.8	1.6/4.1	2.4/4.6	2.2/4.5	2.3/4.1	2.1/4.1	1.7/4.9	0.8/4.8	1.5/3.7	2.6/3.3
1-Axis, N-S	Average	2.0	2.8	3.2	3.7	4.1	4.5	4.4	4.3	3.9	3.7	2.6	1.9	3.4
Horiz Axis	Min/Max	1.2/3.2	1.5/4.3	1.7/4.8	2.1/6.5	2.1/5.4	3.0/6.0	3.0/6.0	3.1/5.4	2.8/5.3	1.9/5.4	0.8/4.5	1.2/3.1	3.1/3.9
1-Axis, N-S	Average	2.6	3.3	3.6	3.7	3.9	4.2	4.1	4.3	4.1	4.3	3.3	2.5	3.7
Tilt=Latitude	Min/Max	1.6/4.1	1.8/5.1	1.9/5.3	2.2/6.6	2.0/5.2	2.8/5.6	2.8/5.6	3.0/5.3	2.9/5.6	2.2/6.3	1.0/5.6	1.6/4.0	3.3/4.2
2-Axis	Average	2.8	3.4	3.6	3.8	4.1	4.6	4.4	4.4	4.2	4.3	3.4	2.8	3.8
	Min/Max	1.7/4.3	1.8/5.2	1.9/5.3	2.2/6.7	2.1/5.5	3.1/6.1	3.0/6.0	3.1/5.5	3.0/5.6	2.2/6.3	1.0/5.9	1.8/4.4	3.4/4.3

Average Climatic Conditions

Element	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
Temperature (°C)	10.2	12.2	15.9	20.2	23.6	26.9	28.1	27.9	25.7	20.9	16.1	11.9	19.9
Daily Minimum Temp	4.3	5.9	10.0	14.5	18.0	21.4	22.4	22.2	19.9	14.2	9.8	5.7	14.1
Daily Maximum Temp	16.1	18.5	21.7	25.8	29.2	32.3	33.7	33.6	31.3	27.6	22.4	18.2	25.9
Record Minimum Temp	-11.1	-6.7	-5.6	-0.6	6.7	11.1	16.7	16.7	8.9	0.0	-7.2	-13.9	-13.9
Record Maximum Temp	28.9	32.8	32.8	35.0	36.1	39.4	40.0	41.7	38.9	35.6	31.7	28.3	41.7
HDD, Base 18.3°C	260	179	104	20	0	0	0	0	0	17	101	208	888
CDD, Base 18.3°C	9	6	28	75	164	257	303	298	220	97	34	10	1500
Relative Humidity (%)	75	73	73	74	75	75	75	75	76	74	75	75	75
Wind Speed (m/s)	4.0	4.2	4.4	4.4	3.9	3.6	3.2	3.0	3.3	3.4	3.8	3.9	3.8

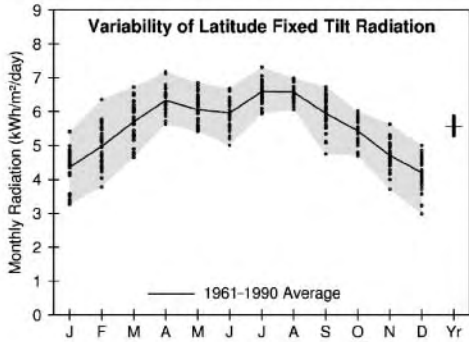
FIGURE 19.15 Solar radiation data manual table for Houston, TX.

Los Angeles, CA

WBAN NO. 23174

LATITUDE: 33.93° N
LONGITUDE: 118.40° W
ELEVATION: 32 m
MEAN PRESSURE: 1012 millibars

STATION TYPE: Primary



Solar Radiation for Flat-Plate Collectors Facing South at a Fixed Tilt (kWh/m²/day), Uncertainty ±9%

Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	2.8	3.6	4.8	6.1	6.4	6.6	7.1	6.5	5.3	4.2	3.2	2.6	4.9
	Min/Max	2.3/3.3	3.0/4.4	4.0/5.6	5.5/6.8	5.7/7.2	5.6/7.7	6.4/8.0	6.1/7.0	4.4/5.8	3.8/4.5	2.7/3.6	2.1/3.0	4.7/5.1
Latitude -15	Average	3.8	4.5	5.5	6.4	6.4	6.4	7.1	6.8	5.9	5.0	4.2	3.6	5.5
	Min/Max	2.9/4.6	3.6/5.7	4.5/6.4	5.8/7.3	5.7/7.3	5.4/7.3	6.3/7.9	6.3/7.2	4.7/6.6	4.4/5.6	3.4/4.9	2.7/4.3	5.2/5.7
Latitude	Average	4.4	5.0	5.7	6.3	6.1	6.0	6.6	6.6	6.0	5.4	4.7	4.2	5.6
	Min/Max	3.3/5.4	3.8/6.4	4.7/6.7	5.6/7.2	5.4/6.8	5.0/6.7	5.9/7.3	6.1/7.0	4.8/6.7	4.7/6.0	3.7/5.6	3.0/5.0	5.3/5.9
Latitude +15	Average	4.7	5.1	5.6	5.9	5.4	5.2	5.8	6.0	5.7	5.5	5.0	4.5	5.4
	Min/Max	3.4/5.9	3.8/6.6	4.5/6.6	5.2/6.7	4.8/6.1	4.4/5.8	5.2/6.3	5.5/6.4	4.5/6.5	4.7/6.1	3.9/6.0	3.1/5.4	5.1/5.7
90	Average	4.1	4.1	3.8	3.3	2.5	2.2	2.4	3.0	3.6	4.2	4.3	4.1	3.5
	Min/Max	2.9/5.2	3.0/5.4	3.1/4.5	2.9/3.6	2.3/2.7	2.1/2.3	2.3/2.5	2.8/3.2	2.9/4.1	3.5/4.7	3.2/5.2	2.7/5.0	3.3/3.7

Solar Radiation for 1-Axis Tracking Flat-Plate Collectors with a North–South Axis (kWh/m²/day), Uncertainty ±9%

Axis Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	3.9	5.0	6.4	7.9	8.0	8.1	9.0	8.3	6.9	5.6	4.4	3.7	6.4
	Min/Max	3.0/4.9	3.7/6.4	5.1/7.8	6.9/9.3	6.7/9.4	6.6/9.7	7.6/10.8	7.6/9.3	5.3/8.0	4.8/6.3	3.4/5.3	2.6/4.3	6.1/6.8
Latitude -15	Average	4.7	5.6	6.9	8.2	8.0	8.0	9.0	8.6	7.3	6.3	5.2	4.4	6.9
	Min/Max	3.5/5.9	4.1/7.4	5.4/8.5	7.1/9.7	6.8/9.4	6.5/9.6	7.6/10.8	7.8/9.5	5.6/8.5	5.4/7.0	3.9/6.3	3.0/5.3	6.5/7.3
Latitude	Average	5.1	6.0	7.1	8.2	7.8	7.7	8.7	8.4	7.4	6.6	5.6	4.9	7.0
	Min/Max	3.7/6.5	4.3/7.9	5.5/8.7	7.0/9.6	6.5/9.2	6.3/9.3	7.3/10.4	7.7/9.3	5.6/8.6	5.6/7.4	4.2/6.9	3.3/5.9	6.5/7.4
Latitude +15	Average	5.4	6.1	7.0	7.8	7.3	7.1	8.1	8.0	7.2	6.6	5.8	5.2	6.8
	Min/Max	3.8/6.8	4.3/8.1	5.4/8.6	6.7/9.3	6.1/8.6	5.8/8.6	6.8/9.8	7.3/8.9	5.4/8.4	5.6/7.5	4.4/7.2	3.4/6.3	6.4/7.3

Solar Radiation for 2-Axis Tracking Flat-Plate Collectors (kWh/m²/day), Uncertainty ±9%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
2-Axis	Average	5.4	6.1	7.1	8.3	8.1	8.2	9.1	8.6	7.4	6.7	5.8	5.3	7.2
	Min/Max	3.9/6.9	4.4/8.1	5.6/8.7	7.1/9.7	6.8/9.6	6.7/9.9	7.7/11.0	7.9/9.5	5.7/8.6	5.6/7.5	4.4/7.2	3.5/6.4	6.7/7.7

Direct Beam Solar Radiation for Concentrating Collectors (kWh/m²/day), Uncertainty ±8%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
1-Axis, E-W Horiz Axis	Average	3.2	3.4	3.6	4.1	3.9	3.9	4.8	4.5	3.8	3.6	3.5	3.2	3.8
	Min/Max	2.0/4.5	1.9/5.0	2.4/4.9	3.3/5.4	3.0/5.1	2.4/5.3	3.7/6.1	4.0/5.3	2.2/4.8	2.8/4.4	2.3/4.7	1.8/4.1	3.4/4.2
1-Axis, N-S Horiz Axis	Average	2.6	3.2	4.1	5.3	5.0	5.0	6.1	5.7	4.6	3.7	2.9	2.4	4.2
	Min/Max	1.6/3.6	1.8/4.8	2.8/5.8	4.2/7.0	3.7/6.6	3.2/6.9	4.6/8.0	4.9/6.8	2.7/5.8	2.9/4.5	1.9/4.0	1.3/3.1	3.8/4.7
1-Axis, N-S Tilt=Latitude	Average	3.6	4.1	4.7	5.5	4.8	4.6	5.8	5.7	5.0	4.5	3.9	3.5	4.6
	Min/Max	2.2/4.9	2.3/6.1	3.2/6.6	4.3/7.2	3.5/6.3	3.0/6.5	4.3/7.6	4.9/6.8	2.9/6.3	3.5/5.4	2.6/5.3	1.8/4.4	4.1/5.1
2-Axis	Average	3.8	4.2	4.7	5.5	5.1	5.0	6.2	5.9	5.0	4.6	4.1	3.8	4.8
	Min/Max	2.3/5.3	2.3/6.2	3.2/6.6	4.4/7.3	3.7/6.7	3.2/7.0	4.7/8.1	5.1/7.0	2.9/6.3	3.5/5.5	2.8/5.6	2.0/4.7	4.3/5.3

Average Climatic Conditions

Element	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
Temperature (°C)	13.8	14.2	14.4	15.6	17.1	18.7	20.6	21.4	21.1	19.3	16.4	13.8	17.2
Daily Minimum Temp	8.8	9.6	10.3	11.6	13.5	15.3	17.1	17.9	17.3	15.1	11.6	8.8	13.1
Daily Maximum Temp	18.7	18.8	18.6	19.7	20.6	22.2	24.1	24.8	24.8	23.6	21.3	18.8	21.3
Record Minimum Temp	-5.0	0.0	1.1	3.9	6.1	8.9	9.4	10.6	8.3	5.0	1.1	0.0	-5.0
Record Maximum Temp	31.1	33.3	35.0	38.9	36.1	40.0	36.1	36.7	43.3	41.1	38.3	34.4	43.3
HDD, Base 18.3°C	143	119	124	88	53	30	5	3	12	18	71	143	810
CDD, Base 18.3°C	0	4	4	6	14	42	76	98	94	49	14	4	404
Relative Humidity (%)	63	68	71	71	74	76	77	77	74	71	65	63	71
Wind Speed (m/s)	3.1	3.4	3.8	4.0	3.9	3.8	3.7	3.7	3.5	3.3	3.2	3.0	3.5

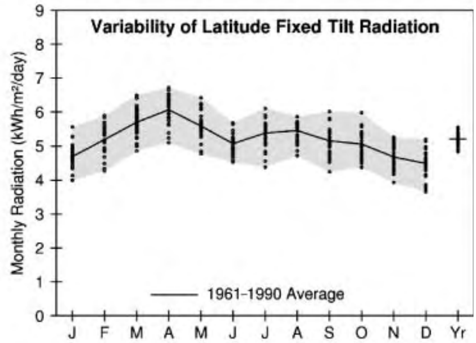
FIGURE 19.16 Solar radiation data manual table for Los Angeles, CA.

Miami, FL

WBAN NO. 12839

LATITUDE: 25.80° N
LONGITUDE: 80.27° W
ELEVATION: 2 m
MEAN PRESSURE: 1017 millibars

STATION TYPE: Primary



Solar Radiation for Flat-Plate Collectors Facing South at a Fixed Tilt (kWh/m²/day), Uncertainty ±9%

Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	3.5	4.2	5.2	6.0	6.0	5.6	5.8	5.6	4.9	4.4	3.7	3.3	4.8
	Min/Max	3.1/4.0	3.6/4.7	4.5/5.8	5.1/6.7	5.1/6.9	4.9/6.3	4.7/6.7	4.9/6.0	4.1/5.7	3.9/4.9	3.2/4.0	2.8/3.7	4.5/5.1
Latitude -15	Average	4.1	4.7	5.5	6.2	5.9	5.5	5.7	5.6	5.1	4.7	4.2	3.9	5.1
	Min/Max	3.5/4.8	4.0/5.3	4.7/6.2	5.2/6.8	5.0/6.8	4.8/6.1	4.7/6.5	4.9/6.1	4.2/5.9	4.2/5.5	3.6/4.6	3.2/4.4	4.8/5.4
Latitude	Average	4.7	5.2	5.7	6.1	5.6	5.1	5.4	5.5	5.1	5.1	4.7	4.5	5.2
	Min/Max	4.0/5.6	4.3/5.9	4.9/6.5	5.1/6.7	4.8/6.4	4.5/5.7	4.4/6.1	4.7/5.9	4.2/6.0	4.4/6.0	3.9/5.3	3.7/5.2	4.8/5.5
Latitude +15	Average	5.0	5.4	5.6	5.7	5.0	4.5	4.8	5.0	4.9	5.1	4.9	4.9	5.1
	Min/Max	4.2/6.0	4.4/6.2	4.8/6.4	4.8/6.3	4.3/5.7	4.0/5.0	3.9/5.4	4.3/5.4	4.1/5.8	4.4/6.1	4.1/5.6	3.9/5.7	4.7/5.4
90	Average	4.1	3.9	3.4	2.6	1.9	1.6	1.7	2.1	2.7	3.5	3.9	4.1	3.0
	Min/Max	3.4/5.1	3.1/4.6	2.9/3.9	2.3/2.8	1.6/2.0	1.5/1.8	1.6/2.0	2.0/2.2	2.3/3.1	3.0/4.2	3.2/4.5	3.2/4.9	2.7/3.1

Solar Radiation for 1-Axis Tracking Flat-Plate Collectors with a North–South Axis (kWh/m²/day), Uncertainty ±9%

Axis Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	4.8	5.7	6.8	7.8	7.4	6.7	7.0	6.8	6.0	5.6	4.8	4.5	6.2
	Min/Max	4.0/5.8	4.5/6.7	5.6/7.8	6.5/9.1	6.1/8.9	5.6/7.7	5.3/8.2	5.8/7.5	4.7/7.1	4.7/6.8	4.0/5.5	3.6/5.3	5.6/6.7
Latitude -15	Average	5.2	6.1	7.0	7.9	7.4	6.6	7.0	6.9	6.2	5.9	5.2	4.9	6.4
	Min/Max	4.3/6.3	4.8/7.1	5.8/8.2	6.6/9.2	6.1/8.9	5.5/7.6	5.3/8.2	5.8/7.6	4.8/7.3	4.9/7.2	4.3/6.0	3.9/5.8	5.8/6.9
Latitude	Average	5.7	6.4	7.2	7.8	7.2	6.3	6.7	6.7	6.2	6.1	5.6	5.4	6.5
	Min/Max	4.6/6.9	5.0/7.6	5.9/8.4	6.5/9.2	5.9/8.6	5.3/7.3	5.1/7.9	5.7/7.4	4.8/7.4	5.1/7.6	4.6/6.5	4.2/6.4	5.8/7.0
Latitude +15	Average	5.9	6.5	7.1	7.5	6.7	5.9	6.3	6.4	6.1	6.2	5.8	5.7	6.3
	Min/Max	4.8/7.3	5.0/7.8	5.9/8.3	6.3/8.8	5.5/8.1	4.9/6.8	4.8/7.3	5.4/7.1	4.7/7.3	5.1/7.7	4.7/6.7	4.4/6.8	5.7/6.9

Solar Radiation for 2-Axis Tracking Flat-Plate Collectors (kWh/m²/day), Uncertainty ±9%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
2-Axis	Average	6.0	6.6	7.2	7.9	7.4	6.7	7.1	6.9	6.2	6.2	5.9	5.8	6.7
	Min/Max	4.8/7.4	5.1/7.8	6.0/8.4	6.6/9.3	6.1/8.9	5.6/7.8	5.4/8.3	5.8/7.6	4.8/7.4	5.1/7.7	4.7/6.8	4.5/6.9	6.0/7.2

Direct Beam Solar Radiation for Concentrating Collectors (kWh/m²/day), Uncertainty ±8%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
1-Axis, E-W	Average	3.3	3.4	3.4	3.7	3.2	2.7	2.9	2.8	2.5	2.8	3.1	3.2	3.1
Horiz Axis	Min/Max	2.3/4.6	2.1/4.3	2.5/4.4	2.6/4.7	2.2/4.5	1.8/3.4	1.2/3.8	1.8/3.3	1.5/3.5	1.9/4.0	2.1/3.8	2.0/4.4	2.4/3.6
1-Axis, N-S	Average	3.1	3.7	4.3	4.9	4.2	3.4	3.7	3.6	3.2	3.2	2.9	2.8	3.6
Horiz Axis	Min/Max	2.0/4.3	2.2/4.8	3.1/5.6	3.5/6.5	2.8/5.9	2.2/4.3	1.6/4.9	2.4/4.3	1.9/4.4	2.1/4.6	2.0/3.7	1.7/3.8	2.8/4.2
1-Axis, N-S	Average	3.8	4.3	4.6	4.9	4.0	3.2	3.5	3.5	3.3	3.6	3.5	3.6	3.8
Tilt=Latitude	Min/Max	2.5/5.3	2.6/5.5	3.3/6.0	3.5/6.5	2.7/5.7	2.1/4.0	1.5/4.6	2.3/4.3	1.9/4.6	2.3/5.2	2.4/4.4	2.1/4.8	3.0/4.4
2-Axis	Average	4.1	4.4	4.6	5.0	4.3	3.4	3.7	3.6	3.3	3.7	3.7	3.9	4.0
	Min/Max	2.7/5.6	2.6/5.7	3.3/6.0	3.5/6.6	2.8/6.0	2.2/4.3	1.6/4.9	2.4/4.4	2.0/4.6	2.4/5.2	2.5/4.7	2.3/5.2	3.1/4.6

Average Climatic Conditions

Element	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
Temperature (°C)	19.6	20.3	22.1	24.0	25.9	27.4	28.1	28.2	27.7	25.7	23.1	20.6	24.4
Daily Minimum Temp	15.1	15.8	17.9	19.9	22.3	23.9	24.6	24.8	24.4	22.3	19.3	16.4	20.6
Daily Maximum Temp	24.0	24.7	26.2	28.0	29.6	30.9	31.7	31.7	31.0	29.2	26.9	24.8	28.2
Record Minimum Temp	-1.1	0.0	0.0	7.8	11.7	15.6	20.6	20.0	20.0	10.6	3.9	-1.1	-1.1
Record Maximum Temp	31.1	31.7	33.3	35.6	35.0	36.7	36.7	36.7	36.1	35.0	31.7	30.6	36.7
HDD, Base 18.3°C	49	28	8	0	0	0	0	0	0	0	3	23	111
CDD, Base 18.3°C	87	83	123	170	236	273	303	307	282	229	147	93	2332
Relative Humidity (%)	73	71	69	67	72	76	75	76	78	75	74	72	73
Wind Speed (m/s)	4.3	4.7	4.9	4.8	4.4	3.8	3.7	3.7	3.8	4.4	4.5	4.3	4.2

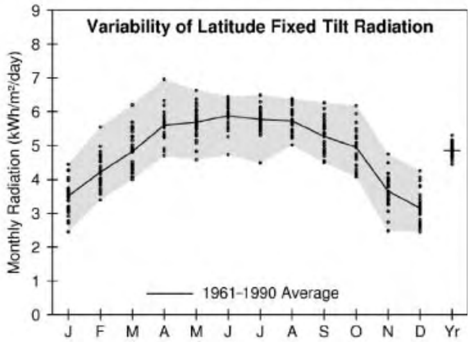
FIGURE 19.17 Solar radiation data manual table for Miami, FL.

Nashville, TN

WBAN NO. 13897

LATITUDE: 36.12° N
LONGITUDE: 86.68° W
ELEVATION: 180 m
MEAN PRESSURE: 996 millibars

STATION TYPE: Primary



Solar Radiation for Flat-Plate Collectors Facing South at a Fixed Tilt (kWh/m²/day), Uncertainty ±9%

Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	2.3	3.1	4.1	5.4	6.0	6.5	6.3	5.7	4.7	3.8	2.5	2.0	4.4
	Min/Max	1.8/2.7	2.7/3.8	3.5/5.0	4.6/6.5	4.9/7.1	5.2/7.2	4.9/7.1	5.0/6.3	4.1/5.4	3.3/4.5	2.0/3.0	1.6/2.4	4.0/4.7
Latitude -15	Average	3.1	3.9	4.7	5.7	6.0	6.4	6.2	5.9	5.2	4.6	3.3	2.8	4.8
	Min/Max	2.3/3.9	3.2/5.0	3.9/5.9	4.8/7.0	4.9/7.1	5.1/7.0	4.8/7.0	5.2/6.6	4.5/6.2	3.9/5.7	2.3/4.2	2.2/3.6	4.4/5.2
Latitude	Average	3.5	4.2	4.8	5.6	5.7	5.9	5.8	5.7	5.3	4.9	3.6	3.1	4.9
	Min/Max	2.5/4.4	3.4/5.6	4.0/6.2	4.7/7.0	4.6/6.6	4.7/6.4	4.5/6.5	5.0/6.4	4.5/6.3	4.1/6.2	2.5/4.8	2.5/4.3	4.5/5.3
Latitude +15	Average	3.7	4.3	4.7	5.2	5.0	5.1	5.1	5.2	5.0	5.0	3.8	3.4	4.6
	Min/Max	2.5/4.8	3.4/5.8	3.9/6.1	4.3/6.5	4.1/5.8	4.2/5.6	4.0/5.7	4.6/5.8	4.3/6.0	4.1/6.3	2.5/5.0	2.6/4.6	4.3/5.1
90	Average	3.3	3.5	3.3	3.1	2.5	2.3	2.4	2.9	3.3	3.9	3.3	3.0	3.1
	Min/Max	2.1/4.3	2.8/4.8	2.7/4.3	2.6/3.8	2.1/2.8	2.1/2.5	2.1/2.6	2.5/3.1	2.8/4.0	3.1/4.9	2.0/4.4	2.2/4.3	2.9/3.4

Solar Radiation for 1-Axis Tracking Flat-Plate Collectors with a North–South Axis (kWh/m²/day), Uncertainty ±9%

Axis Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	3.1	4.2	5.4	7.0	7.6	8.2	7.8	7.3	6.1	5.1	3.4	2.7	5.7
	Min/Max	2.2/3.9	3.3/5.5	4.4/7.1	5.7/8.9	6.0/9.5	5.9/9.3	5.7/9.4	6.3/8.5	5.0/7.5	4.1/6.5	2.3/4.5	2.2/3.6	5.1/6.3
Latitude -15	Average	3.8	4.8	5.8	7.3	7.7	8.1	7.8	7.5	6.5	5.8	4.0	3.3	6.0
	Min/Max	2.5/4.8	3.7/6.4	4.7/7.8	5.9/9.4	6.0/9.6	5.8/9.3	5.7/9.4	6.4/8.8	5.3/8.0	4.6/7.4	2.6/5.3	2.5/4.5	5.4/6.7
Latitude	Average	4.1	5.0	6.0	7.2	7.5	7.8	7.5	7.3	6.5	6.0	4.2	3.6	6.1
	Min/Max	2.7/5.2	3.9/6.9	4.8/8.0	5.8/9.3	5.8/9.3	5.6/8.9	5.5/9.1	6.3/8.6	5.3/8.1	4.8/7.8	2.7/5.8	2.7/5.1	5.5/6.8
Latitude +15	Average	4.2	5.1	5.9	6.9	7.0	7.3	7.0	6.9	6.3	6.1	4.4	3.8	5.9
	Min/Max	2.7/5.5	3.9/7.0	4.7/7.9	5.5/9.0	5.4/8.8	5.2/8.3	5.1/8.5	6.0/8.2	5.2/7.9	4.7/7.9	2.7/6.0	2.8/5.4	5.3/6.6

Solar Radiation for 2-Axis Tracking Flat-Plate Collectors (kWh/m²/day), Uncertainty ±9%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
2-Axis	Average	4.3	5.1	6.0	7.3	7.8	8.3	7.9	7.5	6.5	6.1	4.4	3.8	6.3
	Min/Max	2.8/5.6	3.9/7.1	4.8/8.0	5.9/9.4	6.1/9.7	5.9/9.5	5.8/9.6	6.5/8.8	5.4/8.1	4.8/7.9	2.7/6.0	2.8/5.4	5.7/7.0

Direct Beam Solar Radiation for Concentrating Collectors (kWh/m²/day), Uncertainty ±8%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
1-Axis, E-W	Average	2.4	2.7	2.8	3.4	3.5	3.7	3.4	3.2	2.9	3.2	2.5	2.2	3.0
Horiz Axis	Min/Max	1.3/3.5	1.5/4.4	1.9/4.4	2.3/5.0	2.3/5.2	1.9/4.8	1.8/4.9	2.5/4.3	2.2/4.2	2.2/4.6	1.1/3.8	1.4/3.7	2.6/3.6
1-Axis, N-S	Average	1.8	2.5	3.3	4.4	4.6	4.8	4.4	4.2	3.6	3.3	2.0	1.6	3.4
Horiz Axis	Min/Max	1.0/2.7	1.4/4.1	2.2/5.2	2.9/6.5	3.0/6.8	2.3/6.3	2.4/6.5	3.1/5.7	2.5/5.1	2.2/4.8	0.9/3.3	1.0/2.8	2.8/4.1
1-Axis, N-S	Average	2.6	3.2	3.8	4.6	4.4	4.5	4.2	4.2	3.9	4.1	2.8	2.4	3.7
Tilt=Latitude	Min/Max	1.4/3.8	1.8/5.3	2.5/5.9	3.0/6.8	2.9/6.6	2.2/5.8	2.3/6.1	3.2/5.8	2.8/5.6	2.7/5.8	1.2/4.4	1.5/4.0	3.2/4.6
2-Axis	Average	2.8	3.3	3.8	4.7	4.7	4.9	4.5	4.4	3.9	4.1	2.9	2.6	3.9
	Min/Max	1.5/4.1	1.9/5.5	2.5/5.9	3.1/6.9	3.1/7.0	2.4/6.3	2.4/6.6	3.3/5.9	2.8/5.6	2.7/5.9	1.3/4.6	1.6/4.3	3.3/4.8

Average Climatic Conditions

Element	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
Temperature (°C)	2.3	4.7	10.1	15.1	19.8	24.2	26.3	25.6	22.1	15.8	10.0	4.7	15.1
Daily Minimum Temp	-3.1	-1.2	3.9	8.6	13.7	18.2	20.5	19.8	16.2	9.1	4.2	-0.6	9.1
Daily Maximum Temp	7.7	10.4	16.2	21.6	26.0	30.3	31.9	31.3	28.1	22.5	15.8	10.1	21.0
Record Minimum Temp	-27.2	-25.0	-16.7	-5.0	1.1	5.6	10.6	8.3	2.2	-3.3	-18.3	-23.3	-27.2
Record Maximum Temp	25.6	28.9	30.0	32.8	36.1	41.1	41.7	40.0	40.6	34.4	28.9	26.1	41.7
HDD, Base 18.3°C	496	383	261	107	33	0	0	0	12	108	250	422	2072
CDD, Base 18.3°C	0	0	6	11	79	177	246	226	125	29	0	0	898
Relative Humidity (%)	70	69	65	63	69	70	73	73	74	69	70	72	70
Wind Speed (m/s)	4.1	4.2	4.5	4.2	3.5	3.3	3.0	2.9	3.0	3.2	3.8	4.2	3.7

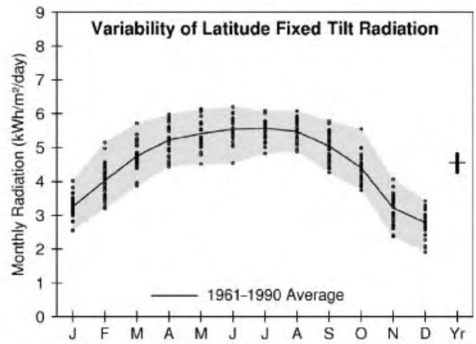
FIGURE 19.18 Solar radiation data manual table for Nashville, TN.

New York City, NY

WBAN NO. 94728

LATITUDE: 40.78° N
LONGITUDE: 73.97° W
ELEVATION: 57 m
MEAN PRESSURE: 1016 millibars

STATION TYPE: Primary



Solar Radiation for Flat-Plate Collectors Facing South at a Fixed Tilt (kWh/m²/day), Uncertainty ±9%

Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	1.9	2.7	3.8	4.9	5.7	6.1	6.0	5.4	4.3	3.2	2.0	1.6	4.0
	Min/Max	1.7/2.2	2.3/3.3	3.3/4.4	4.3/5.5	4.8/6.5	5.0/6.8	5.3/6.6	4.8/5.9	3.9/4.9	2.9/3.8	1.7/2.4	1.4/1.8	3.7/4.2
Latitude -15	Average	2.9	3.7	4.6	5.3	5.8	6.0	6.0	5.7	5.0	4.1	2.9	2.4	4.5
	Min/Max	2.3/3.5	3.0/4.6	3.8/5.5	4.6/6.1	4.8/6.5	4.9/6.7	5.2/6.6	5.1/6.3	4.3/5.7	3.6/5.1	2.2/3.6	1.8/3.0	4.3/4.8
Latitude	Average	3.2	4.0	4.8	5.2	5.4	5.5	5.6	5.5	5.0	4.4	3.2	2.8	4.6
	Min/Max	2.5/4.0	3.2/5.2	3.9/5.7	4.4/6.0	4.5/6.1	4.5/6.2	4.8/6.1	4.9/6.1	4.3/5.8	3.7/5.6	2.4/4.1	1.9/3.4	4.3/4.8
Latitude +15	Average	3.4	4.1	4.6	4.8	4.8	4.8	4.9	5.0	4.8	4.4	3.3	3.0	4.3
	Min/Max	2.7/4.3	3.3/5.4	3.7/5.6	4.1/5.6	4.0/5.4	4.0/5.3	4.2/5.3	4.4/5.5	4.0/5.5	3.7/5.6	2.4/4.3	2.0/3.7	4.0/4.6
90	Average	3.2	3.6	3.5	3.1	2.7	2.6	2.7	3.0	3.4	3.6	3.0	2.7	3.1
	Min/Max	2.4/4.2	2.8/4.9	2.8/4.3	2.7/3.7	2.4/3.1	2.3/2.8	2.4/2.9	2.8/3.3	2.8/3.9	3.0/4.6	2.1/3.8	1.7/3.5	2.9/3.4

Solar Radiation for 1-Axis Tracking Flat-Plate Collectors with a North–South Axis (kWh/m²/day), Uncertainty ±9%

Axis Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	2.7	3.7	5.1	6.3	7.1	7.5	7.5	6.8	5.7	4.3	2.7	2.2	5.1
	Min/Max	2.1/3.3	3.0/4.9	4.1/6.2	5.2/7.4	5.5/8.4	5.8/8.8	6.2/8.4	5.8/7.7	4.6/6.6	3.7/5.6	2.0/3.4	1.6/2.7	4.7/5.4
Latitude -15	Average	3.4	4.5	5.7	6.7	7.2	7.5	7.5	7.1	6.2	5.1	3.4	2.8	5.6
	Min/Max	2.6/4.3	3.5/5.9	4.5/7.1	5.4/7.9	5.5/8.6	5.8/8.8	6.2/8.5	6.0/8.0	5.0/7.2	4.2/6.6	2.4/4.4	1.9/3.6	5.1/5.9
Latitude	Average	3.7	4.7	5.8	6.6	6.9	7.2	7.2	6.9	6.2	5.3	3.7	3.1	5.6
	Min/Max	2.8/4.7	3.7/6.3	4.5/7.3	5.3/7.9	5.3/8.3	5.5/8.4	5.9/8.2	5.9/7.9	5.0/7.3	4.3/6.9	2.5/4.7	2.0/3.9	5.2/6.0
Latitude +15	Average	3.9	4.8	5.7	6.3	6.5	6.7	6.7	6.6	6.0	5.3	3.8	3.3	5.5
	Min/Max	2.9/5.0	3.8/6.4	4.4/7.2	5.1/7.6	5.0/7.8	5.1/7.8	5.5/7.6	5.6/7.5	4.8/7.1	4.3/7.0	2.6/4.9	2.1/4.2	5.0/5.9

Solar Radiation for 2-Axis Tracking Flat-Plate Collectors (kWh/m²/day), Uncertainty ±9%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
2-Axis	Average	3.9	4.8	5.8	6.7	7.3	7.7	7.6	7.1	6.2	5.3	3.8	3.3	5.8
	Min/Max	2.9/5.0	3.8/6.4	4.5/7.3	5.4/8.0	5.6/8.7	5.9/9.0	6.3/8.6	6.1/8.1	5.0/7.3	4.4/7.0	2.6/4.9	2.1/4.2	5.4/6.2

Direct Beam Solar Radiation for Concentrating Collectors (kWh/m²/day), Uncertainty ±8%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
1-Axis, E-W	Average	2.2	2.5	2.7	2.9	3.0	3.2	3.2	3.0	2.8	2.7	2.1	1.9	2.7
Horiz Axis	Min/Max	1.4/2.9	1.7/3.5	1.7/3.9	2.0/3.8	1.8/4.2	1.9/4.4	2.2/4.0	2.2/3.9	1.9/3.5	2.0/4.1	1.1/3.0	0.9/2.7	2.3/3.0
1-Axis, N-S	Average	1.5	2.2	3.0	3.7	3.9	4.1	4.1	3.8	3.3	2.6	1.5	1.2	2.9
Horiz Axis	Min/Max	0.9/2.1	1.4/3.1	1.8/4.3	2.4/4.9	2.2/5.4	2.4/5.6	2.7/5.2	2.8/4.9	2.2/4.2	1.9/4.0	0.8/2.3	0.6/1.8	2.4/3.3
1-Axis, N-S	Average	2.3	2.9	3.6	3.9	3.9	3.9	3.9	3.9	3.7	3.4	2.3	2.0	3.3
Tilt=Latitude	Min/Max	1.5/3.2	2.0/4.2	2.2/5.2	2.6/5.2	2.2/5.3	2.3/5.3	2.6/5.0	2.8/5.0	2.5/4.8	2.4/5.1	1.2/3.3	0.9/2.8	2.8/3.7
2-Axis	Average	2.5	3.0	3.6	4.0	4.1	4.2	4.2	4.1	3.8	3.4	2.4	2.1	3.5
	Min/Max	1.6/3.4	2.1/4.3	2.2/5.2	2.6/5.3	2.3/5.7	2.5/5.8	2.8/5.3	2.9/5.2	2.5/4.8	2.5/5.1	1.3/3.5	1.0/3.1	2.9/3.8

Average Climatic Conditions

Element	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
Temperature (°C)	-0.3	0.9	5.8	11.4	17.1	22.0	24.9	24.2	20.1	14.2	8.7	2.6	12.6
Daily Minimum Temp	-3.7	-2.8	1.6	6.6	12.1	17.2	20.2	19.6	15.6	9.8	5.1	-0.7	8.4
Daily Maximum Temp	3.1	4.6	10.0	16.2	22.1	26.7	29.6	28.7	24.6	18.5	12.2	5.8	16.8
Record Minimum Temp	-21.1	-26.1	-16.1	-11.1	0.0	6.7	11.1	10.0	3.9	-2.2	-15.0	-25.0	-26.1
Record Maximum Temp	22.2	23.9	30.0	35.6	37.2	38.3	41.1	40.0	38.9	34.4	28.9	22.2	41.1
HDD, Base 18.3°C	577	488	389	208	69	0	0	0	19	139	290	489	2669
CDD, Base 18.3°C	0	0	0	0	30	113	203	181	72	9	0	0	609
Relative Humidity (%)	61	60	60	59	64	65	65	67	67	65	64	63	63
Wind Speed (m/s)	6.1	6.1	6.1	5.8	5.2	4.9	4.7	4.7	5.0	5.1	5.8	6.0	5.5

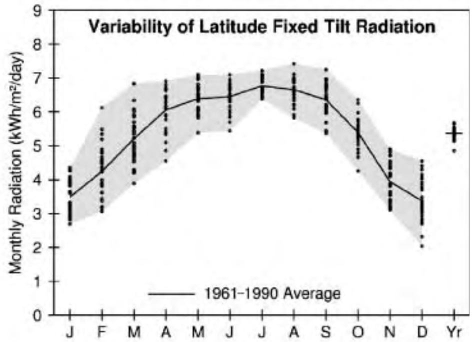
FIGURE 19.19 Solar radiation data manual table for New York City, NY.

San Francisco, CA

WBAN NO. 23234

LATITUDE: 37.62° N
LONGITUDE: 122.38° W
ELEVATION: 5 m
MEAN PRESSURE: 1017 millibars

STATION TYPE: Secondary



Solar Radiation for Flat-Plate Collectors Facing South at a Fixed Tilt (kWh/m²/day), Uncertainty ±9%

Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	2.2	3.0	4.2	5.7	6.7	7.2	7.3	6.5	5.4	3.9	2.5	2.0	4.7
	Min/Max	1.8/2.5	2.3/3.9	3.3/5.3	4.4/6.4	5.7/7.4	6.1/7.9	6.9/7.9	5.7/7.3	4.7/6.0	3.2/4.4	2.1/2.9	1.4/2.4	4.4/4.9
Latitude -15	Average	3.1	3.9	5.0	6.2	6.8	7.0	7.3	6.9	6.2	5.0	3.5	2.9	5.3
	Min/Max	2.4/3.8	2.9/5.5	3.8/6.5	4.7/7.0	5.7/7.5	5.9/7.7	6.9/7.8	6.0/7.7	5.3/7.1	4.0/5.8	2.8/4.3	1.9/3.9	4.9/5.6
Latitude	Average	3.5	4.2	5.2	6.1	6.4	6.5	6.8	6.7	6.4	5.4	3.9	3.4	5.4
	Min/Max	2.7/4.3	3.1/6.1	3.9/6.8	4.6/6.9	5.4/7.1	5.4/7.1	6.4/7.2	5.8/7.4	5.4/7.3	4.3/6.4	3.1/4.9	2.0/4.6	4.9/5.7
Latitude +15	Average	3.7	4.4	5.1	5.6	5.7	5.6	5.9	6.1	6.1	5.5	4.1	3.6	5.1
	Min/Max	2.8/4.7	3.1/6.4	3.8/6.8	4.2/6.4	4.8/6.3	4.7/6.1	5.6/6.3	5.3/6.7	5.2/7.0	4.3/6.5	3.2/5.2	2.1/5.0	4.6/5.5
90	Average	3.3	3.6	3.7	3.4	2.8	2.5	2.7	3.3	4.1	4.3	3.6	3.3	3.4
	Min/Max	2.5/4.2	2.4/5.4	2.7/4.9	2.6/3.8	2.5/3.0	2.3/2.6	2.6/2.8	3.0/3.6	3.5/4.7	3.4/5.2	2.8/4.6	1.9/4.7	2.9/3.7

Solar Radiation for 1-Axis Tracking Flat-Plate Collectors with a North–South Axis (kWh/m²/day), Uncertainty ±9%

Axis Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	3.0	4.1	5.7	7.6	8.8	9.3	9.7	8.7	7.4	5.5	3.5	2.8	6.3
	Min/Max	2.3/3.7	2.8/6.0	4.1/7.7	5.6/8.9	7.0/10.1	7.5/10.8	9.0/10.7	7.4/10.2	6.2/8.8	4.3/6.5	2.8/4.4	1.7/3.8	5.7/6.7
Latitude -15	Average	3.7	4.7	6.3	8.0	8.9	9.2	9.7	9.0	8.1	6.3	4.3	3.5	6.8
	Min/Max	2.8/4.6	3.3/7.1	4.5/8.6	5.8/9.4	7.1/10.3	7.4/10.7	9.0/10.8	7.7/10.6	6.7/9.6	4.8/7.6	3.3/5.4	2.0/4.8	6.1/7.2
Latitude	Average	4.0	5.0	6.5	8.0	8.7	8.9	9.4	8.8	8.2	6.6	4.6	3.9	6.9
	Min/Max	3.0/5.1	3.4/7.6	4.6/8.9	5.7/9.4	6.8/10.0	7.1/10.3	8.7/10.4	7.5/10.4	6.8/9.7	5.1/8.0	3.5/5.9	2.2/5.4	6.1/7.3
Latitude +15	Average	4.2	5.1	6.4	7.7	8.1	8.2	8.7	8.4	8.0	6.7	4.8	4.1	6.7
	Min/Max	3.1/5.4	3.4/7.8	4.5/8.8	5.5/9.0	6.4/9.4	6.6/9.6	8.1/9.7	7.1/9.9	6.6/9.5	5.1/8.1	3.6/6.1	2.3/5.7	5.9/7.1

Solar Radiation for 2-Axis Tracking Flat-Plate Collectors (kWh/m²/day), Uncertainty ±9%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
2-Axis	Average	4.2	5.1	6.5	8.1	9.0	9.4	9.9	9.0	8.2	6.7	4.8	4.1	7.1
	Min/Max	3.1/5.4	3.4/7.8	4.6/8.9	5.8/9.5	7.2/10.4	7.6/11.0	9.2/11.0	7.7/10.6	6.8/9.7	5.1/8.1	3.7/6.2	2.3/5.8	6.3/7.5

Direct Beam Solar Radiation for Concentrating Collectors (kWh/m²/day), Uncertainty ±8%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
1-Axis, E-W	Average	2.5	2.8	3.3	4.0	4.7	5.1	5.5	4.9	4.6	3.9	2.8	2.5	3.9
Horiz Axis	Min/Max	1.6/3.5	1.5/4.9	1.9/5.2	2.4/5.2	3.3/5.7	3.6/6.3	4.9/6.4	3.7/6.1	3.3/5.6	2.6/5.0	1.8/3.9	1.1/3.9	3.2/4.2
1-Axis, N-S	Average	1.9	2.5	3.8	5.2	6.1	6.5	7.1	6.3	5.4	3.9	2.3	1.8	4.4
Horiz Axis	Min/Max	1.2/2.6	1.3/4.6	2.1/6.0	3.1/6.7	4.2/7.5	4.6/8.3	6.2/8.3	4.7/8.0	4.0/6.8	2.5/5.0	1.5/3.2	0.8/2.8	3.7/4.7
1-Axis, N-S	Average	2.7	3.3	4.4	5.4	5.9	6.1	6.8	6.4	6.0	4.8	3.2	2.7	4.8
Tilt=Latitude	Min/Max	1.8/3.8	1.7/6.0	2.5/6.9	3.2/7.0	4.1/7.3	4.3/7.7	5.9/7.9	4.8/8.1	4.4/7.6	3.2/6.3	2.0/4.4	1.2/4.2	4.0/5.3
2-Axis	Average	2.9	3.4	4.4	5.5	6.3	6.7	7.3	6.6	6.1	4.9	3.3	2.9	5.0
	Min/Max	1.9/4.0	1.7/6.1	2.5/6.9	3.3/7.1	4.3/7.7	4.7/8.4	6.4/8.5	5.0/8.3	4.4/7.6	3.2/6.3	2.2/4.7	1.3/4.5	4.2/5.5

Average Climatic Conditions

Element	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
Temperature (°C)	9.3	11.2	11.8	13.1	14.5	16.4	17.1	17.6	18.1	16.1	12.7	9.7	13.9
Daily Minimum Temp	5.4	7.2	7.7	8.4	9.8	11.4	12.2	12.8	12.9	11.0	8.4	5.9	9.4
Daily Maximum Temp	13.1	15.2	16.0	17.7	19.2	21.3	22.0	22.4	23.1	21.2	16.9	13.4	18.4
Record Minimum Temp	-4.4	-3.9	-1.1	-0.6	2.2	5.0	6.1	5.6	3.3	1.1	-3.9	-6.7	-6.7
Record Maximum Temp	22.2	25.6	29.4	33.3	36.1	41.1	40.6	36.7	39.4	37.2	29.4	23.9	41.1
HDD, Base 18.3°C	281	199	202	159	121	67	51	38	44	75	170	269	1676
CDD, Base 18.3°C	0	0	0	3	0	9	12	16	36	6	0	0	81
Relative Humidity (%)	78	76	73	71	71	72	73	74	72	72	75	77	74
Wind Speed (m/s)	3.3	4.0	4.8	5.5	6.2	6.2	6.2	5.7	5.0	4.3	3.7	3.5	4.9

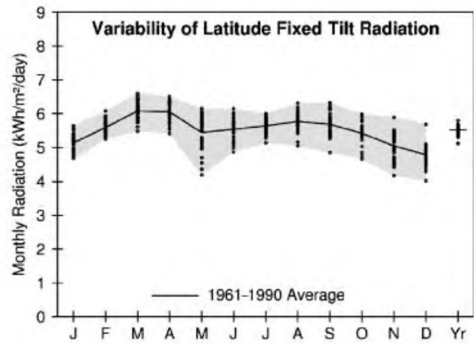
FIGURE 19.20 Solar radiation data manual table for San Francisco, CA.

San Juan, PR

WBAN NO. 11641

LATITUDE: 18.43° N
LONGITUDE: 66.00° W
ELEVATION: 19 m
MEAN PRESSURE: 1014 millibars

STATION TYPE: Primary



Solar Radiation for Flat-Plate Collectors Facing South at a Fixed Tilt (kWh/m²/day), Uncertainty ±9%

Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	4.3	4.9	5.7	6.1	5.8	6.1	6.1	6.0	5.5	4.9	4.3	4.0	5.3
	Min/Max	4.0/4.6	4.7/5.3	5.2/6.2	5.4/6.5	4.4/6.6	5.3/6.8	5.5/6.5	5.2/6.5	4.8/6.1	4.3/5.4	3.7/4.9	3.4/4.5	4.9/5.5
Latitude -15	Average	4.5	5.1	5.8	6.1	5.7	6.0	6.0	6.0	5.6	5.0	4.5	4.1	5.4
	Min/Max	4.1/4.8	4.8/5.5	5.3/6.3	5.5/6.6	4.4/6.5	5.2/6.7	5.5/6.4	5.2/6.5	4.8/6.2	4.4/5.5	3.8/5.1	3.5/4.8	5.0/5.6
Latitude	Average	5.1	5.6	6.1	6.1	5.4	5.5	5.6	5.8	5.7	5.4	5.1	4.8	5.5
	Min/Max	4.7/5.6	5.3/6.1	5.5/6.6	5.4/6.5	4.2/6.2	4.9/6.2	5.1/6.0	5.1/6.3	4.8/6.3	4.7/6.0	4.2/5.9	4.0/5.7	5.1/5.8
Latitude +15	Average	5.5	5.8	6.0	5.7	4.9	4.8	5.0	5.3	5.5	5.5	5.4	5.2	5.4
	Min/Max	5.0/6.1	5.4/6.3	5.4/6.5	5.1/6.1	3.8/5.5	4.3/5.3	4.5/5.3	4.6/5.8	4.7/6.1	4.7/6.1	4.3/6.3	4.3/6.3	5.0/5.7
90	Average	4.2	3.9	3.1	2.0	1.5	1.5	1.5	1.7	2.5	3.3	3.9	4.1	2.8
	Min/Max	3.8/4.8	3.6/4.2	2.8/3.3	1.9/2.1	1.4/1.6	1.4/1.5	1.4/1.5	1.6/1.7	2.2/2.7	2.8/3.7	3.1/4.7	3.3/5.1	2.7/3.0

Solar Radiation for 1-Axis Tracking Flat-Plate Collectors with a North–South Axis (kWh/m²/day), Uncertainty ±9%

Axis Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	5.6	6.4	7.4	7.8	7.2	7.6	7.6	7.5	7.0	6.3	5.5	5.1	6.8
	Min/Max	5.1/6.3	5.8/7.0	6.4/8.2	6.5/8.7	5.0/8.5	6.3/8.9	6.8/8.5	6.4/8.6	5.8/8.2	5.2/7.2	4.4/6.6	4.2/6.4	6.1/7.2
Latitude -15	Average	5.8	6.5	7.5	7.8	7.2	7.5	7.6	7.5	7.0	6.4	5.7	5.2	6.8
	Min/Max	5.2/6.5	5.9/7.2	6.5/8.3	6.5/8.7	5.0/8.4	6.3/8.9	6.7/8.5	6.4/8.6	5.8/8.2	5.3/7.3	4.5/6.8	4.3/6.5	6.1/7.3
Latitude	Average	6.3	6.9	7.7	7.8	6.9	7.2	7.3	7.4	7.1	6.7	6.1	5.7	6.9
	Min/Max	5.6/7.1	6.3/7.6	6.7/8.5	6.4/8.6	4.9/8.2	6.0/8.5	6.5/8.1	6.2/8.4	5.8/8.4	5.5/7.6	4.8/7.3	4.6/7.2	6.2/7.4
Latitude +15	Average	6.6	7.1	7.6	7.5	6.5	6.7	6.8	7.0	7.0	6.7	6.3	6.0	6.8
	Min/Max	5.8/7.4	6.4/7.8	6.6/8.4	6.2/8.3	4.6/7.7	5.6/7.9	6.0/7.6	5.9/8.0	5.7/8.2	5.5/7.7	4.9/7.7	4.8/7.7	6.2/7.4

Solar Radiation for 2-Axis Tracking Flat-Plate Collectors (kWh/m²/day), Uncertainty ±9%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
2-Axis	Average	6.6	7.1	7.7	7.9	7.2	7.7	7.7	7.6	7.1	6.7	6.4	6.1	7.2
	Min/Max	5.9/7.5	6.4/7.8	6.7/8.5	6.5/8.8	5.1/8.5	6.4/9.1	6.8/8.6	6.4/8.6	5.9/8.4	5.5/7.7	5.0/7.7	4.9/7.8	6.5/7.7

Direct Beam Solar Radiation for Concentrating Collectors (kWh/m²/day), Uncertainty ±8%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
1-Axis, E-W	Average	3.5	3.5	3.7	3.7	3.2	3.6	3.6	3.5	3.2	3.1	3.2	3.2	3.4
Horiz Axis	Min/Max	2.8/4.3	2.9/4.2	2.9/4.3	2.6/4.5	1.6/4.3	2.6/4.8	2.9/4.3	2.5/4.4	2.4/4.0	2.2/4.1	2.0/4.5	2.2/4.8	2.9/4.0
1-Axis, N-S	Average	3.5	4.0	4.8	4.9	4.1	4.5	4.6	4.6	4.2	3.7	3.3	3.0	4.1
Horiz Axis	Min/Max	2.8/4.3	3.2/4.8	3.6/5.6	3.4/6.1	2.1/5.5	3.3/6.2	3.7/5.6	3.3/5.9	3.1/5.4	2.6/4.9	2.0/4.6	2.1/4.7	3.5/4.8
1-Axis, N-S	Average	4.0	4.4	5.0	4.9	4.0	4.3	4.4	4.5	4.3	4.0	3.7	3.5	4.2
Tilt=Latitude	Min/Max	3.2/4.9	3.5/5.2	3.8/5.8	3.3/6.0	2.0/5.3	3.1/5.8	3.5/5.3	3.2/5.7	3.2/5.6	2.8/5.3	2.3/5.2	2.5/5.4	3.6/5.0
2-Axis	Average	4.3	4.5	5.0	5.0	4.2	4.6	4.7	4.6	4.3	4.0	3.9	3.8	4.4
	Min/Max	3.4/5.3	3.6/5.4	3.8/5.8	3.4/6.1	2.1/5.6	3.3/6.3	3.7/5.7	3.3/5.9	3.2/5.6	2.8/5.3	2.5/5.5	2.7/5.9	3.7/5.2

Average Climatic Conditions

Element	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
Temperature (°C)	25.0	25.1	25.6	26.3	27.2	27.9	28.1	28.2	28.1	27.7	26.7	25.6	26.8
Daily Minimum Temp	21.6	21.4	22.0	22.7	23.6	24.5	24.9	24.8	24.6	24.2	23.3	22.4	23.3
Daily Maximum Temp	28.4	28.7	29.1	29.9	30.7	31.4	31.4	31.5	31.6	31.3	29.9	28.8	30.2
Record Minimum Temp	16.1	16.7	15.6	17.8	18.9	20.6	20.6	21.1	20.6	19.4	18.9	17.2	15.6
Record Maximum Temp	33.3	35.6	35.6	36.1	35.6	36.1	35.0	36.1	36.1	36.7	35.6	34.4	36.7
HDD, Base 18.3°C	0	0	0	0	0	0	0	0	0	0	0	0	0
CDD, Base 18.3°C	207	188	224	240	274	288	303	305	292	291	250	226	3088
Relative Humidity (%)	74	72	71	71	75	76	76	76	76	77	76	75	75
Wind Speed (m/s)	3.5	3.8	4.0	3.8	3.5	3.8	4.1	3.7	3.2	2.9	3.2	3.5	3.6

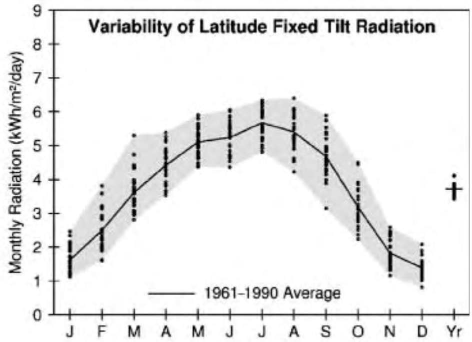
FIGURE 19.21 Solar radiation data manual table for San Juan, PR.

Seattle, WA

WBAN NO. 24233

LATITUDE: 47.45° N
LONGITUDE: 122.30° W
ELEVATION: 122 m
MEAN PRESSURE: 1001 millibars

STATION TYPE: Primary



Solar Radiation for Flat-Plate Collectors Facing South at a Fixed Tilt (kWh/m²/day), Uncertainty ±9%

Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	1.0	1.7	2.8	4.1	5.3	5.8	6.1	5.2	3.8	2.2	1.2	0.8	3.3
	Min/Max	0.8/1.2	1.2/2.1	2.4/3.6	3.5/4.8	4.6/6.1	4.8/6.7	5.2/6.7	4.2/5.9	2.9/4.5	1.8/2.8	0.9/1.4	0.6/1.0	3.2/3.6
Latitude -15	Average	1.5	2.3	3.5	4.6	5.4	5.7	6.1	5.6	4.7	3.0	1.7	1.3	3.8
	Min/Max	1.1/2.2	1.5/3.5	2.8/5.1	3.7/5.5	4.7/6.3	4.7/6.6	5.2/6.8	4.4/6.6	3.2/5.8	2.2/4.2	1.1/2.3	0.8/1.8	3.5/4.2
Latitude	Average	1.6	2.5	3.6	4.4	5.1	5.2	5.7	5.4	4.7	3.2	1.8	1.4	3.7
	Min/Max	1.1/2.5	1.6/3.8	2.8/5.3	3.5/5.4	4.4/5.9	4.4/6.1	4.8/6.3	4.2/6.4	3.1/5.9	2.2/4.5	1.1/2.6	0.8/2.1	3.4/4.1
Latitude +15	Average	1.7	2.5	3.5	4.1	4.5	4.5	4.9	4.9	4.5	3.2	1.8	1.4	3.5
	Min/Max	1.1/2.6	1.6/3.9	2.7/5.2	3.2/5.0	3.9/5.2	3.8/5.2	4.2/5.5	3.8/5.8	3.0/5.7	2.2/4.5	1.1/2.7	0.8/2.2	3.2/3.8
90	Average	1.5	2.2	2.8	3.0	3.0	2.8	3.1	3.4	3.4	2.7	1.7	1.3	2.6
	Min/Max	1.0/2.5	1.4/3.5	2.2/4.3	2.3/3.6	2.6/3.3	2.4/3.2	2.8/3.4	2.7/4.0	2.3/4.4	1.8/3.9	1.0/2.5	0.7/2.1	2.3/2.9

Solar Radiation for 1-Axis Tracking Flat-Plate Collectors with a North–South Axis (kWh/m²/day), Uncertainty ±9%

Axis Tilt (°)		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
0	Average	1.2	2.2	3.7	5.2	6.7	7.3	8.0	6.9	5.1	3.0	1.5	1.0	4.3
	Min/Max	0.9/1.8	1.5/3.3	2.9/5.4	4.1/6.3	5.3/8.2	5.8/9.1	6.1/9.0	5.0/8.4	3.4/6.6	2.1/4.2	0.9/2.0	0.7/1.4	3.8/4.8
Latitude -15	Average	1.6	2.7	4.2	5.6	6.9	7.3	8.2	7.3	5.8	3.6	1.9	1.4	4.7
	Min/Max	1.1/2.5	1.7/4.3	3.2/6.5	4.3/7.0	5.3/8.6	5.8/9.3	6.2/9.2	5.3/9.0	3.7/7.6	2.5/5.2	1.1/2.7	0.8/2.1	4.1/5.3
Latitude	Average	1.8	2.8	4.3	5.5	6.7	7.0	7.9	7.2	5.9	3.7	2.0	1.5	4.7
	Min/Max	1.2/2.8	1.8/4.6	3.3/6.7	4.2/6.9	5.1/8.3	5.5/8.9	5.9/8.9	5.2/8.8	3.7/7.7	2.5/5.5	1.1/2.9	0.8/2.3	4.1/5.3
Latitude +15	Average	1.8	2.8	4.2	5.3	6.3	6.6	7.4	6.8	5.7	3.7	2.0	1.5	4.5
	Min/Max	1.2/2.9	1.7/4.6	3.2/6.6	4.0/6.6	4.7/7.8	5.1/8.3	5.5/8.4	4.9/8.4	3.6/7.5	2.4/5.5	1.1/3.0	0.8/2.4	3.9/5.1

Solar Radiation for 2-Axis Tracking Flat-Plate Collectors (kWh/m²/day), Uncertainty ±9%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
2-Axis	Average	1.8	2.9	4.3	5.6	7.0	7.5	8.3	7.4	5.9	3.7	2.0	1.5	4.9
	Min/Max	1.2/2.9	1.8/4.7	3.3/6.7	4.3/7.0	5.4/8.7	5.9/9.5	6.3/9.4	5.3/9.0	3.8/7.7	2.5/5.5	1.1/3.0	0.8/2.4	4.2/5.5

Direct Beam Solar Radiation for Concentrating Collectors (kWh/m²/day), Uncertainty ±8%

Tracker		Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
1-Axis, E-W	Average	0.9	1.4	1.9	2.3	3.0	3.3	4.0	3.5	2.9	1.8	1.0	0.8	2.2
Horiz Axis	Min/Max	0.4/1.8	0.6/2.9	1.2/3.8	1.3/3.3	1.5/4.3	1.7/4.9	2.3/5.1	2.0/4.7	1.4/4.2	1.0/3.2	0.3/1.7	0.3/1.5	1.6/2.6
1-Axis, N-S	Average	0.5	1.1	1.9	2.8	3.8	4.2	5.2	4.4	3.2	1.6	0.6	0.4	2.5
Horiz Axis	Min/Max	0.2/1.1	0.5/2.3	1.1/3.8	1.6/4.0	1.7/5.7	2.1/6.4	2.8/6.5	2.4/5.9	1.6/4.8	0.7/2.8	0.2/1.2	0.2/0.8	1.8/3.0
1-Axis, N-S	Average	0.9	1.6	2.5	3.1	3.9	4.1	5.1	4.7	3.9	2.2	1.1	0.8	2.8
Tilt=Latitude	Min/Max	0.4/2.0	0.7/3.4	1.4/4.9	1.8/4.4	1.7/5.7	2.0/6.2	2.8/6.4	2.6/6.3	1.9/5.7	1.1/3.9	0.3/1.9	0.3/1.5	2.0/3.3
2-Axis	Average	1.0	1.7	2.5	3.1	4.1	4.4	5.5	4.8	3.9	2.3	1.1	0.9	2.9
	Min/Max	0.4/2.1	0.7/3.5	1.5/4.9	1.8/4.5	1.8/6.0	2.2/6.7	3.0/6.9	2.7/6.5	1.9/5.7	1.1/4.0	0.3/2.0	0.3/1.6	2.1/3.5

Average Climatic Conditions

Element	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec	Year
Temperature (°C)	4.5	6.4	7.6	9.6	12.8	16.1	18.4	18.6	15.9	11.6	7.4	4.7	11.1
Daily Minimum Temp	1.8	3.0	3.6	5.1	7.9	11.1	12.9	13.2	11.1	7.7	4.5	2.1	7.0
Daily Maximum Temp	7.2	9.7	11.5	14.0	17.7	21.1	24.0	24.0	20.7	15.4	10.3	7.3	15.2
Record Minimum Temp	-17.8	-17.2	-11.7	-1.7	-2.2	3.3	6.1	6.7	1.7	-2.2	-14.4	-14.4	-17.8
Record Maximum Temp	17.8	21.1	23.9	29.4	33.9	35.6	37.2	37.2	36.7	31.7	23.3	17.2	37.2
HDD, Base 18.3°C	429	334	334	263	171	80	32	36	87	210	328	422	2727
CDD, Base 18.3°C	0	0	0	0	0	12	36	45	13	0	0	0	106
Relative Humidity (%)	78	75	74	71	69	67	66	68	73	79	80	80	73
Wind Speed (m/s)	4.1	4.0	4.0	4.0	3.7	3.7	3.5	3.4	3.4	3.5	4.0	4.0	3.8

FIGURE 19.22 Solar radiation data manual table for Seattle, WA.

temperatures, record minimum and maximum temperatures, average heating and cooling degree days, average relative humidity, and average wind speed.

The manual is organized alphabetically by state and, within each state, alphabetically by station. It explains how to interpret data in the tables and discusses the methods used to determine the data values. Manual and data tables are out of print, but can be accessed at NREL's Renewable Resource Data Center (see Section 19.1.4.4 on online data products and maps).

19.1.4.3 Development of a New National Solar Radiation Database

The National Renewable Energy Laboratory is currently updating the NSRDB to extend the database through the period 1991–2000. A recent paper (Wilcox et al. 2005) provides an overview of the procedures being undertaken to develop the extended database. Since the National Weather Service eliminated the human weather observation activities in the mid-1990s and converted weather service stations to automated surface observation systems (ASOS), a key data input to the METSTAT model—the percentage of cloud cover that is opaque—is no longer available. To adapt the model to currently available datasets, NREL derived equivalent sky cover inputs (total and opaque cloud cover) from a combination of ASOS and ASOS supplemental cloud measurements (the latter derived from GOES satellite data). ASOS detects clouds to 12,000 ft. (3660 m), whereas the ASOS supplemental cloud measurements provide sky cover estimates for heights above 12,000 ft. for a 50×50 km^2 area centered upon the ASOS station.

Combining the low cloud amounts determined from the ASOS data, and the middle and high cloud amounts determined from the ASOS supplemental data, produced an equivalent total cloud estimate. Applying opacity factors based on cloud types inferred from cloud height produced estimates of opaque cloud amounts (low clouds opacity factor = 1.00; middle clouds opacity factor = 0.93; and high clouds opacity factor = 0.44). Although availability of ASOS data for the 33 test sites was 95% or greater for 1999 and 2000, the ASOS supplemental cloud data was only available for about 70% of the period, with no data available for Alaska. About 1% of the ASOS supplemental data was unusable due to file format errors.

The original METSTAT included an algorithm to compensate for human perception of cloud amounts nearer the horizon being greater than actual. This feature was eliminated because of the strictly vertical view of the ASOS cloud detector.

Other data sources are also being examined to enhance the spatial coverage beyond the 239 stations available for the original NSRDB. Toward that end, the updated NSRDB will make use of a model from the Atmospheric Sciences Research Center (ASRC) at SUNYA (Perez et al. 2002). This SUNYA model derives 10-km pixel solar estimates based on differences between a pixel's clear sky reflectance as seen by the satellite and the brighter values that occur with increasing cloud reflectance of incoming solar radiation.

The model has been evaluated using special model runs for 1999 and 2000 for 33 surface sites chosen for testing various aspects of the NSRDB, as well as the full resolution grids for the 48 lower states, Hawaii and the portion of Alaska below 60° latitude. Figure 19.23 shows a map of annual DNI averages from this dataset.

Because the ASRC satellite imagery archive begins in mid-1998, a large data gap for most of the 1990s still exists. An approach being considered for prior years is to relate monthly averaged post-1998 model runs to time-coincident NASA Surface Meteorology and Solar Energy (SSE) gridded data.

Other aspects to enhance the overall database and assure its continuity in decades to come are also being explored during the current updating process. These include the acquisition of more recent high-quality solar measurement data, improvements to the clear sky algorithm that goes into solar models such as METSTAT, and improved water vapor, aerosol, and ozone data. In addition, the Northeast Regional Climate Center (NRCC), through funding from American Society of Heating, Refrigerating and Air-Conditioning Engineers, produced a solar radiation model aimed primarily at supporting applications for building energy design (Belcher and DeGaetano 2004). That model computes GHI and DNI from the product of extraterrestrial direct beam, modified by the cosine of the zenith angle, and transmittance functions for typical atmospheric attenuation, Rayleigh scattering, gas and water vapor

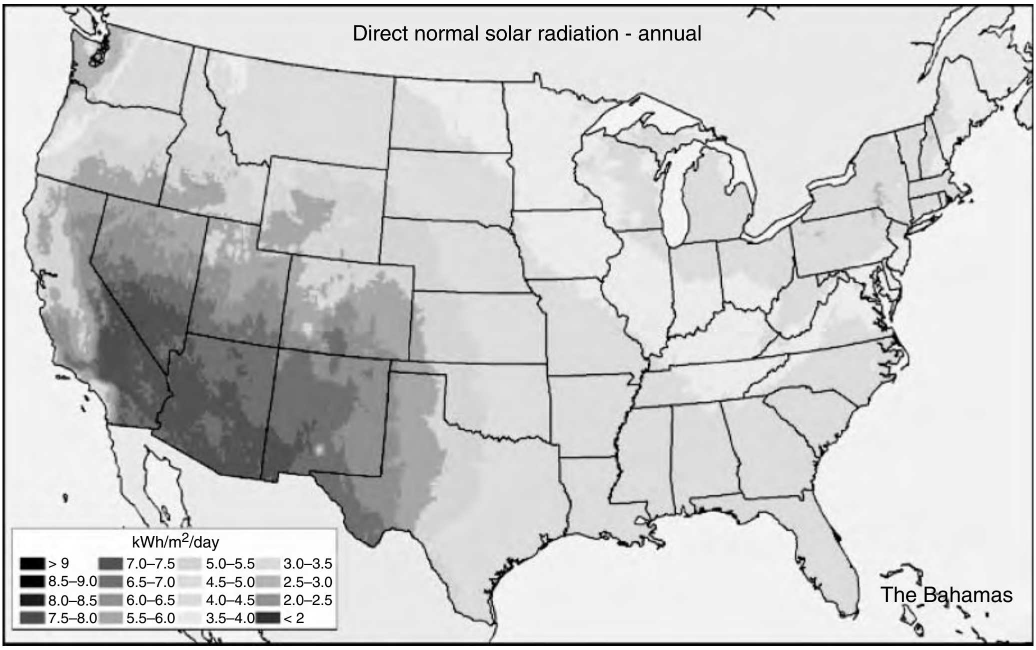


FIGURE 19.23 **(See color insert following page 774.)** Annual direct normal irradiance on 10-km grid, derived from the SUNYA satellite model.

absorption, aerosol absorption, and cloud absorption. This model is currently being tested for use in the NSRDB updates.

By using different models and different methods of cloud observations, discontinuities or inconsistencies may exist in data quality over the decade of the 1990s. We recommend characterizing the differences between the old and new NSRDB by (1) comparing model performance against measured data for both the original METSTAT model and the model used for the update and (2) using the NASA SSE data grid for the 1990s as a reference for identifying differences among the hybrid methods. These characterizations will also help illuminate the limitations of the NSRDB for certain critical applications (such as climate change).

The preliminary work that has been conducted thus far shows that it is feasible to produce an updated NSRDB that not only extends the period of record for the existing database, but also exploits new technologies that pave the way for greatly enhanced solar resource assessment applications. The release of the updated NSRDB is currently planned for 2007. NREL's production plans include a provision to produce incremental annual updates, which will be immediately applied to years 2001–2005, then annually through 2009. Current plans also call for a 2001–2010 decade update to be produced in 2011.

19.1.4.4 Data Products and Maps Online

Many of the database products described in this chapter as well as new products that will come out of the updated NSRDB are accessible via the internet from the Renewable Resource Data Center (RReDC) located at the National Renewable Energy Laboratory. The data center contains information developed by the Laboratory's Resource Assessment Program and allows users to comment on the program's products and services. It is accessible directly at http://rredc.nrel.gov. Solar resource maps have also been incorporated into Geographic Information Systems (GIS) software, and additional maps can be found under the National Solar Atlas site at http://www.nrel.gov/GIS.

19.1.5 Solar Design Tools

Since producing the solar radiation database, the National Renewable Energy Laboratory has developed new design tools to assist people interested in using the database. This section describes some of these tools. The TMY2s present a single, representative year of resource data for each of the 239 stations in the database.

The section also presents a basic set of equations required to determine solar radiation on a horizontal surface and then translate this information into data for tilted surfaces.

19.1.5.1 Typical Meteorological Years 2

A typical meteorological year (TMY) is a dataset of hourly values of solar radiation and meteorological elements for a one-year period. A TMY consists of months selected from different individual years and linked together to form a complete year. For example, if 30 years of data are available, all 30 Januarys would be examined and the one judged most typical would be selected for inclusion in the TMY. The remaining months in the TMY would be chosen in a similar fashion, with particular attention paid to five statistical elements: global horizontal radiation, direct normal radiation, dry bulb temperature, dew point temperature, and wind speed.

The TMY2 datasets were derived from Version 1.1 of the National Solar Radiation Database, and are designated as TMY2s to distinguish them from an earlier TMY developed in the late 1970s for the earlier SOLMET.ERSATZ solar radiation database.

Designers and planners often find such a representative dataset useful in predicting the performance and economic viability of a solar energy system. For instance, the TMY2s were designed for computer simulations of solar energy conversion and building systems. They provide a standard for hourly data for solar radiation and other meteorological elements that supports performance comparisons of system types and configurations for one or more locations.

A TMY is not specific to any particular year and should not be used as an indicator of actual conditions expected over the next year, or even 5 years. Rather, a TMY represents typical conditions over a long time period such as 30 years. Because TMY data represent typical rather than extreme conditions at a specific location, they should be used with caution for component and system design. System hardware should be deigned to withstand a location's worst-case conditions, rather than typical conditions.

The TMY2 datasets are available via the Internet from the National Renewable Energy Laboratory's Renewable Resource Data Center (RreDC) at http://rredc.nrel.gov. TMY2 datasets for all 239 stations may also be obtained on a CD-ROM from

NREL Document Distribution Service
1617 Cole Boulevard
Golden, CO 80401-3393
Telephone: 303-275-4363
Facsimile: 303-275-4053
Internet: Robert_finger@nrel.gov

A printed user's manual (Marion and Urban 1995) is also available from the document distribution center.

19.1.5.2 Estimating Daily and Hourly Radiation from Monthly Averages

The only extant solar radiation data for developing countries is often monthly averages of daily total global horizontal radiation. Given sufficient time and money, an engineer could conduct a multiyear measurement program. Lacking such resources and time, a statistical correlation approach such as the developed by Reddy (1987) can be used. This approach assumes that correlations between monthly averages and daily and hourly distributions at locations having measured data can be used to derive daily and hourly data for locations exhibiting similar clearness indices. This section summarizes Reddy's approach.

19.1.5.3 Horizontal Surfaces

The first step in estimating daily and hourly values of solar radiation on horizontal surfaces is the calculation of the clearness index, $\bar{K}$, using available monthly average daily total global horizontal data and Equation 19.8. $\bar{H}_d$ can be used to predict the monthly average daily total horizontal diffuse radiation. The monthly average daily total horizontal beam radiation can then be obtained from the global solar radiation by subtraction.

The correlation between the clearness index and the average daily total horizontal diffuse radiation takes account of the effect of seasonal variation by including the sunset hour angle, W_S. The correlation for a clearness index in the range of $0.3 < \bar{K} < 0.8$ is

$$\begin{aligned}\frac{\bar{H}_d}{\bar{H}_0} &= 1.391 - 3.560\bar{K} + 4.189\bar{K}^2 - 2.137\bar{K}^3 \text{ for } W_s \leq 81.4° \\ &= 1.311 - 3.022\bar{K} + 3.427\bar{K}^2 - 1.821\bar{K}^3 \text{ for } W_s > 81.4°, \qquad (19.9)\end{aligned}$$

where W_S is the monthly mean value of the sunset hour angle in degrees on a horizontal surface.

When subjected to statistical analysis, different locations with the same mean monthly clearness index, $\bar{K}$, often display cumulative frequency curves that are more or less identical. This gives rise to the long-term implication that the monthly average distribution of the daily total global horizontal radiation may be independent of location and month. This relationship is most consistent for temperate regions. Thus, if one assumes that $\bar{K}$ is known and that the random variable K will also have known minimum and maximum values, then a generalized probability function can be represented as follows:

$$P(K) = C\left(\frac{K_{\max} - K}{K_{\max}}\right)e^{\gamma K} \quad \text{for} \quad K_{\min} \leq K \leq K_{\max}, \qquad (19.10a)$$

where C and γ can be calculated from

$$\begin{aligned}&\int P(K)\mathrm{d}K = 1, \\ &\int KP(K)\mathrm{d}K = \bar{K},\end{aligned} \qquad (18.10b)$$

and $K_{\max}$ can be calculated as follows:

$$K_{\max} = 0.6313 + 0.267\bar{K} - 11.9(\bar{K} - 0.75)^8. \qquad (19.10c)$$

The monthly average daily total diffuse radiation, $\bar{H}_d$, can be predicted from the daily total horizontal global radiation and the clearness index using the following relationship, which includes the sunset angle W_S to account for seasonal variation.

For $W_S < 81.4°$,

$$\frac{\bar{H}_d}{\bar{H}_0} = 1.0 - 0.2727K + 2.449K^2 - 11.9541K^3 + 9.3879K^4 \quad \text{for} \quad K < 0.715, \qquad (19.11a)$$

and

$$\frac{\bar{H}_d}{\bar{H}_0} = 0.143 \quad \text{for} \quad K \geq 0.715, \qquad (19.11b)$$

For $W_S > 81.4°$,

$$\frac{\bar{H}_d}{\bar{H}_0} = 1.0 + 0.2832K - 2.5557K^2 + 0.8448K^3 \text{ for } K < 0.722, \qquad (19.11c)$$

and

$$\frac{\bar{H}_\mathrm{d}}{\bar{H}_0} = -0.175 \quad \text{for} \quad K > 0.722. \tag{19.11d}$$

The hourly horizontal global radiation can also be predicted from daily total horizontal global radiation using a statistical correlation. The correlation is possible if average diurnal solar radiation is assumed to be symmetrical about solar noon. In the relationship, r is defined as the ratio of the monthly mean hourly global radiation to the monthly mean daily global radiation on a horizontal surface (i.e., $r=\bar{I}/\bar{H}$).The statistical correlation is expressed as

$$r(W) = \frac{\pi}{24}(a + b\cos W)\left[\frac{\cos W - \cos W_\mathrm{S}}{\left(\sin W_\mathrm{S} - \frac{\pi}{180} W_\mathrm{S} \cos W_\mathrm{S}\right)}\right], \tag{19.12}$$

Where $a=0.409+0.5016\sin(W_\mathrm{S}-60)$, $b=0.6609-0.4767\sin(W_\mathrm{S}-60)$, W_S is the sunset hour angle, W is the hour angle (in degrees) corresponding to the midpoint of the hour.

The correlation is valid for long-term averages and may be used for determining monthly mean values. It can also be applied to individual days so long as the days are clear. However, curve fitting is typically better when monthly timescales are employed.

Similarly, it is possible to predict hourly horizontal diffuse radiation. In this case r_d is defined as the ratio of the monthly mean hourly diffuse radiation to the monthly mean daily total diffuse radiation on a horizontal surface (i.e., $r_\mathrm{d} = I_\mathrm{d}/\bar{H}_\mathrm{d}$). The corresponding relationship is given by

$$r_\mathrm{d}(W) = \frac{\pi}{24}\left[\frac{\cos W - \cos W_\mathrm{S}}{\left(\sin W_\mathrm{S} - \frac{\pi}{180} W_\mathrm{S} \cos W_\mathrm{S}\right)}\right]. \tag{19.13}$$

This correlation can also be applied to clear days, but works better for monthly timescales.

If the daily total global radiation on a horizontal surface (H) is known, then the hourly diffuse radiation (I_d) can be estimated by first calculating the average global diffuse radiation (H_d) from Equation 19.11 and then using Equation 19.13. Alternatively, one can determine the hourly diffuse radiation (I_d) by calculating the hourly global radiation (I) from Equation 19.12 and then using the following relationships:

$$\frac{I_\mathrm{d}}{I} = 1.0 - 0.09K \quad \text{for} \quad K \le 0.22, \tag{19.14a}$$

$$\frac{I_\mathrm{d}}{I} = 0.9511 - 0.1604K + 4.388K^2 - 16.638K^3 + 12.336K^4 \quad \text{for} \quad 0.22 \le K \le 0.8, \tag{19.14b}$$

$$\frac{I_\mathrm{d}}{I} = 0.162 \quad \text{for} \quad K > 0.8, \tag{19.14c}$$

where K is the hourly clearness index and is equal to I/I_0.

Equation 19.14 should give the most consistent results because the interdependence between the hourly global radiation and the hourly diffuse radiation is generally greater than that between the hourly diffuse radiation and the average global diffuse radiation.

19.1.5.4 Tilted Surfaces

Radiation incident on a tilted surface is composed of beam radiation, diffuse radiation from the sky, and radiation reflected from the ground. An estimate of the beam radiation on a tilted surface can be obtained by first subtracting the diffuse radiation on a horizontal surface from the global horizontal radiation and then multiplying the resulting number by a geometrical conversion factor. Diffuse and reflected radiation values with respect to the tilted surfaces are obtained by multiplying their horizontal values by the fraction of the total hemisphere intersected by the incline plane.

The conversion fact for hourly global radiation (I_T) on a surface with a tilt angle β is

$$I_T = (I - I_d) r_{b,T} + I_d \left(\frac{1 + \cos \beta}{2} \right) + I \rho \left(\frac{1 - \cos \beta}{2} \right), \tag{19.15}$$

Where ρ is the surface albedo, β is the tilt angle of the inclined surface with respect to the horizontal, and $R_{b,T}$ is the ratio of hourly beam radiation on the tilted surface to that on a horizontal surface.

Similarly, the daily total global radiation (H_T) on a tilted surface can be determined as follows:

$$H_T = (H - H_d) R_{b,T} + H_d \left(\frac{1 + \cos \beta}{2} \right) + H \rho \left(\frac{1 - \cos \beta}{2} \right), \tag{19.16}$$

where $R_{b,T}$ is the ratio of the daily beam radiation on the tilted surface to that on the horizontal surface and ρ and β are as previously.

19.1.6 International Solar Radiation Data

The scarcity of solar radiation measurements in some countries has led to the use of satellite data for estimating solar radiation. An international solar radiation database, update surface radiation budget developed by the National Aeronautics and Space Administration's Langley Research Center has become available. The dataset uses a model to convert satellite images to daily total or monthly averages of the daily total surface solar radiation estimates with a grid resolution of 100 km. The dataset covering the period from July 1983–June 1993 is currently available at http://eosweb.larc.nasa.gov/cgi-bin/sse.

The World Radiation Data Center, located at the Main Geophysical Institute in St. Petersburg, Russia, was established in 1964 under the auspices of the World Meteorological Organization. Since then, it has received summarized solar radiation data from as many as 1250 stations throughout the world. Most of the measurement data consist of daily total global horizontal radiation, but some measurements of direct beam and diffuse radiation are also included.

The center's role has been to assess the quality of data submitted by different countries and to archive them. For many years the center has produced quarterly reports providing data summaries from each station that submits data. More recently, the center has established a Web site: http://wrdc.mgo.rssi.ru. In addition, a Web site of the centers data up through about 1994 was established at NREL as a collaborative activity with the WRDC. This site can be accessed at http://wrdc-mgo.nrel.gov.

A major international program, the Solar and Wind Energy Resource Assessment Project, funded by the Global Environment Facility and managed by the United Nations Environment Programme's Division of Technology, Industry and the Environment, has produced satellite-derived solar resource datasets and maps for 13 countries around the world (Guatemala, El Salvador, Honduras, Nicaragua, Cuba, Brazil, Ghana, Kenya, Ethiopia, Sri Lanka, Bangladesh, Nepal, and China). Several approaches, summarized by Renné et al. (2005) were employed to produce these data and maps, such that maps for specific countries, as well as large regions of the earth, are available. In addition, typical meteorological year data for selected ground stations in each of the countries were also developed. All data and maps can be found on the SWERA Web site: http:/.swera.unep.net.

19.2 Wind Energy

Dale E. Berg

19.2.1 Wind Origins

The primary causes of horizontal atmospheric air motion, or wind, are the uneven heating of the earth and its atmosphere by solar radiation and the earth's rotation. The earth's atmosphere reflects about 43%

of the incident solar radiation back into space, absorbs about 17% of it in the lower portions of atmosphere, and transmits the remaining 40% to the surface of the earth, where much of it is then reradiated into the atmosphere. The radiation from the hot sun is at short wavelengths (0.15–4 μm) and passes readily through the atmosphere, whereas the radiation from the cooler earth is at longer wavelengths (5–20 μm) and is readily absorbed by the water vapor in the atmosphere. Thus, the radiation from the earth is primarily responsible for the warmth of the atmosphere near the earth's surface. Heat is also transferred from the earth's surface to the atmosphere by conduction and convection.

On average, the total amount of energy radiated to space from the earth and its atmosphere must be equivalent to the total amount of solar radiation absorbed, or the temperature of the earth and its atmosphere would steadily increase or decrease. The more nearly perpendicular the sun's rays strike the earth, the more solar radiation is transferred through the atmosphere. Thus, during the year, tropical regions receive much more solar energy than do polar regions. Winds and ocean currents level out this imbalance in thermal energy, preventing the tropical regions from getting progressively hotter and the polar regions from getting progressively colder. In addition, the lack of homogeneity of the earth's surface—the land, water, desert, forest, rock, sand, black loam soil, etc.—leads to differences in solar radiation absorption and reflection back to the atmosphere, creating differences in atmospheric temperature, density, and pressure. These differences, in turn, create forces that contribute to wind. For example, the difference in radiation absorption and emission by land and water along a coastline is the dominant cause of the light shore winds or breezes; the difference in radiation absorption and emission by mountains and valleys is a significant contribution to the light upslope and downslope breezes often found in mountainous terrain. The earth's rotation results in Coriolis forces that accelerate each moving particle of air. This acceleration moves an air particle to the right of its direction of motion in the northern hemisphere, and to the left in the southern hemisphere. The geostrophic winds in the upper atmosphere (above 600 m or so) are due to the balance of these Coriolis forces and the pressure gradient forces resulting from the uneven heating of the atmosphere.

The earth's rotation also imparts an angular momentum to each particle of air in a west-to-east direction. Conservation of angular momentum results in an increase in its west-to-east velocity as the particle moves from the equator towards the poles. In the temperate zones, this causes the westerlies, winds opposite to the general flow, in both hemispheres.

The other long-term, large-scale global wind patterns such as equatorial doldrums, trade winds, easterlies, and subtropical and polar jets, illustrated in Figure 19.24, are also caused by the combination of differential solar heating and the rotation of the earth. These wind patterns are often referred to as the *general circulation patterns*. In actuality, these patterns are complicated by seasonal effects due to changes in the earth's position relative to the sun during the course of a year and geographical effects due to the uneven distribution of, and physical properties of, water and land surfaces. Centers of high or low pressure, caused by heating or cooling of the lower atmosphere, include hurricanes, monsoons, and extratropical cyclones. Small-scale, local scale phenomena characterized by local wind include land and sea breezes, valley and mountain winds, monsoon-like flow, foehn winds (dry, high-temperature winds on the downwind side of mountain ranges commonly referred to in the western United States as "chinooks"), thunderstorms, and tornadoes.

Variations of wind speed in time can be divided into the categories of interannual, annual, diurnal and short-term. Interannual variations in wind speed occur over timescales greater than one year. They can have a large effect on long-term wind turbine production. Meteorologists generally conclude that it takes 30 years of data to determine long-term values of weather or climate and that it takes at least five years to arrive at a reliable average annual wind speed at a given location. However, that does not mean that data spanning shorter periods of time are useless.

Annual variations refer to significant variations in seasonal or monthly averaged wind speeds that are common over most of the world. For example, for the eastern one-third of the United States, maximum average wind speeds occur during the winter and early spring. Spring maximums occur over the Great Plains, the North Central States, the Texas Coast, in the basins and valleys of the West and the coastal areas of Central and Southern California. Winter maximums occur over all U.S. mountainous regions,

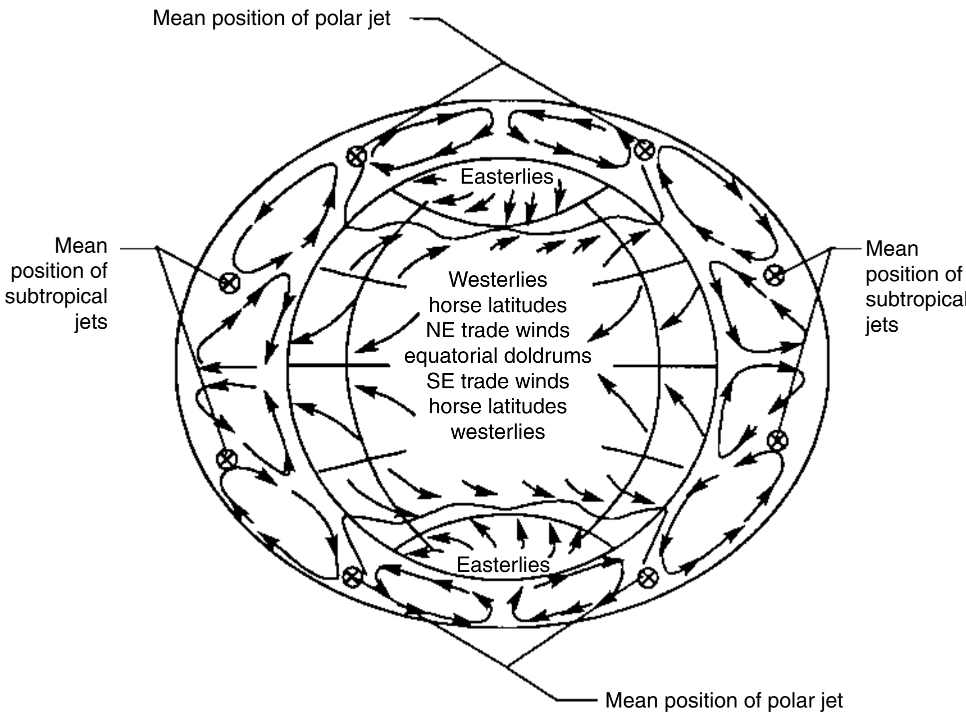


FIGURE 19.24 Semipermanent global wind patterns. (From Spera, D., ed. 1994. *Wind Turbine Technology, Fundamental Concepts of Wind Turbine Engineering,* ASME Press, New York, With Permission.)

except for some areas in the lower southwest, where spring maximums occur. Spring and summer maximums occur in the wind corridors of Oregon, Washington, and California. Figure 19.25 illustrates how the monthly average wind speed at a location can vary significantly from year to year.

Large diurnal or time-of-day variations in wind occur in both tropical and temperate latitudes. This type of wind speed variation is due to differential heating of the earth's surface during the daily radiation cycle. A typical diurnal variation is an increase in wind speed during the day, with the wind speeds lowest

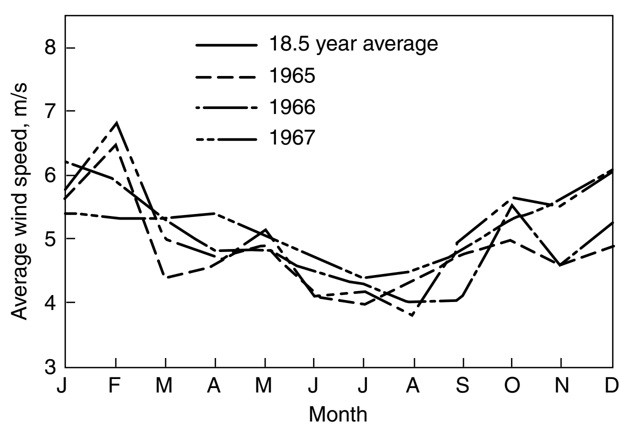


FIGURE 19.25 Seasonal changes of monthly average wind speeds for Billings, Montana. (From Heister, T. R., and Pennell, W. T. 1981. The meteorological aspects of siting large wind turbines, U.S. Department of Energy Report PNL-2522, Pacific Northwest Laboratory, Richland, WA.)

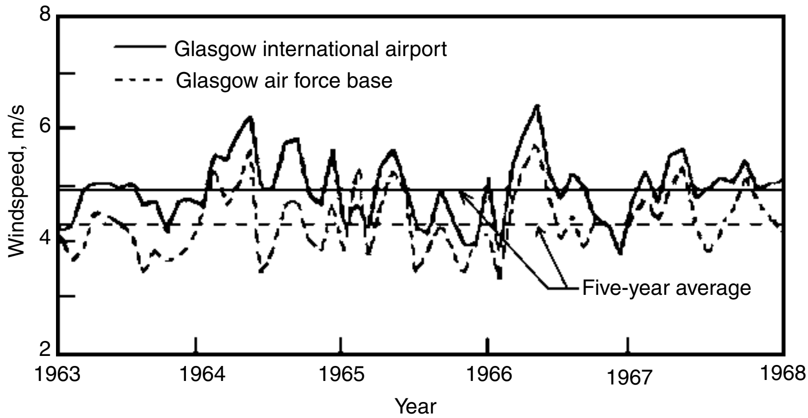


FIGURE 19.26 Time series of monthly wind speed for Glasgow, Montana International Airport and Air Force Base. (From Heister, T. R., and Pennell, W. T. 1981. The meteorological aspects of siting large wind turbines, U.S. Department of Energy Report PNL-2522, Pacific Northwest Laboratory, Richland, WA.)

during the hours from midnight to sunrise. Diurnal wind variations in temperate latitudes over relatively flat land areas are usually due to daily variations in solar radiation. The largest diurnal changes generally occur in spring and summer, and the smallest in winter. The diurnal variation may vary with location and altitude above sea level. For example, at altitudes high above surrounding terrain, e.g., mountains or ridges above a plain, the diurnal pattern may be very different from the pattern for the surrounding terrain. This variation is due to mixing or transfer of momentum from the upper air to the lower air. There may be significant year-to-year differences in diurnal behavior, even at fairly windy locations. Although gross features of the diurnal cycle can be established with a single year of data, a good characterization of more detailed features, such as the amplitude of the diurnal oscillation and the time of day that the maximum winds occur, requires multiple years of data.

Variations in wind speed with periods of less than 10 min and that have a stochastic or random character are generally considered to represent turbulence, or fluctuations imposed on the mean wind speed. A gust is a short-term discrete event within a turbulent wind field. The common method of characterizing a gust is to measure or specify the amplitude, the rise time, the maximum gust variation, and the lapse time associated with it. Wind speed is very dependent on local topographical and ground cover variations. Figure 19.26 illustrates differences between two sites that are located in nominally flat terrain only 21 km apart. In spite of the close proximity and the flat terrain, the five year average mean wind speeds differ by about 12%.

19.2.2 Energy Available from the Wind

Given the fact that the wind blows everywhere, the idea of harnessing it to provide power is very appealing. Mankind has harnessed the power of the wind to produce mechanical power for well over 1000 years; the use of the wind to mill grain in Persia in the tenth century is well documented, and many experts speculate that the Chinese may have invented the windmill as much as 2000 years ago. How much power is available in the wind? Is it enough to be a viable source of energy in the modern world? Answers to these questions require knowledge of several wind energy basics.

19.2.2.1 Wind Power

Consider the air with an average velocity U and density ρ in a cylinder of cross section A and length L, as shown in Figure 19.27. The mass of the air in that cylinder is ρAL, and the kinetic energy (KE) of that air is $\frac{1}{2}\rho ALU^2$ (watt-hours or Wh). Now assume that there is a wind turbine rotor at the downwind end of that cylinder of air. The cylinder of air will pass through that rotor in a period of time, T, where $T = L/U$.

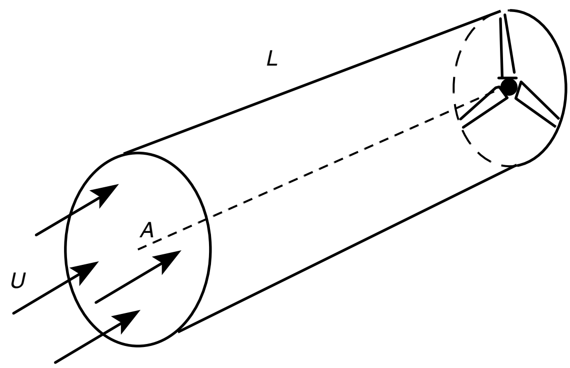


FIGURE 19.27 Steady wind passing through turbine rotor disk.

The power, or time rate of change of the kinetic energy, available at the rotor then is $\frac{1}{2}\rho ALU^2/T$, which becomes

$$\text{Power} = \tfrac{1}{2}\rho AU^3 \text{(W)}. \tag{19.17}$$

This is often rearranged as

$$\text{WPD (wind power density)} = \text{Power}/A = \tfrac{1}{2}\rho U^3 (\text{W/m}^2). \tag{19.18}$$

From the above equation, it is obvious that the most important factor in the available wind power is the velocity of the wind. Increasing the wind velocity by only 20%, from 5 to 6 m/s, for example, increases

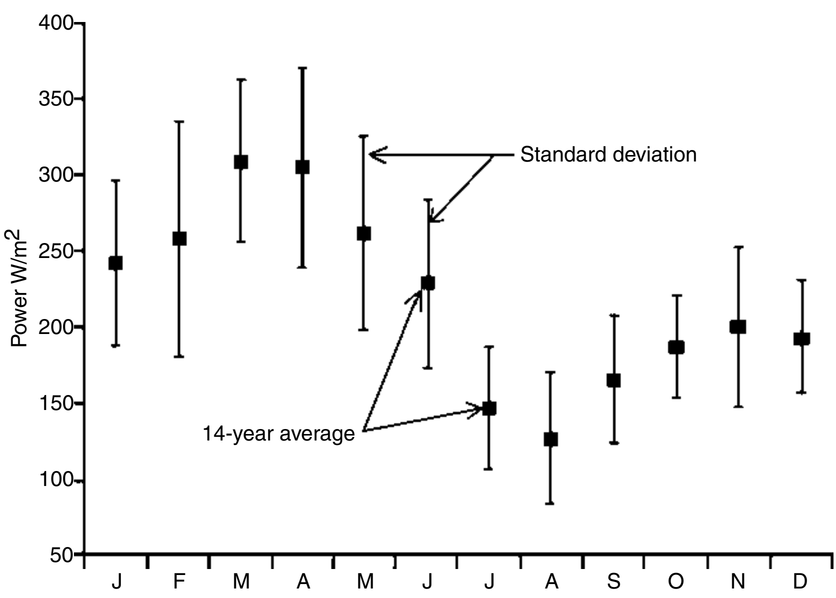


FIGURE 19.28 Seasonal variation of available wind power density for Amarillo, Texas. (From Rohatgi, J. S. and Nelson, V. 1994. *Wind Characteristics: An Analysis for the Generation of Wind Power*, Alternative Energy Institute, West Texas A&M University, Canyon, TX. With permission.)

the available wind power by 73%. Figure 19.28 illustrates the impact of normal annual wind speed variations at a location on the available wind power at that location.

The standard value for air density at sea-level reference conditions of 101,325 Pa pressure and 15°C is 1.225 kg/m^3. The actual density depends on the moisture content or humidity of the air, the temperature and the atmospheric pressure; however, the influence of humidity is very small and is normally neglected. Under this condition, the air density can be calculated from the perfect gas law:

$$\rho = \frac{P}{RT}\,(\mathrm{kg/m^3}), \tag{19.19}$$

where P is the atmospheric pressure in Pa or N/m^2, R is the specific gas constant for air (287 J/kg·K), and T is the absolute air temperature in Kelvin.

If site air pressure is not available, air density can be estimated as a function of site elevation, z, and absolute temperature as (AWS Scientific 1997):

$$\rho = \left(\frac{353.05}{T}\right) \mathrm{e}^{-0.034\left(\frac{z}{T}\right)}\,(\mathrm{kg/m^3}). \tag{19.20}$$

For example, the air density at Denver, Colorado (elevation 1600 m or 5300 ft. above sea level) is approximately 15% lower than that at sea level, so wind of a given velocity at Denver contains 15% less power than wind of the same velocity at sea level (assuming the temperature is the same).

19.2.2.2 Wind Shear

Wind moving across the earth's surface is slowed by trees, buildings, grass, rocks, and other obstructions in its path, resulting in a wind velocity that varies with height above the earth's surface—a phenomena known as the *vertical wind profile* or *vertical wind shear*. In most locations, wind shear is positive (wind speed increases with height), but situations in which the wind shear is negative or inverse are not unusual. In the absence of actual data for a specific site, a commonly used approximation for wind shear in an open area is:

$$U/U_{\mathrm{o}} = (h/h_{\mathrm{o}})^{\alpha}, \tag{19.21}$$

where U is the velocity at a height h, U_{o} is the measured velocity at height h_{o}, and α is the wind shear exponent.

The above equation is only an estimate; a specific site may display much different wind shear behavior, and that will dramatically affect site energy capture; it is therefore important to measure the wind resource at the specific site and height where a wind turbine will be located.

The instantaneous wind shear exponent varies widely with elevation, time of day, season of the year, wind speed, temperature, and nature of the terrain. It could be less than 0.14 during the day and then reach 0.5 at night (Golding 1977). The time-averaged shear exponent is the value that is normally used (averaged over several weeks). It varies with terrain characteristics, but usually falls between 0.10 and 0.25. Wind over a body of open water is normally well modeled by a value of α of about 0.10; wind over a smooth, level, grass-covered terrain such as the U.S. Great Plains by a value of α of about 0.14; wind over row crops or low bushes with a few scattered trees by a value of α of 0.20; and wind over a heavy stand of trees, several buildings, or hilly or mountainous terrain by a value of α of about 0.25. Short-term time-averaged shear factors as large as 1.25 have been documented in rare, isolated cases.

The U.S. Department of Energy is supporting an ongoing program to acquire additional data from tall towers (80 m or taller) around the United States as an initial stage in developing a comprehensive climatology at heights in excess of 50 m for continued wind energy development in the United States. National Renewable Energy Laboratory (NREL) scientists recently analyzed wind data obtained from several of these towers at heights of up to 110 m. They concluded that the time-averaged wind shear exponent for heights between 50 and 100 m may be significantly higher than the 0.14 value often used for estimating the wind resource at an elevation above that at which it was actually measured. Time-averaged shear exponents at well-exposed sites commonly range from 0.2 to 0.23; and the windiest sites have slightly lower shear exponents of 0.18–0.20 (Schwartz and Elliott 2005). As a result of wind shear, the available wind power at a site may vary dramatically with height. For example, for $\alpha = 0.20$, Equation 19.17 and Equation 19.21 reveal that the available wind power density at a height of 50 m is approximately $\{(50/10)^{0.2}\}^3 = 2.63$ times the available wind power density at a height of 10 m.

19.2.2.3 Available Resources

The amount of energy available in the wind (the wind energy resource) at a site is the average amount of power available in the wind over a specified period of time, commonly one year. If the wind speed is 20 m/s, the available power is very large at that instant, but if it only blows at that speed for 10 h per year and the rest of the time the wind speed is near zero, the resource for the year is small. Therefore, the site wind speed distribution (the relative frequency of occurrence for each wind speed), or the wind speed probability density function (pdf) is very important in determining the resource.

If the actual wind speed probability density distribution is not available, it is commonly approximated with the generalized two-parameter Weibull distribution given by:

$$f(U) = \frac{k}{C}\left(\frac{U}{C}\right)^{k-1} \exp\left[-\left(\frac{U}{C}\right)^{k}\right], \tag{19.22}$$

where $f(U)$ is the frequency of occurrence of wind speed U, k is the Weibull shape factor, and C is the Weibull scale factor.

For the special case with parameter k equal to 2, this reduces to the commonly used Rayleigh distribution:

$$f(U) = \frac{\pi}{2}\frac{U}{\bar{U}^2} \exp\left[-\frac{\pi}{4}\frac{U^2}{\bar{U}^2}\right], \tag{19.23}$$

where $\bar{U}$ is the yearly average wind speed.

It is readily apparent that the average wind speed at a site yields a unique Rayleigh distribution for that site. Experimental data from a site also yields the best Weibull distribution for that site, but the process of determining k and C from the experimental data is not straightforward.

The measured wind speed distribution at the Amarillo, Texas airport (yearly average wind speed of 6.6 m/s) is plotted in Figure 19.29, together with a Weibull and the Rayleigh distributions for that wind speed. It is obvious that the Rayleigh distribution is not a good representation for these data; the data has a higher peak, has a lower probability of lower wind speeds, and a significantly higher probability of winds in the 4–12 m/s range. The Weibull distribution (the shape factor k was picked as 2.6, yielding an estimate for the scale factor C of 7.4) is a much better estimation of the data. It shows a good match to the peak probability, but the data consistently has a somewhat lower probability for wind speeds below about 10 m/s and a somewhat higher probability for wind speeds above 10 m/s.

With a wind speed distribution and the local air density, the wind energy resource (WER), or average wind power availability, for a specific site can be estimated as:

$$\text{WER} = \frac{1}{2} A \sum_{i=1}^{n} \rho f(U_i) \Delta U_i U_i^3, \tag{19.24}$$

where n is the number of wind speeds considered and $f(U_i)\Delta U_i$ is the probability of wind speed U_i occurring in the wind speed range ΔU_i.

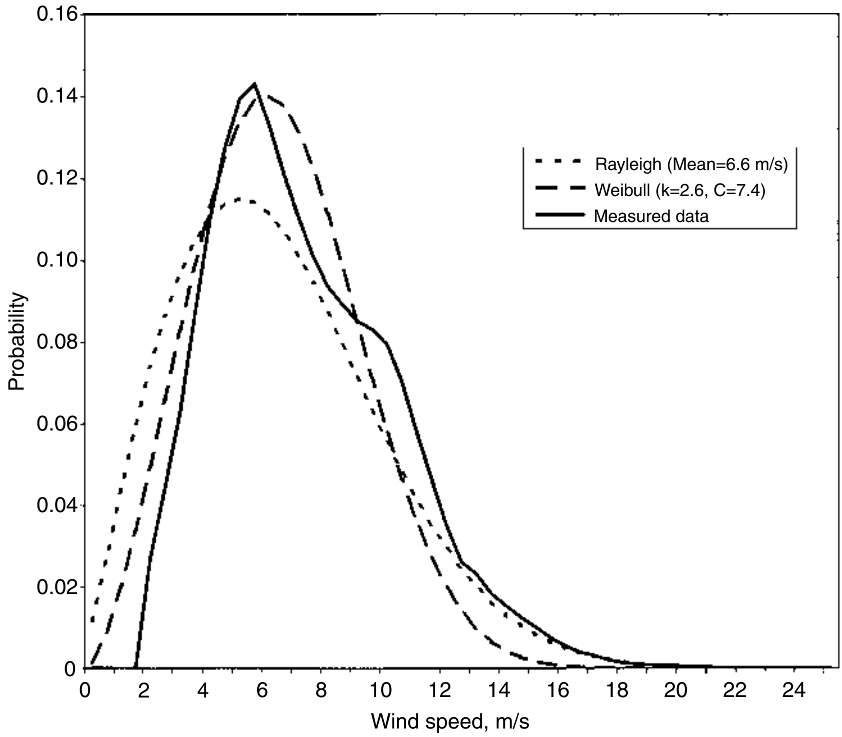


FIGURE 19.29 Measured and analytical wind speed distributions for Amarillo, Texas airport.

If an analytical pdf (such as the Rayleigh or Weibull) is used, then

$$\mathrm{WER} = \frac{1}{2} A \int \rho f(U) U^3 \mathrm{d}U. \tag{19.25}$$

The wind energy potential (*WEP*) or gross annual wind energy production for a specific site and a specific wind turbine can be calculated with a wind speed distribution and the turbine power curve (the electrical power generated by the turbine at each wind speed), properly adjusted for the local air density:

$$\mathrm{WEP} = 8760 \sum_{i=1}^{n} f(U_i) \Delta U_i P(U_i), \tag{19.26}$$

where 8760 is the number of hours in a year and $P(U_i)$ is the electrical power produced by the turbine at wind speed U (the center of the range ΔU_i). With an analytical pdf (the Rayleigh or Weibull), this becomes:

$$\mathrm{WEP} = 8760 \int f(U) P(U)\, \mathrm{d}U. \tag{19.27}$$

It is quite obvious that using an analytical pdf that is not a good match to the actual distribution at a site (such as the Weibull and Rayleigh distributions shown in Figure 19.28) will yield values for *WER* and *WEP* that are not good estimates of the actual values for that site. To obtain good estimates for these quantities, the actual wind speed distribution must be used.

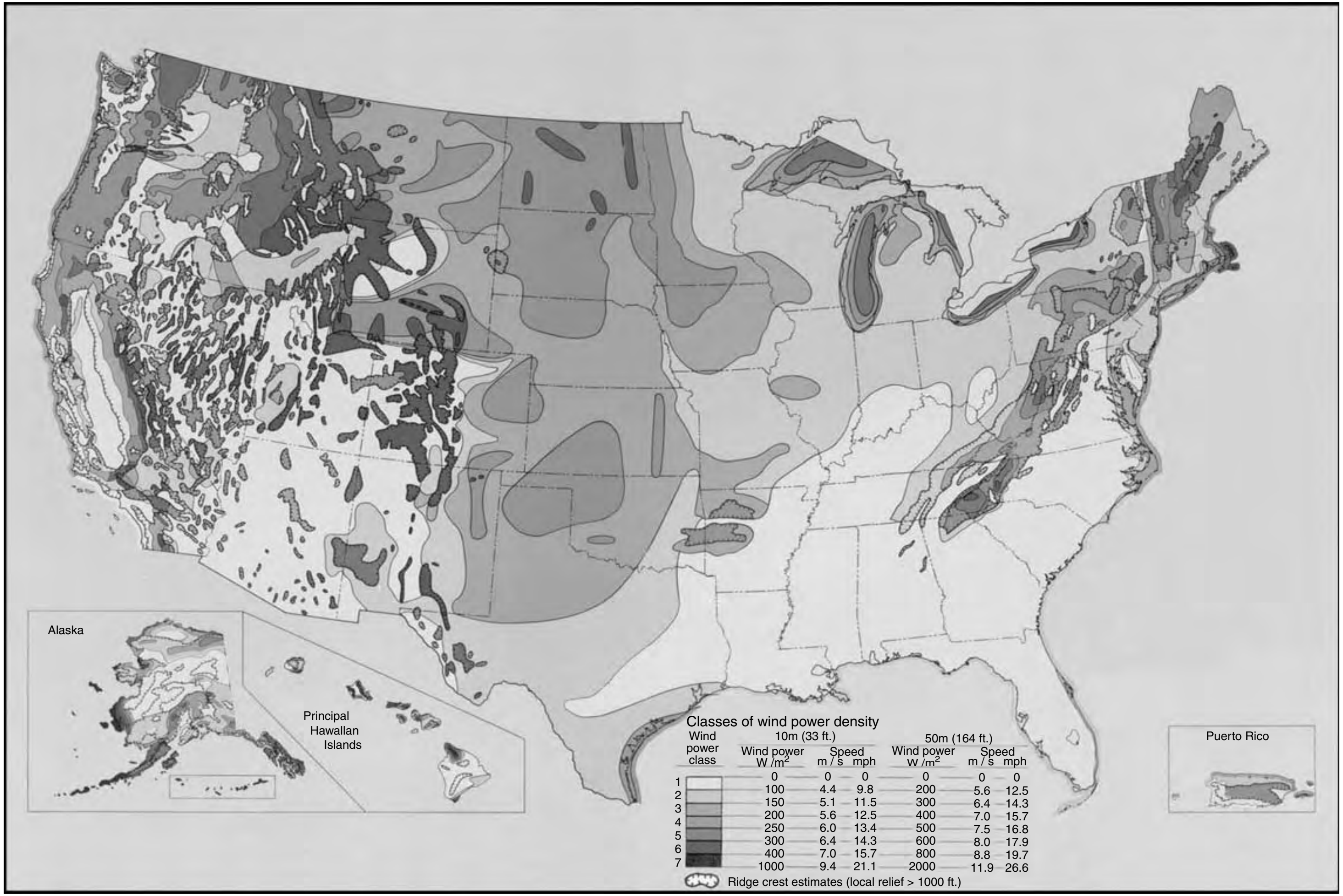


FIGURE 19.30 United States wind energy resources. (From Elliott, D. L., Holladay, C. G., Barchet, W. R., Foote, H. P. and Sandusky, W. F. Wind. 1987. Energy resource atlas of the United States, DOE/CH10094-4, Solar Energy Research Institute, Golden, CO. With permission.)

In 1987, Elliott et al. (1987) at Batelle Pacific Northwest Laboratory (PNL) in the United States carefully analyzed and interpreted the available long-term wind data for the United States and summarized their estimate of the wind energy resources in the *Wind Energy Resource Atlas of the United States*. Their summary for the entire United States is reproduced in Figure 19.30. The results are presented in terms of wind power classes based on the annual average power available per square meter of intercepted area (see the legend on Figure 19.30). Sites of wind power class 3 or higher (at least 150 W/m^2 at 10 m height or 300 W/m^2 at 50 m height) are usually considered economic for utility-scale wind power development with available wind technology. Sites of wind power class 2 or lower (less than 150 W/m^2 at 10 m height or 300 W/m^2 at 50 m height) are usually considered economic only for remote or hybrid wind power systems. Troen and Petersen (1989) at Denmark's Risø National Laboratory produced a European wind atlas that used somewhat different techniques to estimate the wind resources of the European Community countries. This work summarizes the resource available at a 50 m height for five different topographic conditions.

The estimates shown in Figure 19.30 and those presented in Troen and Peterson (1989) are quite crude and have been superseded in recent years by much higher-resolution maps, made possible by improvements in wind resource computer modeling programs and increases in computer speed. Many countries around the world have recently embarked on high-resolution mapping efforts to accurately quantify their wind resources and identify those areas of highest resource. The resultant resource maps are frequently available to the public, but in some cases a payment is required to obtain them. High-resolution wind resource maps of the individual states in the United States may be found on the Web at http://www.eere.energy.gov/windandhydro/windpoweringamerica/wind_maps.asp. Some states have developed even more detailed maps in recent years, and these may be found on Web sites for the individual states. Similar maps for some other countries may be found at http://www.rsvp.nrel.gov/wind_resources.html, and information on where to find maps and/or data for other countries may be found at http://www.windatlas.dk/index.htm. These newer wind resource evaluations frequently reveal far greater wind resource than earlier evaluations. In some cases, this is due to the higher resolution now available, but in other cases this is because the early evaluations sometimes used unverified long-term data from existing weather stations where the anemometers had not been properly maintained—the bearings had deteriorated over time and the anemometers registered lower winds than what actually existed. The verification procedures used in the newer evaluations are designed to help identify and eliminate this type of biased data. But even the highest resolution resource estimates are just that—estimates. The actual wind resources in any specific area can vary dramatically from those estimates and should be determined with long-term, site-specific measurements.

Archer and Jacobson (2005) from Stanford University have recently estimated the average global wind power potential for all land-based locations with wind class 3 or better (mean annual wind speeds in excess of 6.9 m/s at 80 m height) to be on the order of 72 TW (one $TW = 10^{12}$ W) for the year 2000, yielding an average wind energy potential of 627,000 TWh/y. In reality, the use of approximately 90% of the land mass is restricted, leaving only about 10% available for wind power generation. Thus, a realistic wind energy potential from the Stanford study is on the order of 62,700 TWh/y. This figure is consistent with the recently published European Wind Energy Association (2003) figure of 53,000 TWh/y of wind energy potential, and corresponds to over three times the world's total electricity needs in 2001 of 14,000–16,000 TWh/y. Many countries have adequate wind resources to supply their entire energy consumption. Of course, this resource is not necessarily available at the right time or the right place, and actually capturing wind energy on this vast a scale is not apt to happen.

The wind resource over the water-covered portion of the earth's surface is many, many times greater than that over the land, but current wind energy technology is restricted to shallow water (water depths on the order of 30 m or less). The available wind resource over water restricted to these depths is insignificant in comparison with that over the land mass.

19.2.2.4 Environmental Concerns

The fact that a site has great wind resources and access to transmission lines with excess capacity does not mean that it is suitable for wind power development. Potential environmental and social issues must be considered and resolved before a wind facility can be built. In the best possible scenario, such issues will add expense and development time to a project. In the worst scenario, such issues could cause cancellation of a project. The greatest environmental issues that the wind industry has had to face are the issues of bird and bat deaths due to collisions with wind turbines. The bird collision issue has been mostly resolved—unless there is a high level of raptor activity in the area, bird collisions can be addressed by following industry siting guidelines. If large numbers of raptors are present, additional restrictions on turbine placement, style, height, etc. may be required by local authorities to minimize the potential for collisions. At the time of this writing, the bat problem is being actively studied, but not much is known about why bats in a small number of locations are colliding with wind turbines, whereas bats in other locations do not seem to have the problem. Until this problem is better understood, the prudent course of action would be to postpone development of promising sites that have significant numbers of bats nearby. At the very least, the latest information from the appropriate national wind energy association regarding bat collisions with wind turbines should be obtained and carefully studied. Other potential environmental issues include the presence of endangered or protected species, nearby residences or airports, nearby scenic areas, recreational use or other specific use restrictions of the candidate site, or religious significance of the site. Additional details on environmental concerns may be found in the "Wind Energy Conversion" chapter of this handbook (Chapter 22).

19.2.3 Wind Resource Assessment

Three things must be present before a commercial wind energy generation plant can be successfully developed: a site with a good wind resource, access to a transmission line with the capacity to accept the plant output, and a buyer to purchase the energy generated. Only the identification of sites with adequate wind resource will be considered here.

Many different approaches are available when investigating the wind resource in a given land area. The preferred approach will depend on the specific wind energy program objectives and on previous experience with wind resource assessment. In any case, the process normally consists of three basic scales or stages of assessment:

1. Prospecting or identifying general areas of high wind resource. In this stage, wind developers typically rely on existing wind resource information to narrow the search region to the most promising areas.
2. Area wind resource evaluation to determine the actual wind resource. Wind measurement programs are normally undertaken in areas under serious consideration for wind power development. Wind developers use these measurements to determine the actual quality of the resource, compare different candidate sites, estimate the performance of specific turbines, analyze the economic potential of wind development, and make preliminary determinations of turbine placements.
3. Micrositing, in which the siting of the individual wind turbines is optimized. By carefully studying the small-scale variability of the wind resource at a particular location, the developer can position the wind turbines to get maximum wind exposure with minimal interference from other turbines or obstructions and thus maximize the overall energy output. There are many computer codes available to perform this function. The level of sophistication of these codes varies widely. The better ones use computational fluid dynamics to model the wind farm. They utilize digital elevation data to define the terrain of the wind farm and site measurement data to provide calibration and wind input information. In some cases, professional wind energy development companies may be willing to perform this service gratis, in hopes of getting the rights to develop the site. Advertisements for micrositing programs can be found in most wind energy trade magazines. Further discussion of micrositing is beyond the scope of this contribution.

19.2.3.1 Prospecting

The goal of prospecting for wind energy development sites is to identify areas within a fairly large region such as a utility service area, a county, or even a multistate region, that are likely to have good wind resources and that are located near existing power transmission lines with the capacity to handle the power generated by the new development. A good initial step in this process is to obtain the highest resolution available wind resource maps of the entire area of interest (see "Available Resource", above, for links to some map sources) and use those to identify areas of high winds. Keep in mind that values on the wind resource maps are just estimates—some areas of good resource may not be identified, and some areas identified as having good resource may actually have very poor resources. Use these maps in conjunction with topographic maps that show the location of major transmission lines, roads, etc. to narrow the general search area. Use of geographic information system (GIS) technology will permit the overlaying of the resource information with topographic information and road and transmission line information, greatly facilitating the identification of promising sites. Even with all the information that can be gleaned from GIS systems and hard copy and online maps, field visits to prospective sites are very worthwhile because observation of biological indicators and topographic features often help identify the best wind resource areas.

19.2.3.1.1 Biological Indicators

Persistent winds can cause plant deformation; careful observation of these plant deformations can be used to compare candidate sites and, in at least some cases, to estimate the average wind speed. Figure 19.31 illustrates various levels of tree deformity, corresponding to increasing levels of wind from one prevailing direction. The Griggs–Putnam index, explained in Hewson et al. (1978), correlates these degrees of deformation to specific wind speeds for one particular type of tree. While the vegetation in an area of interest may not permit the estimation of actual wind speeds based upon the sketches in this figure, the relative wind speeds in the area might well be established by observing the amount of flagging or throwing that is present. It should be noted that although wind-flagged trees (i.e., trees with branches bent away from a prevailing wind) may indicate that the annual average wind speed is quite strong, trees that are not flagged do not necessarily indicate that the winds are light; those trees may be exposed to strong winds from several directions, with insufficient persistence in any one direction to cause flagging.

19.2.3.1.2 Effects of Topography

The effects of surrounding terrain on the wind speed at a specific site are discussed in various wind turbine siting handbooks including Troen and Petersen (1989); Hiester and Pennell (1981); Wegley et al. (1980), and Rohatgi and Nelson (1994); the following discussion borrows heavily from these sources. Numerous researchers emphasize that the influence of terrain features on the energy output from a turbine may be so great that the economics of the whole project may depend on the proper selection of the site. The ready availability of detailed GIS maps and mapping technology today has greatly simplified this aspect of prospecting.

Terrain Classification. The most basic classification divides terrain into flat and nonflat categories. In a strict sense, the earth's surface is never truly flat; there are always some irregularities such as forest, shelterbelts (or wind breaks), etc. and/or gentle slopes. However, according to Frost and Nowak (1979), the terrain can be considered flat (for the purpose of wind turbine siting) if it meets the following conditions:

- Elevation differences between the wind turbine site and the surrounding terrain are not greater than about 60 m anywhere in an 11.5 km diameter circle around the turbine site.
- No hill has an aspect ratio (height to width) greater than 1/50 within 4 km upwind and downwind of the site.
- The elevation difference within 4 km upwind is small compared to the rotor ground clearance.

Flat terrain is obviously the simplest type of terrain for siting a turbine—the wind speed at a given height is nearly the same over the entire area. In the absence of any obstructions to accelerate wind flow,

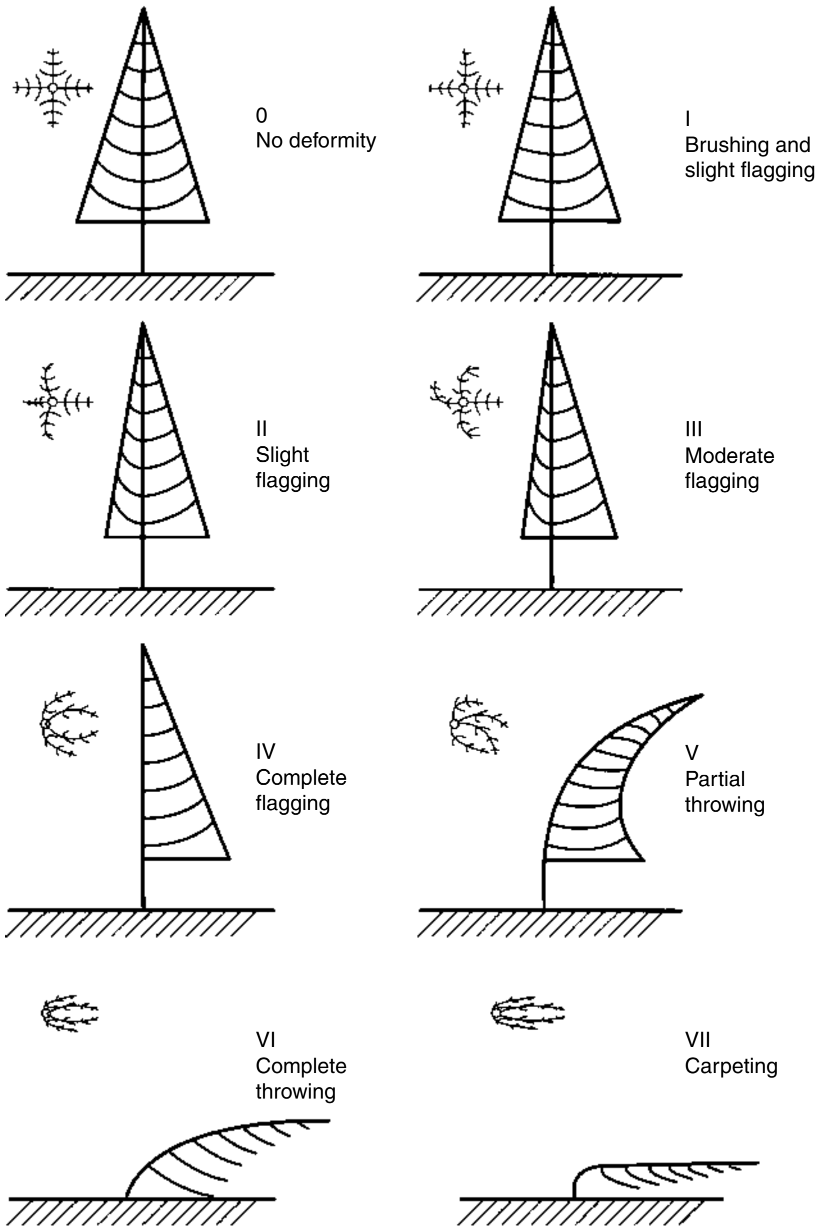


FIGURE 19.31 Griggs–Putnam index of tree deformation. (From Hewson, E. W., Wade, J. E. and Baker, R. W. 1978. *Vegetation as an Indicator of High Wind Velocity*, RLO/2227-T24-78-2, Oregon State University, Corvallis, OR.)

the best way to increase the available wind power is to raise the rotor higher above the ground to take advantage of positive wind shear. In most natural terrain, however, the surface of the earth is not uniform but changes significantly from location to location, affecting the local wind profile. Figure 19.32 illustrates the significant change in a vertical wind profile that results from wind flow going from a smooth to a rough surface—the shape of the profile shifts quite dramatically over a relatively short horizontal distance. In addition, most flat terrain has a variety of man-made (buildings, silos, etc.) and natural obstacles that affect the flow of wind. The effects of these obstacles have been studied extensively;

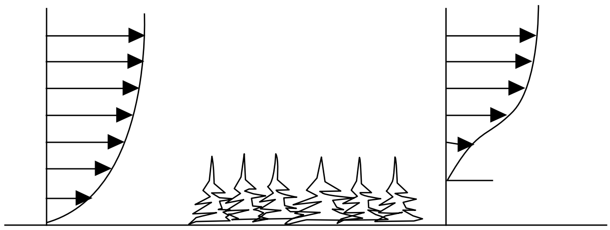


FIGURE 19.32 Effect of change in surface roughness from smooth to rough. (From Wegley, H. L., Ramsdell, J. V., Orgill, M. M. and Drake, R. L. 1980. *A Siting Handbook for Small Wind Energy Conversion Systems*, PNL-2521, Pacific Northwest Laboratory, Richland, WA.)

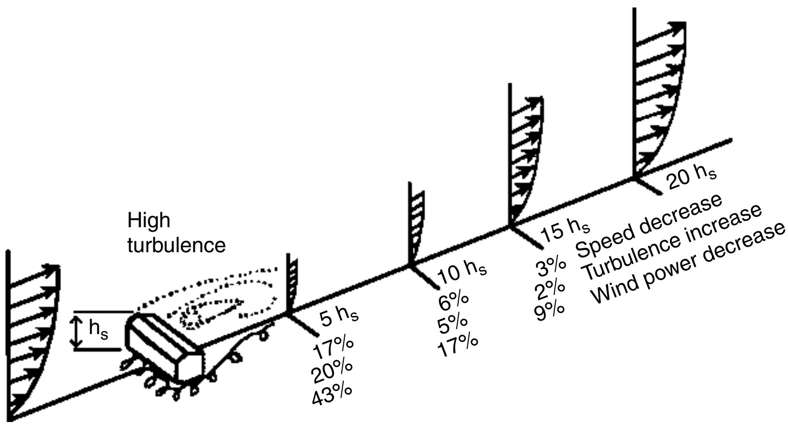


FIGURE 19.33 Wind speed, power, and turbulence effects downstream of a building. (From Wegley, H. L., Ramsdell, J. V., Orgill, M. M. and Drake, R. L. 1980. *A Siting Handbook for Small Wind Energy Conversion Systems*, Pacific Northwest Laboratory, Richland, WA.)

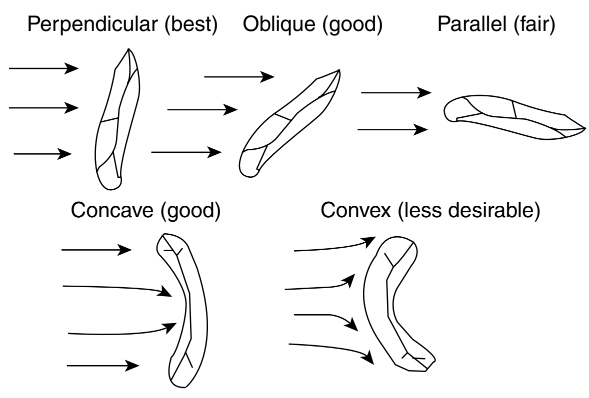


FIGURE 19.34 Effect of ridge orientation and shape on site suitability. (From Wegley, H. L., Ramsdell, J. V., Orgill, M. M. and Drake, R. L. 1980. *A Siting Handbook for Small Wind Energy Conversion Systems*, PNL-2521, Pacific Northwest Laboratory, Richland, WA.)

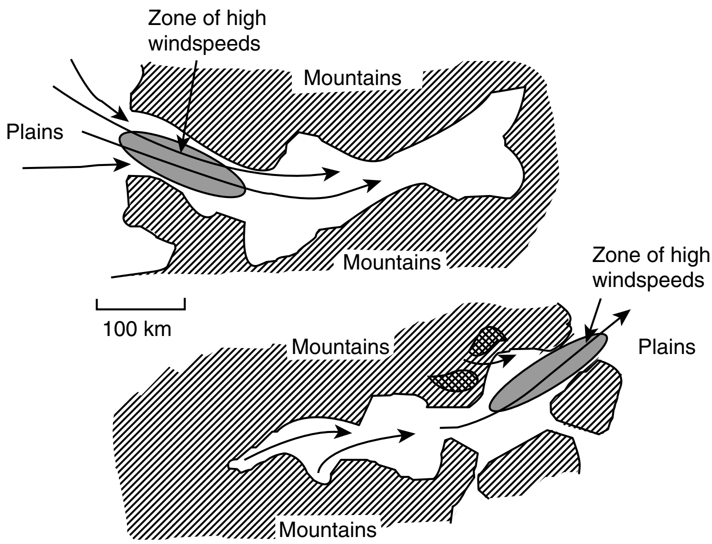


FIGURE 19.35 Increased wind speeds due to channeling of prevailing winds by mountains. (From Rohatgi, J. S. and Nelson, V. 1994. *Wind Characteristics: An Analysis for the Generation of Wind Power*, Alternative Energy Institute, West Texas A&M University, Canyon, TX. With Permission.)

an example of the effect of a building on wind speed, available power, and turbulence is shown in Figure 19.33. Note that the estimates in the figure apply at a height equal to one building height above the ground, and that power losses become small (less than 10%) after a downwind distance equal to 15 building heights.

Ridges are elongated hills with a length-to-height ratio of at least 10 that have little or no flat area on the summit. As illustrated in Figure 19.34, the ideal prevailing wind direction for wind turbine siting should be perpendicular to the ridge axis. When the prevailing wind is not perpendicular to the axis, the ridge will not be as attractive a site. Concavity in the windward direction enhances speedup and convexity reduces speedup by deflecting the wind flow around the ridge. The slope of a ridge is also an important parameter; steeper slopes give rise to stronger wind flow, but they also give rise to high turbulence in the lee of the ridge.

Depressions are characterized by terrain features lower than the surroundings and include valleys, canyons, basins, and passes. These can cause significant speedup of the wind if they effectively channel the wind. The factors that influence the flow in depressions, in addition to diurnal flow variations, include orientation of the wind in relation to the depression, atmospheric stability, the width, length, slope, and roughness of the depression, and the regularity of the section of valley or canyon. Canyons in mountainous terrain, such as those illustrated in Figure 19.35, can be very effective in creating high wind speeds.

Nonflat or complex terrain has large-scale elevations or depressions such as hills, ridges, valleys, and canyons. Keep in mind that information on wind direction should be considered when defining the terrain classification. For example, if an isolated hill (200 m high and 1000 m wide) were situated 1 km south of a proposed site, the site could be classified as complex. If, however, the wind blows only 5% of the time from this direction and the average speed from this direction is low, 2 m/s, for example, then this terrain would be classified as flat. Additional summaries of the effects on wind flow of many types of large scale features, including mountains, large cliffs and escarpments, valleys, and canyons (including slope winds, prevailing winds in alignment and prevailing wind in nonalignment), gaps and gorges, passes and saddles, and large basins, are given in Rohatgi and Nelson (1994).

19.2.3.2 Wind Resource Evaluation

Once general areas of high wind resources have been identified, the actual resource available at one or more of those sites can be determined by performing a full wind resource evaluation or site assessment. This is a costly, complex, and time-consuming activity. Numerous companies perform this service; a good method for locating some is to access the appropriate national wind energy association membership list (typically from their Web site) and search for "meteorology consultants." Even if the service is contracted out, a cursory knowledge of the process is important to ensure that the contractor does a good job. Several handbooks detailing the steps for conducting successful measurement programs are available. The following discussion borrows heavily from AWS Scientific, Inc. (1997).

Long-term variability at a site is nearly as important as mean wind speed, as far as long-term energy production is concerned. Meteorologists generally agree that a well-done evaluation/assessment project will take a minimum of one year to complete (Wegley et al. 1980). A project duration of two or more years will produce more reliable results and is recommended by Schwartz and Elliott (2005) and many other experts. One year is usually sufficient to determine the diurnal and seasonal variability of the wind with an accuracy of 10% at a confidence level of 90%, according to Aspliden et al. (1986) and Corotis (1977). In some cases, high-quality data may be available from a nearby representative site and this might be used to shorten the duration of the assessment project and estimate the interannual variability of the wind. In particular, long-term data from the nearest airport or weather recording station can help determine whether the data obtained at a site is representative of normal winds for the site or whether it is representative of higher or lower than average winds. Wegley et al. (1980) and Gipe (1993) give suggestions on methods of using available data from nearby sites to estimate site wind speed with minimal onsite data.

A single site assessment with a very basic single 50-m monitoring station operated for two years can be expected to run in excess of \$20,000; a 60-m monitoring station will run about \$2500 more. Multiple sites will see some economies of scale, depending on how close the sites are to each other. The total cost to operate a second site can be expected to be 10%–5% less than the cost for the first site. Five sites in fairly close proximity can probably be operated for an average site cost about 25% less than for a single isolated site. Most savings will be realized in the labor and travel categories, although equipment cost savings may be realized as a result of quantity discounts, by sharing installation equipment, and by sharing parts inventories. Conducting an abbreviated resource assessment program to minimize expenses may well turn out to be a case of "penny wise, pound foolish" in that the return on investment of the entire wind farm project will be directly affected by the quality of the actual wind resource at each turbine site. Without a proper site assessment, the resource estimates may not be a good reflection of the actual resource and the production of the wind project may be much different than projected.

For small wind turbines, the expense of anemometers, data logger, and data analysis may be more than the price of the wind turbine. In this case, historical, regional data may be judged to be adequate for estimating the wind resource. Inexpensive digital weather stations, including rudimentary data loggers that work with a personal computer, are available for under \$1000, and one of these could be used in conjunction with the historical data to improve the accuracy of the resource assessment. Although these weather stations may be adequate for this particular application, they are not highly accurate and are not designed for long-term durability, so they are not suitable for collecting long-term data for accurate wind resource assessment as is needed for commercial enterprises. Wind speed, direction, distribution, and shear can vary significantly over fairly short distances in either the horizontal or vertical direction; therefore, to get the best possible estimate of the wind energy resource at a particular location, it is important to measure the wind resource at the specific site and height of interest. In complex terrain, this probably means the use of several monitoring stations to adequately assess a single site.

The keys to a successful assessment program are the early identification of the objectives of the program and the timely development of a detailed plan of action to ensure that the data needed to meet the objectives are acquired. Such a plan should include, at a minimum:

- Quality control measures, including a quality assurance program
- Site data to be measured
- The program duration, minimum measurement accuracy and target data recovery rate (these heavily impact the equipment type and quality)
- Number and tentative location of monitoring towers, together with sensor measurement heights
- Data sampling and recording intervals
- Data storage, handling, and processing procedures

The data recovery rate is defined as the percentage of possible data records that have actually been collected over a reporting period:

$$\text{Data recovery rate} = \frac{\text{Data records collected}}{\text{Data records possible}} * 100(\%),$$

where

$$\text{Data records collected} = \text{Data records possible} - \text{Number of invalid records.}$$

For example, the total possible number of 10-minute records in December is 4464. If 264 records were deemed invalid, the number of valid data records collected would be 4200 (4464—264). The data recovery rate for this example would be:

$$\text{Data recovery rate} = 4200/4464 * (100) = 94.1\%.$$

A data recovery rate of at least 90% (95% or better should be possible) for all measured parameters over the duration of the program, with any data gaps kept to a minimum (less than one week), should be a major goal of the measurement program.

A quality assurance program for flagging and handling suspect data is imperative to ensure the acquisition of high-quality data and the successful completion of the assessment program. The specifics of the components of this program should be determined and documented early in the assessment program and should include:

- Acquisition of equipment that meets accuracy and reliability specifications
- Equipment calibration methods and frequency of calibration
- Installation and maintenance instructions for equipment
- Data validation methods, including specifics on evaluating, removing and/or replacing suspect data, and reporting all of these actions
- Data analysis instructions, including specific calculations to be performed

Quality assurance will also help to minimize the uncertainties that are always inherent in the data. If the assessment process is carefully followed, these uncertainties can be characterized and controlled to maximize the usefulness of the assessment program conclusions.

19.2.3.2.1 Data Measurement

The core of the monitoring program is the collection of wind speed, wind direction, and air temperature data. Keep in mind that it is generally less expensive to provide and monitor extra sensors than to conduct an unscheduled site visit to replace or repair a failed sensor that is the sole source of an essential measurement. Care must be used in mounting the various sensors to minimize any interference of one sensor on another. It is especially important to avoid any interference with the wind measurements.

If the rotor of the turbine is to be relatively small, wind speed measurements at the turbine hub height may suffice. For larger rotors, acquiring wind speed data at multiple measurement heights is necessary for determining the site vertical wind shear characteristics, for conducting turbine performance simulations at several turbine hub heights, and to assist in data validation (multiple anemometers make spotting a bad one fairly easy). Redundant anemometers are sometimes used to minimize the risk of wind speed data loss due to a failed primary anemometer and to provide substitution data when the primary anemometer is shadowed by the tower (when it is directly downwind of the tower). Typical anemometer heights are every 10 m, starting at 20 or 30 m and going up to maximum tower height (50–60 m for a tilt-up tower). The vertical distances between anemometers mounted at different heights are referred to as *height layers*. For example, for anemometers installed at heights of 30, 40, and 50 m, height layers would be 30–50, 30–40, and 40–50 m.

Cup or propeller anemometers are the sensor types most commonly used for the measurement of near-horizontal wind speed.

- Cup anemometer: This instrument consists of a cup assembly (three or four cups) centrally connected to a vertical shaft for rotation. At least one cup always faces the oncoming wind. The cups convert wind pressure force to rotational torque and the transducer in the anemometer produces an electrical signal that is proportional to wind speed.
- Propeller anemometer: This instrument consists of a propeller mounted on a horizontal shaft that is oriented into the wind through the use of a tail vane. The propeller anemometer also generates an electrical signal proportional to wind speed.

Although the two sensor types differ somewhat in their responsiveness to wind speed fluctuations, there is no clear advantage of one type over the other. In practice, the cup type is most commonly used for resource assessment. When selecting an anemometer model, the following should be considered:

- Intended application: Anemometers intended for low wind speed applications, such as air pollution studies, are usually made from lightweight materials. These are probably not suited for very windy or icy environments.
- Survival wind speed: Be sure the anemometer is capable of withstanding the maximum wind speed that it is likely to see. An anemometer with a survival wind speed of 25 m/s is not apt to survive in most windy sites. A survival speed of 50 m/s should be adequate for most sites.
- Starting threshold: This is the minimum wind speed at which the anemometer starts and maintains rotation. For wind resource assessment purposes, it is more important for the anemometer to survive a 25 m/s wind gust than to be responsive to winds under 1 m/s.
- Distance constant: This is the distance the air travels past the anemometer during the time it takes the cups or propeller to reach 63% of the equilibrium speed after a step change in wind speed (the "response time" of the anemometer to a change in wind speed). Longer distance constants are usually associated with heavier anemometers; inertia causes them to take longer to slow down when the wind decreases. These instruments may overestimate the wind speed.
- Reliability and maintenance: Wind sensors are mechanical and eventually wear out, although most have special, long-life bearings that will normally last for at least two years. Be certain to obtain units with bearings that will last for the entire duration of the measurement project.

Anemometers are subject to a variety of errors in the determination of true wind speed, and equations which may be used to estimate the size of these errors are given in Justus (1978). When anemometers are calibrated in steady air flows in a wind tunnel, they may measure the true wind within $\pm 1\%$. In gusty winds, however, anemometers generally speed up faster than they slow down, so that accuracies of $\pm 5\%$

may be more realistic in application. Additional information on anemometers may be found in AWEA (1986) and ASME (1989).

Wind direction vanes should be installed at all significant monitoring levels. Wind direction information is important for identifying preferred terrain shapes and orientations and for optimizing the layout of wind turbines within a wind farm. The most familiar type of vane uses a fin connected to a vertical shaft. The vane constantly seeks a position of force equilibrium by aligning itself into the wind, and produces an electrical signal proportional to the position of the vane relative to some reference direction (usually selected as true—not magnetic—north). Some wind vanes exhibit "dead bands," a narrow section of the rotation where the sensor transitions from the full rotation reading of nearly 360° to the initial reading of 0°. The output of the sensor in this section is usually unpredictable. The position of this dead band should be carefully noted for reference when the vane is mounted to the tower. Newer direction vanes have eliminated this dead band.

Air temperature is used to calculate air density, a quantity required to estimate the wind power density and the power that a wind turbine will generate. It is normally measured either near ground level (2–3 m), or near hub height. In most locations, the average near-ground-level air temperature will be within 1°C of the average temperature at hub height. Ambient air temperature sensors are readily available. The temperature sensor must be protected from direct solar radiation by mounting it within a radiation shield.

Once the basic measurement system is installed, additional resource-related parameters can be acquired at minimal additional cost. The most common additional parameters are vertical wind speed, change in temperature with height (commonly referred to as ΔT), barometric pressure, and solar radiation.

The vertical wind speed provides more detail about site turbulence and can be a good predictor of wind turbine loads. Historically, this parameter has been a research measurement, but as wind energy development spreads into new regions of the country, regional information on vertical wind velocity may become important. The propeller anemometer is especially well suited for measuring the vertical wind component. For this application, the rotation axis would be mounted vertically. The polarity of the DC output signal indicates rotational direction and the signal magnitude indicates actual vertical speed. The vertical wind speed anemometer should be located near the upper basic wind speed monitoring level.

ΔT provides information about turbulence and atmospheric stability. This is measured with a matched set of temperature sensors located near the lower and upper measurement levels. The existing air temperature sensor may be matched with an identical sensor and used to measure ΔT, or a separate pair of matched sensors may be used. Sensors for this application are usually tested over a specified range and matched by the manufacturer. Be certain to use identical equipment (e.g., radiation shield, mounting hardware, etc.) with both sensors so the inherent errors in the signals will cancel out when the difference

TABLE 19.2 Typical Specifications for Sensors

Specification	Wind Speed	Wind Direction	Air Temperature	Vertical Wind Speed	ΔT	Atmospheric Pressure	Solar Radiation
Measurement range	0–50 m/s	0°–360°	−40°–60°C	0–50 m/s	−40°–60°C	94–106 kPa (sea level equivalent)	0–1500 W/m^2
Starting threshold	≤1.0 m/s	≤1.0 m/s	N/A	≤1.0 m/s	N/A	N/A	N/A
Distance constant	≤4.0 m/s	N/A	N/A	≤4.0 m/s	N/A	N/A	N/A
Allowable sensor error	≤3%	≤5°	≤1°C	≤3%	≤1°C	≤1 kPa	≤5%
Sensor resolution	≤0.1 m/s	≤1°	≤0.1°C	≤0.1 m/s	≤0.1°C	≤0.2 kPa	≤1 W/m^2

Note: All sensors should have an operating temperature range of −40°C–60°C and an operating humidity range of 0% to 100%.

Source: From AWS Scientific, Inc. 1997. Wind resource assessment handbook: Fundamentals for conducting a successful monitoring program, SR-440-22223, National Renewable Energy Laboratory, Golden, CO.

between the two values is taken. Radiation shields that use either forced (mechanical) or natural (passive) aspiration are normally used on both sensors to reduce the radiation-induced errors.

Barometric pressure is used with air temperature to determine air density. Because it does not vary much over relatively short distances, many resource assessment programs do not measure barometric pressure; they use elevation-adjusted data taken at a nearby regional National Weather Service station. Several atmospheric (barometric) pressure sensors that are suitable for this application are commercially available. Be certain to select one that will give accurate readings in a windy environment.

Solar radiation, when used in conjunction with wind speed and time of day, can be an indicator of atmospheric stability and is often used in numerical wind flow modeling. These measurements may also be useful for later solar energy evaluation studies. A pyranometer is used to measure global (or total) solar radiation, the combination of direct sunlight and diffuse sky radiation. Remember that the pyranometer must be in a position where it will never be shaded in order to measure accurately. The recommended measurement height is 3–4 m above ground.

Table 19.2 lists nominal specifications for the most common types of sensors. Some sensors require the use of separate signal conditioners, electronic packages that supply power to the sensor and process the signal received from the sensor to convert it into a form that can be used by the data logger, the device which actually acquires the raw data and calculates and saves the average statistics. Data loggers (or data recorders) come in a variety of types, sizes, and capabilities; most include peripheral storage and data-transfer devices. Be certain that the data logger is compatible with the sensor types, number of sensors, and desired sampling and recording intervals. The data logger should also:

- Be capable of recording the time and date corresponding to each data record with that data record
- Contribute negligible errors to the signals received from the sensors
- Operate over the temperature range of −40°C–60°C and over a relative humidity range of 0%–100%
- Offer retrievable data storage media
- Operate on battery power (with an AC adaptor to permit use of AC power when it is available)

The amount of data logger storage capacity that is needed depends on the averaging interval, the number of active data channels, the need for calculating derived quantities such as wind shear exponent and turbulence intensity, and the maximum timespan between data retrievals, including potential delays. Manufacturers usually provide tables or methods to calculate the approximate available storage capacity in days for various memory configurations. To be safe, obtain enough storage capacity to store an additional week of data, in case a data retrieval or site access problem develops.

While all data loggers allow manual retrieval of stored data, many either come equipped with or can easily be equipped with communications equipment to enable remote retrieval of data, typically via cell phone or satellite phone link. The manual method promotes frequent visual onsite inspection of the equipment during the site visits required to retrieve the data, which may be beneficial. However, this method also requires frequent expensive site visits and additional data handling steps (thus increasing potential data loss).

Remote data retrieval permits more frequent data retrieval and inspection than is feasible with manual data retrieval. This allows for prompt identification and resolution of site problems and enhances the data recovery rate. The disadvantages include the upfront cost of the required additional equipment the cost of monthly service, and the risk of communication system problems. The additional costs are usually quickly offset by the savings from not having to make site visits; choosing good quality equipment helps minimize the communication problem risk.

Data loggers with cell phone capability are extremely popular due to their ease of use and reasonable cost. A major concern with these units is whether the cell phone coverage at the measurement site will be adequate to support a solid communication link. Keep in mind that the data logger cell phone is much more powerful than a typical cell phone—it will be attached to a high-gain antenna to permit communication at a much greater range. A general rule of thumb is that a good connection is usually possible as long as there is a cell phone tower in line of sight within about 120 km. Replacement of the standard antenna with one with higher gain may be an option if the nearest cell tower is further away. Additional information regarding cell phone links and accounts should be available from the data logger supplier or manufacturer.

Satellite telephone service may all that is available in really remote locations, but that service is usually quite expensive—the transceiver is much more costly and the monthly fees are usually much higher than is the case for cell phone service. In addition, utilizing a satellite communication link may require the development of specialized data transfer software for the specific application. Again, additional information should be available from the data logger supplier or manufacturer.

If phone service (either cellular or satellite) is either unavailable at the site or is prohibitively expensive, try to arrange for manual data retrieval by someone who lives near the monitoring site. That individual may then be able to send the data media by mail, or transmit the data via e-mail or landline phone transmission. If this means of data retrieval is utilized, be sure to incorporate additional procedures to protect against data loss in the event that the data media gets damaged or lost in the mail or the transmitted data becomes corrupted.

The overall accuracy of any system is determined by its weakest link, or least accurate component. It is also influenced by its complexity, the total number of components or links. The measurement of any parameter (wind speed, wind direction, etc.) requires that several components (sensor, signal conditioner, cabling, and data logger), each potentially contributing an error to the measured parameter, be interconnected. The combination of these errors will determine the system error (the difference between the measurement result and the actual value sensed) for that parameter. Errors contributed by the sensors represent the main concern, because those associated with the electronic subsystem (data logger, signal conditioner, and associated wiring and connectors) are typically negligible (less than 0.1%). Using wind speed as an example, the allowable system error in the measured wind speed might be specified as less than or equal to 3% of the true wind speed value, allowing for a 6% error window ($\pm$3%) centered about the true wind speed. This means that the wind speed sensor must measure the wind with an error of 3% or less.

The resolution of a data measurement is the smallest change in a measured quantity that can be detected by the measuring system. Again, this is primarily a function of the sensor, as data loggers have far higher resolution than is normally required. For example, an anemometer system with a 10-bit data logger analog-to-digital converter or digitizer (a common resolution for a data logger) that is set up to measure a maximum wind speed of 50 m/s has a data logger resolution of $50 / (2^{10}) = 0.05$ m/s. Some data loggers may have higher-precision digitizers (12-bit or 16-bit are common) with a resolution of much less than this (0.01 or 0.001 m/s, respectively). Thus, to achieve the resolution of 0.1 m/s specified for anemometers in Table 19.2, it is necessary for the anemometers themselves to have a resolution of 0.1 m/s or better.

System reliability is the measure of how well an instrumentation system will consistently provide valid data for a measured parameter over its measurement range, i.e., how well it performs in the long run. In selecting instrumentation, it is important to identify and select components that are designed to reliably measure the selected parameters at the prescribed heights for the full monitoring duration and at the required levels of data recovery and accuracy. The instrumentation must also be capable of withstanding the environment of the specific location (e.g., weather extremes, dust, salt, etc.) and be tailored to the selected mode of data retrieval (manually or via cell phone or satellite phone communication link). Although vendors often provide reliability information in terms of a mean time between failures under certain conditions, the best indication of a product's reliability is its performance history. Ask the vendor for a few references and check with those references to determine their satisfaction or lack of satisfaction with the product of interest. Comprehensive quality assurance procedures and the use of redundant sensors are two of the best ways that the user can maximize system reliability. However, there is little that can be done to improve the reliability of sensors or data loggers that are prone to failure.

The equipment should also be proven, affordable, and user-friendly. Complete monitoring systems can be purchased from a single vendor or components from different vendors can be combined. If components from different vendors are used, be sure the individual components are compatible with each other.

Lists of wind resource assessment equipment vendors may be found in the member lists of the various national wind associations. At the American Wind Energy Association (AWEA) Web site (http://www.awea.org), for instance, they may be found by accessing the member directory, selecting "search," "consultants," and then "meteorology."

19.2.3.2.2 Sampling Rates and Statistical Quantities

All data sensors should be sampled once every one or two seconds. The resultant data are not typically recorded, but stored into data logger accumulators for the specified averaging period (10 minutes is the default international averaging period for wind measurements). At the end of the averaging period, the statistics are calculated and stored, the accumulators are cleared, and the storage of data into the accumulators begins anew. The data logger should contain built-in programming to calculate and store, as a minimum, the following statistics:

- The average, standard deviation, and maximum and minimum values for the wind speed at each anemometer level, together with the wind directions associated with each maximum and minimum value
- The average and standard deviation for the wind direction at each level
- The average, standard deviation, and maximum and minimum wind speed difference for each height layer, together with the wind direction associated with each maximum and minimum difference
- The air temperature
- The vertical wind speed (if measured)
- The ΔT (if measured)
- The barometric pressure (if measured)
- The solar radiation (if measured)

These statistics, together with a corresponding time and date stamp, constitute the data to be recorded; as mentioned above, the individual data samples are not normally saved.

Except for wind direction, the average is defined as the numeric mean of all samples. For wind direction, the average is defined as the mean direction on a 0°–360° scale. The standard deviation is defined as the true population standard deviation for all samples within the averaging interval.

19.2.3.2.3 Lightning Protection Devices

One should keep in mind that a single lightning strike can destroy the entire site monitoring system and all associated electronics. Repairing lightning damage will require a large investment in both time and replacement parts. Although no amount of protection can ensure that lightning will not strike the system, appropriate protection can minimize that risk. Consult with the data logger and instrumentation suppliers to ensure that adequate lightning protection is incorporated in the data logger and in all sensors, signal conditioners, and power supplies that will be used. If the supplied lightning protection is not adequate for the application, determine what additional protection is needed and add it before installing the monitoring system.

19.2.3.2.4 Tower Sensor Mounting

Towers for mounting sensors are available in either tubular or lattice types. Both types are available in tilt-up, telescoping, and fixed versions. For new sites, the tubular, tilt-up, guyed tower, which makes possible the assembly of the tower and the mounting of the sensors on the ground, is an excellent and convenient

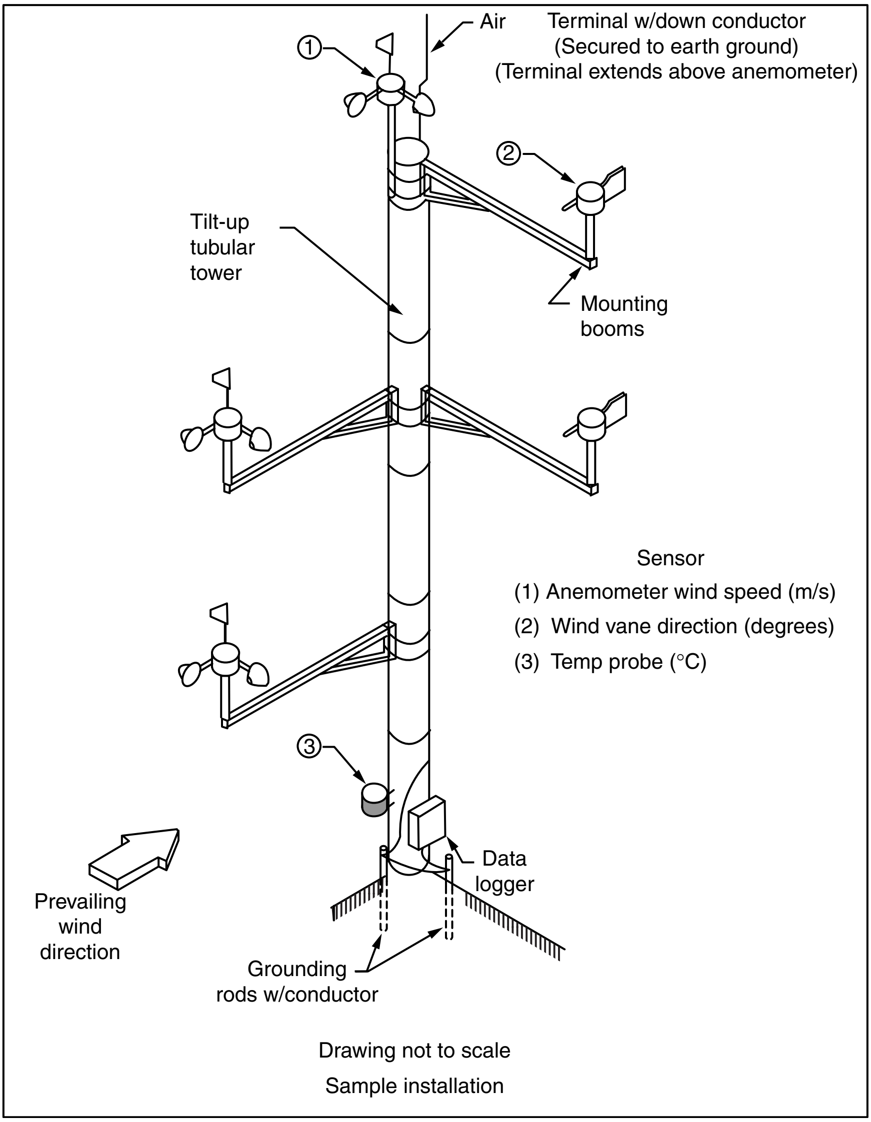


FIGURE 19.36 Typical instrumentation placement on meteorological tower.

choice. It also requires very little ground preparation, and is relatively low cost. The current maximum height of available tilt-up towers is about 60 m; if the tower is to be used in icing conditions, the maximum height is reduced to 50 m. Keep in mind that raising a tower of this size is not a simple task, and neither is replacing tower-mounted sensors; they both require experienced personnel and proper equipment.

The sensor support hardware includes the masts (vertical extensions) and mounting booms (horizontal extensions). Tubing, not solid stock, masts and booms should be used. These are used to position the sensors away from the support tower so as to minimize any influence of the tower, mounting hardware, and other equipment and sensors on each of the measured parameters. This can be achieved by consulting specific manufacturers' instructions, and referring to the sample installation configuration shown in Figure 19.36. Detailed information may be found in the AWEA (1986) procedure and virtually any other wind measurement handbook.

Refer to the manufacturer's instructions for the proper sensor and data logger wiring configurations. Wiring to connect the sensors, signal conditioners, and data logger should be shielded and/or twisted pair cable, whenever possible, to prevent ambient electrical noise from affecting the accuracy of the measurements. Be sure to use insulation and conductor types that are flexible over the full temperature range expected at the site, and use wire with UV-resistant insulation. Rodents and raptors, in particular, seem to have an affinity for wire insulation, so consider using armored cable or protective conduit in areas that are accessible by them. Try to mount the data logger and communications equipment in a substantial, locked container. Mount the communications antenna high enough to discourage vandalism, but where it can still be easily accessed by service personnel. Vandals are a concern, even in very remote locations.

Seal all sensor terminal connections with silicone caulking and protect them from direct exposure to sunlight and water with rubber boots or electrical tape. Wrap the sensor cabling along the length of the support arm and tower and secure it with UV-resistant wire ties or electrical tape. Wrap tape around the sensor wire and leave sufficient slack in the wire wherever chafing can occur between the sensor wire(s) and the support structure (e.g., tilt-up tower anchor collars).

19.2.3.2.5 Data Collection and Handling

All of these efforts devoted to selecting and mounting sensors and ensuring that acquired data are free of interference are for naught if the resultant data is subsequently lost or contaminated. The data collection and handling elements of the monitoring system must incorporate procedures that offer a high level of data protection. In general, the procedures should comply with those specified by the data logger manufacturer and reflect good common sense. A few pertinent comments are listed below.

Data Retrieval Frequency. A key factor in achieving a high level of data recovery is the ability to identify potential data acquisition problems and to quickly determine if a problem exists (a failed sensor, icing, possible loss of electrical ground, etc.) and, if necessary, initiate the steps required to fix the problem. Data transfer and review are the first-order means of achieving this end. A schedule of regular site data transfers or downloads should be developed and maintained. The maximum recommended manual download interval is every two weeks. For remote data transfer systems, a weekly retrieval rate may suffice, but a shorter interval, such as every other day, may be preferable to minimize potential loss of data if a problem arises and to efficiently transfer the large amount of data resulting from 10-minute data averaging. Situations may arise that warrant additional data transfers. For example, sensor data irregularities may become apparent during the review of site data, or the site may experience severe weather, such as icing or thunderstorms. Either of these situations merits a prompt follow-up data transfer (either manual or remote) and review.

Any time a site visit is made, the first order of business should be retrieval of the raw data from the data logger, either by manually downloading it to an infield laptop computer or by transferring it over a telephone link to a central site computer. This will minimize the risk of potential data loss from operator error, static discharges, or electrical surges during handling and/or checking of system components. The last order of business before leaving the site should be a verification that the monitoring system is functioning properly.

Data Protection and Storage. Sensor data that have not been subjected to a validation or verification process are commonly referred to as *raw data*. There is a constant risk of raw data loss or alteration during any measurement program. Aside from the data logger programming requirements, the actual data collection process requires minimal human intervention and data is adequately protected by following recommended installation procedures, including grounding all equipment, as recommended. This field data will eventually be transferred to a personal computer for analysis. Although this will be the primary location of the working database, it should not be the storage area for the archived or raw database, as frequent usage of a computer increases the likelihood of electrical surges, static discharges and other events that may damage hard drives and destroy any databases. Preserve the original raw data; make at least two copies of that dataset on removable media and store the original and all but one of the copies in separate locations (not in the same building). Then apply the validation and processing steps to the remaining copy. Back up this active database on a regular schedule during the validation process. Once the database is fully validated, create multiple copies of it and again store each copy in a separate location (not in the same building).

Improper data handling procedures represent a high risk for data loss. The data reduction and analysis staff will be handling the data medium and be in constant contact with significant numbers of raw and processed databases. Ensure that all personnel are fully trained and understand the data retrieval software and computer operating system, are well aware of all instances in which data can be accidentally overwritten or erased, and that they employ good handling practices of all data storage media.

Data Validation. After the field data are collected, transferred to an office computing environment, and appropriate copies are made, the next steps are to validate and process that data. Again, these steps should be performed on a copy of the database, not on the original data. Data validation is the inspection of all the collected data for completeness and reasonableness, and the elimination of erroneous values. This step transforms raw data into validated data, and is crucial to maintaining high rates of data recovery during the course of the monitoring program. There are many possible causes of erroneous data; faulty or damaged sensors, loose wire connections, broken wires, damaged mounting hardware, data logger malfunctions, static discharges, sensor calibration drift, and icing conditions are some of the contributors. The goal of data validation is to detect as many significant errors from as many causes as possible; catching all the subtle ones is impossible. For example, a disconnected wire can be easily detected by a long string of zero (or random) values, but a loose wire that becomes disconnected intermittently may only partly reduce the recorded value, yet produce data that appears reasonable. Therefore, slight deviations in the data can escape detection (although the use of redundant sensors can reduce this possibility). Properly exercising the other quality assurance components of the monitoring program will also reduce the chances of data problems.

Data should be validated as soon as possible after they are transferred from the site to the office; the sooner a potential measurement problem is spotted, the quicker it can be addressed, and the lower the risk of losing large amounts of data. Data can be validated either manually or automatically, with computer processing. Obviously, manual verification can be extremely tedious and time consuming, but it is a good practice to validate the initial data from a site in this manner, in order to learn the characteristics of the data and become familiar with the types of suspect data that can be expected. This knowledge then makes possible the tailoring of the computer routines to optimize the automated validation process. Validation software is available from several sources, including from data logger vendors, but it is still commonly home-grown and tailored for particular applications.

Data validation can be split into two distinct operations; data screening and data verification.

- Data screening:

  This operation uses a series of validation routines or algorithms to screen all the data to identify suspect or questionable values—values that deserve scrutiny but are not necessarily erroneous. For example, an unusually high hourly wind speed caused by a locally severe thunderstorm may appear on an otherwise average windy day. The result of this data screening is a report that lists the suspect values and which validation routine each of those suspect values failed. General system checks ensure that each data record contains the appropriate number of data fields and that records are contiguous in time (i.e., time and date stamps are in order and none are missing). Measured data checks, on the other hand, ensure that the actual data are reasonable. These normally include range tests, relational tests, and trend tests.

  - Range tests are the simplest and most commonly used validation tests. The measured data are compared to upper and lower limiting values that include nearly (but not absolutely) all of the expected values for the site for each data parameter. For example, a reasonable range for average wind speeds for most sites is 0–25 m/s. Negative values clearly indicate a problem; speeds above 25 m/s are possible, but should be verified with other information. Data reduction and analysis personnel can fine-tune these limiting values as they gain experience. In addition, the limits for appropriate data parameters should be adjusted seasonally. For instance, the limits for air temperature and solar radiation should be lower in winter than in summer. In general, a single item of data should be subjected to more than one range check before it is judged to be valid, because a single check is unlikely to detect all problems. For example, if a frozen wind vane reports an average direction of exactly 180° for six consecutive 10-minute intervals, the values would pass the 0–360° range test, but a check on the standard deviation would reveal a value of zero and the values should be flagged as suspect.

- Relational tests are based on expected physical relationships between various parameters. These ensure that physically improbable situations are flagged as suspect. For example, a significantly higher wind speed at the 25 m level than at the 40 m level should be flagged as suspect.
- Trend tests are based on the rate of change in a parameter over time. An example of a trend that indicates an unusual circumstance and a potential problem is a change in air temperature greater than 5°C in 1 h.

Some data loggers include the capability to record the system battery voltage for each averaging interval. Range and relational tests for a reduction in battery voltage may be used to give early warning of site hardware problems and ensure that data are not lost due to a bad battery or a blown fuse.With experience, the data analysis personnel directly involved in the validation process will become very familiar with the local wind climatology and will learn which criteria are most often triggered and under which conditions. The behavior of the wind under various weather conditions will become apparent, as will the relationship between various parameters. This is an invaluable experience that cannot be gained solely by poring over monthly summary tables, and it may prove to be important for evaluating the impact of the local meteorology on wind turbine operation and maintenance. For example, some validation tests may almost always be failed under light wind conditions, yet the data are valid. This occurrence may argue for one set of test criteria under light wind conditions (below 4 m/s perhaps) and another set for stronger winds. The data analysis personnel should be authorized and encouraged to modify validation test criteria and create new ones as needed, based on their experience with the site data. Be sure to establish operating procedures to ensure appropriate documentation and reporting of any such changes that are made.

- Data verification:

Once suspect data are identified, case-by-case decisions must be made on what to do with the suspect values: retain them as valid, reject them as invalid, or replace them with redundant, valid values (if available). This operation requires the application of judgment of a qualified person familiar with the monitoring equipment and local meteorology. The disposition of each suspect value should be noted in a data verification report. This report should include, for each suspect value, the sensor from which the value was obtained, the date and time that the value was obtained and the disposition of the suspect value, including the source for the replacement value or the validation code for the rejected value.

If a suspect data value is judged to be valid, leave the value as is. If a suspect value is judged to be invalid, but valid data from a redundant sensor is available, replace the invalid value with that from the redundant sensor. If a suspect value is judged to be invalid, and no data from a redundant sensor is available, replace the value with a unique error code that will serve as both a flag for later data processing programs and as an indication of the specific problem with the original value. Selection of the specific error code may require review of the site log or other site data. The data processing and reporting software must incorporate means for handling these error codes. The results of this process are a validated database and a data verification report itemizing the disposition of each suspect data value.

Data Processing and Reporting. After the data validation step is complete, the validated dataset is ready to be processed to quantify the wind resource. This typically involves performing calculations on the dataset, as well as binning or sorting the recorded 10-min average data values into useful subsets based on other chosen averaging intervals, such as hourly or weekly. Hourly averages are normally used for reporting purposes. The processed data are then summarized into weekly or monthly informative reports of summary tables and performance graphs. Items included in these reports typically include mean wind speed, wind direction frequency distribution, maximum gust, mean turbulence intensity, mean power density, and diurnal wind speed and power density (by time of day) for each anemometer and wind direction level; mean shear for each height layer; daily and monthly wind speed distribution for the primary height anemometer and hourly temperature. Data processing and reporting software are available from several sources, including many data logger manufacturers and vendors of spreadsheet, database, and statistical software. Whatever method is used, procedures must be developed to ensure that flagged data points or invalid data codes are excluded from the computations of hourly averages and other quantities. These procedures should be developed and implemented before the first data is recorded.

The wind shear exponent, turbulence intensity, and wind power density are items that are usually included in wind resource reports, but are not routinely produced by most data loggers. These parameters can be easily calculated using a spreadsheet software application to obtain hourly and monthly averages. A description of each parameter and calculation method is presented in detail below.

- Vertical wind shear exponent:

  Wind shear is defined as the change in horizontal wind speed with a change in height. The wind shear exponent (α) must be determined for set of anemometry levels at each site, because the magnitude of α is influenced by site-specific characteristics. Solving the power law equation (Equation 19.21) for α gives

$$\alpha = \frac{\ln(U/U_0)}{\ln(h/h_0)}, \tag{19.28}$$

  where U is the wind speed at height h, and U_0 is the wind speed at height h_0.

- Turbulence intensity:

  Wind turbulence is the rapid disturbances or irregularities in the wind speed, direction, and vertical component. It is an important site characteristic, because high turbulence levels may decrease power output and cause extreme loading on wind turbine components. The most common indicator of turbulence for siting purposes is the standard deviation (σ) of wind speed. Normalizing this value with the mean wind speed gives the turbulence intensity (TI), defined as

$$TI = \sigma/\bar{U}, \tag{19.29}$$

  where σ is the standard deviation of wind speed, and $\bar{U}$ is the mean wind speed (m/s).

- Wind power density:

  WPD is defined in Equation 19.18 as the wind power available per unit area swept by the turbine blades. It combines the effects of the wind speed distribution and its dependence on air density and wind speed. For experimental data, WPD may be calculated as

$$\text{WPD} = \frac{1}{2n}\sum_{i=1}^{n} \rho U_i^3 (\text{W/m}^2), \tag{19.30}$$

  where n is the number of records in the averaging interval, ρ is the air density (kg/m^3), and U_i^3 is the cube of the ith wind speed (m/s) value.

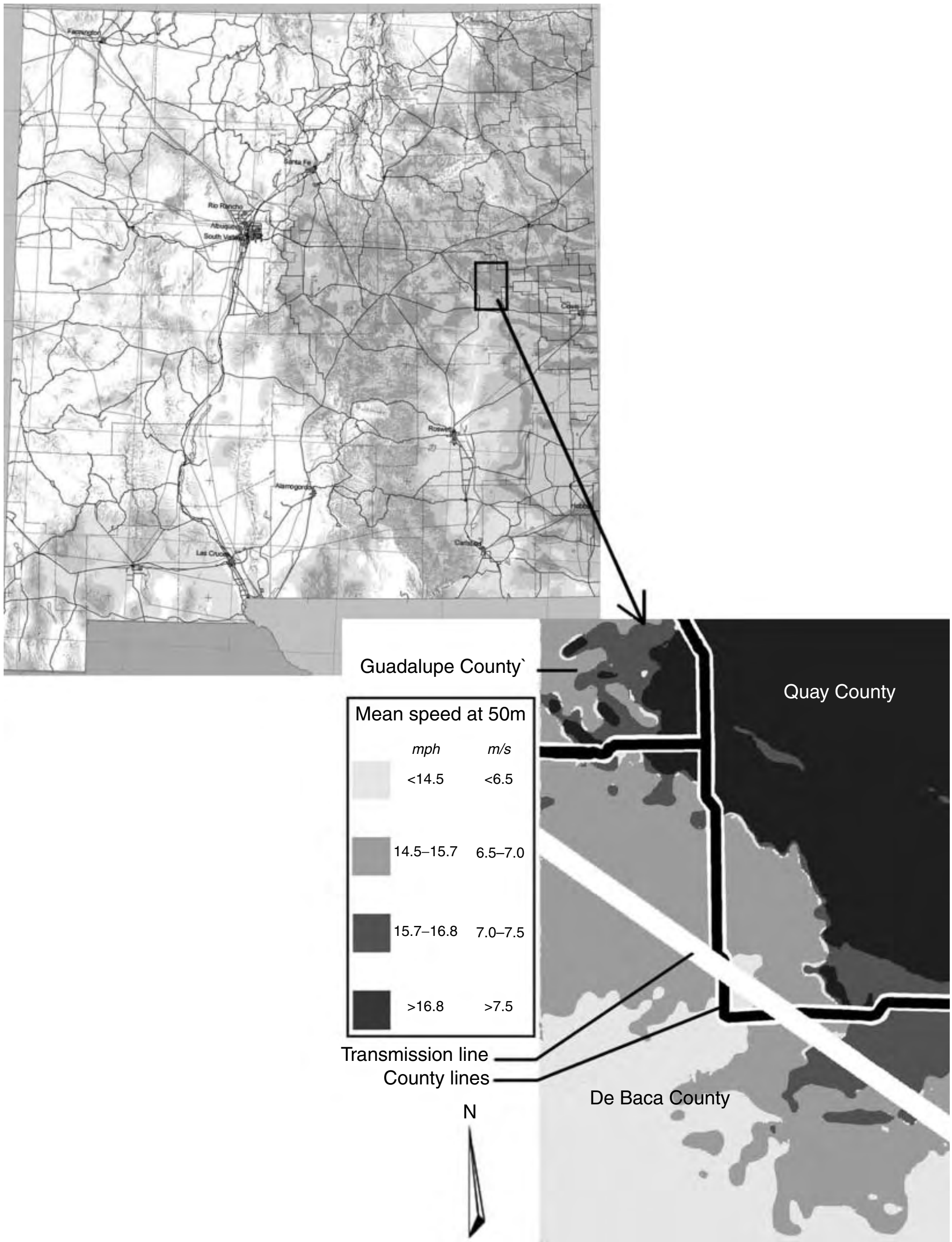


FIGURE 19.37 Wind resource map of the western edge of Tiaban Mesa in eastern New Mexico. (AWS Truewind and the New Mexico Energy, Minerals, and Natural Resources Department.)

19.2.4 Example: Initial Wind Farm Development in New Mexico

By 2001, the New Mexico Public Utility Commission (PUC) had, for some time, been looking at the idea of implementing a Renewable Portfolio Standard (RPS) requiring the Public Service Company of New

Mexico (PNM), the state electrical utility, to obtain some of its electricity from renewable sources. The state legislature was also seriously considering enacting a RPS, but it had not yet done so. PNM had been investigating wind energy and monitoring the wind resource at some promising wind sites for several years, but had not yet been seriously considering adding wind to its generation mix. With the RPS under consideration in both the PUC and the legislature, PNM started seriously considering renewable energy, in general, and in wind energy, in particular.

PNM is the obvious customer for any large wind farm that is developed in the state of New Mexico. New Mexico has fairly good wind resources to the east of the north/south central mountain range and in the eastern part of the state, so those are logical areas in which to start prospecting. The state has worked with NREL and True Solutions (now AWS TrueWind) to develop a high-resolution wind map (available at http://www.emnrd.state.nm.us/ecmd/html/NM%20Wind%20Speed%20June03.pdf). This map also shows the route of high-power transmission lines across the state. A map with transmission line details reveals that a PNM transmission line from Clovis, in eastern New Mexico, to Albuquerque, is a 345-kV line. That line has significant excess capacity, and it directly feeds the 1 GW Albuquerque load.

This line passes through or near several areas with reasonable winds (annual average of 7 m/s or greater), mostly just east of Albuquerque. In the eastern part of the state, the only area near the line with reasonable winds appears to be near the intersection of Guadalupe, De Baca, and Quay Counties. The statewide map in Figure 19.37 locates this area of interest, and the expanded scale map shows the estimated resource from the wind map. The wind map indicates average wind speeds of 7 m/s in the area of the transmission line, with somewhat higher winds just a little to the north.

Closer inspection of the topography in the area shows that the western edge of this high resource corresponds to the western edge of Tiaban Mesa. The mesa edge runs generally north/south in this area, and the top of the mesa is 60–90 m above the plains just to the west. This topography can certainly be expected to enhance the wind speed near the mesa edge. A site visit reveals heavy flagging of the vegetation from the southwest, indicative of a predominant wind direction. The land is privately owned,

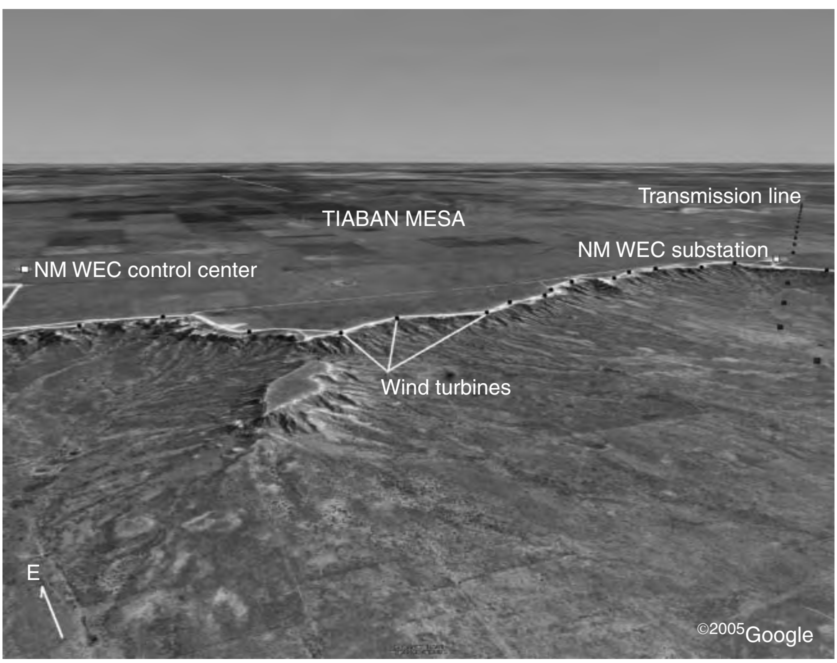


FIGURE 19.38 Perspective view of western edge of Tiaban Mesa. Every second wind turbine and transmission line tower indicated.

with no use restrictions, no large bird populations, and no environmental concerns. A measurement program quickly confirmed that the winds are predominantly southwest, and that the annual average wind speeds along the mesa edge, from just south of the transmission line crossing to several miles north, are 7 m/s or better. Prospective developers thus had the three essential components for a wind farm development: access to a transmission line with excess capacity, a customer for the power (PNM), and a good wind resource.

In August 2002, FPL Energy and PNM announced their agreement to develop the first large wind energy project in the state of New Mexico. FPL Energy built the 204-MW New Mexico Wind Energy Center (NM WEC) on the western edge of Tiaban mesa in 2003. A perspective view of the mesa edge, with the turbine locations and transmission line indicated, is given in Figure 19.38. The PNM-built substation lies right under the existing transmission line and immediately adjacent to the line of turbines on the mesa edge; this ready access to the transmission line was a key driver in the decision to locate the wind farm in this spot. PNM purchases the entire output of the wind farm and sells it a premium to residential and small business customers through their "Sky Blue" renewable energy program. Two years of operation have confirmed that the site resource is a good one: the yearly average wind speed at the 65 m turbine hub height has been about 8 m/s. At the wind farm elevation of 1460 m, this wind speed yields a power density of 460 W/m^2, which corresponds to an NREL wind class 4 site.

19.3 Municipal Solid Waste

Marjorie A. Franklin

This chapter has been extracted from the most recent in a series of reports released by the United States Environmental Protection Agency (EPA) to characterize Municipal Solid Waste (MSW) in the United States (United States Environmental Protection Agency). This report characterizes the national municipal waste stream based on data through 2003. As characterized in the EPA report, MSW includes wastes from residences and commercial establishments. No construction and demolition debris or industrial wastes are included. (Some wastes from industrial establishments, such as packaging and offices wastes, are included, however.)

Identifying the components of the MSW stream is an important step toward addressing the issues associated with using it for energy generation. MSW characterizations that analyze the quantity and composition of the waste stream involve estimating how much MSW is generated, recycled, combusted, and disposed of in landfills. This chapter characterizes the municipal solid waste stream of the nation as a whole; local and regional variations are not addressed.

The methodology used for the characterization of MSW for this chapter estimates the waste stream on a nationwide basis by a "material flows methodology." EPA's Office of Solid Waste and its predecessors in the Public Health Service sponsored work in the 1960s and early 1970s to develop this methodology that is based on production data (by weight) for the materials and products in the waste stream, with adjustments for imports, exports, and product lifetimes.

19.3.1 Materials and Products in MSW

In 2003, generation of municipal solid waste totaled 236 million tons. A breakdown by percentage of the materials generated in MSW in 2003 is shown in Figure 19.39. Paper and paperboard products are the largest component of municipal solid waste by weight (36% of generation) and yard trimmings and food scraps are the second largest components (each 12% of generation). Plastics come next, at 11% of MSW generation. Inorganic portions of the waste stream—metals and glass—total about 13% of generation. (The "other" category also contains some inorganic materials such as broken pottery, kitty litter, etc.) Rubber, leather, textiles, and wood also make up about 13% of MSW generation.

Most of the materials in MSW have some level of recovery for recycling or composting. This is illustrated in Table 19.3. Because each materials category (except for yard trimmings and food scraps) is

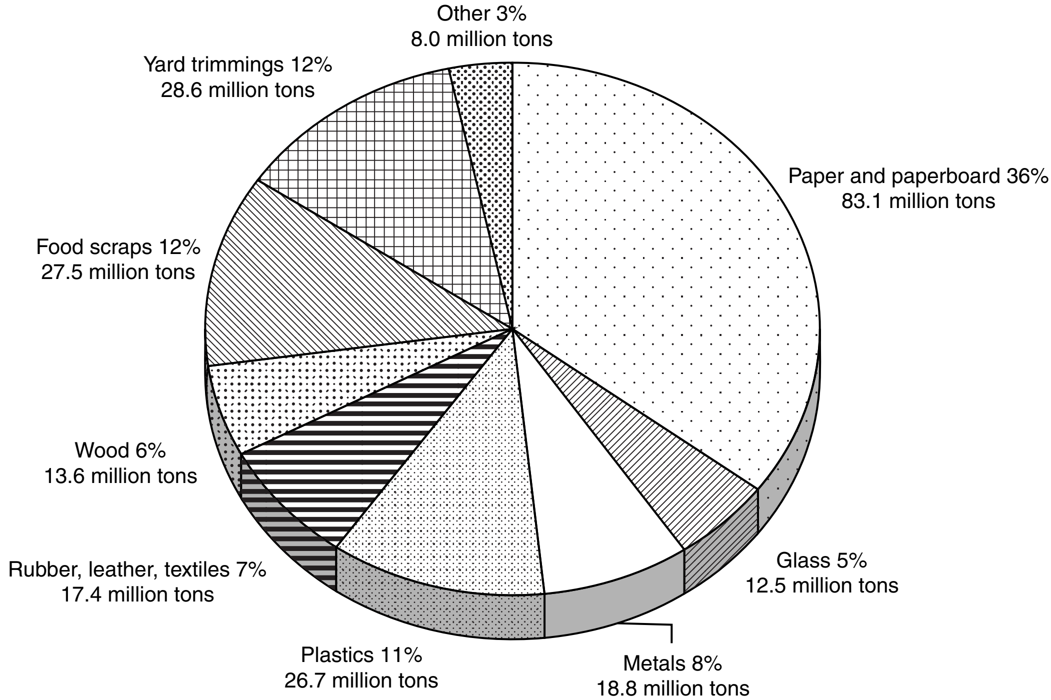


FIGURE 19.39 Materials generated in MSW by weight and percentage, 2003.

TABLE 19.3 Generation and Recovery of Materials in MSW, 2003

	Weight Generated	Weight Recovered	Percent of Generation
Paper and paperboard	83.1	40.0	48.1%
Glass	12.5	2.4	18.8%
Metals			
Steel	14.0	5.1	36.4%
Aluminum	3.2	0.7	21.4%
Other nonferrous metals	1.6	1.1	66.7%
Total metals	18.8	6.8	36.3%
Plastics	26.7	1.4	5.2%
Rubber and leather	6.8	1.1	16.1%
Textiles	10.6	1.5	14.4%
Wood	13.6	1.3	9.4%
Other materials	4.3	1.0	22.7%
Total Materials in Products	176.4	55.4	31.4%
Other wastes			
Food, other	27.6	0.8	2.7%
Yard trimmings	28.6	16.1	56.3%
Misc. inorganic wastes	3.6	Neg.	Neg.
Total other wastes	59.8	16.9	28.2%
Total municipal solid waste	236.2	72.3	30.6%

In millions of tons and percentage of generation of each material. Neg. = Less than 5000 tons or 0.05%. Numbers in this table have been rounded to the first decimal place.

Source: From U.S. Environmental Protection Agency, *Municipal Solid Waste in the United States: 2003 Facts and Figures*, U.S. Environmental Protection Agency.

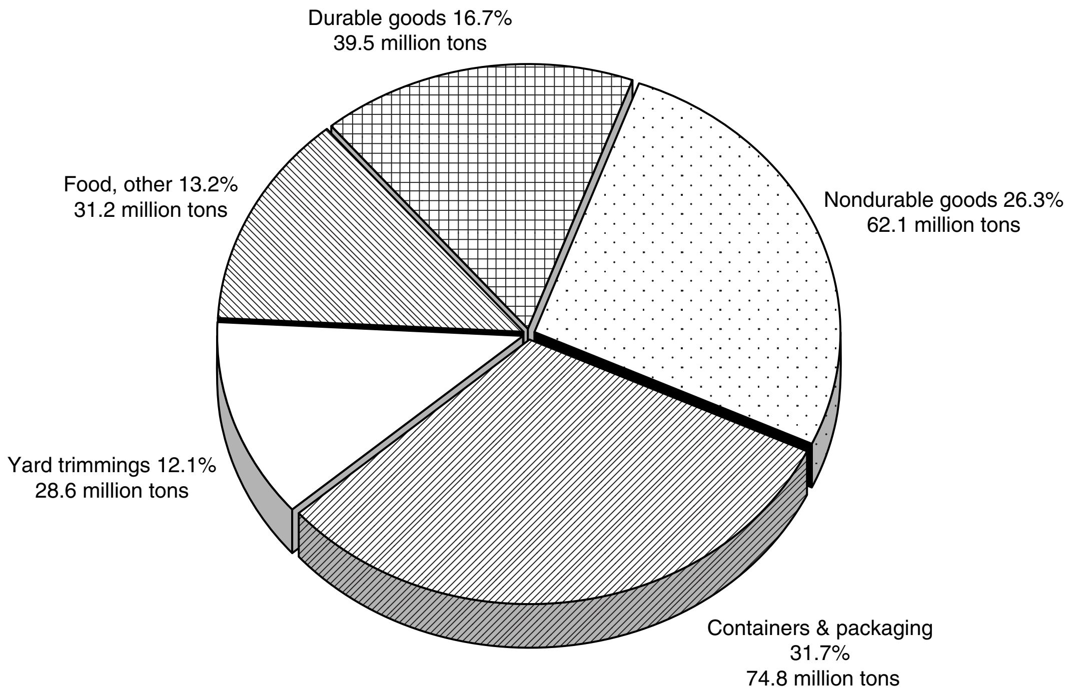


FIGURE 19.40 Products generated in MSW by weight and percentage, 2003.

made up of many different products, some of which may not be recovered at all, the overall recovery rate for any particular material will be lower than recovery rates for some products within the materials category.

The highest recovery rate shown in Table 19.3 is that for nonferrous metals other than aluminum (67% of generation). This is because the lead in lead-acid batteries is recovered at very high rates. Paper and paperboard was recovered at 48% of generation in 2003, and paper and paperboard had by far the highest recovered tonnage. Within that category, newspapers were recovered at 82% and corrugated boxes at 71%

TABLE 19.4 Generation and Recovery of Products in MSW, 2003

	Weight Generated	Weight Recovered	Percent of Generation
Durable goods	39.5	7.2	18.1%
Nondurable goods	62.1	19.3	31.0%
Containers and packaging	74.8	29.0	38.8%
Other wastes			
Food, other	27.6	0.8	2.7%
Yard trimmings	28.6	16.1	56.3%
Misc. inorganic wastes	3.6	Neg.	Neg.
Total other wastes	59.8	16.9	28.2%
Total Municipal Solid Waste	236.2	72.3	30.6%

In millions of tons and percentage of generation of each product category. Neg. = Less than 5000 tons or 0.05%. Numbers in this table have been rounded to the first decimal place.

Source: From U.S. Environmental Protection Agency, *Municipal Solid Waste in the United States: 2003 Facts and Figures*, U.S. Environmental Protection Agency.

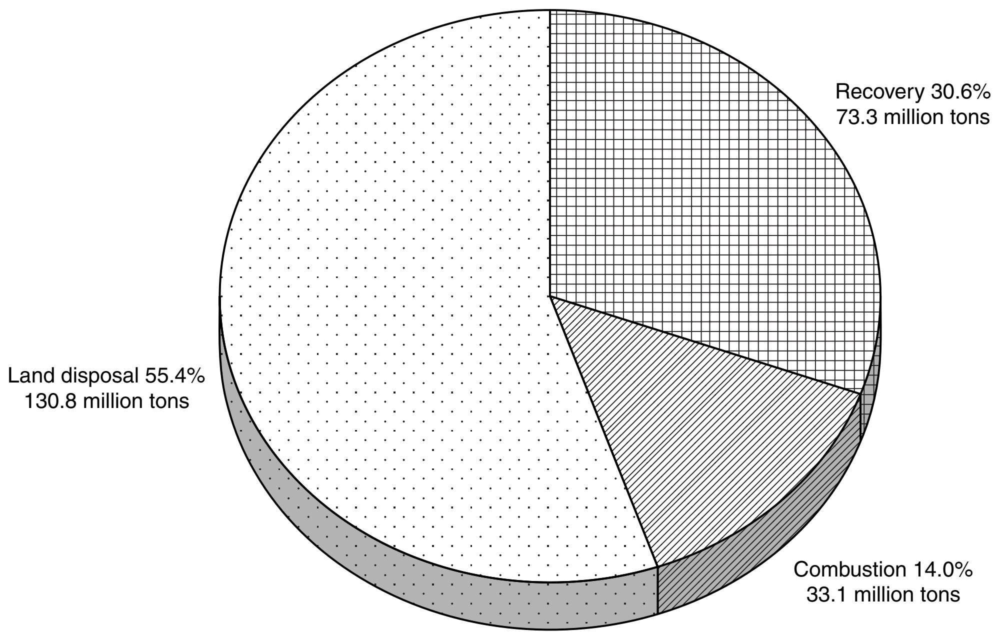


FIGURE 19.41 Management of MSW in the United States, 2003.

of generation. Yard trimmings were recovered for composting at a rate of 56% in 2003. Recovery rates for the other materials are shown in the table.

The many products in MSW are grouped into three main categories: durable goods (e.g., appliances), nondurable goods (e.g., newspapers), and containers and packaging (e.g., beverage cans and corrugated boxes) (see Figure 19.40). The materials in MSW are generally made up of products from each category. There are, however, exceptions. The "durable goods" category contains no paper and paperboard. The "nondurable goods" category includes only small amounts of metals and essentially no glass or wood. The "containers and packaging" category includes only very small amounts of rubber, leather, and textiles.

Generation and recovery of MSW by product category is shown in Table 19.4. Overall, the materials in durable goods were recovered at a rate of about 18% in 2003. Recovery of materials (lead and plastic) from lead-acid batteries was at 93% in 2003. Major appliances were recovered at an overall rate of 67% because of the high rate of recovery of steel in appliances. Recovery of tires at 36% is due to recovery of rubber and some steel.

The overall recovery rate for nondurable goods was estimated at 31% in 2003. Recovery of paper products such as newspapers and office papers accounts for most of this. Recovery of containers and packaging is at the highest rate: 39% in 2003. Large tonnages of corrugated boxes were recovered, at a rate of 71%. Steel cans were recovered at a rate of 60%, and aluminum cans at 44%. Other packaging made of glass, plastics, and wood was also recovered.

19.3.2 Management of MSW

The breakdown of how much MSW went to recycling and composting, combustion, and land disposal[1] in 2003 is shown in Figure 19.41. Recovery of materials for recycling and composting was estimated to have been 73 million tons, or 30.6% of generation, in 2003. Combustion of MSW (nearly all with energy

[1]Land disposal is calculated as the remainder after recycling, composting, and combustion are deducted from generation. This disposal is overwhelmingly landfilled; however, small amounts are littered, self-disposed (e.g., by on-site burning), or otherwise not taken to landfills.

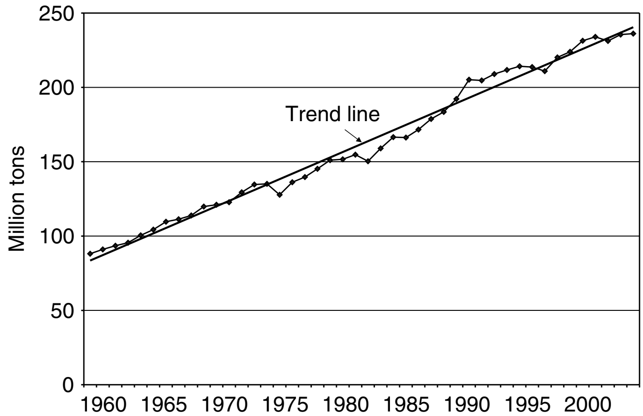


FIGURE 19.42 The trend of municipal solid waste generation, 1960–2003.

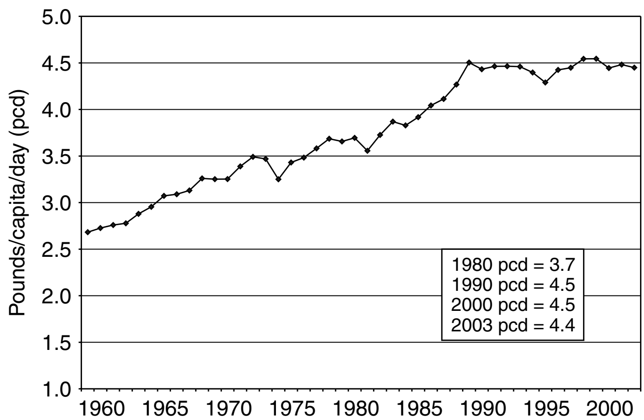


FIGURE 19.43 Generation of municipal solid waste in pounds per capita per day, 1960–2003.

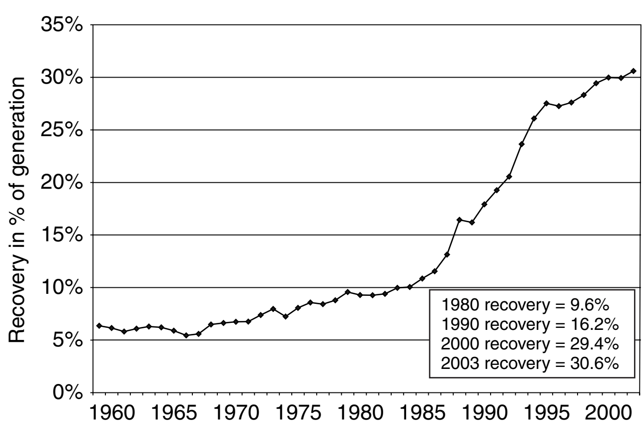


FIGURE 19.44 Recovery of MSW for recycling and composting, in percent of generation, 1960–2003.

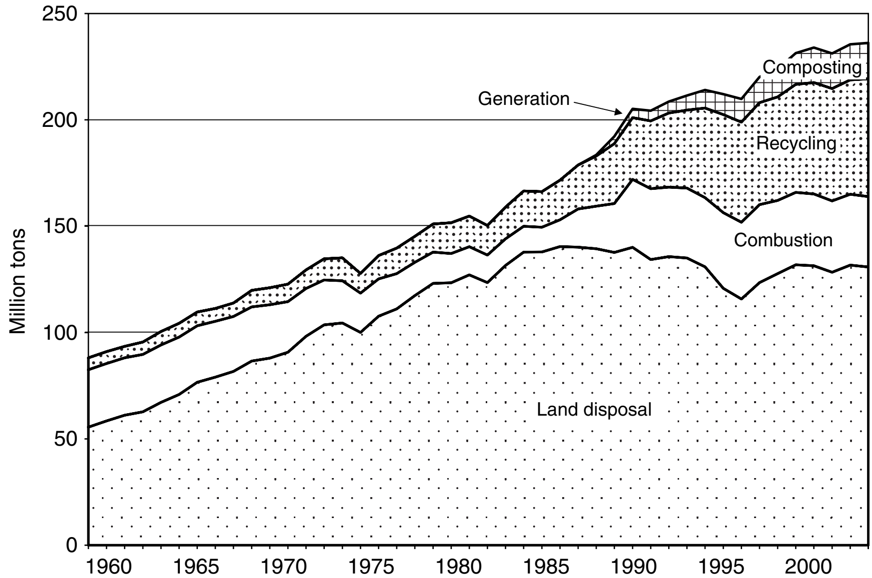


FIGURE 19.45 Generation and management of municipal solid waste, 1960–2003.

recovery) was estimated to have been 33 million tons, or 14% of generation. The remainder—131 million tons, or 55% of generation—was land disposed.

Generation of MSW grew steadily from 1960 to 2003, from 88 million tons to 236 million tons per year (Figure 19.42). As illustrated by the graph, the growth of generation is not continuous, but fluctuates with the economy and, of course, with population growth. It has been demonstrated that there is a high degree of correlation between MSW generation and gross domestic product, and recession years can be identified on the graph. Another way to look at generation is in pounds per capita (pcd) (Figure 19.43). After years of steady growth, pcd has been fairly stable at 4.4–4.5 pounds per capita per day, with a dip in the recession year of 1996.

Recovery for recycling and composting has increased dramatically since the late 1980s (Figure 19.44). Recovery of MSW was minimal in the 1960s and early 1970s. The percentage recovered crept up to about 10% by 1980. Interest in recovery grew rapidly in the late 1980s as concerns were raised about diminishing landfill space in parts of the United States, especially the northeast. Recovery reached 16% in 1990, 26% in 1995, and 29% in 2000. Since then, recovery has been about 30%–31%. Although most recovered material is made up of products in MSW, there also has been an increase in composting of yard trimmings, and, to a much lesser extent, food scraps.

Combustion handled an estimated 30% of MSW generated in 1960, mostly through incinerators with no energy recovery and no air pollution controls. In the 1960s and 1970s combustion dropped steadily as the old incinerators were closed, reaching a low of less than 10% of MSW generated by 1980. The percentage of MSW managed by combustion reached about 17% in 1995; it has been about 14% since 2000.

Land disposal peaked in the 1980s and began to decline after 1990. Land disposal was 81% of generation in 1980; it was 55% of generation in 2003.

19.3.3 Summary

The history of municipal solid waste generation and management is shown in Figure 19.45. The top line of the area graph is MSW generation, whereas management methods are shown as area plots. Major findings from the referenced EPA report are:

- MSW generation in the United States has grown from 88 million tons in 1960 to 236 million tons in 2003. On a per capita basis, generation of MSW has stabilized at 4.4–4.5 pounds per capita per day since 1990.
- Recovery of materials in MSW for recycling and composting has increased dramatically, from about 10% of generation in 1980 to over 30% in 2003. Although recovery percentages increased rapidly until the mid-1990s, incremental percentage growth has been slower in recent years.
- Combustion of MSW reached a low point of about 9% of generation around 1980. Since 2000, combustion has held steady at about 14% of generation.
- Land disposal of MSW peaked at about 140 million tons in the late 1980s. The increase in recovery for recycling and composting along with stable combustion tonnage have lowered the percentage of MSW land disposed to about 55%–56% of generation since 2000.

19.4 Biomass

Robert C. Brown

19.4.1 Defining the Resource

Biorenewable resources, sometimes referred to as *biomass*, are organic materials of recent biological origin (Brown 2003). This definition is deliberately broad with the intent of only excluding fossil fuel resources from the wide variety of organic materials that arise from the biotic environment. Biorenewable resources are generally classified as either *wastes* or *dedicated energy crops*. A waste is a material that has been traditionally discarded because it has no apparent value or represents a nuisance or even a pollutant to the local environment. Dedicated energy crops are plants grown specifically for production of bio-based products, i.e., for purposes other than food or feed.

19.4.2 Waste Materials

Categories of waste materials that qualify as biorenewable resources include municipal solid waste, agricultural residues, agricultural processing by-products, and manure. Municipal solid waste (MSW) is whatever is thrown out in the garbage, and clearly includes materials that do not qualify as biorenewable resources, such as glass, metal, and plastics. This "mixed waste" resource is treated separately in Section 19.3. Agricultural residues are simply that part of a crop discarded after harvest such as corn stover (husks and stalks), rice hulls, wheat straw, and bagasse (fibrous material remaining after the milling of sugar cane). Food processing waste is the effluent from a wide variety of industries ranging from breakfast cereal manufacturers to alcohol breweries. These wastes may be dry solids or watery liquids. The recent concentration of animals into giant livestock facilities has led to calls to treat animal wastes in a manner similar to that for human wastes.

Waste materials share few common traits other than the difficulty of characterizing them because of their variable and complex composition. Thus, waste biomass presents special problems to engineers who are tasked with converting this sometimes unpredictable feedstock into reliable power or high-quality fuels and chemicals. The major virtue of waste materials is their low cost. By definition, waste materials have little apparent economic value and often can be acquired for little more than the cost of transporting the material from its point of origin to a processing plant. Increasing costs for solid waste disposal and sewer discharges and restrictions on landfilling certain wastes allow some wastes to be acquired at negative cost; that is, a biorenewable resource processing plant is paid by a company seeking to dispose of a waste stream. For this reason, many of the most economically attractive opportunities in biorenewable resources involve waste feedstocks.

19.4.3 Dedicated Energy Crops

Dedicated energy crops are typically high fiber crops grown specifically for their high productivity of holocellulose (cellulose and hemicellulose). Harvesting may occur on an annual basis, as with switchgrass, or on a 5–7 year cycle, as with certain strains of fast-growing trees such as hybrid poplar. Lignocellulosic crops are conveniently divided into herbaceous energy crops (HEC) and short-rotation woody crops (SRWC) (Wright and Hohenstein 1994).

Herbaceous crops are plants that have little or no woody tissue. The aboveground growth of these plants usually lives for only a single growing season. However, herbaceous crops include both annuals and perennials. Annuals die at the end of a growing season and must be replanted in the spring. Perennials die back each year in temperate climates but reestablish themselves each spring from rootstock. Both annual and perennial herbaceous energy crops are harvested on at least an annual basis, if not more frequently, with yields averaging 5.5–11 Mg/ha/y, with maximum yields of 20–25 Mg/ha/y in temperate regions (Wright and Hohenstein 1994). As with trees, yields can be much higher in tropical and subtropical regions.

Herbaceous crops more closely resemble hardwoods in their chemical properties than they do softwoods. Their low lignin content makes them relatively easy to delignify, which improves accessibility of the carbohydrate in the lignocellulose. The hemicellulose contains mostly xylan, which is highly susceptible to acid hydrolysis compared to the cellulose. As a result, agricultural residues are susceptible to microbial degradation, destroying their processing potential in a matter of days if exposed to the elements. Herbaceous crops have relatively high silica content compared to woody crops, which can present problems during processing.

Short-rotation woody crop (SRWC) describes woody biomass that is fast growing and suitable for use in dedicated feedstock supply systems. Desirable SRWC candidates display rapid juvenile growth, wide site adaptability, and pest and disease resistance. Woody crops grown on a sustainable basis are harvested on a rotation of 3–10 years.

Woody crops include hardwoods and softwoods. Hardwoods are trees classified as *angiosperms*, which are also known as *flowering plants*. Examples include willow, oak, and poplar. Hardwoods can resprout from stumps, a process known as coppicing, which reduces their production costs compared to softwoods. Advantages of hardwoods in processing include: high density for many species; relative ease of delignification and accessibility of wood carbohydrates; the presence of hemicellulose high in xylan, which can be removed relatively easily; low content of ash, particularly silica, compared to softwoods and herbaceous crops; and high acetyl content compared to most softwoods and herbaceous crops, which is an advantage in the recovery of acetic acid.

Softwoods are trees classified as *gymnosperms*, which encompass most trees known as *evergreens*. Examples include pine, spruce, and cedar. Softwoods are generally fast growing but their carbohydrate is not as accessible for chemical processing as the carbohydrates in hardwood. Because softwoods have considerable value as construction lumber and pulpwood, it is more readily available as waste material in the form of logging and manufacturing residues than are hardwoods. Logging residues, consisting of a high proportion of branches and tops, contain considerable high-density compression wood, which is not easily delignified. Logging residues are more suitable as boiler fuel or other thermochemical treatments than as feedstock for chemical or enzymatic processing.

19.4.4 Properties of Biomass

Evaluation of biomass resources as potential feedstocks generally requires information about plant composition, heating value, bulk density, and production yields. Compositional information can be

TABLE 19.5 Properties of Selected Biomass

Feedstock	Organic Composition (wt %)				Elemental Analysis (dry wt %)					Proximate Analysis (dry wt %)			HHV (MJ/kg)	Bulk density (kg/m^3)	Yield (Mg/ha)
	Cellulose	Hemicellulose	Lignin	Other	C	H	O	N	Ash	Volatile Matter	Fixed C	Ash			
Corn stover	53	15	16	16	44	5.6	43	0.6	6.8	75	19	6	17.7	160–300	8,400
Herbaceous crop	45	30	15	10	47	5.8	42	0.7	4.5	81	15	4	18.7	160–300	14,000
Woody crop	50	23	22	5	48	5.9	44	0.5	1.6	82	16	1.3	19.4	280–480	14,000

Source: From Brown, R. C. 2003. Biorenewable Resources: Engineering New Products from Agriculture, Iowa State Press, Ames, IA.

reported in terms of organic components, proximate analysis, or ultimate analysis. This information is widely scattered and often show wide variations. A table of properties for a few representative kinds of fibrous biomass is given in Table 19.5. Compilations of relevant biomass properties are found in Jenkins et al. (1998), Klass (1998), Brown (2003), and Hofbauer (2004).

Analysis in terms of organic components reports the kinds and amounts of plant chemicals including proteins, oils, sugars, starches, and lignocellulose (fiber). The composition varies widely among plant parts. For example, corn grain is mostly starch (72 wt %) with a relatively small amount of fiber (13 wt %) while corn stover, that part of the crop left on the field, is mostly fiber (84%) with very little starch content. This information is particularly useful in designing biological processes that convert plant components into commodity chemicals and transportation fuels. Often the fiber, which is a polymeric composite of cellulose, hemicellulose, and lignin, is reported in terms of these three constituents, especially when the goal is to break down the carbohydrate fractions into simple sugars.

Proximate analysis is important in developing thermochemical conversion processes for biomass. Proximate analysis reports the yields (% mass basis) of various products obtained upon heating the material under controlled conditions; these products include moisture, volatile matter, fixed carbon, and ash. Because moisture content of biomass is so variable and can be easily determined by gravimetric methods (weighing, heating at 100°C, and reweighing), the proximate analysis of biomass is commonly reported on a dry basis. Volatile matter is that fraction of biomass that decomposes and escapes as gases upon heating a sample at moderate temperatures (about 400°C) in an inert (nonoxidizing) environment. Knowledge of volatile matter is important in designing burners and gasifiers for biomass. The remaining fraction is a mixture of solid carbon (fixed carbon) and mineral matter (ash), which can be distinguished by further heating the sample in the presence of oxygen: the carbon is converted to carbon dioxide leaving only the ash.

Ultimate analysis is simply the (major) elemental composition of the biomass on a gravimetric basis: carbon, hydrogen, oxygen, nitrogen, sulfur, and chlorine along with moisture and ash. Sometimes this information is presented on a dry, ash-free (daf) basis. This information is very important in performing mass balances on biomass conversion processes. Compared to fossil fuels, biomass is characterized by relatively high oxygen content of biomass (typically 40–45 wt %), which detracts from its heating value and represents new challenges in converting these compounds into substitutes for the hydrocarbons that currently dominate our economy. In many instances, generic molecular formula based on one mole of carbon is convenient for performing mass balances on a process. Cellulose and starch have the generic molecular formula $CH_{1.7}O_{0.83}$, hemicellulose can be represented by $CH_{1.6}O_{0.8}$, and wood is $CH_{1.4}O_{0.66}$ (Klass 1998).

Heating value is the net enthalpy released upon reacting a particular fuel with oxygen under isothermal conditions (the starting and ending temperatures are the same). If water vapor formed during reaction condenses at the end of the process, the latent enthalpy of condensation contributes to what is known as the higher heating value (HHV). Otherwise, the latent enthalpy does not contribute and the lower heating value (LHV) prevails. These measurements are typically performed in a bomb calorimeter and yield the higher heating value for the fuel. Heating values of biomass can be conveniently estimated from the percent of carbon in the biomass on a dry basis using the empirical relationship (Klass 1998):

$$\text{HHV in MJ/dry kg} = 0.4571 \times (\%\ \text{C on dry basis}) - 2.70, \tag{19.31}$$

For estimating purposes, assuming a higher heating value of 18 MJ/kg is a good approximation for many kinds of (dry) biomass.

Enthalpies of formation are very useful in thermodynamic calculations but these data are rarely tabulated for biomass because of the wide variability in its composition. However, if higher heating value for a biomass fuel has been determined in a bomb calorimeter, its enthalpy of formation can be determined by summing the enthalpies of formation of the products of combustion and subtracting from this sum the higher heating value. For example, suppose a sample of switchgrass has an elemental analysis that gives a generic molecular formula of $CH_{1.4}O_{0.8}$ (molecular weight of 26.2 kg/kmol) and its higher heating value is measured to be 18.1 MJ/kg. The combustion of one kilomole of switchgrass can be represented by:

$$CH_{1.4}O_{0.8} + 0.95O_2 \rightarrow CO_2 + 0.7\ H_2O(\text{liq}).$$

The enthalpy of reaction, ΔH_R, is calculated from the various enthalpies of formation, h_f°:

$$\begin{aligned}\Delta H_R &= h^o_{f\ CO_2} + 0.7h^o_{f\ H_2O(\text{liq})} - \left(h^o_{f\ CH_{1.4}O_{0.8}} + 0.95h^o_{f\ O_2}\right)\\ &= -393.5\text{MJ/kmol} + 0.7(-285.8\text{MJ/kmol}) - \left(h^o_{f\ CH_{1.4}O_{0.8}} + 0\right)\\ &= 18.1\ \text{MJ/kg}\ \left(26.2\ \text{kg/kmol}\right).\end{aligned}$$

Solving for the enthalpy of formation for $CH_{1.4}O_{0.8}$ yields:

$$h^o_{f\ CH_{1.4}O_{0.8}} = -119.3\text{MJ/kmol}.$$

This value can be used in various thermodynamic calculations for this switchgrass sample, whether gasification to hydrogen or hydrolysis to fermentable sugars.

Bulk density is determined by weighing a known volume of biomass that is packed or baled in the form anticipated for its transportation or use. Clearly, solid logs will have higher bulk density than the same wood chipped. Bulk density will be an important determinant of transportation costs and the size of fuel storage and handling equipment. Volumetric energy content is also important in transportation and storage issues. Volumetric energy content, which is simply the enthalpy content of fuel per unit volume, is calculated by multiplying the higher heating value of a fuel by its bulk density.

Knowledge of the annual yields of biomass crops (Mg per hectare) is important in planning a biomass energy project. Yield depends on many factors including plant variety, crop management (fertilization and pest control), soil type, landscape, climate, weather, and water drainage. Thus, each project will require site-specific information obtained through discussions with state extension agents and local agronomists in combination with field trials in advance of detailed manufacturing plant design.

19.4.5 Size of Resource Base

The amount of biomass that could be produced annually in the United States is very difficult to estimate. In addition to the uncertainties in estimating yields on the wide diversity of landscapes that could be planted to biomass crops, a variety of social, political, economic, and environmental factors influence decisions on putting land into biomass production. Because these factors are not static, assessments of land available can change with time. It is not surprising that there is a wide opinion on the size of the biomass supply.

Two studies in the 1990s illustrate the divergence of these views. Graham (1994) published an analysis of potential biomass supply that foresaw 159 million hectares of land suitable for dedicated energy crops. Because "suitable" was defined as lands capable of producing in excess of 11.2 Mg/ha/y of dry biomass, the total supply of biomass was projected to be 1.8 billion Mg of dry biomass. Assuming a heating value of 18 MJ/kg, this analysis, which ignores the contribution of agricultural residues to biomass supply, yields 32 billion GJ of energy.[2] In contrast, a study by Wyman and Goodman (1993) a year earlier estimated dedicated energy crops could supply 26–52 billion GJ of energy. They also included agricultural residues, forestry residues, and municipal solid wastes in their analysis, which contributed another 8 billion GJ of biomass energy, yielding a total annual energy supply from biomass of up to 60 billion GJ. Because energy consumption in the United States is about 98 billion GJ per year (Energy Information

[2]In the United States, the unit of national energy consumption is the quad, which is 1 quadrillion British thermal units (Btu), equal to 1.054 billion GJ. Thus, in the above discussions, a billion GJ can be thought of as one quad.

TABLE 19.6 Present Biomass Supply in the United States (2003)

	Annual Biomass Supply (million dry Mg/yr)
Forestry products industry	
Wood residues	40
Pulping liquors	47
Urban woods and processing residues	32
Fuelwood	32
Grains for biofuels	16
Bioproducts	5
Biomass Total	172

Source: From Energy Information Administration, 2004. Annual energy outlook 2004: With projections to 2025.

TABLE 19.7 Potential Biomass Supply for the United States

	Annual Biomass Supply (million Mg/y)
Forestry resources	
Fuelwood	47
Milling residues	132
Urban wood residue	43
Logging residues	58
Forest thinning	55
Forestry total	335
Agricultural resources	
Crop residues	389
Perennial crops	343
Grains for biofuels	79
Processing residues and manure	96
Agricultural total	907
Biomass total	1240

Source: From Perlack, R. D., Wright, L. L., Turhollow, A. F., Graham, R. L., Stokes, B. J. and Erbach, D. C. 2005. Biomass as feedstock for a bioenergy and bioproducts industry: The technical feasibility of a billion ton annual supply, Technical Report GO-102995-2135, U.S. Department of Energy.

Administration 1994), the Graham study suggests that biomass could supply up to 33% of U.S. energy demand while the Wyman and Goodman study places this percentage as high as 60%.

More recently, the United States Department of Agriculture sponsored a study of biomass supply, including both dedicated crops and agricultural residues (Perlack et al. 2005). The conclusion of this study is that in excess of 1.2 billion tons of dry biomass, representing 25 billion GJ of energy could be produced in a sustainable manner. Because this report is the most recent and takes in consideration information from earlier studies, it will be examined here in more detail.

In 2003, the harvested biomass supply in the United States was 172 million dry Mg in the form of wood residues and pulping liquors from the forest products industry, urban wood and processing residues, fuelwood, grains for biofuels, and bioproducts (Energy Information Administration 2004). Table 19.6 details the biomass supply from these various sources, with pulping liquors representing the single largest source (27% of the total). With an average heating value of 18 MJ/kg, biomass contributed nearly 3 billion GJ to the nation's energy supply, which is only 3% of the total U.S. consumption. This is only a small fraction of what could be produced and harvested in a sustainable fashion.

The land base of the United States encompasses nearly 2263 million acres. About 33% is classified as forest land, 26% as grassland pasture and range, 20% as cropland, 8% special uses (e.g., public facilities), and 13% miscellaneous such as urban areas, swamps, and deserts (Vesterby and Krupa 2001; Alig et al. 2003). The USDA study considered how the first three categories of land could be employed in biomass production, carefully excluding inaccessible and environmentally sensitive lands from consideration. It also accounted for use of these lands in the production of conventional forest products and agricultural commodities.

The study included some forward looking projections on agricultural technology: small grain yields increases by 50%; residue ratio for soybeans increases to 2:1; harvest technology recovers 75% of crop residues (when sustainable removal is possible); all cropland is managed by no-till methods; 55 million acres of cropland, idle cropland, and pasture are dedicated to perennial bioenergy crops; and all manure in excess of allowable on-farm soil application is used for biofuel.

Table 19.7 summarizes the findings of the USDA study, which projects at least 1.2 billion Mg of dry biomass could be harvested annually in the United States. Forestry resources, in the form of fuelwood, milling residues, urban wood residue, logging residues, and wood recovered from forest thinning (for forest fire control) could yield 335 million Mg of dry wood per year or 28% of the total. Agricultural resources, in the form of crop residues, perennial crops, grains for biofuels, and processing residues and manure, could yield 907 million Mg of dry biomass per year. Crop residues and perennial crops represent 31% and 28%, respectively, of the total biomass supply. Assuming an average heating value of 18 MJ/kg, this 1.2 billion Mg supply of biomass represents 21 billion GJ of energy, or 21% of the total U.S. energy consumption. Depending upon the technology employed to produce cellulosic renewable fuels and to utilize it in the transportation sector, this supply could satisfy two-thirds or more of current U. S. gasoline demand. Thus, biomass could be a significant contribution to the future U.S. energy supply.

References

Alig, R. J., Plantinga, A. J., Ahn, S. and Kline, J. D. 2003. Land use changes involving forestry in the United States: 1952 to 1997, with projections to 2050. General Technical PNW-GTR-587. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Corvallis, OR, September.

Archer, C. L. and Jacobson, M. Z. 2005. Evaluation of global wind power. *Journal of Geophysical Research*, 110, D12110, doi: 10.1029/2004JD005462.

Aspliden, C. I., Elliot, D. L., and Wendall, L. L. 1986. Resource assessment methods, siting and performance evaluation, In *Physical Climatology for Solar and Wind Energy*, R. Guzzi and C. G. Justus, eds., pp. 321–326. World Scientific, NJ.

American Society of Mechanical Engineers (ASME), 1989. Performance test code for wind urbines. ASME/ANSI PTC 42-1988, American Society of Mechanical Engineers, New York.

American Society for Testing and Materials (ASTM) 1984. Standard for terrestrial solar spectral irradiance tables at air mass 1.5 for a 37° tilted surface. ASTM Standard E892-82. In *Annual Book of ASTM Standards*, Vol. 12.02, pp. 72–719. The American Society of Testing and Materials, Philadelphia.

American Wind Energy Association (AWEA). 1986. Standard procedures for meteorological Measurements at a potential wind turbine site. D-FC02-86CH10302. American Wind Energy Association, Washington, DC.

AWS Scientific, Inc. 1997. Wind resource assessment handbook: Fundamentals for conducting a successful monitoring program, SR-440-22223. National Renewable Energy Laboratory, Golden, CO.

Belcher, B. N. and DeGaetano, A. T. 2004. Integration of ASOS weather data into building energy calculations with emphasis on model derived solar radiation. Publication ASHRAE 1226-RP. Northeast Regional Climate Center, Ithaca NY.

Brown, R. C. 2003. *Biorenewable Resources: Engineering New Products from Agriculture*. Iowa State Press, Ames, IA.

Cooper, P. I. 1969. The absorption of solar radiation on solar stills. *Solar Energy*, 12, 3, 333–346.

Corotis, R. B. 1977. Stochastic modeling of site wind characteristics. RLO/2342-77/2. Northwestern University, Evanston, IL.

Duffie, J. A. and Beckman, W. A. 1980. *Solar Engineering of Thermal Processes*. Wiley, New York.

Elliott, D. L., Holladay, C. G., Barchet, W. R., Foote, H. P., Sandusky, W. F. 1987. Wind energy resource atlas of the United States. DOE/CH10094-4. Solar Energy Research Institute, Golden, CO.

Energy Information Administration. 2004. Annual energy outlook 2004: With projections to 2025, January.

European Wind Energy Association (EWEC)/Greenpeace 2003. *Wind Force 12: A Blueprint to Achieve 12% of the World's Electricity from Wind Power by 2020*. EWEC and Greenpeace, Brussels.

Frost, W. and Nowak, D. K. 1979. Summary of guidelines for siting wind turbine generators relative to small-scale two-dimensional terrain features. Report RLO-2443-77/1, Battelle Pacific Northwest Laboratory, Richland, WA.

Gipe, P. 1993. *Wind Power for Home & Business—Renewable Energy for the 1990s and Beyond*. Chelsea Green Publishing, Post Mills.

Golding, E. 1977. *The Generation of Electricity by Wind Power*. E. & F.N. Spon, London.

Graham, R. L. 1994. An analysis of the potential land base for energy crops in the conterminous United States. *Biomass and Bioenergy*, 6(3), 175–189.

Heister, T. R., Pennell, W. T. 1981. The meteorological aspects of siting large wind turbines. U.S. Department of Energy Report PNL-2522. Pacific Northwest Laboratory, Richland, WA.

Hewson, E. W., Wade, J. E., Baker, R. W. 1978. Vegetation as an indicator of high wind velocity. RLO/2227-T24-78-2. Oregon State University, Corvallis, OR.

Hofbauer, H. 2004. BIOBIB—A database for biofuels, Institute of Chemical Engineering, Vienna University of Technology, Vienna, Austria, http://www.vt.tuwien.ac.at/biobib/biobib.html (accessed on october 4, 2006).

Iqbal, Muhammad 1983. *An Introduction to Solar Radiation*. Academic Press, Toronto.

Jenkins, B. M., Baxter, L. L., Miles, T. R. Jr., and Miles, T. R. 1998. Combustion properties of biomass. *Fuel Processing Technology*, 54, 1-3, 17–46.

Justus, C. G. 1978. *Winds and Wind System Performance*. Franklin Institute Press, Philadelphia.

Klass, D. L. 1998. *Biomass for Renewable Energy, Fuels, and Chemicals*. Academic Press, New York.

Kreider, J. F. and Kreith, F. 1981. *Solar Energy Handbook*. McGraw-Hill, New York.

Marion, W., Urban, K. 1995. User's *Manual for TMY2s*. Derived from the 1961–1990 National Solar Radiation Data Base. NREL/SP 463-7668, National Renewable Energy Laboratory, Golden, CO.

Marion, W., Wilcox, S. 1994. Solar radiation data manual for flat-plate and concentrating collectors. NREL/TP-463-5607. National Renewable Energy Laboratory, Golden, CO.

Maxwell, E. L., Marion, W., Myers, D. R. 1995. Comparisons of NSRDB and SOLMER Data, In *National Solar Radiation Data Base*, *Vol. 2*, (1961–1990), pp. 216–260, National Renewable Energy Laboratory, Golden, CO.

Maxwell, E. L., Myers, D. R. 1992. Daily estimates of aerosol optical depth for solar radiation models. In *Proceedings of the 1992 Annual Conference of the American Solar Energy Society*, pp. 323–327. Pacific Northwest Research Station, Corvallis, OR.

National Renewable Energy Laboratory. 1992. *NSRDB-Vol. 1, National Solar Radiation Data Base (1961–1990): User's Manual*, National Renewable Energy Laboratory, Golden, CO.

Perez, R., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R., and Vignola, F. 2002. A new operational satellite-to-irradiance model. *Solar Energy*, 73, 5, 307–317.

Perlack, R. D., Wright, L. L., Turhollow, A. F., Graham, R. L., Stokes, B. J., Erbach, D. C. 2005. Biomass as feedstock for a bioenergy and bioproducts industry: The technical feasibility of a billion-ton annual supply. Technical Report GO-102995-2135, U.S. Department of Energy.

Reddy, T. A. 1987. *The Design and Sizing of Active Solar Thermal Systems*. Clarendon Press, Oxford.

Renné, D., George, R., Marion, B., Perez, R., Schillings, C., Trieb, F., Meyer, R. et al. 2005. Results of Solar Resource Assessments in the UNEP/SWERA Project. In *Proceedings of the 2005 Solar World Congress*.

Rohatgi, J. S. and Nelson, V. 1994. *Wind Characteristics: An Analysis for the Generation of Wind Power*. Alternative Energy Institute, West Texas A&M University, Canyon, TX.

Schwartz, M., Elliott, D. 2005. Towards a wind energy climatology at advanced turbine hub heights. NREL/CP-500-38109. National Renewable Energy Laboratory, Golden, CO.

Troen, I. and Petersen, E. L. 1989. *European Wind Atlas*. Risø National Laboratory, Roskilde.

U.S. Environmental Protection Agency. 2003. Municipal solid waste in the United States: 2003 facts and figures. http://www.epa.gov/epaoswer/non-hw/muncpl/msw99.htm (accessed on october 4, 2006).

Vesterby, M., Krupa, L. 2001. Major uses of land in the United States, 1997. Statistical Bulletin No. 973, U.S. Department of Agriculture, Economic Research Service.

Wegley, H. L., Ramsdell, J. V., Orgill, M. M., Drake, R. L. 1980. A siting handbook for small wind energy conversion systems. PNL-2521. Pacific Northwest Laboratory, Richland, WA.

Wilcox, S., Anderberg, M., George, R., Marion, W., Myers, D., Renné, D., Beckman, W. et al. 2005. Progress on an updated National Solar Radiation Data Base for the United States. In *Proceedings of the 2005 Solar World Congress*. Pacific Northwest Research Station, Corvallis, OR.

Wright, L., Hohenstein, L., eds. 1994. Dedicated feedstock supply systems: Their current status in the U.S.A. *Biomass and Bioenergy*, 6(3), 161–173.

Wyman, C. E. and Goodman, B. J. 1993. Biotechnology for production of fuels, chemicals, and materials from biomass. *Applied Biochemistry & Biotechnology*, 39/40, 41–59.

20

Solar Thermal Energy Conversion

T. Agami Reddy
Drexel University

Riccardo Battisti
University of Rome "La Sapienza"

Hans Schweiger
Active Solar System Group

Werner Weiss
AEE INTEC

Jeffrey H. Morehouse
University of South Carolina

Sanjay Vijayaraghavan
Intel Technology India Pvt. Ltd.

D. Yogi Goswami
University of South Florida

20.1 Active Solar Heating Systems

T. Agami Reddy

20.1.1 Introduction

This section defines the scope of the entire chapter and presents a brief overview of the types of applications that solar thermal energy can potentially satisfy.

20.1.1.1 Motivation and Scope

Successful solar system design is an iterative process involving consideration of many technical, practical, reliability, cost, code, and environmental considerations (Mueller Associates 1985). The success of a project involves identification of and intelligent selection among trade-offs, for which a proper understanding of goals, objectives, and constraints is essential. Given the limited experience available in the solar field, it is advisable to keep solar systems as simple as possible and not be lured by the promise of higher efficiency offered by more complex systems. Because of the location-specific variability of the solar resource, solar systems offer certain design complexities and concerns not encountered in traditional energy systems.

The objective of this chapter is to provide energy professionals which a fundamental working knowledge of the scientific and engineering principles of solar collectors and solar systems relevant to both the prefeasibility study and the feasibility study of a solar project. Conventional equipment such as heat exchangers, pumps, and piping layout are but briefly described. Because of space limitations, certain equations/correlations had to be omitted, and proper justice could not be given to several concepts and design approaches. Effort has been made to provide the reader with pertinent references to textbooks, manuals, and research papers.

A detailed design of solar systems requires in-depth knowledge and experience in (i) the use of specially developed computer programs for detailed simulation of solar system performance, (ii) designing conventional equipment, controls, and hydronic systems, (iii) practical aspects of equipment installation, and (iv) economic analysis. These aspects are not addressed here, given the limited scope of this chapter. Readers interested in acquiring such details can consult manuals such as Mueller Associates (1985) or SERI (1989).

The lengthy process outlined above pertains to large solar installations. The process is much less involved when a small domestic hot-water system, or unitary solar equipment or single solar appliances such as solar stills, solar cookers, or solar dryers are to be installed. Not only do such appliances differ in engineering construction from region to region, there are also standardized commercially available units whose designs are already more or less optimized by the manufacturers, normally as a result of previous experimentation, both technical or otherwise. Such equipment is not described in this chapter for want for space.

The design concepts described in this chapter are applicable to domestic water heating, swimming pool heating, active space heating, industrial process heat, convective drying systems, and solar cooling systems.

20.1.2 Solar Collectors

20.1.2.1 Collector Types

A solar thermal collector is a heat exchanger that converts radiant solar energy into heat. In essence this consists of a receiver that absorbs the solar radiation and then transfers the thermal energy to a working fluid. Because of the nature of the radiant energy (its spectral characteristics, its diurnal and seasonal variability, changes in diffuse to global fraction, etc.), as well as the different types of applications for which solar thermal energy can be used, the analysis and design of solar collectors present unique and unconventional problems in heat transfer, optics, and material science. The classification of solar collectors can be made according to the type of working fluid (water, air, or oils) or the type of solar receiver used (nontracking or tracking).

Most commonly used working fluids are water (glycol being added for freeze protection) and air. Table 20.1 identifies the relative advantages and potential disadvantages of air and liquid collectors and associated systems. Because of the poorer heat transfer characteristics of air with the solar absorber, the air collector may operate at a higher temperature than a liquid-filled collector, resulting in greater thermal losses and, consequently, a lower efficiency. The choice of the working fluid is usually dictated by the application. For example, air collectors are suitable for space heating and convective drying applications, while liquid collectors are the obvious choice for domestic and industrial hot-water applications. In certain high-temperature applications, special types of oils are used that provide better heat transfer characteristics.

The second criterion of collector classification is according to the presence of a mechanism to track the sun throughout the day and year in either a continuous or discreet fashion (see Table 20.2). The

TABLE 20.1 Advantages and Disadvantages of Liquid and Air Systems

Characteristics	Liquid	Air
Efficiency	Collectors generally more efficient for a given temperature difference	Collectors generally operate at slightly lower efficiency
System configuration	Can be readily combined with service hot-water and cooling systems	Space heat can be supplied directly but does not adapt easily to cooling. Can preheat hot-water
Freeze protection	May require antifreeze and heat exchangers that add cost and reduce efficiency	None needed
Maintenance	Precautions must be taken against leakage, corrosion and boiling	Low maintenance requirements. Leaks repaired readily with duct tape, but leaks may be difficult to find
Space requirements	Insulated pipes take up nominal space and are more convenient to install in existing buildings	Duct work and rock storage units are bulky, but ducting is a standard HVAC installation technique
Operation	Less energy required to pump liquids	More energy required by blowers to move air; noisier operation
Cost	Collectors cost more	Storage costs more
State of the art	Has received considerable attention from solar industry	Has received less attention from solar industry

Source: From SERI, *Engineering Principles and Concepts for Active Solar Systems*, Hemisphere Publishing Company, New York, 1989.

stationary flat-plate collectors are rigidly mounted, facing toward the equator with a tilt angle from the horizontal roughly equal to the latitude of the location for optimal year-round operation. The compound parabolic concentrators (CPCs) can be designed either as completely stationary devices or as devices that need seasonal adjustments only. On the other hand, Fresnel reflectors, paraboloids, and heliostats need two-axis tracking. Parabolic troughs have one axis tracking either along the east–west direction or the north–south direction. These collector types are described by Kreider (1979a, 1979b) and Rabl (1985).

A third classification criterion is to distinguish between nonconcentrating and concentrating collectors. The main reason for using concentrating collectors is not that *more energy* can be collected but that the thermal energy is obtained at higher temperatures. This is done by decreasing the area from which heat losses occur (called the receiver area) with respect to the aperture area (i.e., the area that intercepts the solar radiation). The ratio of the aperture to receiver area is called the *concentration ratio*.

20.1.2.2 Flat-Plate Collectors

20.1.2.2.1 Description

The flat-plate collector is the most common conversion device in operation today, since it is most economical and appropriate for delivering energy at temperatures up to about 100°C. The construction of flat-plate collectors is relatively simple, and many commercial models are available.

Figure 20.1 shows the physical arrangements of the major components of a conventional flat-plate collector with a liquid working fluid. The blackened absorber is heated by radiation admitted via the transparent cover. Thermal losses to the surroundings from the absorber are contained by the cover,

TABLE 20.2 Types of Solar Thermal Collectors

Nontracking Collectors	Tracking Collectors
Basic flat-plate	Parabolic troughs
Flat-plate enhanced with side reflectors or V-troughs	Fresnel reflectors
Tubular collectors	Paraboloids
Compound parabolic concentrators (CPCs)	Heliostats with central receivers

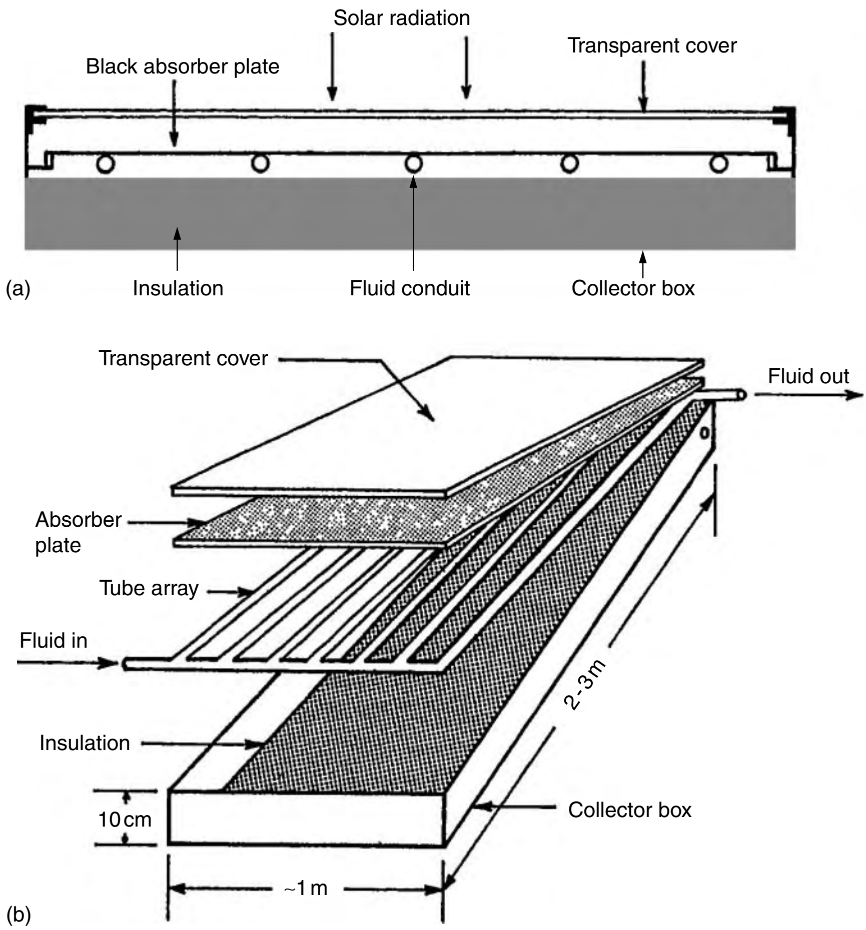


FIGURE 20.1 Cross-section and isometric view of a flat-plate collector.

which acts as a black body to the infrared radiation (this effect is called the *greenhouse* effect), and by insulation provided under the absorber plate. Passages attached to the absorber are filled with a circulating fluid, which extracts energy from the hot absorber. The simplicity of the overall device makes for long service life.

The absorber is the most complex portion of the flat-plate collector, and a great variety of configurations are currently available for liquid and air collectors. Figure 20.2 illustrates some of these concepts in absorber design for both liquid and air absorbers. Conventional materials are copper, aluminum, and steel. The absorber is either painted with a dull black paint or can be coated with a *selective surface* to improve performance (see "Improvements to Flat-Plate Collector Performance" for more details). Bonded plates having internal passageways perform well as absorber plates because the hydraulic passageways can be designed for optimal fluid and thermal performance. Such collectors are called *roll-bond* collectors. Another common absorber consists of tubes soldered or brazed to a single metal sheet, and mechanical attachments of the tubes to the plate have also been employed. This type of collector is called a *tube-and-sheet* collector. Heat pipe collectors have also been developed, though these are not as widespread as the previous two types. The so-called *trickle type* of flat-plate collector, with the fluid flowing directly over the corrugated absorber plate, dispenses entirely with fluid passageways. Tubular collectors have also been used because of the relative ease by which air can be evacuated from such collectors, thereby reducing convective heat losses from the absorber to the ambient air.

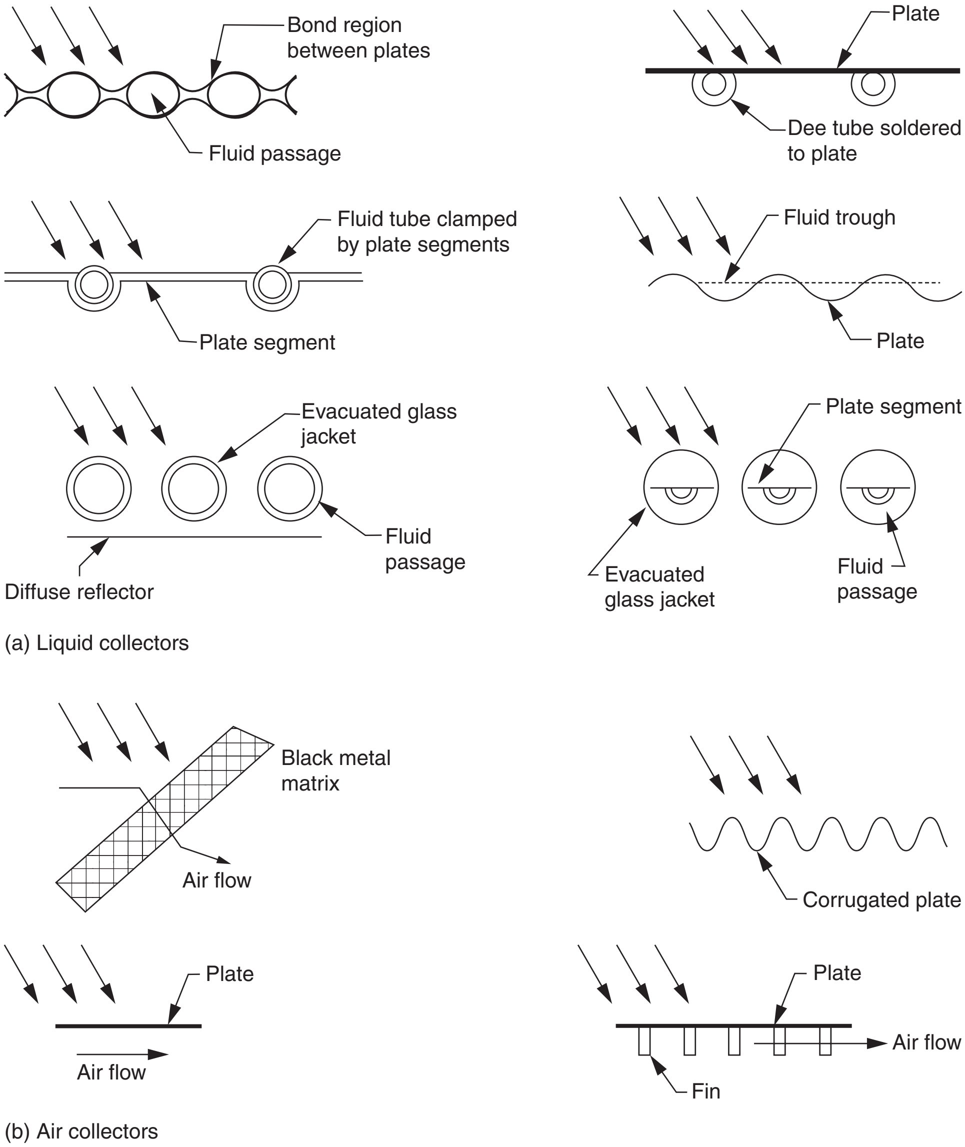


FIGURE 20.2 Typical flat-plate absorber configurations.

The absorber in an air collector normally requires a larger surface than in a liquid collector because of the poorer heat transfer coefficients of the flowing air stream. Roughness elements and producing turbulence by way of devices such as expanded metal foil, wool, and overlapping plates have been used as a means for increasing the heat transfer from the absorber to the working fluid. Another approach to enhance heat transfer is to use packed beds of expanded metal foils or matrices between the glazing and the bottom plate.

20.1.2.2.2 Modeling

A particular modeling approach and the corresponding degree of complexity in the model are dictated by the objective as well as by experience gained from past simulation work. For example, it has been found that transient collector behavior has insignificant influence when one is interested in determining the long-term performance of a solar thermal system. For complex systems or systems meant for nonstandard applications, detailed modeling and careful simulation of system operation are a must initially, and simplifications in component models and system operation can subsequently be made. However, in the case of solar thermal systems, many of the possible applications have been studied to date and a backlog of experience is available not only concerning system configurations but also with reference to the degree of component model complexity.

Because of low collector time constants (about 5–10 min), heat capacity effects are usually small. Then the instantaneous (or hourly, because radiation data are normally available in hourly time increments only) steady-state useful energy q_C in watts delivered by a solar flat-plate collector of surface area A_C is given by

$$q_C = A_C F'[I_T \eta_0 - U_L(T_{Cm} - T_a)]^+ \tag{20.1}$$

where F' is the plate efficiency factor, which is a measure of how good the heat transfer is between the fluid and the absorber plate; η_0 is the optical efficiency, or the product of the transmittance and absorptance of the cover and absorber of the collector; U_L is the overall heat loss coefficient of the collector, which is dependent on collector design only and is normally expressed in W/(m^{2}°C); T_{Cm} is the *mean* fluid temperature in the collector (in °C); and I_T is the radiation intensity on the plane of the collector (in W/m^2). The + sign denotes that negative values are to be set to zero, which physically implies that the collector should not be operated when q_C is negative (i.e., when the collector loses more heat than it can collect).

However, because T_{Cm} is not a convenient quantity to use, it is more appropriate to express collector performance in terms of the fluid *inlet* temperature to the collector (T_{Ci}). This equation is known as the classical Hottel–Whillier–Bliss (HWB) equation and is most widely used to predict instantaneous collector performance:

$$q_C = A_C F_R[I_T \eta_0 - U_L(T_{Ci} - T_a)]^+ \tag{20.2}$$

where F_R is called the heat removal factor and is a measure of the solar collector performance as a heat exchanger, since it can be interpreted as the ratio of actual heat transfer to the maximum possible heat transfer. It is related to F' by

$$\frac{F_R}{F'} = \frac{(mc_p)_C}{A_C F' U_L}\left\{1 - \exp\left[-\frac{A_C U_L F'}{(mc_p)_C}\right]\right\} \tag{20.3}$$

where m_C is the total fluid flow rate through the collectors and c_{pc} is the specific heat of the fluid flowing through the collector. The variation of (F_R/F') with $[(mC_p)_C/A_C U_L F']$ is shown graphically in Figure 20.3. Note the asymptotic behavior of the plot, which suggests that increasing the fluid flow rate more than a certain amount results in little improvement in F_R (and hence in q_C) while causing a quadratic increase in the pressure drop.

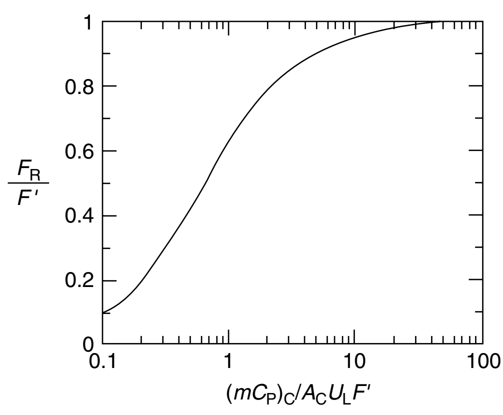


FIGURE 20.3 Variation of F_R/F' as a function of $[(mc_p)_c/(A_c U_L F')]$. (From Duffie, J. A. and Beckman, W. A., *Solar Engineering of Thermal Processes*, Wiley Interscience, New York, 1980.)

Factors influencing solar collector performance are of three types: (i) constructional, that is, related to collector design and materials used, (ii) climatic, and (iii) operational, that is, fluid temperature, flow rate, and so on. The plate efficiency factor F' is a factor that depends on the physical constructional features and is essentially a constant for a given liquid collector. (This is not true for air collectors, which require more careful analysis.) Operational features involve changes in m_C and T_{Ci}. While changes in m_C affect F_R as per Equation 20.3, we note from Equation 20.2 that to enhance q_C, T_{Ci} needs to be kept as low as possible. For solar collectors that are operated under more or less constant flow rates, specifying $F_R\eta_0$ and F_RU_L is adequate to predict collector performance under varying climatic conditions.

There are a number of procedures by which collectors have been tested. The most common is a *steady-state procedure*, where transient effects due to collector heat capacity are minimized by performing tests only during periods when radiation and ambient temperature are steady. The procedure involves simultaneous and accurate measurements of the mass flow rate, the inlet and outlet temperatures of the collector fluid, and the ambient conditions (incident solar radiation, air temperature, and wind speed). The most widely used test procedure is the ASHRAE Standard 93–77 (1978), whose test setup is shown in Figure 20.4. Though a solar simulator can be used to perform indoor testing, outdoor testing is always more realistic and less expensive. The procedure can be used for nonconcentrating collectors using air or liquid as the working fluid (but not two phase mixtures) that have a single inlet and a single outlet and contain no integral thermal storage.

Steady-state procedures have been in use for a relatively long period and though the basis is very simple the engineering setup is relatively expensive (see Figure 20.4). From an overall heat balance on the collector fluid and from Equation 20.2, the expressions for the instantaneous collector efficiency under normal solar incidence are

$$\eta_C \equiv \frac{q_C}{A_C I_T} = \frac{(mc_p)_C(T_{Co} - T_{Ci})}{A_C I_T} \tag{20.4}$$

$$= \left[F_R \eta_n - F_R U_L \left(\frac{T_{Ci} - T_a}{I_T} \right) \right] \tag{20.5}$$

where η_n is the optical efficiency at normal solar incidence.

From the test data, points of η_c against reduced temperature $[(T_{Ci} - T_a)/I_T]$ are plotted as shown in Figure 20.5. Then a linear fit is made to these data points by regression, from which the values of $F_R\eta_n$ and F_RU_L are easily deduced. It will be noted that if the reduced term were to be taken as $[(T_{Cm} - T_a)/I_T]$, estimates of $F'\eta_n$ and $F'U_L$ would be correspondingly obtained.

20.1.2.2.3 Incidence Angle Modifier

The optical efficiency η_0 depends on the collector configuration and varies with the angle of incidence as well as with the relative values of diffuse and beam radiation. The incidence angle modifier is defined as $K_\eta \equiv (\eta_0/\eta_n)$. For flat-plate collectors with 1 or 2 glass covers, K_η is almost unchanged up to incidence angles of 60°, after which it abruptly drops to zero.

A simple way to model the variation of K_η with incidence angle for flat-plate collectors is to specify η_n, the optical efficiency of the collector at normal beam incidence, to assume the entire radiation to be beam, and to use the following expression for the angular dependence (ASHRAE 1978)

$$K_\eta = 1 + b_0 \left(\frac{1}{\cos\theta} - 1 \right) \tag{20.6}$$

where θ is the solar angle of incidence on the collector plate (in degrees) and b_0 is a constant called the incidence angle modifier coefficient. Plotting K_η against $[(1/\cos\theta) - 1]$ results in linear plots (see Figure 20.6), thus justifying the use of Equation 20.6. We note that for one-glass and two-glass covers, approximate values of b_0 are -0.10 and -0.17, respectively.

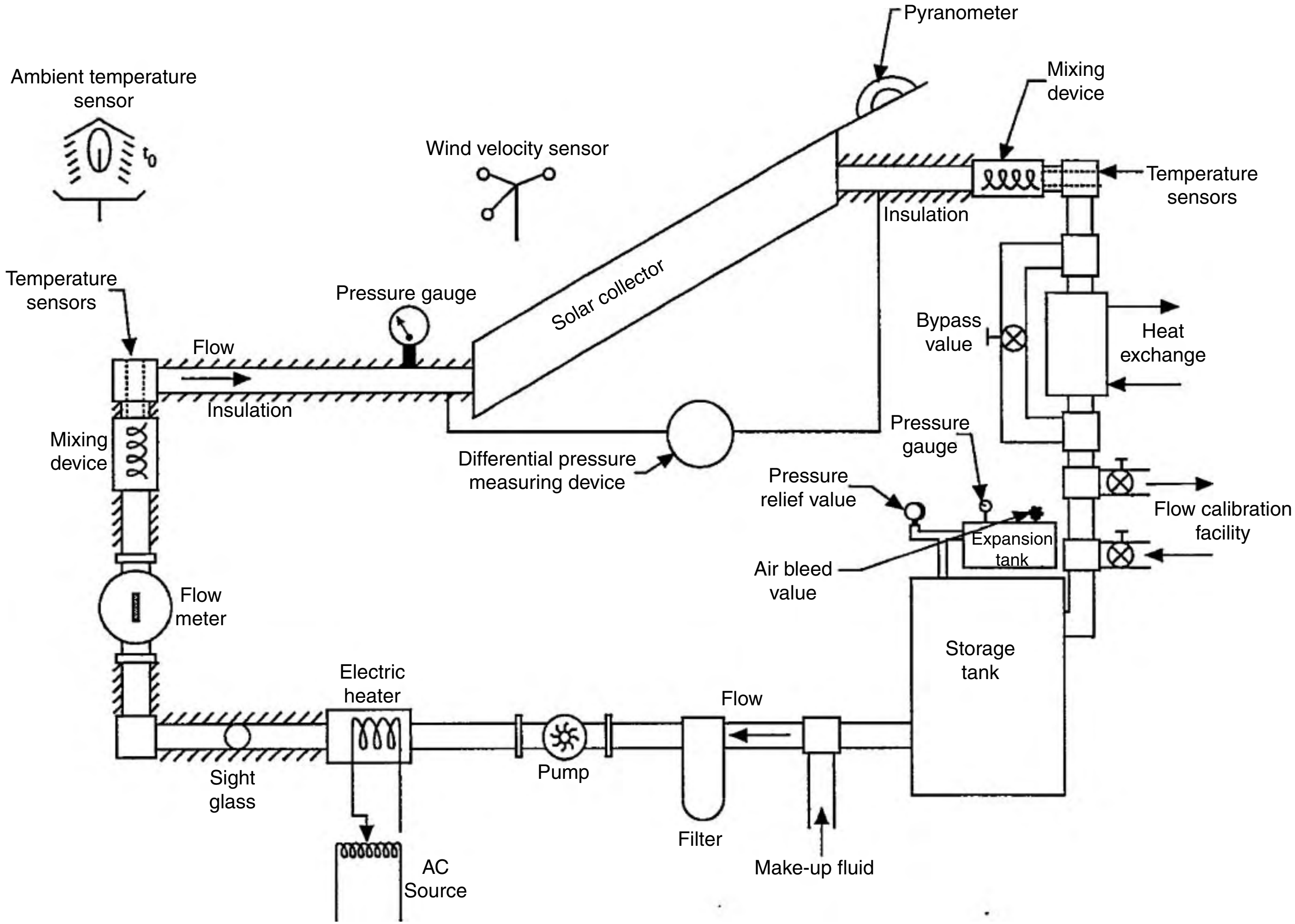


FIGURE 20.4 Set up for testing liquid collectors according to ASHRAE Standard 93–72.

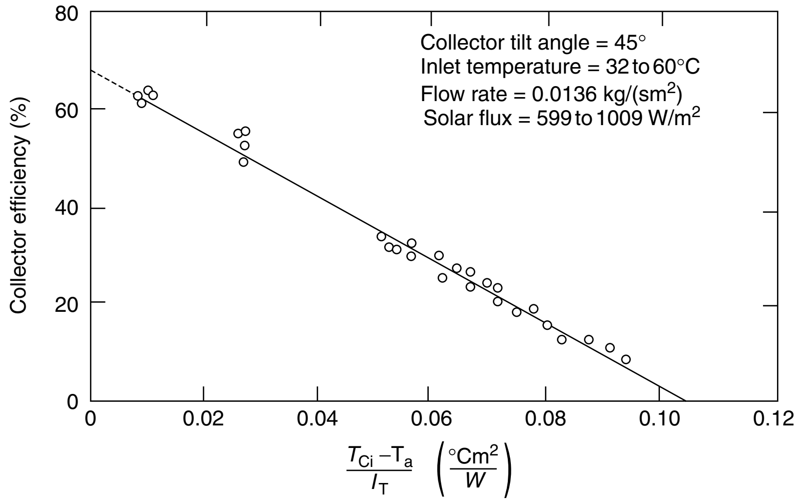


FIGURE 20.5 Thermal efficiency curve for a double glazed flat-plate liquid collector. Test conducted outdoors on a 1.2 m by 1.25 m panel with 10.2 cm of glass fiber back insulation and a flat copper absorber with black coating of emissivity of 0.97. (From ASHRAE Standard 93–77, *Methods of Testing to Determine the Thermal Performance of Solar Collectors*, American Society of Heating, Refrigeration and Air Conditioning Engineers, New York, 1978.)

In case the diffuse solar fraction is high, one needs to distinguish between beam, diffuse, and ground-reflected components. Diffuse radiation, by its very nature, has no single incidence angle. One simple way is to assume an equivalent incidence angle of 60° for diffuse and ground-reflected components. One would then use Equation 20.6 for the beam component along with its corresponding value of θ and account for the contribution of diffuse and ground reflected components by assuming a value of $\theta = 60°$ in Equation 20.6. For more accurate estimation, one can use the relationship between the effective diffuse solar incidence angle versus collector tilt given in Duffie and Beckman (1980). It should be noted that the preceding equation gives misleading results with incidence angles close to 90°. An alternative functional form for the incidence angle modifier for both flat-plate and concentrating collectors has been proposed by Rabl (1981).

Example 20.1.1

From the thermal efficiency curve given in Figure 20.5 determine the performance parameters of the corresponding solar collector.

Extrapolating the curve yields y-intercept = 0.69, x-intercept = 0.105 (m^{2}°C/W). Since the reduced temperature in Figure 20.5 is in terms of the inlet fluid temperature to the collector, Equation 20.5 yields $F_R\eta_n = 0.69$ and $F_R U_L = 0.69/0.105 = 6.57$ W/(m^{2}°C). Alternatively, the collector parameters in terms of the plate efficiency factor can be deduced. From Figure 20.5, the collector area $= 1.22 \times 1.25 = 1.525$ m^2, while the flow rate $(m/A_C) = 0.0136$ kg/(s m^2). From Equation 20.3,

$$F'/F_R = -(0.0136 \times 4190/6.57)\ln[-6.57/(0.0136 \times 4190)] = 1.0625$$

Thus $F'U_L = 6.57 \times 1.0625 = 6.98$ W/(m^{2}°C) and $F'\eta_n = 0.69 \times 1.0625 = 0.733$.

Example 20.1.2

How would the optical efficiency be effected at a solar incidence angle of 60° for a flat-plate collector with two glass covers?

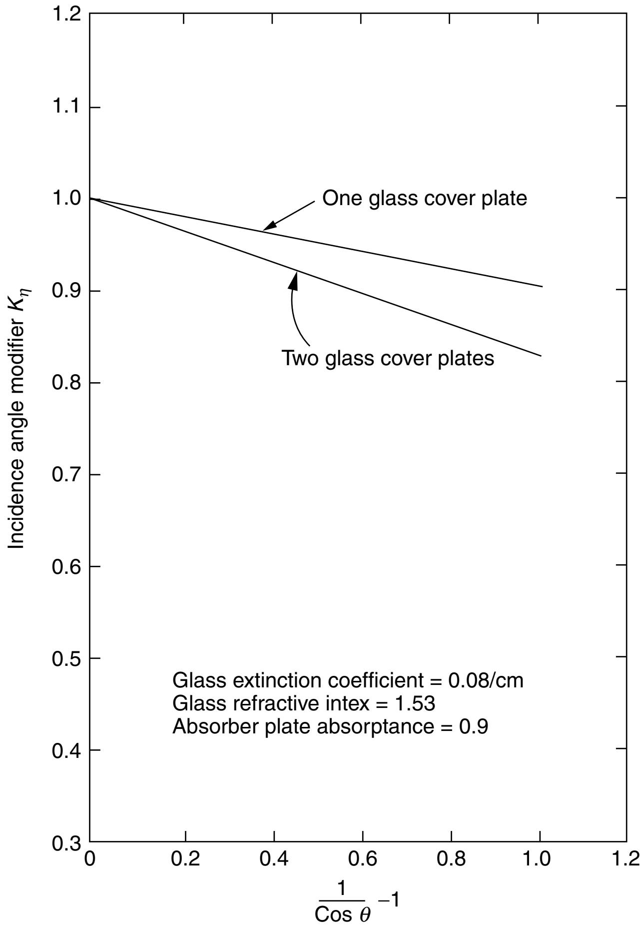


FIGURE 20.6 Incidence angle modifiers for two flat-plate collectors with nonselective coating on the absorber. (Adapted From ASHRAE Standard 93–77, *Methods of Testing to Determine the Thermal Performance of Solar Collectors*, American Society of Heating, Refrigeration and Air Conditioning Engineers, New York, 1978.)

Assume a value of $b_0 = -0.17$. From Equation 20.6, $K_\eta = 0.83$. Thus

$$F_R\eta_0 = F_R\eta_n K_\eta = 0.69 \times 0.83 = 0.57$$

20.1.2.2.4 Other Collector Characteristics

There are three collector characteristics that a comprehensive collector testing process should also address. The collector *time constant* is a measure that determines how intermittent sunshine affects collector performance and is useful in defining an operating control strategy for the collector array that avoids instability. Collector performance is usually enhanced if collector time constants are kept low. ASHRAE 93–77 also includes a method for determining this value. Commercial collectors usually have time constants of about 5 min or less, and this justifies the use of the HWB model (see Equation 20.2).

Another quantity to be determined from collector tests is the collector *stagnation temperature*. This is the equilibrium temperature reached by the absorber plate when no heat is being extracted from the collector. Determining the maximum stagnation temperature, which occurs under high I_T and T_a values, is useful in order to safeguard against reduced collector life due to thermal damage to collectors (namely irreversible thermal expansion, sagging of covers, physical deterioration, optical changes, etc.) in the field when not in use. Though the stagnation temperature could be estimated from Equation 20.2 by setting $q_C = 0$ and solving for T_{Ci}, it is better to perform actual tests on collectors before field installation.

The third collector characteristic of interest is the *pressure drop* across the collector for different fluid flow rates. This is an important consideration for liquid collectors, and more so for air collectors, in order to keep parasitic energy consumption (namely electricity to drive pumps and blowers) to a minimum in large collector arrays.

20.1.2.3 Improvements to Flat-Plate Collector Performance

There are a number of ways by which the performance of the basic flat-plate collectors can be improved. One way is to enhance optical efficiency by treatment of the glass cover thereby reducing reflection and enhancing performance. As much as a 4% increase has been reported (Anderson 1977). Low-iron glass can also reduce solar absorption losses by a few percent.

These improvements are modest compared to possible improvements from reducing losses from the absorber plate. Essentially, the infrared upward reradiation losses from the heated absorber plate have to be decreased. One could use a second glass cover to reduce the losses, albeit at the expense of higher cost and lower optical efficiency. Usually for water heating applications, radiation accounts for about two-thirds of the losses from the absorber to the cover with convective losses making up the rest (conduction is less than about 5%). The most widely used manner of reducing these radiation losses is to use selective surfaces whose emissivity varies with wavelength (as against matte-black painted absorbers, which are essentially gray bodies). Note that 98% of the solar spectrum is at wavelengths less than 3.0 μm, whereas less than 1% of the black body radiation from a 200°C surface is at wavelengths less than 3.0 μm. Thus selective surfaces for solar collectors should have high-solar absorptance (i.e., low reflectance in the solar spectrum) and low long-wave emittance (i.e., high reflectance in the long-wave spectrum). The spectral reflectance of some commonly used selective surfaces is shown in Figure 20.7. Several commercial collectors for water heating or low-pressure steam (for absorption cooling or process heat applications) are available that use selective surfaces.

Another technique to simultaneously reduce both convective and radiative losses between the absorber and the transparent cover is to use honeycomb material (Hollands 1965). The honeycomb material can be reflective or transparent (the latter is more common) and should be sized properly. Glass honeycombs have had some success in reducing losses in high-temperature concentrating receivers, but plastics are

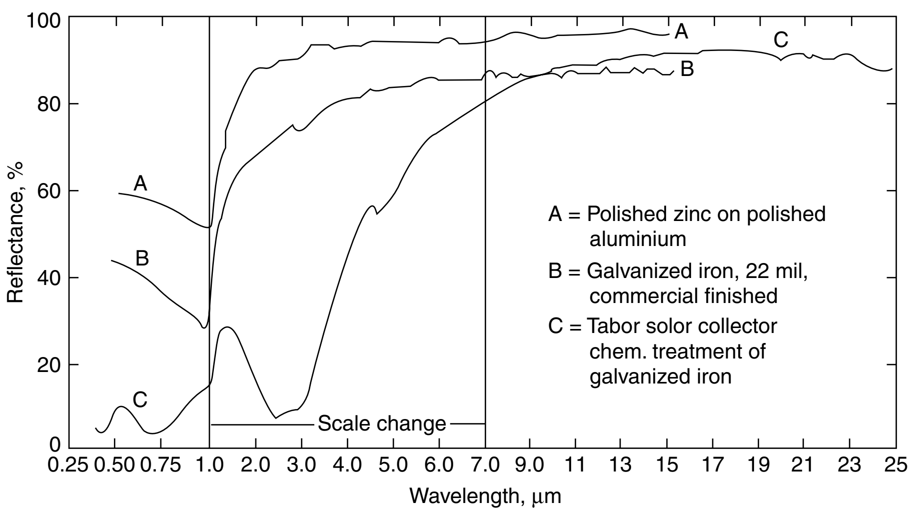


FIGURE 20.7 Spectral reflectance of several surfaces. (From Edwards, D. K., Nelson, K. E., Roddick, R. D., and Gier, J. T., Basic Studies on the Use of Solar Energy, Report no. 60–93, Department of Engineering. University of California at Los Angeles, CA, 1960.)

usually recommended for use in flat-plate collectors. Because of the poor thermal aging properties, honeycomb flat-plate collectors have had little commercial success. Currently the most promising kind seems to be the simplest (both in terms of analysis and construction), namely collectors using horizontal rectangular slats (Meyer 1978). Convection can be entirely suppressed provided the slats with the proper aspect ratio are used.

Finally, collector output can be enhanced by using side reflectors, for instance a sheet of anodized aluminum. The justification in using these is their low cost and simplicity. For instance, a reflector placed in front of a tilted collector cannot but increase collector performance because losses are unchanged and more solar radiation is intercepted by the collector. Reflectors in other geometries may cast a shadow on the collector and reduce performance. Note also that reflectors would produce rather nonuniform illumination over the day and during the year, which, though not a problem in thermal collectors, may drastically penalize the electric output of photovoltaic modules. Whether reflectors are cost-effective depends on the particular circumstances and practical questions such as aesthetics and space availability. The complexity involved in the analysis of collectors with planar reflectors can be reduced by assuming the reflector to be long compared to its width and treating the problem in two dimensions only. How optical performance of solar collectors are affected by side planar reflectors is discussed in several papers, for example Larson (1980), Chiam (1981).

20.1.2.4 Other Collector Types

20.1.2.4.1 Evacuated Tubular Collectors

One method of obtaining temperatures between 100 and 200°C is to use evacuated tubular collectors. The advantage in creating and being able to maintain a vacuum is that convection losses between glazing

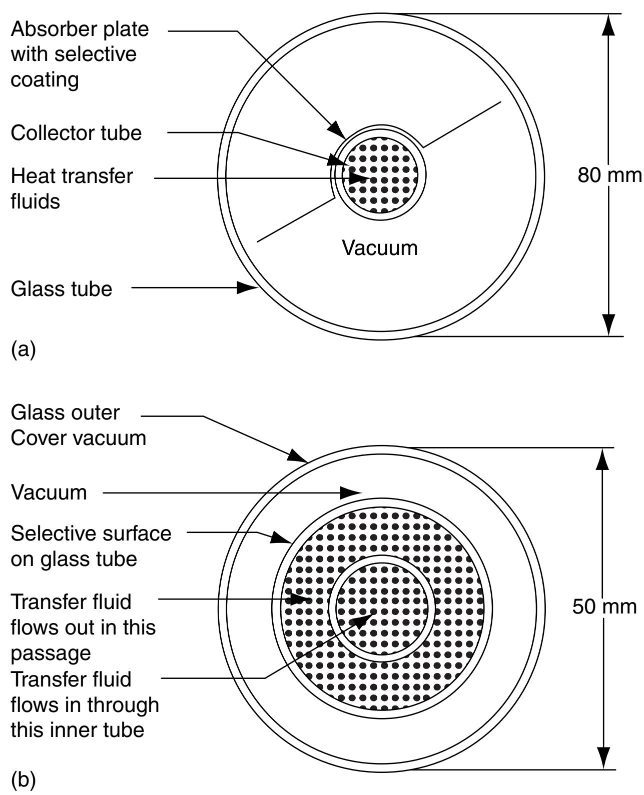


FIGURE 20.8 Evacuated tubular collectors. (From Charters, W. W. S. and Pryor, T. L., *An Introduction to the Installation of Solar Energy Systems*, Victoria Solar Energy Council, Melbourne, Australia, 1982.)

and absorber can be eliminated. There are different possible arrangements of configuring evacuated tubular collectors. Two designs are shown in Figure 20.8. The first is like a small flat-plate collector with the liquid to be heated making one pass through the collector tube. The second uses an all-glass construction with the glass absorber tube being coated selectively. The fluid being heated passes up the middle of the absorber tube and then back through the annulus. Evacuated tubes can collect both direct and diffuse radiation and do not require tracking. Glass breakage and leaking joints due to thermal expansion are some of the problems which have been experienced with such collector types. Various reflector shapes (like flat-plate, V-groove, circular, cylindrical, involute, etc.) placed behind the tubes are often used to usefully collect some of the solar energy, which may otherwise be lost, thus providing a small amount of concentration.

20.1.2.4.2 Compound Parabolic Concentrators

The CPC collector, discovered in 1966, consists of parabolic reflectors that funnel radiation from aperture to absorber rather than focusing it. The right and left halves belong to different parabolas (hence the name *compound*) with the edges of the receiver being the foci of the opposite parabola (see Figure 20.9). It has been proven that such collectors are *ideal* in that any solar ray, be it beam or diffuse, incident on the aperture within the acceptance angle will reach the absorber while all others will bounce back to and fro and reemerge through the aperture. CPCs are also called *nonimaging* concentrators because they do not form clearly defined images of the solar disk on the absorber surface as achieved in classical concentrators. CPCs can be designed both as low-concentration devices with large acceptance angles or as high-concentration devices with small acceptance angles. CPCs with low-concentration ratios (of about 2) and with east–west axes can be operated as stationary devices throughout the year or at most

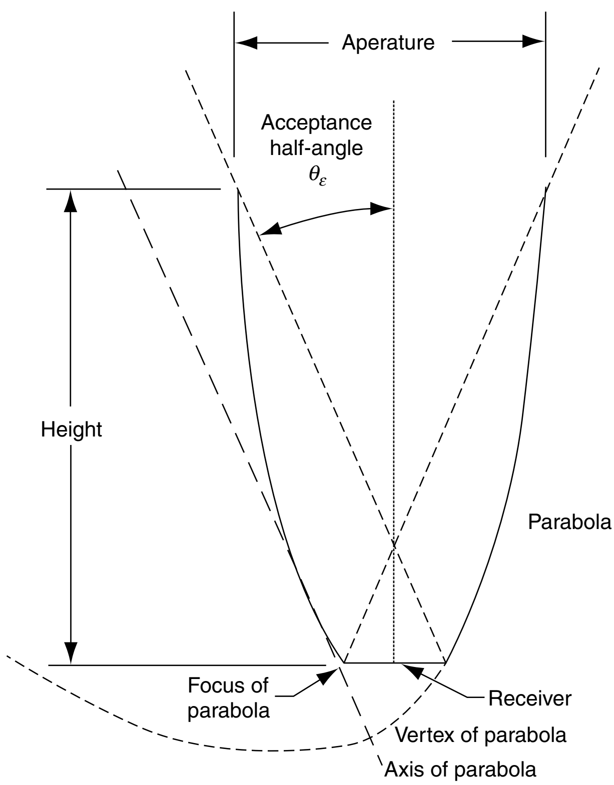


FIGURE 20.9 Cross-section of a symmetrical nontruncated CPC. (From Duffie, J. A. and Beckman, W. A., *Solar Engineering of Thermal Processes*, Wiley Interscience, New York, 1980.)

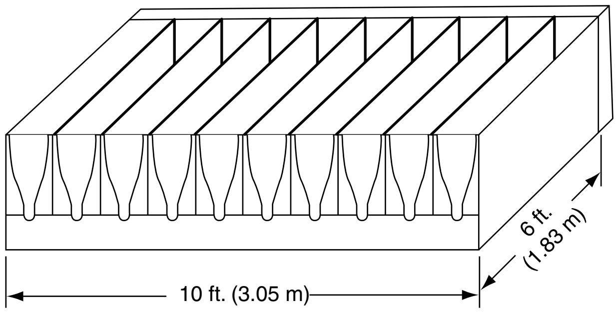


FIGURE 20.10 A CPC collector module. (From SERI, *Engineering Principles and Concepts for Active Solar Systems*, Hemisphere Publishing Company, New York, 1989.)

with seasonal adjustments only. CPCs, unlike other concentrators, are able to collect all the beam and a large portion of the diffuse radiation. Also they do not require highly specular surfaces and can thus better tolerate dust and degradation. A typical module made up of several CPCs is shown in Figure 20.10. The absorber surface is located at the bottom of the trough, and a glass cover may also be used to encase the entire module. CPCs show considerable promise for water heating close to the boiling point and for low-pressure steam applications. Further details about the different types of absorber and receiver shapes used, the effect of truncation of the receiver and the optics, can be found in Rabl (1985).

20.1.3 Long-Term Performance of Solar Collectors

20.1.3.1 Effect of Day-to-Day Changes in Solar Insolation

Instantaneous or hourly performance of solar collectors has been discussed in "Flat-Plate Collectors." For example, one would be tempted to use the HWB Equation 20.2 to predict long-term collector performance at a prespecified and constant fluid inlet temperature T_{Ci} merely by assuming average hourly values of I_T and T_a. Such a procedure would be erroneous and lead to underestimation of collector output because of the presence of the control function, which implies that collectors are turned on only when $q_C > 0$, that is, when radiation I_T exceeds a certain critical value I_C. This critical radiation value is found by setting q_C in Equation 20.2 to zero:

$$I_C = U_L(T_{Ci} - T_a)/\eta_0 \tag{20.7a}$$

To be more rigorous, a small increment δ to account for pumping power and stability of controls can also be included if needed by modifying the equation to

$$I_C = U_L(T_{Ci} + \delta - T_a)/\eta_0 \tag{20.7b}$$

Then, Equation 20.2 can be rewritten in terms of I_C as

$$q_C = A_C F_R \eta_0 [I_T - I_C]^+ \tag{20.8}$$

Why one cannot simply assume a mean value of I_T in order to predict the mean value of q_C will be illustrated by the following simple concept (Klein 1978). Consider the three identical day sequences shown in sequence A of Figure 20.11. If I_{C1} is the critical radiation intensity and if it is constant over the whole day, the useful energy collected by the collector is represented by the sum of the shaded areas. If a higher critical radiation value shown as I_{C2} in Figure 20.11 is selected, we note that no useful energy is collected at all. Actual weather sequences would not look like that in sequence A but rather like that in sequence B, which is comprised of an excellent, a poor, and an average day. Even if both sequences have

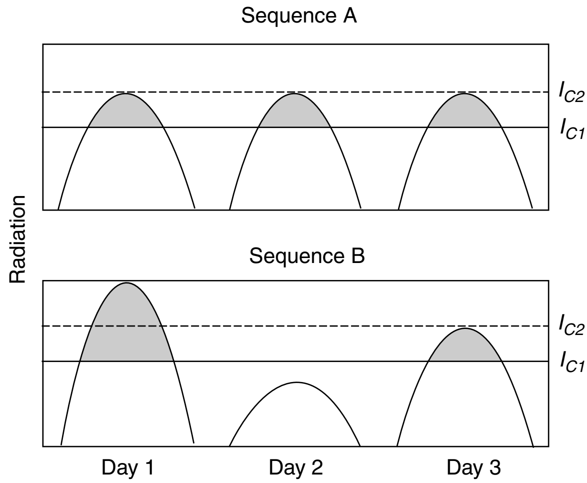


FIGURE 20.11 Effect of radiation distribution on collector long-term performance. (From Klein, S. A., Calculation of flat-plate collector utilizability, *Solar Energy*, 21, 393, 1978.)

the same average radiation over 3 days, a collector subjected to sequence B will collect useful energy when the critical radiation is I_{C2}. Thus, neglecting the variation of radiation intensity from day-to-day over the long term and dealing with mean values would result in an underestimation of collector performance.

Loads are to a certain extent repetitive from day-to-day over a season or even the year. Consequently, one can also expect collectors to be subjected to a known diurnal repetitive pattern or mode of operation, that is, the collector inlet temperature T_{Ci} has a known repetitive pattern.

20.1.3.2 Individual Hourly Utilizability

In this mode, T_{Ci} is assumed to very over the day but has the same variation for all the days over a period of N days (where $N=30$ days for monthly and $N=365$ for yearly periods). Then from Equation 20.8, *total* useful energy collected over N days during individual hour i of the day is

$$q_{CN}(i) = A_C F_R \bar{\eta}_0 \bar{I}_{Ti} \sum_{i=1}^{N} \frac{[I_{Ti} - I_C]^+}{\bar{I}_{Ti}} \tag{20.9}$$

Let us define the radiation ratio

$$X_i = I_{Ti}/\bar{I}_{Ti} \tag{20.10}$$

and the critical radiation ratio

$$X_C = I_C/\bar{I}_{Ti}$$

The modified HWB Equation 20.8 can be rewritten as

$$q_{CN}(i) = A_C F_R \bar{\eta}_0 \bar{I}_{Ti} N \phi_i(x_c) \tag{20.11}$$

where the individual hourly utilizability factor ϕ_i is identified as

$$\phi_i(X_C) = \frac{1}{N} \sum^{N} (X_i - X_C)^+ \tag{20.12}$$

Thus ϕ_i can be considered to be the fraction of the incident solar radiation that can be converted to useful heat by an ideal collector (i.e., whose $F_R\eta_0 = 1$). The utilizability factor is thus a *radiation statistic* in the sense that it depends solely on the radiation values at the specific location. As such, it is in no way dependent on the solar collector itself. Only after the radiation statistics have been applied is a collector dependent significance attached to X_C.

Hourly utilizability curves on a *monthly* basis that are independent of location were generated by Liu and Jordan (1963) over 30 years ago for flat-plate collectors (see Figure 20.12). The key climatic parameter which permits generalization is the *monthly clearness index* $\bar{K}$ of the location defined as

$$\bar{K} = \bar{H}/\bar{H}_0 \tag{20.13}$$

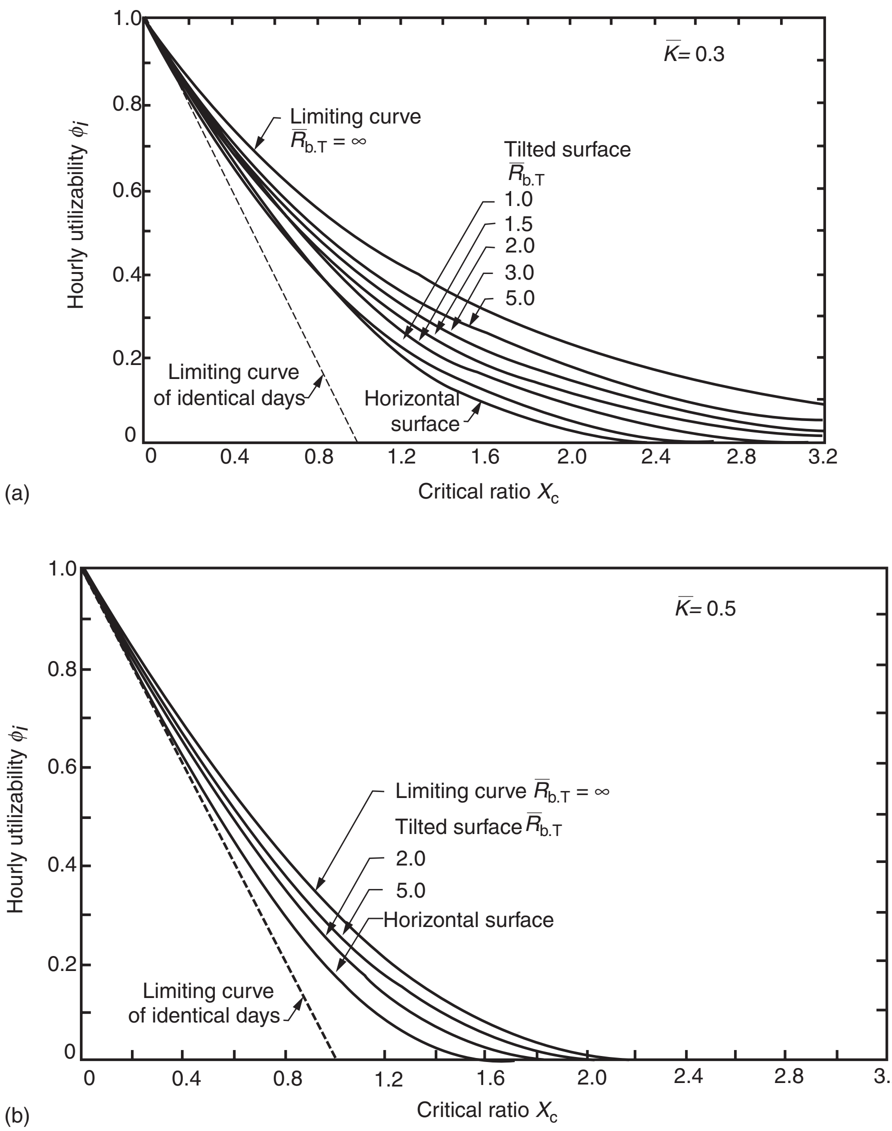


FIGURE 20.12 Generalized hourly utilizability curves of Liu and Jordan (1963) for three different monthly mean clearness indices K. (a) $\bar{K} = 0.3$, (b) $\bar{K} = 0.5$, (c) $\bar{K} = 0.7$. (From Liu, B. Y. H. and Jordan, R. C., A rational procedure for predicting the long-term average performance of flat-plate solar energy collectors, *Solar Energy*, 7, 53, 1963.)

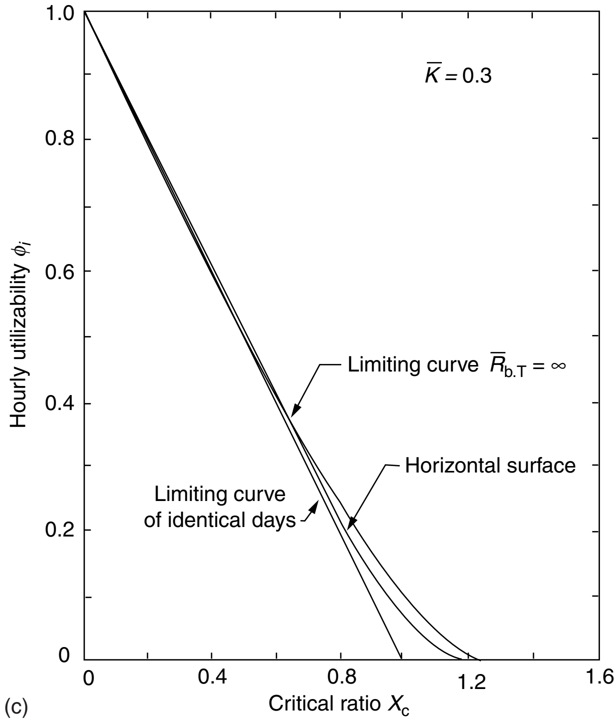


FIGURE 20.12 (*continued*)

where $\bar{H}$ is the monthly mean daily global radiation on the horizontal surface and $\bar{H}_0$ is the monthly mean daily extraterrestrial radiation on a horizontal surface.

Extensive tables giving monthly values of $\bar{K}$ for several different locations worldwide can be found in several books, for example, Duffie and Beckman (1980) or Reddy (1987). The curves apply to equator-facing tilted collectors with the effect of collector tilt accounted for by the factor $\bar{R}_{b,T}$ which is the ratio of the monthly mean daily extraterrestrial radiation on the tilted collector to that on a horizontal surface. Monthly mean daily calculations can be made using the 15th of the month, though better accuracy is achieved using slightly different dates (Reddy 1987). Clark, Klein, and Beckman (1983), working from measured data from several U.S. cities, have proposed the following correlation for individual hourly utilizability over monthly time scales applicable to flat-plate collectors only:

$$\begin{aligned} \phi_i &= 0 \quad \text{for } X_C \geq X_{max} \\ &= (1 - X_C/X_{max})^2 \quad \text{for } X_{max} = 2 \\ &= \left| |a| - [a^2 + (1 + 2a)(1 - X_C/X_{max})^2]^{1/2} \right| \quad \text{otherwise} \end{aligned} \tag{20.14}$$

where

$$a = (X_{max} - 1)(2 - X_{max}) \tag{20.15}$$

and

$$X_{max} = 1.85 + 0.169(\bar{r}_T/\bar{k}^2) - 0.0696 \cos \beta/\bar{k}^2 - 0.981\bar{k}/(\cos \delta)^2 \tag{20.16}$$

where $\bar{k}$ is the monthly mean *hourly* clearness index for the particular hour, δ is the solar declination, β is the tilt angle of the collector plane with respect to the horizontal, and $\bar{r}_T$ is the ratio of monthly average hourly global radiation on a tilted surface to that on a horizontal surface for that particular hour. For an isotropic sky assumption, $\bar{r}_T$ is given by

$$\bar{r}_T = (1 - \bar{I}_d \bar{I}) r_{b,T} + \left(\frac{1 + \cos\beta}{2}\right) \bar{I}_d \bar{I} + \left(\frac{1 - \cos\beta}{2}\right) \rho \tag{20.17}$$

where $\bar{I}_d$ and $\bar{I}$ are the hourly diffuse and global radiation on the horizontal surface, $r_{b,T}$ is the ratio of hourly beam radiation on the tilted surface to that on a horizontal surface (this is a purely astronomical quantity and can be calculated accurately from geometric considerations), and ρ is the ground albedo.

Example 20.1.3

Compute the total energy collected during 11:30–12:30 for the month of September in New York, NY (latitude: 40.75°N, $T_a = 20°C$) by a flat-plate solar collector of 5 m^2 area having zero tilt. The collector performance parameters are $F_R\eta_0 = 0.54$ and $F_R U_L = 3.21$ W/(m^{2}°C) and the collector inlet temperature is 80°C. The corresponding hourly mean clearness index $\bar{k}$ is 0.44, and the monthly mean hourly radiation on a horizontal surface $\bar{I}_{Ti}$ (11:30–12:30) is 6.0 MJ/(m^2 h).

From Equation 20.7a, critical radiation $I_C = 3.21 \times (80 - 20)/0.54 = 356.7$ W/m^2 = 1.28 MJ/(m^2 h). For the average day of September, solar declination $\delta = 2.2°$. Also, because the collector is horizontal $\bar{r}_T = 1$ and $\beta = 0$. Thus from Equation 20.16

$$X_{max} = 1.85 + 0.169/0.44^2 - 0.0696/0.44^2 - 0.981 \times 0.44/(\cos 2.2)^2 = 1.93.$$

Also from Equation 20.15, $a = (1.93 - 1)/(2 - 1.93) = 13.29$.
The critical radiation ratio $X_C = 1.28/1.93 = 0.663$.
Because $X_C < X_{max}$, from Equation 20.14 we have

$$\phi_i(X_C) = \left|13.29 - [13.29^2 + (1 + 2 \times 13.29)(1 - 0.663/1.93)^2]^{1/2}\right| = |13.29 - 13.73| = 0.44.$$

Finally, the total energy collected is given by Equation 20.11

$$q_{CN}(11:30\text{–}12:30) = 5 \times 0.54 \times 60 \times 30 \times 0.44 = 214 \text{ MJ/h}$$

20.1.3.3 Daily Utilizability

20.1.3.3.1 Basis

In this mode, T_{Ci}, and hence the critical radiation level, is assumed constant during all hours of the day. The *total* useful energy over N days that can be collected by solar collectors operated all day over n hours is given by

$$Q_{CN} = A_C F_R \bar{\eta}_0 \bar{H}_T N \bar{\phi} \tag{20.18}$$

where $\bar{H}_T$ is the average daily global radiation on the collector surface, and $\bar{\phi}$ (called Phibar) is the daily utilizability factor, defined as

$$\bar{\phi} = \sum^{N} \sum^{n} (I_T - I_C)^+ / \sum^{N} \sum^{n} I_T = \frac{1}{Nn} \sum^{N} \sum^{n} (X_i - X_C)^+ \tag{20.19}$$

Generalized correlations have been developed both at monthly time scales and for annual time scales based on the parameter $\bar{K}$. Generalized (i.e., location and month independent) correlations for $\bar{\phi}$ on a *monthly* time scale have been proposed by Theilacker and Klein (1980). These are strictly applicable for flat-plate collectors only. Collares-Pereira and Rabl (1979) have also proposed generalized correlations for $\bar{\phi}$ on a monthly time scale which, though a little more tedious to use are applicable to concentrating collectors as well. The reader may refer to Rabl (1985) or Reddy (1987) for complete expressions.

20.1.3.3.2 Monthly Time Scales

The Phibar method of determining the daily utilizability fraction proposed by Theilacker and Klein (1980) correlates $\bar{\phi}$ to the following factors:

1. A geometry factor $\bar{R}_T/\bar{r}_{T,noon}$, which incorporates the effects of collector orientation, location, and time of year. $\bar{R}_T$ is the ratio of monthly average global radiation on the tilted surface to that on a horizontal surface. $\bar{r}_{T,noon}$ is the ratio of radiation at noon on the tilted surface to that on a horizontal surface for the average day of the month. Geometrically, $\bar{r}_{T,noon}$ is a measure of the maximum height of the radiation curve over the day, whereas $\bar{R}_T$ is a measure of the enclosed area. Generally the value $(\bar{R}_T/\bar{r}_{T,noon})$ is between 0.9 and 1.5.
2. A dimensionless critical radiation level $\bar{X}_{C,K}$ where

$$\bar{X}_{C,K} = I_C/\bar{I}_{T,noon} \tag{20.20}$$

with $\bar{I}_{T,noon}$, the radiation intensity on the tilted surface at noon, given by

$$\bar{I}_{T,noon} = \bar{r}_{noon}\bar{r}_{T,noon}\bar{H} \tag{20.21}$$

where $\bar{r}_{noon}$ is the ratio of radiation at noon to the daily global radiation on a horizontal surface during the mean day of the month which can be calculated from the following correlation proposed by Liu and Jordan (1960):

$$r(W) = \frac{I(W)}{H} = \frac{\pi}{24}(a + b\cos W)\frac{(\cos W - \cos W_S)}{\left(\sin W_S - \frac{\pi}{180}W_S\cos W_S\right)} \tag{20.22}$$

with

$$a = 0.409 + 0.5016\sin(W_S - 60)$$
$$b = 0.6609 - 0.4767\sin(W_S - 60)$$

where W is the hour angle corresponding to the midpoint of the hour (in degrees) and W_S is the sunset hour angle given by

$$\cos W_S = -\tan L \tan\delta \tag{20.23}$$

where L is the latitude of the location. The fraction r is the ratio of hourly to daily global radiation on a horizontal surface. The factors $\bar{r}_{T,noon}$ and $\bar{r}_{noon}$ can be determined from Equation 20.17 and Equation 20.22, respectively, with $W = 0°$.

The Theilacker and Klein correlation for the daily utilizability for equator-facing flat-plate collectors is

$$\bar{\phi}(X_{C,K}) = \exp\{[a' + b'(\bar{r}_{T,noon}/\bar{R}_T)][X_{C,K} + c'X_{C,K}^2]\} \tag{20.24}$$

where

$$\begin{aligned} a' &= 7.476 - 20.00\bar{K} + 11.188\bar{K}^2 \\ b' &= -8.562 + 18.679\bar{K} - 9.948\bar{K}^2 \\ c' &= -0.722 + 2.426\bar{K} + 0.439\bar{K}^2 \end{aligned} \tag{20.25}$$

How $\bar{\phi}$ varies with the critical radiation ratio $\bar{X}_{C,K}$ for three different values of $\bar{K}$ is shown in Figure 20.13.

Example 20.1.4

A flat-plate collector operated horizontally at Fort Worth, Texas ($L = 32.75°N$), has a surface area of 20 m^2. It is used to heat 10 kg/min. of water entering the collector at a constant temperature of 80°C each day from 6 a.m. to 6 p.m. The collector performance parameters are $F_R\eta_0 = 0.70$ and $F_RU_L = 5.0$ W/(m^{2}°C). Use Klein's correlation to compute the energy collected by the solar collectors during September.

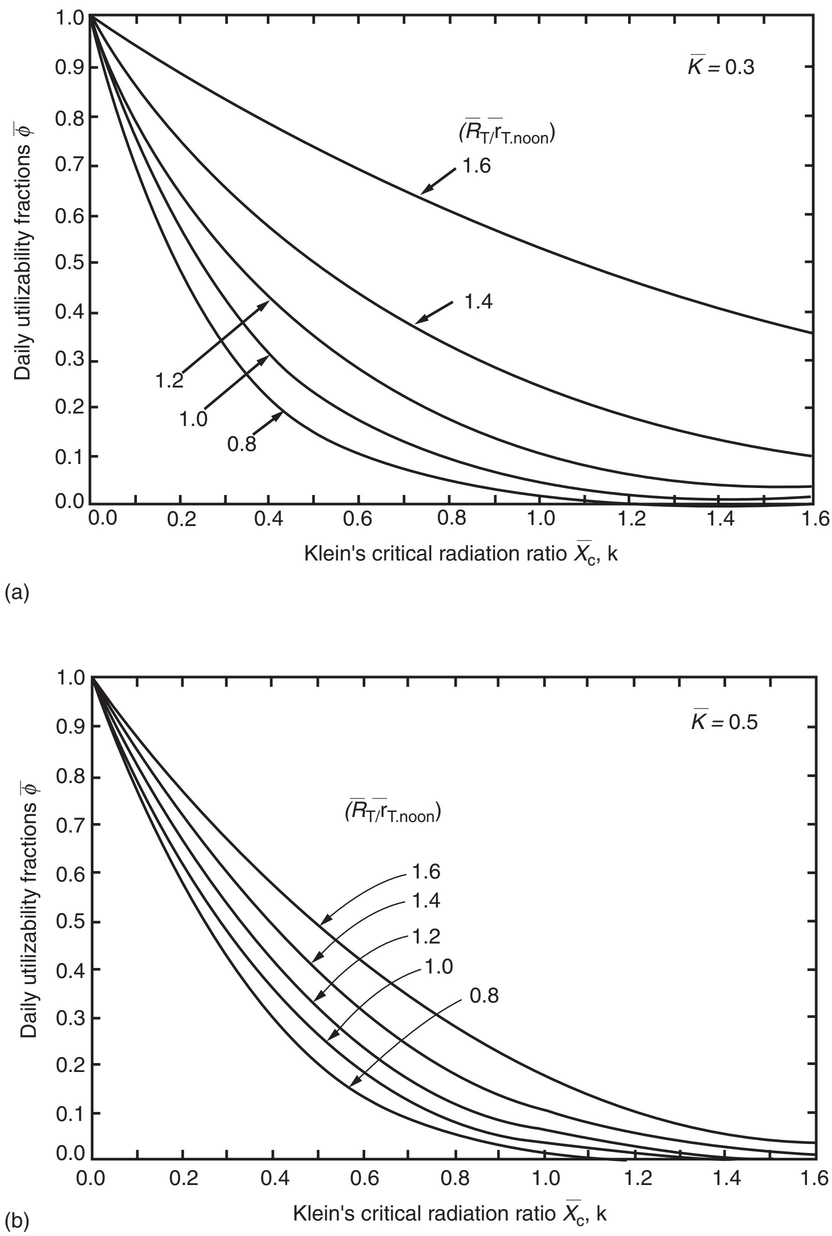


FIGURE 20.13 Generalized daily utilizability curves of Theilacker and Klein (1980) for three different K values. (a) $\bar{K} = 0.3$, (b) $\bar{K} = 0.5$, (c) $\bar{K} = 0.7$. (From Theilacker, J. C. and Klein, S. A., Improvements in the utilizability relationships, *American Section of the International Solar Energy Society Meeting Proceedings*, p. 271. Phoenix, AZ, 1980.)

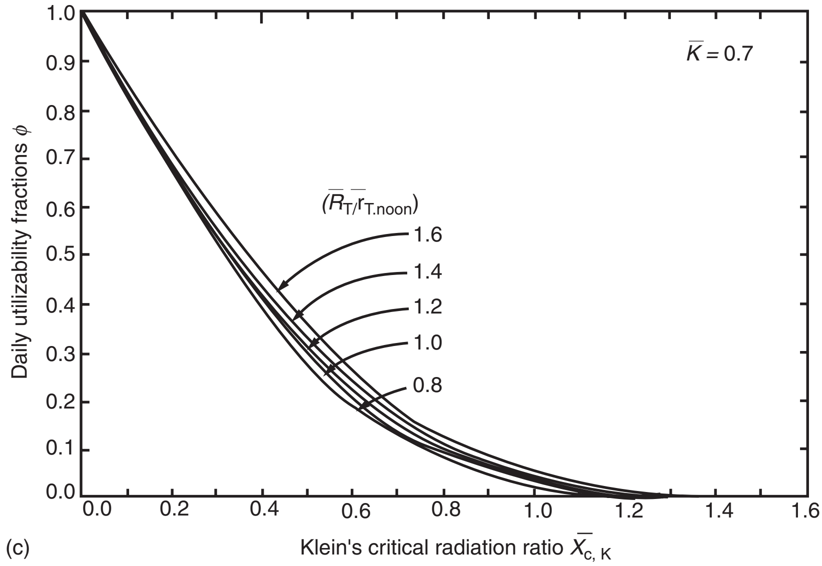


FIGURE 20.13 (*continued*)

Assume $\bar{H} = 18.28\ \text{MJ/(m}^2\text{d)}$, $\bar{K} = 0.57$ and $\bar{T}_a = 25°\text{C}$. Assume the mean sunset hour angle for September to be 90°.

The critical radiation is calculated first:

$$I_C = (5/0.7)(80 - 25) = 393\ \text{W/m}^2 = 1.414\ \text{MJ/(m}^2\text{h)}$$

For a horizontal surface, $\bar{R}_T = \bar{r}_{T,noon} = 1$. From Equation 20.22, $r(W = 0) = \pi/24(a + b) = 0.140$. Klein's critical radiation ratio (20.1.20) $\bar{X}_{C,K} = 1.414/(18.28 \times 0.140) = 0.553$. From Equation 20.24, $\bar{\phi} = 0.318$. Finally, from Equation 20.18, the total monthly energy collected by the solar collectors is $Q_{CM} = 20 \times 0.7 \times 30 \times 0.318 \times 18.28 = 2.44$ GJ/month.

20.1.3.3.3 Annual Time Scales

Generalized expressions for the *yearly* average energy delivered by the principal collector types with constant radiation threshold (i.e., when the fluid inlet temperature is constant for all hours during the day over the entire year) have been developed by Rabl (1981) based on data from several U.S. locations. The correlations are basically quadratic of the form

$$\frac{Q_{CY}}{A_C F_R \eta_n} = \tilde{a} + \tilde{b} I_C + \tilde{c} I_C^2 \qquad (20.26)$$

where the coefficients $\tilde{a}$, $\tilde{b}$, and $\tilde{c}$ are functions of collector type and/or tracking mode, climate, and in some cases, latitude. The complete expressions as revised by Gordon and Rabl (1982) are given in Reddy (1987). Note that the yearly *daytime* average value of T_a should be used to determine I_C. If this is not available, the yearly mean *daily* average value can be used. Plots of Q_{CY} versus I_C for flat-plate collectors that face the equator with tilt equal to the latitude are shown in Figure 20.14. The solar radiation enters these expressions as $\tilde{I}_{bn}$, the annual average beam radiation at normal incidence. This can be estimated from the following correlation

$$\tilde{I}_{bn} = 1.37\tilde{K} - 0.34 \qquad (20.27)$$

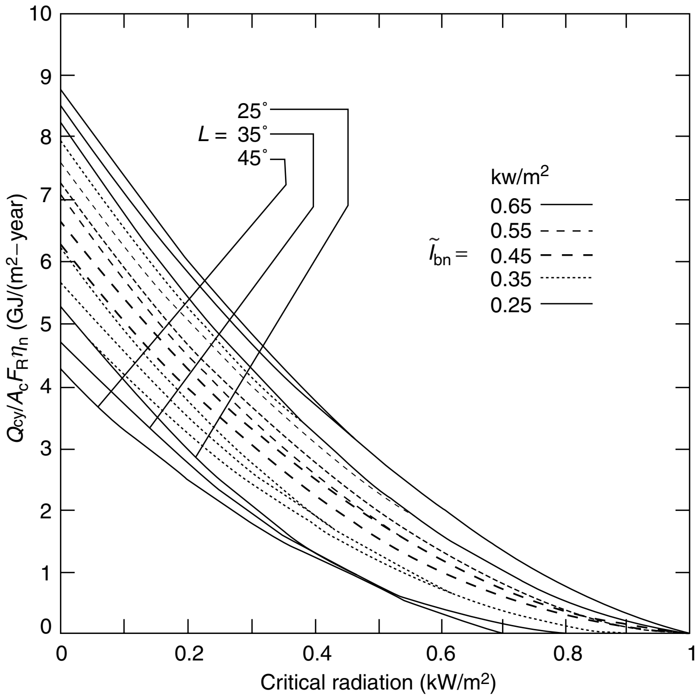


FIGURE 20.14 Yearly total energy delivered by flat-plate collectors with tilt equal to latitude. (From Gordon, J. M. and Rabl, A., Design, analysis and optimization of solar industrial process heat plants without storage, *Solar Energy*, 28, 519, 1982.)

where $\tilde{I}_{bn}$ is in kW/m^2 and $\tilde{K}$ is the annual average clearness index of the location. Values of $\tilde{K}$ for several locations worldwide are given in Reddy (1987).

This correlation is strictly valid for latitudes ranging from 25 to 48°. If used for lower latitudes, the correlation is said to lead to overprediction. Hence, it is recommended that for such lower latitudes a value of 25° be used to compute Q_{Cy}.

A direct comparison of the yearly performance of different collector types is given in Figure 20.15 (from Rabl 1981). A latitude of 35°N is assumed and plots of Q_{Cy} U.S. ($T_{Ci} - T_a$) have been generated in a sunny climate with $l_{bn} = 0.6$ kW/m^2. Relevant collector performance data are given in Figure 20.15. The crossover point between flat-plate and concentrating collectors is approximately 25°C above ambient temperature whether the climate is sunny or cloudy.

20.1.4 Solar Systems

20.1.4.1 Classification

Solar thermal systems can be divided into two categories: standalone or solar supplemented. They can be further classified by means of energy collection as active or passive, by their use as residential or industrial. Further, they can be divided by collector type into liquid or air systems, and by the type of storage they use into seasonal or daily systems.

20.1.4.1.1 Standalone and Solar Supplemented Systems

Standalone systems are systems in which solar energy is the only source of energy input used to meet the required load. Such systems are normally designed for applications where a certain amount of tolerance is permissible concerning the load requirement; in other words, where it is not absolutely imperative that

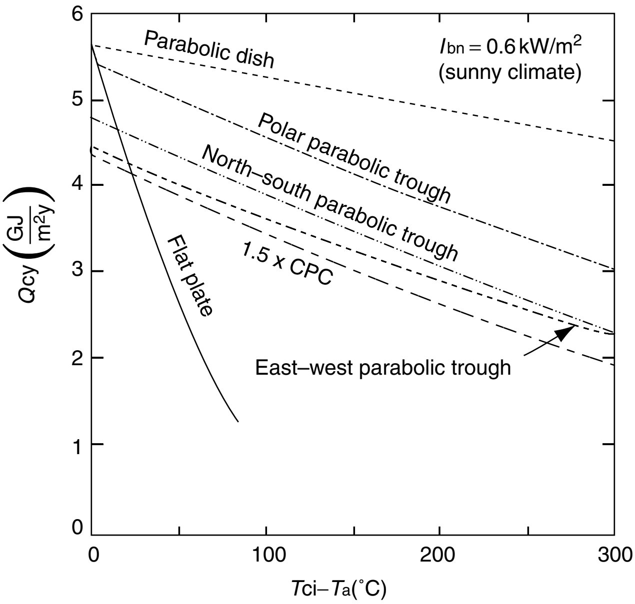


FIGURE 20.15 Figure illustrating the comparative performance (yearly collectible energy) of different collector types as a function of the difference between collector inlet temperature and ambient collector performance parameters $F'\eta_0$ and $F'U_L$ in W/(m^2 °C) are: flat plate (0.70 and 5.0), CPC (0.60 and 0.75), parabolic trough (0.65 and 0.67), and parabolic dish (0.61 and 0.27). (From Rabl, A., Yearly average performance of the principal solar collector types, *Solar Energy*, 27, 215, 1981.)

the specified load be met each and every instant. This leniency is generally admissible in the case of certain residential and agricultural applications. The primary reasons for using such systems are their low cost and simplicity of operation.

Solar-supplemented systems, widely used for both industrial and residential purposes, are those in which solar energy supplies part of the required heat load, the rest being met by an auxiliary source of heat input. Due to the daily variations in incident solar radiation, the portion of the required heat load supplied by the solar energy system may vary from day-to-day. However, the auxiliary source is so designed that at any instant it is capable of meeting the remainder of the required heat load. It is normal practice to incorporate an auxiliary heat source large enough to supply the entire heat load required. Thus, the benefit in the solar subsystem is not in its capacity credit (i.e., not that a smaller capacity conventional system can be used), but rather that a part of the conventional fuel consumption is displaced. The solar subsystem thus acts as a fuel economizer.

Solar-supplemented energy systems will be the primary focus of this chapter. Designing such systems has acquired a certain firm scientific rationale, and the underlying methodologies have reached a certain maturity and diversity, which may satisfy professionals from allied fields. On the other hand, unitary solar apparatus are not discussed here, since these are designed and sized based on local requirements, material availability, construction practices, and practical experience. Simple rules of thumb based on prior experimentation are usually resorted to for designing such systems.

20.1.4.1.2 Active and Passive Systems

Active systems are those systems that need electric pumps or blowers to collect solar energy. It is evident that the amount of solar energy collected should be more than the electrical energy used. Active systems are invariably used for industrial applications and for most domestic and commercial applications as well. *Passive systems* are those systems that collect or use solar energy without direct recourse to any source of conventional power, such as electricity, to aid in the collection. Thus, either such systems operate by natural thermosyphon (for example, domestic water heating systems) between collector, storage, and load or, in the case of space heating, the architecture of the building is such as to favor optimal use of solar energy. Use of a passive system for space heating applications, however, in no way precludes the use of a backup auxiliary system. This chapter deals with active solar systems only.

20.1.4.1.3 Residential and Industrial Systems

Basically, the principles and the components used in these two types of systems are alike, the difference being in the load distribution, control strategies, and relative importance of the components with respect to each other. Whereas *residential* loads have sharp peaks in the early morning or in the evening and have significant seasonal variations, industrial loads tend to be fairly uniform over the year. Constant loads favor the use of solar energy because good equipment utilization can be achieved. Because of differences in load distribution, the role played by the storage differs for both applications. Residential loads often occur at times when solar radiation is no longer available. Thus the collector and the storage subsystems interact in a mode without heat withdrawal from the storage. Finally, for economic reasons, many residential systems are designed to operate by natural thermosyphon, in which case no pumps or controls are needed.

On the other hand, for *industrial and commercial* applications, there is no a priori relationship between the time dependence of the load and the period of sunshine. Moreover, a high reliability has to be assured, so the solar system will have to be combined with a conventional system. Very often, a significant portion of the load can be directly supplied by the solar system even without storage. Another option is to use buffer storage for short periods, on the order of a few hours, in case of discontinuous batch process loads. Thus, the proper design of the storage component has to be given adequate consideration. At present, due to economic constraints as well as the fact that proper awareness of the various installations and operational difficulties associated with larger solar thermal systems is still lacking, solar thermal systems are normally designed either (i) with the no-storage option, or (ii) with buffer storage where a small fraction of the total heat demand is only supplied by the solar system.

20.1.4.1.4 Liquid and Air Collectors

Although air has been the primary fluid for space heating and drying applications, solar air heating systems have until recently been relegated to second place, mainly as a result of the engineering difficulties associated with such systems. Also, applications involving hot air are probably less common than those needing hot water. Air systems for space heating are well described by Löf (1981).

Even with liquid solar collectors, various configurations are possible, and these can be classified basically as *nontracking* (which include flat-plate collectors and CPCs) or *tracking* collectors (which include various types of concentrating collectors). For low-grade thermal heat, for which solar energy is most suited, flat-plate collectors are far more appropriate than concentrating collectors, not only because of their lower cost but also because of their higher thermal efficiencies at low temperature levels. Moreover, their operation and maintenance costs are lower. Finally, for locations having a high fraction of diffuse radiation, as in the tropics, flat-plate collectors are considered to be thermally superior because they can make use of diffuse radiation as well as beam radiation. Although the system design methodologies presented in this chapter explicitly assume flat-plate collector systems, these design approaches can be equally used with concentrating collectors.

20.1.4.1.5 Daily and Seasonal Storage

By *daily storage* is meant systems having capacities equivalent to at most a few days of demand (i.e., just enough to tide over day-to-day climatic fluctuations). In *seasonal storage*, solar energy is stored during the summer for use in winter. Industrial demand loads, which are more or less uniform over the year, are badly suited for seasonal-storage systems. This is also true of air-conditioning for domestic and commercial applications because the load is maximum when solar radiation is also maximum, and vice versa. The present-day economics of seasonal storage units do not usually make such systems an economical proposition except for community heating in cold climates.

20.1.4.2 Closed-Loop and Open-Loop Systems

The two possible configurations of solar thermal systems with daily storage are classified as closed-loop or open-loop systems. Though different authors define these differently, we shall define these as follows. A *closed-loop system* has been defined as a circuit in which the performance of the solar collector is directly dependent on the storage temperature. Figure 20.16 gives a schematic of a closed-loop system in which the fluid circulating in the collectors does not mix with the fluid supplying thermal energy to the load. Thus, these two subsystems are distinct in the sense that any combination of fluids (water or air) is theoretically feasible (a heat exchanger, as shown in the figure, is of course imperative when the fluids are different). However, in practice, only water-water, water-air, or air-air combinations are used. From the point of system performance, the storage temperature normally varies over the day and, consequently, so does collector performance. Closed-loop system configurations have been widely used to date for domestic hot water and space heating applications. The flow rate per unit collector area is generally around 50 kg/(h m^2) for liquid collectors. The storage volume makes about 5–10 passes through the collector during a typical sunny day, and this is why such systems are called *multipass* systems. The temperature rise for each pass is small, of the order of 2°C–5°C for systems with circulating pumps and about 10°C for thermosyphon systems. An expansion tank and a check valve to prevent reverse thermosyphoning at nights, although not shown in the figure, are essential for such system configurations.

Figure 20.17 illustrates one of the possible configurations of *open-loop systems*. Open-loop systems are defined as systems in which the collector performance is independent of the storage temperature. The working fluid may be rejected (or a heat recuperator can be used) if contaminants are picked up during its passage through the load. Alternatively, the working fluid could be directly recalculated back to the entrance of the solar collector field. In all these open-loop configurations, the collector is subject to a given or known inlet temperature specified by the load requirements.

If the working fluid is water, instead of having a continuous flow rate (in which case the outlet temperature of the water will vary with isolation), a solenoid valve can be placed just at the exit of the collector, set so as to open when the desired temperature level of the fluid in the collector is reached. The water is then discharged into storage, and fresh water is taken into the collector. The solar collector will thus operate in a discontinuous manner, but this will ensure that the temperature in the storage is always at the desired level. An alternative way of ensuring uniform collector outlet temperature is to vary the flow rate according to the incident radiation. One can collect a couple of percent more energy

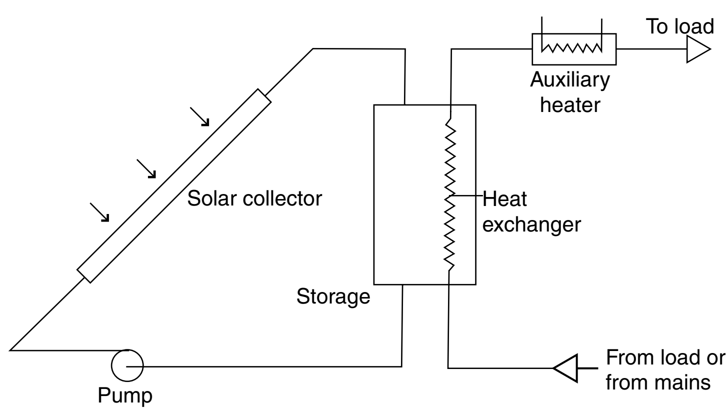


FIGURE 20.16 Schematic of a closed-loop solar system.

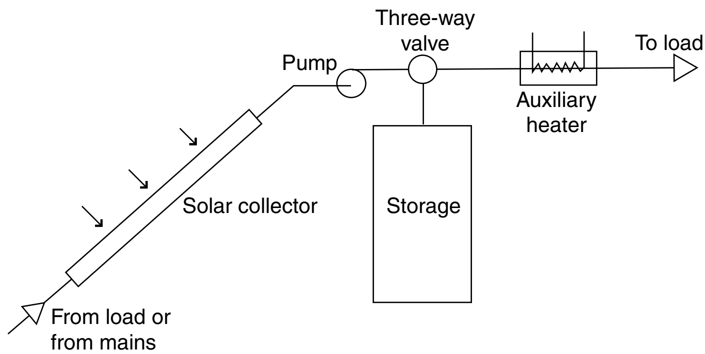


FIGURE 20.17 Schematic of an open-loop solar system.

than with constant rate single-pass designs (Gordon and Zarmi 1985). However, this entails changing the flow rate of the pump more or less continuously, which is injurious to the pump and results in reduced life. Of all the three variants of the open-loop configuration, the first one, namely the single-pass open-loop solar thermal system configuration with constant flow rate and without a solenoid valve, is the most common.

As stated earlier, closed-loop systems are appropriate for domestic applications. Until recently, industrial process heat systems were also designed as large solar domestic hot-water systems with high collector flow rates and with the storage tank volume making several passes per day through the collectors. Consequently, the storage tank tends to be fairly well mixed. Also the tank must be strong enough to withstand the high pressure from the water mains. The open-loop single-pass configuration, wherein the required average daily fluid flow is circulated just once through the collectors with the collector inlet temperature at its lowest value, has been found to be able to deliver as much as 40% more yearly energy for industrial process heat applications than the multipass designs (Collares-Pereira et al. 1984). Finally, in a closed-loop system where an equal amount of fresh water is introduced into storage whenever a certain amount of hot water is drawn off by the load, it is not possible to extract the entire amount of thermal energy contained in storage since the storage temperature is continuously reduced due to mixing. This *partial depletion effect* in the storage tank is not experienced in open-loop systems. The penalty in yearly energy delivery ranges typically from about 10% for daytime-only loads to around 30% for nighttime-only loads compared to a closed-loop multipass system where the storage is depleted every day. Other advantages of open-loop systems are (i) the storage tank need not be pressurized (and hence is less costly), and (ii) the pump size and parasitic power can be lowered.

A final note of caution is required. The single-pass design is not recommended for *variable* loads. The tank size is based on yearly daily load volumes, and efficient use of storage requires near-total depletion of the daily collected energy each day. If the load draw is markedly lower than its average value, the storage would get full relatively early the next day and solar collection would cease. It is because industrial loads tend to be more uniform, both during the day and over the year, than domestic applications that the single-pass open-loop configuration is recommended for such applications.

20.1.4.2.1 Description of a Typical Closed-Loop System

Figure 20.18 illustrates a typical closed-loop solar-supplemented liquid heating system. The useful energy is often (but not always) delivered to the storage tank via a collector-heat exchanger, which separates the collector fluid stream and the storage fluid. Such an arrangement is necessary either for antifreeze protection or to avoid corrosion of the collectors by untreated water containing gases and salts. A safety relief valve is provided because the system piping is normally nonpressurized, and any steam produced in the solar collectors will be let off from this valve. When this happens, energy dumping is said to take place.

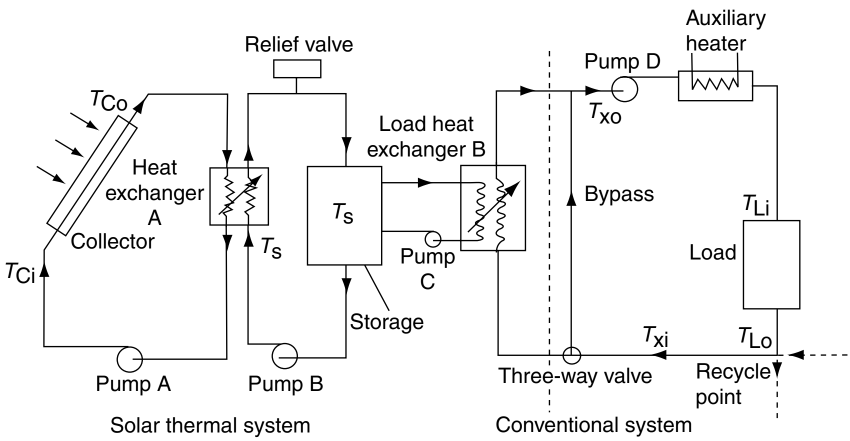


FIGURE 20.18 Schematic of a typical closed-loop system with auxiliary heater placed in series (also referred to as a topping-up type).

Fluid from storage is withdrawn and made to flow through the load-heat exchanger when the load calls for heat. Whenever possible, one should withdraw fluid directly from the storage and pass it through the load, and avoid incorporating the load-heat exchanger, since it introduces additional thermal penalties and involves extra equipment and additional parasitic power use. Heat is withdrawn from the storage tank at the top and reinjected at the bottom in order to derive maximum benefit from the thermal stratification that occurs in the storage tank. A bypass circuit is incorporated prior to the load heat exchanger and comes into play

1. when there is no heat in the storage tank (i.e., storage temperature T_S is less than the fluid temperature entering the load heat exchanger T_{Xi})
2. when T_S is such that the temperature of the fluid leaving the load heat exchanger is greater than that required by the load (i.e., $T_{Xo} > T_{Li}$, in which case the three-way valve bypasses part of the flow so that $T_{Xo} = T_{Li}$). The bypass arrangement is thus a differential control device which is said to modulate the flow such that the above condition is met. Another operational strategy for maintaining $T_{Xo} = T_{Li}$ is to operate the pump in a "bang-bang" fashion (i.e., by short cycling the pump). Such an operation is not advisable, however, since it would lead to premature pump failure.

An auxiliary heater of the *topping-up type* supplies just enough heat to raise T_{Xo} to T_{Li}. After passing through the load, the fluid (which can be either water or air) can be recirculated or, in case of liquid contamination through the load, fresh liquid can be introduced. The auxiliary heater can also be placed in parallel with the load (see Figure 20.19), in which case it is called an *all-or-nothing type*. Although such an arrangement is thermally less efficient than the topping-up type, this type is widely used during the solar retrofit of heating systems because it involves little mechanical modifications or alterations to the auxiliary heater itself.

It is obvious that there could also be solar-supplemented energy systems that do not include a storage element in the system. Figure 20.20 shows such a system configuration with the auxiliary heater installed in series. The operation of such systems is not very different from that of systems with storage, the primary difference being that whenever instantaneous solar energy collection exceeds load requirements (i.e., $T_{Co} > T_{Li}$), energy dumping takes place. It is obvious that by definition there cannot be a closed-loop, no-storage solar thermal system. Solar thermal systems without storage are easier to construct and operate, and even though they may be effective for 8–10 h a day, they are appropriate for applications such as process heat in industry.

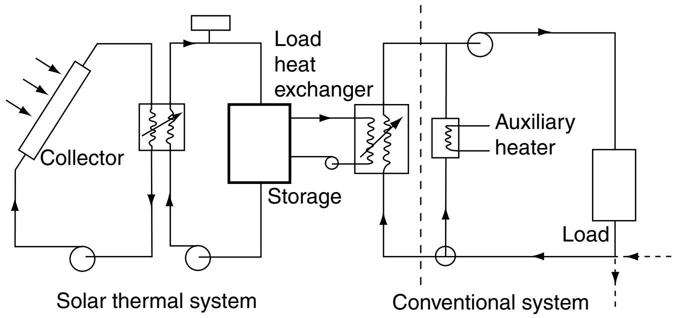


FIGURE 20.19 Schematic of a typical closed-loop system with auxiliary heater placed in parallel (also referred to as an all-or-nothing type).

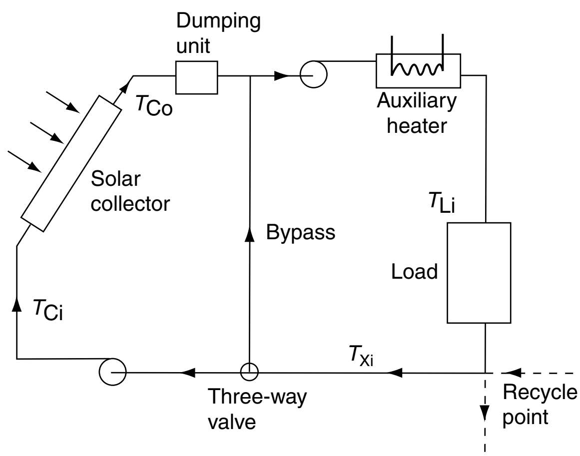


FIGURE 20.20 Simple solar thermal system without storage.

Active closed-loop solar systems as described earlier are widely used for service hot-water systems, that is, for domestic hot water and process heat applications as well as for space heat. There are different variants to this generic configuration. A system without the collector-heat exchanger is referred to having collectors *directly coupled* to the storage tank (as against *indirect coupling* as in Figure 20.16). For domestic hot-water systems, the system can be simplified by placing the auxiliary heater (which is simply an electric heater) directly inside the storage tank. One would like to maintain stratification in the tank so that the coolest fluid is at the bottom of the storage tank, thereby enhancing collection efficiency. Consequently, the electric heater is placed at about the upper third portion of the tank so as to assure good collection efficiency while assuring adequate hot water supply to the load. A more efficient but expensive option is widely used in the United States: the *double tank system*, shown in Figure 20.21. Here the functions of solar storage and auxiliary heating are separated, with the solar tank acting as a preheater for the conventional gas or electric unit. Note that a further system simplification can be achieved for domestic applications by placing the load heat exchanger directly inside the storage tank. In certain cases, one can even eliminate the heat exchanger completely.

Another system configuration is the *drain-back* (also called drain-out) system, where the collectors are emptied each time the solar system shuts off. Thus the system invariably loses collector fluid at least once, and often several times, each day. No collector-heat exchanger is needed, and freeze protection is inherent in such a configuration. However, careful piping design and installation, as well as a two-speed pump, are

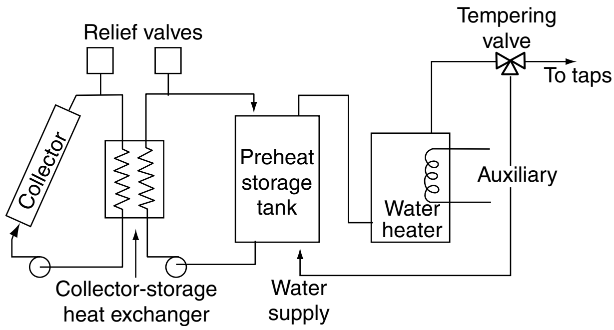


FIGURE 20.21 Schematic of a standard domestic hot-water system with double tank arrangement. (From Duffie, J. A. and Beckman, W. A., *Solar Engineering of Thermal Processes*, Wiley Interscience, New York, 1980.)

needed for the system to work properly (Newton and Gilman 1981). The drain-back configuration may be either open (vented to atmosphere) or closed (for better corrosion protection). Long-term experience in the United States with the drain-back system has shown it to be very reliable if engineered properly. A third type of system configuration is the *drain-down* system, where the fluid from the collector array is removed only when adverse conditions, such as freezing or boiling, occur. This design is used when freezing ambient temperatures are only infrequently encountered.

Active solar systems of the type described above are mostly used in countries such as the United States and Canada. Countries such as Australia, India, and Israel (where freezing is rare) usually prefer thermosyphon systems. No circulating pump is needed, the fluid circulation being driven by density difference between the cooler water in the inlet pipe and the storage tank and the hotter water in the outlet pipe of the collector and the storage tank. The low fluid flow in thermosyphon systems enhances thermal stratification in the storage tank. The system is usually fail-proof, and a study by Liu and Fanney (1980) reported that a thermosyphon system performed better than several pumped service hot-water systems. If operated properly, thermosyphon and active solar systems are comparable in their thermal performance. A major constraint in installing thermosyphon systems in already existing residences is the requirement that the bottom of the storage tank be at least 20 cm or more higher than the top of the solar collector in order to avoid reverse thermosyphoning at night. To overcome this, spring-loaded one-way valves have been used, but with mixed success.

20.1.5 Controls

There are basically five categories to be considered when designing automatic controls (Mueller Associates 1985): (i) collection to storage, (ii) storage to load, (iii) auxiliary energy to load, (iv) miscellaneous (i.e., heat dumping, freeze protection, overheating, etc.), and (v) alarms. The three major control system components are sensors, controllers, and actuating devices. Sensors are used to detect conditions (such as temperatures, pressures, etc.). Controllers receive output from the sensors, select a course of action, and signal a system component to adjust the condition. Actuated devices are components such as pumps, valves, and alarms that execute controller commands and regulate the system.

The sensors for the controls must be set, operated, and located correctly if the solar system is to collect solar energy effectively, reduce operating time, wear and tear of active components, and minimize auxiliary and parasitic energy use. Moreover, sensors also need to be calibrated frequently. For diagnostic purposes, it may be advisable to add extra sensors and data acquisition equipment in order to verify system operation and keep track of long-term system operation. Potential problems can be then rectified in time. The reader may refer to manuals by Mueller Associates (1985) or by SERI (1989) for more details on controls pertaining to solar energy systems.

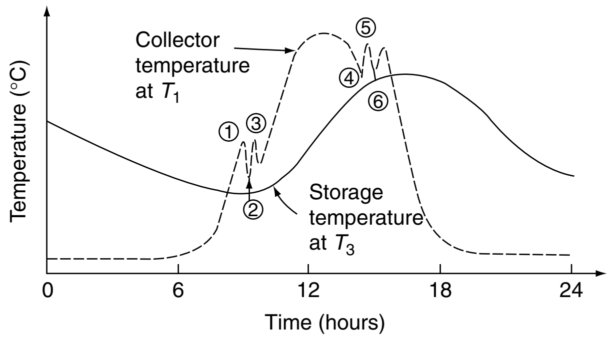


FIGURE 20.22 Typical diurnal variation of collector and storage temperatures. (From CSU, Solar Heating and Cooling of Residential Buildings—Design of Systems, manual prepared by the Solar Energy Applications Laboratory, Colorado State University, 1980.)

Though single-point temperature controllers or solar-cell-activated controls have been used for activated solar collectors, the best way to do so is by differential temperature controllers. Temperature sensors are used to measure the fluid temperature at collector outlet and at the bottom of the storage tank. When the difference is greater than a set amount, say 5°C, then the controller turns the pump on. If the pump is running and the temperature difference falls below another preset value, say 1°C, the controller stops the pump. The temperature deadband between switching-off and reactivating levels should be set with care, since too high a deadband would adversely affect collection efficiency and too low a value would result in short cycling of the collector pump. Figure 20.22 taken from CSU (1980), shows typical diurnal temperature variations of the liquids at collector exit T_1 and in the storage bottom T_3 as a result of heat withdrawal and/or heat losses from the storage. At about 8:30 a.m., $T_1 > T_3$ and, since there is no flow in the collector, T_1 increases rapidly until the difference $(T_1 - T_3)$ reaches the preset activation level (shown as point 1). The collector pump A comes on, and liquid circulation through the collector begins. Because of this cold water surge, T_1 decreases, resulting in a drop of $(T_1 - T_3)$ to the preset deactivating level (shown as point 2). The pump switches off, and liquid flow through the collectors stops. Gradually T_1 increases again, and so on. The number of on–off cycles at system start-up depends on solar intensity, fluid flow rate, volume of water in the collector loop, and the differential controller setting. A similar phenomenon of cycling also occurs in the afternoon. However, the error introduced in solar collector long-term performance predictions by neglecting this cycling effect in the modeling equations is usually small.

20.1.5.1 Corrections to Collector Performance Parameters

20.1.5.1.1 Combined Collector-Heat Exchanger Performance

The use of the heat exchanger A in Figure 20.18 imposes a penalty on the performance of the solar system because T_{Ci} is always higher than T_S, thereby decreasing q_C (see Figure 20.23). The collector-heat exchanger can be implicitly accounted for by suitably modifying the collector performance parameters. Recall from basic heat transfer the concept of heat exchanger effectiveness E defined as the ratio of the actual heat transfer rate to the maximum possible heat transfer rate, that is,

$$E = (mc_p)_a(T_{ai} - T_{ao})/(mc_p)_{min}(T_{ai} - T_{bi}) \quad (20.28a)$$

$$= (mc_p)_b(T_{bo} - T_{bi})/(mc_p)_{min}(T_{ai} - T_{bi}) \quad (20.28b)$$

where $(mC_p)_x$ = capacitance rate of fluid X (with $X = a$ for the warmer fluid, or $X = b$ for the cooler fluid) and $(mC_p)_{min}$ is the lower heat capacitance value of either stream. The advantage of this modeling approach is that, to a good approximation, E can be considered constant in spite of variations in

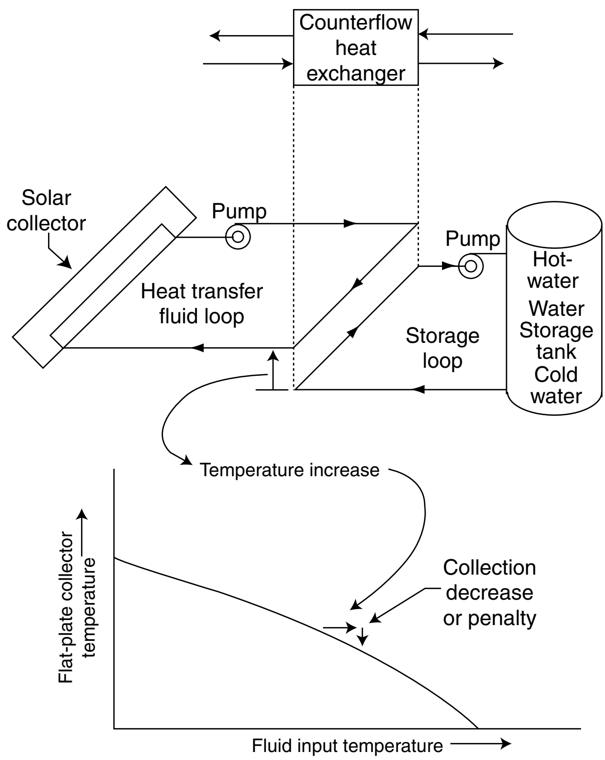


FIGURE 20.23 Heat collection decrease caused by double-loop heat exchangers. (From Cole, R. L., Nield, K. J., Rohde, R. R., and Wolosewicz, R. M. eds., *Design and Installation Manual for Thermal Energy Storage*, ANL-79-15, Argonne National Laboratory, Argonne, IL, 1979.)

temperature levels provided the mass flow rates of both fluids remain constant. Thus knowing the two flow rates, E, T_{ai}, and T_{bi}, both the exit fluid temperatures can be conveniently deduced. De Winter (1975) has shown that the combined performance of the solar collector and the heat exchanger can be conveniently modeled by replacing the collector heat removal factor F_R by a combined collector-exchanger heat removal factor F'_R such that

$$\frac{F'_R}{F_R} = \left[1 + \frac{F_R U_L A_C}{(mc_p)_C}\left\{\frac{(mc_p)_C}{E_A(mc_p)\min} - 1\right\}\right]^{-1} \tag{20.29}$$

where $(mC_p)_c$ is the capacitance rate of the fluid through the collector and E_A is the effectiveness of heat exchanger A. The variation of F'_R/F_R is shown in Figure 20.24. The plots exhibit the same type of asymptotic behavior with mass flow rate as in Figure 20.3.

The design of the collector-heat exchanger also requires care if the penalty imposed by it on the solar collection is to be minimized. Using a large heat exchanger increases the effectiveness and lowers this penalty; that is, the ratio (F'_R/F_R) is high, but the associated initial and operating costs may be higher. Both these considerations need to be balanced for optimum design (see Figure 20.25). Optimum heat exchanger area A_X can be found from the following equation proposed by Cole et al. (1979):

$$A_X = A_C\left[\frac{F_R U_L C_C}{U_X C_X}\right]^{1/2} \tag{20.30}$$

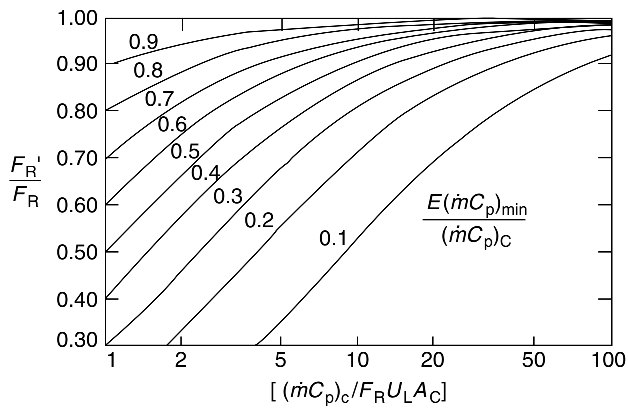


FIGURE 20.24 Variation of collector-heat exchanger correction factor. (From Duffie, J. A. and Beckman, W. A., *Solar Engineering of Thermal Processes*, Wiley Interscience, New York, 1980.)

where A_C is the collector area, C_C is the cost per unit collector area, C_X is the cost per unit heat exchanger area, and U_X is the heat loss per unit area of the heat exchanger.

20.1.5.1.2 Collector Piping and Shading Losses

Other corrections that can be applied to collector performance parameters include those for thermal losses from the piping (or from ducts) between the collection subsystem and the storage unit. Beckman (1978) has shown that these losses can be conveniently taken into consideration by suitably modifying the η_n and U_L terms of the solar collectors as follows:

$$\frac{\eta'_n}{\eta_n} = \left[1 + \frac{u_d A_0}{(mc_p)_C}\right]^{-1} \quad \text{and} \quad \frac{U'_L}{U_L} = \frac{1 - \frac{U_d A_i}{(mc_p)_C} + \frac{U_d (A_i + A_0)}{A_C F_R U_L}}{1 + \frac{U_d A_0}{(mc_p)_C}} \tag{20.31}$$

where U_d is the heat coefficient from the pipe or duct, A_0 is the heat loss area of the outlet pipe or duct, and A_i is the heat loss area of the inlet pipe or duct.

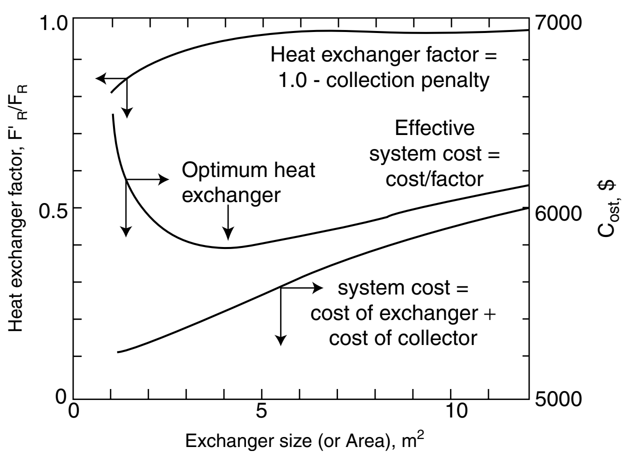


FIGURE 20.25 Typical heat-exchanger optimization plot. (From Cole, R. L., Nield, K. J., Rohde, R. R., and Wolosewicz, R. M. eds., *Design and Installation Manual for Thermal Energy Storage*, ANL-79-15, Argonne National Laboratory, Argonne, IL, 1979.)

When large collector arrays are mounted on flat roofs or level ground, multiple rows of collectors are usually arranged in a sawtooth fashion. These multiple rows must be spaced so that they do not shade each other at low sun angles. Unlimited space is rarely available, and it is desirable to space the rows as close as possible to minimize piping and to keep land costs low. Some amount of shading, especially during early mornings and late evenings during the winter months is generally acceptable. Detailed analysis of shading losses is cumbersome though not difficult and equations presented in standard text books such as Duffie and Beckman (1980) can be used directly.

20.1.6 Thermal Storage Systems

Low-temperature solar thermal energy can be stored in liquids, solids, or phase change materials (PCMs). Water is the most frequently used liquid storage medium because of its low cost and high specific heat. The most widely used solid storage medium is rocks (usually of uniform circular size 25–40 mm in diameter). PCM storage is much less bulky because of the high-latent heat of the PCM material, but this technology has yet to become economical and safe for widespread use.

Water storage would be the obvious choice when liquid collectors are used to supply hot water to a load. When hot air is required (for space heat or for convective drying), one has two options: an air collector with a pebble-bed storage or a system with liquid collectors, water storage, and a load heat exchanger to transfer heat from the hot water to the working air stream. Though a number of solar air systems have been designed and operated successfully (mainly for space heating), water storage is very often the medium selected. Water has twice the heat capacity of rock, so water storage tanks will be smaller than rock-bed containers. Moreover, rock storage systems require higher parasitic energy to operate, have higher installation costs, and require more sophisticated controls. Water storage permits simultaneous charging and discharging while such an operation is not possible for rock storage systems. The various types of materials used as containers for water and rock-bed storage and the types of design, installation, and operation details one needs to take care of in such storage systems are described by Mueller Associates (1985), SERI (1989).

Sensible storage systems, whether water or rock-bed, exhibit a certain amount of thermal stratification. Standard textbooks present relevant equations to model such effects. In the case of active closed-loop multipass hot-water systems, storage stratification effects can be neglected for long-term system performance with little loss of accuracy. Moreover, this leads to conservative system design (i.e., solar contribution is underpredicted if stratification is neglected). A designer who wishes to account for the effect of stratification in the water storage can resort to a formulation by Phillips and Dave (1982), who showed that this effect can be fairly well modeled by introducing a *stratification coefficient* (which is a system constant that needs to be determined only once) and treating the storage subsystem as fully mixed. However, this approach is limited to the specific case of no (or very little) heat withdrawal from storage during the collection period. Even when water storage systems are highly stratified, simulation studies seem to indicate that modeling storage as a one-dimensional plug-flow three-node heat transfer problem yields satisfactory results of long-term solar system performance.

The thermal losses q_w from the storage tank can be modeled as

$$q_w = (UA_S)(T_S - T_{env}) \tag{20.32}$$

where (UA_S) is the storage overall heat loss per unit temperature difference and T_{env} is the temperature of the air surrounding the storage tank. Note that (UA_S) depends (i) on the storage size, which is a parameter to be sized during system design, and (ii) on the configuration of the storage tank (i.e., on the length by diameter ratio in case of a cylindrical tank). For storage tanks, this ratio is normally in the range of 1.0–2.0.

20.1.7 Solar System Simulation

A system model is nothing but an assembly of appropriate component modeling equations that are to be solved over time subject to certain forcing functions (i.e., the meteorological data and load data). The resulting set of simultaneous equations can be solved either analytically or numerically.

The analytical method of resolution is appropriate, or possible, only for simplified system configurations and operating conditions. This approach has had some success in the analysis and design of open-loop systems (refer to Reddy 1987, and Gordon and Rabl 1986, for more details). On the other hand, numerical simulation can be performed for any system configuration and operating strategy, however, complex. However, this is time-consuming and expensive in computer time and requires a high level of operator expertise.

We shall illustrate the approach of numerical simulation by considering the simple solar system shown in Figure 20.18. Assuming a fully mixed storage tank, the instantaneous energy balance equation is

$$(Mc_p)_S(\mathrm{d}T_S/\mathrm{d}t) = q_C - q_u - q_w \tag{20.33}$$

where

q_C is the useful energy delivered by the solar collector (given by Equation 20.2.)
q_w is the thermal loss from the storage tank (given by Equation 20.32)
q_u is the useful heat transferred through the load heat exchanger, which can be determined as follows:

The maximum hourly rate of energy transfer through the load heat exchanger is

$$q_{max} = E_B(mc_p)_{min}(T_S - T_{Xi})\delta_L \tag{20.34}$$

where δ_L is a control function whose value is either 1 or 0 depending upon whether there is a heat demand or not. Since q_{max} can be greater than the amount of thermal energy q_L actually required by the load, the bypass arrangement can be conveniently modeled as

$$q_u = \min(q_{max}, q_L) \tag{20.35}$$

where

$$q_L = (mc_p)_L(T_{Li} - T_{Xi}) \tag{20.36}$$

for water heating and industrial process heat loads. Space heating and cooling loads can be conveniently determined by one of the several variants of the bin-type methods (ASHRAE 1985).

The amount of energy q_{max} supplied by a topping-up type of auxiliary heater is

$$q_{max} = q_L - q_u \tag{20.37}$$

Assuming $T_{env} = T_a$, Equation 20.33 can be expanded into

$$(Mc_p)_S\frac{\mathrm{d}T_S}{\mathrm{d}t} - A_C F_R[I_T\eta_0 - U_L(T_S - T_a)]^+ - (mc_p)_s(T_S - T_{Xi})\delta_L - (UA)_S(T_S - T_a) \tag{20.38}$$

The presence of control functions and time dependence of I_T and T_a prevent a general analytical treatment, though, as mentioned earlier, specific cases can be handled. The numerical approach involves expressing this differential equation in finite difference form. After rearranging, one gets

$$= T_{S,b} + \frac{\Delta t}{(Mc_p)_S}\{A_C F_R[I_T\eta_0 - U_L(T_{S,b} - T_a)]^+ - (mc_p)_S(T_{S,b} - T_{Xi})\delta_L - (UA)_S(T_{S,b} - T_a)\}T_{S,f} \tag{20.39}$$

where $T_{S,b}$ and $T_{S,f}$ are the storage temperatures at the beginning and the end of the time step Δt. The time step is sufficiently small (say 1 h) that I_T and T_a can be assumed constant. This equation is repeatedly used over the time period in question (day, month, or year), and the total energy supplied by the collector or to the load can be estimated.

Such methods of simulation, referred as stepwise steady-state simulations, implicitly assume that the solar thermal system operates in a steady-state manner during one time step, at the end of which it undergoes an abrupt change in operating conditions as a result of changes in the forcing functions, and thereby attains a new steady-state operating level. Although in reality, the system performance varies smoothly over time and is consequently different from that outlined earlier, it has been found that, in most cases, taking time steps of the order of 1 h yields acceptable results of long-term performance.

The objective of solar-supplemented energy systems is to displace part of the conventional fuel consumption of the auxiliary heater. The index used to represent the contribution of the solar thermal system is the *solar fraction*, which is the fraction of the total energy required by the load that is supplied by the solar system. The solar fraction could be expressed over any time scale, with month and year being the most common. Two commonly used definitions of the monthly solar fraction are

1. Thermal solar fraction:

$$f_Y = Q_{UM}/Q_{LM} = 1 - Q_{aux,M}/Q_{LM} \tag{20.40}$$

where

Q_{UM} is the monthly total thermal energy supplied by the solar system
Q_{LM} is the monthly total thermal requirements of the load
$Q_{aux,M}$ is the monthly total auxiliary energy consumed

2. Energy solar fraction (i.e., thermal plus parasitic energy):

$$f'_M = Q'_{UM}/Q'_{LM} \tag{20.41}$$

where Q'_{UM} is Q_{UM} *minus* the parasitic energy consumed by the solar system and Q'_{UM} is Q_{LM} *plus* the parasitic energy consumed by the load.

Example 20.1.5

Simulate the closed-loop solar thermal system shown in Figure 20.18 for each hour of a day assuming both collector and load heat exchangers to be absent (i.e., $E_A = E_B = 1$). Assume the following data as input for the simulation: $A_C = 10$ m^2, $F_R U_L = 5.0$ W/m$^{2\circ}$C), $F_R \eta_0 = 0.7$, $(Mc_p) = 2.0$ MJ/°C and $(UA)_S =$ 3 W/°C. Water is withdrawn to meet a load from 9 a.m. to 7 p.m. (solar time) at a constant rate of 60 kg/h and is replenished from the mains at a temperature of 25°C. The storage temperature at the start (i.e., at 6 a.m.) is 40°C, and the environment temperature is equal to the ambient temperature. The temperature of the water entering the load should not exceed 55°C. The hourly values of the solar radiation on the plane of the collector are given in column 2 of Table 20.3 and the ambient temperature is assumed constant over the day and equal to 25°C. The variation of the optical efficiency with angle of incidence can be neglected.

The results of the simulation are given in Table 20.3. The following equations should permit the reader to verify for himself the results obtained. Simulating the system entails solving the following equations in the sequence given here:

Column 4. Useful energy delivered by the collector (Equation 20.2)

$$q_C = 10[0.7 I_T - 5(3{,}600/10^6) T_{S,b} - 25]^+ \text{ (MJ/h)}$$

The term $(3600/10^6)$ is introduced to convert W/m^2 (the units in which I_T is expressed) into MJ/(h m^2). Note that $T_{S,b}$ is taken to be equal to $T_{S,f}$ of the final hour.

Column 5. Thermal losses from the storage tank (Equation 20.32)

$$q_w = 3(3{,}600/10^6)(T_{S,b} - 25) \text{ (MJ/h)}$$

TABLE 20.3 Simulation Results of Example 20.1.5

(1) Solar Time (h)	(2) I_T (MJ/m^{2}h)	(3) $T_{S,f}$ (°C)	(4) q_C (MJ/h)	(5) q_w (MJ/h)	(6) q_{max} (MJ/h)	(7) q_L (MJ/h)	(8) q_U (MJ/h)	(9) q_{aux} (MJ/h)
Start		40.00						
6–7	0.37	39.92	0.00	0.16	0.00	0.00	0.00	0.00
7–8	0.95	41.82	3.96	0.16	0.00	0.00	0.00	0.00
8–9	1.54	45.61	7.75	0.18	0.00	0.00	0.00	0.00
9–10	2.00	48.05	10.29	0.22	5.18	754	5.18	2.36
10–11	2.27	50.90	11.74	0.25	5.79	7.54	5.79	1.75
11–12	2.46	53.78	12.56	0.28	6.51	7.54	6.51	1.03
12–13	2.50	56.17	12.32	0.31	7.24	7.54	7.24	0.31
13–14	2.24	57.26	10.07	0.34	7.84	7.54	7.54	0.00
14–15	2.12	57.84	9.03	0.35	8.11	7.54	7.54	0.00
15–16	1.37	55.73	3.68	0.35	8.25	7.54	7.54	0.00
16–17	0.76	51.79	0.00	0.33	7.72	7.54	7.54	0.00
17–18	0.23	48.28	0.00	0.29	6.73	7.54	6.73	0.81
18–19	0.00	45.23	0.00	0.25	5.85	7.54	5.85	1.69
Total	18.81	—	81.41	3.48	69.24	75.40	67.48	7.94

Column 6. The maximum rate of energy that can be transferred from the load can be calculated from Equation 20.34

$$q_{max} = 60(4{,}190/10^6)(T_{S,b} - 25)\ \text{(MJ/h)}$$

Column 7. The thermal energy required by the load (from Equation 20.36)

$$q_L = 60(4{,}190/10^6)(55 - 25) = 7.54\ \text{MJ/h}$$

Column 8. The actual amount of heat withdrawn from storage (Equation 20.35)

$$q_w = \min[\text{column 6}, \text{column 7}]$$

Column 9. The amount of energy supplied by the auxiliary heater (Equation 20.2.37)

$$q_{aux} = \text{column 7} - \text{column 8}$$

The final storage temperature $T_{S,f}$ is now calculated from Equation 20.39

$$T_{S,f} = T_{S,b} + [\text{column 4} - \text{column 8} - \text{column 5}]/2.0$$

From Table 20.3, we note that the solar collector efficiency over the entire day is $[81.41/(18.81 \times 10)] = 0.43$. The corresponding daily solar fraction $= (67.48/75.40) = 0.895$.

20.1.8 Solar System Sizing Methodology

Sizing of solar systems primarily involves determining the collector area and storage size that are most cost effective. Standalone and solar-supplemented systems have to be treated separately since the basic design problem is somewhat different. The interested reader can refer to Gordon (1987) for sizing standalone systems.

20.1.8.1 Solar-Supplemented Systems

20.1.8.1.1 Production Functions

Because of the annual variation of incident solar radiation, it is not normally economical to size a solar subsystem such that it provides 100% of the heat demand. Most solar energy systems follow the *law of* *diminishing returns*. This implies that increasing the size of the solar collector subsystem results in a less than proportional increase in the annual fuel savings (or alternatively, in the annual solar fraction).

Any model has two types of variables: exogenous and endogenous. The *exogenous parameters* are also called the input variables, and these in turn may be of two kinds. *Variable exogenous parameters* are the collector area A_C, the collector performance parameters $F_R\eta_n$ and F_RU_L, the collector tilt, the thermal storage capacity $(Mc_p)_S$, the heat exchanger size, and the control strategies of the solar thermal system. On the other hand, the climatic data specified by radiation and the ambient temperature, as well as the end use thermal demand characteristics, are called *constrained exogenous parameters* because they are imposed externally and cannot be changed. The *endogenous parameters* are the output parameters whose values are to be determined, the annual solar fraction being one of the parameters most often sought.

Figure 20.26 illustrates the law of diminishing results. The annual solar fraction f_Y is seen to increase with collector area but at a decreasing rate and at a certain point will reach saturation. Variation of any of the other exogenous parameters also exhibits a similar trend. The technical relationship between f_Y and one or several variable exogenous parameters for a given location is called the *yearly production function*.

It is only for certain simple types of solar thermal systems that an analytical expression for the production can be deduced directly from theoretical considerations. The most common approach is to carry out computer simulations of the particular system (solar plus auxiliary) over the complete year for several combinations of values of the exogenous parameters. The production function can subsequently be determined by an empirical curve fit to these discrete sets of points.

Example 20.1.6

Kreider (1979a, 1979b) gives the following expression for the production function of an industrial solar water heater for a certain location:

$$f_Y = Q_{UY}/Q_{LY} = \left(0.35 - \frac{F_R U_L}{100 F_R \eta_n}\right) \ln\left(1 + \frac{20 F_R \eta_n A_C}{Q_{LY}}\right) \tag{20.42}$$

where Q_{UY} is the thermal energy delivered by the solar thermal system over the year in GJ/y; Q_{LY} is the yearly thermal load demand, also in GJ/y; and F_RU_L is in W/(m$^{2\circ}$C). Note that only certain solar system exogenous parameters figure explicitly in this expression, thereby implying that other exogenous parameters (for example, storage volume) have not been varied during the study. As an illustration,

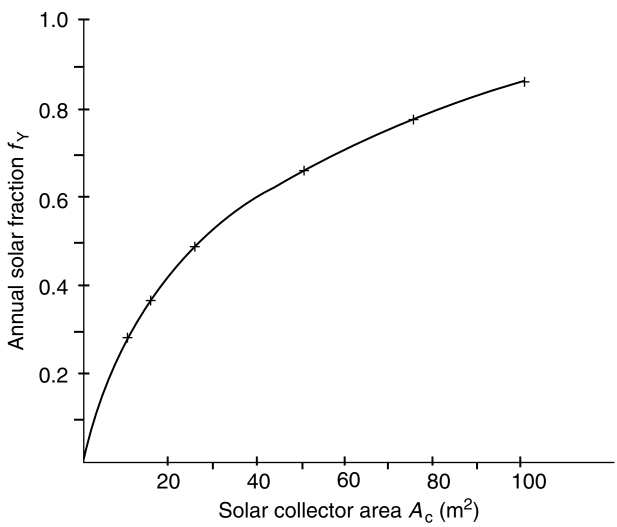


FIGURE 20.26 A typical solar system production function (see Example 20.3.6).

let us assume the following nominal values: $Q_{LY} = 100$ GJ/y, $F_R U_L = 2.0$ W/(m$^{2\circ}$C), and $F_R \eta_n = 0.7$. For a 1% increase in collector area A_C, the corresponding percentage increase in Q_{UY} (called elasticity) can be determined:

$$\frac{dQ_{UY}}{Q_{UY}} = \frac{dA_C}{A_C}\left[\left(\frac{Q_{LY}}{20F_R\eta_n A_C} + 1\right)\ln\left(1 + \frac{20F_R\eta_n A_C}{Q_{LY}}\right)\right]^{-1} \tag{20.43}$$

From this, we obtain the expression for *marginal productivity*

$$\frac{dQ_{UY}}{dA_C} = \frac{Q_{UY}}{A_C}\left[\left(\frac{Q_{LY}}{20F_R\eta_n A_C} + 1\right)\ln\left(1 + \frac{20F_R\eta_n A_C}{Q_{LY}}\right)\right]^{-1} \tag{20.44}$$

Numerical values can be obtained from the preceding expression. Though Q_{UY} increases with A_C, the marginal productivity of Q_{UY} goes on decreasing with increasing A_C, thus illustrating the law of diminishing returns. A qualitative explanation of this phenomenon is that as A_C increases, the mean operating temperature level of the collector increases, thus leading to decreasing solar collection rates. Figure 20.26 illustrates the variation of f_Y with A_C as given by Equation 20.42 when the preceding numerical values are used.

The objective of the sizing study in its widest perspective is to determine, for a given specific thermal end use, the size and configuration of the solar subsystem that results in the most economical operation of the entire system. This economical optimum can be determined using the production function along with an appropriate economic analysis. Several authors—for example, Duffie and Beckman (1980) or Rabl (1985)—have presented fairly rigorous methodologies of economic analysis, but a simple approach is adequate to illustrate the concepts and for preliminary system sizing.

20.1.8.1.2 Simplified Economic Analysis

It is widely recognized that *discounted cash flow analysis* is most appropriate for applications such as sizing an energy system. This analysis takes into account both the initial cost incurred during the installation of the system and the annual running costs over its entire life span.

The economic objective function for optimal system selection can be expressed in terms of either the energy cost incurred or the energy savings. These two approaches are basically similar and differ in the sense that the objective function of the former has to be minimized while that of the latter has to be maximized. In our analysis, we shall consider the latter approach, which can further be subdivided into the following two methods:

1. Present worth or life cycle savings, wherein all running costs are discounted to the beginning of the first year of operation of the system.
2. Annualized life cycle savings, wherein the initial expenditure incurred at the start as well as the running costs over the life of the installation are expressed as a yearly mean value.

20.1.9 Solar System Design Methods

20.1.9.1 Classification

Design methods may be separated into three generic classes. The *simple* category, usually associated with the prefeasibility study phase involves quick manual calculations of solar collector/system performance and rule-of-thumb engineering estimates. For example, the generalized yearly correlations proposed by Rabl (1981) and described in Section 20.1.2 could be conveniently used for year-round, more or less constant loads. The approach is directly valid for open-loop solar systems, while it could also be used for closed-loop systems if an *average* collector inlet temperature could be determined. A simple manner of selecting this temperature $\bar{T}_m$ for domestic closed-loop multipass systems is to assume the following empirical relation:

$$\bar{T}_{\mathrm{m}} - T_{\mathrm{mains}}/3 + (2/3)T_{\mathrm{set}} \tag{20.45}$$

where T_{mains} is the average annual supply temperature and T_{set} is the required hot-water temperature (about 60°C–80°C in most cases).

These manual methods often use general guidelines, graphs, and/or tables for sizing and performance evaluation. The designer should have a certain amount of knowledge and experience in solar system design in order to make pertinent assumptions and simplifications regarding the operation of the particular system.

Mid-level design methods are resorted to during the feasibility phase of a project. The main focus of this chapter has been toward this level, and a few of these design methods will be presented in this section. A personal computer is best suited to these design methods because they could be conveniently programmed to suit the designer's tastes and purpose (spreadsheet programs, or better still one of the numerous equation-solver software packages, are most convenient). Alternatively, commercially available software packages such as f-chart (Beckman, Klein, and Duffie 1977) could also be used for certain specific system configurations.

Detailed design methods involve performing hourly simulations of the solar system over the entire year from which accurate optimization of solar collector and other equipment can be performed. Several simulation programs for active solar energy systems are available, TRNSYS (Klein et al. 1975, 1979) developed at the University of Wisconsin–Madison being perhaps the best known. This public-domain software has technical support and is being constantly upgraded. TRNSYS contains simulation models of numerous subsystem components (solar radiation, solar equipment, loads, mechanical equipment, controls, etc.) that comprise a solar energy system. A user can conveniently hook up components representative of a particular solar system to be analyzed and then simulate that system's performance at a level of detail that the user selects. Thus TRNSYS provides the design with large flexibility, diversity, and convenience of usage.

As pointed out by Rabl (1985), the detailed computer simulations approach, though a valuable tool, has several problems. Judgment is needed both in the selection of the input and in the evaluation of the output. The very flexibility of big simulation programs has drawbacks. So many variables must be specified by the user that errors in interpretation or specification are common. Also, learning how to use the program is a time-consuming task. Because of the numerous system variables to be optimized, the program may have to be run for numerous sets of combinations, which adds to expense and time. The inexperienced user can be easily misled by the second-order details while missing first-order effects. For example, uncertainties in load, solar radiation, and economic variables are usually very large, and long-term performance simulation results are only accurate to within a certain degree. Nevertheless, detailed simulation programs, if properly used by experienced designers, can provide valuable information on system design and optimization aspects at the final stages of a project design.

There are basically three types of mid-level design approaches: the empirical correlation approach, the analytical approach, and the one-day repetitive methods (described fully in Reddy 1987). We shall illustrate their use by means of specific applications.

20.1.9.2 Active Space Heating

The solar system configuration for this particular application has become more or less standardized. For example, for a liquid system, one would use the system shown in Figure 20.27. One of the most widely used design methods is the f-chart method (Beckman, Klein, and Duffie 1977; Duffie and Beckman 1980), which is applicable for standardized liquid and air heating systems as well as for standardized domestic hot-water systems. The f-chart method basically involves using a simple algebraic correlation that has been deduced from numerous TRNSYS simulation runs of these standard solar systems subject to a wide range of climates and solar system parameters (see Figure 20.28). Correlations were

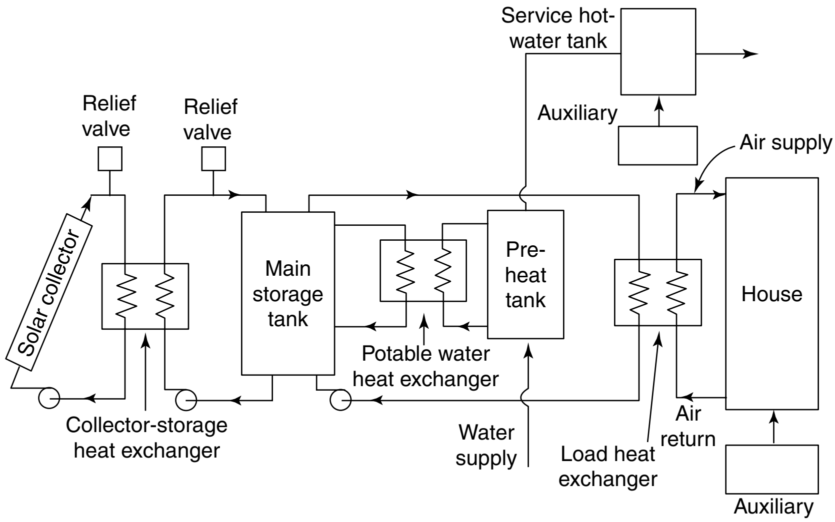


FIGURE 20.27 Schematic of the standard space heating liquid system configuration for the f-chart method. (From Duffie, J. A. and Beckman, W. A., *Solar Engineering of Thermal Processes*, Wiley Interscience, New York, 1980.)

developed between monthly solar fractions and two easily calculated dimensionless variables X and Y, where

$$X = (A_C F'_R U_L (T_{Ref} - \bar{T}_a)\Delta t)/Q_{LM} \tag{20.46}$$

$$Y = A_C F'_R \bar{\eta}_0 \bar{H}_T N / Q_{LM} \tag{20.47}$$

where

A_C collector area (m^2)
F'_R collector-heat exchanger heat removal factor (given by Equation 20.29)
U_L collector overall loss coefficient (W/(m$^{2\circ}$C))
Δt total number of seconds in the month $= 3600 \times 24 \times N = 86{,}400 \times N$

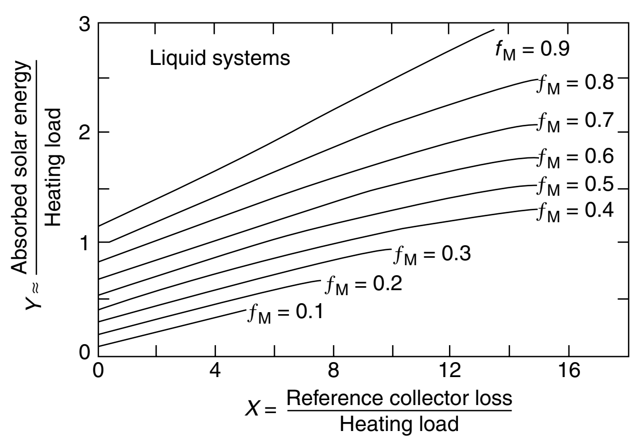


FIGURE 20.28 The f-chart correlation for liquid system configuration. (From Duffie, J. A. and Beckman, W. A., *Solar Engineering of Thermal Processes*, Wiley Interscience, New York, 1980.)

$\bar{T}_a$ monthly average ambient temperature (°C)
T_{Ref} an empirically derived reference temperature, taken as 100°C
Q_{LM} monthly total heating load for space heating and/or hot water (J)
$\bar{H}_T$ monthly average daily radiation incident on the collector surface per unit area (J/m^2)
N number of days in the month
$\bar{\eta}_0$ monthly average collector optical efficiency

The dimensionless variable X is the ratio of reference collector losses over the entire month to the monthly total heat load; the variable Y is the ratio of the monthly total solar energy absorbed by the collectors to the monthly total heat load. It will be noted that the collector area and its performance parameters are the predominant exogenous variables that appear in these expressions. For changes in secondary exogenous parameters, the following corrective terms X_C and Y_C should be applied for liquid systems:

1. for changes in storage capacity:

$$X_C/X = (\text{actual storage capacity/standard storage capacity})^{-0.25} \tag{20.48}$$

where the standard storage volume is 75 L/m^2 of collector area.

2. for changes in heat exchanger size:

$$Y_C/Y = 0.39 + 0.65\exp[-(0.139(UA)_B/(E_L(mc_p)_{min}))] \tag{20.49}$$

The monthly solar fraction for liquid space heating can then be determined from the following empirical correlation:

$$f_M = 1.029Y - 0.065X - 0.245Y^2 + 0.0018X^2 + 0.0215Y^3 \tag{20.50}$$

subject to the conditions that $0 \le X \le 15$ and $0 \le Y \le 3$. This empirical correlation is shown graphically in Figure 20.28.

A similar correlation has also been proposed for space heating systems using air collectors and pebble-bed storage. The procedure for exploiting the preceding empirical correlations is as follows. For a predetermined location, specified by its 12 monthly radiation and ambient temperature values, Equation 20.50 is repeatedly used for each month of the year for a particular set of variable exogenous parameters. The monthly solar fraction f_M and thence the annual thermal energy delivered by the solar thermal system are easily deduced. Subsequently, the entire procedure is repeated for different values and combinations of variable exogenous parameters. Finally, an economic analysis is performed to determine optimal sizes of various solar system components. Care must be exercised that the exogenous parameters considered are not outside the range of validity of the f-chart empirical correlations.

Example 20.1.7

(Adapted from Duffie and Beckman 1980). A solar heating system is to be designed for Madison, Wisconsin (latitude 43°N) using one-cover collectors with $F_R\eta_n = 0.74$ and $F_RU_L = 4$ W/(m^{2}°C). The collector faces south with a slope of 60° from the horizontal. The average daily radiation on the tilted surface in January is 12.9 MJ/m^2, and the average ambient temperature is −7°C. The heat load is 36 GJ for space heating and hot water. The collector-heat exchanger correction factor is 0.97 and the ratio of monthly average to normal incidence optical efficiency is 0.96. Calculate the energy delivered by the solar system in January if 50 m^2 of collector area is to be used.

From Equation 20.46 and Equation 20.47, with $A_C = 50\ m^2$,

$$X = 4.0 \times 0.97[100 - (-7)]31 \times 86{,}400 \times 50/(36 \times 10^9) = 1.54$$
$$Y = 0.74 \times 0.97 \times 0.96 \times 12.9 \times 10^6 \times 31 \times 50/(36 \times 10^9) = 0.38$$

From Equation 20.50, the solar fraction for January is $f_M = 0.26$. Thus the useful energy delivered by the solar system $= 0.26 \times 36 = 9.4$ GJ.

In an effort to reduce the tediousness involved in having to perform 12 monthly calculations, two analogous approaches that enable the annual solar fraction to be determined directly have been developed by Barley and Winn (1978), Lameiro and Bendt (1978). These involve the computation of a few site-specific empirical coefficients, thereby rendering the approach less general. For example, the *relative-area* method suggested by Barley and Winn enable the designer to directly calculate the annual solar fraction of the corresponding system using four site-specific empirical coefficients. The approach involves curve fits to simulation results of the f-chart method for specific locations in order to deduce a correlation such as:

$$f = c_1 + c_2 \ln(A/A_{0.5}) \tag{20.51}$$

where c_1 and c_2 are location-specific parameters that are tabulated for several United States locations, and $A_{0.5}$ is the collector area corresponding to an annual solar fraction of 0.5 given by

$$A_{0.5} = A_S(UA)/(F'_R\eta_0 - F'_R U_L Z) \tag{20.52}$$

where A_S and Z are two more location specific parameters, UA is the overall heat loss coefficient of the building, and $F'_R\eta_0$ and $F'_R U_L$ are the corresponding solar collector performance parameters corrected for the effect of the collector-heat exchanger.

Barley and Winn also proposed a simplified economic life-cycle analysis whereby the optimal collector area could be determined directly. Another well-known approach is the *Solar Load Radio* (SLR) method for sizing residential space heating systems (Hunn 1980).

20.1.9.3 Domestic Water Heating

The f-chart correlation (Equation 20.50) can also be used to predict the monthly solar fraction for domestic hot-water systems represented by Figure 20.21 provided the water mains temperature T_{mains} is between 5 and 20°C and the minimum acceptable hot-water temperature drawn from the storage for end use (called the set water temperature T_w) is between 50 and 70°C. Further, the dimensionless parameter X must be corrected by the following ratio

$$X_w/X = (11.6 + 1.8T_w + 3.86T_{mains} - 2.32\bar{T}_a)/(100 - \bar{T}_a) \tag{20.53}$$

In case the domestic hot-water load is much smaller than the space heat load, it is recommended that Equation 20.50 be used without the above correction.

20.1.9.4 Industrial Process Heat

As discussed in "Description of a Typical Closed-Loop System," two types of solar systems for industrial process heat are currently used: the closed-loop multipass systems (with an added distinction that the auxiliary heater may be placed either in series or in parallel (see Figure 20.18 and Figure 20.19) and the open-loop singlepass system. How such systems can be designed will be described next.

20.1.9.4.1 Closed-Loop Multipass Systems

Auxiliary Heater in Parallel. The Phibar-f chart method (Klein and Beckman 1979; Duffie and Beckman 1980; Reddy 1987) is a generalization of the f-chart method in the sense that no restrictions need be imposed on the temperature limits of the heated fluid in the solar thermal system. However, three basic criteria for the thermal load have to be satisfied for the Phibar-f chart method to be applicable: (i) the thermal load must be constant and uniform over each day and for at least a month, (ii) the thermal energy supplied to the load must be above a minimum temperature that completely specifies the temperature level of operation of the load, and (iii) either there is no conversion efficiency in the load (as in the case of hot water usage) or the efficiency of conversion is constant (either because the load temperature level is constant or because the conversion efficiency is independent of the load temperature level). The approach is strictly applicable to solar systems with the auxiliary heater in parallel (Figure 20.19).

A typical application for the Phibar-f chart method is absorption air-conditioning. The hot water inlet temperature from the collectors to the generator must be above a minimum temperature level (say, 80°C) for the system to use solar heat. If the solar fluid temperature is less (even by a small amount), the entire energy to heat up the water to 80°C is supplied by the auxiliary system.

As a result of continuous interaction between storage and collector in a closed-loop system, the variation of the storage temperature and hence the fluid inlet temperature to the collectors) over the day and over the month is undetermined. The Phibar-f chart method implicitly takes this into account and reduces these temperature fluctuations down to a monthly mean equivalent storage temperature $\bar{T}_S$ The determination of this temperature in conjunction with the daily utilizability approach is the basis of the design approach.

The basic empirical correlation of the Phibar-f chart method, shown graphically in Figure 20.29, is as follows:

$$f_M = Y\bar{\phi} - a[\exp(bf_M) - 1][1 - \exp(cX)] \tag{20.54}$$

with $0 < X < 20$ and $0 < Y < 1.6$, and $\bar{\phi}$ is the Klein daily utilizability fraction described in "Daily Utilizability" and given by Equation 20.24. Y is given by Equation 20.47, and X is now slightly different from Equation 20.46 and is defined as:

$$X = A_C F_R U_L \Delta t(100°C)/Q_{LM} \tag{20.55}$$

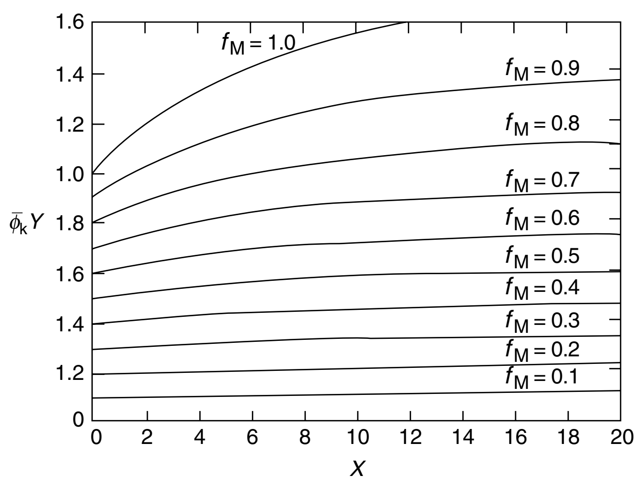


FIGURE 20.29 The Phibar-f chart correlation for a storage capacity of 350 kJ/m^2 and for a 12 h per day thermal load. (From Duffie, J. A. and Beckman, W. A., *Solar Engineering of Thermal Processes*, Wiley Interscience, New York, 1980.)

The values of the constants a, b, and c are given by the following:

1. for an end use load operating between 6 a.m. and 6 p.m. every day of the month,

$$a = 0.015[(Mc_p)_S/350 \text{ kJ/(m}^2\text{°C)}]^{-0.76} \quad \text{for} \quad 175 \leq [(Mc_p)_S/A_C] \leq 1{,}400 \text{ kJ/(m}^2\text{°C)}$$
$$\text{and } b = 3.85 \text{ and } c = -0.15 \tag{20.56}$$

2. for an end use load operating 24 h per day over the entire month,

$$a = 0.043 \quad \text{only for}[(Mc_p)_S/A_C] = 350\text{kJ/(m}^2\text{°C)}, b = 2.81, \text{and } c = -0.18 \tag{20.57}$$

It will be noted that $(Y\bar{\phi})$ denotes the maximum solar fraction that would have resulted had T_{Ci}, the inlet temperature to the collector, been equal to T_{Li} throughout the month. The term in Equation 20.54 that is subtracted from $(Y\bar{\phi})$ represents the decrease in the solar fraction as a result of $T_{Ci} > T_{Xi}$. The solar fraction computed from Equation 20.54 has to be corrected for the effect of thermal losses from the storage as well as the presence of the load-heat exchanger, both of which will decrease the solar fraction. For complete details, refer to Duffie and Beckman (1980) or Reddy (1987). Note that Equation 20.54 needs to be solved for f_m in an iterative manner.

Auxiliary Heater in Series. The Phibar-f chart method has also been modified to include solar systems with the auxiliary heater in series as shown in Figure 20.18. This configuration leads to higher solar fractions but retrofit to existing systems may be more costly.

In this case, the empirical correlation given by Equation 20.54 has been modified by Braun, Klein, and Pearson (1983) as follows:

$$f_M = Y\bar{\phi} - a[\exp(bf_M) - 1][1 - \exp(cX)]\exp(-1.959Z) \tag{20.58}$$

with $Z = Q_{LM}/(C_L \times 100°\text{C})$ and (i) when there is no load-heat exchanger, C_L is the monthly total load heat capacitance, which is the product of the monthly total mass of water used and the specific heat capacity of water, and (ii) when there is a load-heat exchanger present $C_L = E_L \times C_{min}$, where E_L is the effectiveness of the load-heat exchanger and C_{min} is the monthly total heat capacitance, which is the lesser of the two fluids rates across the load heat exchanger.

The modified Phibar-f chart is similar to the original method in respect to load uniformity on a day-to-day basis over the month and in assuming no conversion efficiency. The interested may refer to Braun, Klein, and Pearson (1983) or Reddy (1987) for complete details.

20.1.9.4.2 Open-Loop Single-Pass Systems

The advantages offered by open-loop single-pass systems over closed-loop multipass systems for meeting constant loads has been described in "Closed-Loop and Open-Loop Systems." Because industrial loads operate during the entire sun-up hours or even for 24 h daily, the simplest solar thermal system is one with no heat storage (Figure 20.20). A sizable portion (between 25 and 70%) of the daytime thermal load can be supplied by such systems and consequently, the sizing of such systems will be described below (Gordon and Rabl 1982). We shall assume that T_{Li} and T_{Xi} are constant for all hours during system operation. Because no storage is provided, excess solar energy collection (whenever $T_{Ci} > T_{Li}$) will have to be dumped out.

The maximum collector area $\hat{A}_C$ for which energy dumping does not occur at any time of the year can be found from the following instantaneous heat balance equation:

$$P_L = \hat{A}_C\hat{F}_R[I_{max}\eta_n - U_L(T_{Ci} - T_a)] \tag{20.59}$$

where P_L, the instantaneous thermal heat demand of the load (say, in kW) is given by

$$P_L = m_L c_p(T_{Li} - T_{Xi}) \tag{20.60}$$

and F_R is the heat removal factor of the collector field when its surface area is $\hat{A}_C$. Since $\hat{A}_C$ is as yet unknown, the value of $\hat{F}_R$ is also undetermined. (Note that though the *total* fluid flow rate is known, the flow rate per unit collector area is not known.) Recall that the plate efficiency factor F' for liquid collectors can be assumed constant and independent of fluid flow rate per unit collector area. Equation 20.59 can be expressed in terms of critical radiation level I_C:

$$P_L = \hat{A}_C \hat{F}_R \eta_n (I_{max} - I_C) \tag{20.61a}$$

or

$$\hat{A}_C \hat{F}_R \eta_n = P_L/(I_{max} - I_C) \tag{20.61b}$$

Substituting Equation 20.3 in lieu of F_R and rearranging yields

$$\hat{A}_C = -(m_L c_p/F' U_L)\ln[1 - P_L U_L/(\eta_n (I_{max} - I_C) m_L c_p)] \tag{20.62}$$

If the actual collector area A_C exceeds this value, dumping will occur as soon as the radiation intensity reaches a value I_D, whose value is determined from the following heat balance:

$$P_L = A_C F_R \eta_n (I_D - I_C) \tag{20.63a}$$

Hence

$$I_D = I_C + P_L/(A_C F_R \eta_n) \tag{20.63b}$$

Note that the value of I_D decreases with increasing collector area A_C, thereby indicating that increasing amounts of solar energy will have to be dumped out.

Since the solar thermal system is operational during the entire sunshine hours of the year, the yearly total energy collected can be directly determined by the Rabl correlation given by Equation 20.26. Similarly, the yearly total solar energy collected by the solar system which has got to be dumped out is

$$Q_{DY} = A_C F_R \eta_n (\tilde{a} + \tilde{b} I_D + \tilde{c} I_D^2) \tag{20.64}$$

The yearly total solar energy delivered to the load is

$$\begin{aligned} Q_{UY} &= Q_{CY} - Q_{DY} \\ &= A_C F_R \eta_n [\tilde{b}(I_C - I_D) + \tilde{c}(I_C^2 - I_D^2)] \\ &= -(\tilde{b} + 2\tilde{c} I_C) P_L - \tilde{c} P_L^2/(A_C F_R \eta_n) \end{aligned} \tag{20.65}$$

$$= -(\tilde{b} + 2\tilde{c} I_C) P_L - \tilde{c} P_L^2/(A_C F_R \eta_n) \tag{20.66}$$

Replacing the value of F_R given by Equation 20.3, the annual production function in terms of A_C is

$$Q_{UY} = -(\tilde{b} + 2\tilde{c} I_C) P_L - \frac{\tilde{c} P_L^2}{\left(\frac{F'\eta_n}{F' U_L}\right)(m_L c_p)\left[1 - \exp\left(-\frac{F' U_L A_C}{m_L c_p}\right)\right]} \tag{20.67}$$

subject to the condition that $A_C > \hat{A}_C$. If the thermal load is not needed during all days of the year due to holidays or maintenance shutdown, the production function can be reduced proportionally. This is illustrated in the following example.

Example 20.1.8

Obtain the annual production function of an open-loop solar thermal system without storage that is to be set up in Boston, Massachusetts according to the following load specifications: industrial hot water load for 12 h a day (6 a.m.–6 p.m.) and during 290 days a year, mass flow rate $m_L = 0.25$ kg/s, required inlet temperature $T_{Li} = 60°C$. Contaminants are picked up in the load, so that all used water is to be rejected and fresh water at ambient temperature is taken in. Flat-plate collectors with tilt equal to latitude with the following parameters are used $F'\eta_n = 0.75$ and $F'U_L = 5.5$ W/(m^2°C). The latitude of Boston is 42.36°N = 0.739 radians. The yearly $\tilde{K} = 0.45$ and $\tilde{T}_a = 10.9°C$. Use the following Gordon and Rabl (1981) correlation:

$$\begin{aligned} Q_{CY}/A_c F_R \eta_n = {} & [(5.215 + 6.973 I_{bn}) + (-5.412 + 4.293 I_{bn})L + (1.403 - 0.899 I_{bn})L^2] + \\ & [(-18.596 - 5.931 I_{bn}) + (15.468 + 18.845 I_{bn})L + (-0.164 - 35.510 I_{bn})L^2] I_C + \\ & [(-14.601 - 3.570 I_{bn}) + (13.675 - 15.549 I_{bn})L + (-1.620 + 30.564 I_{bn})L^2] I_C^2 \end{aligned}$$

From Equation 20.27, $\tilde{I}_{bn} = 1.37 \times 0.45 - 0.34 = 0.276$ kW/m^2. The critical radiation level $I_C = 0$, since $T_{Ci} = T_a$. Consequently, Equation 20.26, using the above expression reduces to

$$\begin{aligned} Q_{CY}/(A_C F_R \eta_n) &= 5.215 + 6.973 \times 0.276 + (-5.412 + 4.293 \times 0.276)0.739 + (1.403 - 0.899 \\ &\quad \times 0.276)0.739^2 \\ &= 4.646 \text{ GJ/(m}^2\text{y)}. \end{aligned}$$

The expression for the dumped out energy is found from Equation 20.64 and the previous expression by replacing I_C by I_D:

$$\begin{aligned} Q_{CY,dump}/(A_C F_R \eta_n) &= 4.646 + [(-18.596 - 5.931 \times 0.276) + (15.468 + 18.845 \times 0.276)0.739 + \\ &\quad (-0.164 - 35.510 \times 0.276)0.739^2] I_D + [(14.601 - 3.57 \times 0.276) + \\ &\quad (-13.675 - 15.549 \times 0.276)0.739 + (1.620 + 30.564 \times 0.276)0.739^2] I_D^2 \\ &= 4.646 - 10.40 I_D + 5.83 I_D^2 \text{(GJ/m}^2\text{y)} \end{aligned}$$

The thermal energy demand $P_L = 0.25 \times 4.19(60 - 10.9) = 51.43$ kW.

The annual production function is

$$\begin{aligned} Q_{UY}(365/290) &= -(-10.40 + 2\tilde{c} \times 0)51.43 - (5.83 \times 51.43^2)/\{(0.75/5.5)(0.25 \times 4.19) \\ &\quad [1 - \exp[-(5.5 \times A_C)/(0.25 \times 4190)]]\} \end{aligned}$$

$$\text{or } Q_{UY} = 424.96 - 85.78/[1 - \exp(-A_C/190.45)] \text{(GJ/y)}$$

Complete details as well as how this approach can be extended to solar systems with storage (see Figure 20.30) can be found in Rabl (1985) or Reddy (1987).

20.1.10 Design Recommendations and Costs

20.1.10.1 Design Recommendations

As mentioned earlier, design methods reduce computational effort compared to detailed computer simulations. Even with this decrease, the problem of optimal system design and sizing remains formidable because of

a. The presence of several solar thermal system configuration alternatives.
b. The determination of optimal component sizes for a given system.

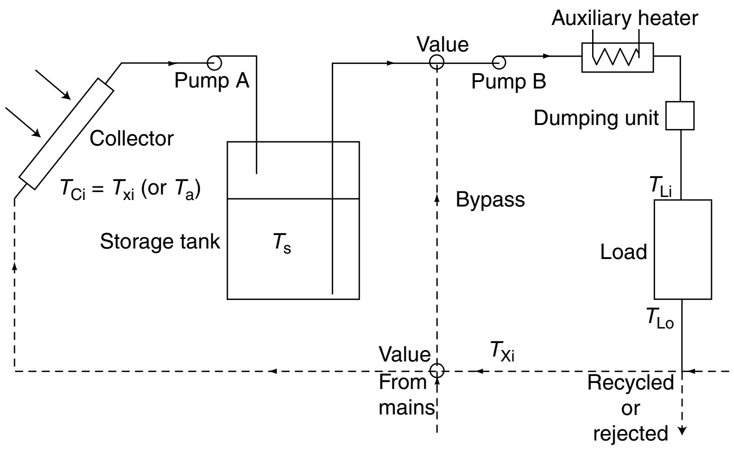


FIGURE 20.30 Open-loop solar industrial hot-water system with storage.

c. The presence of certain technical and economic constraints.
d. The choice of proper climatic, technical, and economic input parameters.
e. The need to perform sensitivity analysis of both technical and economic parameters.

For most practical design work, a judicious mix of theoretical expertise and practical acumen is essential. Proper focus right from the start on the important input variables as well as the restriction of the normal range of variation would lead to a great decrease in design time and effort several examples of successful case studies and system design recommendations are described in the published literature (see, for example, Kutcher et al. 1982).

20.1.10.2 Solar System Costs

How the individual components of the solar system contribute to the total cost can be gauged from Table 20.4. We note that collectors constitute the major fraction (from 15 to 30%), thus suggesting that collectors should be selected and sized with great care. Piping costs are next with other collector-related costs like installation and support structure being also important.

TABLE 20.4 Percentage of Total System Cost by Component

Cost Component	Percentage Range
Collectors	15–30
Collector installation	5–10
Collector support structure	5–20[a]
Storage tanks	5–7
Piping and specialties	10–30
Pumps	1–3
Heat exchangers	0–5[b]
Chiller	5–10
Miscellaneous	2–10
Instrumentation	1–3
Insulation	2–8
Control subsystem	4–9
Electrical	2–6

[a] For collectors mounted directly on a tilted roof.
[b] For systems without heat exchangers.
Source: From Mueller Associates. 1985. Active Solar Thermal Design Manual, funded by U.S. DOE (no. EG-77-C-01-4042), SERI (XY-2-02046-1) and ASHRAE (project no. 40). Baltimore, MD.

References

Anderson, B. 1977. *Solar Energy: Fundamentals in Building Design*, McGraw-Hill, New York.

ASHRAE Standard 93-77. 1978. *Methods of Testing to Determine the Thermal Performance of Solar Collectors*. American Society of Heating, Refrigeration and Air Conditioning Engineers, New York.

ASHRAE. 1985. *Fundamentals*, American Society of Heating, Refrigeration and Air Conditioning Engineers, New York.

Barley, C. D. and Winn, C. B. 1978. Optimal sizing of solar collectors by the method of relative areas. *Solar Energy*, 21, 279.

Beckman, W. A. 1978. Technical note: Duct and pipe losses in solar energy systems. *Solar Energy*, 21, 531.

Beckman, W. A., Klein, S. A., and Duffie, J. A. 1977. *Solar Heating Design by the f-Chart Method*, Wiley Interscience, New York.

Braun, J. E., Klein, S. A., and Pearson, K. A. 1983. An improved design method for solar water heating systems. *Solar Energy*, 31, 597.

Charters, W. W. S. and Pryor, T. L. 1982. In *An Introduction to the Installation of Solar Energy Systems*, pp. 503–127. Victoria Solar Energy Council, Melbourne, Australia.

Chiam, H. F. 1981. Planar concentrators for flat-plate solar collectors, *Solar Energy*, 26, 503.

Clark, D. R., Klein, S. A., and Beckman, W. A. 1983. Algorithm for evaluating the hourly radiation utilizability function. *ASME Journal of Solar Energy Engineering*, 105, 281.

Cole, R. L., Nield, K. J., Rohde, R. R., and Wolosewicz, R. M. eds. 1979. *Design and Installation Manual for Thermal Energy Storage*, ANL-79-15, pp. 223–397. Argonne National Laboratory, Argonne, IL.

Collares-Pereira, M. and Rabl, A. 1979. Derivation of method for predicting the long-term average energy delivery of solar collectors. *Solar Energy*, 23, 223.

Collares-Pereira, M., Gordon, J. M., Rabl, A., and Zarmi, Y. 1984. Design and optimization of solar industrial hot water systems with storage. *Solar Energy*, 32, 121.

CSU. 1980. *Solar Heating and Cooling of Residential Buildings*—Design of Systems, manual prepared by the Solar Energy Applications Laboratory, Colorado State University. Colorado State University, Fort Collins, CO.

de Winter, F. 1975. Heat exchanger penalties in double loop solar water heating systems. *Solar Energy*, 17, 335.

Duffie, J. A. and Beckman, W. A. 1980. *Solar Engineering of Thermal Processes*, Wiley Interscience, New York.

Edwards, D. K., Nelson, K. E., Roddick, R. D., and Gier, J. T. 1960. Basic Studies on the Use of Solar Energy, Report no. 60-93. Department of Engineering. University of California, Los Angeles, CA.

Gordon, J. M. and Rabl, A. 1982. Design, analysis and optimization of solar industrial process heat plants without storage. *Solar Energy*, 28, 519.

Gordon, J. M. and Zarmi, Y. 1985. An analytic model for the long-term performance of solar thermal systems with well-mixed storage. *Solar Energy*, 35, 55.

Gordon, J. M. and Rabl, A. 1986. Design of solar industrial process heat systems, In *Reviews of Renewable Energy Sources* chapter 6. Sodha, M. S. Mathur, S. S. Malik, M. A.S and Kandpal, T. C. eds., pp. 55–177. Wiley Eastern, New Delhi.

Gordon, J. M. 1987. Optimal sizing of stand-alone photovoltaic systems. *Solar Cells*, 20, 295.

Hollands, K. G. T. 1965. Honeycomb devices in flat-plate solar collectors. *Solar Energy*, 9, 159.

Hunn, B. D. 1980. A simplified method for sizing active solar space heating systems, *In Solar Energy Technology Handbook, Part B: Applications, System Design and Economics*, Dickinson, W. C. and Cheremisinoff, P. N. eds., pp. 639–255. Marcel Dekker, New York.

Klein, S. A., Cooper, P. I., Freeman, T. L., Beekman, D. M., Beckman, W. A. and Doffie, J. A. 1975. A method of simulation of solar processes and its applications, *Solar Energy*, 17, 29.

Klein, S. A. 1978. Calculation of flat-plate collector utilizability. *Solar Energy*, 21, 393.

Klein, S. A. et al. 1979. TRNSYS-A Transient System Simulation User's Manual, University of Wisconsin-Madison Engineering Experiment Station Report. University of Wisconsin-Madison Engineering Experiment Station Report 38-10.

Klein, S. A. and Beckman, W. A. 1979. A general design method for closed-loop solar energy systems. *Solar Energy*, 22, 269.

Kreider, J. F. 1979. Medium and High Temperature Solar Energy Processes. Academic Press, New York.

Kutcher, C. F., Davenport, R. L., Dougherty, D. A., Gee, R. C., Masterson, P. M., and May, E. K. 1982. Design Approaches for Solar Industrial Process Heat Systems, SERI/TR-253-1356. Solar Energy Research Institute, Golden, CO.

Lameiro, G. F. and Bendt, P., 1978. The GFL method for designing solar energy space heating and domestic hot water systems, In Proceedings of American Solar Energy Society Conference, Vol. 2, p. 113. Boulder, CO.

Larson, D. C. 1980. Optimization of flat-plate collector flat mirror system. *Solar Energy*, 24, 203.

Larson, R. W., Vignola, F., and West, R. 1992a. *Economics of Solar Energy Technologies*. American Solar Energy Society Report, Boulder, CO.

Liu, B. Y. H. and Jordan, R. C. 1960. The inter-relationship and characteristic distribution of direct, diffuse and total solar radiation. *Solar Energy*, 4, 1.

Liu, B. Y. H. and Jordan, R. C. 1963. A rational procedure for predicting the long-term average performance of flat-plate solar energy collectors. *Solar Energy*, 7, 53.

Liu, S. T. and Fanney, A. H. 1960. Comparing experimental and Computer-predicted performance for solar hot water systems, *ASHRAE Journal*, 22, 5, 34.

Löf, G. O. G. 1981. Air based solar systems for space heating, In *Solar Energy Handbook*, J. F. Kreider and F. Kreith, eds., pp. 39–44. McGraw-Hill, New York.

Meyer, B. A. 1978. Natural convection heat transfer in small and moderate aspect ratio enclosures—An application to flat-plate collectors, In *Thermal Storage and Heat Transfer in Solar Energy Systems*, Kreith, F. Boehm, R. Mitchell, J. and Bannerot, R. eds., pp. 555–558. American Society of Mechanical Engineers, New York.

Mitchell, J. C., Theilacker, J. C., and Klein, S. A. 1981. Technical note: Calculation of monthly average collector operating time and parasitic energy requirements. *Solar Energy*, 26, 555–558.

Mueller Associates. 1985. *Active Solar Thermal Design Manual*, funded by U. S. DOE (no. EG-77-C-01-4042). SERI(XY-2-02046-l) and ASHRAE (project no. 40). Baltimore, MD.

Newton, A. B. and Gilman, S. H. 1981. *Solar Collector Performance Manual*, funded by U. S. DOE (no. EG-77-C-01-4042), SERI(XH-9-8265-1) and ASHRAE (project no. 32, Task 3).

Phillips, W. F. and Dave, R. N. 1982. Effect of stratification on the performance of liquid-based solar heating systems. *Solar Energy*, 29, 111.

Rabl, A. 1981. Yearly average performance of the principal solar collector types. *Solar Energy*, 27, 215.

Rabl, A. 1985. *Active Solar Collectors and Their Applications*. Oxford University Press, New York.

Reddy, T. A. 1987. *The Design and Sizing of Active Solar Thermal Systems*. Oxford University Press, Oxford, U. K.

SERI 1989. Engineering Principles and Concepts for Active Solar Systems. Hemisphere Publishing Company, New York.

Theilacker, J. C. and Klein, S. A. 1980. Improvements in the utilizability relationships. *American Section of the International Solar Energy Society Meeting Proceedings*, p. 271. Phoenix, AZ.

Liu, S. T. and Fanney, A. H. 1980. Comparing experimental and computer-prodicted performance for solar hot water systems. *ASHRAE Journal*, 22, 5, 34.

Klein, S. A., Cooper, P. I., Freeman, T. L., Beekman, D. M., Beckman, W. A. and Doffie, J. A. 1975. A method of simulation of solar processes and its applications, *Solar Energy*, 17, 29.

20.2 Solar Heat for Industrial Processes

Riccardo Battisti, Hans Schweiger, and Werner Weiss

20.2.1 The Potential for Solar Process Heat

Currently, the widespread use of residential solar thermal energy has focused almost exclusively on swimming pools, domestic hot water preparation, and space heating, while its use in the service sector and in industrial applications is insignificant. On the other hand, the industrial sector accounts for a large share of the total energy consumption in the OECD (Organization for Economic Co-operation and Development) countries at approximately 30%, and a significant portion of industrial heat demand falls within a temperature range compatible with solar thermal collectors.

As a matter of fact, 30%–50% of the thermal energy needed in commercial and industrial companies for production processes is below 250°C (Kreider 1979; European Commission 2001). In this temperature range, the heat demand in the European Union (EU) for industrial processes can be estimated with about 300 TWh, or 7% of the total energy demand. The total potential for industrial process heat at below 150°C was estimated to be 202.8 TWh for the 12 countries that formed the EU in 1994 (Laue and Reichert 1994).

Studies for the application potential for industrial solar thermal systems were carried out in Austria (Müller 2004), Germany and Greece (PROCESOL 2000), Italy (IEA 2005), Netherlands (KWA 2001), Portugal and Spain (European Commission 2001). The potential for solar low temperature heat ranges between 3% and 4% of the total industrial heat demand in Italy, Spain, Portugal, and Austria. These studies primarily showed that solar thermal plants can readily provide the required low- and medium-temperature process heat.

The most promising industrial sectors and processes are shown in Table 20.5, and the distribution of heat demand by temperature range is shown in Figure 20.31.

Another important result showed by the studies is that the available surface area on factory roofs is often a limiting factor for the installation of a solar plant for industrial use.

TABLE 20.5 Industrial Sectors and Processes Suitable for Solar Thermal Use

Sector	Processes	Temperature (°C)
Brewing and malting	Wort boiling	100
	Bottle washing	60
	Drying	90
	Cooling	60
Milk	Pasteurization	60–85
	Sterilization	130–150
Food preservation	Pasteurization	110–125
	Sterilization	<80
	Cooking	70–100
	Scalding	95–100
	Bleaching	<90
Meat	Washing, sterilization, cleaning	<90
	Cooking	90–100
Wine and beverage	Bottle washing	60–90
	Cooling (single effect absorption cooling)	85
Textile	Washing, bleaching, dyeing	<90
	Cooking	140–200
Automobile	Paint drying	160–220
	Degreasing	35–55
Paper	Paper pulp: cooking	170–180
	Boiler feed water	<90
	Bleaching	130–150
	Drying	130–160
Tanning	Water heating for damp processes	165–180 (steam)
Cork	Drying, cork baking	40–155

Source: From European Commission, *The Potential of Solar Heat for Industrial Processes*, Final report, EC Project, Contract No. NNE5-1999-0308, 2001, http://www.solarpaces.org. With permission.

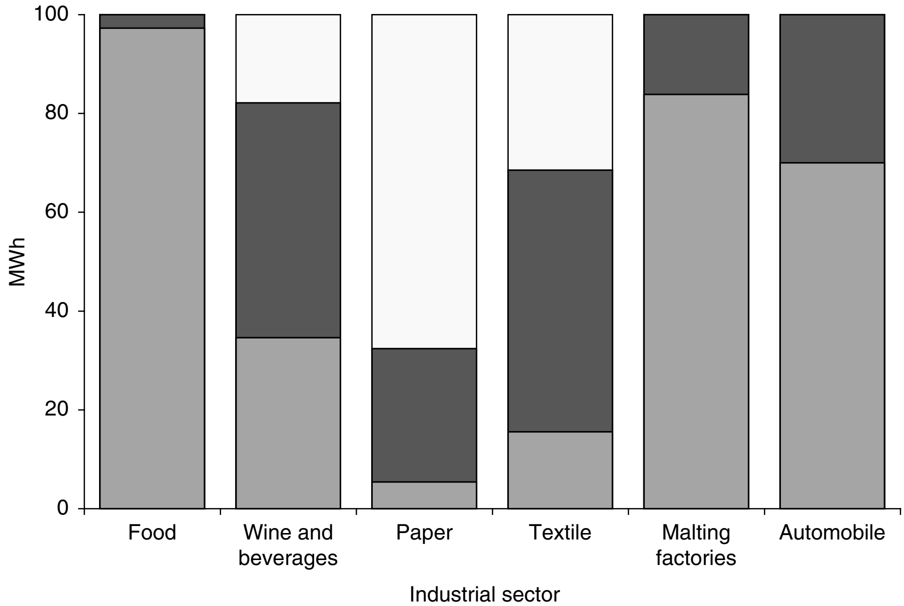


FIGURE 20.31 Distribution of the heat demand by temperature range in some selected companies studied within the POSHIP project, grouped by industrial sectors. (From European Commission, *The Potential of Solar Heat for Industrial Processes*, Final Report, EC Project, Contract No. NNE5-1999-0308, 2001, http://www.aiguasol.com/poship.htm. With permission.)

20.2.2 Solar Thermal Systems in Industrial Processes: Integration and Basic Design Guidelines

20.2.2.1 Which Solar Collectors?

For process temperatures up to about 60°C, flat-plate collectors with selective absorbers would be most appropriate and could be the most economical solution even up to a temperature range of 90°C. For temperatures above this range, other collector types should be considered: evacuated tubes, high efficiency flat plate, CPC or line-axis concentrating collectors. In the framework of IEA Task 33/IV, "Solar Heat for Industrial Processes," which is carried out within the Solar Heating and Cooling Program of the IEA, new "medium temperature collectors" (i.e., with operating temperatures between 80 and 250°C) are being developed, including:

- Improved flat-plate collectors
- Stationary low-concentration collectors
- Small parabolic trough collectors

Different collector technologies already available on the market can be used for applications in the 80°C–120°C temperature range. For example, flat-plate collectors with double antireflection glazings and hermetically sealed collectors with inert-gas fillings, or even a combination of both, reduce collector heat losses without significantly sacrificing the optical performance. Figure 20.32 shows estimated efficiency curves of single-, double-, and triple-glazed flat-plate collectors when newly developed antireflection glazings (AR-glass) are used.

Another solution for medium temperature collectors is to reduce heat losses by concentration by, for example, using stationary CPC collector without vacuum and with a low concentration factor (in the range of two).

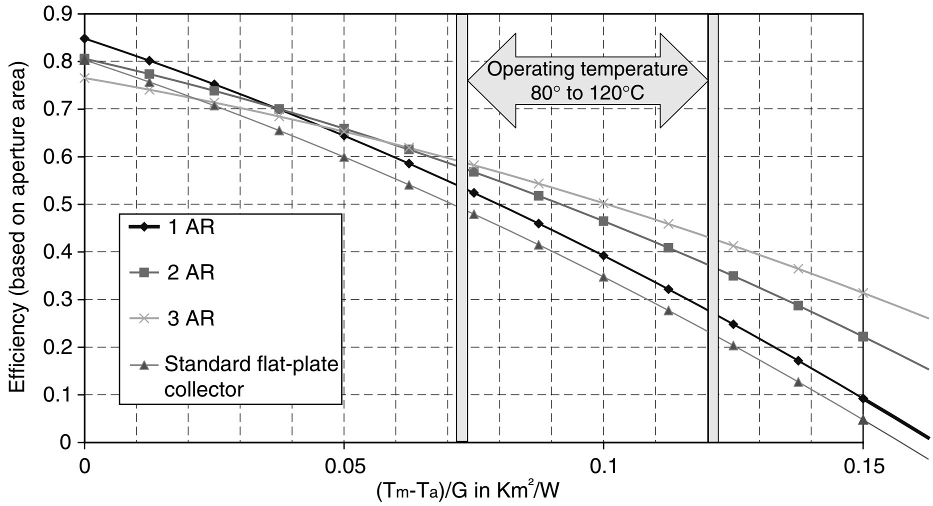


FIGURE 20.32 Efficiency curves of a single-, double-, and triple-glazed antireflection collector in comparison with a standard flat-plate collector with ordinary solar glass. (From IEA., *Task 33/IV, Solar Heat for Industrial Processes*, International Energy Agency, 2005.)

Between 150 and 250°C, it is appropriate to consider the parabolic trough collector technology. Much is known about high-temperature applications (400°C–600°C) using parabolic trough collectors for electric power production, but adjustments must be made for the medium temperature range. Current developments involved in Task 33/IV have been carried out in Spain, Austria, and Germany. The first results showed that parabolic troughs could even be an appropriate alternative for large systems at low temperatures (about 60°C). It is noteworthy that small parabolic trough are readily available in the market and are a reliable technology (e.g., in the U.S., some plants have been operating since the early 1990s).

20.2.2.2 Coupling the Solar Thermal System with the Processes

The integration of solar heat into industrial production processes is challenging for both the process engineer and the solar designer. Existing heating systems based on steam or hot water from boilers are normally designed for much higher temperatures (150°C–180°C) compared to those that the processes need (100°C or lower) to keep temperature differences small. On the contrary, the solar thermal system should always be coupled to the existing heat supply at the lowest possible temperature. Nevertheless, for fluid preheating, solar heat should be introduced only after preheating by waste heat recovery systems, and not as an alternative to these systems (Figure 20.33).

Even if the waste heat recovery raises the working temperature in the solar thermal system, the combination of both systems yields better results than a solar thermal system at lower temperature but without heat recovery.

Today, one of the most used heat recovery assessment methodologies is the "pinch" analysis. The discovery of the heat recovery "pinch" was a major breakthrough in the development of design methods for energy efficient industrial processes. Based on the analysis of hot and cold streams within the process, the pinch methodology gives fundamental hints about the possibility and the right position of heat exchangers for waste heat recovery. This allows developing integral solutions for solar thermal energy applications in given industrial processes. From the point of view of energy, the streams that constitute a process flow sheet may be classified into two groups: "hot" streams, i.e., streams that must be cooled, and "cold" streams, i.e., streams that must be heated.

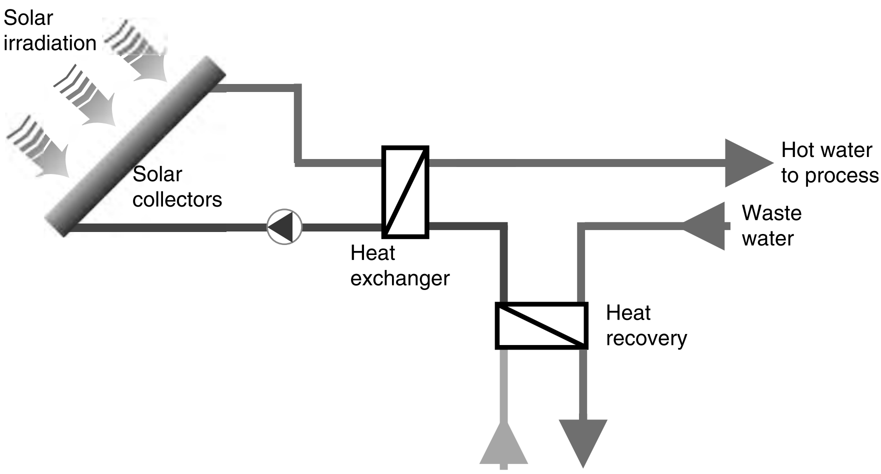


FIGURE 20.33 Combination of solar thermal system and waste heat recovery. (From European Commission, *The Potential of Solar Heat for Industrial Processes*, Final Report, EC Project, Contract No. NNE5-1999-0308, 2001, http://www.aiguasol.com/poship.htm. With permission.)

When there are a large number of streams, the selection of the best match between these streams is not obvious. The pinch methodology provides a systematic way to find the optimal solutions for the implementation of heat recovery techniques.

The solar thermal system may be coupled with the conventional heat supply system in several ways, including direct coupling to a specific process, preheating of water, and steam generation in the central system (Figure 20.34).

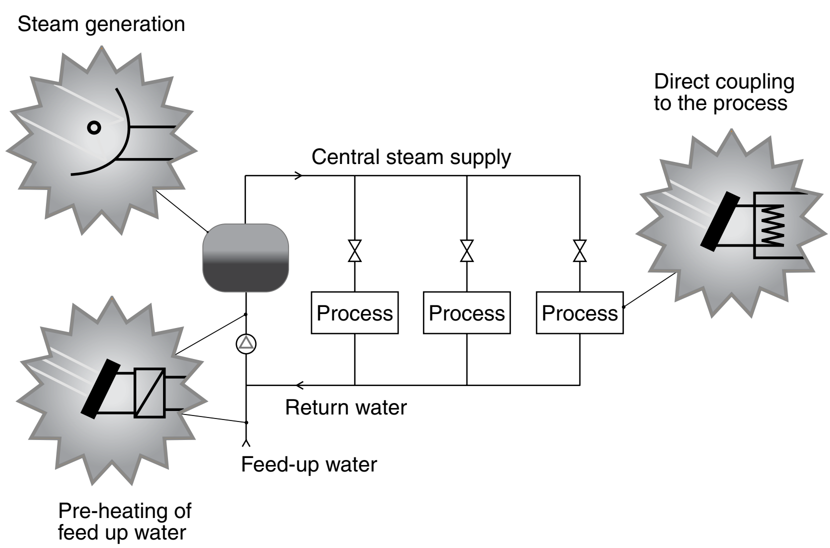


FIGURE 20.34 Coupling the solar thermal system with the conventional heat supply. (From European Commission, *The Potential of Solar Heat for Industrial Processes*, Final Report, EC Project, Contract No. NNE5-1999-0308, 2001, http://www.aiguasol.com/poship.htm. With permission.)

Whenever possible, a direct coupling of the solar thermal systems to one or several processes is preferred because the working temperatures are lower. Direct coupling to a process can mainly be carried out in the following two ways:

- Preheating of a circulating fluid (e.g., feed-up water, return of closed circuits, air preheating). This solution is feasible if fluid circulation is either continuous or periodic (e.g., periodic replacement of bath water). If circulation is discontinuous, a storage tank must be introduced. The mean working temperature of the solar thermal system is lower than the required final process temperature. The smaller the solar fraction, the lower the mean working temperature. For very low solar fractions, the mean working temperature may be close to the fluid inlet (or return) temperature.
- Heating of liquid baths or hot (e.g., drying) chambers. The energy demand is both for heating-up at the operational start-up, either concentrated in the early morning hours or periodically each time the used fluid is replaced with fresh fluid, and for maintaining the operating temperature, which is generally a nearly constant load.

The existing heat exchangers for bath heating generally require steam at temperatures that are too high for a solar thermal system. The introduction of additional heat exchangers with a larger exchange area into existing baths is not always possible due to lack of space or other technical restrictions. In some cases, an external heat exchanger in combination with a circulation pump can be used.

If the process baths are well-insulated, they can be used for solar heat storage. For example, maintaining the temperature of the solar thermal system during a weekend without operation can reduce the heat demand for start-up on Monday morning.

In almost all industries, coupling of a solar thermal system to the central heat supply system is possible. This can be done either by preheating the feed-up water for the steam boilers (the temperature level rises with increasing condensate recovery) or by a solar steam generator. The latter is only recommended at sites with a high level of solar radiation and if concentrating collectors are used.

20.2.2.3 Heat Demand and Storage

Another challenge in applying solar thermal energy to industrial production processes is the time dependency of the solar energy supply and the heat demand of the processes (Figure 20.35). Very few

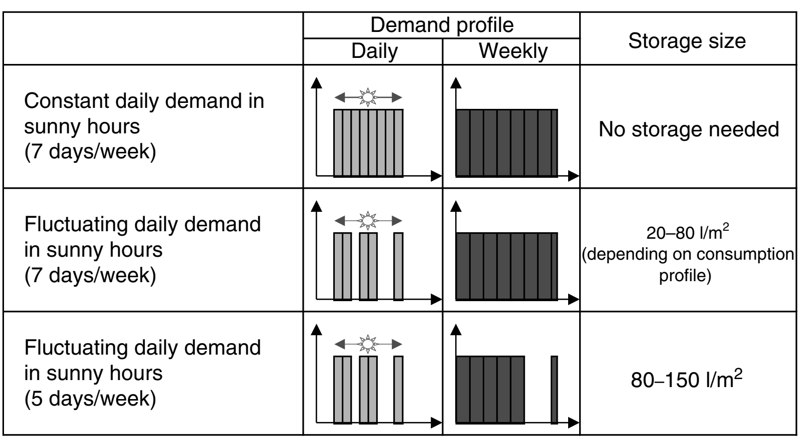


FIGURE 20.35 Heat storage size depending on daily and weekly heat demand profiles.

production lines run at constant load throughout the day. Most processes in smaller companies run for one or two shifts per day and show a batch operation mode.

When the process heat demand is continuous during sunny hours with no weekend breaks, the load is always higher than the solar gains. Therefore, the solar thermal system can be designed without storage allowing the solar heat to be fed directly to the process or to the heat supply system. This enables building lower cost solar thermal systems by eliminating storage related costs.

In case the total weekly demand is constant, but there are strong fluctuations in the daily demand during operational periods (e.g., demand peaks, short breaks of operation), storage of 20–80 l/m^2 of collectors is necessary, depending on the process heat demand profile.

If the daily fluctuations come together with weekend breaks (five days of operation per week), then the storage size is recommended to be 80–150 l/m^2. Weekend storage is generally not recommended for small systems. The larger the system size, the more effective the heat storage over longer periods (e.g., weekends). A weekend storage becomes economic for systems with an installed capacity of about 350 kW_{th}, corresponding to 500 m^2 or more of collector array.

Storage for longer periods (seasonal storage) can only be considered for very large systems (greater than 3.5 MW_{th}), but systems with only seasonal utilization (less than six months of operation a year) are generally not economically viable.

From the previous considerations, key criteria for the feasibility of a solar process heat plant can be drawn:

- Temperature level: solar heat at temperatures above 150°C is technically feasible but economically reasonable only at favorable locations; applications at low temperature (less than 60°C) offer the best economics
- Continuous or quasicontinuous demand (otherwise storage is needed and plant costs increase)
- Technical possibility of introducing a heat exchanger in the existing equipment or heat supply circuit for the solar thermal system

The different possible schemes of solar thermal plants for process heat production are summarized in Figure 20.36. This classification scheme, currently under development and improvement, allows choosing the most suitable plant scheme depending on key process features (open/closed, collector fluid same as or different from heat distribution fluid, need for storage). The scheme is divided into three parts: energy supply (both solar and conventional), heat transfer/storage, and process load. For each column corresponding to a class of processes, two possibilities are given depending on the heat distribution medium: water/air or steam. These concepts are currently being developed and tested in demonstration plants.

20.2.3 Overview of Existing Solar Process Heat Plants

Since the 1980s, several solar thermal systems for industrial applications have been developed and are currently operating. At present, 84 plants are reported worldwide, with a total installed thermal power of about 24 MW_{th} (34,000 m^2).

The majority of the operational plants are in the sectors of food and beverage, textile, and transport (e.g., washing and painting of car components). The main applications differ among countries. In Greece, for instance, several "solar dairies" are operating, whereas in Germany the most common application is for car-washing facilities.

Most of the reported plants supply heat at temperature levels between 60 and 100°C and therefore, standard flat-plate collectors are suitable. As a matter of fact, flat-plate collectors are used in about 65% of the operational systems. Some plants are working at temperatures above 160°C, whereas there is only one project operating in the intermediate range from 100 to 160°C. Analysis of the reported plants reveals that they appear to encompass a broad range of working temperatures, but with no significant correlation with the solar field size. Figure 20.37 through Figure 20.40 show some examples of solar industrial process heat plants.

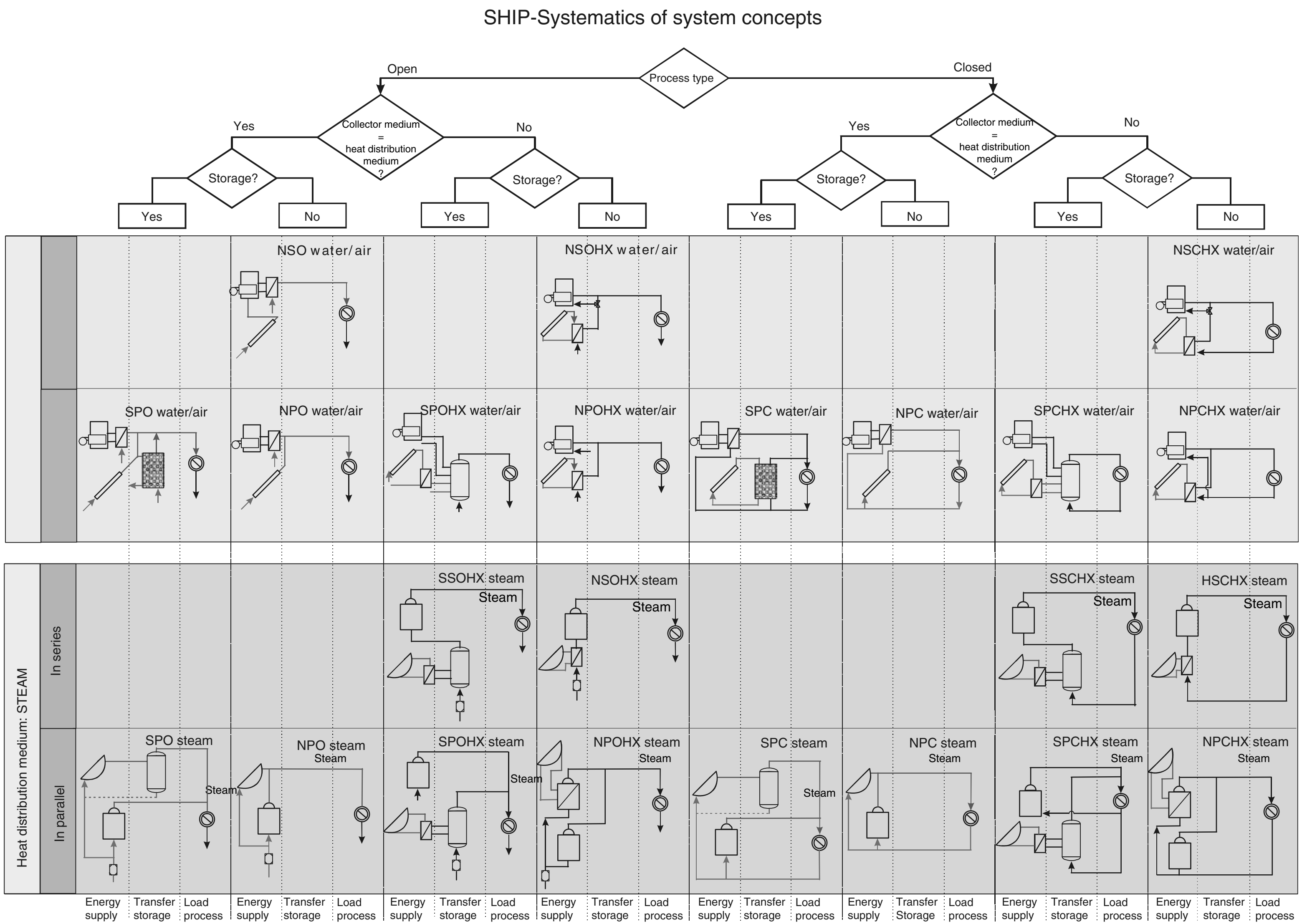


FIGURE 20.36 Systematic of systems concepts. (From IEA., *Task 33/IV, Solar Heat for Industrial Processes*, International Energy Agency, 2005.)

FIGURE 20.37 El NASR Pharmaceutical Chemicals, Egypt. Installed capacity: 1.33 MW_{th}. (From Fichtner Solar GmbH, Germany. With permission.)

FIGURE 20.38 Alpino SA, dairy industry, Greece. Installed capacity: 518 kW_{th}. (From Alpino SA, Greece. With permission.)

FIGURE 20.39 Wine cooling and bottle washing, Austria. Installed capacity: 70 kW_{th}. (From S.O.L.I.D. GmbH, Austria. With permission.)

FIGURE 20.40 Parking service Castellbisbal SA, container washing, Spain. Installed capacity: 357 kW_{th}. (From Aiguasol Engineering, Spain. With permission.)

References

European Commission. 2001. The Potential of Solar Heat for Industrial Processes. Final Report, EC Project, Contract No. NNE5-1999-0308, 2001, http://www.aiguasol.com/poship.htm (accessed on 2006/10/11).

IEA (International Energy Agency). 2005. *Task 33/IV. Solar heat for industrial processes*, International Energy Agency.

Kreider, J. F. 1979. *Medium and High Temperature Solar Energy Processes*, Academic Press, New York.

KWA. 2001. Bedrijfsadviseurs B. V., *Onderzoek naar het potentieel van zonthermishce energie in de industrie*. Relatienummer 8543.00, Rapportnummer 2009740DR01.DOC.

Laue, H. J. and Reichert, J. 1994. *Potential for medium and large sized industrial heat pumps in Europe*. Contract No. XVII/7001/90-8, Final report, European Commission, Directorate General for Energy (DGXII).

Müller, T. et al. 2004. *PROMISE-Produzieren mit Sonnenenergie, Projekt im Rahmen der Programmlinien "Fabrik der Zukunft" des Bundesministeriums für Verkehr*. Innovation und Technologie, Endbericht, Gleisdorf.

PROCESOL. 2000. *Solar Thermal Process Heating in Industrial Applications: A Stimulate Plan*, Final Report, EC ALTENTER Project, Contract No 4.1030/Z/98–205.

20.3 Passive Solar Heating, Cooling, and Daylighting

Jeffrey H. Morehouse

20.3.1 Introduction

Passive systems are defined, quite generally, as systems in which the thermal energy flow is by natural means: by conduction, radiation, and natural convection. A *passive heating system* is one in which the sun's radiant energy is converted to heat upon absorption by the building. The absorbed heat can be transferred to thermal storage by natural means or used to directly heat the building. *Passive cooling systems* use natural energy flows to transfer heat to the environmental sinks: the ground, air, and sky.

If one of the major heat transfer paths employs a pump or fan to force flow of a heat transfer fluid, then the system is referred to as having an active component or subsystem. Hybrid systems—either for heating or cooling—are ones in which there are both passive and active energy flows. The use of the sun's radiant energy for the natural illumination of a building's interior spaces is called *daylighting*. Daylighting design approaches use both solar beam radiation (referred to as *sunlight*) and the diffuse radiation scattered by the atmosphere (referred to as *skylight*) as sources for interior lighting, with historical design emphasis on utilizing skylight.

20.3.1.1 Distinction Between a Passive System and Energy Conservation

A distinction is made between energy conservation techniques and passive solar measures. Energy conservation features are designed to reduce the heating and cooling energy required to thermally condition a building: the use of insulation to reduce either heating and cooling loads, and the use of window shading or window placement to reduce solar gains, reducing summer cooling loads. Passive features are designed to increase the use of solar energy to meet heating and lighting loads, plus the use of ambient "coolth" for cooling. For example, window placement to enhance solar gains to meet winter heating loads and/or to provide daylighting is passive solar use, and the use of a thermal chimney to draw air through the building to provide cooling is also a passive cooling feature.

20.3.1.2 Key Elements of Economic Consideration

The distinction between passive systems, active systems, or energy conservation is not critical for economic calculations, as they are the same in all cases: a trade-off between the life-cycle cost of the energy saved (performance) and the life-cycle cost of the initial investment, operating, and maintenance costs (cost).

20.3.1.2.1 Performance: Net Energy Savings

The key performance parameter to be determined is the net annual energy saved by the installation of the passive system. The basis for calculating the economics of any solar energy system is to compare it against a "normal" building; thus, the actual difference in the annual cost of fuel is the difference in auxiliary energy that would be used with and without solar. Therefore, the energy saved rather than energy delivered, energy collected, useful energy, or some other energy measure, must be determined.

20.3.1.2.2 Cost: Over and Above "Normal" Construction

The other significant part of the economic trade-off involves determining the difference between the cost of construction of the passive building and of the "normal" building against which it is to be compared. The convention, adopted from the economics used for active solar systems, is to define a "solar add-on cost." Again, this may be a difficult definition in the case of passive designs because the building can be significantly altered compared to typical construction since, in many cases, it is not just a one-to-one replacement of a wall with a different wall, but it is more complex and involves assumptions and simulations concerning the "normal" building.

20.3.1.2.3 General System Application Status and Costs

Almost 500,000 buildings in the U.S. were constructed or retrofitted with passive features in the 20 years after 1980. Passive heating applications are primarily in single-family dwellings and secondarily in small commercial buildings. Daylighting features that reduce lighting loads and the associated cooling loads are usually more appropriate for large office buildings.

A typical passive heating design in a favorable climate might supply up to one-third of a home's original load at a cost of \$5 to \$10 per million Btu net energy saved. An appropriately designed daylighting system can supply lighting at a cost of 2.5–5 ¢ per kWh (Larson, Vignola, and West 1992a, 1992b, 1992c).

20.3.2 Solar Thermosyphon Water Heating

Solar hot-water heating systems are composed of a collector and a storage tank. When the flow between the collector and tank is by natural circulation, these passive solar hot-water systems are referred to as *thermosyphon systems*. This ability of thermosyphon systems to heat water without an externally powered pump has spurred its use in both regions where power is unavailable and where power is very expensive.

20.3.2.1 Thermosyphon Concept

The natural tendency of a less dense fluid to rise above a more dense fluid can be used in a simple solar water heater to cause fluid motion through a collector. The density difference is created within the solar collector where heat is added to increase the temperature and decrease the density of the liquid. This collection concept is called a *thermosyphon*, and Figure 20.41 schematically illustrates the major components of such a system.

The flow pressure drop in the fluid loop (ΔP_{FLOW}) must equal the buoyant force "pressure difference" ($\Delta P_{\mathrm{BUOYANT}}$) caused by the differing densities in the "hot" and "cold" legs of the fluid loop:

$$\begin{aligned} \Delta P_{\mathrm{FLOW}} &= \Delta P_{\mathrm{BUOYANT}} \\ &= \rho_{\mathrm{stor}} g H - \left[\int_O^L \rho(x) g \, \mathrm{d}x + \rho_{\mathrm{out}} g (H - L) \right], \end{aligned} \tag{20.68}$$

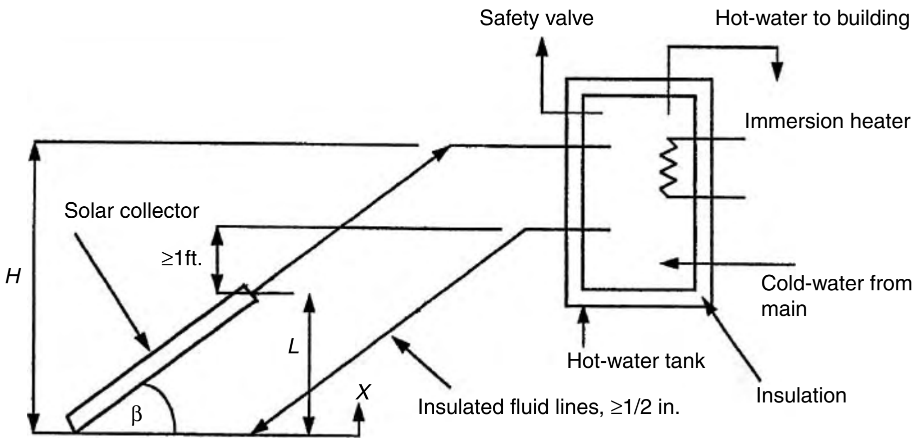


FIGURE 20.41 Schematic diagram of thermosyphon loop used in a natural circulation, service water-heating system. The flow pressure drop in the fluid loop must equal the bouyant force "pressure" $[\int_0^L g\rho(x)dx - \rho_{stor}gL]$ where $\rho(x)$ is the local collector fluid density and ρ_{stor} is the tank fluid density, assumed uniform.

where H is the height of the "legs," L is the height of the collector (see Figure 20.41), $\rho(x)$ is the local collector fluid density,ρ_{stor} is the tank fluid density, andρ_{out} is the collector outlet fluid density; the latter two densities are assumed to be uniform. The flow pressure term, ΔP_{FLOW}, is related to the flow loop system headloss that is in turn directly connected to friction and fitting losses and the loop flow rate:

$$\Delta P_{FLOW} = \oint_{LOOP} \rho \, d(h_L), \tag{20.69}$$

where $h_L = KV^2$, with K being the sum of the component loss "velocity" factors (see any fluid mechanics text), and V is the flow velocity.

20.3.2.2 Thermo-Fluid System Design Considerations

Because the driving force in a thermosyphon system is only a small density difference and not a pump, larger than normal plumbing fixtures must be used to reduce pipe friction losses. In general, one pipe size larger than would be used with a pump system is satisfactory. Under no conditions should piping smaller than 1/2-in (12-mm) national pipe thread (NPT) be used. Most commercial thermosyphons use 1-in (25-mm) NPT pipe. The flow rate through a thermosyphon system is about 1 gal/ft.^{2} h (40 L/m^2 h) in bright sun, based on collector area.

Because the hot-water system loads vary little during a year, the best angle to tilt the collector is that equal to the local latitude. The temperature difference between the collector inlet water and the collector outlet water is usually 15–20°F (8–11°C) during the middle of a sunny day (Close 1962). After sunset, a thermosyphon system can reverse its flow direction and lose heat to the environment during the night. To avoid reverse flow, the top header of the absorber should be at least 1 ft. (30 cm) below the cold leg fitting on the storage tank, as shown.

To provide heat during long cloudy periods, an electrical immersion heater can be used as a backup for the solar system. The immersion heater is located near the top of the tank to enhance stratification and so that the heated fluid is at the required delivery temperature at the delivery point. Tank stratification is desirable in a thermosyphon to maintain flow rates as high as possible. Insulation must be applied over

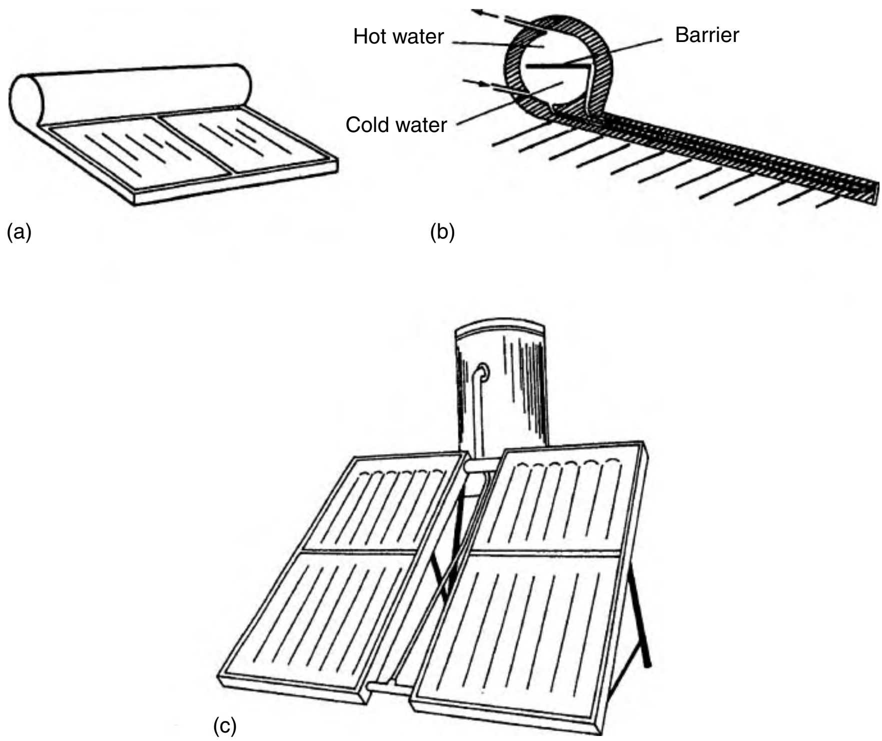


FIGURE 20.42 Passive solar water heaters; (a) compact model using combined collector and storage, (b) section view of the compact model, and (c) tank and collector assembly.

the entire tank surface to control heat loss. Figure 20.42 illustrates two common thermosyphon system designs.

Several features inherent in the thermosyphon design limit its utility. If it is to be operated in a freezing climate, a nonfreezing fluid must be used, which in turn requires a heat exchanger between collector and potable water storage. (If potable water is not required, the collector can be drained during cold periods instead.) Heat exchangers of either the shell-and-tube type or the immersion-coil type require higher flow rates for efficient operation than a thermosyphon can provide. Therefore, the thermosyphon is generally limited to nonfreezing climates. A further restriction on thermosyphon use is the requirement for an elevated tank. In many cases structural or architectural constraints prohibit raised-tank locations. In residences, collectors are normally mounted on the roof, and tanks mounted above the high point of the collector can easily become the highest point in a building. Practical considerations often do not permit this application.

Example 20.3.1

Determine the "pressure difference" available for a thermosyphon system with 1-m high collector and 2-m high "legs." The water temperature input to the collector is 25°C and the collector outputtemperature is 35°C. If the overall system loss velocity factor (K) is 15.6, estimate the system flow velocity.

Solution. Equation 20.68 is used to calculate the pressure difference, with the water densities being found from the temperatures (in steam tables):

$$\rho_{\text{stor}}(25°\text{C}) = 997.009\ \text{kg/m}^3;$$

$$\rho_{\text{out}}(35°\text{C}) = 994.036\ \text{kg/m}^3;$$
$$\rho_{\text{coll.ave.}}(30°\text{C}) = 996.016\ \text{kg/m}^3$$

(note: average collector temperature used in "temperature") and with $H = 2$ m and $L = 1$ m,

$$\begin{aligned}\Delta P_{\text{BUOYANT}} &= (997.009)9.81(2) - [(996.016)9.81(1) + (994.036)9.81(1)] \\ &= 38.9\ \text{N/m}^2(\text{Pa}).\end{aligned}$$

The system flow velocity is estimated from the "system K" given, the pressure difference calculated above, taking the average density of the water around the loop (at 30°C), and substituting into Equation 20.69:

$$\Delta P_{\text{BUOYANT}} = (\rho_{\text{loop.ave}})(h_L)_{\text{loop}} = (\rho_{\text{loop.ave}})KV^2,$$

$$V^2 = 38.9/(996.016)(15.6),$$

$$V = 0.05\ \text{m/s}.$$

20.3.3 Passive Solar Heating Design Fundamentals

Passive heating systems contain the five basic components of all solar systems, as described in the previous chapter on Active Solar Systems. Typical passive realizations of these components are:

1. Collector: windows, walls and floors
2. Storage: walls and floors, large interior masses (often these are integrated with the collector absorption function)
3. Distribution system: radiation, free convection, simple circulation fans
4. Controls: moveable window insulation, vents both to other inside spaces or to ambient
5. Backup system: any nonsolar heating system

The design of passive systems requires the strategic placement of windows, storage masses, and the occupied spaces themselves. The fundamental principles of solar radiation geometry and availability are instrumental in the proper location and sizing of the system's "collectors" (windows). Storage devices are usually more massive than those used in active systems and are frequently an integral part of the collection and distribution system.

20.3.3.1 Types of Passive Heating Systems

A commonly used method of cataloging the various passive system concepts is to distinguish three general categories: direct, indirect, and isolated gain. Most of the physical configurations of passive heating systems are seen to fit within one of these three categories.

For direct gain (Figure 20.43), sunlight enters the heated space and is converted to heat at absorbing surfaces. This heat is then distributed throughout the space and to the various enclosing surfaces and room contents.

For indirect gain category systems, sunlight is absorbed and stored by a mass interposed between the glazing and the conditioned space. The conditioned space is partially enclosed and bounded by this thermal storage mass, so a natural thermal coupling is achieved. Examples of the indirect approach are the thermal storage wall, the thermal storage roof, and the northerly room of an attached sunspace.

In the thermal storage wall (Figure 20.44), sunlight penetrates the glazing and is absorbed and converted to heat at a wall surface interposed between the glazing and the heated space. The wall is usually masonry (Trombe wall) or containers filled with water (water wall), although it might contain phase-change material. The attached sunspace (Figure 20.45) is actually a two-zone combination of direct gain and thermal storage wall. Sunlight enters and heats a direct gain southerly "sunspace" and also heats a

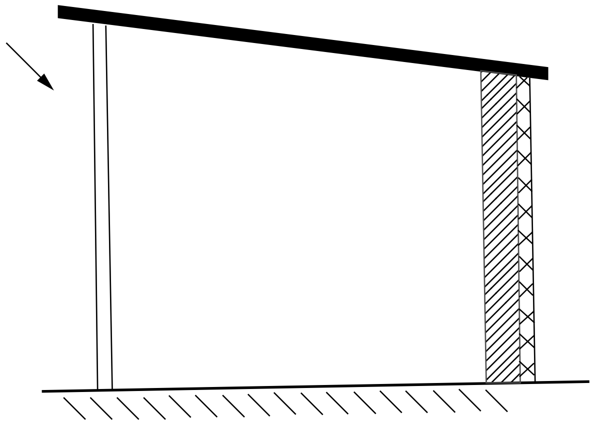


FIGURE 20.43 Direct gain.

mass wall separating the northerly buffered space, which is heated indirectly. The “sunspace” is frequently used as a greenhouse, in which case, the system is called an “attached greenhouse.” The thermal storage roof (Figure 20.46) is similar to the thermal storage wall except that the interposed thermal storage mass is located on the building roof.

The isolated gain category concept is an indirect system, except that there is a distinct thermal separation (by means of either insulation or physical separation) between the thermal storage and the heated space. The convective (thermosyphon) loop, as depicted in Figure 20.41, is in this category and, while often used to heat domestic water, is also used for building heating. It is most akin to conventional active systems in that there is a separate collector and separate thermal storage. The thermal storage wall, thermal storage roof, and attached sunspace approaches can also be made into isolated systems by insulating between the thermal storage and the heated space.

20.3.3.2 Fundamental Concepts for Passive Heating Design

Figure 20.47 is an equivalent thermal circuit for the building illustrated in Figure 20.44, the Trombe wall-type system. For the heat transfer analysis of the building, three temperature nodes can be identified: room temperature, storage wall temperature, and the ambient temperature. The circuit responds to

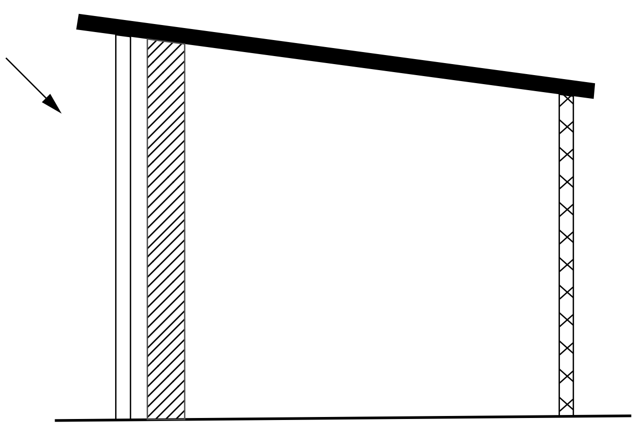


FIGURE 20.44 Thermal storage wall.

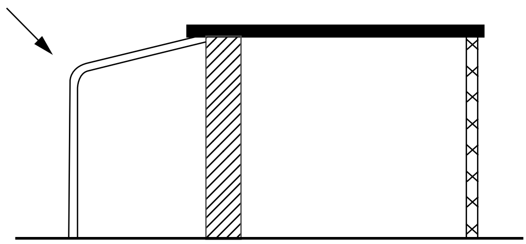


FIGURE 20.45 Attached sunspace.

climatic variables represented by a current injection I_s (solar radiation) and by the ambient temperature T_a. The storage temperature, T_s, and room temperature, T_r, are determined by current flows in the equivalent circuit. By using seasonal and annual climatic data, the performance of a passive structure can be simulated and the results of many such simulations correlated to give the design approaches described below.

20.3.3.3 Passive Design Approaches

Design of a passive heating system involves selection and sizing of the passive feature type(s), determination of thermal performance, and cost estimation. Ideally, a cost/performance optimization would be performed by the designer. Owner and architect ideas usually establish the passive feature type, with general size and cost estimation available. However, the thermal performance of a passive heating system has to be calculated.

There are several "levels" of methods that can be used to estimate the thermal performance of passive designs. First-level methods involve a rule of thumb and/or generalized calculation to get a starting estimate for size and/or annual performance. A second-level method involves climate, building, and passive system details, which allow annual performance determination, plus some sensitivity to passive system design changes. Third-level methods involve periodic calculations (hourly, monthly) of performance and permit more detailed variations of climatic, building, and passive solar system design parameters.

These three levels of design methods have a common basis in that they all are derived from correlations of a multitude of computer simulations of passive systems (PSDH 1980, 1984). As a result, a similar set of defined terms is used in many passive design approaches:

- A_p, solar projected area, m^2 ($ft.^2$): the net south-facing passive solar glazing area projected onto a vertical plane

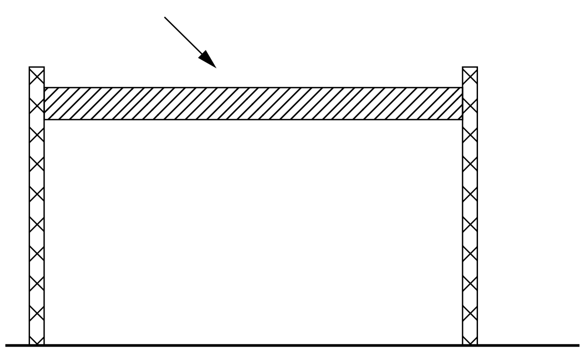


FIGURE 20.46 Thermal storage roof.

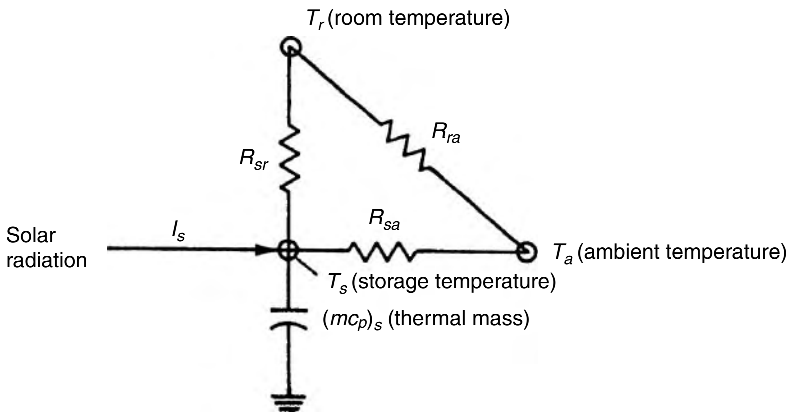


FIGURE 20.47 Equivalent thermal circuit for passively heated solar building in Figure 20.44.

- NLC, net building load coefficient, kJ/CDD (Btu/FDD): net load of the nonsolar portion of the building per degree day of indoor–outdoor temperature difference. The CDD and FDD terms refer to Celsius and Fahrenheit degree days, respectively

- Q_{net}, net reference load, Wh (Btu): heat loss from nonsolar portion of building as calculated by

$$Q_{net} = \text{NLC} \times (\text{Number of degree days}). \tag{20.70}$$

- LCR, load collector ratio, kJ/m^2 CDD (Btu/ft.^{2} FDD): ratio of NLC to A_p,

$$\text{LCR} = \text{NLC}/A_p \tag{20.71}$$

- SSF, solar savings fraction, %: percentage reduction in required auxiliary heating relative to net reference load,

$$\text{SSF} = 1 - \frac{\text{Auxiliary heat required } (Q_{aux})}{\text{Net reference load } (Q_{net})} \tag{20.72}$$

Therefore, using Equation 20.70, the auxiliary heat required, Q_{aux}, is given by

$$Q_{aux} = (1 - \text{SSF}) \times \text{NLC} \times (\text{Number of degree days}). \tag{20.73}$$

The amount of auxiliary heat required is often a basis of comparison between possible solar designs as well as being the basis for determining building energy operating costs. Thus, many of the passive design methods are based on determining SSF, NLC, and the number of degree days in order to calculate the auxiliary heat required for a particular passive system by using Equation 20.73.

20.3.3.4 The First Level: Generalized Methods

A first estimate or starting value is needed to begin the overall passive system design process. Generalized methods and rules of thumb have been developed to generate initial values for solar aperture size, storage size, solar savings fraction, auxiliary heat required, and other size and performance characteristics. The following rules of thumb are meant to be used with the defined terms presented above.

20.3.3.5 Load

A rule of thumb used in conventional building design is that a design heating load of 120–160 kJ/CDD per m^2 of floor area (6–8 Btu/FDD ft.^{2}) is considered an energy conservative design. Reducing these non-solar values by 20% to solarize the proposed south-facing solar wall gives rule-of-thumb NLC values per unit of floor area:

$$\text{NLC/Floor area} = 100\text{–}130\ \text{kJ/CDD m}^2\ (4.8\text{–}6.4\ \text{Btu/FDD ft.}^2). \tag{20.74}$$

20.3.3.6 Solar Savings Fraction

A method of getting starting-point values for the solar savings fraction is presented in Figure 20.48 (PSDH 1984). The map values represent optimum SSF (in percent) for a particular set of conservation and passive-solar costs for different climates across the United States. With the Q_{net} generated from the NLC rule of thumb (see above) and the SSF read from the map, the Q_{aux} can be determined.

20.3.3.7 Load Collector Ratio (LCR)

The A_{p} can be determined using the NLC from above if the LCR is known. The rule of thumb associated with "good" values of LCR (PSDH 1984) differs depending on whether the design is for a "cold" or "warm" climate:

$$\text{"Good" LCR} = \begin{cases} \text{For cold climate : 410 kJ/m}^2\ \text{CDD (20 Btu/ft.}^2\ \text{FDD)} \\ \text{For warm climate : 610 kJ/m}^2\ \text{CDD (30 Btu/ft.}^2\ \text{FDD)} \end{cases} \tag{20.75}$$

20.3.3.8 Storage

Rules of thumb for thermal mass storage relate storage material total heat capacity to the solar projected area (PSDH 1984). The use of the storage mass is to provide for heating on cloudy days and to regulate sunny day room air temperature swing. When the thermal mass directly absorbs the solar radiation, each

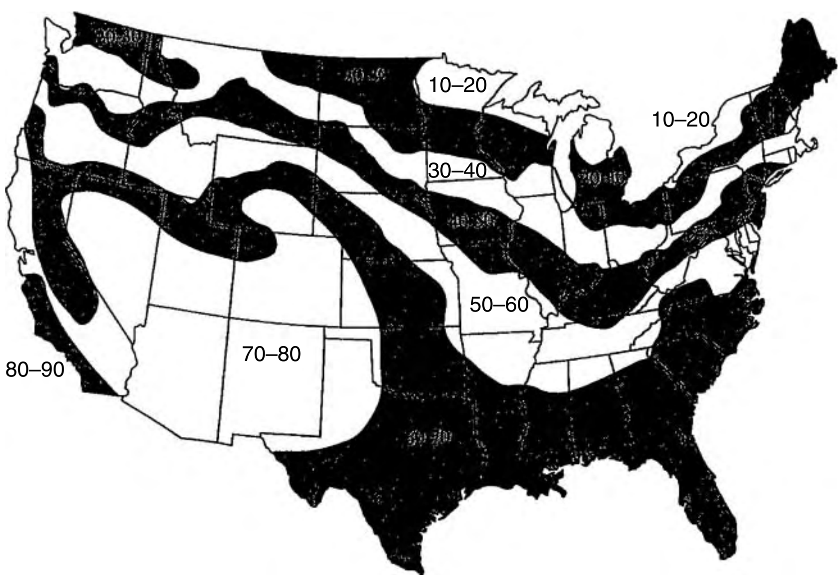


FIGURE 20.48 Starting-point values of solar savings fraction (SSF) in percent. (From PSDH, *Passive Solar Design Handbook*. Part One: Total Environmental Action, Inc., Part Two: Los Alamos Scientific Laboratory, Part Three: Los Alamos National Laboratory. Van Nostrand Reinhold, New York, 1984.)

square meter of the projected glazing area requires enough mass to store 613 kJ/°C. If the storage material is not in direct sunlight, but heated from room air only, then four times as much mass is needed. In a room with a directly sunlight-heated storage mass, the room air temperature swing will be approximately one-half the storage mass temperature swing. For room air heated storage, the air temperature swing is twice that of the storage mass.

Example 20.3.2

A Denver, Colorado, building is to have a floor area of 195 m^2 (2100 $ft.^2$). Determine rule-of-thumb size and performance characteristics.

Solution. From Equation 20.72, the NLC is estimated as

$$\begin{aligned}\text{NLC} &= (115\ \text{kJ/CDD m}^2) \times (195\ \text{m}^2)\\ &= 22,400\ \text{kJ/CDD}\ (11,800\ \text{Btu/FDD}).\end{aligned}$$

Using the "cold" LCR value and Equation 20.71, the passive solar projected area is

$$\begin{aligned}A_{\text{p}} &= \text{NLC/LCR} = (22,400\ \text{kJ/CDD})/(410\ \text{kJ/m}^2\ \text{CDD})\\ &= 54.7\ \text{m}^2(588\ \text{ft.}^2)\end{aligned}$$

Locating Denver on the map of Figure 20.48 gives an SSF value in the 70%–80% range (use 75%). An annual °C-degree-day value can be found in city climate tables (PSDH 1984; NCDC 1992), and is 3491 CDD (6283 FDD) for Denver. Thus, the auxiliary heat required, Q_{aux}, is found using Equation 20.73:

$$\begin{aligned}Q_{\text{aux}} &= (1-0.75)(22,400\ \text{kJ/CDD})(3491\ \text{CDD})\\ &= 19,600\ \text{MJ}\ (18.5\times 10^6\ \text{Btu})\ \text{annually}.\end{aligned}$$

The thermal storage can be sized using directly solar-heated and/or room air heated mass by using the projected area. Assuming brick with a specific heat capacity of 840 J/kg°C, the storage mass is found by

$$\begin{aligned}A_p \times (613\ \text{kJ/C}) &= m \times (840\ \text{J/kg°C});\\ m_{\text{d}} &= 40,000\ \text{kg}\ (88,000\ \text{lbm})\ [\text{Direct sun}]\\ \text{or}\quad m_{\text{a}} &= 160,000\ \text{kg}\ (351,000\ \text{lbm})\ [\text{Air heated}]\end{aligned}$$

A more location-dependent set of rules of thumb is presented in PSDH (1980). The first rule of thumb relates solar projected area as a percentage of floor area to solar savings fraction, with and without night insulation of the solar glazing:

"A solar projected area of (B1)% to (B2)% of the floor area can be expected to produce a SSF in (location) of (S1)% to (S2)%, or, if R9 night insulation is used, of (S3)% to (S4)%."

The values of B1, B2, S1, S2, S3 and S4 are found using Table 20.6 for the location. The thermal storage mass rule of thumb is again related to the solar projected area:

"A thermal storage wall should have 14 kg × SSF (%) of water or 71 kg × SSF (%) of masonry for each square meter of solar projected area. For a direct gain space, the mass above should be used with a surface area of at least three times the solar projected area, and masonry no thicker than 10–15 cm. If the mass is located in back rooms, then four times the above mass is needed."

TABLE 20.6 Values to Be Used in the Glazing Area and SSF Relations Rules of Thumb

City	B1	B2	S1	S2	S3	S4
Birmingham, Alabama	0.09	0.18	22	37	34	58
Mobile, Alabama	0.06	0.12	26	44	34	60
Montgomery, Alabama	0.07	0.15	24	41	34	59
Phoenix, Arizona	0.06	0.12	37	60	48	75
Prescott, Arizona	0.10	0.20	29	48	44	72
Tucson, Arizona	0.06	0.12	35	57	45	73
Winslow, Arizona	0.12	0.24	30	47	48	74
Yuma, Arizona	0.04	0.09	43	66	51	78
Fort Smith, Arkansas	0.10	0.20	24	39	38	64
Little Rock, Arkansas	0.10	0.19	23	38	37	62
Bakersfield, California	0.08	0.15	31	50	42	67
Baggett, California	0.07	0.15	35	56	46	73
Fresno, California	0.09	0.17	29	46	41	65
Long Beach, California	0.05	0.10	35	58	44	72
Los Angeles, California	0.05	0.09	36	58	44	72
Mount Shasta, California	0.11	0.21	24	38	42	67
Needles, California	0.06	0.12	39	61	49	76
Oakland, California	0.07	0.15	35	55	46	72
Red Bluff, California	0.09	0.18	29	46	41	65
Sacramento, California	0.09	0.18	29	47	41	66
San Diego, California	0.04	0.09	37	61	46	74
San Francisco, California	0.06	0.13	34	54	45	71
Santa Maria, California	0.05	0.11	31	53	42	69
Colorado Springs, Colorado	0.12	0.24	27	42	47	74
Denver, Colorado	0.12	0.23	27	43	47	74
Eagle, Colorado	0.14	0.29	25	35	53	77
Grand Junction, Colorado	0.13	0.27	29	43	50	76
Pueblo, Colorado	0.11	0.23	29	45	48	75
Hartford, Connecticut	0.17	0.35	14	19	40	64
Wilmington, Delaware	0.15	0.29	19	30	39	63
Washington, District Of Columbia	0.12	0.23	18	28	37	61
Apalachicola, Florida	0.05	0.10	28	47	36	61
Daytona Beach, Florida	0.04	0.08	30	51	36	63
Jacksonville, Florida	0.05	0.09	27	47	35	62
Miami, Florida	0.01	0.02	27	48	31	54
Orlando, Florida	0.03	0.06	30	52	37	63
Tallahassee, Florida	0.05	0.11	26	45	35	60
Tampa, Florida	0.03	0.06	30	52	36	63
West Palm Beach, Florida	0.01	0.03	30	51	34	59
Atlanta, Georgia	0.06	0.17	22	36	34	58
Augusta, Georgia	0.06	0.16	24	40	35	60
Macon, Georgia	0.07	0.15	25	41	35	59
Savannah, Georgia	0.06	0.13	25	43	35	60
Boise, Idaho	0.14	0.28	27	38	48	71
Lewiston, Idaho	0.15	0.29	22	29	44	65
Pocatello, Idaho	0.13	0.26	25	35	51	74
Chicago, Illinois	0.17	0.35	17	23	43	67
Moline, Illinois	0.20	0.39	17	22	46	70
Springfield, Illinois	0.15	0.30	19	26	42	67
Evansville, Indiana	0.14	0.27	19	29	37	61
Fort Wayne, Indiana	0.16	0.33	13	17	37	60
Indianapolis, Indiana	0.14	0.28	15	21	37	60
South Bend, Indiana	0.18	0.35	12	15	39	61
Burlington, Iowa	0.18	0.36	20	27	47	71
Des Moines, Iowa	0.21	0.43	19	25	58	75

(continued)

TABLE 20.6 (*Continued*)

City	B1	B2	S1	S2	S3	S4
Mason City, Iowa	0.22	0.44	18	19	56	79
Sioux City, Iowa	0.23	0.46	20	24	53	76
Dodge City, Kansas	0.12	0.23	27	42	46	73
Goodland, Kansas	0.13	0.27	26	39	47	74
Topeka, Kansas	0.14	0.26	24	35	45	71
Wichita, Kansas	0.14	0.26	26	41	45	72
Lexington, Kentucky	0.13	0.27	17	26	35	58
Louisville, Kentucky	0.13	0.27	18	27	35	59
Baton Rouge, Louisiana	0.06	0.12	26	43	34	59
Lake Charles, Louisiana	0.06	0.11	24	41	32	57
New Orleans, Louisiana	0.05	0.11	27	46	35	61
Shreveport, Louisiana	0.08	0.15	26	43	36	61
Caribou, Maine	0.25	0.30	NR	NR	53	74
Portland, Maine	0.17	0.34	14	17	45	69
Baltimore, Maryland	0.14	0.27	19	30	38	62
Boston, Massachusetts	0.15	0.29	17	25	40	64
Alpena, Michigan	0.21	0.42	NR	NR	47	69
Detroit, Michigan	0.17	0.34	13	17	39	61
Flint, Michigan	0.15	0.31	11	12	40	62
Grand Rapids, Michigan	0.19	0.38	12	13	39	61
Sault Ste. Marie, Michigan	0.25	0.50	NR	NR	50	70
Traverse City, Michigan	0.18	0.36	NR	NR	42	62
Duluth, Minnesota	0.25	0.50	NR	NR	50	70
International Falls, Minnesota	0.25	0.50	NR	NR	47	66
Minneapolis-St. Paul, Minnesota	0.25	0.50	NR	NR	55	76
Rochester, Minnesota	0.24	0.49	NR	NR	54	76
Jackson, Mississippi	0.06	0.15	24	48	34	59
Meridian, Mississippi	0.08	0.15	23	39	34	58
Columbia, Missouri	0.13	0.26	20	30	41	66
Kansas City, Missouri	0.14	0.29	22	32	44	70
Saint Louis, Missouri	0.15	0.29	21	33	41	65
Springfield, Missouri	0.13	0.26	22	34	40	65
Billings, Montana	0.16	0.32	24	31	53	76
Cut Bank, Montana	0.24	0.49	22	23	62	81
Dillon, Montana	0.16	0.32	24	32	54	77
Glasgow, Montana	0.25	0.50	NR	NR	55	75
Great Falls, Montana	0.18	0.37	23	26	56	77
Helena, Montana	0.20	0.39	21	25	55	77
Lewistown, Montana	0.19	0.38	21	25	54	76
Miles City, Montana	0.23	0.47	21	23	60	80
Missoula, Montana	0.18	0.36	15	16	47	68
Grand Island, Nebraska	0.18	0.36	24	33	51	76
North Omaha, Nebraska	0.20	0.48	21	29	51	76
North Platte, Nebraska	0.17	0.34	25	36	50	76
Scottsbluff, Nebraska	0.16	0.31	24	36	49	74
Elko, Nevada	0.12	0.25	27	39	52	76
Ely, Nevada	0.12	0.23	27	41	50	77
Las Vegas, Nevada	0.09	0.18	35	56	48	75
Lovelock, Nevada	0.13	0.25	32	48	53	78
Reno, Nevada	0.11	0.22	31	48	49	76
Tonopah, Nevada	0.11	0.23	31	48	51	77
Winnemucca, Nevada	0.13	0.26	28	42	49	75
Concord, New Hampshire	0.17	0.34	13	15	45	68
Newark, New Jersey	0.13	0.25	19	29	39	64
Albuquerque, New Mexico	0.11	0.22	29	47	46	73

(continued)

TABLE 20.6 (*Continued*)

City	B1	B2	S1	S2	S3	S4
Clayton, New Mexico	0.10	0.20	28	45	45	73
Farmington, New Mexico	0.12	0.24	29	45	49	76
Los Alamos, New Mexico	0.11	0.22	25	40	44	72
Roswell, New Mexico	0.10	0.19	30	49	45	73
Truth or Consequences, New Mexico	0.09	0.17	32	51	46	73
Tucumcari, New Mexico	0.10	0.20	30	48	45	73
Zuni, New Mexico	0.11	0.21	27	43	45	73
Albany, New York	0.21	0.41	13	15	43	66
Binghamton, New York	0.15	0.30	NR	NR	35	56
Buffalo, New York	0.19	0.37	NR	NR	36	57
Massena, New York	0.25	0.50	NR	NR	50	71
New York (Central Park), New York	0.15	0.30	16	25	36	59
Rochester, New York	0.18	0.37	NR	NR	37	58
Syracuse, New York	0.19	0.38	NR	NR	37	59
Asheville, North Carolina	0.10	0.20	21	35	36	61
Cape Hatteras, North Carolina	0.09	0.17	24	40	36	60
Charlotte, North Carolina	0.08	0.17	23	38	36	60
Greensboro, North Carolina	0.10	0.20	23	37	37	63
Raleigh-Durham, North Carolina	0.09	0.19	22	37	36	61
Bismarck, North Dakota	0.25	0.50	NR	NR	56	77
Fargo, North Dakota	0.25	0.50	NR	NR	51	72
Minot, North Dakota	0.25	0.50	NR	NR	52	72
Akron-Canton, Ohio	0.15	0.31	12	16	35	57
Cincinnati, Ohio	0.12	0.24	15	23	35	57
Cleveland, Ohio	0.15	0.31	11	14	34	55
Columbus, Ohio	0.14	0.28	13	18	35	57
Dayton, Ohio	0.14	0.28	14	20	36	59
Toledo, Ohio	0.17	0.34	13	17	38	61
Youngstown, Ohio	0.16	0.32	NR	NR	34	54
Oklahoma City, Oklahoma	0.11	0.22	25	41	41	67
Tulsa, Oklahoma	0.11	0.22	24	38	40	65
Astoria, Oregon	0.09	0.19	21	34	37	60
Burns, Oregon	0.13	0.25	23	32	47	71
Medford, Oregon	0.12	0.24	21	32	38	60
North Bend, Oregon	0.09	0.17	25	42	38	64
Pendleton, Oregon	0.14	0.27	22	30	43	64
Portland, Oregon	0.13	0.26	21	31	38	60
Redmond, Oregon	0.13	0.27	26	38	47	71
Salem, Oregon	0.12	0.24	21	32	37	59
Allentown, Pennsylvania	0.15	0.29	16	24	39	63
Erie, Pennsylvania	0.17	0.34	NR	NR	35	55
Harrisburg, Pennsylvania	0.13	0.26	17	26	38	62
Philadelphia, Pennsylvania	0.15	0.29	19	29	38	62
Pittsburgh, Pennsylvania	0.14	0.28	12	16	33	55
Wilkes-Barre-Scranton, Pennsylvania	0.16	0.32	13	18	37	60
Providence, Rhode Island	0.15	0.30	17	24	40	64
Charleston, South Carolina	0.07	0.14	25	41	34	59
Columbia, South Carolina	0.08	0.17	25	41	36	61
Greenville-Spartanburg, South Carolina	0.08	0.17	23	38	36	60
Huron, South Dakota	0.25	0.50	NR	NR	58	79
Pierre, South Dakota	0.22	0.43	21	23	58	80
Rapid City, South Dakota	0.15	0.30	23	32	51	76
Sioux Falls, South Dakota	0.22	0.45	18	19	57	79
Chattanooga, Tennessee	0.09	0.19	19	32	33	56
Knoxville, Tennessee	0.09	0.18	20	33	33	56

(*continued*)

TABLE 20.6 (*Continued*)

City	B1	B2	S1	S2	S3	S4
Memphis, Tennessee	0.09	0.19	22	36	36	60
Nashville, Tennessee	0.10	0.21	19	30	33	55
Abilene, Texas	0.09	0.18	29	47	41	68
Amarillo, Texas	0.11	0.22	29	46	45	72
Austin, Texas	0.06	0.13	27	46	37	63
Brownsville, Texas	0.03	0.06	27	46	32	57
Corpus Christi, Texas	0.05	0.09	29	49	36	63
Dallas, Texas	0.08	0.17	27	44	38	64
Del Rio, Texas	0.06	0.12	30	50	39	66
El Paso, Texas	0.09	0.17	32	53	45	72
Forth Worth, Texas	0.09	0.17	26	44	38	64
Houston, Texas	0.06	0.11	25	43	34	59
Laredo, Texas	0.05	0.09	31	52	39	64
Lubbock, Texas	0.09	0.19	30	49	44	72
Lufkin, Texas	0.07	0.14	26	43	35	61
Midland-Odessa, Texas	0.09	0.18	32	52	44	72
Port Arthur, Texas	0.06	0.11	26	44	34	60
San Angelo, Texas	0.08	0.15	29	48	40	67
San Antonio, Texas	0.06	0.12	28	48	38	64
Sherman, Texas	0.10	0.20	25	41	38	64
Waco, Texas	0.06	0.15	27	45	38	64
Wichita Falls, Texas	0.10	0.20	27	45	41	67
Bryce Canyon, Utah	0.13	0.25	26	39	52	78
Cedar City, Utah	0.12	0.24	28	43	48	75
Salt Lake City, Utah	0.13	0.26	27	39	48	72
Burlington, Vermont	0.22	0.43	NR	NR	46	68
Norfolk, Virginia	0.09	0.19	23	38	37	62
Richmond, Virginia	0.11	0.22	21	34	37	61
Roanoke, Virginia	0.11	0.23	21	34	37	61
Olympia, Washington	0.12	0.23	20	29	38	59
Seattle-Tacoma, Washington	0.11	0.22	21	30	39	59
Spokane, Washington	0.20	0.39	20	24	48	68
Yakima, Washington	0.18	0.36	24	31	49	70
Charleston, West Virginia	0.13	0.25	16	24	32	54
Huntington, West Virginia	0.13	0.25	17	27	34	57
Eau Claire, Wisconsin	0.25	0.50	NR	NR	53	75
Green Bay, Wisconsin	0.23	0.46	NR	NR	53	75
La Crosse, Wisconsin	0.21	0.43	NR	NR	52	75
Madison, Wisconsin	0.20	0.40	15	17	51	74
Milwaukee, Wisconsin	0.18	0.35	15	18	48	71
Casper, Wyoming	0.13	0.26	27	39	53	78
Cheyenne, Wyoming	0.11	0.21	25	39	47	74
Rock Springs, Wyoming	0.14	0.28	26	38	54	79
Sheridan, Wyoming	0.16	0.31	22	30	52	75
		Canada				
Edmonton, Alberta	0.25	0.50	NR	NR	54	72
Suffield, Alberta	0.25	0.50	28	30	67	85
Nanaimo, British Columbia	0.13	0.26	26	35	45	66
Vancouver, British Columbia	0.13	0.26	20	28	48	60
Winnipeg, Manitoba	0.25	0.50	NR	NR	54	74
Dartmouth, Nova Scotia	0.14	0.28	17	24	45	70
Moosonee, Ontario	0.25	0.50	NR	NR	48	67
Ottawa, Ontario	0.25	0.50	NR	NR	59	80
Toronto, Ontario	0.18	0.36	17	23	44	68
Normandie, Quebec	0.25	0.50	NR	NR	54	74

Note: NR, not recommended.
Source: From PSDH, *Passive Solar Design Handbook*, U.S. Department of Energy, Washington, DC, 1980.

Example 20.3.3

Determine size and performance passive solar characteristics with the location-dependent set of rules of thumb for the house of the previous example.

Solution. Using Table 20.6 with the 195 m^2 house in Denver yields:

$$\text{Solar projected area} = 12\% \text{ to } 23\% \text{ of floor area}$$

$$= 23.4\ \text{m}^2 \text{ to } 44.9\ \text{m}^2.$$

$$\text{SSF (no night insulation)} = 27\% \text{ to } 43\%.$$

$$\text{SSF (R9 night insulation)} = 47\% \text{ to } 74\%.$$

Using the rule of thumb for the thermal storage mass:

$$m = 17\ \text{kg} \times 43\% \times 44.9\ \text{m}^2$$

$$= 33,000\ \text{kg}\ (72,000\ \text{lbm})[\text{Thermal wall or direct gain}]$$

Comparing the results of this example to those of the previous example, the two rules of thumb are seen to produce "roughly" similar answers. General system cost and performance information can be generated with results from rule-of-thumb calculations, but a more detailed level of information is needed to determine design-ready passive system type (direct gain, thermal wall, sunspace), size, performance, and costs.

20.3.3.9 The Second Level: LCR Method

The LCR method is useful for making estimates of the annual performance of specific types of passive system(s) combinations. The LCR method was developed by calculating the annual SSF for 94 reference passive solar systems for 219 U.S. and Canadian locations over a range of LCR values. Table 20.7 includes the description of these 94 reference systems for use both with the LCR method and with the SLR method described below. Tables were constructed for each city with LCR versus SSF listed for each of the 94 reference passive systems. (Note that the solar load ratio (SLR) method was used to make the LCR calculations, and this SLR method is described in the next section as the third-level method.) Although the complete LCR tables (PSDH 1984) include 219 locations, Table 20.8 only includes six "representative" cities (Albuquerque, Boston, Madison, Medford, Nashville, Santa Maria), purely due to space restrictions. The LCR method consists of the following steps (PSDH 1984):

1. Determine the building parameters:
   a. Building load coefficient, NLC
   b. Solar projected area, A_p
   c. Load collector ratio, $\text{LCR} = \text{NLC}/A_p$
2. Find the short designation of the reference system closest to the passive system design (Table 20.7)
3. Enter the LCR Tables (Table 20.8)
   a. Find the city
   b. Find the reference system listing
   c. Determine annual SSF by interpolation using the LCR value from above
   d. Note the annual heating degree days (Number of degree days)
4. Calculate the annual auxiliary heat required:

$$\text{Auxiliary heat required} = (1 - \text{SSF}) \times \text{NLC} \times (\text{Number of degree days}).$$

If more than one reference solar system is being used, then find the "aperture area weighted" SSF for the combination. Determine each individual reference system SSF using the total aperture area LCR, then take the "area weighted" average of the individual SSFs.

TABLE 20.7 Designations and Characteristics for 94 Reference Systems

(a) Overall System Characteristics	
Masonry Properties	
Thermal conductivity (k)	
Sunspace floor	0.5 Btu/h/ft./°F
All other masonry	1.0 Btu/h/ft./°F
Density (Q)	150 lb/ft.^{3}
Specific heat (c)	0.2 Btu/lb/°F
Infrared emittance of normal surface	0.9
Infrared emittance of selective surface	0.1
Solar Absorptances	
Waterwall	1.0
Masonry, Trombe wall	1.0
Direct gain and sunspace	0.8
Sunspace: water containers	0.9
Lightweight common wall	0.7
Other lightweight surfaces	0.3
Glazing Properties	
Transmission characteristics	Diffuse
Orientation	Due south
Index of refraction	1.526
Extinction coefficient	0.5 in.$^{-1}$
Thickness of each pane	⅛ in.
Gap between panes	½ in.
Ared emittance	0.9
Control Range	
Room temperature	65°F–75°F
Sunspace temperature	45°F–95°F
Internal heat generation	0
Thermocirculation Vents (when used)	
Vent area/projected area (sum of both upper and lower vents)	0.06
Height between vents	8 ft.
Reverse flow	None
Nighttime Insulation (when used)	
Thermal resistance	R9
In place, solar time	5:30 P.M. to 7:30 A.M.
Solar Radiation Assumptions	
Shading	None
Ground diffuse reflectance	0.3

(b) Direct-Gain (DG) System Types

Designation	Thermal Storage Capacity[a] (Btu/ft.^{2}/°F)	Mass Thickness[a] (in.)	Mass-Area-to-Glazing-Area Ratio	No. of Glazings	Nighttime Insulation
A1	30	2	6	2	No
A2	30	2	6	3	No
A3	30	2	6	2	Yes
B1	45	6	3	2	No
B2	45	6	3	3	No
B3	45	6	3	2	Yes
C1	60	4	6	2	No
C2	60	4	6	3	No
C3	60	4	6	2	Yes

(continued)

TABLE 20.7 (*Continued*)

(c) Vented Trombe Wall (TW) System Types						
Designation	Thermal Storage Capacity[a] (Btu/ft.^{2}/°F)	Wall Thickness[a] (in.)	ρck (Btu^2/h/ft.^{4}/°F^2)	No. of Glazings	Wall Surface	Nighttime Insulation
A1	15	6	30	2	Normal	No
A2	22.5	9	30	2	Normal	No
A3	30	12	30	2	Normal	No
A4	45	18	30	2	Normal	No
B1	15	6	15	2	Normal	No
B2	22.5	9	15	2	Normal	No
B3	30	12	15	2	Normal	No
B4	45	18	15	2	Normal	No
C1	15	6	7.5	2	Normal	No
C2	22.5	9	7.5	2	Normal	No
C3	30	12	7.5	2	Normal	No
C4	45	18	7.5	2	Normal	No
D1	30	12	30	1	Normal	No
D2	30	12	30	3	Normal	No
D3	30	12	30	1	Normal	Yes
D4	30	12	30	2	Normal	Yes
D5	30	12	30	3	Normal	Yes
E1	30	12	30	1	Selective	No
E2	30	12	30	2	Selective	No
E3	30	12	30	1	Selective	Yes
E4	30	12	30	2	Selective	Yes

(d) Unvented Trombe Wall (TW) System Types						
Designation	Thermal Storage Capacity[a] (Btu/ft.^{2}/°F)	Wall Thickness[a] (in.)	ρck (Btu^2/h/ft.^{4}/°F^2)	No. of Glazings	Wall Surface	Nighttime Insulation
F1	15	6	30	2	Normal	No
F2	22.5	9	30	2	Normal	No
F3	30	12	30	2	Normal	No
F4	45	18	30	2	Normal	No
G1	15	6	15	2	Normal	No
G2	22.5	9	15	2	Normal	No
G3	30	12	15	2	Normal	No
G4	45	18	15	2	Normal	No
H1	15	6	7.5	2	Normal	No
H2	22.5	9	7.5	2	Normal	No
H3	30	12	7.5	2	Normal	No
H4	45	18	7.5	2	Normal	No
I1	30	12	30	1	Normal	No
I2	30	12	30	3	Normal	No
I3	30	12	30	1	Normal	Yes
I4	30	12	30	2	Normal	Yes
I5	30	12	30	3	Normal	Yes
J1	30	12	30	1	Selective	No
J2	30	12	30	2	Selective	No
J3	30	12	30	1	Selective	Yes
J4	30	12	30	2	Selective	Yes

(continued)

TABLE 20.7 (*Continued*)

(e) Waterwall (WW) System Types					
Designation	Thermal Storage Capacity[a] (Btu/ft.^{2}/°F)	Wall Thickness (in.)	No. of Glazings	Wall Surface	Nighttime Insulation
A1	15.6	3	2	Normal	No
A2	31.2	6	2	Normal	No
A3	46.8	9	2	Normal	No
A4	62.4	12	2	Normal	No
A5	93.6	18	2	Normal	No
A6	124.8	24	2	Normal	No
B1	46.8	9	1	Normal	No
B2	46.8	9	3	Normal	No
B3	46.8	9	1	Normal	Yes
B4	46.8	9	2	Normal	Yes
B5	46.8	9	3	Normal	Yes
C1	46.8	9	1	Selective	No
C2	46.8	9	2	Selective	No
C3	46.8	9	1	Selective	Yes
C4	46.8	9	2	Selective	Yes

(f) Sunspace (SS) System Types					
Designation	Type	Tilt (°)	Common Wall	End Walls	Nighttime Insulation
A1	Attached	50	Masonry	Opaque	No
A2	Attached	50	Masonry	Opaque	Yes
A3	Attached	50	Masonry	Glazed	No
A4	Attached	50	Masonry	Glazed	Yes
A5	Attached	50	Insulated	Opaque	No
A6	Attached	50	Insulated	Opaque	Yes
A7	Attached	50	Insulated	Glazed	No
A8	Attached	50	Insulated	Glazed	Yes
B1	Attached	90/30	Masonry	Opaque	No
B2	Attached	90/30	Masonry	Opaque	Yes
B3	Attached	90/30	Masonry	Glazed	No
B4	Attached	90/30	Masonry	Glazed	Yes
B5	Attached	90/30	Insulated	Opaque	No
B6	Attached	90/30	Insulated	Opaque	Yes
B7	Attached	90/30	Insulated	Glazed	No
B8	Attached	90/30	Insulated	Glazed	Yes
C1	Semienclosed	90	Masonry	Common	No
C2	Semienclosed	90	Masonry	Common	Yes
C3	Semienclosed	90	Insulated	Common	No
C4	Semienclosed	90	Insulated	Common	Yes
D1	Semienclosed	50	Masonry	Common	No
D2	Semienclosed	50	Masonry	Common	Yes
D3	Semienclosed	50	Insulated	Common	No
D4	Semienclosed	50	Insulated	Common	Yes
E1	Semienclosed	90/30	Masonry	Common	No
E2	Semienclosed	90/30	Masonry	Common	Yes
E3	Semienclosed	90/30	Insulated	Common	No
E4	Semienclosed	90/30	Insulated	Common	Yes

[a] The thermal storage capacity is per unit of projected area, or, equivalently, the quantity ρck. The wall thickness is listed only as an appropriate guide by assuming $\rho c = 30$ Btu/ft.^{3}/°F.

Source: From PSDH, Passive Solar Design Handbook. Part One: Total Environmental Action, Inc., Part Two: Los Alamos Scientific Laboratory, Part Three: Los Alamos National Laboratory. Van Nostranal Laboratory. Van Nostranal Reinhold, New York, 1984.

TABLE 20.8 LCR Tables for Six Representative Cities (Albuquerque, Boston, Madison, Medford, Nashville, and Santa Maria)

SSF	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
Santa Maria, California									3053 DD
WW A1	1776	240	119	73	50	35	25	18	12
WW A2	617	259	154	103	74	54	39	28	19
WW A3	523	261	164	114	82	61	45	33	22
WW A4	482	260	169	119	87	65	48	35	24
WW A5	461	263	175	125	92	69	52	38	26
WW A6	447	263	177	128	95	72	54	40	27
WW B1	556	220	128	85	60	43	32	23	15
WW B2	462	256	168	119	88	66	49	36	25
WW B3	542	315	211	151	112	85	64	47	32
WW B4	455	283	197	144	109	83	63	47	32
WW B5	414	263	184	136	103	79	60	45	31
WW C1	569	330	221	159	118	89	67	49	33
WW C2	478	288	197	143	107	81	61	45	31
WW C3	483	318	228	170	130	100	77	57	40
WW C4	426	280	200	149	114	88	68	51	35
TW A1	1515	227	113	70	48	34	24	17	11
TW A2	625	234	134	89	63	46	33	24	16
TW A3	508	231	140	95	68	50	37	27	18
TW A4	431	217	137	95	69	51	38	28	19
TW B1	859	212	112	71	49	35	25	18	12
TW B2	502	209	124	83	59	43	32	23	15
TW B3	438	201	123	84	60	44	33	24	16
TW B4	400	184	112	76	55	40	30	22	14
TW C1	568	188	105	69	48	35	25	18	12
TW C2	435	178	105	70	50	36	27	19	13
TW C3	413	165	97	64	46	33	25	18	12
TW C4	426	146	82	54	38	27	20	14	10
TW D1	403	170	101	67	48	35	25	18	12
TW D2	488	242	152	105	76	57	42	31	21
TW D3	509	271	175	123	90	67	50	36	25
TW D4	464	266	177	127	94	71	53	39	27
TW D5	425	250	169	122	91	69	52	38	26
TW E1	581	309	199	140	102	76	57	42	28
TW E2	512	283	186	132	97	73	55	40	27
TW E3	537	328	225	164	123	94	71	53	36
TW E4	466	287	199	145	109	83	63	47	32
TW F1	713	198	107	68	47	34	25	18	12
TW F2	455	199	120	81	58	42	31	22	15
TW F3	378	190	120	83	60	45	33	24	16
TW F4	311	169	110	77	57	42	32	23	16
TW G1	450	170	98	65	46	33	24	17	12
TW G2	331	163	102	70	51	38	28	20	14
TW G3	278	147	94	66	48	36	27	20	13
TW G4	222	120	78	55	40	30	22	16	11
TW H1	295	137	84	57	41	30	22	16	11
TW H2	226	118	75	52	38	28	21	15	10
TW H3	187	99	64	44	33	24	18	13	9
TW H4	143	75	48	33	24	18	14	10	7
TW I1	318	144	88	59	42	31	23	16	11
TW I2	377	203	132	93	68	51	38	28	19
TW I3	404	226	149	106	78	58	44	32	22
TW I4	387	230	156	113	84	64	48	36	24
TW I5	370	226	155	113	85	65	49	36	25

(continued)

TABLE 20.8 (*Continued*)

SSF	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
TW J1	483	271	179	127	94	71	53	39	26
TW J2	422	246	165	119	88	67	50	37	25
TW J3	446	283	199	146	111	85	65	48	33
TW J4	400	254	178	132	100	77	58	43	30
DG A1	392	188	117	79	55	38	26	16	7
DG A2	389	190	121	85	61	45	32	22	14
DG A3	443	220	142	102	77	58	44	31	19
DG B1	384	191	122	86	64	48	35	24	13
DG B2	394	196	127	91	69	53	40	29	19
DG B3	445	222	145	105	80	62	49	37	25
DG C1	451	225	146	104	78	61	47	34	21
DG C2	453	226	148	106	80	63	49	37	25
DG C3	509	254	167	121	92	73	58	45	31
SS A1	1171	396	220	142	98	69	49	34	22
SS A2	1028	468	283	190	135	98	71	50	33
SS A3	1174	380	209	133	91	64	45	31	20
SS A4	1077	481	289	193	136	98	71	50	32
SS A5	1896	400	204	127	86	60	42	29	18
SS A6	1030	468	283	190	135	97	71	50	32
SS A7	2199	359	178	109	72	50	35	24	15
SS A8	1089	478	285	190	133	96	69	48	31
SS B1	802	298	170	111	77	55	40	28	18
SS B2	785	366	224	152	108	79	57	41	27
SS B3	770	287	163	106	74	52	37	26	17
SS B4	790	368	224	152	108	78	57	40	26
SS B5	1022	271	144	91	62	44	31	22	14
SS B6	750	356	219	149	106	77	56	40	26
SS B7	937	242	127	80	54	38	27	19	12
SS B8	750	352	215	146	103	75	55	39	25
SS C1	481	232	144	99	71	52	39	28	19
SS C2	482	262	170	120	88	66	49	36	24
SS C3	487	185	107	71	50	36	27	19	13
SS C4	473	235	147	102	74	55	41	30	20
SS D1	1107	477	282	188	132	95	68	48	31
SS D2	928	511	332	232	169	125	92	66	43
SS D3	1353	449	248	160	110	78	56	39	25
SS D4	946	500	319	222	160	117	86	61	40
SS E1	838	378	227	153	108	78	56	40	26
SS E2	766	419	272	190	138	102	75	54	36
SS E3	973	322	178	115	79	56	40	28	18
SS E4	780	393	247	170	122	89	65	47	31
Albuquerque, New Mexico									4292 DD
WW A1	1052	130	62	38	25	18	13	9	6
WW A2	354	144	84	56	39	29	21	15	10
WW A3	300	146	90	62	45	33	24	18	12
WW A4	276	146	93	65	47	35	26	19	13
WW A5	264	148	97	69	50	38	28	21	14
WW A6	256	148	99	70	52	39	30	22	15
WW B1	293	111	63	41	28	20	15	11	7
WW B2	270	147	96	67	49	37	28	20	14
WW B3	314	179	119	84	62	47	35	26	18
WW B4	275	169	116	85	64	49	37	28	19
WW B5	252	159	110	81	61	47	36	27	19
WW C1	333	190	126	89	66	50	38	28	19
WW C2	287	171	115	83	62	47	36	27	18
WW C3	293	191	136	101	77	59	46	34	24

(continued)

TABLE 20.8 (*Continued*)

SSF	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
WW C4	264	172	122	91	69	54	41	31	22
TW A1	900	124	60	37	25	17	12	9	6
TW A2	361	130	73	48	33	24	18	13	8
TW A3	293	129	77	52	37	27	20	15	10
TW A4	249	123	76	52	38	28	21	15	10
TW B1	502	117	60	38	26	18	13	9	6
TW B2	291	118	68	45	32	23	17	12	8
TW B3	254	114	68	46	33	24	18	13	9
TW B4	233	104	63	42	30	22	16	12	8
TW C1	332	106	58	37	26	19	14	10	6
TW C2	255	101	58	39	27	20	15	11	7
TW C3	243	94	54	36	25	18	13	10	7
TW C4	254	84	46	30	21	15	11	8	5
TW D1	213	86	50	33	23	17	12	9	6
TW D2	287	139	86	59	43	32	24	17	12
TW D3	294	153	97	68	49	37	27	20	14
TW D4	281	158	104	74	55	41	31	23	16
TW D5	260	151	101	73	54	41	31	23	16
TW E1	339	177	113	78	57	43	32	23	16
TW E2	308	168	109	77	56	42	32	23	16
TW E3	323	195	133	96	72	55	42	31	21
TW E4	287	175	120	88	66	50	38	28	20
TW F1	409	108	57	36	24	17	13	9	6
TW F2	260	110	65	43	31	22	17	12	8
TW F3	216	106	66	45	33	24	10	13	9
TW F4	178	95	61	42	31	23	17	13	9
TW G1	256	93	53	34	24	17	13	9	6
TW G2	189	91	56	38	27	20	15	11	7
TW G3	159	82	52	36	26	20	15	11	7
TW G4	128	68	43	30	22	16	12	9	6
TW H1	168	76	45	31	22	16	12	9	6
TW H2	130	66	41	29	21	15	11	8	6
TW H3	108	56	35	25	8	13	10	7	5
TW H4	83	42	27	19	13	10	7	5	4
TW I1	166	73	43	29	20	15	11	8	5
TW I2	221	117	75	52	30	28	21	16	11
TW I3	234	128	83	59	43	32	24	10	12
TW I4	234	137	92	66	49	37	28	21	14
TW I5	226	136	93	67	50	38	29	22	15
TW J1	282	156	102	72	53	40	30	22	15
TW J2	254	146	97	69	51	39	29	22	15
TW J3	269	169	118	86	65	50	38	29	20
TW J4	247	155	106	80	60	46	35	26	18
DG A1	211	97	57	36	22	13	5	—	—
DG A2	227	107	67	46	32	23	16	10	5
DG A3	274	131	83	59	44	34	25	18	10
DG B1	210	97	60	42	30	21	13	6	—
DG B2	232	110	69	49	37	28	21	14	8
DG B3	277	134	85	61	47	37	28	21	14
DG C1	253	120	74	53	39	30	22	14	—
DG C2	271	130	82	59	45	35	26	19	12
DG C3	318	155	96	71	54	43	34	26	18
SS A1	591	187	101	64	44	31	22	16	10
SS A2	531	232	137	92	65	47	34	25	16
SS A3	566	170	90	56	38	27	19	13	8

(continued)

TABLE 20.8 (*Continued*)

SSF	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
SS A4	537	230	135	89	63	45	33	23	15
SS A5	980	187	92	56	37	26	18	13	8
SS A6	529	231	136	91	64	47	34	24	16
SS A7	1103	158	74	44	29	20	14	10	6
SS A8	540	226	131	87	61	44	32	23	15
SS B1	403	141	78	50	35	25	18	13	8
SS B2	412	186	111	75	53	39	28	20	14
SS B3	372	130	71	46	31	22	16	11	7
SS B4	403	181	106	72	51	37	27	20	13
SS B5	518	127	65	40	27	19	13	9	6
SS B6	390	179	106	73	52	38	28	20	13
SS B7	457	108	54	33	22	16	11	8	5
SS B8	379	171	102	69	49	35	26	19	12
SS C1	270	126	77	52	37	27	20	15	10
SS C2	282	150	97	68	49	37	28	20	14
SS C3	276	101	57	37	26	19	14	10	7
SS C4	277	135	83	57	41	31	23	17	11
SS D1	548	225	130	85	59	43	31	22	14
SS D2	474	253	162	113	82	61	45	33	22
SS D3	683	212	113	72	49	35	25	17	11
SS D4	484	248	156	107	77	57	42	30	20
SS E1	410	176	103	68	48	35	25	18	12
SS E2	390	208	133	92	67	50	37	27	18
SS E3	487	151	80	51	35	25	18	12	8
SS E4	400	195	120	82	59	43	32	23	15
Nashville, Tennessee									3696 DD
WW A1	588	60	24	13	8	5	3	2	1
WW A2	192	70	38	23	15	11	7	5	3
WW A3	161	72	42	27	18	13	9	6	4
WW A4	148	72	43	29	20	14	10	7	5
WW A5	141	74	46	31	22	16	11	8	5
WW A6	137	74	47	32	22	16	12	8	5
WW B1	135	41	19	10	6	3	2	—	—
WW B2	152	78	48	33	23	17	12	9	6
WW B3	179	97	61	42	30	22	16	12	8
WW B4	164	97	65	46	34	25	19	14	9
WW B5	153	93	63	45	33	25	19	14	9
WW C1	193	105	67	46	33	24	18	13	8
WW C2	169	97	63	44	32	24	18	13	8
WW C3	181	115	79	58	43	33	25	18	12
WW C4	164	104	72	53	39	30	23	17	11
TW A1	509	59	25	13	8	5	3	2	1
TW A2	199	64	33	20	13	9	6	4	3
TW A3	160	65	36	23	15	11	8	5	3
TW A4	136	62	36	23	16	11	8	6	4
TW B1	282	57	26	15	9	6	4	3	2
TW B2	161	59	32	20	13	9	6	4	3
TW B3	141	58	32	21	14	10	7	5	3
TW B4	131	54	30	19	13	9	7	5	3
TW C1	188	53	27	16	10	7	5	3	2
TW C2	144	52	28	18	12	8	6	4	2
TW C3	139	49	27	17	11	8	5	4	2
TW C4	149	45	23	14	9	7	5	3	2
TW D1	99	33	16	9	5	3	2	1	—
TW D2	164	75	44	29	20	14	10	7	5

(continued)

TABLE 20.8 (*Continued*)

SSF	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
TW D3	167	82	49	33	23	17	12	8	5
TW D4	168	91	58	40	29	21	15	11	7
TW D5	160	89	58	40	29	22	16	12	8
TW E1	198	98	59	40	28	20	15	10	7
TW E2	182	95	59	40	29	21	15	11	7
TW E3	197	115	76	54	39	29	22	16	11
TW E4	178	105	70	50	37	27	20	15	10
TW F1	221	50	23	13	8	5	4	2	1
TW F2	139	53	29	18	12	8	6	4	2
TW F3	116	52	30	19	13	9	7	5	3
TW F4	96	47	28	19	13	9	7	5	3
TW G1	137	44	22	13	9	6	4	3	2
TW G2	101	44	25	16	11	8	5	4	2
TW G3	86	41	24	16	11	8	6	4	2
TW G4	69	34	21	14	10	7	5	3	2
TW H1	89	36	20	13	8	6	4	3	2
TW H2	69	33	19	12	9	6	4	3	2
TW H3	59	28	17	11	8	5	4	3	2
TW H4	46	22	13	9	6	4	3	2	1
TW I1	74	26	13	7	4	2	1	—	—
TW I2	125	62	38	25	18	13	9	7	4
TW I3	133	69	43	29	20	15	11	8	5
TW I4	139	78	51	35	26	19	14	10	7
TW I5	137	80	53	37	27	20	15	11	7
TW J1	164	86	54	36	26	19	14	10	6
TW J2	150	82	53	36	26	19	14	10	7
TW J3	165	101	68	49	36	27	20	15	10
TW J4	153	93	63	46	34	25	19	14	10
DG A1	98	34	—	—	—	—	—	—	—
DG A2	130	55	31	19	11	6	—	—	—
DG A3	173	78	47	32	23	16	11	7	2
DG B1	100	36	17	—	—	—	—	—	—
DG B2	134	58	33	22	15	10	6	—	—
DG B3	177	81	49	33	24	18	14	10	6
DG C1	131	52	28	17	9	—	—	—	—
DG C2	161	71	42	28	20	14	10	6	—
DG C3	205	94	57	39	29	22	17	12	8
SS A1	351	100	50	29	19	13	9	6	4
SS A2	328	135	76	49	33	24	17	12	8
SS A3	330	87	41	24	15	10	6	4	2
SS A4	331	133	74	47	32	22	16	11	7
SS A5	595	98	43	24	15	10	7	4	2
SS A6	324	132	75	48	32	23	16	11	7
SS A7	668	79	32	17	10	6	4	2	1
SS A8	330	129	71	45	30	21	15	10	6
SS B1	236	74	38	23	15	10	7	5	3
SS B2	258	110	63	41	28	20	14	10	6
SS B3	212	65	32	19	12	8	5	3	2
SS B4	251	105	60	39	27	19	13	9	6
SS B5	307	65	30	17	10	7	4	3	2
SS B6	241	104	60	39	27	19	14	10	6
SS B7	264	52	23	12	7	5	3	2	—
SS B8	233	98	56	36	25	17	12	9	5
SS C1	141	60	33	21	14	10	7	5	3
SS C2	161	81	50	33	23	17	12	9	6
SS C3	149	48	25	15	10	7	4	3	2

(continued)

TABLE 20.8 (*Continued*)

SSF	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
SS C4	160	73	43	28	19	14	10	7	5
SS D1	317	119	64	39	26	18	13	8	5
SS D2	287	147	90	61	43	31	23	16	10
SS D3	405	113	55	33	21	14	10	6	4
SS D4	295	144	87	58	40	29	21	15	10
SS E1	229	89	48	29	19	13	9	6	4
SS E2	233	118	72	48	34	24	18	12	8
SS E3	283	77	37	22	14	9	6	4	2
SS E4	242	111	65	43	29	21	15	11	7
Medford, Oregon									4930 DD
WW A1	708	64	24	11	—	—	—	—	—
WW A2	212	73	38	22	13	7	3	—	—
WW A3	174	75	41	25	16	9	5	2	—
WW A4	158	74	43	27	17	11	6	3	1
WW A5	149	75	45	29	19	12	7	4	2
WW A6	144	75	46	30	20	13	8	4	2
WW B1	154	43	16	—	—	—	—	—	—
WW B2	162	80	48	31	21	14	9	6	3
WW B3	190	100	62	41	28	19	13	8	5
WW B4	171	99	65	45	32	23	16	11	7
WW B5	160	95	63	45	32	23	17	12	7
WW C1	205	108	67	45	31	21	15	10	6
WW C2	178	99	63	43	30	22	15	10	6
WW C3	189	117	80	57	42	31	23	16	10
WW C4	170	106	72	52	38	28	21	15	9
TWAI	607	63	25	12	5	—	—	—	—
TW A2	222	68	33	19	11	6	2	—	—
TW A3	175	67	36	21	13	8	4	2	—
TW A4	147	64	36	22	14	9	5	3	1
TW B1	327	61	27	14	7	3	—	—	—
TW B2	178	62	32	19	12	7	4	2	—
TW B3	154	60	33	20	12	8	4	2	1
TW B4	143	56	31	19	12	8	5	2	1
TW C1	212	56	27	15	9	5	2	—	—
TW C2	159	55	28	17	11	7	4	2	—
TW C3	154	52	27	16	10	6	4	2	1
TW C4	167	48	24	14	9	5	3	2	—
TW D1	112	34	14	—	—	—	—	—	—
TW D2	177	77	44	28	18	12	8	5	3
TW D3	180	85	50	32	21	14	9	6	3
TW D4	177	93	58	39	27	19	13	9	5
TW D5	168	92	58	40	28	20	14	10	6
TW E1	213	101	60	39	26	18	12	8	4
TW E2	194	98	59	39	27	19	13	9	5
TW E3	208	118	77	53	38	27	20	13	8
TW E4	186	108	71	49	36	26	19	13	8
TW F1	256	53	23	12	5	—	—	—	—
TW F2	153	56	29	17	10	5	2	—	—
TW F3	125	54	30	18	11	7	3	1	—
TW F4	102	48	28	18	11	7	4	2	1
TW G1	153	46	22	12	7	—	—	—	—
TW G2	109	46	25	15	9	5	3	1	—
TW G3	92	42	24	15	9	6	3	2	—
TW G4	74	35	20	13	8	5	3	2	—
TW H1	97	38	20	12	7	4	1	—	—

(continued)

TABLE 20.8 (***Continued***)

SSF	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
TW H2	75	34	19	12	7	5	3	1	—
TW H3	63	29	17	10	7	4	3	1	—
TW H4	49	23	13	8	5	3	2	1	—
TW I1	83	27	10	—	—	—	—	—	—
TW I2	133	64	38	24	16	11	7	4	2
TW I3	142	71	43	28	19	13	9	5	3
TW I4	146	80	51	35	25	17	12	8	5
TW I5	144	82	53	37	26	19	13	9	6
TW J1	175	89	54	36	24	17	11	7	4
TW J2	158	85	53	36	25	18	12	8	5
TW J3	173	103	69	48	35	26	18	13	8
TW J4	160	96	64	45	33	24	17	12	8
DG A1	110	35	—	—	—	—	—	—	—
DG A2	142	58	32	18	9	—	—	—	—
DG A3	187	82	48	32	22	15	9	5	—
DG B1	110	40	15	—	—	—	—	—	—
DG B2	146	61	35	21	13	7	—	—	—
DG B3	193	84	51	34	24	17	12	7	3
DG C1	144	57	29	13	—	—	—	—	—
DG C2	177	75	44	28	19	12	6	—	—
DG C3	224	98	60	41	29	21	14	10	5
SS A1	415	110	51	28	16	9	4	2	—
SS A2	372	146	79	48	31	21	14	8	5
SS A3	397	96	42	21	10	—	—	—	—
SS A4	379	144	76	46	29	19	12	7	4
SS A5	732	111	45	23	12	5	—	—	—
SS A6	368	143	77	47	30	20	13	8	4
SS A7	846	90	33	14	—	—	—	—	—
SS A8	379	140	73	44	27	17	11	6	3
SS B1	274	81	38	21	12	6	3	—	—
SS B2	288	117	65	40	26	18	12	7	4
SS B3	249	71	33	17	8	—	—	—	—
SS B4	282	113	62	38	25	16	11	7	4
SS B5	368	72	30	15	7	—	—	—	—
SS B6	269	111	62	30	25	17	11	7	4
SS B7	323	58	23	10	—	—	—	—	—
SS B8	262	106	57	35	23	15	9	6	3
SS C1	153	62	33	19	11	5	—	—	—
SS C2	172	83	50	32	22	15	10	6	3
SS C3	166	51	24	13	7	3	—	—	—
SS C4	173	76	43	27	18	12	8	5	3
SS D1	367	129	65	37.	22	13	7	3	1
SS D2	318	156	92	60	40	27	18	12	7
SS D3	480	124	57	31	18	10	5	2	—
SS D4	328	153	89	57	38	26	17	11	6
SS E1	262	95	48	27	15	7	—	—	—
SS E2	257	124	73	47	31	21	14	9	5
SS E3	334	84	38	20	10	4	—	—	—
SS E4	269	118	67	42	27	18	12	7	4
Boston, Massachusetts									5621 DD
WW A1	368	28	9	—	—	—	—	—	—
WW A2	119	41	20	12	7	5	3	2	—
WW A3	101	43	24	15	10	6	4	3	1
WW A4	93	44	26	16	11	7	5	3	2
WW A5	89	45	27	18	12	8	6	4	2

(continued)

TABLE 20.8 (*Continued*)

SSF	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
WW A6	87	46	28	19	13	9	6	4	3
WW B1	59	—	—	—	—	—	—	—	—
WW B2	103	52	31	21	15	10	7	5	3
WW B3	123	66	41	28	20	14	10	7	5
WW B4	118	70	46	33	24	18	13	9	6
WW B5	113	69	46	33	25	18	14	10	7
WW C1	135	72	46	31	22	16	12	8	5
WW C2	121	68	44	31	22	16	12	9	6
WW C3	136	86	60	44	33	25	19	14	9
WW C4	124	78	54	40	30	23	17	12	8
TW A1	324	30	11	4	—	—	—	—	—
TW A2	126	37	18	10	6	4	2	1	—
TW A3	102	39	21	13	8	5	3	2	1
TW A4	88	38	22	14	9	6	4	3	2
TW B1	180	32	13	7	4	2	—	—	—
TW B2	104	36	19	11	7	5	3	2	1
TW B3	92	36	19	12	8	5	3	2	1
TW B4	86	34	19	12	8	5	4	2	1
TW C1	122	32	15	9	5	3	2	1	—
TW C2	95	33	17	10	7	4	3	2	1
TW C3	93	31	16	10	6	4	3	2	1
TW C4	102	29	15	9	6	4	3	2	1
TW D1	45	—	—	—	—	—	—	—	—
TW D2	112	49	28	18	12	9	6	4	3
TW D3	113	54	32	21	15	10	7	5	3
TW D4	121	64	41	28	20	15	11	8	5
TW D5	118	66	42	30	21	16	12	8	6
TW E1	138	67	40	27	18	13	9	7	4
TW E2	130	66	41	28	20	14	10	7	5
TW E3	146	84	56	39	29	21	16	11	8
TW E4	133	78	52	37	27	20	15	11	7
TW F1	134	25	10	4	—	—	—	—	—
TW F2	86	30	16	9	5	3	2	1	—
TW F3	72	31	17	11	7	4	3	2	1
TW F4	61	29	17	11	7	5	3	2	1
TW G1	83	24	11	6	3	2	—	—	—
TW G2	63	26	14	9	5	4	2	1	—
TW G3	54	25	14	9	6	4	3	2	1
TW G4	45	21	12	8	5	4	3	2	1
TW H1	54	21	11	6	4	2	1	—	—
TW H2	44	20	11	7	5	3	2	1	—
TW H3	38	17	10	6	4	3	2	1	—
TW H4	30	14	8	5	3	2	2	1	—
TW I1	30	—	—	—	—	—	—	—	—
TW I2	84	41	24	16	11	8	6	4	2
TW I3	91	46	28	19	13	9	7	5	3
TW I4	100	56	36	25	18	13	10	7	5
TW I5	101	58	38	27	20	15	11	8	5
TW J1	114	59	37	25	17	12	9	6	4
TW J2	107	58	37	25	18	13	10	7	4
TW J3	123	75	51	36	27	20	15	11	7
TW J4	115	70	47	34	25	19	14	10	7
DG A1	43	—	—	—	—	—	—	—	—
DG A2	85	34	18	9	—	—	—	—	—
DG A3	125	56	33	22	16	11	7	4	—

(continued)

TABLE 20.8 (*Continued*)

SSF	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
DG B1	44	—	—	—	—	—	—	—	—
DG B2	87	36	20	12	7	—	—	—	—
DG B3	129	58	35	24	17	13	9	6	3
DG C1	71	23	—	—	—	—	—	—	—
DG C2	109	47	27	17	12	8	4	—	—
DG C3	151	68	41	28	21	16	12	8	5
SS A1	230	61	29	16	10	6	4	2	1
SS A2	231	93	52	33	22	15	11	7	5
SS A3	205	48	20	10	4	—	—	—	—
SS A4	229	90	49	31	20	14	9	6	4
SS A5	389	58	23	11	6	3	—	—	—
SS A6	226	91	50	32	21	15	10	7	4
SS A7	420	40	12	—	—	—	—	—	—
SS A8	226	86	46	28	19	12	8	6	3
SS B1	151	44	21	12	7	4	2	1	—
SS B2	183	77	43	28	19	13	9	6	4
SS B3	129	36	16	8	3	—	—	—	—
SS B4	176	73	41	26	17	12	8	6	4
SS B5	193	36	15	7	3	—	—	—	—
SS B6	169	72	41	26	18	12	9	6	4
SS B7	157	25	7	—	—	—	—	—	—
SS B8	160	66	37	23	16	11	7	5	3
SS C1	84	33	17	10	6	4	2	1	—
SS C2	110	54	33	22	15	11	8	5	3
SS C3	91	26	12	7	4	2	—	—	—
SS C4	109	48	28	18	12	9	6	4	3
SS D1	206	73	38	22	14	9	5	3	2
SS D2	203	103	63	42	29	21	15	10	6
SS D3	264	69	32	18	10	6	4	2	1
SS D4	208	100	60	39	27	19	14	9	6
SS E1	140	51	25	14	8	4	2	—	—
SS E2	161	80	48	32	22	15	11	7	5
SS E3	177	44	19	10	5	2	—	—	—
SS E4	166	75	43	28	19	13	9	6	4
Madison, Wisconsin									7730 DD
WW A1	278	—	—	—	—	—	—	—	—
WW A2	91	27	12	—	—	—	—	—	—
WW A3	77	30	15	8	3	—	—	—	—
WW A4	72	32	17	10	5	—	—	—	—
WW A5	69	33	19	11	7	4	—	—	—
WW A6	67	34	19	12	7	4	2	—	—
WW B1	—	—	—	—	—	—	—	—	—
WW B2	84	41	24	15	10	7	5	3	2
WW B3	102	53	32	21	15	10	7	5	3
WW B4	101	59	39	27	19	14	10	7	5
WW B5	98	59	39	28	20	15	11	8	5
WW C1	113	59	37	25	17	12	8	6	3
WW C2	103	57	37	25	18	13	9	6	4
WW C3	119	75	51	37	28	21	15	11	7
WW C4	109	68	47	34	25	19	14	10	7
TW A1	249	16	—	—	—	—	—	—	—
TW A2	97	26	11	4	—	—	—	—	—
TW A3	79	28	13	7	3	—	—	—	—
TW A4	69	28	15	9	5	3	—	—	—

(*continued*)

TABLE 20.8 (*Continued*)

SSF	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
TW B1	139	20	5	—	—	—	—	—	—
TW B2	81	26	12	6	3	—	—	—	—
TW B3	72	27	13	7	4	2	—	—	—
TW B4	69	26	13	8	5	3	1	—	—
TW C1	96	23	10	4	—	—	—	—	—
TW C2	76	25	12	7	4	2	—	—	—
TW C3	75	24	12	7	4	2	1	—	—
TW C4	84	23	11	6	4	2	1	—	—
TW D1	—	—	—	—	—	—	—	—	—
TW D2	91	39	22	13	9	6	4	2	1
TW D3	93	43	25	16	10	7	5	3	1
TW D4	103	54	34	23	16	12	8	6	4
TW D5	102	56	36	25	18	13	10	7	4
TW E1	115	54	32	21	14	10	7	4	3
TW E2	110	55	34	22	16	11	8	5	3
TW E3	126	72	47	33	24	18	13	9	6
TW E4	116	68	45	32	23	17	13	9	6
TW F1	99	13	—	—	—	—	—	—	—
TW F2	65	20	8	—	—	—	—	—	—
TW F3	55	22	11	5	—	—	—	—	—
TW F4	47	21	11	7	4	2	—	—	—
TW G1	61	14	—	—	—	—	—	—	—
TW G2	47	18	8	4	—	—	—	—	—
TW G3	42	18	9	5	3	—	—	—	—
TW G4	35	16	9	5	3	2	—	—	—
TW H1	41	13	6	—	—	—	—	—	—
TW H2	34	14	7	4	2	—	—	—	—
TW H3	29	13	7	4	2	1	—	—	—
TW H4	24	10	6	3	2	1	—	—	—
TW I1	—	—	—	—	—	—	—	—	—
TW I2	68	32	18	12	8	5	3	2	1
TW I3	75	37	22	14	10	7	4	3	2
TW I4	85	47	30	21	15	11	8	5	3
TW I5	87	50	33	23	16	12	9	6	4
TW J1	95	48	29	19	13	9	6	4	3
TW J2	91	48	30	21	14	10	7	5	3
TW J3	106	65	43	31	23	17	12	9	6
TW J4	100	61	41	29	21	16	12	9	6
DG A1	—	—	—	—	—	—	—	—	—
DG A2	68	25	11	—	—	—	—	—	—
DG A3	109	47	28	18	12	8	5	—	—
DG B1	—	—	—	—	—	—	—	—	—
DG B2	70	27	14	6	—	—	—	—	—
DG B3	114	50	30	20	14	10	7	4	—
DG C1	47	—	—	—	—	—	—	—	—
DG C2	91	37	21	13	7	—	—	—	—
DG C3	133	59	35	24	17	13	9	6	3
SS A1	192	47	20	9	3	—	—	—	—
SS A2	200	78	42	26	17	12	8	5	3
SS A3	166	32	—	—	—	—	—	—	—
SS A4	197	74	39	23	15	10	6	4	2
SS A5	329	42	13	—	—	—	—	—	—
SS A6	195	75	40	25	16	11	7	5	3
SS A7	349	22	—	—	—	—	—	—	—
SS A8	192	69	36	21	13	8	5	3	2

(continued)

TABLE 20.8 (***Continued***)

SSF	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
SS B1	122	32	13	5	—	—	—	—	—
SS B2	158	64	36	22	15	10	7	5	3
SS B3	100	22	—	—	—	—	—	—	—
SS B4	150	60	33	29	13	9	6	4	2
SS B5	156	24	—	—	—	—	—	—	—
SS B6	145	59	33	20	13	9	6	4	2
SS B7	122	—	—	—	—	—	1—	—	—
SS B8	136	54	29	18	11	7	5	3	2
SS C1	61	20	7	—	—	—	—	—	—
SS C2	90	43	25	16	11	7	5	3	2
SS C3	67	16	—	—	—	—	—	—	—
SS C4	90	38	22	13	9	6	4	2	1
SS D1	169	56	26	13	6	—	—	—	—
SS D2	175	86	51	34	23	16	11	7	5
SS D3	221	52	21	10	—	—	—	—	—
SS D4	179	84	49	32	21	15	10	7	4
SS E1	108	34	12	—	—	—	—	—	—
SS E2	135	65	38	24	16	11	7	5	3
SS E3	141	29	8	—	—	—	—	—	—
SS E4	140	61	34	21	14	9	6	4	2

Source: From PSDH, *Passive Solar Design Handbook*, Los Alamos National Laboratory, Van Nostrand Reinhold, New York, 1984.

The LCR method allows no variation from the 94 reference passive designs. To treat off-reference designs, sensitivity curves have been produced that illustrate the effect on SSF of varying one or two design variables. These curves were produced for the six "representative" cities, chosen for their wide geographical and climatological ranges. Several of these SSF "sensitivity curves" are presented in Figure 20.49 for storage wall (a, b, c) and sunspace (d) design variations.

Example 20.3.4

The previously used 2100 ft.^{2} building with NLC=11,800 Btu/FDD is preliminarily designed to be located in Medford, Oregon, with 180 ft.^{2} of 12-in thick vented Trombe wall and 130 ft.^{2} of direct gain, both systems with double glazing, nighttime insulation, and 30 Btu/ft.^{2} thermal storage capacity. Determine the annual auxiliary energy needed by this design.

Solution. Step 1 yields:

$$\text{NLC} = 11,800\ \text{Btu/FDD}.$$

$$A_\text{p} = 180 + 130 = 320\ \text{ft.}^2$$

$$\text{LCR} = 11{,}800/320 = 36.8\ \text{Btu/FDD ft.}^2$$

Step 2 yields: From Table 20.7 the short designations for the appropriate systems are

TWD4 (Trombe wall)

DGA3 (Direct gain)

Step 3 yields: From Table 20.8 for Medford, Oregon, with LCR=36.8,

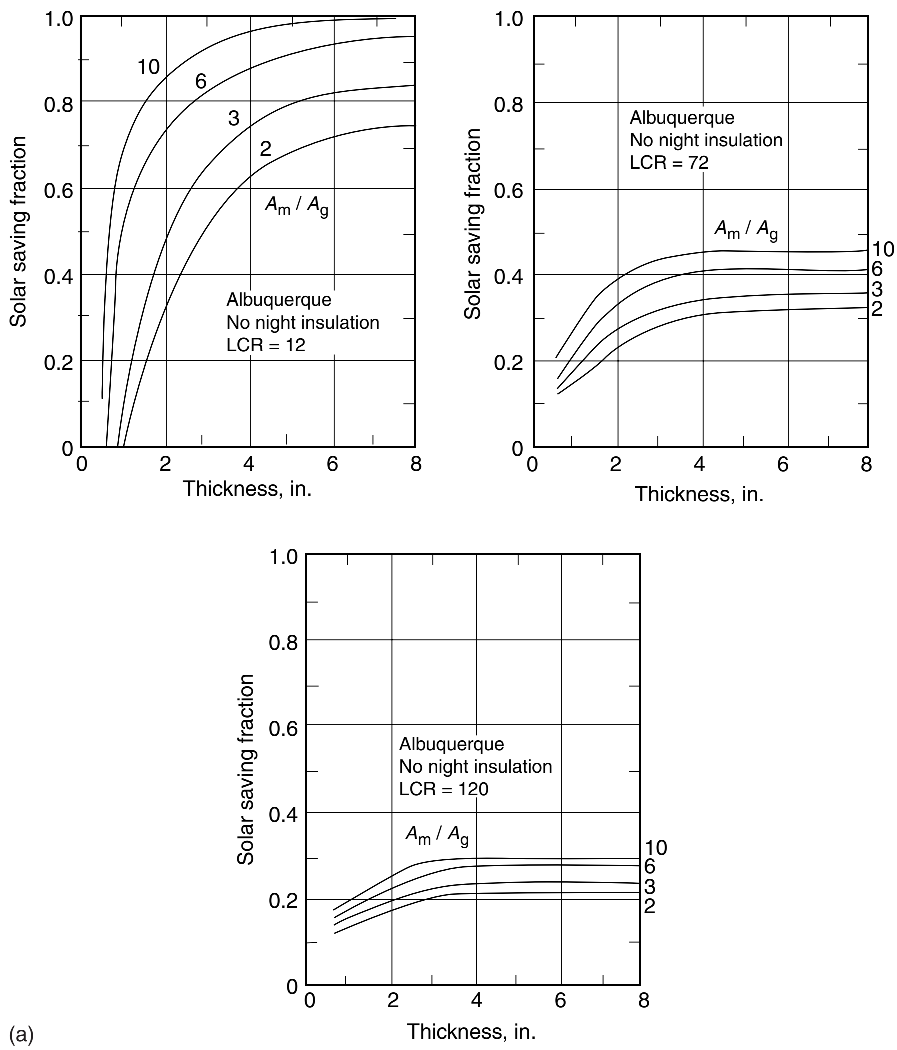


FIGURE 20.49 (a) Storage wall: Mass thickness. Sensitivity of SSF to Off-Reference conditions. (b) Storage wall: *ρck* product. (c) Storage wall: Number of glazings. (d) Sunspace : Storage volume to projected area ratio. (From PSDH. *Passive Solar Design Handbook*. Volume One: Passive Solar Design Concepts, DOE/CS-0127/1, March 1980. Prepared by Total Environmental Action, Inc. (B. Anderson, C. Michal, P. Temple, and D. Lewis); Volume Two: *Passive Solar Design Analysis*, DOE/CS-0127/2, January 1980. Prepared by Los Alamos Scientific Laboratory (J. D. Balcomb, D. Barley, R McFarland, J. Perry, W. Wray and S. Noll). U.S. Department of Energy, Washington, DC, 1980.)

$$\text{TWD4}:\text{SSF(TW)} = 0.42$$

$$\text{DGA3}:\text{SSF(DG)} = 0.37$$

Determine the "weighted area" average SSF:

$$\text{SSF} = \frac{180(0.42) + 130(0.37)}{320} = 0.39.$$

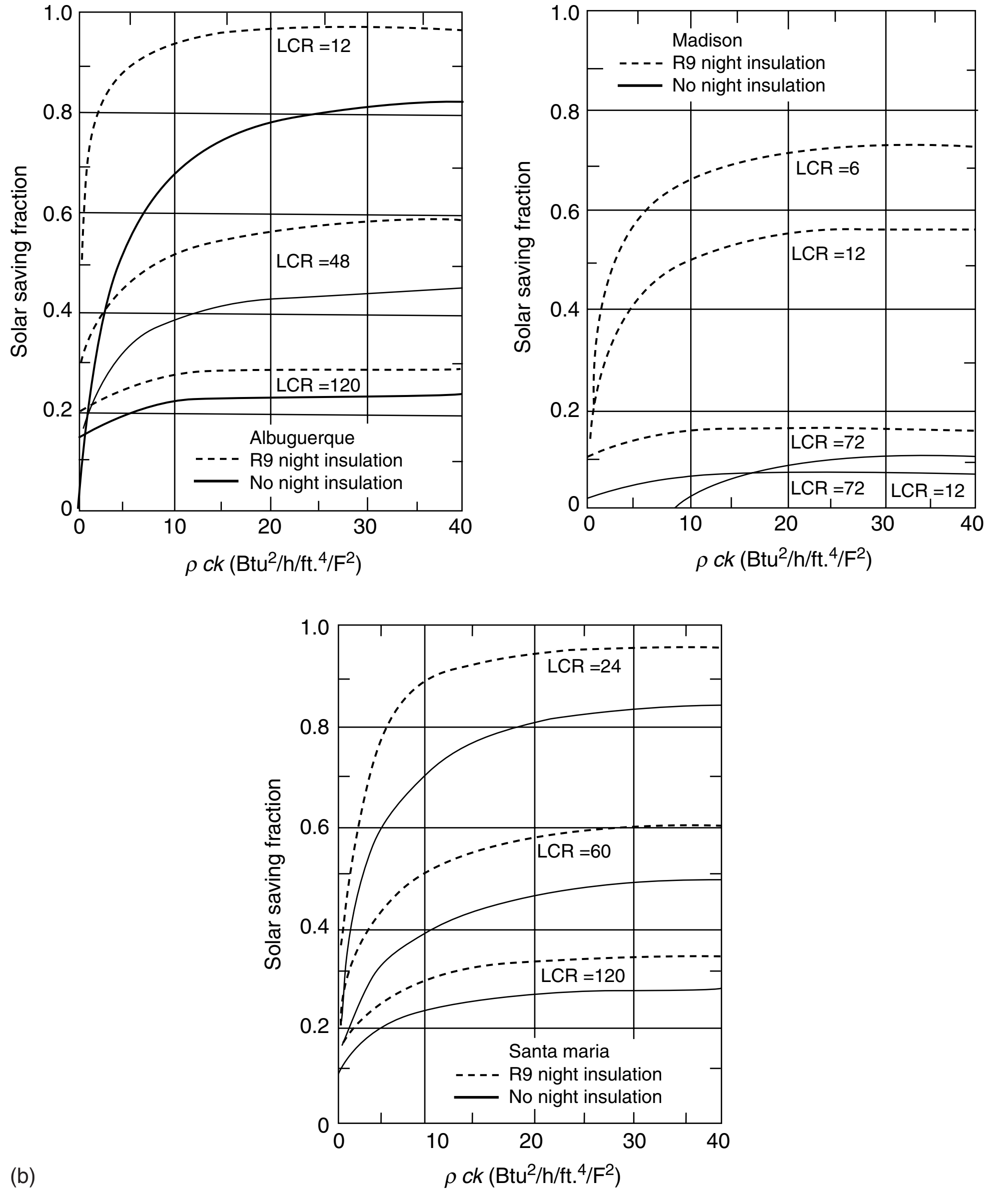


FIGURE 20.49 (*continued*)

Step 4 yields: Using Equation 20.73 and reading 4930 FDD from Table 20.8,

$$Q_{\text{aux}} = (1-0.39) \times 11,800 \text{ Btu} \times 4,930 \text{ FDD} = 35.5 \times 106 \text{ Btu annually.}$$

Using the reference system characteristics yields the thermal storage size: Trombe wall ($\rho ck = 30$, concrete properties from Table 20.7c):

$$\begin{aligned} m\,(\text{TW}) &= \text{density} \times \text{area} \times \text{thickness} \\ &= 150\ \text{lbm/ft.}^3 \times 180\ \text{ft.}^2 \times 1\ \text{ft.} \\ &= 27,000\ \text{lbm.} \end{aligned}$$

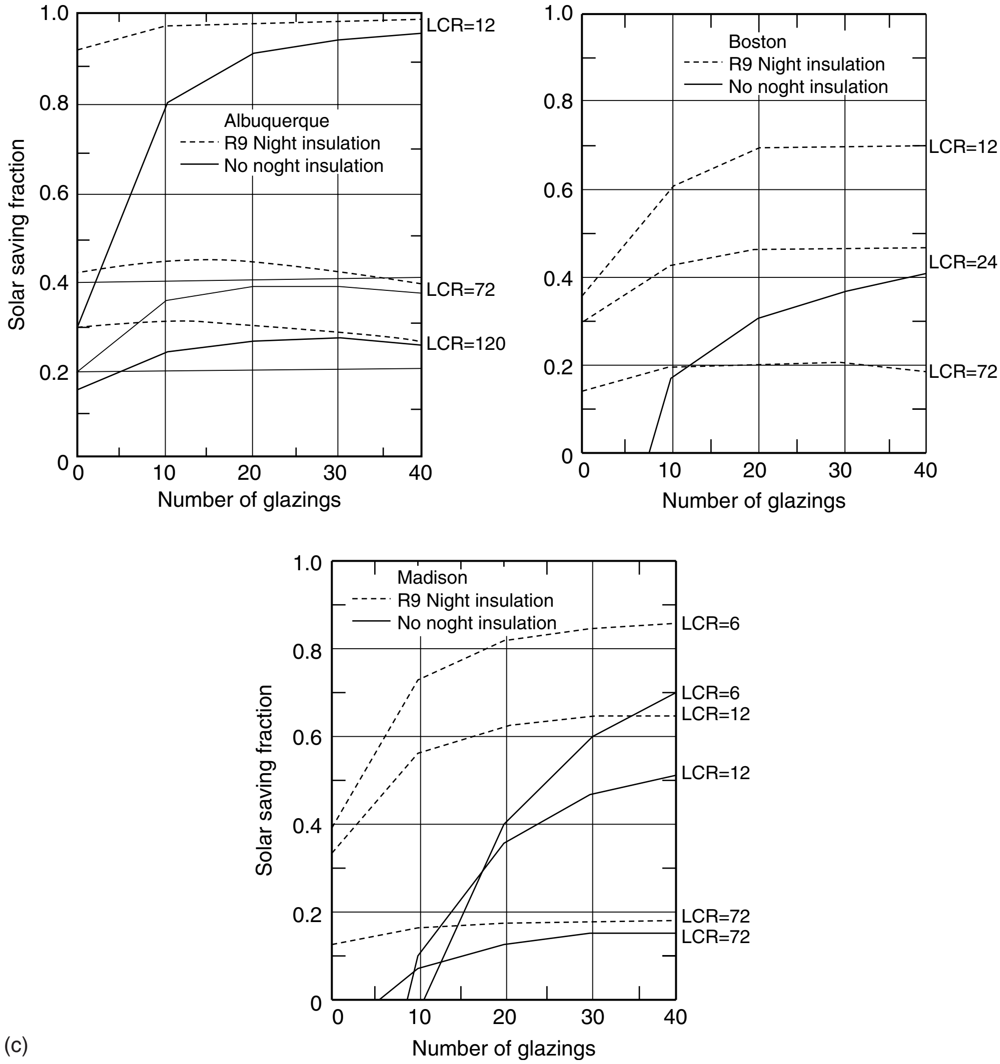


FIGURE 20.49 (*continued*)

Direct Gain (ρck=30, concrete properties), using mass area to glazing area ratio of 6:

$$\text{mass area} = 6 \times 130 = 780\ \text{ft.}^2\ \text{of}\ 2''\ \text{thick concrete}$$

$$m(\text{DG}) = 150\ \text{lbm/ft.}^3 \times 780\ \text{ft.}^2 \times 1/6\ \text{ft.}$$
$$= 19{,}500\ \text{lbm}$$

Using the LCR method allows a basic design of passive system types for the 94 reference systems, and the resulting annual performance. A bit more design variation can be obtained by using the sensitivity curves of Figure 20.49 to modify the SSF of a particular reference system. For instance, a direct gain system SSF of 0.37 would increase by approximately 0.03 if the mass-glazing-area ratio (assumed 6) were increased to 10, and would decrease by about 0.04 if the mass-glazing-area ratio were decreased to 3. This information provides a designer with quantitative information for making trade-offs.

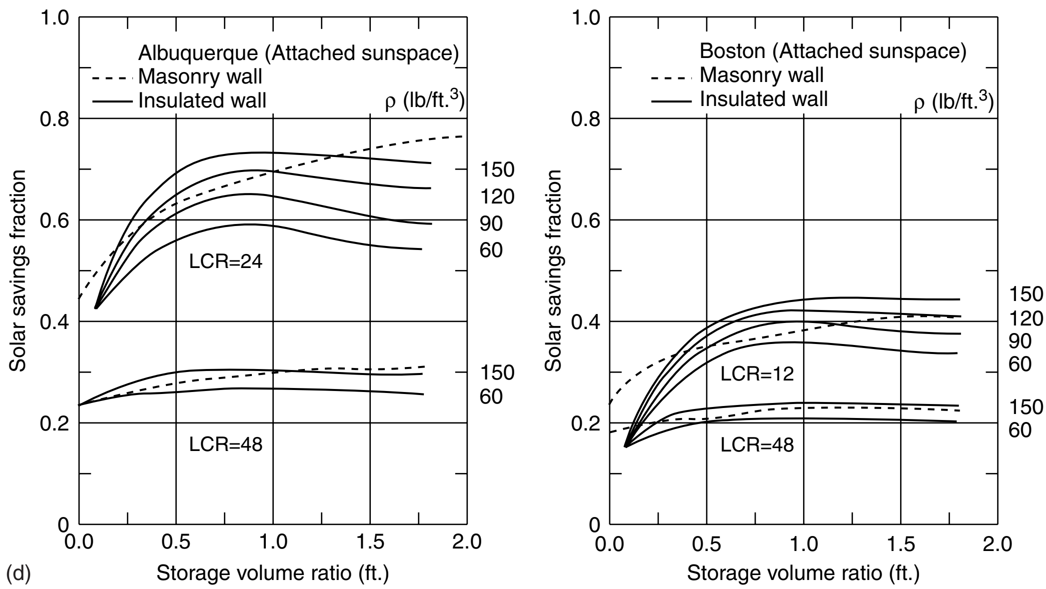


FIGURE 20.49 (*continued*)

20.3.3.10 The Third Level: SLR Method

The solar-load ratio (SLR) method calculates monthly performance, and the terms and values used are monthly based. The method allows the use of specific location weather data and the 94 reference design passive systems (Table 20.7). In addition, the sensitivity curves (Figure 20.49) can again be used to define performance outside the reference design systems. The result of the SLR method is the determination of the monthly heating auxiliary energy required that is then summed to give the annual requirement for auxiliary heating energy. Generally, the SLR method gives annual values within $\pm$ 3% of detailed simulation results, but the monthly values may vary more (PSDH 1984; Duffie and Beckman 1991). Thus, the monthly SLR method is more "accurate" than the rule-of-thumb methods, plus providing the designer with system performance on a month-by-month basis.

The SLR method uses equations and correlation parameters for each of the 94 reference systems combined with the insolation absorbed by the system, the monthly degree days, and the system's LCR to determine the monthly SSF. These correlation parameters are listed in Table 20.9 as *A*, *B*, *C*, *D*, *R*, *G*, *H*, and LCRs for each reference system (PSDH 1984). The correlation equations are

$$\text{SSF} = 1 - K(1 - F), \tag{20.76}$$

where

$$K = 1 + G/\text{LCR}, \tag{20.77}$$

$$F = \begin{cases} AX, \text{ when } X < R \\ B - C\exp(-DX), \text{ when } X > R \end{cases} \tag{20.78}$$

$$X = \frac{S/DD - (LCRs)\,H}{(LCR)K}. \tag{20.79}$$

and X is called the *generalized solar load ratio*. The term S is the monthly insolation absorbed by the system per unit of solar projected area. Monthly average daily insolation data on a vertical south facing

TABLE 20.9 SLR Correlation Parameters for the 94 Reference Systems

Type	A	B	C	D	R	G	H	LCRs	STDV
WW A1	0.0000	1.0000	0.9172	0.4841	−9.0000	0.00	1.17	13.0	0.053
WW A2	0.0000	1.0000	0.9833	0.7603	−9.0000	0.00	0.92	13.0	0.046
WW A3	0.0000	1.0000	1.0171	0.8852	−9.0000	0.00	0.85	13.0	0.040
WW A4	0.0000	1.0000	1.0395	0.9569	−9.0000	0.00	0.81	13.0	0.037
WW A5	0.0000	1.0000	1.0604	1.0387	−9.0000	0.00	0.78	13.0	0.034
WW A6	0.0000	1.0000	1.0735	1.0827	−9.0000	0.00	0.76	13.0	0.033
WW B1	0.0000	1.0000	0.9754	0.5518	−9.0000	0.00	0.92	22.0	0.051
WW B2	0.0000	1.0000	1.0487	1.0851	−9.0000	0.00	0.78	9.2	0.036
WW B3	0.0000	1.0000	1.0673	1.0087	−9.0000	0.00	0.95	8.9	0.038
WW B4	0.0000	1.0000	1.1028	1.1811	−9.0000	0.00	0.74	5.8	0.034
WW B5	0.0000	1.0000	1.1146	1.2771	−9.0000	0.00	0.56	4.5	0.032
WW C1	0.0000	1.0000	1.0667	1.0437	−9.0000	0.00	0.62	12.0	0.038
WW C2	0.0000	1.0000	1.0846	1.1482	−9.0000	0.00	0.59	8.7	0.035
WW C3	0.0000	1.0000	1.1419	1.1756	−9.0000	0.00	0.28	5.5	0.033
WW C4	0.0000	1.0000	1.1401	1.2378	−9.0000	0.00	0.23	4.3	0.032
TW A1	0.0000	1.0000	0.9194	0.4601	−9.0000	0.00	1.11	13.0	0.048
TW A2	0.0000	1.0000	0.9680	0.6318	−9.0000	0.00	0.92	13.0	0.043
TW A3	0.0000	1.0000	0.9964	0.7123	−9.0000	0.00	0.85	13.0	0.038
TW A4	0.0000	1.0000	1.0190	0.7332	−9.0000	0.00	0.79	13.0	0.032
TW B1	0.0000	1.0000	0.9364	0.4777	−9.0000	0.00	1.01	13.0	0.045
TW B2	0.0000	1.0000	0.9821	0.6020	−9.0000	0.00	0.85	13.0	0.038
TW B3	0.0000	1.0000	0.9980	0.6191	−9.0000	0.00	0.80	13.0	0.033
TW B4	0.0000	1.0000	0.9981	0.5615	−9.0000	0.00	0.76	13.0	0.028
TW C1	0.0000	1.0000	0.9558	0.4709	−9.0000	0.00	0.89	13.0	0.039
TW C2	0.0000	1.0000	0.9788	0.4964	−9.0000	0.00	0.79	13.0	0.033
TW C3	0.0000	1.0000	0.9760	0.4519	−9.0000	0.00	0.76	13.0	0.029
TW C4	0.0000	1.0000	0.9588	0.3612	−9.0000	0.00	0.73	13.0	0.026
TW D1	0.0000	1.0000	0.9842	0.4418	−9.0000	0.00	0.89	22.0	0.040
TW D2	0.0000	1.0000	1.0150	0.8994	−9.0000	0.00	0.80	9.2	0.036
TW D3	0.0000	1.0000	1.0346	0.7810	−9.0000	0.00	1.08	8.9	0.036
TW D4	0.0000	1.0000	1.0606	0.9770	−9.0000	0.00	0.85	5.8	0.035
TW D5	0.0000	1.0000	1.0721	1.0718	−9.0000	0.00	0.61	4.5	0.033
TW E1	0.0000	1.0000	1.0345	0.8753	−9.0000	0.00	0.68	12.0	0.037
TW E2	0.0000	1.0000	1.0476	1.0050	−9.0000	0.00	0.66	8.7	0.035
TW E3	0.0000	1.0000	1.0919	1.0739	−9.0000	0.00	0.61	5.5	0.034
TW E4	0.0000	1.0000	1.0971	1.1429	−9.0000	0.00	0.47	4.3	0.033
TW F1	0.0000	1.0000	0.9430	0.4744	−9.0000	0.00	1.09	13.0	0.047
TW F2	0.0000	1.0000	0.9900	0.6053	−9.0000	0.00	0.93	13.0	0.041
TW F3	0.0000	1.0000	1.0189	0.6502	−9.0000	0.00	0.86	13.0	0.036
TW F4	0.0000	1.0000	1.0419	0.6258	−9.0000	0.00	0.80	13.0	0.032
TW G1	0.0000	1.0000	0.9693	0.4714	−9.0000	0.00	1.01	13.0	0.042
TW G2	0.0000	1.0000	1.0133	0.5462	−9.0000	0.00	0.88	13.0	0.035
TW G3	0.0000	1.0000	1.0325	0.5269	−9.0000	0.00	0.82	13.0	0.031
TW G4	0.0000	1.0000	1.0401	0.4400	−9.0000	0.00	0.77	13.0	0.030
TW H1	0.0000	1.0000	1.0002	0.4356	−9.0000	0.00	0.93	13.0	0.034
TW H2	0.0000	1.0000	1.0280	0.4151	−9.0000	0.00	0.83	13.0	0.030
TW H3	0.0000	1.0000	1.0327	0.3522	−9.0000	0.00	0.78	13.0	0.029
TW H4	0.0000	1.0000	1.0287	0.2600	−9.0000	0.00	0.74	13.0	0.024
TW I1	0.0000	1.0000	0.9974	0.4036	−9.0000	0.00	0.91	22.0	0.038
TW I2	0.0000	1.0000	1.0386	0.8313	−9.0000	0.00	0.80	9.2	0.034
TW I3	0.0000	1.0000	1.0514	0.6886	−9.0000	0.00	1.01	8.9	0.034
TW I4	0.0000	1.0000	1.0781	0.8952	−9.0000	0.00	0.82	5.8	0.032
TW I5	0.0000	1.0000	1.0902	1.0284	−9.0000	0.00	0.65	4.5	0.032
TW J1	0.0000	1.0000	1.0537	0.8227	−9.0000	0.00	0.65	12.0	0.037
TW J2	0.0000	1.0000	1.0677	0.9312	−9.0000	0.00	0.62	8.7	0.035

(continued)

TABLE 20.9 (***Continued***)

Type	A	B	C	D	R	G	H	LCRs	STDV
TW J3	0.0000	1.0000	1.1153	0.9831	−9.0000	0.00	0.44	5.5	0.034
TW J4	0.0000	1.0000	1.1154	1.0607	−9.0000	0.00	0.38	4.3	0.033
DG A1	0.5650	1.0090	1.0440	0.7175	0.3931	9.36	0.00	0.0	0.046
DG A2	0.5906	1.0060	1.0650	0.8099	0.4681	5.28	0.00	0.0	0.039
DG A3	0.5442	0.9715	1.1300	0.9273	0.7068	2.64	0.00	0.0	0.036
DG B1	0.5739	0.9948	1.2510	1.0610	0.7905	9.60	0.00	0.0	0.042
DG B2	0.6180	1.0000	1.2760	1.1560	0.7528	5.52	0.00	0.0	0.035
DG B3	0.5601	0.9839	1.3520	1.1510	0.8879	2.38	0.00	0.0	0.032
DG C1	0.6344	0.9887	1.5270	1.4380	0.8632	9.60	0.00	0.0	0.039
DG C2	0.6763	0.9994	1.4000	1.3940	0.7604	5.28	0.00	0.0	0.033
DG C3	0.6182	0.9859	1.5660	1.4370	0.8990	2.40	0.00	0.0	0.031
SS A1	0.0000	1.0000	0.9587	0.4770	−9.0000	0.00	0.83	18.6	0.027
SS A2	0.0000	1.0000	0.9982	0.6614	−9.0000	0.00	0.77	10.4	0.026
SS A3	0.0000	1.0000	0.9552	0.4230	−9.0000	0.00	0.83	23.6	0.030
SS A4	0.0000	1.0000	0.9956	0.6277	−9.0000	0.00	0.80	12.4	0.026
SS A5	0.0000	1.0000	0.9300	0.4041	−9.0000	0.00	0.96	18.6	0.031
SS A6	0.0000	1.0000	0.9981	0.6660	−9.0000	0.00	0.86	10.4	0.028
SS A7	0.0000	1.0000	0.9219	0.3225	−9.0000	0.00	0.96	23.6	0.035
SS A8	0.0000	1.0000	0.9922	0.6173	−9.0000	0.00	0.90	12.4	0.028
SS B1	0.0000	1.0000	0.9683	0.4954	−9.0000	0.00	0.84	16.3	0.028
SS B2	0.0000	1.0000	1.0029	0.6802	−9.0000	0.00	0.74	8.5	0.026
SS B3	0.0000	1.0000	0.9689	0.4685	−9.0000	0.00	0.82	19.3	0.029
SS B4	0.0000	1.0000	1.0029	0.6641	−9.0000	0.00	0.76	9.7	0.026
SS B5	0.0000	1.0000	0.9408	0.3866	−9.0000	0.00	0.97	16.3	0.030
SS B6	0.0000	1.0000	1.0068	0.6778	−9.0000	0.00	0.84	8.5	0.028
SS B7	0.0000	1.0000	0.9395	0.3363	−9.0000	0.00	0.95	19.3	0.032
SS B8	0.0000	1.0000	1.0047	0.6469	−9.0000	0.00	0.87	9.7	0.027
SS C1	0.0000	1.0000	1.0087	0.7683	−9.0000	0.00	0.76	16.3	0.025
SS C2	0.0000	1.0000	1.0412	0.9281	−9.0000	0.00	0.78	10.0	0.027
SS C3	0.0000	1.0000	0.9699	0.5106	−9.0000	0.00	0.79	16.3	0.024
SS C4	0.0000	1.0000	1.0152	0.7523	−9.0000	0.00	0.81	10.0	0.025
SS D1	0.0000	1.0000	0.9889	0.6643	−9.0000	0.00	0.84	17.8	0.028
SS D2	0.0000	1.0000	1.0493	0.8753	−9.0000	0.00	0.70	9.9	0.028
SS D3	0.0000	1.0000	0.9570	0.5285	−9.0000	0.00	0.90	17.8	0.029
SS D4	0.0000	1.0000	1.0356	0.8142	−9.0000	0.00	0.73	9.9	0.028
SS E1	0.0000	1.0000	0.9968	0.7004	−9.0000	0.00	0.77	19.6	0.027
SS E2	0.0000	1.0000	1.0468	0.9054	−9.0000	0.00	0.76	10.8	0.027
SS E3	0.0000	1.0000	0.9565	0.4827	−9.0000	0.00	0.81	19.6	0.028
SS E4	0.0000	1.0000	1.0214	0.7694	−9.0000	0.00	0.79	10.8	0.027

Source: From PSDH, *Passive Solar Design Handbook*, Los Alamos National Laboratory, Van Nostrand Reinhold, New York, 1984.

surface can be found and/or calculated using various sources (PSDH 1984; McQuiston and Parker 1994) and the S term can be determined by multiplying by a transmission and an absorption factor and the number of days in the month. Absorption factors for all systems are close to 0.96 (PSDH 1984), whereas the transmission is approximately 0.9 for single glazing, 0.8 for double glazing, and 0.7 for triple glazing.

Example 20.3.5

For a vented, 180 ft.^{2}, double-glazed with night insulation, 12-in thick Trombe wall system (TWD4) in a NLC = 11,800 Btu/FDD house in Medford, Oregon, determine the auxiliary energy required in January.

Solution. Weather data for Medford, Oregon (PSDH 1984) yields for January ($N = 31$, days): daily vertical surface insolation = 565 Btu/ft.^{2} and 880 FDD, so $S = (31)(565)(0.8)(0.96) =$ 13,452 Btu/ft.^{2} month.

$$\text{LCR} = \text{NLC}/A_\text{p} = 11,800/180 = 65.6\ \text{Btu/FDD ft.}^2.$$

From Table 20.9 at TWD4: $A=0$, $B=1$, $C=1.0606$, $D=0.977$, $R=-9$, $G=0$, $H=0.85$, LCRs$=$ 5.8 Btu/FDD ft.^{2}.

Substituting into Equation 20.77 gives

$$K = 1 + 0/65.6 = 1.$$

Equation 20.79 gives

$$X = \frac{(13,452/880) - (5.8 \times 0.85)}{65.6 \times 1} = 0.16.$$

Equation 20.78 gives

$$F = 1 - 1.0606\ \text{e}^{-0.977 \times 0.16} = 0.09,$$

and Equation 20.76 gives

$$\text{SSF} = 1 - 1(1 - 0.09) = 0.09.$$

The January auxiliary energy required can be calculated using Equation 20.73:

$$\begin{aligned} Q_\text{aux}(\text{Jan}) &= (1 - \text{SSF}) \times \text{NLC} \times (\text{Number of degree days}) \\ &= (1 - 0.09) \times 11,800 \times 880 \\ &= 9,450,000\ \text{Btu}. \end{aligned}$$

As mentioned, the use of sensitivity curves (PSDH 1984) as in Figure 20.49 will allow SSF to be determined for many off-reference system design conditions involving storage mass, number of glazings, and other more esoteric parameters. Also, the use of multiple passive system types within one building would be approached by calculating the SSF for each type system individually using a "combined area" LCR, and then a weighted-area (aperture) average SSF would be determined for the building.

20.3.4 Passive Space Cooling Design Fundamentals

Passive cooling systems are designed to use natural means to transfer heat from buildings, including convection/ventilation, evaporation, radiation, and conduction. However, the most important element in both passive and conventional cooling design is to prevent heat from entering the building in the first place. Cooling conservation techniques involve building surface colors, insulation, special window glazings, overhangs and orientation, and numerous other architectural/engineering features.

20.3.4.1 Solar Control

Controlling the solar energy input to reduce the cooling load is usually considered a passive (versus conservation) design concern because solar input may be needed for other purposes, such as daylighting throughout the year and/or heating during the winter. Basic architectural solar control is normally "designed in" via the shading of the solar windows, where direct radiation is desired for winter heating and needs to be excluded during the cooling season.

The shading control of the windows can be of various types and "controllability," ranging from drapes and blinds, use of deciduous trees, to the commonly used overhangs and vertical louvers. A rule of thumb design for determining proper south-facing window overhang for both winter heating and summer

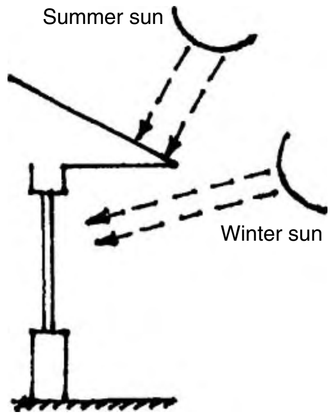
TABLE 20.10 South-Facing Window Overhang Rule of Thumb

Length of the Overhang $= \frac{\text{Window Height}}{F}$

(a) Overhang Factors

North Latitude	F^{a}
28	5.6–11.1
32	4.0–6.3
36	3.0–4.5
40	2.5–3.4
44	2.0–2.7
48	1.7–2.2
52	1.5–1.8
56	1.3–1.5

(b) Roof Overhang Geometry



Properly sized overhangs shade out hot summer sun but allow winter sun (which is lower in the sky) to penetrate windows

[a] Select a factor according to your latitude. Higher values provide complete shading at noon on June 21; lower values, until August 1.

Source: From Halacy, 1984.

shading is presented in Table 20.10. Technical details on calculating shading from various devices and orientations are found in Olgyay and Olgyay (1976) and ASHRAE (1989, 1993, 1997).

20.3.4.2 Natural Convection/Ventilation

Air movement provides cooling comfort through convection and evaporation from human skin. ASHRAE (1989) places the comfort limit at 79°F for an air velocity of 50 ft./min (fpm), 82°F for 160 fpm, and 85°F for 200 fpm. To determine whether or not comfort conditions can be obtained, a designer must calculate the volumetric flow rate, Q, which is passing through the occupied space. Using the cross-sectional area, A_x, of the space and the room air velocity, V_a, required, the flow is determined by

$$Q = A_x V_a. \tag{20.80}$$

The proper placement of windows, "narrow" building shape, and open landscaping can enhance natural wind flow to provide ventilation. The air flow rate through open windows for wind-driven ventilation is given by ASHRAE (1989, 1993, 1997):

$$Q = C_v V_w A_w, \tag{20.81}$$

where Q is air flow rate (m^3/s), A_w is free area of inlet opening (m^2), V_w is wind velocity (m/s), and C_v is effectiveness of opening that is equal to 0.5–0.6 for wind perpendicular to opening, and 0.25–0.35 for wind diagonal to opening.

The stack effect can induce ventilation when warm air rises to the top of a structure and exhausts outside, while cooler outside air enters the structure to replace it. Figure 20.50 illustrates the solar

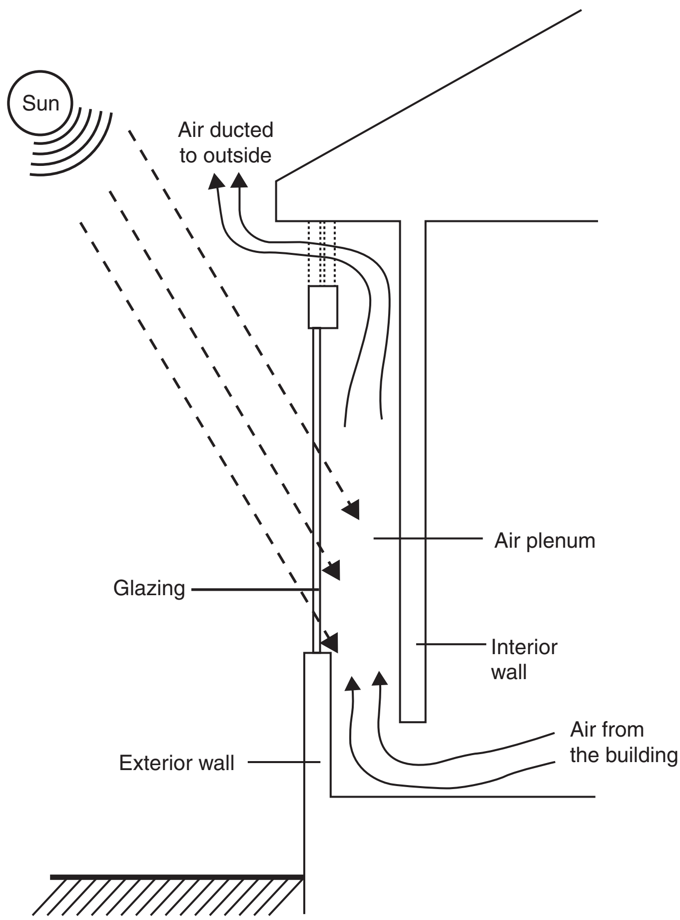


FIGURE 20.50 The stack-effect/solar chimney concept to induce convection/ventilation. (From PSDH. *Passive Solar Design Handbook*. Volume One: Passive Solar Design Concepts, DOE/CS-0127/1, March 1980. Prepared by Total Environmental Action, Inc. (B. Anderson, C. Michal, P. Temple, and Lewis); Volume Two: *Passive* Solar Design Analysis, DOE/CS-0127/2, January 1980. Prepared by Los Alamos Scientific Laboratory (J. D. Balcomb, D. Barley, R McFarland, J. Perry, W. Wray and S. Noll). U.S. Department of Energy, Washington, DC, 1980.)

chimney concept, which can easily be adapted to a thermal storage wall system. The greatest stack-effect flow rate is produced by maximizing the stack height and the air temperature in the stack, as given by

$$Q = 0.116\, A_j \sqrt{h(T_s - T_o)} \qquad (20.82)$$

where Q is stack flow rate (m^3/s), A_j is the area of inlets or outlets, whichever is smaller (m^2), h is the inlet to outlet height (m), T_s is the average temperature in stack (°C), and T_o is the outdoor air temperature (°C).

If inlet or outlet area is twice the other, the flow rate will increase by 25%, and by 35% if the areas' ratio is 3:1 or larger (Table 20.11).

Example 20.3.6

A two-story (5-m) solar chimney is being designed to produce a flow of 0.25 m^3/s through a space. The preliminary design features include a 25 cm × 1.5 m inlet, a 50 cm × 1.5 m outlet, and an estimated 35°C average stack temperature on a sunny 30°C day. Can this design produce the desired flow?

TABLE 20.11 Ground Reflectivities

Material	ρ (%)
Cement	27
Concrete	20–40
Asphalt	7–14
Earth	10
Grass	6–20
Vegetation	25
Snow	70
Red brick	30
Gravel	15
White paint	55–75

Source: From Murdoch, J. B., *Illumination Engineering—From Edison's Lamp to the Laser*, Macmillan, New York, 1985.

Solution. Substituting the design data into Equation 20.82,

$$Q = 0.116(0.25 \times 1.5)[5(5)]^{1/2}$$
$$= 0.2 \text{ m}^3/\text{s}.$$

Because the outlet area is twice the inlet area, the 25% flow increase can be used:

$$Q = 0.2(1.25) = 0.25\text{m}^3/\text{s}$$

(answer: Yes, the proper flow rate is obtained).

20.3.4.3 Evaporative Cooling

When air with less than 100% relative humidity moves over a water surface, the evaporation of water causes both the air and the water itself to cool. The lowest temperature that can be reached by this direct evaporative cooling effect is the wet-bulb temperature of the air, which is directly related to the relative humidity, with lower wet-bulb temperature associated with lower relative humidity. Thus, dry air (low relative humidity) has a low wet-bulb temperature and will undergo a large temperature drop with evaporative cooling, while humid air (high relative humidity) can only be slightly cooled evaporatively. The wet-bulb temperature for various relative humidity and air temperature conditions can be found via the "psychrometric chart" available in most thermodynamic texts. Normally, an evaporative cooling process cools the air only part of the way down to the wet-bulb temperature. To get the maximum temperature decrease, it is necessary to have a large water surface area in contact with the air for a long time, and interior ponds and fountain sprays are often used to provide this air-water contact area.

The use of water sprays and open ponds on roofs provides cooling primarily via evaporation. The hybrid system involving a fan and wetted mat, the "swamp cooler," is by far the most widely used evaporative cooling technology. Direct, indirect, and combined evaporative cooling system design features are described in ASHRAE (1989, 1991, 1993, 1995, 1997).

20.3.4.4 Nocturnal and Radiative Cooling Systems

Another approach to passive convective/ventilative cooling involves using cooler night air to reduce the temperature of the building and/or a storage mass. Thus, the building/storage mass is prepared to accept part of the heat load during the hotter daytime. This type of convective system can also be combined with evaporative and radiative modes of heat transfer, utilizing air and/or water as the convective fluid. Work in Australia (Close et al. 1968) investigated rock storage beds that were chilled using evaporatively cooled night air. Room air was then circulated through the bed during the day to provide space cooling. The use of encapsulated roof ponds as a thermal cooling mass has been tried by

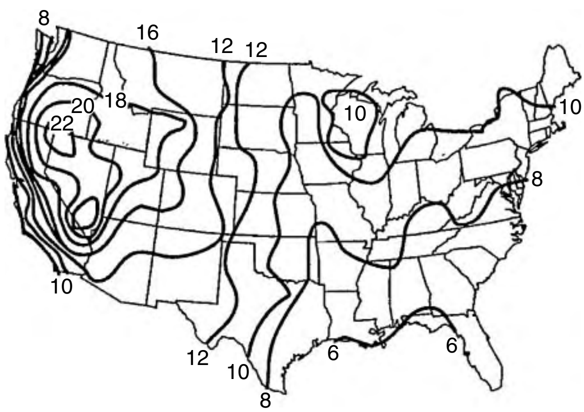


FIGURE 20.51 Average monthly sky temperature depression ($I_{AIR} - I_{SKY}$) for July in °F. (Adapted from Martin, M. and Berdahl, P., *Solar Energy*, 33(314), 321–336, 1984.

several investigators (Hay and Yellot 1969; Marlatt, Murray, and Squire 1984; Givoni 1994) and is often linked with nighttime radiative cooling.

All warm objects emit thermal infrared radiation; the hotter the body, the more energy it emits. A passive cooling scheme is to use the cooler night sky as a sink for thermal radiation emitted by a warm storage mass, thus chilling the mass for cooling use the next day. The net radiative cooling rate, Q_r, for a horizontal unit surface (ASHRAE 1989, 1993, 1997) is

$$Q_r = \varepsilon\sigma\left(T_{\text{body}}^4 - T_{\text{sky}}^4\right), \tag{20.83}$$

where Q_r is the net radiative cooling rate, W/m^2 (Btu/h ft.^{2}), ε is the surface emissivity fraction (usually 0.9 for water), σ is 5.67×10^{-8} W/m^2 K^4 (1.714×10^{-9} Btu/h ft.^{2} R^4), T_{body} is the warm body temperature, Kelvin (Rankine), and T_{sky} is the effective sky temperature, Kelvin (Rankine).

The monthly average air–sky temperature difference has been determined (Martin and Berdahl 1984) and Figure 20.51 presents these values for July (in °F) for the United States.

Example 20.3.7

Estimate the overnight cooling possible for a 10 m^2, 85°F water thermal storage roof during July in Los Angeles.

Solution. Assume the roof storage unit is black with $\varepsilon=0.9$. From Figure 20.51, $T_{\text{air}}-T_{\text{sky}}$ is approximately 10°F for Los Angeles. From weather data for LA airport, (PSDH 1984; ASHRAE 1989), the July average temperature is 69°F with a range of 15°F. Assuming night temperatures vary from the average (69°F) down to half the daily range (15/2), then the average nighttime temperature is chosen as $69-(1/2)(15/2)=65$°F. Therefore, $T_{\text{sky}}=65-10=55$°F. From Equation 20.83,

$$Q_r = 0.9(1.714\times10^{-9})[(460+85)^4-(460+55)^4]$$
$$= 27.6 \text{ Btu/h ft.}^2.$$

For a 10-h night and 10 m^2 (107.6 ft.^{2}) roof area,

$$\text{Total radiative cooling} = 27.6(10)(107.6)$$
$$= 29,700 \text{ Btu.}$$

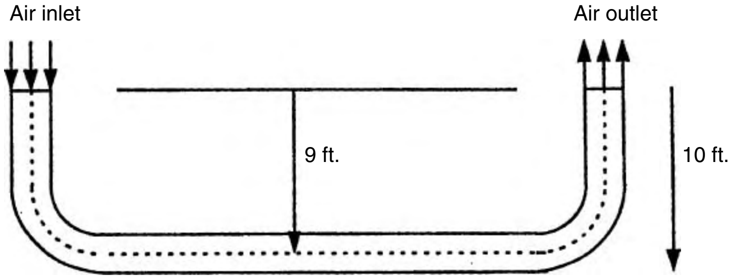


FIGURE 20.52 Open loop underground air tunnel system.

Note that this does not include the convective cooling possible, which can be approximated (at its maximum rate) for still air (ASHRAE 1989, 1993, 1997) by

$$\begin{aligned}\text{Maximum total } Q_{\text{conv}} &= hA(T_{\text{roof}} - T_{\text{air}})(\text{Time}) \\ &= 5(129)(85-55)(10) \\ &= 161{,}000 \text{ Btu.}\end{aligned}$$

This is a maximum since the 85°F storage temperature will drop as it cools; this is also the case for the radiative cooling calculation. However, convection is seen to usually be the more dominant mode of nighttime cooling.

20.3.4.5 Earth Contact Cooling (or Heating)

Earth contact cooling or heating is a passive summer cooling and winter heating technique that utilizes underground soil as the heat sink or source. By installing a pipe underground and passing air through the pipe, the air will be cooled or warmed depending on the season. A schematic of an open loop system and a closed loop air-conditioning system are presented in Figure 20.52 and Figure 20.53, respectively (Goswami and Biseli 1994).

The use of this technique can be traced back to 3000 BC when Iranian architects designed some buildings to be cooled by natural resources only. In the nineteenth century, Wilkinson (USDA 1960) designed a barn for 148 cows where a 500-ft. long underground passage was used for cooling during the summertime. Since that time, a number of experimental and analytical studies of this technique have

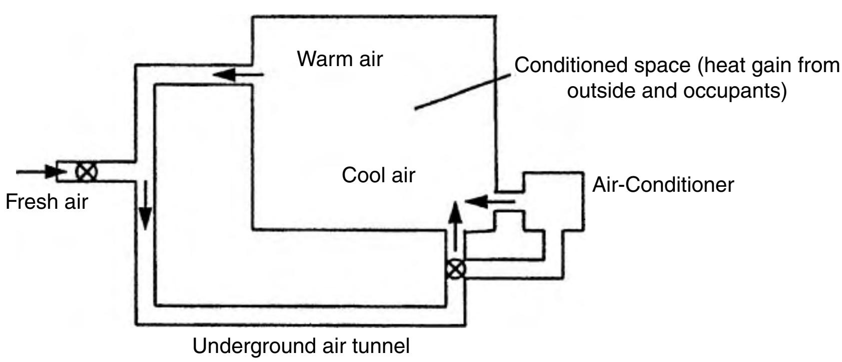


FIGURE 20.53 Schematic of closed loop air-conditioning system using air-tunnel.

continued to appear in the literature (Krarti and Kreider 1996; Hollmuller and Lachal 2001; De Paepe and Janssens 2003). Goswami and Dhaliwal (1985) have given a brief review of the literature, as well as presenting an analytical solution to the problem of transient heat transfer between the air and the surrounding soil as the air is made to pass through a pipe buried underground.

20.3.4.5.1 Heat Transfer Analysis

The transient thermal analysis of the air and soil temperature fields (Goswami and Dhaliwal 1985) is conducted using finite elements with the convective heat transfer between the air and the pipe and using semi-infinite cylindrical conductive heat transfer to the soil from the pipe. It should be noted that the thermal resistance of the pipe (whether of metal, plastic or ceramic) is negligible relative to the surrounding soil.

Air and Pipe Heat Transfer—The pipe is divided into a large number of elements and a psychrometric energy balance written for each, depending on whether the air leaves the element (1) unsaturated, or (2) saturated.

1. If the air leaves an element as unsaturated, the energy balance on the element is

$$mC_p(T_1 - T_2) = hA_p(T_{air} - T_{pipe}). \tag{20.84}$$

   T_{air} can be taken as $(T_1 + T_2)/2$. Substituting and simplifying,

$$T_2 = \left[\left(1 - \frac{U}{2}\right)T_1 + UT_{pipe}\right] / \left(1 + \frac{U}{2}\right), \tag{20.85}$$

   where U is defined as

$$U = \frac{A_p h}{mC_p}$$

2. If the air leaving the element is saturated, the energy balance is

$$mC_pT_1 + m(W_1 - W_2)H_{fg} = mC_pT_2 + hA_p(T_{air} - T_{pipe}). \tag{20.86}$$

   Simplifying gives:

$$T_2 = \left(1 - \frac{U}{2}T_1\right) + \frac{W_1 - W_2}{C_p}H_{fg} + UT_{pipe} / \left(1 + \frac{U}{2}\right). \tag{20.87}$$

   The convective heat transfer coefficient h in the preceding equations depends on Reynolds number, the shape, and roughness of the pipe.

Using the exit temperature from the first element as the inlet temperature for the next element, the exit temperature for the element can be calculated in a similar way. Continuing this way from one element to the next, the temperature of air at the exit from the pipe can be calculated.

Soil Heat Transfer—The heat transfer from the pipe to the soil is analyzed by considering the heat flux at the internal radius of a semi-infinite cylinder formed by the soil around the pipe. For a small element the problem can be formulated as

$$\frac{\partial^2 T(r,t)}{\partial r^2} + \frac{1}{r}\frac{\partial T(r,t)}{\partial r} = \frac{1}{\alpha}\frac{\partial T(r,t)}{\partial t}, \tag{20.88}$$

with initial and boundary conditions as

$$T(r,0) = T_e,$$

$$T(\infty,t) = T_e,$$

$$-K\frac{\partial T}{\partial r}(r,t) = q'',$$

where T_e is the bulk earth temperature and q'' is also given by the amount of heat transferred to the pipe from the air by convection, i.e., $q^{11} = h(T_{air} - T_{pipe})$.

20.3.4.5.2 Soil Temperatures and Properties

Kusuda and Achenback (1965) and Labs (1981) studied the earth temperatures in the United States. According to both of these studies, temperature swings in the soil during the year are dampened with depth below the ground. There is also a phase lag between the soil temperature and the ambient air temperature, and this phase lag increases with depth below the surface. For example, the soil temperature for light dry soil at a depth of about 10 ft. (3.05 m) varies by approximately $\pm$5°F (2.8°C) from the mean temperature (approximately equal to mean annual air temperature) and has a phase lag of approximately 75 days behind ambient air temperature (Labs 1981).

The thermal properties of the soil are difficult to determine. The thermal conductivity and diffusivity both change with the moisture content of the soil itself, which is directly affected by the temperature of and heat flux from and to the buried pipe. Most researchers have found that using constant property values for soil taken from standard references gives reasonable predictive results (Goswami and Ileslamlou 1990).

20.3.4.5.3 Generalized Results from Experiments

Figure 20.54 presents data from Goswami and Biseli (1994) for an open system, 100-ft. long, 12-in diameter pipe, buried 9 ft. deep. The figure shows the relationship between pipe inlet-to-outlet temperature reduction ($T_{in} - T_{out}$) and the initial soil temperature with ambient air inlet conditions of 90°F, 55% relative humidity for various pipe flow rates.

Other relations from this same report which can be used with the Figure 20.54 data include: (1) the effect of increasing pipe/tunnel length on increasing the inlet-to-outlet air temperature difference is fairly

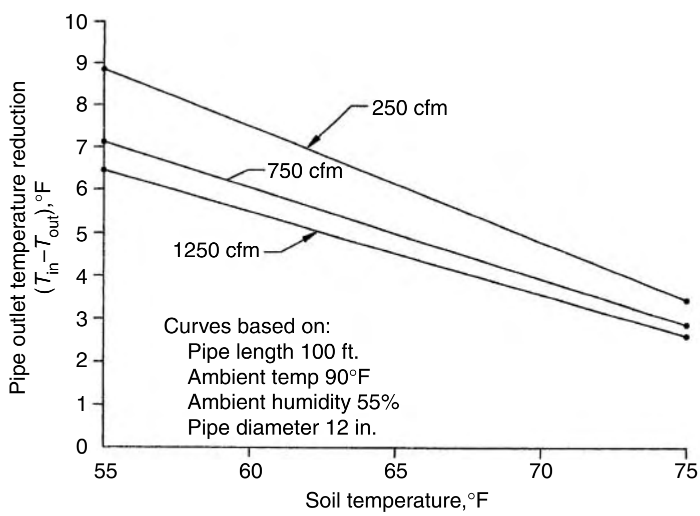


FIGURE 20.54 Air temperature drop through a 100-ft. long, 12-in. diameter pipe buried 9 ft. underground.

linear up to 250 ft.; and (2) the effect of decreasing pipe diameter on lowering the outlet air temperature is slight, and only marginally effective for pipes less than 12-in. in diameter.

Example 20.3.8

Provide the necessary 12-in diameter pipe length(s) that will deliver 1500 cfm of 75°F air if the ambient temperature is 85°F and the soil at 9 ft. is 65°F.

Solution. From Figure 20.54, for 100 ft. of pipe at 65°F soil temperature, the pipe temperature reduction is

$$\begin{aligned} T_{in} - T_{out} &= 6°\text{F (at 250 cfm)} \\ &= 5°\text{F (at 750 cfm)} \\ &= 4.5°\text{F (at 1250 cfm)}. \end{aligned}$$

Because the "length versus temperature reduction" is linear (see text above), the 10°F reduction required (85 down to 75) would be met by the 750 cfm case (5°F for 100 ft.) if 200 ft. of pipe is used. Then, two 12-in diameter pipes would be required to meet the 1500-cfm requirement.

Answer: Two 12-in diameter pipes, each 200 ft. long. (Note: see what would be needed if the 250 cfm or the 1250 cfm cases had been chosen. Which of the three flow rate cases leads to the "cheapest" installation?)

20.3.5 Daylighting Design Fundamentals

Daylighting is the use of the sun's radiant energy to illuminate the interior spaces in a building. In the 19th century, electric lighting was considered an alternative technology to daylighting. Today the situation is reversed, primarily due to the economics of energy use and conservation. However, there are good physiological reasons for using daylight as an illuminant. The quality of daylight matches the human eye's response, thus permitting lower light levels for task comfort, better color rendering, and clearer object discrimination (Robbins 1986; McCluney 1998; Clay 2001).

20.3.5.1 Lighting Terms and Units

Measurement of lighting level is based on the "standard candle", where the lumen (lm), the unit of luminous flux (ϕ), is defined as the rate of luminous energy passing through a 1-m^2 area located 1 m from the candle. Thus, a standard candle generates 4π lumens, which radiate away in all directions. The illuminance (E) on a surface is defined as the luminous flux on the surface divided by the surface area, $E = \phi/A$. Illuminance is measured in either lux (lx), as lm/m^2, or footcandles (fc), as lm/ft.^{2}.

Determination of the daylighting available at a given location in a building space at a given time is important to evaluate the reduction possible in electric lighting and the associated impact on heating and cooling loads. Daylight provides about 110 lm/W of solar radiation, fluorescent lamps about 75 lm/W of electrical input, and incandescent lamps about 20 lm/W; thus daylighting generates only 1/2–1/5 the heating that equivalent electric lighting does, significantly reducing the building cooling load.

20.3.5.2 Approach to Daylighting Design

Aperture controls such as blinds and drapes are used to moderate the amount of daylight entering the space, as are the architectural features of the building itself (glazing type, area, and orientation; overhangs and wingwalls; lightshelves; etc.). Many passive and "active" reflective, concentrating, and diffusing devices are available to specifically gather and direct both the direct and diffuse components of

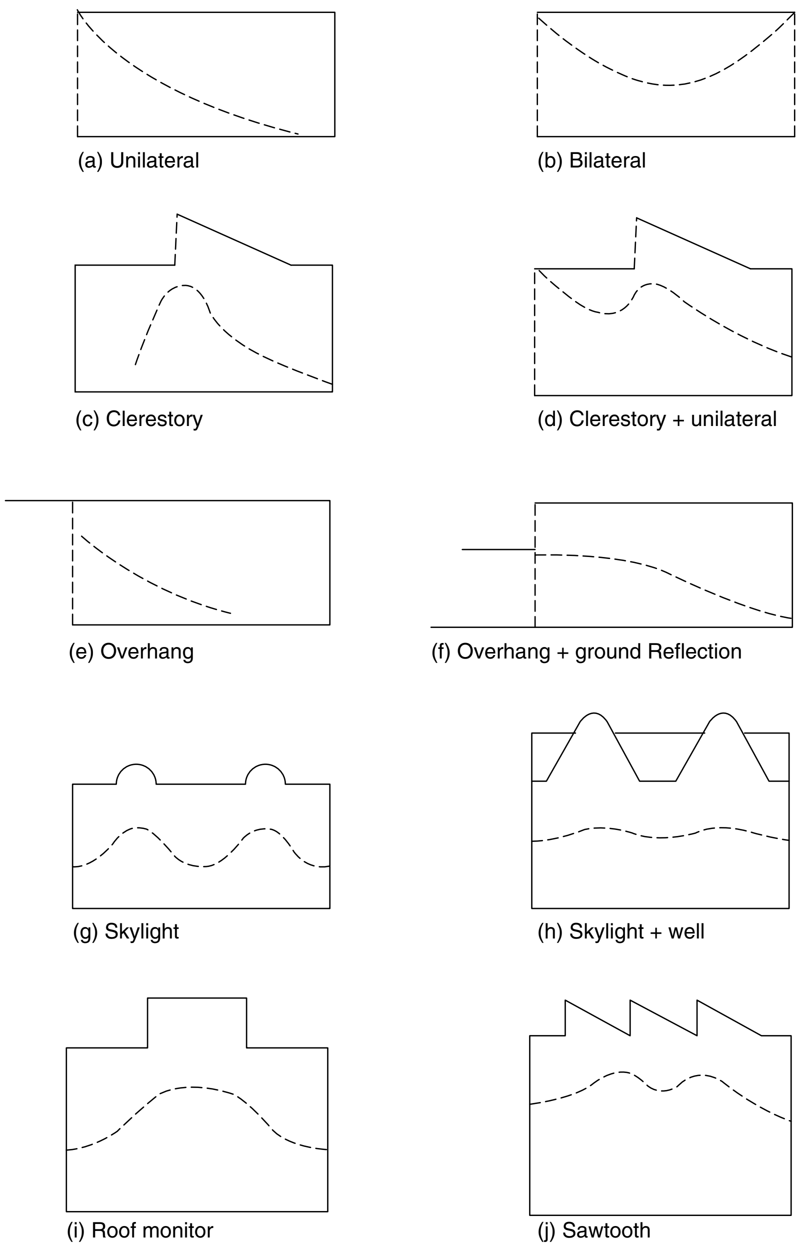


FIGURE 20.55 Examples of sidelighting and toplighting architectural features (dashed lines represent illuminance distributions). (From Murdoch, J. B., *Illumination Engineering from Edison's Lamp to the Laser*. Macmillan, New York, 1985)

daylight to areas within the space (Kinney et al. 2005). Electric-lighting dimming controls are used to adjust the electric light level based on the quantity of the daylighting. With these two types of controls (aperture and lighting), the electric lighting and cooling energy use and demand, as well as cooling system sizing, can be reduced. However, the determination of the daylighting position and time illuminance value within the space is required before energy usage and demand reduction calculations can be made.

Daylighting design approaches use both solar beam radiation (referred to as sunlight) and the diffuse radiation scattered by the atmosphere (referred to as skylight) as sources for interior lighting, with historical design emphasis being on utilizing skylight. Daylighting is provided through a variety of glazing features, which can be grouped as sidelighting (light enters via the side of the space) and toplighting (light enters from the ceiling area). Figure 20.55 illustrates several architectural forms producing sidelighting and toplighting, with the dashed lines representing the illuminance distribution within the space. The calculation of work-plane illuminance depends on whether sidelighting and/or toplighting features are used, and the combined illuminance values are additive.

20.3.5.3 Sun-Window Geometry

The solar illuminance on a vertical or horizontal window depends on the position of the sun relative to that window. In the method described here, the sun and sky illuminance values are determined using the sun's altitude angle (α) and the sun-window azimuth angle difference (Φ). These angles need to be determined for the particular time of day, day of year, and window placement under investigation.

20.3.5.3.1 Solar Altitude Angle

The solar altitude angle, α, is the angle swept out by a person's arm when pointing to the horizon directly below the sun and then raising the arm to point at the sun. The equation to calculate solar altitude, α, is

$$\sin \alpha = \cos L \cos \delta \cos H + \sin L \sin \delta, \tag{20.89}$$

where L is the local latitude (degrees), δ is the earth-sun declination (degrees) given by $\delta = 23.45 \sin[360(n-81)/365]$, n is the day number of the year, and H is the hour angle (degrees) given by

$$H = \frac{(12 \text{ noon} - \text{time})(\text{in minutes})}{4}; \; (+\text{ morning}, \; -\text{afternoon}). \tag{20.90}$$

20.3.5.3.2 Sun-Window Azimuth Angle Difference

The difference between the sun's azimuth and the window's azimuth, Φ, needs to be calculated for vertical window illuminance. The window's azimuth angle, γ_w, is determined by which way it faces, as measured from south (east of south is positive, westward is negative). The solar azimuth angle, γ_s, is calculated:

$$\sin \gamma_s = \frac{\cos \delta \sin H}{\cos \alpha}. \tag{20.91}$$

The sun-window azimuth angle difference, Φ, is given by the absolute value of the difference between γ_s and γ_w:

$$\Phi = |\gamma_s - \gamma_w|. \tag{20.92}$$

20.3.5.4 Daylighting Design Methods

To determine the annual lighting energy saved (ES_L), calculations using the lumen method described below should be performed on a monthly basis for both clear and overcast days for the space under investigation. Monthly weather data for the site would then be used to prorate clear and overcast lighting energy demands monthly. Subtracting the calculated daylighting illuminance from the design illuminance leaves the supplementary lighting needed, which determines the lighting energy required.

The approach in the method below is to calculate the "sidelighting" and the "skylighting" of the space separately, and then combine the results. This procedure has been computerized (Lumen II/ Lumen Micro) and includes many details of controls, daylighting technologies, and weather. ASHRAE (1989, 1993, 1997) lists many of the methods and simulation techniques currently used with daylighting and its associated energy effects.

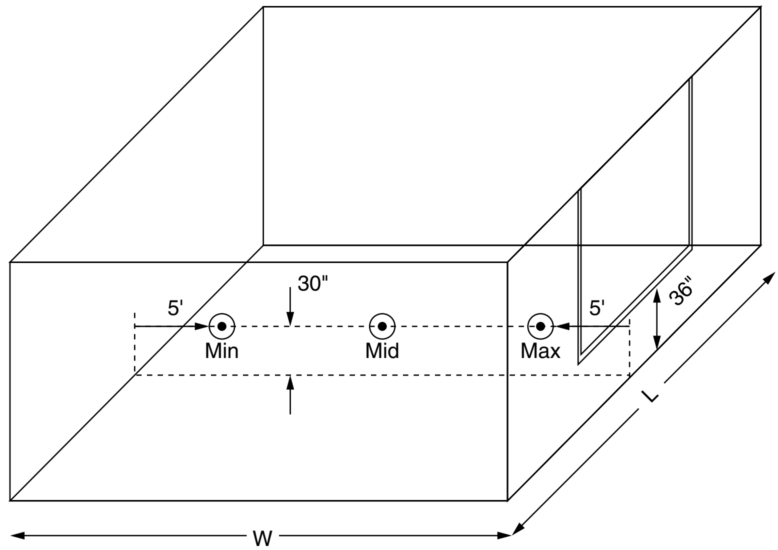


FIGURE 20.56 Location of illumination points within the room (along centerline of window) determined by lumen method of sidelighting.

20.3.5.4.1 Lumen Method of Sidelighting (Vertical Windows)

The lumen method of sidelighting calculates interior horizontal illuminance at three points, as shown in Figure 20.56, at the 30-in (0.76-m) work-plane on the room-and-window centerline. A vertical window is assumed to extend from 36 in (0.91 m) above the floor to the ceiling. The method accounts for both direct and ground-reflected sunlight and skylight, so both horizontal and vertical illuminances from sun and sky are needed. The steps in the lumen method of sidelighting are presented next.

As mentioned, the incident direct and ground-reflected window illuminance are normally calculated for both a cloudy and a clear day for representative days during the year (various months), as well as for clear or cloudy times during a given day. Thus, the interior illumination due to sidelighting and skylighting can then be examined for effectiveness throughout the year.

Step 1: Incident Direct Sky and Sun Illuminances—The solar altitude and sun-window azimuth angle difference are calculated for the desired latitude, date, and time using Equation 20.89 and Equation 20.92, respectively. Using these two angles, the total illuminance on the window (E_{sw}) can be determined by summing the direct sun illuminance (E_{uw}) and the direct sky illuminance (E_{kw}), each determined from the appropriate graph in Figure 20.57.

Step 2: Incident Ground-Reflected Illuminance—The sun illuminance on the ground (E_{ug}), plus the overcast or clear sky illuminance (E_{kg}) on the ground, make up the total horizontal illuminance on the ground surface (E_{sg}). A fraction of the ground surface illuminance is then considered diffusely reflected onto the vertical window surface (E_{gw}), where gw indicates from the ground to the window.

The horizontal ground illuminances can be determined using Figure 20.58, where the clear sky plus sun case and the overcast sky case are functions of solar altitude. The fractions of the ground illuminance diffusely reflected onto the window depends on the reflectivity (ρ) of the ground surface (see Table 20.11.) and the window-to-ground surface geometry.

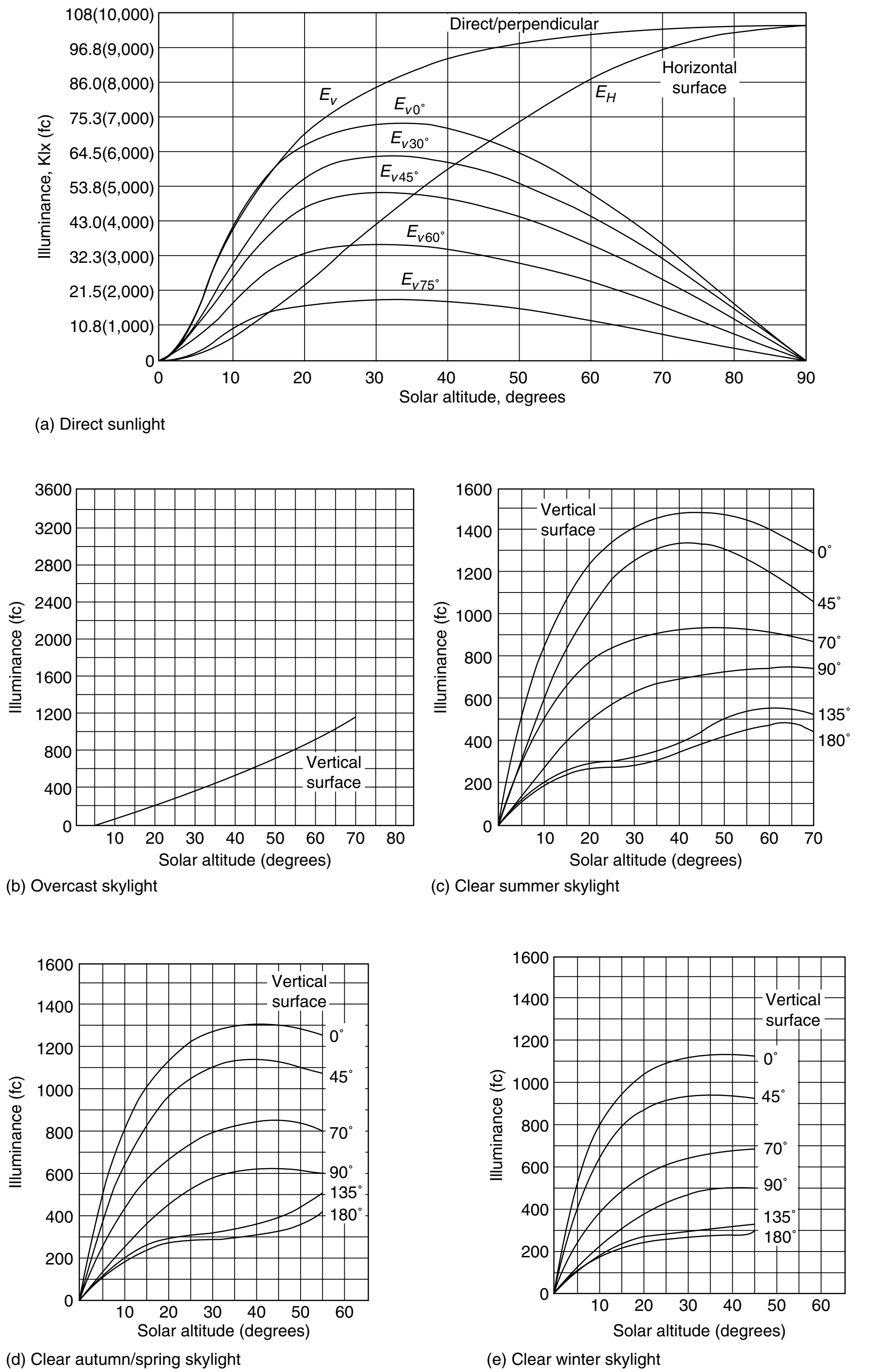


FIGURE 20.57 Vertical illuminance from (a) direct sunlight and (b-c) skylight, for various sun-window azimuth angle differences. (From IES (Illumination Engineering Society), *Lighting Handbook, Applications Volume*. Illumination Engineering Society, New York, 1987.)

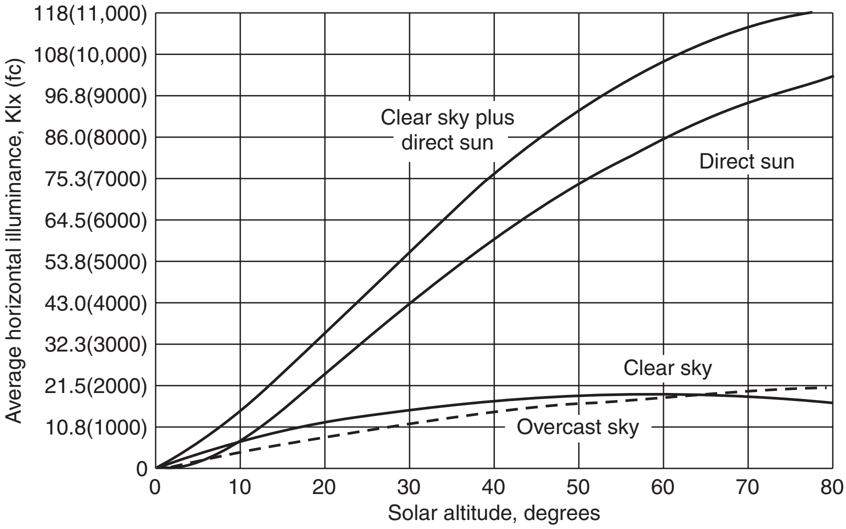


FIGURE 20.58 Horizontal illuminance for overcast sky, clear sky, direct sun, and clear sky plus direct sun. (From Murdoch, J. B., *Illumination Engineering — From Edison's Lamp to the Laser*, Mac Millan, New York, 1985.)

If the ground surface is considered uniformly reflective from the window outward to the horizon, then the illuminance on the window from ground reflection is

$$E_{gw} = \frac{\rho E_{sg}}{2}. \tag{20.93}$$

A more complicated ground-reflection case is illustrated in Figure 20.59, where multiple "strips" of differently reflecting ground are handled using the angles to the window where a strip's illuminance on a window is calculated,

$$E_{gw(strip)} = \frac{\rho_{strip} E_{sg}}{2}(\cos\theta_1 - \cos\theta_2). \tag{20.94}$$

And the total reflected onto the window is the sum of the strip illuminances:

$$E_{gw} = \frac{E_{sg}}{2}[\rho_1(\cos 0 - \cos\theta_1) + \rho_2(\cos\theta_1 - \cos\theta_2) + \cdots + \rho_n(\cos\theta_{n-1} - \cos 90)]. \tag{20.95}$$

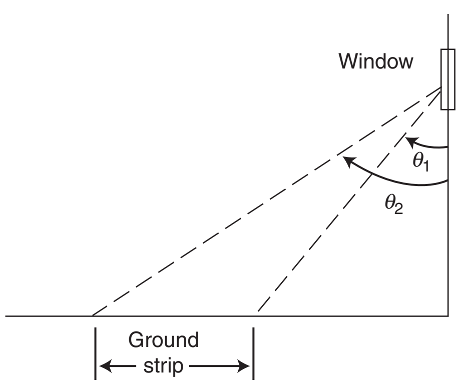


FIGURE 20.59 Geometry for ground "strips". (From Murdoch, J. B., *Illumination Engineering — From Edison's Lamp to the Laser*, Mac Millan, New York, 1985.)

TABLE 20.12 Glass Transmittances

Glass	Thickness (in.)	τ (%)
Clear	1/3	89
Clear	3/16	88
Clear	1/4	87
Clear	5/16	86
Grey	1/8	61
Grey	3/16	51
Grey	1/4	44
Grey	5/16	35
Bronze	1/8	68
Bronze	3/16	59
Bronze	1/4	52
Bronze	5/16	44
Thermopane	1/8	80
Thermopane	3/16	79
Thermopane	1/4	77

Source: From Murdoch, J. B., *Illumination Engineering—From Edison's Lamp to the Laser*, Macmillan, New York, 1985.

Step 3: Luminous Flux Entering Space—The direct sky–sun and ground-reflected luminous fluxes entering the building are attenuated by the transmissivity of the window. Table 20.12 presents the transmittance fraction (τ) of several window glasses. The fluxes entering the space are calculated from the total sun–sky and the ground-reflected illuminances by using the area of the glass, A_w:

$$\begin{aligned} \phi_{sw} &= E_{sw}\tau A_w, \\ \phi_{gw} &= E_{gw}\tau A_w, \end{aligned} \tag{20.96}$$

Step 4: Light Loss Factor—The light loss factor (K_m) accounts for the attenuation of luminous flux due to dirt on the window (WDD, window dirt depreciation) and on the room surfaces (RSDD, room surface dirt depreciation). WDD depends on how often the window is cleaned, but a 6-month average for offices is 0.83 and for factories is 0.71 (Murdoch 1985).

The RSDD is a more complex calculation involving time between cleanings, the direct-indirect flux distribution, and room proportions. However, for rooms cleaned regularly, RSDD is around 0.94 and for once-a-year-cleaned dirty rooms, the RSDD would be around 0.84.

The light loss factor is the product of the preceding two fractions:

$$K_m = (\text{WDD})(\text{RSDD}). \tag{20.97}$$

Step 5: Work-Plane Illuminances—As discussed earlier, Figure 20.56 illustrates the location of the work-plane illuminances determined with this lumen method of sidelighting. The three illuminances (max, mid, min) are determined using two coefficients of utilization, the *C* factor and *K* factor. The *C* factor depends on room length and width and wall reflectance. The *K* factor depends on ceiling-floor height, room width, and wall reflectance. Table 20.13 presents *C* and *K* values for the three cases of incoming fluxes: sun plus clear sky, overcast sky, and ground-reflected. Assumed ceiling and floor reflectances are given for this case with no window controls (shades, blinds, overhangs, etc.). These further window control complexities can be found in Libbey-Owens-Ford Company (1976); IES (1987), and others. A reflectance of 70% represents light-colored walls, with 30% representing darker walls.

The work-plane max, mid, and min illuminances are each calculated by adding the sun–sky and ground-reflected illuminances, which are given by

$$E_{sp} = \phi_{sw} C_s K_s K_m, \qquad (20.98)$$
$$E_{gp} = \phi_{gw} C_g K_g K_m,$$

where the "sp" and "gp" subscripts refer to the sky-to-work-plane and ground-to-work-plane illuminances.

Example 20.3.9

Determine the clear-sky illuminances for a 30-ft.-long, 30-ft.-wide, 10-ft.-high room with a 20-ft.-long window with a 3-ft. sill. The window faces 10°E of south, the building is at 32°N latitude, and it is January 15 at 2 p.m. The ground cover outside is grass, the glass is l/4-in clear, and the walls are light colored.

Solution. Following the steps in the "sidelighting" method:

Step 1: With $L=32$, $n=15$, $H=(12-14)60/4=-30$,

$$\delta = 23.45 \sin[360(15-81)/365] = -21.3°.$$

Then, Equation 20.89 yields $\alpha=41.7°$, Equation 20.91 yields $\gamma_s=-38.7°$, and Equation 20.92 yields

$$\Phi = |-38.7-(+10)| = 48.7°.$$

From Figure 20.57 with $\alpha=41.7°$ and $\Phi=48.7°$:

(a) For clear sky (winter, no sun): $E_{kw}=875$ fc.
(b) For direct sun: $E_{uw}=4{,}100$ fc.
(c) Total clear sky plus direct: $E_{sw}=4{,}975$ fc.

(Note: A high E_{uw} value probably indicates a glare situation!)

Step 2: Horizontal illuminances from Figure 20.58: $E_{sg}=4007$ fc.

Then Equation 20.93 yields, with $\rho_{grass}=0.06$, $E_{gw}=222$ fc.

Step 3: From Equation 20.96, with $\tau=0.87$ and $A_w=140$ ft.^{2},

$$\Phi_{sw} = 4975(0.87)(140) = 605,955 \text{ lm},$$

$$\Phi_{gw} = 222(0.87)(140) = 27,040 \text{ lm}.$$

Step 4: For a clean office room:

$$K_m = (0.83)(0.94) = 0.78.$$

Step 5: From Table 20.13, for 30′ width, 30′ length, 10′ ceiling, and wall reflectivity 70%,

(a) Clear sky:

$$C_{s,\,max} = 0.0137;\ K_{s,\,max} = 0.125,$$

$$C_{s,\,mid} = 0.0062;\ K_{s,\,mid} = 0.110,$$

$$C_{s,\,min} = 0.0047; K_{s,\,min} = 0.107.$$

TABLE 20.13a *C* and *K* Factors for No Window Controls for Overcast Sky

Illumination by Overcast Sky

C: Coefficient of Utilization

Room Length (ft.)		20		30		40	
Wall Reflectance (%)		70	30	70	30	70	30
	Room Width (ft.)						
Max	20	0.0276	0.0251	0.0191	0.0173	0.0143	0.0137
	30	0.0272	0.0248	0.0188	0.0172	0.0137	0.0131
	40	0.0269	0.0246	0.0182	0.0171	0.0133	0.0130
Mid	20	0.0159	0.0177	0.0101	0.0087	0.0081	0.0071
	30	0.0058	0.0050	0.0054	0.0040	0.0034	0.0033
	40	0.0039	0.0027	0.0030	0.0023	0.0022	0.0019
Min	20	0.0087	0.0053	0.0063	0.0043	0.0050	0.0037
	30	0.0032	0.0019	0.0029	0.0017	0.0020	0.0014
	40	0.0019	0.0009	0.0016	0.0009	0.0012	0.0008

K: Coefficient of Utilization

Ceiling Height (ft.)		8		10		12		14	
Wall Reflectance (%)		70	30	70	30	70	30	70	30
	Room Width (ft.)								
Max	20	0.125	0.129	0.121	0.123	0.111	0.111	0.0991	0.0973
	30	0.122	0.131	0.122	0.121	0.111	0.111	0.0945	0.0973
	40	0.145	0.133	0.131	0.126	0.111	0.111	0.0973	0.0982
Mid	20	0.0908	0.0982	0.107	0.115	0.111	0.111	0.105	0.122
	30	0.156	0.102	0.0939	0.113	0.111	0.111	0.121	0.134
	40	0.106	0.0948	0.123	0.107	0.111	0.111	0.135	0.127
Min	20	0.0908	0.102	0.0951	0.114	0.111	0.111	0.118	0.134
	30	0.0924	0.119	0.101	0.114	0.111	0.111	0.125	0.126
	40	0.111	0.0926	0.125	0.109	0.111	0.111	0.133	0.130

Source: From IES, 1979.

TABLE 20.13b *C* and *K* Factors for No Window Controls for Clear Sky

Illumination by Clear Sky

C: Coefficient of Utilization

Room Length (ft.)		20		30		40	
Wall Reflectance (%)		70	30	70	30	70	30
	Room Width (ft.)						
Max	20	0.0206	0.0173	0.0143	0.0123	0.0110	0.0098
	30	0.0203	0.00.173	0.0137	0.0120	0.0098	0.0092
	40	0.0200	0.0168	0.0131	0.0119	0.0096	0.0091
Mid	20	0.0153	0.0104	0.0100	0.0079	0.0083	0.0067
	30	0.0082	0.0054	0.0062	0.0043	0.0046	0.0037
	40	0.0052	0.0032	0.0040	0.0028	0.0029	0.0023
Min	20	0.0106	0.0060	0.0079	0.0049	0.0067	0.0043
	30	0.0054	0.0028	0.0047	0.0023	0.0032	0.0021
	40	0.0031	0.0014	0.0027	0.0013	0.0021	0.0012

K: Coefficient of Utilization

Ceiling Height (ft.)		8		10		12		14	
Wall Reflectance (%)		70	30	70	30	70	30	70	30
	Room Width (ft.)								
Max	20	0.145	0.155	0.129	0.132	0.111	0.111	0.101	0.0982
	30	0.141	0.149	0.125	0.130	0.111	0.111	0.0954	0.101
	40	0.157	0.157	0.135	0.134	0.111	0.111	0.0964	0.0991
Mid	20	0.110	0.128	0.116	0.126	0.111	0.111	0.103	0.108
	30	0.106	0.125	0.110	0.129	0.111	0.111	0.112	0.120
	40	0.117	0.118	0.122	0.118	0.111	0.111	0.123	0.122
Min	20	0.105	0.129	0.112	0.130	0.111	0.111	0.111	0.116
	30	0.0994	0.144	0.107	0.126	0.111	0.111	0.107	0.124
	40	0.119	0.116	0.130	0.118	0.111	0.111	0.120	0.118

Source: From IES, 1979.

TABLE 20.13c *C* and *K* Factors for No Window Controls for Ground Illumination (Ceiling Reflectance, 80%; Floor Reflectance, 30%)

Ground Illumination

C: Coefficient of Utilization							
Room Length (ft.)		20		30		40	
Wall Reflectance (%)		70	30	70	30	70	30
	Room Width (ft.)						
Max	20	0.0147	0.0112	0.0102	0.0088	0.0081	0.0071
	30	0.0141	0.0012	0.0098	0.0088	0.0077	0.0070
	40	0.0137	0.0112	0.0093	0.0086	0.0072	0.0069
Mid	20	0.0128	0.0090	0.0094	0.0071	0.0073	0.0060
	30	0.0083	0.0057	0.0062	0.0048	0.0050	0.0041
	40	0.0055	0.0037	0.0044	0.0033	0.0042	0.0026
Min	20	0.0106	0.0071	0.0082	0.0054	0.0067	0.0044
	30	0.0051	0.0026	0.0041	0.0023	0.0033	0.0021
	40	0.0029	0.0018	0.0026	0.0012	0.0022	0.0011

K: Coefficient of Utilization									
Ceiling Height (ft.)		8		10		12		14	
Wall Reflectance (%)		70	30	70	30	70	30	70	30
	Room Width (ft.)								
Max	20	0.124	0.206	0.140	0.135	0.111	0.111	0.0909	0.0859
	30	0.182	0.188	0.140	0.143	0.111	0.111	0.0918	0.0878
	40	0.124	0.182	0.140	0.142	0.111	0.111	0.0936	0.0879
Mid	20	0.123	0.145	0.122	0.129	0.111	0.111	0.100	0.0945
	30	0.0966	0.104	0.107	0.112	0.111	0.111	0.110	0.105
	40	0.0790	0.0786	0.0999	0.106	0.111	0.111	0.118	0.118
Min	20	0.0994	0.108	0.110	0.114	0.111	0.111	0.107	0.104
	30	0.0816	0.0822	0.0984	0.105	0.111	0.111	0.121	0.116
	40	0.0700	0.0656	0.0946	0.0986	0.111	0.111	0.125	0.132

Source: From IES, 1979.

(b) Ground reflected:

$$C_{g,\ max} = 0.0098;\ K_{g,\ max} = 0.140.$$

$$C_{g,\ mid} = 0.0062;\ K_{,\ mid} = 0.107.$$

$$C_{g,\ min} = 0.0041;\ K_{g,\ min} = 0.0984.$$

Then using Equation 20.98,

$$E_{sp,\ max} = 605,955(0.0137)(0.125)(0.78) = 809\ \text{fc},$$

$$E_{sp,\ mid} = 605,955(0.0062)(0.110)(0.78) = 322\ \text{fc},$$

$$E_{sp,min} = 605,955(0.0047)(0.107)(0.78) = 238\ \text{fc},$$

$$E_{gp,\ max} = 27,040(0.0098)(0.140)(0.78) = 29\ \text{fc},$$

$$E_{gp,\ mid} = 27,040(0.0062)(0.107)(0.78) = 14\ \text{fc},$$

$$\mathrm{E}_{gp,\ min} = 27,040(0.0041)(0.984)(0.78) = 9\ \text{fc}.$$

Thus,

$$E_{max} = 838\ \text{fc},$$

$$E_{mid} = 336\ \text{fc},$$

$$E_{min} = 247\ \text{fc}.$$

20.3.5.4.2 Lumen Method of Skylighting

The lumen method of skylighting calculates the average illuminance at the interior work-plane provided by horizontal skylights mounted on the roof. The procedure for skylighting is generally the same as that described above for sidelighting. As with windows, the illuminance from both overcast sky and clear sky plus sun cases are determined for specific days in different seasons and for different times of the day, and a judgment is then made as to the number and size of skylights and any controls needed.

The procedure is presented in four steps: (1) finding the horizontal illuminance on the outside of the skylight; (2) calculating the effective transmittance through the skylight and its well; (3) figuring the interior space light loss factor and the utilization coefficient; and finally, (4) calculating illuminance on the work-plane.

Step 1: Horizontal Sky and Sun Illuminances—The horizontal illuminance value for an overcast sky or a clear sky plus sun situation can be determined from Figure 20.58 knowing only the solar altitude.

Step 2: Net Skylight Transmittance—The transmittance of the skylight is determined by the transmittance of the skylight cover(s), the reflective efficiency of the skylight well, the net-to-gross skylight area, and the transmittance of any light-control devices (lenses, louvers, etc.).

The transmittance for several flat-sheet plastic materials used in skylight domes is presented in Table 20.14. To get the effective dome transmittance (T_D) from the flat-plate transmittance (T_F) value (AAMA 1977), use

TABLE 20.14 Flat-Plate Plastic Material Transmittance for Skylights

Type	Thickness (in.)	Transmittance (%)
Transparent	1/8–3/16	92
Dense translucent	1/3	32
Dense translucent	3/16	24
Medium translucent	1/8	56
Medium translucent	3/16	52
Light translucent	1/8	72
Light translucent	3/16	68

Source: From Murdoch, J. B., *Illumination Engineering—From Edison's Lamp to the Laser*, Macmillan, New York, 1985.

$$T_{\mathrm{D}} = 1.25T_{\mathrm{F}}(1.18 - 0.416T_{\mathrm{F}}). \tag{20.99}$$

If a double-domed skylight is used, then the single-dome transmittances are combined as follows (Pierson 1962):

$$T_{\mathrm{D}} = \frac{T_{\mathrm{D_1}} T_{\mathrm{D_2}}}{T_{\mathrm{D_1}} T_{\mathrm{D_2}} - T_{\mathrm{D_1}} T_{\mathrm{D_2}}} \tag{20.100}$$

If the diffuse and direct transmittances for solar radiation are available for the skylight glazing material, it is possible to follow this procedure and determine diffuse and direct dome transmittances separately. However, this difference is usually not a significant factor in the overall calculations.

The efficiency of the skylight well (N_{w}) is the fraction of the luminous flux from the dome that enters the room from the well. The well index (WI) is a geometric index (height, h; length, l; width, w) given by

$$\mathrm{WI} = \frac{h(w + l)}{2wl}, \tag{20.101}$$

and WI is used with the well–wall reflectance value in Figure 20.60 to determine well efficiency, N_{w}.

With T_{D} and N_{w} determined, the net skylight transmittance for the skylight and well is given by:

$$T_{\mathrm{n}} = T_{\mathrm{D}} N_{\mathrm{w}} R_{\mathrm{A}} T_{\mathrm{C}}, \tag{20.102}$$

where R_{A} is the ratio of net to gross skylight areas and T_{C} is the transmittance of any light-controlling devices.

Step 3: Light Loss Factor and Utilization Coefficient—The light loss factor (K_{m}) is again defined as the product of the room surface dirt depreciation (RSDD) and the skylight direct depreciation (SDD) fractions, similar to Equation 20.97. Following the reasoning for the sidelighting case, the RSDD value for clean rooms is around 0.94 and 0.84 for dirty rooms. Without specific data indicating otherwise, the SDD fraction is often taken as 0.75 for office buildings and 0.65 for industrial areas.

The fraction of the luminous flux on the skylight that reaches the work-plane (K_{u}) is the product of the net transmittance (T_{n}) and the room coefficient of utilization (RCU). Dietz et al. (1981) developed RCU equations for office and warehouse interiors with ceiling, wall, and floor reflectances of 75%, 50%, and 30%, and 50%, 30%, and 20%, respectively.

$$\mathrm{RCU} = \frac{1}{1 + A(\mathrm{RCR})^{B}} \text{ if RCR} < 8, \tag{20.103}$$

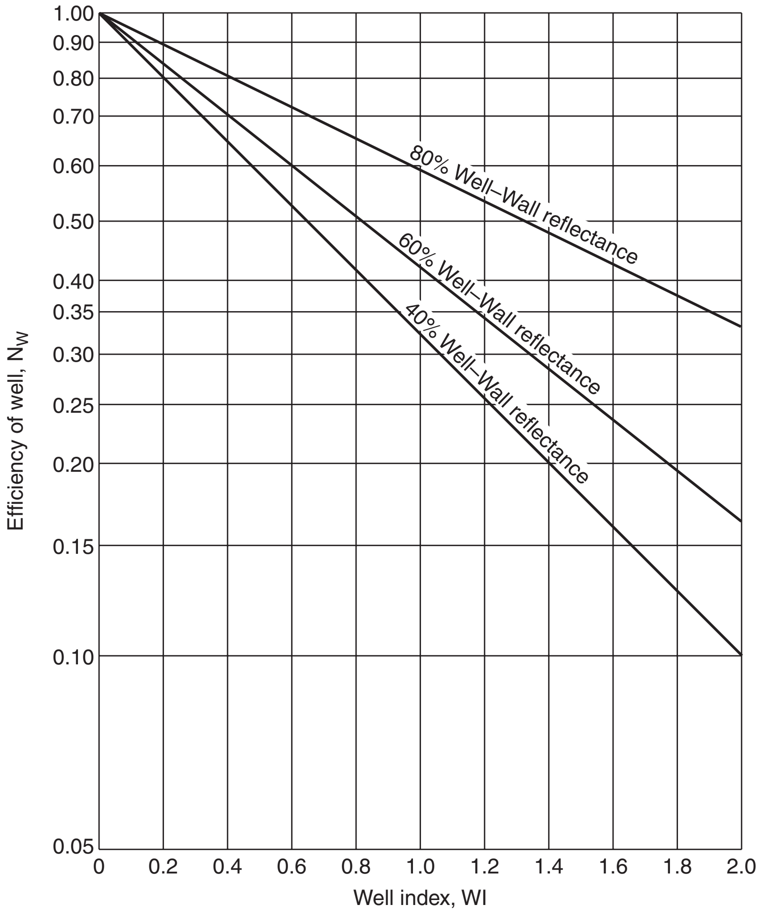


FIGURE 20.60 Efficiency of well versus well index. (From IES (Illumination Engineering Society), *Lighting Handbook, Applications Volume*. Illumination Engineering Society, New York, 1987.)

where A is 0.0288 and B is 1.560 for offices, and A is 0.0995 and B is 1.087 for warehouses. Room cavity ratio (RCR) is given by

$$RCR = \frac{5h_c(l + w)}{lw}, \tag{20.104}$$

where h_c is the ceiling height above the work-plane and l and w are the room length and width, respectively.

The RCU is then multiplied by the previously determined T_n to give the fraction of the external luminous flux passing through the skylight and incident on the workplace:

$$K_u = T_n(\text{RCU}). \tag{20.105}$$

Step 4: Work-Plane Illuminance—The illuminance at the work-plane (E_{TWP}) is given by

$$E_{TWP} = E_H\left(\frac{A_T}{A_{WP}}\right)K_u K_m, \tag{20.106}$$

where E_H is the horizontal overcast or clear sky plus sun illuminance from Step 1, A_T is total gross area of the skylights (number of skylights times skylight gross area), and A_{WP} is the work-plane area (generally room length times width). Note that in Equation 20.106, it is also possible to fix the E_{TWP} at some desired value and determine the required skylight aread.

Rules of thumb for skylight placement for uniform illumination include 4%–8% of roof area and spacing less than 1.5 times ceiling-to-work-plane distance between skylights (Murdoch 1984).

Example 20.3.10

Determine the work-plane "clear sky plus sun" illuminance for a $30 \times 30 \times 10$ ft.^{3} office with 75% ceiling, 50% wall, and 30% floor reflectance with four 4×4 ft.^{2} double-domed skylights at 2:00 p.m. on January 15 at 32° latitude. The skylight well is 1 ft. deep at with 60% reflectance walls, and the outer- and inner-dome flat-plastic transmittances are 0.85 and 0.45, respectively. The net skylight area is 90%.

Solution. Follow the four steps in the lumen method for skylighting.

Step 1: Use Figure 20.58 with the solar altitude of 41.7° (calculated from Equation 20.93) for the clear sky plus sun curve to get horizontal illuminance:

$$E_H = 7400 \text{ fc.}$$

Step 2: Use Equation 20.99 to determine domed transmittances from the flat plate plastic transmittances given,

$$T_{D1} = 1.25(0.85)[1.18 - 0.416(0.85)] = 0.89,$$

$$T_{D_2}(T_F = 0.45) = 0.56,$$

and use Equation 20.100 to get total dome transmittance from the individual dome transmittances:

$$T_D = \frac{(0.89)(0.56)}{(0.89) + (0.56) - (0.89)(0.56)} = 0.52.$$

To determine well efficiency, use WI = 0.25 from Equation 20.101 with 60% wall reflectance in Figure 20.60 to give $N_w = 0.80$. With $R_A = 0.90$, use Equation 20.102 to calculate net transmittance:

$$T_n = (0.52)(0.80)(0.90)(1.0) = 0.37.$$

Step 3: The light loss factor is assumed to be from "typical" values in Equation 20.97: $K_m = (0.75)(0.94) = 0.70$. The room utilization coefficient is determined using Equation 20.103 and Equation 20.104:

$$\text{RCR} = \frac{5(7.5)(30 + 30)}{(30)(30)} = 2.5,$$

$$\text{RCU} = [1 + 0.0288(2.5)^{1.560}]^{-1} = 0.89.$$

Equation 20.104 yields $K_u = (0.37)(0.89) = 0.33$.

Step 4: The work-plane illuminance is calculated by substituting the above values into Equation 20.106:

$$E_{TWP} = 7400 \left[\frac{4(16)}{30(30)}\right] 0.33(0.70),$$
$$E_{TWP} = 122 \text{ fc.}$$

20.3.5.5 Daylighting Controls and Economics

The economic benefit of daylighting is directly tied to the reduction in lighting electrical energy operating costs. Also, lower cooling-system operating costs are possible due to the reduction in heating caused by the reduced electrical lighting load. The reduction in lighting and cooling system electrical power during peak demand periods could also beneficially affect demand charges.

The reduction of the design cooling load through the use of daylighting can also lead to the reduction of installed or first-cost cooling system dollars. Normally, economics dictate that an automatic lighting control system must take advantage of the reduced lighting/cooling effect, and the control system cost minus any cooling system cost savings should be expressed as a "net" first cost. A payback time for the lighting control system ("net" or not) can be calculated from the ratio of first costs to yearly operating savings. In some cases, these paybacks for daylighting controls have been found to be in the range of 1–5 years for office building spaces (Rundquist 1991).

Controls, both aperture and lighting, directly affect the efficacy of the daylighting system. As shown in Figure 20.61, aperture controls can be architectural (overhangs, light shelves, etc.) and/or window shading devices (blinds, automated louvers, etc.). The aperture controls generally moderate the sunlight entering the space to maximize/minimize solar thermal gain, permit the proper amount of light for visibility, and prevent glare and beam radiation onto the workplace. Photosensor control of electric lighting allows the dimming (or shutting off) of the lights in proportion to the amount of available daylighting illuminance.

In most cases, increasing the solar gain for daylighting purposes, with daylighting controls, saves more in electrical lighting energy and the cooling energy associated with the lighting than is incurred with the added solar gain (Rundquist 1991). In determining the annual energy savings total from daylighting, ES_T, the annual lighting energy saved from daylighting, ES_L, is added with the reduction in cooling system energy, ΔES_C, and with the negative of the heating system energy increase ΔES_H:

$$ES_T = ES_L + \Delta ES_C - \Delta ES_H. \tag{20.107}$$

A simple approach to estimating the heating and cooling energy changes associated with the lighting energy reduction is by using the fraction of the year associated with the cooling or heating season (f_C, f_H) and the seasonal COP of the cooling or heating equipment. Thus, Equation 20.107 can be

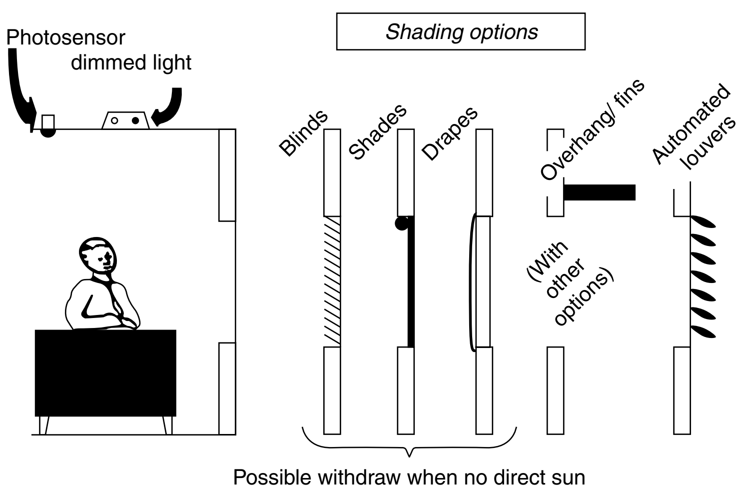


FIGURE 20.61 Daylighting system controls. (From Rundquist, R. A., Daylighting controls: Orphan of HVAC design. *ASHRAE Journal*, 11 (November), 30–340, 1991.)

expressed as

$$\mathrm{ES_T} = \mathrm{ES_L} + \frac{f_\mathrm{C}\mathrm{ES_L}}{\mathrm{COP_c}} - \frac{f_\mathrm{H}\mathrm{ES_L}}{\mathrm{COP_H}},$$

$$\mathrm{ES_T} = \mathrm{ES_L}\left(1 + \frac{f_\mathrm{C}}{\mathrm{COP_C}} - \frac{f_\mathrm{H}}{\mathrm{COP_H}}\right). \tag{20.108}$$

It should be noted that the increased solar gain due to daylighting has not been included here but would reduce summer savings and increase winter savings. If it is assumed that the increased wintertime daylighting solar gain approximately offsets the reduced lighting heat gain, then the last term in Equation 20.108 becomes negligible.

To determine the annual lighting energy saved ($\mathrm{ES_L}$), calculations using the lumen method described earlier should be performed on a monthly basis for both clear and overcast days for the space under investigation. Monthly weather data for the site would then be used to prorate clear and overcast lighting energy demands monthly. Subtracting the calculated (controlled) daylighting illuminance from the design illuminance leaves the supplementary lighting needed, which determines the lighting energy required.

This procedure has been computerized and includes many details of controls, daylighting methods, weather, and heating and cooling load calculations. ASHRAE (1989) lists many of the methods and simulation techniques currently used with daylighting and its associated energy effects.

Example 20.3.11

A 30×20 ft.^{2} space has a photosensor dimmer control with installed lighting density of 2.0 W/ft.^{2}. The required workplace illuminance is 60 fc and the available daylighting illuminance is calculated as 40 fc on the summer peak afternoon. Determine the effect on the cooling system (adapted from Rundquist 1991).

Solution. The lighting power reduction is (2.0 W/ft.^{2}) (30×20) ft.$^2\times$(40 fc/60 fc)$=$800 W. The space cooling load would also be reduced by this amount (assuming CLF$=$1.0):

$$\frac{800\ \mathrm{W} \times 3.413\ \mathrm{Btu\ h/W}}{12{,}000\ \mathrm{Btu\ h/tn.}} = 0.23\ \mathrm{tn.}$$

Assuming 1.5 tn. nominally installed for 600 ft.^{2} of space at \$2200/tn., the 0.23-tn. reduction is "worth" $0.23\times$\$2200/tn.$=$\$506. The lighting controls cost about \$1/ft.^{2} of controlled area, so the net installed first cost is

$$\text{Net first cost} = \$\ 600\ \text{controls} - \$\ 500\ \text{A/C savings} = \$\ 100.$$

Assuming the day-to-monthly-to-annual illuminance calculations gave a 30% reduction in annual lighting, the associated operating savings can be determined. Lighting energy savings are

$$\mathrm{ES_L} = 0.30 \times 2.0\ \mathrm{W/ft.^2} \times 600\ \mathrm{ft.^2} \times 2500\ \mathrm{h/year} = 900\ \mathrm{kWh}.$$

Using Equation 20.108 to also include cooling energy saved due to lighting reduction (with $\mathrm{COP_c}=$ 2.5, $f_c=0.5$, and neglecting heating) gives

$$\mathrm{ES_T} = 900(1 + 0.5/2.5 - 0) = 1080\ \mathrm{kWh}.$$

At \$0.10 per kWh, the operating costs savings are \$0.10/kWh$\times$1080 kWh$=$\$108/year.

Thus, the simple payback is approximately 1 year (100/108) for the "net" situation, and a little over 5.5 years (600/108) against the "controls" cost alone. It should also be noted that the 800 W lighting electrical reduction at peak hours, with an associated cooling energy reduction of 800 W/2.5 COP = 320 W, provides a peak demand reduction for the space of 1.1 kW, which can be used as a “first-cost savings” to offset control system costs.

Glossary

Active system: A system employing a forced (pump or fan) convection heat transfer fluid flow.

Daylighting: The use of the sun’s radiant energy for illumination of a building’s interior space.

Hybrid system: A system with parallel passive and active flow systems or one using forced convection flow to distribute from thermal storage.

Illuminance: The density of luminous flux incident on a unit surface. Illuminance is calculated by dividing the luminous flux (in lumens) by the surface area (m^2, $ft.^2$). Units are lux (lx) (lumens/m^2) in SI and footcandles (fc) (lumens/$ft.^2$) in English systems.

Luminous flux: The time rate of flow of luminous energy (lumens). A lumen (lm) is the rate which luminous energy from a 1 candela (cd) intensity source is incident on a 1-m^2 surface 1 m from the source.

Passive cooling system: A system using natural energy flows to transfer heat to the environmental sinks (ground, air, and sky).

Passive heating system: A system in which the sun’s radiant energy is converted to heat by absorption in the system, and the heat is distributed by naturally occurring processes.

Sidelighting: Daylighting by light entering through the wall/side of a space.

Skylight: The diffuse solar radiation from a clear or overcast sky, excluding the direct radiation from the sun.

Sunlight: The direct solar radiation from the sun.

Toplighting: Daylighting by light entering through the ceiling area of a space.

References

AAMA (Architectural Aluminum Manufacturers Association). 1977. Publication 1602.1.1977, *Voluntary Standard Procedure for Calculating Skylight Annual Energy Balance.* AAMA, Chicago, IL.

ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers). 1989. Fundamentals. In *ASHRAE Handbook*. ASHRAE, Atlanta, GA.

ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers), 1993. Fundamentals. In *ASHRAE Handbook*. ASHRAE, Atlanta, GA

ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers). 1991. Heating, ventilating, and air-conditioning applications, In *ASHRAE Handbook*. ASHRAE, Atlanta, GA.

ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers), 1995. Heating, ventilating, and air-conditioning applications, In *ASHRAE Handbook*, pp. 279–127. ASHRAE, Atlanta, GA.

Clay, R. A. 2001. Green is good for you. *Monitor on Psychology*, 32, 4, 40–42.

Close, D. J., Dunkle, R. V., and Robeson, K. A. 1968. Design and performance of a thermal storage air conditioning system. *Mechanical and Chemical Engineering Transactions*, MC4, 45.

De Paepe, M. and Janssens, A. 2003. Thermo-hydraulic design of earth-air heat exchanger. *Energy and Buildings*, 35, 389–397.

Dietz, P., Murdoch, J., Pokoski, J., and Boyle, J. 1981. A skylight energy balance analysis procedure. *Journal of the Illuminating Engineering Society*, 11, October, 27–34.

Duffie, J. A. and Beckman, W. A. 1991. *Solar Engineering of Thermal Processes. 2nd Ed.*, Wiley, New York.

Goswami, D. Y. and Biseli, K. M. 1994. Use of underground air tunnels for headting and coooling agricultural residential buildings. Report EES-78, Florida Energy Extension service, University of Florida, Gainesville, Fl. August.

Goswami, D. Y. and Dhaliwal, A. S. 1985. Heat transfer analysis in environmental control using an underground air tunnel. *J. of Solar Energy Eng.*, 107 (May): 141–45.

Goswami, D. Y. and Ileslamlou, S. 1990. Performance analysis of a closed-loop climate control system using undergroung air tunnel. *J. of Solar Energy Eng.* 112 (May): 76–81.

Givoni, B. 1994. *Passive and Low Energy Cooling of Buildings*, Van Nostrand Reinhold, New York.

Hay, H. and Yellott, J. 1969. Natural air conditioning with roof ponds and movable insulation. *ASHRAE Transactions*, 75 (1), 165–177.

Hollmuller, P. and Lachal, B. 2001. Cooling and preheating with buried pipe systems: Monitoring, simulation and economic aspects. *Energy and Buildings*, 33, 509–518.

IES (Illumination Engineering Society). 1987. *Lighting Handbook, Applications Volume*. Illumination Engineering Society, New York.

Kinney, L., McCluney, R., Cler, G., and Hutson, J. 2005. New designs in active daylighting: Good ideas whose time has (finally) come. In *Proceedings of the 2005 Solar World Congress*. August 6–12, 2005, ISES, Orlando, FL.

Krarti, M. and Kreider, J. F. 1996. Analytical model for heat transfer in an underground air tunnel. *Energy Conversion Management*, 37, 10, 1561–1574.

Kusuda, T. and Achenbach, P. R. 1965. Earth temperature and thermal diffusivity at selected stations in the United States. ASHRAE Transactions, 71 (1), 965.

Labs, K. 1981. Regional analysis of ground and above ground climate. Report ORNL/Sub-81/40451/1, Oak Ridge National Laboratory, Oak Ridge, TN.

Larson, R., Vignola, F., and West, R. eds. 1992. *Economics of Solar Energy Technologies*. American Solar Energy Society, Orlando, FL.

Libbey-Owens-Ford Company 1976. *How to Predict Interior Daylight Illumination*. Libbey-Owens-Ford Company, Toledo, OH.

Martin, M. and Berdahl, P. 1984. Characteristics of infrared sky radiation in the United States. *Solar Energy*, 33, 314, 321–336.

Marlatt,W., Murray, C., and Squire, S. 1984. Roof Pond Systems Energy Technology Engineering Center. Report No. ETEC 6, April. Rockwell International, New York.

McCluney, R. 1998. Advanced fenestration daylighting systems. *In International Daylighting Conference '98*, Natural Resources Canada/CETC, Ottawa, Canada.

McQuiston, P. C. and Parker, J. D. 1994. *Heating, Ventilating, and Air Conditioning. 4th Ed.*, Wiley, New York.

Murdoch, J. B. 1985. *Illumination Engineering—From Edison's Lamp to the Laser*. Macmillan, New York.

NCDC (National Climactic Data Center). 1992. *Climatography of the U.S. #81*. NCDC, Asheville, NC.

Olgyay, A. and Olgyay, V. 1967. *Solar Control and Shading Devices*, Princeton University Press, Princeton, NJ.

Pierson, O. 1962. *Acrylics for the Architectural Control of Solar Energy*. Rohm and Haas, Philadelphia, PA.

Robbins, C. L. 1986. *Daylighting—Design and Analysis*, Van Nostrand, New York.

Rundquist, R. A. 1991. Daylighting controls: Orphan of HVAC design. *ASHRAE Journal*, 11 (November), 30–34.

PSDH. 1980. *Passive Solar Design Handbook*. Volume One: Passive Solar Design Concepts, DOE/CS-0127/1, March 1980. Prepared by Total Environmental Action, Inc. (B. Anderson, C. Michal, P. Temple, and D. Lewis); Volume Two: *Passive Solar Design Analysis*, DOE/CS-0127/2, January 1980. Prepared by Los Alamos Scientific Laboratory (J.D. Balcomb, D. Barley, R. McFarland, J. Perry, W. Wray and S. Noll). U.S. Department of Energy, Washington, DC.

PSDH. 1984. *Passive Solar Design Handbook*. Part One: Total Environmental Action, Inc., Part Two: Los Alamos Scientific Laboratory, Part Three: Los Alamos National Laboratory. Van Nostrand Reinhold, New York.

For Further Information

The most complete basic reference for passive system heating design is still the Passive Solar Design Handbook, all three parts. Solar Today magazine, published by the American Solar Energy Society, is the most available source for current practice designs and economics, as well as a source for passive system equipment suppliers.The ASHRAE Handbook of Fundamentals is a general introduction to passive cooling techniques and calculations, with an emphasis on evaporative cooling. Passive Solar Buildings and Passive Cooling, both published by MIT Press, contain a large variety of techniques and details concerning passive system designs and economics. All the major building energy simulation codes (DOE-2, EnergyPlus, TRNSYS, TSB13, etc.) now include passive heating and cooling technologies.The Illlumination Engineering Society's Lighting Handbook presents the basis for and details of daylighting and artificial lighting design techniques. However, most texts on illumination present simplified format day lighting procedures. Currently used daylighting computer programs include various versions of Lumen Micro, Lightscape, and Radiance Passive Solar Design Strategies: Guidelines for Homebuilders (Passive Solar Industries Council, Washington, DC, 1989) presents a user-friendly approach to passive solar design.

20.4 Solar Cooling

D. Yogi Goswami and Sanjay Vijayaraghavan

In some ways, solar energy is better suited to space cooling and refrigeration than to space heating. The seasonal variation of solar energy is extremely well-suited to the space-cooling requirements of buildings. The principal factors affecting the temperature in a building are the average quantity of radiation received and the environmental air temperature. Because the warmest seasons of the year correspond to periods of high insolation, solar energy is most available when comfort cooling is most needed. Moreover, the efficiency of solar collectors increases with increasing insolation and increasing environmental temperature. Consequently, in the summer, the amount of energy delivered per unit surface area of collector can be larger than that in winter.

Solar cooling using various refrigeration cycles is technically feasible and has been demonstrated several times over past few decades. However, application of these systems has not become popular due to the unfavorable economics. The most widely used methods applied to solar cooling and air conditioning are vapor compression cycles, absorption-cooling cycles, and desiccant cooling. The vapor compression refrigeration cycle is probably the most widely used refrigeration cycle. The vapor compression refrigeration cycle requires energy input into the compressor which may be provided as electricity from a photovoltaic system or as mechanical energy from a solar driven heat engine. Referring to Figure 20.62, the compressor raises the pressure of the refrigerant, which also increases the temperature. The compressed high-temperature refrigerant vapor then transfers heat to the ambient environment in the condenser, where it condenses to a high-pressure liquid at a temperature close to the environmental temperature. The liquid refrigerant is then passed through the expansion valve where the pressure is suddenly reduced, resulting in a vapor–liquid mixture at a much lower temperature. The low-temperature refrigerant is then used to cool air or water in the evaporator where the liquid refrigerant evaporates by absorbing heat from the medium being cooled. The cycle is completed by the vapor returning to the compressor. If water is cooled in the evaporator, the device is usually called a *chiller*. The chilled water could then be used to cool air in a building.

In an absorption system, the refrigerant is evaporated or distilled from a less volatile absorbent, the vapor is condensed in a water- or air-cooled condenser, and the resulting liquid is passed through a pressure-reducing valve to the cooling section (evaporator) of the unit. The refrigerant from the evaporator flows into the absorber, where it is reabsorbed in the stripped absorbing liquid and

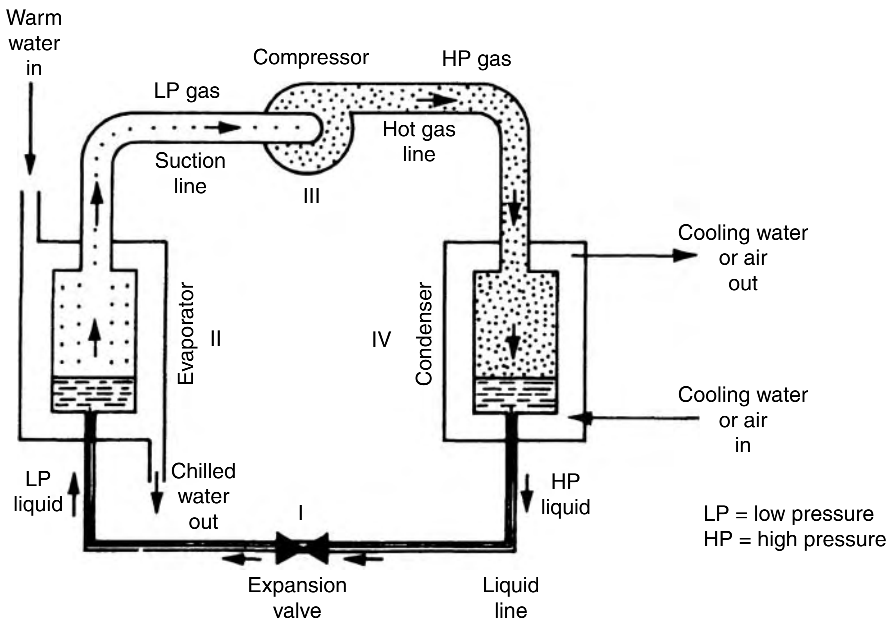


FIGURE 20.62 A schematic diagram showing a typical vapor compression refrigeration cycle.

pumped back to the heated generator. The heat required to evaporate the refrigerant in the generator can be supplied directly from solar energy as shown in Figure 20.63.

In humid climates, removal of moisture from the air represents a major portion of the air-conditioning load. In such climates, desiccant systems can be used for dehumidification, in which solar energy can provide most of the energy requirements. There are several passive space cooling techniques that are

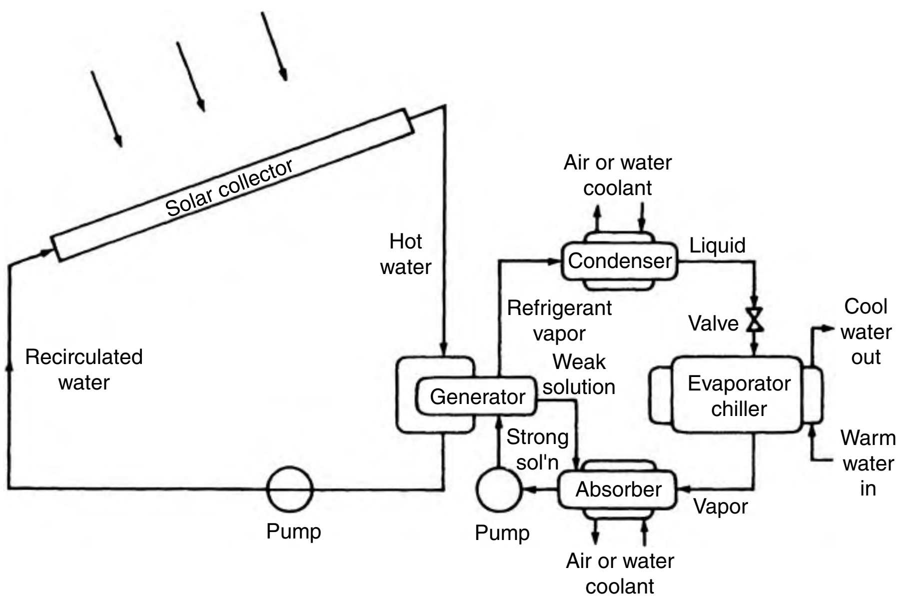


FIGURE 20.63 Figure shows the basic arrangement of a solar driven absorption cycle.

described elsewhere. The present section covers the active solar cooling techniques based on vapor compression and vapor-absorption refrigeration cycles and desiccant humidification.

20.4.1 Vapor Compression Cycle

The principle of operation of a vapor compression refrigeration cycle can be illustrated conveniently with the aid of a pressure–enthalpy diagram as shown in Figure 20.64. The ordinate is the pressure of the refrigerant in N/m^2 absolute, and the abscissa its enthalpy in kJ/kg. The roman numerals in Figure 20.64 correspond to the physical locations in the schematic diagram of Figure 20.62. Process I is a throttling process in which hot liquid refrigerant at the condensing pressure p_c passes through the expansion valve, where its pressure is reduced to the evaporator pressure, p_e. This is an isenthalpic (constant enthalpy) process in which the temperature of the refrigerant decreases. In this process, some vapor is produced and the state of the mixture of liquid refrigerant and vapor entering the evaporator is shown by point A. Because the expansion process is isenthalpic, the following relation holds:

$$h_{ve}f + h_{le}(1-f) = h_{lc},$$

where f is the fraction of mass in vapor state, subscripts "v" and "l" refer to vapor and liquid states, respectively, and "c" and "e" refer to states corresponding to condenser and evaporator pressures, respectively.

Process II represents the vaporization of the remaining liquid. This is the process during which heat is removed from the chiller. Thus, the specific refrigeration effect (per kilogram of refrigerant flow), q_r, is

$$q_r = h_{ve} - h_{lc}, \text{ in kJ/kg.}$$

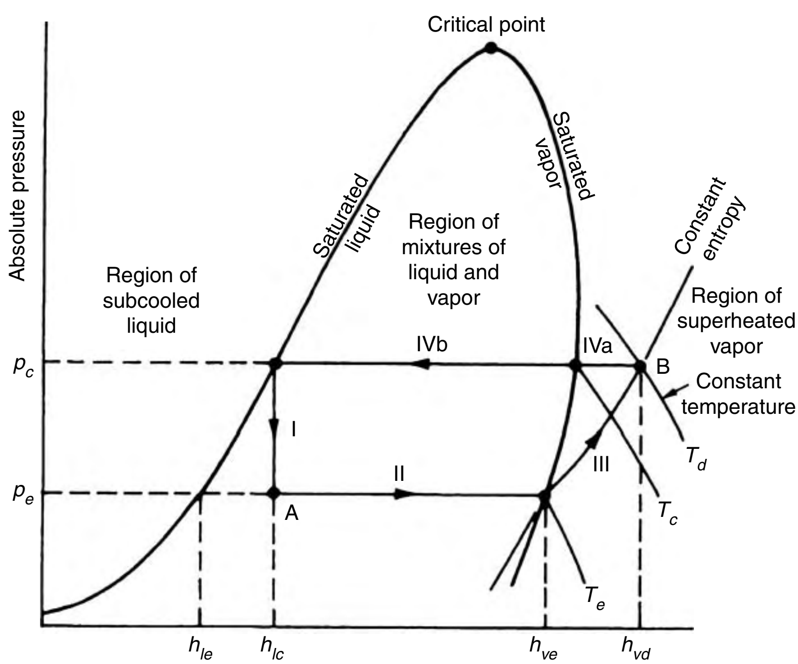


FIGURE 20.64 The thermodynamic state processes of the vapor compression refrigeration cycle shown on a pressure–enthalpy (p–h) diagram.

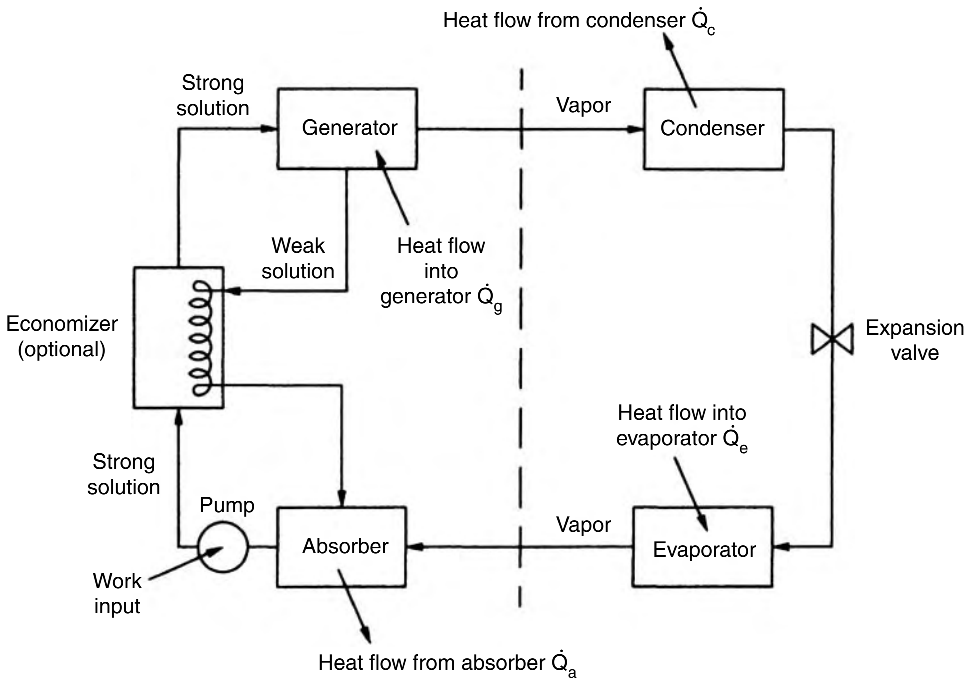


FIGURE 20.65 A typical absorption refrigeration cycle.

In the United States, it is still common practice to measure refrigeration in terms of tons. One ton is the amount of cooling produced if 1 ton of ice is melted over a period of 24 h. One ton of cooling is equivalent to 3.516 kW, or 12,000 Btu/h.

Process III in Figure 20.64 represents the compression of refrigerant from pressure p_e to pressure p_c. The process requires work input from an external source, which may be obtained from a solar-driven expander turbine or a solar electrical system. In general, if the heated vapor leaving the compressor is at the condition represented by point B in Figure 20.64, the work of compression is

$$W_c = \dot{m}_r(h_{vd} - h_{ve}).$$

In an idealized cycle analysis, the compression process is usually assumed to be isentropic.

Process IV represents the condensation of the refrigerant. Actually, sensible heat is first removed in subprocess IVa as the vapor is cooled at constant pressure from T_d to T_c and latent heat is removed at the condensation temperature T_c corresponding to the saturation pressure, p_c, in the condenser. The heat transfer rate in the condenser, $\dot{Q}_c$, is

$$\dot{Q}_c = \dot{m}_r(h_{vd} - h_{lc}).$$

This heat must be rejected to the environment, either to cooling water or to the atmosphere if no water is available.

The overall performance of a refrigeration machine is usually expressed as the coefficient of performance that is defined as the ratio of the heat transferred in the evaporator, $\dot{Q}_r$, to the shaft work supplied by the compressor:

$$\text{COP} = \frac{\dot{Q}_r}{W_c} = \frac{h_{ve} - h_{lc}}{h_{vd} - h_{ve}}.$$

20.4.2 Absorption Air Conditioning

Absorption air conditioning is compatible with solar energy because a large fraction of the energy required is thermal energy at temperatures that solar collectors can easily provide. Low- and medium-temperature solar collectors have been used to drive several absorption air conditioning systems (Macriss and Zawacki 1989; Mathur 1989; Manrique 1991; Chinnappa and Wijeysundera 1992; Siddiqui 1993; Thornbloom and Nimmo 1994; Hewett 1995). Although single-effect absorption refrigeration systems can be run using solar heated hot water at 80°C, higher temperatures are preferred for better refrigeration cycle performance. The key difference between a conventional gas-fired absorption chiller and one used for solar applications is the larger heat transfer area required to make the cycle work using the lower driving temperatures available in solar applications. Figure 20.65 shows a schematic of an absorption refrigeration system. Absorption refrigeration differs from vapor compression air conditioning only in the method of compressing the refrigerant (left of the dashed line in Figure 20.65). In absorption air conditioning systems, the pressurization is accomplished by first dissolving the refrigerant in a liquid (the absorbent) in the absorber section, then pumping the solution to a high pressure with a liquid pump. The low-boiling refrigerant is then driven from solution by the addition of heat in the generator. By this means, the refrigerant vapor is compressed without the large input of high-grade shaft work that the vapor compression cycle demands.

The effective performance of an absorption cycle depends on the two materials that comprise the refrigerant–absorbent pair. Desirable characteristics for the refrigerant–absorbent pair are as follows:

1. Absence of a solid-phase sorbent
2. A refrigerant more volatile than the absorbent so that separation from the absorbent occurs easily in the generator
3. An absorbent that has a strong affinity for the refrigerant under conditions in which absorption takes place
4. A high degree of stability for long-term operations
5. Nontoxic and nonflammable fluids for residential applications; this requirement is less critical in industrial refrigeration
6. A refrigerant that has a large latent heat so that the circulation rate can be kept low
7. A low fluid viscosity that improves heat and mass transfer and reduces pumping power
8. Fluids that do not have long-term environmental effects

Lithium bromide–water (LiBr–H_2O) and ammonia–water (NH_3–H_2O) are the two pairs that meet most of the requirements and have been used commercially in several applications. In the LiBr–H_2O system, water is the refrigerant and LiBr is the absorbent, whereas in the ammonia–water system, ammonia is the refrigerant and water is the absorbent. Because the LiBr–H_2O system has a high-volatility ratio, it can operate at lower pressures and therefore, at the lower generator temperatures achievable by flat-plate collectors. A disadvantage of this system is that LiBr has a tendency to crystallize in the stream returning from the generator. Crystallization is avoided by careful system design and by the use of additives. Furthermore, because the refrigerant is water, the system evaporator cannot be operated at or below the freezing point of water. Therefore, the LiBr–H_2O system is operated at evaporator temperatures of 5°C or higher. Using a mixture of LiBr with some other salt as the absorbent can overcome the crystallization problem. The ammonia–water system has the advantage that the evaporator can be maintained at very low temperatures. However, for temperatures much below 0°C, water vapor must be removed from ammonia as much as possible to prevent ice crystals from forming. This requires a rectifying column after the boiler. Also, ammonia is a safety code group B2 fluid (ASHRAE Standard 34-1992) that restricts its use indoors (ASHRAE 1997). Consequently, the ammonia–water system cannot use a direct expansion (DX) evaporator. Other refrigerant–absorbent pairs include (Macriss and Zawacki 1989):

- Ammonia–salt
- Methylamine–salt

- Alcohol–salt
- Ammonia–organic solvent
- Sulfur dioxide–organic solvent
- Halogenated hydrocarbons–organic solvent
- Water–alkali nitrate
- Ammonia–water–salt

If the pump work is neglected, the COP of an absorption air conditioner can be calculated from Figure 20.65:

$$\mathrm{COP} = \frac{\text{cooling effect}}{\text{heat input}} = \frac{\dot{Q}_e}{\dot{Q}_g}.$$

The COP values for absorption air conditioning range from 0.5 for a small, single-stage unit to 0.85 for a double-stage, steam-fired unit. Another figure of merit for absorption systems is the ratio of the cooling effect to work supplied to the system (circulation pumps, fans, etc.).

Explicit procedures for the mechanical and thermal design as well as the sizing of the heat exchangers are presented in standard heat transfer texts. In large commercial units, it may be possible to use higher concentrations of LiBr, operate at a higher absorber temperature, and thus save on heat-exchanger costs. In a solar-driven unit, this approach would require concentrator-type or high-efficiency flat-plate solar collectors.

20.4.2.1 Ammonia–Water Systems

The main difference between an ammonia–water system and a water–lithium bromide system is that a small amount of absorbent (water) also evaporates along with the refrigerant (ammonia) in the vapor generator. Therefore, ammonia–water systems use a rectifier (also called a *dephlagmator*) after the generator to condense as much water vapor out of the vapor mixture as possible. Figure 20.66 shows a schematic of an NH_3–H_2O absorption refrigeration system. Because ammonia has a much lower boiling point than water, a very high fraction of ammonia and a very small fraction of water are boiled off in the

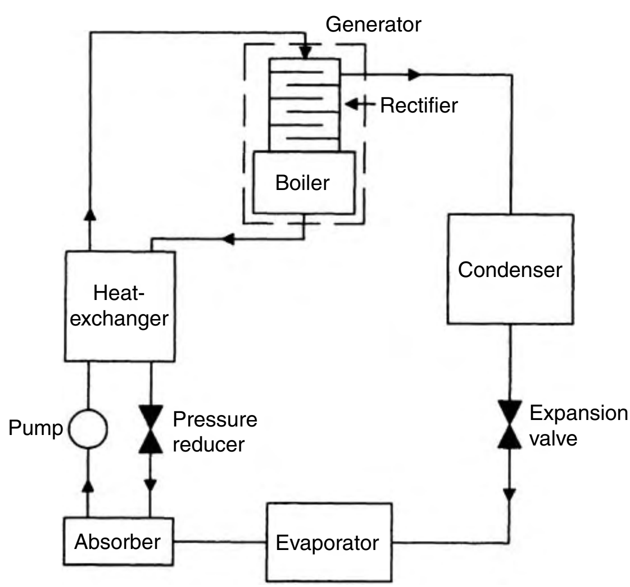


FIGURE 20.66 A diagram showing the arrangement of components for an ammonia-absorption cycle.

boiler. The vapor is cooled as it rises in the rectifier by the countercurrent flow of the strong NH_3–H_2O solution from the absorber; therefore, some moisture is condensed. The weak ammonia–water solution from the boiler goes through a pressure-reducing valve to the absorber, where it absorbs the ammonia vapor from the evaporator. The high-pressure ammonia from the rectifier is condensed by rejecting heat to the atmosphere. It may be further subcooled before expanding in a throttle valve. The two-phase low-temperature ammonia from the throttle valve provides refrigeration in the evaporator. The vapor from the evaporator is recombined with the weak ammonia solution in the absorber. Operating pressures are primarily controlled by the ambient air temperature for an air-cooled condenser, the evaporator temperature, and the concentration of the ammonia solution in the absorber.

20.4.2.2 Multieffect Systems

A major price component in designing absorption systems is the solar collector field. To improve the economics of a solar absorption system, the efficiency of the solar collectors must be improved in addition to the COP of the absorption system, thus reducing the required collector area. A single-effect absorption system has a typical efficiency of around 0.7. For higher COP, double-effect systems are used. Double-effect systems typically operate at higher temperatures than single-effect systems requiring higher concentration solar collectors to provide the heat input. An example of a typical double-effect lithium bromide system is shown in Figure 20.67.

A double-effect lithium bromide cycle has two generators at two different pressure levels. Vapor is generated using the solar heat source in the first generator. This vapor is condensed in the second generator and the heat of condensation is used to produce more vapor (this arrangement is known as a *condenser-coupled system*). Thus the double-effect absorption cycle is a triple pressure cycle. A double-effect ammonia–water system is configured slightly differently (absorber-coupled), but still uses the same

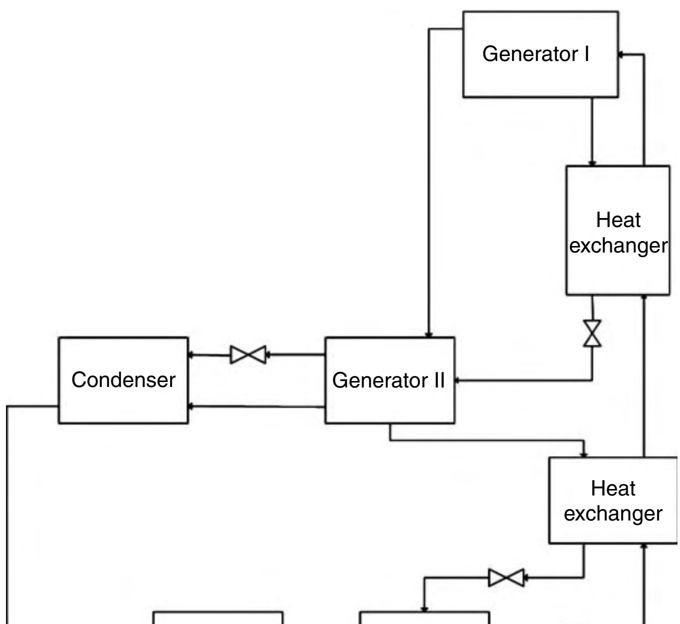


FIGURE 20.67 A schematic diagram of a condenser coupled double-effect absorption cooling cycle. (Adapted from Wahlig, M., In *Active Solar Systems*, Vol. 6 of Solar Heat Technologies: Fundamentals and Applications, MIT Press, Cambridge, MA, 747–854, 1988. With permission.)

principle of internal heat recovery to produce more refrigerant vapor than is possible in a single-effect system. However, it requires a much higher driving temperature (140°C or higher) to operate efficiently.

20.4.3 Solar Desiccant Dehumidification

In hot and humid regions of the world experiencing significant latent cooling demand, solar energy may be used for dehumidification using liquid or solid desiccants. Rangarajan, Shirley, and Raustad (1989) compared a number of strategies for ventilation air conditioning for Miami, Florida, and found that a conventional vapor compression system could not even meet the increased ventilation requirements of ASHRAE Standard 62-1989. By pretreating the ventilation air with a desiccant system, proper indoor humidity conditions could be maintained and significant electrical energy could be saved. A number of researchers have shown that a combination of a solar desiccant and a vapor compression system can save from 15 to 80% of the electrical energy requirements in commercial applications such as supermarkets (Meckler 1988; Meckler, Parent, and Pesaran 1993; Meckler 1994; Meckler 1995; Spears and Judge 1997; Oberg and Goswami 1998a, 1998b).

In a desiccant air conditioning system, moisture is removed from the air by bringing it in contact with the desiccant, followed by sensible cooling of the air by a vapor compression cooling system, vapor-absorption cooling system, or evaporative cooling system. The driving force for the process is the water vapor pressure. When the vapor pressure in air is higher than on the desiccant surface, moisture is transferred from the air to the desiccant until an equilibrium is reached (see Figure 20.67). To regenerate the desiccant for reuse, the desiccant is heated, which increases the water vapor pressure on its surface. If air with lower vapor pressure is brought into contact with this desiccant, the moisture passes from the desiccant to the air (Figure 20.68). Two types of desiccants are used: solids such as silica gel and lithium chloride, or liquids such as salt solutions and glycols.

The two solid desiccant materials that have been used in solar systems are silica gel and molecular sieves, a selective absorber. Figure 20.69 shows the equilibrium absorption capacity of several substances. Note that molecular sieves has the highest capacity up to 30% humidity, and silica gel is optimal between 30 and 75%, the typical humidity range for buildings.

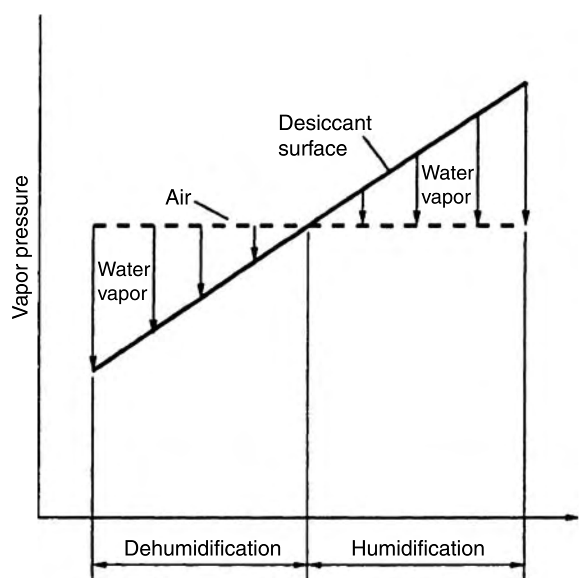


FIGURE 20.68 Vapor pressure vs. temperature and water content for desiccant and air.

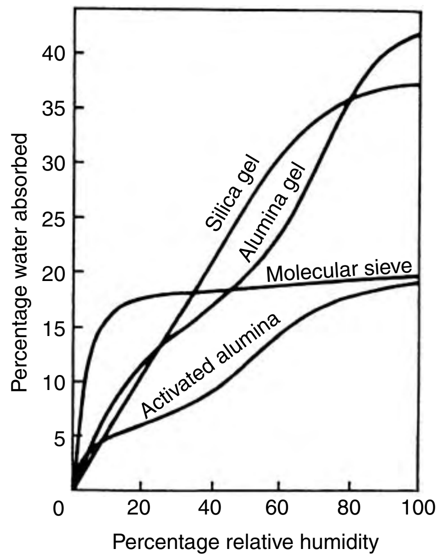


FIGURE 20.69 Equilibrium capacities of common water absorbents.

Figure 20.70 is a schematic diagram of a desiccant cooling ventilation cycle (also known as the Pennington cycle), which achieves both dehumidification and cooling. The desiccant bed is normally a rotary wheel of a honeycomb-type substrate impregnated with the desiccant. As the air passes through the rotating wheel, it is dehumidified while its temperature increases (processes 1 and 2) due to the latent heat of condensation. Simultaneously, a hot air stream passes through the opposite side of the rotating wheel, which removes moisture from the wheel. The hot and dry air at state 2 are cooled in a heat exchanger wheel to condition 3 and further cooled by evaporative cooling to condition 4. Air at condition 3 may be further cooled by vapor compression or vapor absorption systems instead of evaporative cooling. The return air from the conditioned space is cooled by evaporative cooling (processes 5 and 6), which in turn cools the heat exchanger wheel. This air is then heated to condition 7. Using solar heat, it is further heated to condition 8 before going through the desiccant wheel to regenerate the desiccant. A number of researchers have studied this cycle, or an innovative variation of it, and have found thermal COPs in the range of 0.5–2.58 (Pesaran, Penney, and Czandema 1992).

20.4.4 Liquid-Desiccant Cooling System

Liquid desiccants offer a number of advantages over solid desiccants. The ability to pump a liquid desiccant makes it possible to use solar energy for regeneration more efficiently. It also allows several small dehumidifiers to be connected to a single regeneration unit. Because a liquid desiccant does not require simultaneous regeneration, the liquid may be stored for regeneration later when solar heat is available. A major disadvantage is that the vapor pressure of the desiccant itself may be enough to cause some desiccant vapors to mix with the air. This disadvantage, however, may be overcome by proper choice of the desiccant material.

A schematic of a liquid desiccant system is shown in Figure 20.71. Air is brought into contact with concentrated desiccant in a countercurrent flow in a dehumidifier. The dehumidifier may be a spray column or packed bed. The packings provide a very large area for heat and mass transfer between the air and the desiccant. After dehumidification, the air is sensibly cooled before entering the conditioned space. The dilute desiccant exiting the dehumidifier is regenerated by heating and exposing it to a countercurrent flow of a moisture-scavenging air stream. Liquid desiccants commonly used are aqueous solutions of lithium bromide, lithium chloride, calcium chloride, mixtures of these solutions, and

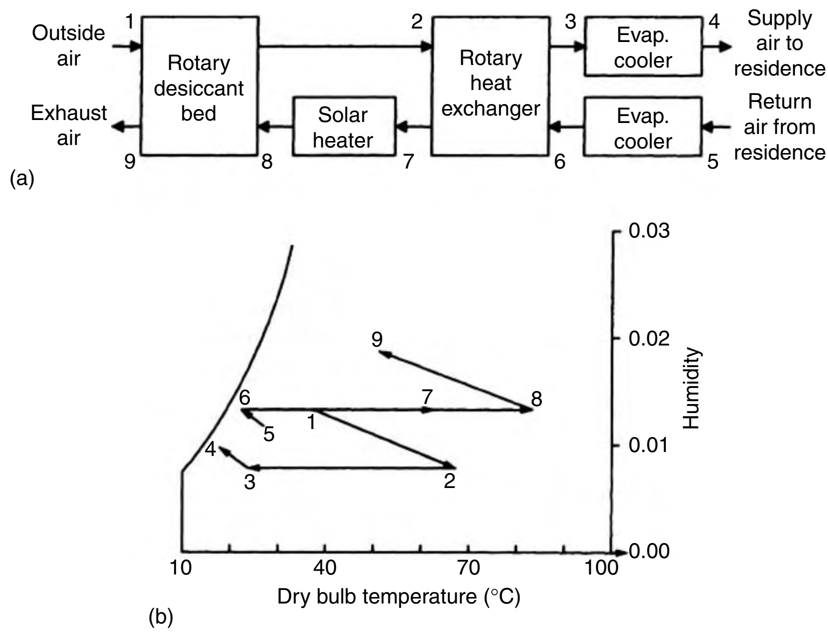


FIGURE 20.70 Schematic of a desiccant cooling ventilation cycle, (a) schematic of air flow, (b) the process on a psychrometric chart.

triethylene glycol (TEG). (See Oberg and Goswami 1998b). Vapor pressures of these common desiccants are shown in Figure 20.72 as a function of concentration and temperature, based on a number of references (Cyprus Foote Mineral Company; Dow Chemical Company 1992, 1996; Ertas, Anderson, and Kiris 1992; Zaytsev and Aseyev 1992). Other physical properties important in the selection of desiccant

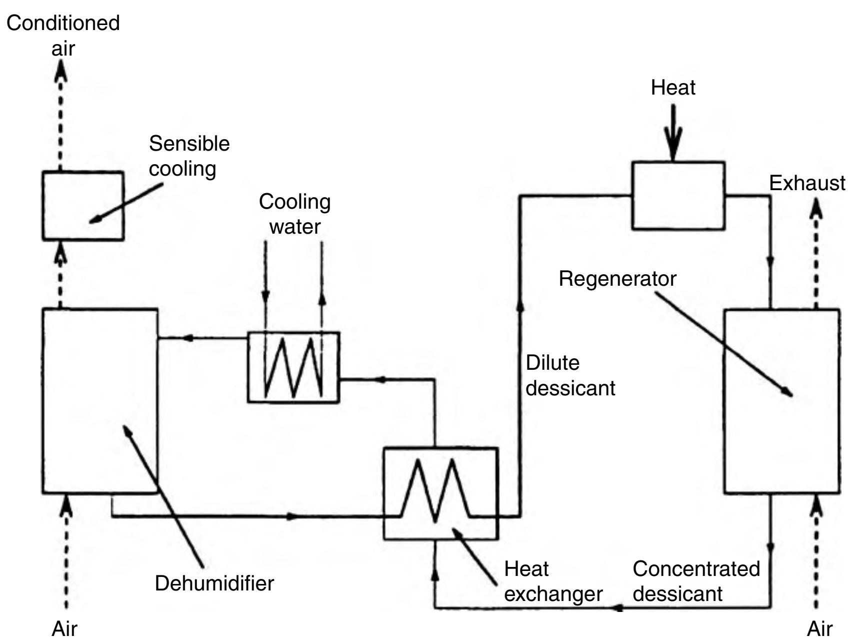


FIGURE 20.71 A conceptual liquid-desiccant cooling system.

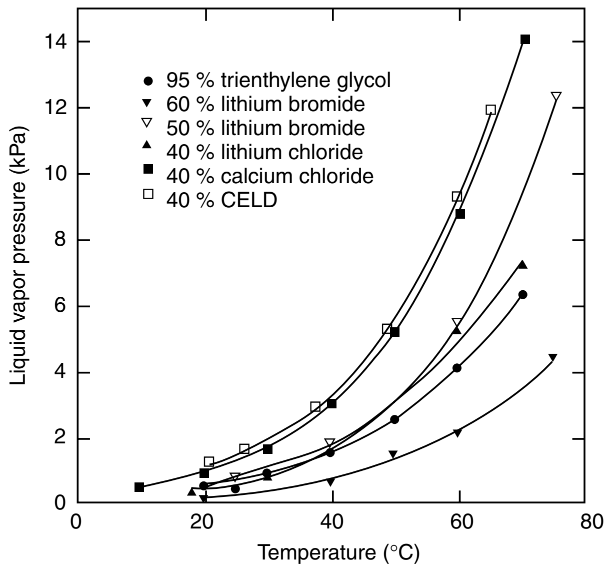


FIGURE 20.72 Vapor pressures of liquid desiccants.

materials are listed in Table 20.15. Although salt solutions and TEG have similar vapor pressures, the salt solutions are corrosive and have higher surface tension. The disadvantage of TEG is that it requires higher pumping power because of its higher viscosity.

Oberg and Goswami (1998b) have presented an in-depth review of liquid-desiccant cooling systems. Based on an extensive numerical modeling and on experimental studies, they have presented correlations for the performance of a packed-bed liquid-desiccant dehumidifier and a regenerator.

The performance of a packed-bed dehumidifier or a regenerator may be represented by a humidity effectiveness, ε_y, defined as the ratio of the actual change in humidity of the air to the maximum possible for the operating conditions (Ullah, Kettleborough, and Gandhidasan 1988; Chung 1989; Khan 1994):

$$\varepsilon_y = \frac{Y_{in} - Y_{out}}{Y_{in} - Y_{eq}},$$

where Y_{in} and Y_{out} are the humidity ratios of the air inlet and outlet, respectively, and Y_{eq} is the humidity ratio in equilibrium with the desiccant solution at the local temperature and concentration (Figure 20.73).

In addition to the humidity effectiveness, an enthalpy effectiveness, ε_H, is also used as a performance parameter (Khan 1994; Kettleborough and Waugaman 1995):

$$\varepsilon_H = \frac{H_{a,in} - H_{a,out}}{H_{a,in} - H_{a,eq}},$$

where $H_{a,in}$, $H_{a,out}$, and $H_{a,eq}$ are the enthalpies of the air at the inlet and outlet, and in equilibrium with the desiccant, respectively.

Oberg and Goswami (1998a) found the following correlation for ε_y and ε_H:

$$\varepsilon_y, \varepsilon_H = 1 - C_1 (L/G)^a (H_{a,in}/H_{L,out})^b (aZ)^c,$$

TABLE 20.15 Physical Properties of Liquid Desiccants at 25°C

Desiccant	Density, $\rho \times 10^{-3}$ (kg/m^3)	Viscosity, $\mu \times 10^3$ (Ns/m^2)	Surface Tension, $\gamma \times 10^3$ (N/m)	Specific Heat, c_p (kJ/kg °C)	Reference
95% by weight triethylene glycol	1.1	28	46	2.3	Thornbloom and Nimmo (1996)
55% by weight lithium bromide	1.6	6	89	2.1	Gordon and Rabl (1986); Oberg and Goswami (1998a, 1998b)
40% calcium chloride	1.4	7	93	2.5	Gordon and Rabl (1986); Siddiqui (1993); Spears and Judge (1997)
40% by weight lithium chloride	1.2	9	96	2.5	Gordon and Rabl (1986)
40% by weight CELD	1.3	5	—	—	Cyprus Foote Mineral Company

Source: From Oberg, V. and Goswami, D. Y., *Advances in Solar Energy*, 12, 431–470, 1998.

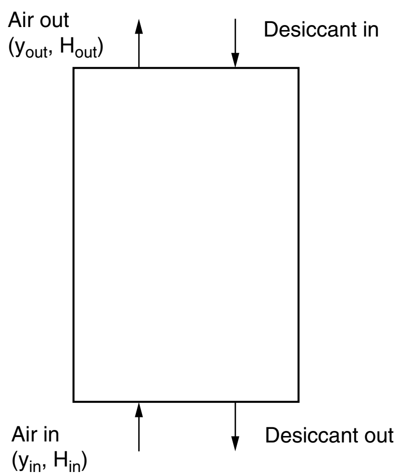


FIGURE 20.73 Exchange of humidity and moisture between desiccant and air in the tower.

TABLE 20.16 Constants for Performance Correlations

	C_1	B	k_1	M_1	k_2	m_2
ε_y	48.345	−0.751	0.396	−1.573	0.033	−0.906
ε_H	3.766	−0.528	0.289	−1.116	0.004	−0.365

where

$$a = k_1 \frac{\gamma_L}{\gamma_c} + m_1,$$

$$c = k_2 \frac{\gamma_L}{\gamma_c} + m_2.$$

Here, C_1, b_1, k_1, m_1, and m_2 are constants listed in Table 20.16. L and G are the liquid and air mass-flow rates, respectively; a is the packing surface area per unit volume for heat and mass transfer in m^2/m^3; Z is the tower height in meters; γ_L is the surface tension of the liquid desiccant; and γ_c is the critical surface tension for the packing material.

Although liquid-desiccant cooling systems are not an off-the-shelf variety currently on the market, there are a number of examples of their use, especially in hybrid combinations with conventional vapor compression systems. In a hybrid system, the liquid desiccant would remove moisture from the air, allowing the vapor compression system to be downsized. Hybrid systems are especially useful in their ability to maintain comfort conditions in hot and humid climates, where conventional high-efficiency systems usually have trouble maintaining low humidity. Mago and Goswami (2003) and Mago (2003a, 2003b) did a simple cost analysis of a hybrid system for a residential building and a supermarket. They found that a house that typically requires a 5-tn. conventional air conditioning unit in Florida could use a hybrid system consisting of a desiccant tower of height 1.1 m, and a vapor compression system of 2 tn.

The total cost of the conventional system was estimated at \$4000, whereas the hybrid system would be \$3250 (\$1000 for the desiccant system + \$2250 for a 2-tn. vapor compression system). A solar system for regeneration of the desiccant was estimated at \$6500, making the total costs of a solar hybrid liquid-desiccant system at \$9780. They estimated that based on the electrical savings, a simple payback period would be 9.5 years. The payback period is high mainly because the solar regeneration system is not utilized throughout the year. For applications where the system can be used year-around, the payback period would be much shorter.

For a commercial application in a supermarket where the hybrid desiccant system and the solar regeneration system would be used throughout the year, the payback period would be less than two years (Mago and Goswami 2003; Mago 2003a). In this application, Mago also estimated that a conventional 13-tn. system, at an estimated cost of \$9300, would be replaced by a hybrid system consisting of a 2-m-high desiccant tower at a cost of \$1500 and a 5-tn. vapor compression system at an estimated cost of \$4000. A solar regeneration system for this desiccant system was estimated at \$9300. The two-year payback period confirmed the estimate of the U.S. Department of Energy (1996) that a simple payback is typically less than five years.

References

ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers). 1997. Fundamentals. In *ASHRAE Handbook*, ASHRAE, Atlanta, GA.

Chinnappa, J. C. V., and Wijeysundera, N. E. 1992. Simulation of solar-powered ammonia-water integrated hybrid cooling system. *Journal of Solar Energy Engineering*, 114, 125–127.

Chung, T. -W. 1989. Predictions of the moisture removal efficiencies for packed-bed dehumidification systems. In *Solar engineering—1989, Proceedings of the 11th Annual ASME Solar Energy Conference*, 371–377, ASMF, New York.

Cyprus Foote Mineral Company. Technical data on lithium bromide and lithium chloride. Bulletins 145 and 151, Cyprus Foote Mineral Company, Kings Mountain, NC.

Dow Chemical Company. 1996. *Calcium chloride handbook*, Dow Chemical Company, Midland, MI.

Dow Chemical Company. 1992. *A guide to glycols*, Dow Chemical Company, Midland, MI.

Ertas, A., Anderson, E. E. and Kiris, I. 1992. Properties of a new liquid desiccant solution—Lithium chloride and calcium chloride mixture. *Solar Energy*, 49, 205–212.

Hewett, R. 1995. Solar absorption cooling: An innovative use of solar energy. In *AlChE Symposium Series, No. 306*, Vol. 91, AICHE, New Year.

Kettleborough, C. F., and Waugaman, D. G. 1995. An alternative desiccant cooling cycle. *Journal of Solar Energy Engineering*, 117, 251–255.

Khan, A. Y. 1994. Sensitivity analysis and component modeling of a packed-type liquid desiccant system at partial load operating conditions. *International Journal of Energy Research*, 18, 643–655.

Macriss, R. A., and Zawacki, T. S. 1989. Absorption fluid data survey: 1989 update. Oak Ridge National Laboratory Report, ORNL/Sub84-47989/4.

Mago, P. J. 2003a. Sistema Hibrido de Enfriamiento como alternative al sistema convencional de aire acondicionado. *Revista SABER*, 15(1), 39–44.

Mago, P. J. 2003b. Analisis economico de la utilización de un sistema hibrido enfriamientote liquido secante en aplicaciones residenciales y comerciales. *Revista SABER*, 15(1), 45–50.

Mago, P. J., and Goswami, D. Y. 2003. Study of the performance of a hybrid liquid desiccant system using lithium chloride. *ASME Journal of Solar Energy Engineering*, 125(1), 129–131.

Manrique, J. A. 1991. Thermal performance of ammonia-water refrigeration system. *International Communications in Heat and Mass Transfer*, 19(6), 779–789.

Mathur, G. D. 1989. Solar-operated absorption coolers. *Heating/Piping/Air Conditioning*. (November). 61, 11, 103–108.

Meckler, M. 1988. Off-peak desiccant cooling and cogeneration combine to maximize gas utilization. *ASHRAE Transactions*, 94(Part 1), 575–596.

Meckler, H. 1994. Desiccant-assisted air conditioner improves IAQ and comfort. *Heating, Piping & Air Conditioning*, 66(10), 75–84.

Meckler, M. 1995. Desiccant outdoor air preconditioners maximize heat recovery ventilation potentials. *ASHRAE Transactions*, 101(Part 2), 992–1000.

Meckler, M., Parent, Y. O. and Pesaran, A. A. 1993. Evaluation of dehumidifiers with polymeric desiccants. Gas Institute Report, Contract No. 5091-246-2247, Gas Research Institute, Chicago, IL.

Oberg, V., and Goswami, D. Y. 1998a. Experimental study of heat and mass transfer in a packed bed liquid desiccant air dehumidifier. In *Solar engineering,* Morehouse J. H. and Hogan, R. E. ed., ASMF, New York, pp. 155–166.

Oberg, V., and Goswami, D. Y. 1998b. A review of liquid desiccant cooling. *Advances in Solar Energy,* 12, 431–470.

Pesaran, A. A., Penney, T. R. and Czandema, A. W. 1992. Desiccant cooling: State-of-the-art assessment. Report No. NREL/TP-254-4147, National Renewable Energy Laboratory, Golden, CO.

Rangarajan, K., Shirley, III, D. B. and Raustad, R. A. 1989. Cost-effective HVAC technologies to meet ASHRAE Standard 62-1989 in hot and human climates. *ASHRAE Transactions*. (Part 1), 166–182.

Siddiqui, A. M. 1993. Optimum generator temperatures in four absorption cycles using different sources of energy. *Energy Conversion and Management*, 34(4), 251–266

Spears, J. W., and Judge, J. 1997. Gas-fired desiccant system for retail super center. *ASHRAE Journal*, 39, 65–69.

Thornbloom, M. and Nimmo, B. 1994. Modification of the absorption cycle for low generator firing temperatures. Solar Engineering 1994, In *Proceedings of the Joint Solar Energy Engineering Conference ASME* 1994, ASME, New York, 367–372.

Thornbloom, M. and Nimmo, B. 1996. Impact of design parameters on solar open cycle liquid desiccant dehumidification system, In *Solar '96 Proceedings of the 1996 Annual Conference of the American Solar Energy Society,* American Solar Energy Society, Boulder, Colorado, pp. 107–111.

Ullah, M. R., Kettleborough, C. F., and Gandhidasan, P. 1988. Effectiveness of moisture removal for an adiabatic counterflow packed tower absorber operating with CaCl-air contact system. *Journal of Solar Energy Engineering*, 110, 98–101.

US Department of Energy. 1996. Desiccant cooling programs. What's new in building energy research. Report DOE/GO-10096-084, Washington, DC.

Zaytsev, 1.0. and Aseyev, G. G. 1992. *Properties of aqueous solutions of electrolytes*, CRC Press, Boca Raton, FL.

A1

The International System of Units, Fundamental Constants, and Conversion Factors

Nitin Goel
Intel Technology India Pvt. Ltd.

The International system of units (SI) is based on seven base units. Other derived units can be related to these base units through governing equations. The base units with the recommended symbols are listed in Table A1.1. Derived units of interest in solar engineering are given in Table A1.2.

Standard prefixes can be used in the SI system to designate multiples of the basic units and thereby conserve space. The standard prefixes are listed in Table A1.3.

Table A1.4 lists some physical constants that are frequently used in solar engineering, together with their values in the SI system of units.

Conversion factors between the SI and English systems for commonly used quantities are given in Table A1.5.

TABLE A1.1 The Seven SI Base Units

Quantity	Name of Unit	Symbol
Length	Meter	m
Mass	Kilogram	kg
Time	Second	s
Electric current	Ampere	A
Thermodynamic temperature	Kelvin	K
Luminous intensity	Candela	cd
Amount of a substance	Mole	mol

TABLE A1.2 SI Derived Units

Quantity	Name of Unit	Symbol
Acceleration	Meters per second squared	m/s^2
Area	Square meters	m^2
Density	Kilogram per cubic meter	kg/m^3
Dynamic viscosity	Newton-second per square meter	$N\ s/m^2$
Force	Newton ($=1$ kg m/s^2)	N
Frequency	Hertz	Hz
Kinematic viscosity	Square meter per second	m^2/s
Plane angle	Radian	rad
Potential difference	Volt	V
Power	Watt($=1$ J/s)	W
Pressure	Pascal ($=1$ N/m^2)	Pa
Radiant intensity	Watts per steradian	W/sr
Solid angle	Steradian	sr
Specific heat	Joules per kilogram–Kelvin	J/kg K
Thermal conductivity	Watts per meter–Kelvin	W/m K
Velocity	Meters per second	m/s
Volume	Cubic meter	m^3
Work, energy, heat	Joule ($=1$ N/m)	J

TABLE A1.3 English Prefixes

Multiplier	Symbol	Prefix	Multiplier	Multiplier Symbol
10^{12}	T	Tera	10^3	M (thousand)
10^9	G	Giga	10^6	MM (million)
10^6	m	Mega		
10^3	k	Kilo		
10^2	h	Hecto		
10^1	da	Deka		
10^{-1}	d	Deci		
10^{-2}	c	Centi		
10^{-3}	m	Milli		
10^{-6}	μ	Micro		
10^{-9}	n	Nano		
10^{-12}	p	Pico		
10^{-15}	f	Femto		
10^{-18}	a	Atto		

TABLE A1.4 Physical Constants in SI Units

Quantity	Symbol	Value
Avogadro constant	N	6.022169×10^{26} kmol^{-1}
Boltzmann constant	k	1.380622×10^{-23} J/K
First radiation constant	$C_1 = 2\pi h C^2$	3.741844×10^{-16} W m^2
Gas constant	R	8.31434×10^{3} J/kmol K
Planck constant	h	6.626196×10^{-34} J s
Second radiation constant	$C_2 = hc/k$	1.438833×10^{-2} m K
Speed of light in a vacuum	C	2.997925×10^{8} m/s
Stefan–Boltzmann constant	σ	5.66961×10^{-8} W/m^2 K^4

TABLE A1.5 Conversion Factors

Physical Quantity	Symbol	Conversion Factor
Area	A	1 ft.^{2} = 0.0929 m^2
		1 acre = 43,560 ft.^{2} = 4047 m^2
		1 hectare = 10,000 m^2
		1 square mile = 640 acres
Density	ρ	1 lb$_m$/ft.^{3} = 16.018 kg/m^3
Heat, energy, or work	Q or W	l Btu = 1055.1 J
		1 kWh = 3.6 MJ
		1 Therm = 105.506 MJ
		l cal = 4.186 J
		1 ft. lb$_f$ = 1.3558 J
Force	F	1 lb$_f$ = 4.448 N
Heat flow rate, refrigeration	q	1 Btu/h = 0.2931 W
		1 ton (refrigeration) = 3.517 kW
		l Btu/s = 1055.1 W
Heat flux	q/A	1 Btu/h ft.^{2} = 3.1525 W/m^2
Heat-transfer coefficient	h	1 Btu/h ft.^{2} F = 5.678 W/m^2 K
Length	L	1 ft. = 0.3048 m
		1 in. = 2.54 cm
		1 mi = 1.6093 km
Mass	m	1 lb$_m$ = 0.4536 kg
		1 ton = 2240 lbm
		1 tonne (metric) = 1000 kg
Mass flow rate	$\dot{m}$	1 lb$_m$/h = 0.000126 kg/s
Power	$\dot{W}$	1 hp = 745.7 W
		1 kW = 3415 Btu/h
		1 ft. lb$_f$/s = 1.3558 W
		1 Btu/h = 0.293 W
Pressure	p	1 lb$_f$/in.^{2} (psi) = 6894.8 Pa (N/m^2)
		1 in. Hg = 3,386 Pa
		1 atm = 101,325 Pa (N/m^2) = 14.696 psi
Radiation	l	1 langley = 41,860 J/m^2
		1 langley/min = 697.4 W/m^2
Specific heat capacity	c	1 Btu/lb$_m$ °F = 4187 J/kg K
Internal energy or enthalpy	e or h	1 Btu/lb$_m$ = 2326.0 J/kg
		1 cal/g = 4184 J/kg
Temperature	T	$T(°R) = (9/5)T(K)$
		$T(°F) = [T(°C)](9/5) + 32$
		$T(°F) = [T(K) - 273.15](9/5) + 32$
Thermal conductivity	k	1 Btu/h ft. °F = 1.731 W/m K
Thermal resistance	R_{th}	1 h °F/Btu = 1.8958 K/W
Velocity	V	1 ft./s = 0.3048 m/s
		1 mi/h = 0.44703 m/s
Viscosity, dynamic	μ	1 lb$_m$/ft. s = 1.488 N s/m^2
		1 cP = 0.00100 N s/m^2
Viscosity, kinematic	ν	1 ft.^{2}/s = 0.09029 m^2/s
		1 ft.^{2}/h = 2.581 $\times 10^{-5}$ m^2/s
Volume	V	1 ft.^{3} = 0.02832 m^3 = 28.32 L
		1 barrel = 42 gal (U.S.)
		1 gal (U.S. liq.) = 3.785 L
		1 gal (U.K.) = 4.546 L
Volumetric flow rate	$\dot{Q}$	1 ft.^{3}/min (cfia) = 0.000472 m^3/s
		1 gal/min (GPM) = 0.0631 l/s

A2

Solar Radiation Data

Nitin Goel
Intel Technology India Pvt. Ltd.

The altitude and azimuth of the sun are given by

$$\sin \alpha = \sin L \sin \delta_s + \cos \phi \cos \delta_s \cos h_s \qquad (1)$$

and

$$\sin a_s = -\cos \delta_s \sin h_s / \cos \alpha \qquad (2)$$

where α = altitude of the sun (angular elevation above the horizon)
L = latitude of the observer
δ_s = declination of the sun
h_s = hour angle of sun (angular distance from the meridian of the observer)
a_s = azimuth of the sun (measured eastward from north)

From Eqs. (1) and (2) it can be seen that the altitude and azimuth of the sun are functions of the latitude of the observer, the time of day (hour angle), and the date (declination).

Figure A2.1(*b*–*g*) provides a series of charts, one for each 5° of latitude (except 5°, 15°, 75°, and 85°) giving the altitude and azimuth of the sun as a function of the true solar time and the declination of the sun in a form originally suggested by Hand. Linear interpolation for intermediate latitudes will give results within the accuracy to which the charts can be read.

On these charts, a point corresponding to the projected position of the sun is determined from the heavy lines corresponding to declination and solar time.

To find the solar altitude and azimuth:

1. Select the chart or charts appropriate to the latitude.
2. Find the solar declination δ corresponding to the date.
3. Determine the *true solar time* as follows:
   (a) To the *local standard time* (zone time) add 4′ for each degree of longitude the station is east of the standard meridian or subtract 4′ for each degree west of the standard meridian to get the *local mean solar time*.
   (b) To the *local mean solar time* add algebraically the equation of time; the sum is the required *true solar time*.
4. Read the required altitude and azimuth at the point determined by the declination and the true solar time. Interpolate linearly between two charts for intermediate latitudes.

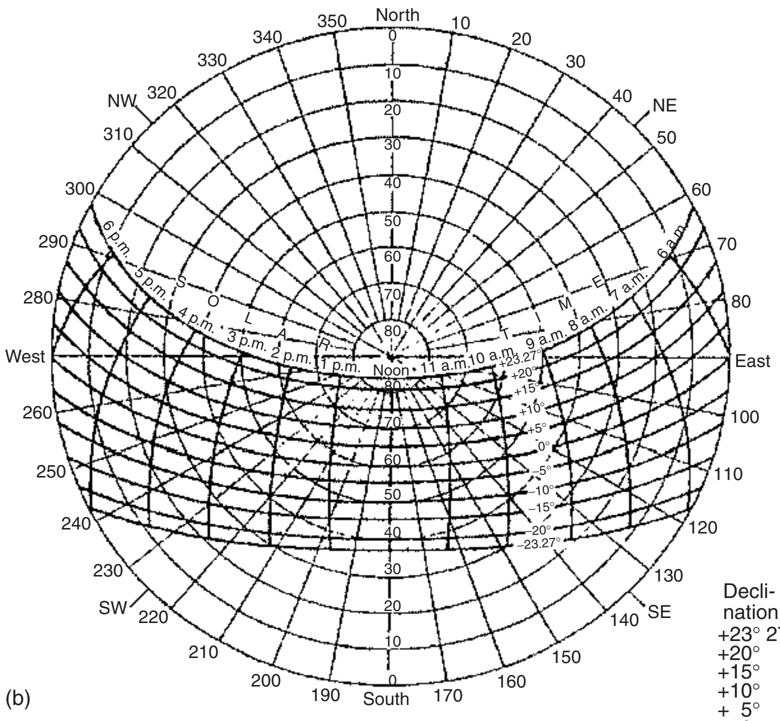
It should be emphasized that the solar altitude determined from these charts is the true geometric position of the center of the sun. At low solar elevations terrestrial refraction may considerably alter the apparent position of sun. Under average atmospheric refraction the sun will appear on the horizon when it actually is about 34′ below the horizon; the effect of refraction decreases rapidly with increasing solar elevation. Since sunset or sunrise is defined as the time when the upper limb of the sun appears on the horizon, and the semidiameter of the sun is 16′, sunset or sunrise occurs under average atmospheric refraction when the sun is 50′ below the horizon. In polar regions especially, unusual atmospheric refraction can make considerable variation in the time of sunset or sunrise.

The 90°N chart is included for interpolation purposes; the azimuths lose their directional significance at the pole.

Altitude and azimuth in southern latitudes. To compute solar altitude and azimuth for southern latitudes, change the sign of the solar declination and proceed as above. The resulting azimuths will indicate angular distance from *south* (measured eastward) rather than from north.

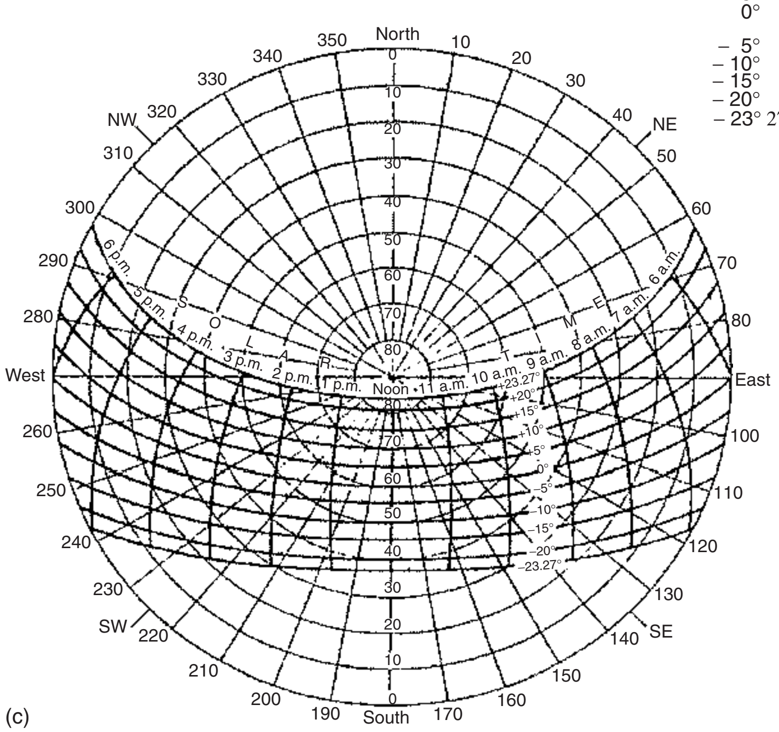
(a)

FIGURE A2.1 Description of method for calculating true solar time, together with accompanying meteorological charts, for computing solar-altitude and azimuth angles, (a) Description of method; (b) chart, 25°N latitude; (c) chart, 30°N latitude; (d) chart, 35°N latitude; (e) chart, 40°N latitude; (f) chart, 45°N latitude; (g) chart, 50°N latitude. Description and charts reproduced from the “Smithsonian Meteorological Tables” with permission from the Smithsonian Institute, Washington, D.C.



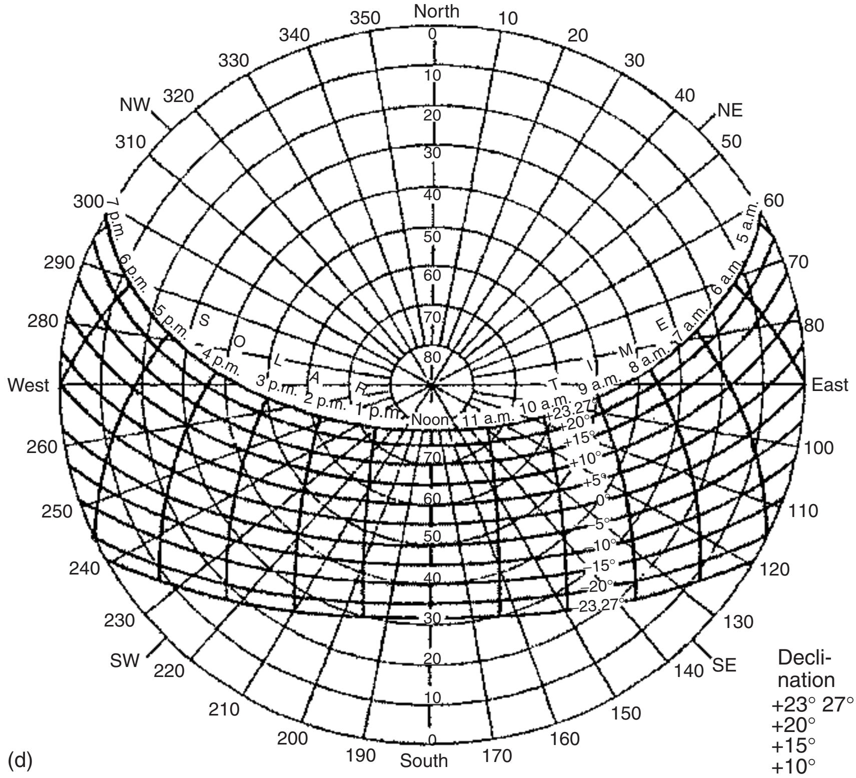
(b)

Declination	Approx. dates
+23° 27°	June 22
+20°	May 21, July 24
+15°	May 1, Aug. 13
+10°	Apr. 16, Aug. 28
+ 5°	Apr. 3, Sept. 10
0°	Mar. 21, Sept. 23
− 5°	Mar. 8, Oct. 6
− 10°	Feb. 23, Oct. 20
− 15°	Feb. 9, Nov. 3
− 20°	Jan. 21, Nov. 22
− 23° 27°	Dec. 22



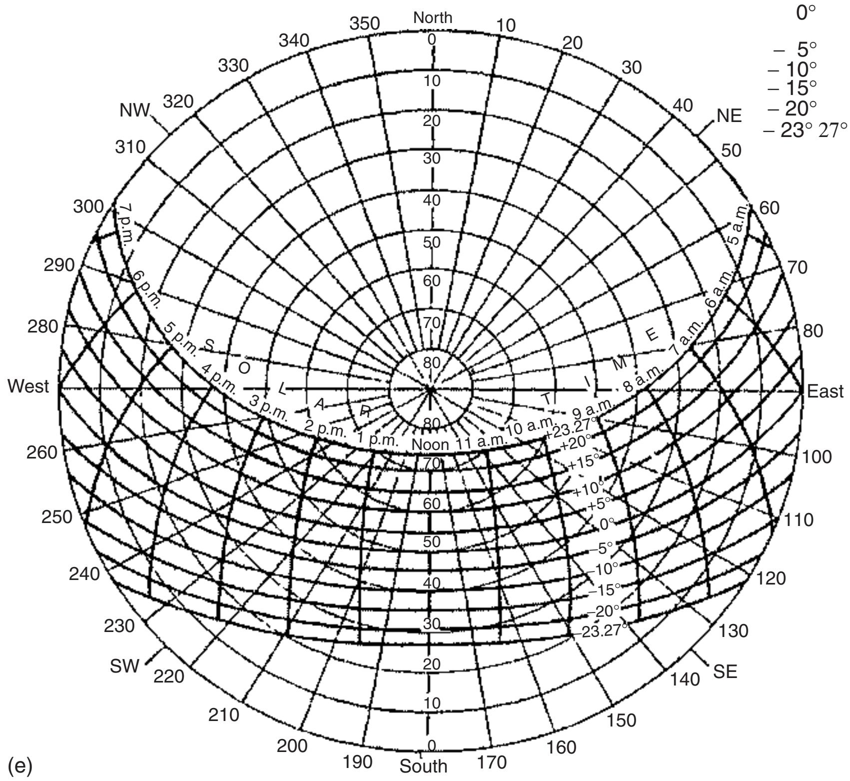
(c)

FIGURE A2.1 (*continued*)



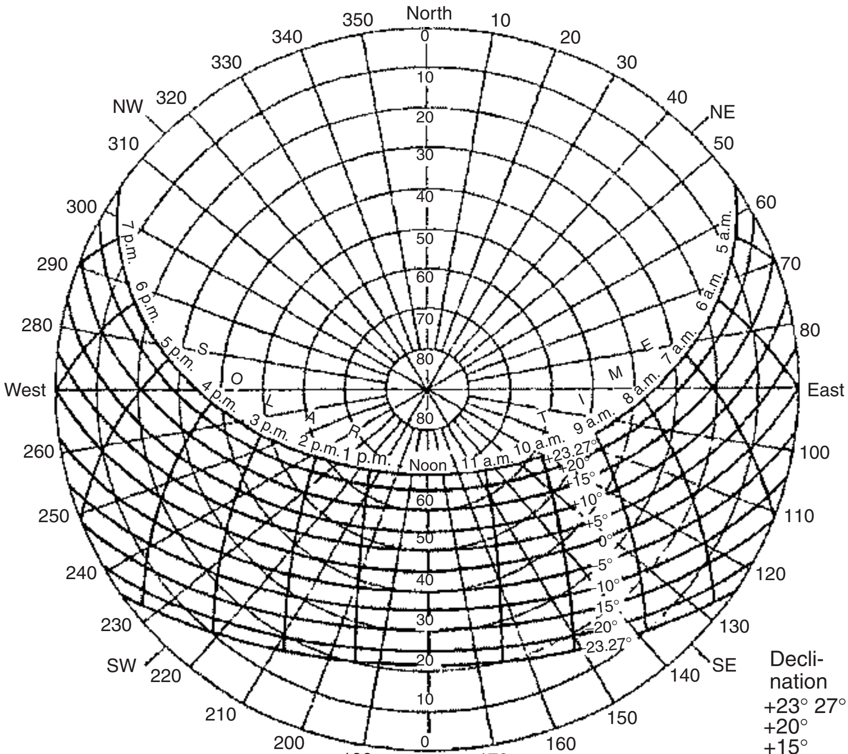
(d)

Declination	Approx. dates
+23° 27°	June 22
+20°	May 21, July 24
+15°	May 1, Aug. 13
+10°	Apr. 16, Aug. 28
+ 5°	Apr. 3, Sept. 10
0°	Mar. 21, Sept. 23
− 5°	Mar. 8, Oct. 6
− 10°	Feb. 23, Oct. 20
− 15°	Feb. 9, Nov. 3
− 20°	Jan. 21, Nov. 22
− 23° 27°	Dec. 22



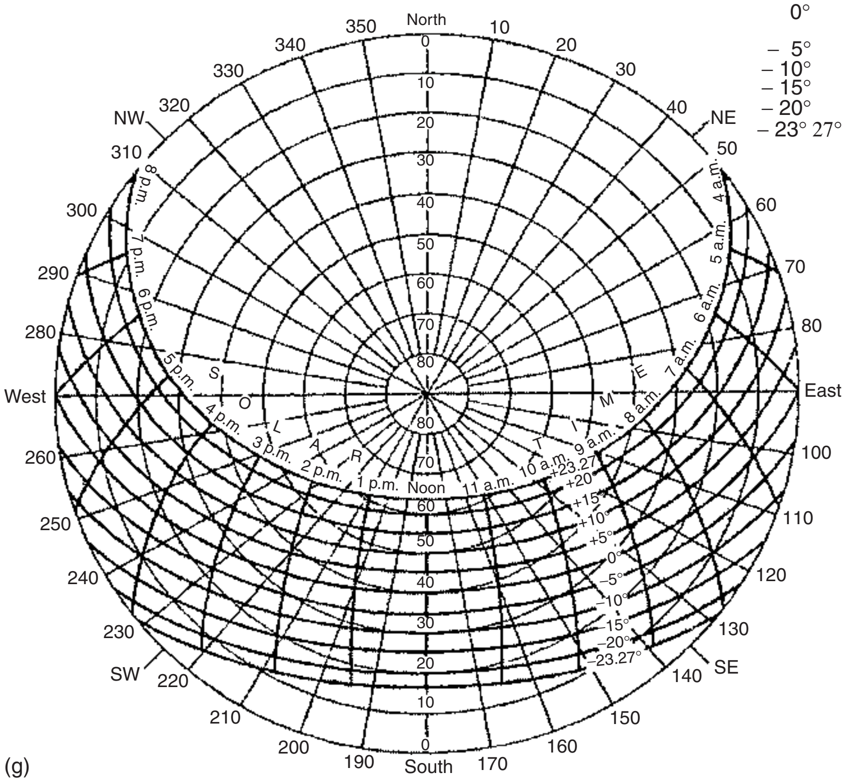
(e)

FIGURE A2.1 (*continued*)



(f)

Declination	Approx. dates
+23° 27°	June 22
+20°	May 21, July 24
+15°	May 1, Aug. 13
+10°	Apr. 16, Aug. 28
+ 5°	Apr. 3, Sept. 10
0°	Mar. 21, Sept. 23
– 5°	Mar. 8, Oct. 6
– 10°	Feb. 23, Oct. 20
– 15°	Feb. 9, Nov. 3
– 20°	Jan. 21, Nov. 22
– 23° 27°	Dec. 22



(g)

FIGURE A2.1 (*continued*)

TABLE A2.1 Solar Irradiance for Different Air Masses

Wavelength	Air Mass; $\alpha = 0.66$; $\beta = 0.085$[a]				
	0	1	4	7	10
0.290	482.0	0.0	0.0	0.0	0.0
0.295	584.0	0.0	0.0	0.0	0.0
0.300	514.0	4.1	0.0	0.0	0.0
0.305	603.0	11.4	0.0	0.0	0.0
0.310	689.0	30.5	0.0	0.0	0.0
0.315	764.0	79.4	0.1	0.0	0.0
0.320	830.0	202.6	2.9	0.0	0.0
0.325	975.0	269.5	5.7	0.1	0.0
0.330	1059.0	331.6	10.2	0.3	0.0
0.335	1081.0	383.4	17.1	0.8	0.0
0.340	1074.0	431.3	24.9	1.8	0.1
0.345	1069.0	449.2	33.3	2.5	0.2
0.350	1093.0	480.5	40.8	3.5	0.3
0.355	1083.0	498.0	48.4	4.7	0.5
0.360	1068.0	513.7	57.2	6.4	0.7
0.365	1132.0	561.3	68.4	8.3	1.0
0.370	1181.0	603.5	80.5	10.7	1.4
0.375	1157.0	609.4	89.0	13.0	1.9
0.380	1120.0	608.0	97.2	15.6	2.5
0.385	1098.0	609.8	104.5	17.9	3.1
0.390	1098.0	623.9	114.5	21.0	3.9
0.395	1189.0	691.2	135.8	26.7	5.2
0.400	1429.0	849.9	178.8	37.6	7.9
0.405	1644.0	992.8	218.7	48.2	10.6
0.410	1751.0	1073.7	247.5	57.1	13.2
0.415	1774.0	1104.5	266.5	64.3	15.5
0.420	1747.0	1104.3	278.9	70.4	17.8
0.425	1693.0	1086.5	287.2	78.9	20.1
0.430	1639.0	1067.9	295.4	81.7	22.6
0.435	1663.0	1100.1	318.4	92.2	26.7
0.440	1810.0	1215.5	368.2	111.5	33.8
0.445	1922.0	1310.4	415.3	131.6	41.7
0.450	2006.0	1388.4	460.3	152.6	50.6
0.455	2057.0	1434.8	486.9	165.2	56.1
0.460	2066.0	1452.2	504.4	175.2	60.8
0.465	2048.0	1450.7	515.7	183.3	65.1
0.470	2033.0	1451.2	527.9	192.0	69.8
0.475	2044.0	1470.3	547.3	203.7	75.8
0.480	2074.0	1503.4	572.6	218.1	83.1
0.485	1976.0	1443.3	562.4	219.2	85.4
0.490	1950.0	1435.2	572.2	228.2	91.0
0.495	1960.0	1453.6	592.9	241.9	98.7
0.500	1942.0	1451.2	605.6	252.7	105.5
0.505	1920.0	1440.1	607.6	256.4	108.2
0.510	1882.0	1416.8	604.4	257.8	110.0
0.515	1833.0	1384.9	597.3	257.6	111.1
0.520	1833.0	1390.0	606.1	264.3	115.2
0.525	1852.0	1409.5	621.3	273.9	120.7
0.530	1842.0	1406.9	626.9	279.4	124.5
0.535	1818.0	1393.6	627.7	282.8	127.4
0.540	1783.0	1371.7	624.5	284.4	129.5
0.545	1754.0	1354.2	623.2	286.8	132.0
0.550	1725.0	1336.6	621.7	289.2	134.5
0.555	1720.0	1335.7	625.5	293.0	137.3

(continued)

TABLE A2.1 ***(Continued)***

Wavelength	Air Mass; $\alpha = 0.66$; $\beta = 0.085$[a]				
	0	1	4	7	10
0.560	1695.0	1319.2	622.0	293.3	138.3
0.565	1705.0	1330.0	631.3	299.6	142.2
0.570	1712.0	1338.4	639.5	305.6	146.0
0.575	1719.0	1346.9	647.8	311.6	149.6
0.580	1715.0	1346.7	652.0	315.7	152.8
0.585	1712.0	1347.3	656.6	320.0	156.0
0.590	1700.0	1340.7	657.7	322.6	158.3
0.595	1682.0	1329.4	656.4	324.1	160.0
0.600	1660.0	1319.6	655.8	325.9	162.0
0.605	1647.0	1311.0	661.3	333.6	168.2
0.610	1635.0	1307.9	669.6	342.8	175.5
0.620	1602.0	1294.2	682.4	359.9	189.7
0.630	1570.0	1280.9	695.6	377.8	205.2
0.640	1544.0	1272.1	711.4	397.9	222.5
0.650	1511.0	1257.1	723.9	416.9	240.1
0.660	1486.0	1244.2	730.2	428.6	251.6
0.670	1456.0	1226.8	733.8	438.9	262.5
0.680	1427.0	1209.9	737.4	449.5	273.9
0.690	1402.0	1196.2	742.9	461.3	286.5
0.698	1374.6	1010.3	546.1	311.8	181.6
0.700	1369.0	1175.3	743.7	470.6	297.7
0.710	1344.0	1157.4	739.2	472.1	301.5
0.720	1314.0	1135.1	731.7	471.6	304.0
0.728	1295.5	1003.1	582.3	351.7	212.5
0.730	1290.0	1117.8	727.1	479.0	307.7
0.740	1260.0	1095.1	718.9	471.9	309.8
0.750	1235.0	1076.6	713.2	472.4	313.0
0.762	1205.5	794.0	357.1	163.6	69.1
0.770	1185.0	1039.2	700.8	472.7	318.8
0.780	1159.0	1019.4	693.6	472.0	321.1
0.790	1134.0	1000.3	686.7	471.4	323.6
0.800	1109.0	981.2	679.4	470.5	325.8
0.806	1095.1	874.4	547.7	355.9	234.4
0.825	1048.0	931.6	654.3	459.6	322.8
0.830	1036.0	921.8	649.3	457.3	322.1
0.835	1024.5	912.4	644.4	455.2	321.5
0.846	998.1	476.2	181.0	85.9	44.2
0.860	968.0	506.4	212.0	107.4	58.3
0.870	947.0	453.8	174.7	84.0	43.8
0.875	436.5	449.2	173.4	83.6	43.7
0.887	912.5	448.6	178.3	87.7	46.7
0.900	891.0	448.9	183.7	92.3	50.0
0.907	882.8	455.2	190.9	97.6	53.7
0.915	874.5	461.5	198.5	103.2	57.5
0.925	863.5	279.0	73.6	28.0	12.1
0.930	858.0	221.8	46.9	15.4	6.0
0.940	847.0	313.4	95.0	39.6	18.5
0.950	837.0	296.5	86.3	35.0	16.0
0.955	828.5	321.1	102.3	44.1	21.2
0.965	811.5	344.4	120.4	55.1	27.8
0.975	794.0	576.9	346.0	224.6	150.1
0.985	776.0	544.6	316.1	201.2	132.4
1.018	719.2	617.5	391.0	247.5	156.7
1.082	620.0	512.9	290.4	164.4	93.1

(continued)

TABLE A2.1 (Continued)

Wavelength	Air Mass; $\alpha = 0.66$; $\beta = 0.085$[a]				
	0	1	4	7	10
1.094	602.0	464.1	303.1	210.8	149.9
1.098	596.0	503.7	304.1	183.6	110.9
1.101	591.8	504.8	362.7	267.3	198.8
1.128	560.5	135.1	27.7	9.1	3.6
1.131	557.0	152.2	35.3	12.6	5.3
1.137	550.1	143.1	31.7	11.0	4.5
1.144	542.0	191.2	57.4	24.2	11.6
1.147	538.5	174.5	48.2	19.3	8.8
1.178	507.0	399.3	195.1	95.4	46.6
1.189	496.0	402.2	214.5	114.4	61.0
1.193	492.0	424.0	310.8	233.3	176.6
1.222	464.3	391.8	235.3	141.3	84.9
1.236	451.2	390.8	254.1	165.2	107.4
1.264	426.5	329.2	209.7	140.0	94.3
1.276	416.7	342.6	238.6	172.6	126.3
1.288	406.8	347.3	216.1	134.4	83.7
1.314	386.1	298.3	137.6	63.5	29.3
1.335	369.7	190.6	85.0	46.7	27.7
1.384	343.7	5.7	0.1	0.0	0.0
1.432	321.0	44.6	5.4	1.3	0.4
1.457	308.6	85.4	20.6	7.7	3.3
1.472	301.4	77.4	17.4	6.2	2.6
1.542	270.4	239.3	165.9	115.0	79.7
1.572	257.3	222.6	168.1	130.4	102.1
1.599	245.4	216.0	166.7	131.5	104.5
1.608	241.5	208.5	157.4	122.1	95.7
1.626	233.6	206.7	160.7	127.5	101.9
1.644	225.6	197.9	152.4	120.1	95.5
1.650	223.0	195.7	150.9	119.1	94.7
1.676	212.1	181.9	114.8	72.4	45.7
1.732	187.9	161.5	102.5	65.1	41.3
1.782	166.6	136.7	75.6	41.8	23.1
1.862	138.2	4.0	0.1	0.0	0.0
1.955	112.9	42.7	14.5	6.8	3.6
2.008	102.0	69.4	35.8	17.7	6.4
2.014	101.2	74.7	45.5	28.8	17.8
2.057	95.6	69.5	41.3	25.3	14.8
2.124	87.4	70.0	35.9	18.4	9.5
2.156	83.8	66.0	32.3	15.8	7.7
2.201	78.9	66.1	49.1	38.0	29.7
2.266	72.4	61.6	46.8	36.8	29.3
2.320	67.6	57.2	43.2	33.8	26.8
2.338	66.3	54.7	39.9	30.4	23.4
2.356	65.1	52.0	36.3	26.5	19.6
2.388	62.8	36.0	18.7	11.7	7.8
2.415	61.0	32.5	15.8	9.4	6.0
2.453	58.3	29.6	13.7	7.9	5.0
2.494	55.4	20.3	6.8	3.2	1.7
2.537	52.4	4.6	0.4	0.1	0.0
2.900	35.0	2.9	0.2	0.0	0.0
2.941	33.4	6.0	1.0	0.3	0.1
2.954	32.8	5.7	0.9	0.3	0.1
2.973	32.1	8.7	2.2	0.9	0.4
3.005	30.8	7.8	1.8	0.7	0.3

(continued)

TABLE A2.1 ***(Continued)***

Wavelength	Air Mass; $\alpha = 0.66$; $\beta = 0.085$[a]				
	0	1	4	7	10
3.045	28.8	4.7	0.7	0.2	0.1
3.056	28.2	4.9	0.8	0.2	0.1
3.097	26.2	3.2	0.4	0.1	0.0
3.132	24.9	6.8	1.7	0.7	0.3
3.156	24.1	18.7	12.6	8.9	6.3
3.204	22.5	2.1	0.2	0.0	0.0
3.214	22.1	3.4	0.5	0.1	0.0
3.245	21.1	3.9	0.7	0.2	0.1
3.260	20.6	3.7	0.6	0.2	0.1
3.285	19.7	14.2	8.5	5.1	2.8
3.317	18.8	12.9	6.9	3.5	1.3
3.344	18.1	4.2	0.9	0.3	0.1
3.403	16.5	12.3	7.8	5.1	3.2
3.450	15.6	12.5	8.9	6.7	5.0
3.507	14.5	12.5	9.9	8.1	6.7
3.538	14.2	11.8	8.8	6.9	5.5
3.573	13.8	10.9	5.4	2.6	1.3
3.633	13.1	10.8	8.3	6.7	5.5
3.673	12.6	9.1	6.1	4.6	3.5
3.696	12.3	10.4	8.2	6.7	5.6
3.712	12.2	10.9	9.0	7.6	6.5
3.765	11.5	9.5	7.2	5.9	4.8
3.812	11.0	8.9	6.7	5.4	4.4
3.888	10.4	8.1	5.6	4.0	2.9
3.923	10.1	8.0	5.6	4.2	3.1
3.948	9.9	7.8	5.5	4.0	3.0
4.045	9.1	6.7	4.1	2.6	1.5
Total Wm^3	1353	889.2	448.7	255.2	153.8

W/m^2 μm; H_2O 20 mm; O_3 3.4 mm.

[a] The parameters α and β are measures of turbidity of the atmosphere. They are used in the atmospheric transmittance equation $\bar{\tau}_{atm} = e^{-(C_1+C_2)_m}$; C_1 includes Rayleigh and ozone attenuation; $C_2 \equiv \beta/\lambda^a$.

Source: From Thekaekara, M. P. 1974. *The Energy Crisis and Energy from the Sun*. Institute for Environmental Sciences.

TABLE A2.2 Monthly Averaged, Daily Extraterrestrial Insolation on a Horizontal Surface (Units: Wh/m^2)

Latitude (deg)	January	February	March	April	May	June	July	August	September	October	November	December
20	7415	8397	9552	10,422	10,801	10,868	10,794	10,499	9791	8686	7598	7076
25	6656	7769	9153	10,312	10,936	11,119	10,988	10,484	9494	8129	6871	6284
30	5861	7087	8686	10,127	11,001	11,303	11,114	10,395	9125	7513	6103	5463
35	5039	6359	8153	9869	10,995	11,422	11,172	10,233	8687	6845	5304	4621
40	4200	5591	7559	9540	10,922	11,478	11,165	10,002	8184	6129	4483	3771
45	3355	4791	6909	9145	10,786	11,477	11,099	9705	7620	5373	3648	2925
50	2519	3967	6207	8686	10,594	11,430	10,981	9347	6998	4583	2815	2100
55	1711	3132	5460	8171	10,358	11,352	10,825	8935	6325	3770	1999	1320
60	963	2299	4673	7608	10,097	11,276	10,657	8480	5605	2942	1227	623
65	334	1491	3855	7008	9852	11,279	10,531	8001	4846	2116	544	97

TABLE A2.3a Worldwide Global Horizontal Average Solar Radiation (Units: MJ/sq.m-day)

Position	Lat	Long	January	February	March	April	May	June	July	August	September	October	November	December
Argentina														
Buenos Aires	34.58 S	58.48 W	24.86	21.75	18.56	11.75	8.71	7.15	7.82	8.75	14.49	16.66	24.90	21.93
Australia														
Adelaide	34.93 S	138.52 E	20.99	17.50	20.15	18.27	17.98	—	18.81	19.64	20.11	20.88	20.57	20.72
Brisbane	27.43 S	153.08 E	25.36	22.22	13.25	16.61	12.23	11.52	9.70	15.10	17.61	19.89	—	—
Canberra	35.30 S	148.18 E	28.20	24.68	20.56	14.89	10.29	6.62	—	12.33	16.88	24.06	26.00	25.77
Darwin	12.47 S	130.83 E	26.92	23.40	18.13	13.62	9.30	7.89	9.41	11.15	14.85	18.87	23.43	22.34
Hobart	42.88 S	147.32 E	—	—	—	10.09	7.26	6.04	5.72	9.21	13.54	18.12	—	—
Laverton	37.85 S	114.08 E	22.96	20.42	15.59	13.40	7.48	6.10	6.54	10.43	13.24	18.76	—	—
Sydney	33.87 S	151.20 E	21.09	21.75	17.63	13.63	9.78	8.79	7.62	12.84	16.93	22.10	—	—
Austria														
Wien	48.20 N	16.57 E	3.54	7.10	8.05	14.72	16.79	20.87	19.89	17.27	12.55	8.45	3.51	2.82
Innsbruck	47.27 N	11.38 E	5.57	9.28	10.15	15.96	14.57	17.65	18.35	17.26	12.98	9.08	4.28	3.50
Barbados														
Husbands	13.15 N	59.62 W	19.11	20.23	—	21.80	19.84	20.86	21.55	22.14	—	—	18.30	16.56
Belgium														
Ostende	51.23 N	2.92 E	2.82	5.75	9.93	15.18	16.74	16.93	18.21	18.29	11.71	6.15	2.69	1.97
Melle	50.98 N	3.83 E	2.40	4.66	8.41	13.55	14.23	13.28	15.71	15.61	10.63	5.82	2.40	1.59
Brunei														
Brunei	4.98 N	114.93 E	19.46	20.12	22.71	20.54	19.74	18.31	19.38	20.08	20.83	17.51	17.39	18.12
Bulgaria														
Chirpan	42.20 N	25.33 E	6.72	6.79	8.54	13.27	17.25	17.39	19.85	14.61	12.53	8.52	5.08	5.09
Sofia	42.65 N	23.38 E	4.05	6.23	7.93	9.36	12.98	19.73	19.40	17.70	14.71	6.44	—	3.14
Canada[a]														
Montreal	45.47 N	73.75 E	4.74	8.33	11.84	10.55	15.05	22.44	21.08	18.67	14.83	9.18	4.04	4.01
Ottawa	45.32 N	75.67 E	5.34	9.59	13.33	13.98	20.18	20.34	19.46	17.88	13.84	7.38	4.64	5.04
Toronto	43.67 N	79.38 E	4.79	8.15	11.96	14.00	18.16	24.35	23.38	—	15.89	9.40	4.72	3.79
Vancouver	49.18 N	123.17 E	3.73	4.81	12.14	16.41	20.65	24.04	22.87	19.08	12.77	7.39	4.29	1.53
Chile														
Pascua	27.17 S	109.43 W	19.64	16.65	—	11.12	9.52	8.81	10.90	12.29	17.19	20.51	21.20	22.44
Santiago	33.45 S	70.70 W	18.61	16.33	13.44	8.32	5.07	3.66	3.35	5.65	8.15	13.62	20.14	23.88
China														
Beijing	39.93 N	116.28 W	7.73	10.59	13.87	17.93	20.18	18.65	15.64	16.61	15.52	11.29	7.25	6.89
Guangzhou	23.13 N	113.32 E	11.01	6.32	4.04	7.89	10.53	12.48	16.14	16.02	15.03	15.79	11.55	9.10

(continued)

TABLE A2.3a ***(Continued)***

Position	Lat	Long	January	February	March	April	May	June	July	August	September	October	November	December
Harbin	45.75 N	126.77 E	5.15	9.54	17.55	20.51	20.33	17.85	19.18	16.09	13.38	14.50	10.50	6.98
Kunming	25.02 N	102.68 E	9.92	11.26	14.38	18.00	18.53	17.37	11.95	18.47	15.94	12.45	11.96	13.62
Lanzhou	36.05 N	103.88 E	7.30	12.47	10.62	18.91	17.40	20.40	20.23	17.37	13.23	10.21	8.22	6.43
Shanghai	31.17 N	121.43 E	7.44	10.31	11.78	14.36	14.23	16.79	14.63	11.85	15.96	12.03	7.73	8.70
Columbia														
Bogota	4.70 N	74.13 W	17.89	—	19.37	16.58	14.86	—	15.42	18.20	17.05	14.58	14.20	16.66
Cuba														
Havana	23.17 N	82.35 W	—	14.70	18.94	20.95	22.63	18.83	21.40	20.19	16.84	16.98	13.19	13.81
Czech														
Kucharovice	48.88 N	16.08 E	3.03	5.85	9.88	14.06	20.84	19.24	21.18	19.41	13.61	6.11	3.47	2.12
Churanov	49.07 N	13.62 E	2.89	5.82	9.24	13.18	21.32	15.68	20.51	19.49	12.84	5.68	3.36	2.99
Hradec Kralov	50.25 N	15.85 E	3.51	5.94	10.58	15.95	20.42	18.43	17.17	17.92	11.86	6.27	2.45	1.89
Denmark														
Copenhagen	55.67 N	12.30 E	1.83	3.32	7.09	11.12	21.39	24.93	—	13.92	10.10	5.20	2.81	1.23
Egypt														
Cairo	30.08 N	31.28 E	10.06	12.96	18.49	23.04	21.91	26.07	25.16	23.09	21.01	—	11.74	9.85
Mersa Matruh	31.33 N	27.22 E	8.38	11.92	18.47	24.27	24.17	—	26.67	26.27	21.92	18.28	11.71	8.76
Ethiopia														
Addis Ababa	8.98 N	38.80 E	—	11.39	—	12.01	—	—	—	6.33	9.35	11.71	11.69	11.50
Fiji														
Nandi	17.75 S	177.45 E	20.82	20.65	20.25	18.81	15.68	14.18	15.08	16.71	19.37	20.11	21.78	25.09
Suva	48.05 S	178.57 E	20.37	17.74	16.22	13.82	10.81	12.48	11.40	—	—	18.49	19.96	20.99
Finland														
Helsinki	60.32 N	24.97 E	1.13	2.94	5.59	11.52	17.60	16.81	20.66	15.44	8.44	3.31	0.97	0.63
France														
Agen	44.18 N	0.60 E	4.83	7.40	10.69	17.12	19.25	20.42	21.63	20.64	15.56	8.41	5.09	5.01
Nice	43.65 N	7.20 E	6.83		11.37	17.79	20.74	24.10	24.85	24.86	15.04	10.99	7.08	6.73
Paris	48.97 N	2.45 E	2.62	5.08	7.21	12.90	14.84	13.04	15.54	16.30	10.17	5.61	3.14	2.20
Germany														
Bonn	50.70 N	7.15 E	2.94	5.82	8.01	14.27	15.67	14.41	18.57	17.80	11.70	6.15	3.42	1.90
Nuremberg	53.33 N	13.20 E	3.23	6.92	9.08	15.69	15.71	18.21	21.14	17.98	12.43	8.15	2.79	2.51
Bremen	53.05 N	8.80 E	2.36	4.93	8.53	14.52	14.94	14.52	19.40	15.02	10.48	6.27	2.80	1.66
Hamburg	53.63 N	10.00 E	1.97	3.96	7.59	12.32	14.11	12.69	19.00	14.11	10.29	6.45	2.33	1.43
Stuttgart	48.83 N	9.20 E	3.59	7.18	9.22	15.81	17.72	17.44	22.21	19.87	12.36	7.81	3.19	2.54
Ghana														
Bole	9.03 N	2.48 W	18.29	19.76	19.71	19.15	16.61	—	—	13.68	16.29	17.27	17.33	15.93

Accra	5.60 N	0.17 W	14.82	16.26	18.27	16.73	18.15	13.96	13.86	13.49	15.32	19.14	18.16	14.23
Great Britain														
Belfast	54.65 N	6.22 W	2.00	3.60	6.85	12.00	15.41	15.09	15.46	13.56	11.49	4.63	2.34	1.24
Jersey	49.22 N	2.20 W	2.76	5.65	9.51	14.98	18.51	17.83	18.14	18.62	12.98	6.16	3.26	2.83
London	51.52 N	0.12 W	2.24	3.87	7.40	12.01	12.38	13.24	16.59	16.23	12.59	5.67	2.87	1.97
Greece														
Athens	37.97 N	23.72 E	9.11	10.94	15.70	20.91	23.85	25.48	24.21	23.08	19.03	13.29	5.98	6.64
Sikiwna	37.98 N	22.73 E	7.60	8.16	11.99	21.06	22.62	24.32	23.56	21.73	17.30	11.75	9.45	6.35
Guadeloupe														
Le Raizet	16.27 N	61.52 W	14.88	18.10	20.55	19.69	20.26	20.65	20.65	20.24	18.47	17.79	13.49	14.38
Guyana														
Cayenne	4.83 N	52.37 W	14.46	14.67	16.28	17.57	—	14.92	17.42	18.24	20.52	—	22.69	17.04
Hong Kong														
King's Park	22.32 N	114.17 W	12.34	7.39	6.94	9.50	11.38	13.60	16.70	17.06	15.91	16.52	14.19	10.00
Hungary														
Budapest	47.43 N	19.18 E	2.61	7.46	11.14	14.46	20.69	19.47	21.46	19.72	12.88	7.96	2.95	2.47
Iceland														
Reykjavik	64.13 N	21.90 W	0.52	2.02	6.25	11.77	13.07	14.58	16.83	11.35	9.70	3.18	1.00	0.65
India														
Bombay	19.12 N	72.85 E	18.44	21.00	22.72	24.52	24.86	19.75	15.84	16.00	18.19	20.38	19.18	17.81
Calcutta	22.53 N	88.33 E	15.69	18.34	20.09	22.34	22.37	17.55	17.07	16.55	16.52	16.90	16.35	15.00
Madras	13.00 N	80.18 E	19.09	22.71	25.14	24.88	23.89	—	18.22	19.68	19.S1	16.41	14.76	15.79
Nagpur	21. 10 N	79.05 E	18.08	21.01	22.25	24.08	24.79	19.84	15.58	15.47	17.66	20.10	18.98	17.33
New Delhi	28.58 N	77.20 E	14.62	18.25	20.15	23.40	23.80	19.16	20.20	19.89	20.08	19.74	16.95	14.22
Ireland														
Dublin	53.43 N	6.25 W	2.51	4.75	7.48	11.06	17.46	19.11	15.64	13.89	9.65	5.77	2.93	—
Israel														
Jerusalem	31.78 N	35.22 E	10.79	13.01	18.08	23.79	29.10	31.54	31.83	28.79	25.19	20.26	12.61	10.71
Italy														
Milan	45.43 N	9.28 E	—	6.48	10.09	13.17	17.55	16.32	18.60	16.86	11.64	5.40	3.52	2.41
Rome	41.80 N	12.55 E	—	9.75	13.38	15.82	15.82	18.89	22.27	21.53	16.08	8.27	6.41	4.49
Japan														
Fukuoka	33.58 N	130.38 E	8.11	8.72	10.95	13.97	14.36	12.81	13.84	16.75	13.92	11.86	10.05	7.30
Tateno	36.05 N	140.13 E	9.06	12.17	11.00	15.78	16.52	15.26	—	—	—	9.60	8.55	8.26
Yonago	35.43 N	133.35 E	6.25	7.16	10.87	17.30	16.72	15.44	17.06	19.93	12.41	10.82	7.50	5.51
Kenya														
Mombasa	4.03 S	39.62 E	22.30	22.17	22.74	18.49	18.31	17.41	—	18.12	21.03	22.97	21.87	21.25
Nairobi	1.32 S	36.92 E	—	24.10	21.20	18.65	14.83	15.00	13.44	14.12	19.14	19.38	16.90	18.27

(continued)

TABLE A2.3a ***(Continued)***

Position	Lat	Long	January	February	March	April	May	June	July	August	September	October	November	December
Lithuania														
Kaunas	54.88 N	23.88 E	1.89	4.43	7.40	12.97	18.88	18.74	21.41	15.79	10.40	5.64	1.80	1.10
Madagascar														
Antanarivo	18.80 S	47.48 E	15.94	13.18	13.07	11.53	9.25	8.21	9.32	—	—	16.43	15.19	15.62
Malaysia														
Kualalumpur	3.12 N	101.55 E	15.36	17.67	18.48	16.87	15.67	16.24	15.32	15.89	14.62	14.13	13.54	11.53
Piang	5.30 N	100.27 E	19.47	21.35	23.24	20.52	18.63	19.32	17.17	16.96	15.93	16.01	18.35	17.37
Martinique														
Le Lamentin	14.60 N	61.00 W	17.76	20.07	22.53	21.95	22.42	21.23	20.86	21.84	20.23	19.87	14.08	16.25
Mexico														
Chihuahua	28.63 N	106.08 W	14.80	—	—	—	26.94	26.28	24.01	24.22	20.25	19.55	10.57	15.79
Orizabita	20.58 N	99.20 E	19.49	23.07	27.44	27.35	26.04	25.05	—	27.53	21.06	17.85	15.48	12.93
Mongolia														
Ulan Bator	47.93 N	106.98 E	6.28	9.22	14.34	18.18	20.50	19.34	16.34	16.65	14.08	11.36	7.19	5.35
Uliasutai	47.75 N	96.85 E	6.43	10.71	14.83	20.32	23.86	20.46	21.66	17.81	15.97	10.92	7.32	5.08
Morocco														
Casablanca	33.57 N	7.67 E	11.46	12.70	15.93	21.25	24.45	25.27	25.53	23.60	19.97	14.68	11.61	9.03
Mozambique														
Maputo	25.97 S	32.60 E	26.35	23.16	19.33	20.54	16.33	14.17	—	—	—	22.55	25.48	26.19
Netherlands														
Maastricht	50.92 N	5.78 E	3.20	5.43	8.48	14.82	14.97	14.32	18.40	17.51	11.65	6.51	3.01	1.72
New Caledonia														
Koumac	20.57 S	164.28 E	24.89	21.15	16.96	18.98	15.67	14.55	15.75	17.62	22.48	15.83	27.53	26.91
New Zealand														
Wilmington	41.28 S	174.77 E	22.59	19.67	14.91	9.52	6.97	4.37	5.74	7.14	12.50	16.34	19.07	24.07
Christchurch	43.48 S	172.55 E	23.46	19.68	13.98	8.96	6.47	4.74	5.38	6.94	13.18	17.45	18.91	24.35
Nigeria														
Benin City	6.32 N	5.60 E	14.89	17.29	19.15	17.21	16.97	15.04	10.24	12.54	14.37	15.99	17.43	15.75
Norway														
Bergen	60.40 N	5.32 E	0.46	1.33	3.18	8.36	19.24	16.70	16.28	10.19	6.53	3.19	1.36	0.35
Oman														
Seeb	23.58 N	58.28 E	12.90	14.86	21.22	22.22	25.30	24.02	23.46	21.66	20.07	18.45	15.49	13.12
Salalah	17.03 N	54.08 E	16.52	16.92	18.49	20.65	21.46	16.92	8.52	11.41	17.14	18.62	16.42	—
Pakistan														
Karachi	24.90 N	67.13 E	13.84	—	—	19.69	20.31	16.62	—	—	—	—	12.94	11.07
Multan	30.20 N	71.43 E	12.29	15.86	18.33	22.35	22.57	21.65	20.31	20.44	20.57	15.91	12.68	10.00

Islamabad	33.62 N	73.10 E	10.38	12.42	16.98	22.65	—	25.49	20.64	18.91	14.20	15.30	10.64	8.30
Peru														
Puno	15.83 S	70.02 W	14.98	12.92	16.08	20.03	17.45	17.42	15.74	15.32	16.11	16.18	14.24	13.90
Poland														
Warszawa	52.28 N	20.97 E	1.73	3.83	7.81	10.53	19.22	17.11	20.18	15.00	10.65	4.95	2.39	1.68
Kolobrzeg	54.18 N	15.58 E	2.50	3.25	8.86	15.21	20.79	20.50	17.19	16.46	7.95	5.75	1.78	1.18
Portugal														
Evora	38.57 N	7.90 W	9.92	12.43	17.81	18.69	23.57	29.23	28.75	23.77	20.17	—	6.81	4.57
Lisbon	38.72 N	9.15 W	9.24	11.60	17.52	18.49	24.64	29.02	28.14	22.20	19.76	13.56	7.18	4.83
Romania														
Bucuresti	44.50 N	26.13 E	7.05	10.22	12.04	16.53	18.97	22.16	23.19	—	17.17	9.55	4.82	—
Constania	44.22 N	28.63 E	5.62	9.28	14.31	20.59	23.23	25.80	27.98	24.22	16.91	11.89	6.19	5.10
Galati	45.50 N	28.02 E	6.09	9.33	14.31	17.75	21.77	22.74	25.55	19.70	14.05	11.26	6.32	5.38
Russia														
Alexandovsko	60.38 N	77.87 E	1.34	4.17	9.16	17.05	21.83	21.34	20.26	13.05	10.16	4.68	1.71	0.68
Moscow	55.75 N	37.57 E	1.45	3.96	8.09	11.69	18.86	18.12	17.51	14.17	10.92	4.03	2.28	1.29
St. Petersburg	59.97 N	30.30 E	1.03	3.11	4.88	12.24	20.59	21.55	20.43	13.27	7.83	2.93	1.16	0.59
Verkhoyansk	67.55 N	133.38 E	0.21	2.25	7.61	15.96	19.64	—	—	14.12	7.59	3.51	0.54	—
St. Pierre & Miquelon														
St. Pierre	46.77 N	56.17 W	4.43	6.61	12.50	17.57	18.55	17.84	19.95	16.46	12.76	8.15	3.69	3.33
Singapore														
Singapore	1.37 N	103.98 E	19.08	20.94	20.75	18.20	14.89	15.22	13.92	16.66	16.51	15.82	13.81	12.67
South Korea														
Seoul	37.57 N	126.97 E	6.24	9.40	10.34	13.98	16.35	17.49	10.65	12.94	11.87	10.35	6.47	5.14
South Africa														
Cape Town	33.98 S	18.60 E	27.47	25.57	—	15.81	11.44	9.08	8.35	13.76	17.30	22.16	26.37	27.68
Port Elizabeth	33.98 S	25.60 E	27.22	22.06	19.01	15.29	11.79	11.13	10.73	13.97	18.52	23.09	23.15	27.26
Pretoria	25.73 S	28.18 E	26.06	22.43	20.52	16.09	15.67	13.67	15.19	18.65	21.62	21.75	24.82	23.43
Spain														
Madrid	40.45 N	3.72 W	7.73	10.53	15.35	21.74	22.81	22.05	26.27	22.90	18.89	10.21	8.69	5.56
Sudan														
Wad Madani	14.40 N	33.48 E	21.92	24.01	23.43	25.17	23.92	23.51	22.40	22.85	21.75	20.47	20.19	19.21
Elfasher	13.62 N	25.33 E	21.56	21.84	24.54	25.29	24.31	24.15	22.87	21.19	22.58	23.85	—	—
Shambat	15.67 N	32.53 E	23.90	27.38	—	27.45	23.21	26.15	23.55	25.46	24.05	23.51	23.82	22.53
Sweden														
Karlstad	59.37 N	13.47 E	1.26	3.13	5.02	14.01	19.90	16.70	20.92	14.14	10.52	3.98	1.47	0.94
Lund	55.72 N	13.22 E	1.97	3.47	6.66	12.48	17.83	13.38	18.74	14.99	10.39	5.45	1.82	1.21
Stockholm	59.35 N	18.07 E	1.32	2.69	4.75	13.21	15.58	14.79	20.52	14.48	10.50	4.04	1.19	0.83

(continued)

TABLE A2.3a ***(Continued)***

Position	Lat	Long	January	February	March	April	May	June	July	August	September	October	November	December
Switzerland														
Geneva	46.25 N	6.13 E	2.56	7.21	9.46	17.07	20.98	19.78	22.38	20.50	13.62	8.44	3.31	2.87
Zurich	47.48 N	8.53 E	2.31	7.02	7.54	15.04	16.33	16.73	20.28	18.32	12.52	7.18	2.64	2.29
Thailand														
Bangkok	13.73 N	100.57 E	16.67	19.34	23.00	22.48	20.59	17.71	18.02	16.04	16.23	16.81	18.60	16.43
Trinidad & Tobago														
Crown Point	11.15 N	60.83 W	13.05	15.61	15.17	16.96	17.61	15.37	13.16	13.08	12.24	8.76	—	—
Tunisia														
Sidi Bouzid	36.87 N	10.35 E	7.88	10.38	13.20	17.98	25.12	26.68	27.43	24.33	18.87	12.11	9.37	6.72
Tunis	36.83 N	10.23 E	7.64	9.88	14.79	31.61	25.31	26.03	26.60	20.37	19.58	12.91	9.35	7.16
Ukraine														
Kiev	50.40 N	30.45 E	2.17	4.87	11.15	12.30	20.49	—	18.99	18.55	9.72	9.84	3.72	2.52
Uzbekistan														
Tashkent	41.27 N	69.27 E	7.27	10.81	15.93	23.60	25.21	29.53	28.50	26.68	20.76	13.25	8.61	4.59
Venezuela														
Caracas	10.50 N	66.88 W	14.25	13.56	16.30	15.56	15.69	15.56	16.28	17.11	17.04	15.14	14.74	13.50
St. Antonio	7.85 N	72.45 W	11.78	10.54	10.65	12.07	12.65	21.20	14.68	15.86	16.62	15.32	12.28	11.28
St. Fernando	7.90 N	67.42 W	14.92	16.82	16.89	—	—	14.09	13.78	14.42	14.86	15.27	14.25	13.11
Vietnam														
Hanoi	21.03 N	105.85 E	5.99	7.48	8.73	13.58	19.10	21.26	19.85	19.78	20.67	14.78	12.44	13.21
Yogoslavia														
Beograd	44.78 N	20.53 E	4.92	6.27	10.64	14.74	20.95	22.80	22.09	20.27	15.57	11.24	6.77	4.99
Kopaonik	43.28 N	20.80 E	7.03	10.93	14.75	12.78	13.54	20.43	22.48	—	20.14	11.61	6.26	4.64
Portoroz	45.52 N	13.57 E	5.11	7.84	13.75	17.30	23.66	22.31	25.14	21.34	13.40	8.98	6.04	3.92
Zambia														
Lusaka	15.42 S	28.32 W	16.10	18.02	20.24	19.84	17.11	16.37	19.45	20.72	21.68	23.83	23.85	20.52
Zimbabwe														
Bulawayo	20.15 S	28.62 N	20.03	22.11	21.03	18.09	17.15	15.36	16.46	19.49	21.55	23.44	25.08	23.46
Harare	17.83 S	31.02 N	19.38	19.00	19.22	17.67	18.35	16.10	14.55	17.87	21.47	23.98	19.92	21.88

Note: Data for 872 locations is available from these sources in 68 countries.

[a] Source for Canadian Data: Environment Canada: Internet address: http://www.ec.gc.ca./envhome.html

Source: From Voeikov Main Geophysical Observatory, Russia: Internee address: http://wrdc-mgo.nrel.gov/html/get_data-ap.html

TABLE A2.3b Average Daily Solar Radiation on a Horizontal Surface in U.S.A. (Units: MJ/sq. m-day)

Position	January	February	March	April	May	June	July	August	September	October	November	December	Average
Alabama													
Birmingham	9.20	11.92	15.67	19.65	21.58	22.37	21.24	20.21	17.15	14.42	10.22	8.40	16.01
Montgomery	9.54	12.49	16.24	20.33	22.37	23.17	21.80	20.56	17.72	14.99	10.90	8.97	16.58
Alaska													
Fairbanks	0.62	2.77	8.31	14.66	17.98	19.65	16.92	12.36	7.02	3.20	1.01	0.23	8.74
Anchorage	1.02	3.41	8.18	13.06	15.90	17.72	16.69	12.72	8.06	3.97	1.48	0.56	8.63
Nome	0.51	2.95	8.29	15.22	18.97	19.65	16.69	11.81	7.72	3.63	0.99	0.09	8.86
St. Paul Island	1.82	4.32	8.52	12.72	14.08	14.42	12.83	10.33	7.84	4.54	2.16	1.25	7.95
Yakutat	1.36	3.63	7.72	12.61	14.76	15.79	14.99	12.15	7.95	3.97	1.82	0.86	8.18
Arizona													
Phoenix	11.58	15.33	19.87	25.44	28.85	30.09	27.37	25.44	21.92	17.60	12.95	10.56	20.56
Tucson	12.38	15.90	20.21	25.44	28.39	29.30	25.44	24.08	21.58	17.94	13.63	11.24	20.44
Arkansas													
Little Rock	9.09	11.81	15.56	19.19	21.80	23.51	23.17	21.35	17.26	14.08	9.77	8.06	16.24
Fort Smith	9.31	12.15	15.67	19.31	21.69	23.39	23.85	24.46	17.26	13.97	9.88	8.29	16.35
California													
Bakersfield	8.29	11.92	16.69	22.15	26.57	28.96	28.73	26.01	21.35	15.90	10.33	7.61	18.74
Fresno	7.61	11.58	16.81	22.49	27.14	29.07	28.96	25.89	21.12	15.56	9.65	6.70	18.62
Long Beach	9.99	12.95	17.03	21.60	23.17	24.19	26.12	24.08	19.31	14.99	11.24	9.31	17.83
Sacramento	6.93	10.68	15.56	21.24	25.89	28.28	28.62	25.32	20.56	14.54	8.63	6.25	17.72
San Diego	11.02	13.97	17.72	21.92	22.49	23.28	24.98	23.51	19.53	15.79	12.26	10.22	18.06
San Francisco	7.72	10.68	15.22	20.44	24.08	25.78	26.46	23.39	19.31	13.97	8.97	7.04	16.92
Los Angeles	10.11	13.06	17.26	21.80	23.05	23.74	25.67	23.51	18.97	14.99	11.36	9.31	17.72
Santa Maria	10.22	13.29	17.49	22.26	25.10	26.57	26.91	24.42	20.10	15.67	11.47	9.54	18.62
Colorado													
Boulder	7.84	10.45	15.64	17.94	17.94	20.47	20.28	17.12	16.07	12.09	8.66	7.10	14.31
Colorado Springs	9.09	12.15	16.13	20.33	22.26	24.98	23.96	21.69	18.51	14.42	9.99	8.18	16.81
Connecticut													
Hartford	6.70	9.65	13.17	16.69	19.53	21.24	21.12	18.51	14.76	10.68	6.59	5.45	13.74
Delaware													
Wilmington	7.27	10.22	13.97	17.60	20.33	22.49	21.80	19.65	15.79	11.81	7.84	6.25	14.65
Florida													
Daytona Beach	11.24	13.85	17.94	22.15	23.17	22.03	21.69	20.44	17.72	14.99	12.15	10.33	17.38

(continued)

TABLE A2.3b *(Continued)*

Position	January	February	March	April	May	June	July	August	September	October	November	December	Average
Jacksonville	10.45	13.17	17.03	21.12	22.03	21.58	21.01	19.42	16.69	14.20	11.47	9.65	16.47
Tallahassee	10.33	13.29	16.92	21.24	22.49	22.03	20.90	19.65	17.72	15.56	11.92	9.77	16.81
Miami	12.72	15.22	18.51	21.58	21.46	20.10	21.10	20.10	17.60	15.67	13.17	11.81	17.38
Key West	13.17	16.01	19.65	22.71	22.83	22.03	22.03	21.01	18.74	16.47	13.85	15.79	18.40
Tampa	11.58	14.42	18.17	22.26	23.05	21.92	20.90	19.65	17.60	16.01	12.83	11.02	17.49
Georgia													
Athens	9.43	12.38	16.01	20.21	22.03	22.83	21.80	20.21	17.26	14.42	10.45	8.40	16.29
Atlanta	9.31	12.26	16.13	20.33	22.37	23.17	22.15	20.56	17.49	14.54	10.56	8.52	16.43
Columbus	9.77	12.72	16.47	20.67	22.37	22.83	21.58	20.33	17.60	14.99	11.02	9.09	16.62
Macon	9.54	12.61	16.35	20.56	22.37	22.83	21.58	20.21	17.26	14.88	10.90	8.86	16.50
Savanna	9.99	12.72	16.81	21.01	22.37	22.60	21.80	19.76	16.92	14.65	11.13	9.20	16.58
Hawaii													
Honolulu	14.08	16.92	19.42	21.24	22.83	23.51	23.74	23.28	21.35	18.06	14.88	13.40	19.42
Idaho													
Boise	5.79	8.97	13.63	18.97	23.51	26.01	27.37	23.62	18.40	12.26	6.70	5.11	15.90
Illinois													
Chicago	6.47	9.31	12.49	16.47	20.44	22.60	22.03	19.31	15.10	10.79	6.47	5.22	13.85
Rockford	6.70	9.77	12.72	16.58	20.33	22.49	22.15	19.42	15.22	10.79	6.59	5.34	14.08
Springfield	7.50	10.33	13.40	17.83	21.46	23.51	23.05	20.56	16.58	12.26	7.72	6.13	15.10
Indiana													
Indianapolis	7.04	9.99	13.17	17.49	21.24	23.28	22.60	20.33	16.35	11.92	7.38	5.79	14.76
Iowa													
Mason City	6.70	9.77	13.29	16.92	20.78	22.83	22.71	19.76	15.33	10.90	6.59	5.45	14.31
Waterloo	6.81	9.77	13.06	16.92	20.56	22.83	22.60	19.76	15.33	10.90	6.70	5.45	14.20
Kansas													
Dodge City	9.65	12.83	16.69	21.01	23.28	25.78	25.67	22.60	18.40	14.42	10.11	8.40	17.49
Goodland	8.97	11.92	16.13	20.44	22.71	25.78	25.55	22.60	18.28	14.08	9.65	7.84	17.03
Kentucky													
Lexington	7.27	9.88	13.51	17.60	20.56	22.26	21.46	19.65	16.01	12.38	7.95	6.25	14.54
Louisville	7.27	10.22	13.63	17.83	20.90	22.71	22.03	20.10	16.35	12.38	7.95	6.25	14.76
Louisiana													
New Orleans	9.77	12.83	16.01	19.87	21.80	22.03	20.67	19.65	17.60	15.56	11.24	9.31	16.35
Lake Charles	9.77	12.83	16.13	19.31	21.58	22.71	21.58	20.33	18.06	15.56	11.47	9.31	16.58
Maine													
Portland	6.70	9.99	13.78	16.92	19.99	21.92	21.69	19.31	15.22	10.56	6.47	5.45	13.97
Maryland													

Baltimore	7.38	10.33	13.97	17.60	20.21	22.15	21.69	19.19	15.79	11.92	8.06	6.36	14.54
Massachusetts													
Boston	6.70	9.65	13.40	16.92	20.21	22.03	21.80	19.31	15.33	10.79	6.81	5.45	14.08
Michigan													
Detroit	5.91	8.86	12.38	16.47	20.33	22.37	21.92	18.97	14.76	10.11	6.13	4.66	13.63
Lansing	5.91	8.86	12.49	16.58	20.21	22.26	21.92	18.85	14.54	9.77	5.91	4.66	13.51
Minnesota													
Duluth	5.68	9.31	13.74	17.38	20.10	21.46	21.80	18.28	13.29	8.86	5.34	4.43	13.29
Minneapolis	6.36	9.77	13.51	16.92	20.56	22.49	22.83	19.42	14.65	9.99	6.13	4.88	13.97
Rochester	6.36	9.65	13.17	16.58	20.10	22.15	22.15	19.08	14.54	10.11	6.25	5.11	13.74
Mississippi													
Jackson	9.43	12.38	16.13	19.87	22.15	23.05	22.15	19.08	14.54	10.11	6.25	5.11	13.74
Missouri													
Columbia	8.06	10.90	14.31	18.62	21.58	23.62	23.85	21.12	16.69	12.72	8.29	6.70	15.56
Kansas City	7.95	10.68	14.08	18.28	21.24	23.28	23.62	20.78	16.58	12.72	8.40	6.70	15.44
Springfield	8.52	11.02	14.65	18.62	21.24	23.05	23.62	21.24	16.81	13.17	8.86	7.27	15.67
St. Louis	7.84	10.56	13.97	18.06	21.12	23.05	22.94	20.44	16.58	12.49	8.18	6.59	15.22
Montana													
Helena	5.22	8.29	12.61	17.15	20.67	23.28	25.21	21.24	15.79	10.45	6.02	4.43	14.20
Lewistown	5.22	8.40	12.72	17.15	20.33	23.05	24.53	20.78	15.10	10.22	5.91	4.32	13.97
Nebraska													
Omaha	7.50	10.33	13.97	18.06	21.24	2.40	23.51	20.56	16.01	11.81	7.61	6.13	15.10
Lincoln	7.33	10.10	13.65	16.22	19.26	21.21	22.15	18.87	15.44	11.54	7.76	6.20	14.16
Nevada													
Elko	7.61	10.56	14.42	18.85	22.71	25.67	26.69	23.62	19.31	13.63	8.29	6.70	16.58
Las Vegas	10.79	14.42	19.42	24.87	28.16	30.09	28.28	25.89	22.15	17.03	12.15	9.88	20.33
Reno	8.29	11.58	16.24	21.24	25.10	27.48	28.16	24.98	20.56	14.88	9.31	7.38	17.94
New Hampshire													
Concord	6.81	10.11	13.97	16.92	20.21	21.80	21.80	19.08	14.99	10.45	6.47	5.45	14.08
New Jersey													
Atlantic City	7.38	10.22	13.97	17.49	20.21	21.92	21.24	19.19	15.79	11.92	8.06	6.36	14.54
Newark	6.93	9.77	13.51	17.26	19.76	21.35	21.01	18.85	15.33	11.36	7.27	5.68	13.97
New Mexico													
Albuquerque	11.47	14.99	19.31	24.53	27.60	29.07	27.03	24.76	21.12	17.03	12.49	10.33	19.99
New York													
Albany	6.36	9.43	12.95	16.69	19.53	21.46	21.58	18.51	14.65	10.11	6.13	5.00	13.51
Buffalo	5.68	8.40	12.15	16.35	19.76	22.03	21.69	18.62	14.08	9.54	5.68	4.54	13.29
New York City	6.93	9.88	13.85	17.72	20.44	22.03	21.69	19.42	15.56	11.47	7.27	5.79	14.31

(continued)

TABLE A2.3b ***(Continued)***

Position	January	February	March	April	May	June	July	August	September	October	November	December	Average
Rochester	5.68	8.52	12.26	16.58	19.87	21.92	21.69	18.51	14.20	9.54	5.68	4.54	13.29
North Carolina													
Charlotte	8.97	11.81	15.67	19.76	21.58	22.60	21.92	19.99	16.92	13.97	9.99	8.06	16.01
Wilmington	9.31	12.15	16.24	20.44	21.92	22.60	21.58	19.53	16.69	14.08	10.56	8.52	16.13
North Dakota													
Fargo	5.79	9.09	13.17	16.92	20.56	22.37	23.17	19.87	14.31	9.54	5.68	4.54	13.74
Bismarck	6.12	9.75	13.88	17.43	21.45	23.01	24.06	20.12	15.21	10.61	6.28	4.84	14.39
Ohio													
Cleveland	5.79	8.63	12.04	16.58	20.10	22.15	21.92	18.97	14.76	10.22	6.02	4.66	13.51
Columbus	6.47	9.09	12.49	16.58	19.76	21.58	21.12	18.97	15.44	11.24	6.81	5.34	13.74
Dayton	6.81	9.43	12.83	17.03	20.33	22.37	22.37	19.65	15.90	11.47	7.04	5.45	14.20
Youngstown	5.79	8.40	11.92	15.90	19.19	21.24	20.78	18.06	14.31	10.11	6.02	4.77	13.06
Oklahoma													
Oklahoma City	9.88	1.25	16.47	20.33	22.26	24.42	24.98	22.49	18.17	14.54	10.45	8.74	17.15
Oregon													
Eugene	4.54	7.04	11.24	15.79	19.99	22.37	24.19	21.01	15.90	9.65	5.11	3.75	13.40
Medford	5.34	8.52	13.17	18.62	23.39	26.23	27.82	23.96	18.62	11.92	6.02	4.43	15.67
Portland	4.20	6.70	10.68	15.10	18.97	21.24	22.60	19.53	14.88	9.20	4.88	3.52	12.61
Pacific Islands													
Guam	16.35	17.38	19.65	20.78	20.56	19.76	18.28	17.49	17.49	16.58	15.79	15.10	17.94
Pennsylvania													
Philadelphia	7.04	9.88	13.63	17.26	19.99	22.03	21.46	19.42	15.67	11.58	7.72	6.02	14.31
Pittsburgh	6.25	8.97	12.61	16.47	19.65	21.80	21.35	18.85	15.10	10.90	6.59	5.00	13.63
Rhode Island													
Providence	6.70	9.65	13.40	16.92	19.99	21.58	21.24	18.85	15.22	11.02	6.93	5.56	13.97
South Carolina													
Charleston	9.77	12.72	16.81	21.12	22.37	22.37	21.92	19.65	16.92	14.54	11.02	9.09	16.58
Greenville	9.20	12.04	15.90	19.99	21.58	22.60	21.58	19.87	16.81	14.08	10.22	8.18	16.01
South Dakota													
Pierre	6.47	9.54	13.85	17.94	21.46	24.08	24.42	21.46	16.35	11.24	7.04	5.45	14.99
Rapid City	6.70	9.88	14.20	18.28	21.46	24.19	24.42	21.80	16.92	11.81	7.50	5.79	15.33
Tennessee													
Memphis	8.86	11.58	15.22	19.42	22.03	23.85	23.39	21.46	17.38	14.20	9.65	7.84	16.24
Nashville	8.29	11.13	14.65	19.31	21.69	23.51	22.49	20.56	16.81	13.51	8.97	7.15	15.67
Texas													
Austin	10.68	13.63	17.03	19.53	21.24	23.74	24.42	22.83	18.85	15.67	11.92	9.99	17.49

Brownsville	10.33	13.17	16.47	19.08	20.78	22.83	23.28	21.58	18.62	16.13	12.38	9.88	17.03
El Paso	12.38	16.24	20.90	25.44	28.05	28.85	26.46	24.30	21.12	17.72	13.63	11.47	20.56
Houston	9.54	12.26	15.22	18.06	20.21	21.69	21.35	20.21	17.49	15.10	11.02	8.97	15.90
San Antonio	10.88	13.53	16.26	17.35	21.10	23.87	24.92	22.81	19.22	15.52	11.50	9.98	17.24
Utah													
Salt Lake City	6.93	10.45	14.76	19.42	23.39	26.46	26.35	23.39	18.85	13.29	8.06	6.02	16.47
Vermont													
Burlington	5.79	9.20	13.06	16.47	19.87	21.69	21.80	18.74	14.42	9.43	5.56	4.43	13.40
Virginia													
Norfolk	8.06	10.90	14.65	18.51	20.78	22.15	21.12	19.42	16.13	12.49	9.09	7.27	15.10
Richmond	8.06	10.90	14.76	18.62	20.90	22.49	21.58	19.53	16.24	12.61	8.97	7.15	15.22
Washington													
Olympia	3.63	6.02	9.99	14.20	18.06	20.10	21.12	18.17	13.63	7.95	4.32	3.07	11.70
Seattle	3.52	5.91	10.11	14.65	19.08	20.78	21.80	18.51	13.51	7.95	4.20	2.84	11.92
Yakima	4.88	7.95	12.83	17.83	22.49	24.87	25.89	22.26	16.92	10.68	5.56	4.09	17.76
West Virginia													
Charleston	7.04	9.65	13.40	17.15	20.21	21.69	20.90	18.97	15.56	11.81	7.72	6.02	14.20
Elkins	6.93	9.43	12.83	16.35	19.08	20.56	19.99	18.06	14.88	11.13	7.27	5.79	13.51
Wisconsin													
Green Bay	6.25	9.31	13.17	16.81	20.56	22.49	22.03	18.85	14.20	9.65	5.79	4.88	13.74
Madison	6.59	9.88	13.29	16.92	20.67	22.83	22.37	19.42	14.76	3.41	6.25	5.22	14.08
Milwaukee	6.47	9.31	12.72	16.69	20.78	22.94	22.60	19.42	14.88	10.22	6.25	5.11	13.97
Wyoming													
Rock Springs	7.61	10.90	15.10	19.42	23.17	26.01	25.78	22.94	18.62	13.40	8.40	6.70	16.58
Sendan	6.47	9.77	13.97	17.94	20.90	23.85	24.64	21.69	16.47	11.24	7.15	5.56	14.99

Source: From National Renewable Energy Laboratory, U.S.A.; Internet Address: http://rredc.nrel.gov/solar.

TABLE A2.4 Reflectivity Values for Characteristic Surfaces (Integrated Over Solar Spectrum and Angle of Incidence)

Surface	Average Reflectivity
Snow (freshly fallen or with ice film)	0.75
Water surfaces (relatively large incidence angles)	0.07
Soils (clay, loam, etc.)	0.14
Earth roads	0.04
Coniferous forest (winter)	0.07
Forests in autumn, ripe field crops, plants	0.26
Weathered blacktop	0.10
Weathered concrete	0.22
Dead leaves	0.30
Dry grass	0.20
Green grass	0.26
Bituminous and gravel roof	0.13
Crushed rock surface	0.20
Building surfaces, dark (red brick, dark paints, etc.)	0.27
Building surfaces, light (light brick, light paints, etc.)	0.60

Source: From Hunn, B. D. and Calafell, D. O. 1977. *Solar Energy,* Vol. 19, p. 87; see also List, R. J. 1949. *Smithsonian Meteorological Tables*, 6th Ed., pp. 442–443. Smithsonian Institution Press.

A3

Properties of Gases, Vapors, Liquids and Solids

Nitin Goel
Intel Technology India Pvt. Ltd.

TABLE A3.1 Properties of Dry Air at Atmospheric Pressures between 250 and 1000 K

T^a (K)	ρ (kg/m^3)	c_p (kJ/kg K)	μ (kg/m s$\times 10^5$)	ν (m^2/s$\times 10^6$)	k (W/m K)	α (m^2/s$\times 10^4$)	dPr
250	1.4128	1.0053	1.488	9.49	0.02227	0.13161	0.722
300	1.1774	1.0057	1.983	15.68	0.02624	0.22160	0.708
350	0.9980	1.0090	2.075	20.76	0.03003	0.2983	0.697
400	0.8826	1.0140	2.286	25.90	0.03365	0.3760	0.689
450	0.7833	1.0207	2.484	28.86	0.03707	0.4222	0.683
500	0.7048	1.0295	2.671	37.90	0.04038	0.5564	0.680
550	0.6423	1.0392	2.848	44.34	0.04360	0.6532	0.680
600	0.5879	1.0551	3.018	51.34	0.04659	0.7512	0.680
650	0.5430	1.0635	3.177	58.51	0.04953	0.8578	0.682
700	0.5030	1.0752	3.332	66.25	0.05230	0.9672	0.684
750	0.4709	1.0856	3.481	73.91	0.05509	1.0774	0.686
800	0.4405	1.0978	3.625	82.29	0.05779	1.1951	0.689
850	0.4149	1.1095	3.765	90.75	0.06028	1.3097	0.692
900	0.3925	1.1212	3.899	99.3	0.06279	1.4271	0.696
950	0.3716	1.1321	4.023	108.2	0.06525	1.5510	0.699
1000	0.3524	1.1417	4.152	117.8	0.06752	1.6779	0.702

[a] Symbols: K = absolute temperature, degrees Kelvin; $\nu = \mu/\rho$; ρ = density; c_p specific heat capacity; $\alpha = c_p\rho/k$; μ = viscosity; k = thermal conductivity; Pr = Prandtl number, dimensionless. The values of μ, k, c_p, and Pr are not strongly pressure-dependent and may be used over a fairly wide range of pressures.

Source: From Natl. Bureau Standards (U.S.) Circ. 564, 1955.

TABLE A3.2 Properties of Water (Saturated Liquid) between 273 and 533 K

T (K)	T (°F)	T (°C)	c_p (kJ/kg °C)	ρ (kg/m^3)	μ (kg/m s)	k (W/m °C)	Pr	$(g\beta\rho^2 c_p/\mu k)(m^{-3}{}^\circ C^{-1})$
273	32	0	4.225	999.8	1.79×10^{-3}	0.566	13.25	
277.4	40	4.44	4.208	999.8	1.55	0.575	11.35	1.91×10^{9}
283	50	10	4.195	999.2	1.31	0.585	9.40	6.34×10^{9}
288.6	60	15.56	4.186	998.6	1.12	0.595	7.88	1.08×10^{10}
294.1	70	21.11	4.179	997.4	9.8×10^{-4}	0.604	6.78	1.46×10^{10}
299.7	80	26.67	4.179	995.8	8.6	0.614	5.85	1.91×10^{10}
302.2	90	32.22	4.174	994.9	7.65	0.623	5.12	2.48×10^{10}
310.8	100	37.78	4.174	993.0	6.82	0.630	4.53	3.3×10^{10}
316.3	110	43.33	4.174	990.6	6.16	0.637	4.04	4.19×10^{10}
322.9	120	48.89	4.174	988.8	5.62	0.644	3.64	4.89×10^{10}
327.4	130	54.44	4.179	985.7	5.13	0.649	3.30	5.66×10^{10}
333.0	140	60	4.179	983.3	4.71	0.654	3.01	6.48×10^{10}
338.6	150	65.55	4.183	980.3	4.3	0.659	2.73	7.62×10^{10}
342.1	160	71.11	4.186	977.3	4.01	0.665	2.53	8.84×10^{10}
349.7	170	76.67	4.191	973.7	3.72	0.668	2.33	9.85×10^{10}
355.2	180	82.22	4.195	970.2	3.47	0.673	2.16	1.09×10^{11}
360.8	190	87.78	4.199	966.7	3.27	0.675	2.03	
366.3	200	93.33	4.204	963.2	3.06	0.678	1.90	
377.4	220	104.4	4.216	955.1	2.67	0.684	1.66	
388.6	240	115.6	4.229	946.7	2.44	0.685	1.51	
399.7	260	126.7	4.250	937.2	2.19	0.685	1.36	
410.8	280	137.8	4.271	928.1	1.98	0.685	1.24	
421.9	300	148.9	4.296	918.0	1.86	0.684	1.17	
449.7	350	176.7	4.371	890.4	1.57	0.677	1.02	
477.4	400	204.4	4.467	859.4	1.36	0.665	1.00	
505.2	450	232.2	4.585	825.7	1.20	0.646	0.85	
533.0	500	260	4.731	785.2	1.07	0.616	0.83	

Source: Adapted from Brown, A. I. and S. M. Marco. 1958. Introduction to Heat Transfer, 3d Ed., McGraw-Hill Book Company, New York.

TABLE A3.3 Emittances and Absorptances of Materials

Substance	Short-Wave Absorptance	Long-Wave Emittance	a ε
Class I substances: Absorptance to emittance ratios less than 0.5			
Magnesium carbonate, $MgCO_3$	0.025–0.04	0.79	0.03–0.05
White plaster	0.07	0.91	0.08
Snow, fine particles, fresh	0.13	0.82	0.16
White paint, 0.017 in. on aluminum	0.20	0.91	0.22
Whitewash on galvanized iron	0.22	0.90	0.24
White paper	0.25–0.28	0.95	0.26–0.29
White enamel on iron	0.25–0.45	0.9	0.28–0.5
Ice, with sparse snow cover	0.31	0.96–0.97	0.32
Snow, ice granules	0.33	0.89	0.37
Aluminum oil base paint	0.45	0.90	0.50
White powdered sand	0.45	0.84	0.54
Class II substances: Absorptance to emittance ratios between 0.5 and 0.9			
Asbestos felt	0.25	0.50	0.50
Green oil base paint	0.5	0.9	0.56
Bricks, red	0.55	0.92	0.60
Asbestos cement board, white	0.59	0.96	0.61
Marble, polished	0.5–0.6	0.9	0.61
Wood, planed oak	–	0.9	–
Rough concrete	0.60	0.97	0.62
Concrete	0.60	0.88	0.68
Grass, green, after rain	0.67	0.98	0.68
Grass, high and dry	0.67–0.69	0.9	0.76
Vegetable fields and shrubs, wilted	0.70	0.9	0.78
Oak leaves	0.71–0.78	0.91–0.95	0.78–0.82
Frozen soil	–	0.93–0.94	–
Desert surface	0.75	0.9	0.83
Common vegetable fields and shrubs	0.72–0.76	0.9	0.82
Ground, dry plowed	0.75–0.80	0.9	0.83–0.89
Oak woodland	0.82	0.9	0.91
Pine forest	0.86	0.9	0.96
Earth surface as a whole (land and sea, no clouds)	0.83	10^{10}	–
Class III substances: Absorptance to emittance ratios between 0.8 and 1.0			
Grey paint	0.75	0.95	0.79
Red oil base paint	0.74	0.90	0.82
Asbestos, slate	0.81	0.96	0.84
Asbestos, paper		0.93–0.96	
Linoleum, red–brown	0.84	0.92	0.91
Dry sand	0.82	0.90	0.91
Green roll roofing	0.88	0.91–0.97	0.93
Slate, dark grey	0.89	–	
Old grey rubber		0.86	–
Hard black rubber	–	0.90–0.95	
Asphalt pavement	0.93	–	–
Black cupric oxide on copper	0.91	0.96	0.95
Bare moist ground	0.9	0.95	0.95
Wet sand	0.91	0.95	0.96
Water	0.94	0.95–0.96	0.98
Black tar paper	0.93	0.93	1.0
Black gloss paint	0.90	0.90	1.0
Small hole in large box, furnace, or enclosure	0.99	0.99	1.0
"Hohlraum," theoretically perfect black body	1.0	1.0	1.0
Class IV substances: Absorptance to emittance ratios greater than 1.0			
Black silk velvet	0.99	0.97	1.02

(continued)

TABLE A3.3 ***(Continued)***

Substance	Short-Wave Absorptance	Long-Wave Emittance	a ε
Alfalfa, dark green	0.97	0.95	1.02
Lampblack	0.98	0.95	1.03
Black paint, 0.017 in. on aluminum	0.94–0.98	0.88	1.07–1.11
Granite	0.55	0.44	1.25
Graphite	0.78	0.41	1.90
High ratios, but absorptances less than 0.80			
Dull brass, copper, lead	0.2–0.4	0.4–0.65	1.63–2.0
Galvanized sheet iron, oxidized	0.8	0.28	2.86
Galvanized iron, clean, new	0.65	0.13	5.0
Aluminum foil	0.15	0.05	3.00
Magnesium	0.3	0.07	4.3
Chromium	0.49	0.08	6.13
Polished zinc	0.46	0.02	23.0
Deposited silver (optical reflector) untarnished	0.07	0.01	
Class V substances: Selective surfaces[a]			
Plated metals:[b]			
Black sulfide on metal	0.92	0.10	9.2
Black cupric oxide on sheet aluminum	0.08–0.93	0.09–0.21	
Copper ($5\times10\sim^{5}$ cm thick) on nickel or silver-plated metal			
Cobalt oxide on platinum			
Cobalt oxide on polished nickel	0.93–0.94	0.24–0.40	3.9
Black nickel oxide on aluminum	0.85–0.93	0.06–0.1	14.5–15.5
Black chrome	0.87	0.09	9.8
Particulate coatings:			
Lampblack on metal			
Black iron oxide, 47 µm grain size, on aluminum			
Geometrically enhanced surfaces:[c]			
Optimally corrugated greys	0.89	0.77	1.2
Optimally corrugated selectives	0.95	0.16	5.9
Stainless-steel wire mesh	0.63–0.86	0.23–0.28	2.7–3.0
Copper, treated with NaClO, and NaOH	0.87	0.13	6.69

[a] Selective surfaces absorb most of the solar radiation between 0.3 and 1.9 µm, and emit very little in the 5–15 µm range–the infrared.

[b] For a discussion of plated selective surfaces, see Daniels, *Direct Use of the Sun's Energy*, especially chapter 12.

[c] For a discussion of how surface selectivity can be enhanced through surface geometry, see K. G. T. Hollands, July 1963. Directional selectivity emittance and absorptance properties of vee corrugated specular surfaces, *J. Sol. Energy Sci. Eng, vol. 3.*

Source: From Anderson, B. 1977. *Solar Energy*, McGraw-Hill Book Company. With permission.

TABLE A3.4 Thermal Properties of Metals and Alloys

Material	k, Btu/(hr)(ft.)(°F)				c, Btu/(lb_m)(°F)	ρ, lb_m/ft.^{3}	α, ft.^{2}/hr
	32°F	212°F	572°F	932°F	32°F	32°F	32°F
Metals							
Aluminum	117	119	133	155	0.208	169	3.33
Bismuth	4.9	3.9	…	…	0.029	612	0.28
Copper, pure	224	218	212	207	0.091	558	4.42
Gold	169	170	…	…	0.030	1,203	4.68
Iron, pure	35.8	36.6	…	…	0.104	491	0.70
Lead	20.1	19	18	…	0.030	705	0.95
Magnesium	91	92	…	…	0.232	109	3.60
Mercury	4.8	…	…	…	0.033	849	0.17
Nickel	34.5	34	32	…	0.103	555	0.60
Silver	242	238	…	…	0.056	655	6.6
Tin	36	34	…	…	0.054	456	1.46
Zinc	65	64	59	…	0.091	446	1.60
Alloys							
Admiralty metal	65	64					
Brass, 70% Cu, 30% Zn	56	60	66	…	0.092	532	1.14
Bronze, 75% Cu, 25% Sn	15	…	…	…	0.082	540	0.34
Cast iron							
Plain	33	31.8	27.7	24.8	0.11	474	0.63
Alloy	30	28.3	27		0.10	455	0.66
Constantan, 60% Cu, 40% Ni	12.4	12.8			0.10	557	0.22
18-8 Stainless steel,							
Type 304	8.0	9.4	10.9	12.4	0.11	488	0.15
Type 347	8.0	9.3	11.0	12.8	0.11	488	0.15
Steel, mild, 1% C	26.5	26	25	22	0.11	490	0.49

Source: From Kreith, F. 1997. *Principles of Heat Transfer*, PWS Publishing Co., Boston.

TABLE A3.5 Thermal Properties of Some Insulating and Building Materials

Material	Average, Temperature, °F	k, Btu/(hr)(ft.)(°F)	c, Btu/(lb$_m$)(°F)	ρ, lb$_m$/ft.^{3}	a, ft.^{2}/hr
		Insulating Materials			
Asbestos	32	0.087	0.25	36	−0.01
	392	0.12		36	−0.01
Cork	86	0.025	0.04	10	−0.006
Cotton, fabric	200	0.046			
Diatomaceous earth, powdered	100	0.030	0.21	14	−0.01
	300	0.036	…		
	600	0.046	…		
Molded pipe covering	400	0.051	…	26	
	1600	0.088	…		
		Glass Wool			
Fine	20	0.022	…		
	100	0.031	…	1.5	
	200	0.043	…		
Packed	20	0.016	…		
	100	0.022	…	6.0	
	200	0.029	…		
Hair felt	100	0.027	…	8.2	
Kaolin insulating brick	932	0.15	…	27	
	2102	0.26	…		
Kaolin insulating firebrick	392	0.05	…	19	
	1400	0.11	…		
85% magnesia	32	0.032	…	17	
	200	0.037	…	17	
Rock wool	20	0.017	…	8	
	200	0.030	…		
Rubber	32	0.087	0.48	75	0.0024
		Building Materials			
Brick					
Fire-clay	392	0.58	0.20	144	0.02
	1832	0.95			
Masonry	70	0.38	0.20	106	0.018
Zirconia	392	0.84	…	304	
	1832	1.13	…		
Chrome brick	392	0.82	…	246	
	1832	0.96	…		
		Concrete			
Stone	−70	0.54	0.20	144	0.019
10% Moisture	−70	0.70	…	140	−0.025
Glass, window	−70	−0.45	0.2	170	0.013
Limestone, dry	70	0.40	0.22	105	0.017
		Sand			
Dry	68	0.20	…	95	
10% H_2O	68	0.60		100	
		Soil			
Dry	70	−0.20	0.44	…	−0.01
Wet	70	−1.5	…		−0.03
		Wood			
Oak ⊥ to grain	70	0.12	0.57	51	0.0041
‖ to grain	70	0.20	0.57	51	0.0069
Pine ⊥ to grain	70	0.06	0.67	31	0.0029
‖ to grain	70	0.14	0.67	31	0.0067
Ice	32	1.28	0.46	57	0.048

Source: From Kreith, R. 1997. *Principles of Heat Transfer*, PWS Publishing Co.

TABLE A3.6 Saturated Steam and Water–SI Units

Temperature (K)	Pressure (MN/m^2)	Specific Volume (m^3/kg)		Specific Energy Internal (kJ/kg)		Specific Enthalpy (kJ/kg)			Specific Entropy (kJ/kg.K)	
		v_f	v_g	u_f	u_g	h_f	h_{fg}	h_g	s_f	s_g
273.15	0.0006109	0.0010002	206.278	−0.03	2375.3	−0.02	2501.4	2501.3	−0.0001	9.1565
273.16	0.0006113	0.0010002	206.136	0	2375.3	+0.01	2501.3	2501.4	0	9.1562
278.15	0.0008721	0.0010001	147.120	+20.97	2382.3	20.98	2489.6	2510.6	+0.0761	9.0257
280.13	0.0010000	0.0010002	129.208	29.30	2385.0	29.30	2484.9	2514.2	0.1059	8.975
283.15	0.0012276	0.0010004	106.379	42.00	2389.2	42.01	2477.7	2519.8	0.1510	8.9008
286.18	0.0015000	0.0010007	87.980	54.71	2393.3	54.71	2470.6	2525.3	0.1957	8.8279
288.15	0.0017051	0.0010009	77.926	62.99	2396.1	62.99	2465.9	2528.9	0.2245	8.7814
290.65	0.0020000	0.0010013	67.004	73.48	2399.5	73.48	2460.0	2533.5	0.2607	8.7237
293.15	0.002339	0.0010018	57.791	83.95	2402.9	83.96	2454.1	2538.1	0.2966	8.6672
297.23	0.0030000	0.0010027	45.665	101.04	2408.5	101.05	2444.5	2545.5	0.3545	8.5776
298.15	0.003169	0.0010029	43.360	104.88	2409.8	104.89	2442.3	2547.2	0.3674	8.5580
302.11	0.004000	0.0010040	34.800	121.45	2415.2	121.46	2432.9	2554.4	0.4226	8.4746
303.15	0.004246	0.0010043	32.894	125.78	2416.6	125.79	2430.5	2556.3	0.4369	8.4533
306.03	0.005000	0.0010053	28.192	137.81	2420.5	137.82	2423.7	2561.5	0.4764	8.3951
308.15	0.005628	0.0010060	25.216	146.67	2423.4	146.68	2418.6	2565.3	0.5053	8.3531
309.31	0.006000	0.0010064	23.739	151.53	2425.0	151.53	2415.9	2567.4	0.5210	8.3304
312.15	0.007000	0.0010074	20.530	163.39	2428.8	163.40	2409.1	2572.5	0.5592	8.2758
313.15	0.007384	0.0010078	19.523	167.56	2430.1	167.57	2406.7	2574.3	0.5725	8.2570
314.66	0.008000	0.0010084	18.103	173.87	2432.2	173.88	2403.1	2577.0	0.5926	8.2287
316.91	0.009000	0.0010094	16.203	183.27	2435.2	183.29	2397.7	2581.0	0.6224	8.1872
318.15	0.009593	0.0010099	15.258	188.44	2436.8	188.45	2394.8	2583.2	0.6387	8.1648
318.96	0.010000	0.0010102	14.674	191.82	2437.9	191.83	2392.8	2584.7	0.6493	8.1502
323.15	0.012349	0.0010121	12.032	209.32	2443.5	209.33	2382.7	2592.1	0.7038	8.0763
327.12	0.015000	0.0010141	10.022	225.92	2448.7	225.94	2373.1	2599.1	0.7549	8.0085
328.15	0.015758	0.0010146	9.568	230.21	2450.1	230.23	2370.7	2600.9	0.7679	7.9913
333.15	0.019940	0.0010172	7.671	251.11	2456.6	251.13	2358.5	2609.6	0.8312	7.9096
333.21	0.020000	0.0010172	7.649	251.38	2456.7	251.40	2358.3	2609.7	0.8320	7.9085
338.15	0.025030	0.0010199	6.197	272.02	2463.1	272.06	2346.2	2618.3	0.8935	7.8310
342.25	0.030000	0.0010223	5.229	289.20	2468.4	289.23	2336.1	2625.3	0.9439	7.7686
343.15	0.031190	0.0010228	5.042	292.95	2469.6	292.98	2333.8	2626.8	0.9549	7.7553
348.15	0.038580	0.0010259	4.131	313.90	2475.9	313.93	2221.4	2635.3	1.0155	7.6824
349.02	0.040000	0.0010265	3.993	317.53	2477.0	317.58	2319.2	2636.8	1.0259	7.6700
353.15	0.047390	0.0010291	3.407	334.86	2482.2	334.91	2308.8	2643.7	1.0753	7.6122
354.48	0.050000	0.0010300	3.240	340.44	2483.9	340.49	2305.4	2645.9	1.0910	7.5939

358.15	0.057830	0.0010325	2.828	355.84	2488.4	355.90	2296.0	2651.9	1.1343	7.5445
359.09	0.060000	0.0010331	2.732	359.79	2489.6	359.86	2293.6	2653.5	1.1453	7.5320
363.10	0.070000	0.0010360	2.365	376.63	2494.5	376.70	2283.3	2660.0	1.1919	7.4797
363.15	0.070140	0.0010360	2.361	376.85	2494.5	376.92	2283.2	2660.1	1.1925	7.4791
366.65	0.080000	0.0010386	2.087	391.58	2498.8	391.66	2274.1	2665.8	1.2329	7.4346
368.15	0.084550	0.0010397	1.9819	397.88	2500.6	397.96	2270.2	2668.1	1.2500	7.4159

Subscripts: *f* refers to a property of liquid in equilibrium with vapor; *g* refers to a property of vapor in equilibrium with liquid; *fg* refers to a change by evaporation. Table from Bolz, R. E. and G. L. Tuve, eds. 1973. *CRC Handbook of Tables for Applied Engineering Science*, 2nd Ed., Chemical Rubber Co., Cleveland, Ohio.

TABLE A3.7 Superheated Steam–SI Units

Pressure (MN/m^2) (Saturation Temperature)[a]		Temperature									
		50°C 323.15 K	100°C 373.15 K	150°C 423.15 K	200°C 473.15 K	300°C 573.15 K	400°C 673.15 K	500°C 773.15 K	700°C 973.15 K	1000°C 1273.15 K	1300°C 1573.15 K
0.001	*v*	149.093	172.187	195.272	218.352	264.508	310.661	356.814	449.117	587.571	726.025
(6.98°C)	*u*	2445.4	2516.4	2588.4	2661.6	2812.2	2969.0	3132.4	3479.6	4053.0	4683.7
(280.13 K)	*h*	2594.5	2688.6	2783.6	2880.0	3076.8	3279.7	3489.2	3928.7	4640.6	5409.7
	s	9.2423	9.5129	9.7520	9.9671	10.3443	10.6705	10.9605	11.4655	12.1019	12.6438
0.002	*v*	74.524	86.081	97.628	109.170	132.251	155.329	178.405	224.558	293.785	363.012
(17.50°C)	*u*	2445.2	2516.3	2588.3	2661.6	2812.2	2969.0	3132.4	3479.6	4053.0	4683.7
(290.65 K)	*h*	2594.3	2688.4	2793.6	2879.9	3076.7	3279.7	3489.2	3928.7	4640.6	5409.7
	s	8.9219	9.1928	9.4320	9.6471	10.0243	10.3506	10.6406	11.1456	11.7820	12.3239
0.004	*v*	37.240	43.028	48.806	54.580	66.122	77.662	89.201	112.278	146.892	181.506
(28.%°C)	*u*	2444.9	2516.1	2588.2	2661.5	2812.2	2969.0	3132.3	3479.6	4053.0	4683.7
(302.11 K)	*h*	2593.9	2688.2	2783.4	2879.8	3076.7	3279.6	3489.2	3928.7	4640.6	5409.7
	s	8.6009	8.8724	9.1118	9.3271	9.7044	10.0307	10.3207	10.8257	11.4621	12.0040
0.006	*v*	24.812	28.676	32.532	36.383	44.079	51.774	59.467	74.852	97.928	121.004
(36.16°C)	*u*	2444.6	2515.9	2588.1	2661.4	2812.2	2969.0	31323	3479.6	4053.0	4683.7
(309.31 K)	*h*	2593.4	2688.0	2783.3	2879.7	3076.6	3279.6	3489.1	3928.7	4640.6	5409.7
	s	8.4128	8.6847	8.9244	9.1398	9.5172	9.8435	10.1336	10.6386	11.2750	11.8168
0.008	*v*	18.598	21.501	24.395	27.284	33.058	38.829	44.599	56.138	73.446	90.753
(41.51°C)	*u*	2444.2	2515.7	2588.0	2661.4	2812.1	2969.0	3132.3	3479.6	4053.0	4683.7
(314.66 K)	*h*	2593.0	2687.7	2783.1	2879.6	3076.6	3279.6	3489.1	3928.7	4640.6	5409.7
	s	8.2790	8.5514	8.7914	9.0069	9.3844	9.7107	10.0008	10.5058	11.1422	11.6841
0.010	*v*	14.869	17.196	19.512	21.825	26.445	31.063	35.679	44.911	58.757	72.602
(45.81°C)	*u*	2443.9	2515.5	2587.9	2661.3	2812.1	2968.9	3132.3	3479.6	4053.0	4683.7
(318.96 K)	*h*	2592.6	2687.5	2783.0	2879.5	3076.5	3279.6	3489.1	3928.7	4640.6	5409.7
	s	8.1749	8.4479	8.6882	8.9038	9.2813	9.6077	9.8978	10.4028	11.0393	11.5811
0.020	*v*	7.412	8.585	9.748	10.907	13.219	15.529	17.838	22.455	29.378	36.301
(60.06°C)	*u*	2442.2	2514.6	2587.3	2660.9	2811.9	2968.8	3132.2	3479.5	4053.0	4683.7
(333.21 K)	*h*	2590.4	2686.2	2782.3	2879.1	3076.3	3279.4	3489.0	3928.6	4640.6	5409.7
	s	7.8498	8.1255	8.3669	8.5831	8.9611	9.2876	9.5778	10.0829	10.7193	11.2612
0.040	*v*	3.683	4.279	4.866	5.448	6.606	7.763	8.918	11.227	14.689	18.151
(75.87°C)	*u*	2438.8	2512.6	2586.2	2660.2	2811.5	2968.6	3132.1	3479.4	4052.9	4683.6
(349.02 K)	*h*	2586.1	2683.8	2780.8	2878.1	3075.8	3279.1	3488.8	3928.5	4640.5	5409.6
	s	7.5192	7.8003	8.0444	8.2617	8.6406	8.9674	9.2577	9.7629	10.3994	10.9412
0.060	*v*	2.440	2.844	3.238	3.628	4.402	5.174	5.944	7.484	9.792	12.100
(85.94°C)	*u*	2435.3	2510.6	2585.1	2659.5	2811.2	2968.4	3131.9	3479.4	4052.9	4683.6
(359.09 K)	*h*	2581.7	2681.3	2779.4	2877.2	3075.3	3278.8	3488.6	3928.4	4640.4	5409.6
	s	7.3212	7.6079	7.8546	8.0731	8.4528	8.7799	9.0704	9.5757	10.2122	10.7541
0.080	*v*	1.8183	2.127	2.425	2.718	3.300	3.879	4.458	5.613	7.344	9.075
(93.50°C)	*u*	2431.7	2508.7	2583.9	2658.8	2810.8	2968.1	3131.7	3479.3	4052.8	4683.5
(366.65 K)	*h*	2577.2	2678.8	2777.9	2876.2	3074.8	3278.5	3488.3	3928.3	4640.4	5409.5
	s	7.1775	7.4698	7.7191	7.9388	8.3194	8.6468	8.9374	9.4428	10.0794	10.6213
0.100	*v*	1.4450	1.6958	1.9364	2.172	2.639	3.103	3.565	4.490	5.875	7.260

(99.63°C)	*u*	2428.2	2506.7	2582.8	2658.1	2810.4	2967.9	3131.6	3479.2	4052.8	4683.5
(372.78 K)	*h*	2572.7	2676.2	2776.4	2875.3	3074.3	3278.2	3488.1	3928.2	4640.3	5409.5
	s	7.0633	7.3614	7.6134	7.8343	8.2158	8.5435	8.8342	9.3398	9.9764	10.5183
0.200	*v*	0.6969	0.8340	0.9596	1.0803	1.3162	1.5493	1.7814	2.244	2.937	3.630
(120.23°C)	*u*	2409.5	2496.3	2576.9	2654.4	2808.6	2966.7	3130.8	3478.8	4052.5	4683.2
(393.38 K)	*h*	2548.9	2663.1	2768.8	2870.5	3071.8	3276.6	3487.1	3927.6	4640.0	5409.3
	s	6.6844	7.0135	7.2795	7.5066	7.8926	8.2218	8.5133	9.0194	9.6563	10.1982
0.300	*v*	0.4455	0.5461	0.6339	0.7163	0.8753	1.0315	1.1867	1.4957	1.9581	2.4201
(133.55°C)	*u*	2389.1	2485.4	2570.8	2650.7	2806.7	2965.6	3130.0	3478.4	4052.3	4683.0
(406.70 K)	*h*	2522.7	2649.2	2761.0	2865.6	3069.3	3275.0	3486.0	3927.1	4639.7	5409.0
	s	6.4319	6.7965	7.0778	7.3115	7.7022	8.0330	8.3251	8.8319	9.4690	10.0110
0.400	*v*	0.3177	0.4017	0.4708	0.5342	0.6548	0.7726	0.8893	1.1215	1.4685	1.8151
(143.63°C)	*u*	2366.3	2473.8	2564.5	2646.8	2804.8	2964.4	3129.2	3477.9	4052.0	4682.8
(416.78 K)	*h*	2493.4	2634.5	2752.8	2860.5	3066.8	3273.4	3484.9	3926.5	4639.4	5408.8
	s	6.2248	6.6319	6.9299	7.1706	7.5662	7.8985	8.1913	8.6987	9.3360	9.8780
0.500	*v*		0.3146	0.3729	0.4249	0.5226	0.6173	0.7109	0.8969	1.1747	1.4521
(151.86°C)	*u*		2461.5	2557.9	2642.9	2802.9	2963.2	328.4	3477.5	4051.8	4682.5
(425.01 K)	*h*		2618.7	2744.4	2855.4	3064.2	3271.9	3483.9	3925.9	4639.1	5408.6
	s		6.4945	6.8111	7.0592	7.4599	7.7938	8.0873	8.5952	9.2328	9.7749

[a] Symbols: v = specific volume, m^3/kg; u = specific internal energy, U/kg; h = specific enthalpy, kJ/kg; s = specific entropy, kJ/K kg.

Source: From Bolz, R. E. and G. L. Tuve, Eds. 1973. *CRC Handbook of Tables for Applied Engineering Science*, 2nd Ed., Chemical Rubber Co., Cleveland, Ohio.

A4

Ultimate Analysis of Biomass Fuels

Nitin Goel
Intel Technology India Pvt. Ltd.

Material	C^a (%)	H_2^a (%)	O_2^a (%)	N_2^a (%)	S^a (%)	A^a (%)	HHV (kJ/kg)[b]
			Agricultural Wastes				
Bagasse (sugarcane refuse)	47.3	6.1	35.3	0.0	0.0	11.3	21,255
Feedlot manure	42.7	5.5	31.3	2.4	0.3	17.8	17,160
Rice hulls	38.5	5.7	39.8	0.5	0.0	15.5	15,370
Rice straw	39.2	5.1	35.8	0.6	0.1	19.2	15,210
			Municipal Solid Waste				
General	33.9	4.6	22.4	0.7	0.4	38.0	13,130
Brown paper	44.9	6.1	47.8	0.0	0.1	1.1	17,920
Cardboard	45.5	6.1	44.5	0.2	0.1	3.6	18,235
Corrugated boxes	43.8	5.7	45.1	0.1	0.2	5.1	16,430
Food fats	76.7	12.1	11.2	0.0	0.0	0.0	38,835
Garbage	45.0	6.4	28.8	3.3	0.5	16.0	19,730
Glass bottles (labels)	0.5	0.1	0.4	0.0	0.0	99.0	195
Magazine paper	33.2	5.0	38.9	0.1	0.1	22.7	12,650
Metal cans (labels, etc.)	4.5	0.6	4.3	0.1	0.0	90.5	1,725
Newspapers	49.1	6.1	43.0	0.1	0.2	1.5	19,720
Oils, paints	66.9	9.6	5.2	2.0	0.0	16.3	31,165
Paper food cartons	44.7	6.1	41.9	0.2	0.2	6.9	17,975
Plastics							
General	60.0	7.2	22.6	0.0	0.0	10.2	33,415
Polyethylene	85.6	14.4	0.0	0.0	0.0	0.0	46,395
Vinyl chloride	47.1	5.9	18.6	(Chlorine = 28.4%)			20,535
Rags	55.0	6.6	31.2	4.6	0.1	2.5	13,955
Rubber	77.7	10.3	0.0	0.0	2.0	10.0	26,350
			Sewage				
Raw sewage	45.5	6.8	25.8	3.3	2.5	16.1	16,465
Sewage sludge	14.2	2.1	10.5	1.1	0.7	71.4	4,745
			Wood and Wood Products				
Hardwoods							
Beech	51.6	6.3	41.5	0.0	0.0	0.6	20,370
Hickory	49.7	6.5	43.1	0.0	0.0	0.7	20,165

(continued)

Material	C^a (%)	$H_2{}^a$ (%)	$O_2{}^a$ (%)	$N_2{}^a$ (%)	S^a (%)	A^a (%)	HHV (kJ/kg)[b]
Maple	50.6	6.0	41.7	0.3	0.0	1.4	19,955
Poplar	51.6	6.3	41.5	0.0	0.0	0.6	20,745
Oak	49.5	6.6	43.4	0.3	0.0	0.2	20,185
Softwoods							
Douglas fir	52.3	6.3	40.5	0.1	0.0	0.8	21,045
Pine	52.6	6.1	40.9	0.2	0.0	0.2	21,280
Redwood	53.5	5.9	40.3	0.1	0.0	0.2	21,025
Western hemlock	50.4	5.8	41.4	0.1	0.1	2.2	20,045
Wood products							
Charcoal (made at 400°C)	76.5	3.9	15.4	0.8	0.0	3.4	28,560
Charcoal (made at 500°C)	81.7	3.2	11.5	0.2	0.0	3.4	31,630
Douglas fir bark	56.2	5.9	36.7	0.0	0.0	1.2	22,095
Pine bark	52.3	5.8	38.8	0.2	0.0	2.9	20,420
Dry sawdust pellets	47.2	5.5	46.3	0.0	0.0	1.0	20,500
Ripe leaves	40.5	6.0	45.1	0.2	0.1	8.1	16,400
Plant wastes							
Brush	42.5	5.9	41.2	2.0	0.1	8.3	18,370
Evergreen trimmings	49.5	6.6	41.2	1.7	0.2	0.8	6,425
Garden plants	48.0	6.8	41.3	1.2	0.3	2.4	8,835
Grass	48.4	6.8	41.6	1.2	0.3	1.7	18,520

[a] All percentages on moisture-free basis.
[b] 1 kJ/kg = 0.43 Btu/lbm.

A5

Thermophysical Properties of Refrigerants

Nitin Goel
Intel Technology India Pvt. Ltd.

R-22 (Chlorodifluoromethane)
R-134a (1,1,12-Tetrafluoroethane)
R-404A [R-125/143a/134a (44/52/4)]
R-407C [R-32/125/134a (23/25/52)
R-410A [R-32/125 (50/50)]
Ammonia/Water
Water/Lithium bromide

The Montreal Protocol, signed in 1987 and later amended in 1990, 1992, 1997, and 1999 controls the production of ozone-depleting substances including refrigerants containing chlorine and/or bromine production chloro-fluoro-carbons (CFC). Pursuent to this treaty, refrigerants such as R-11 and R-12, ceased to exist in 1996 although continued use from existing stocks is permitted. In addition, hydrofluorocarbon (HCFCs) (such as R-22 and R-123) are being phased out, with complete cessation of production by January 1, 2030.

These refrigerants are being replaced by HFC refrigerants which have zero ozone depletion potential. Common HFC refrigerants are R-32, R-125, 134a, and R-143a and their mixtures, such as, R-404A, R-407C, and R-410A.

This appendix gives thermophysical properties of these HFC refrigerants and ammonia water and water–lithium bromide mixtures which are used in absorption refrigeration systems. Properties of R-22 are given to serve as a reference (Figure A5.1 through Figure A5.8).

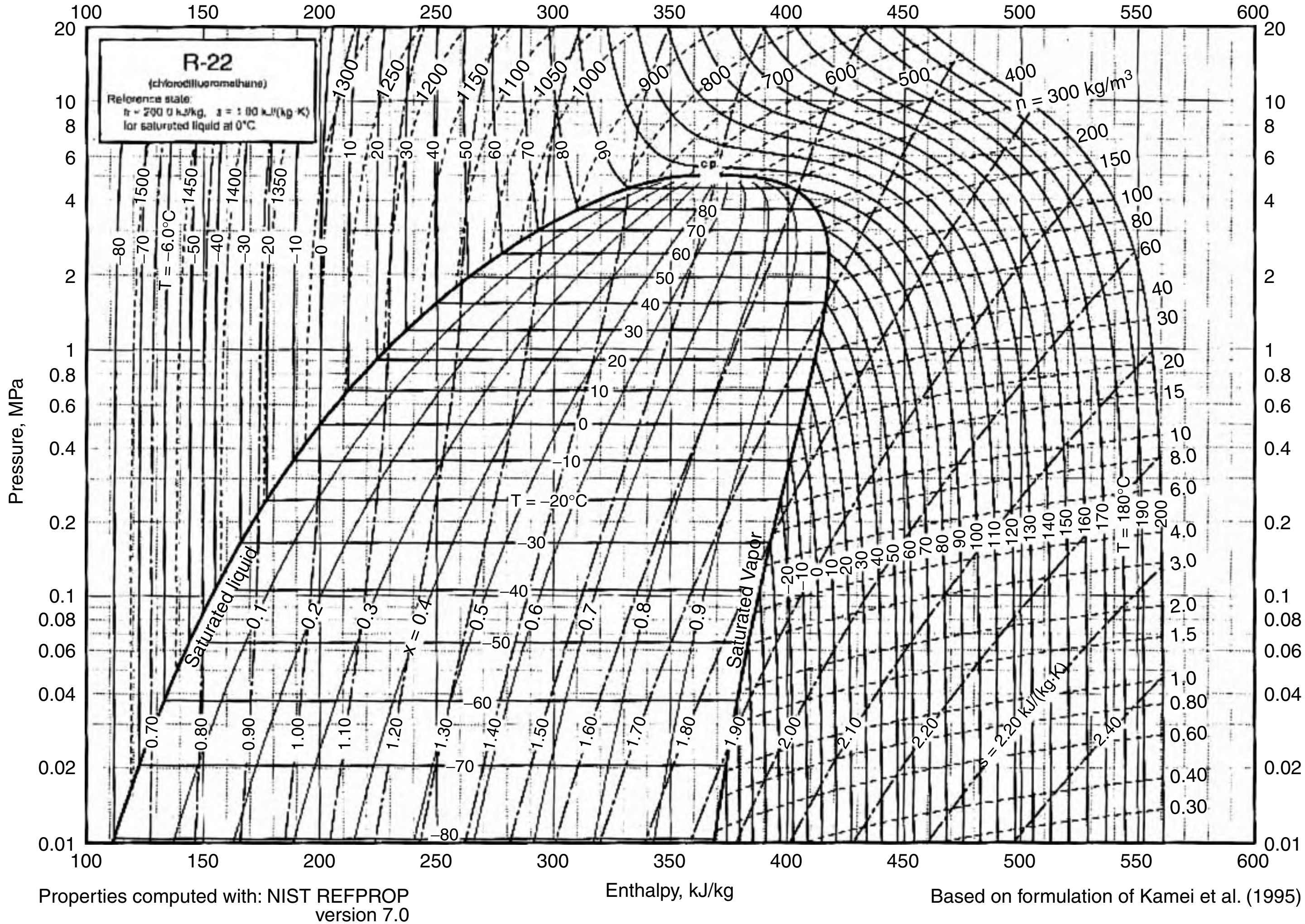


FIGURE A5.1 Pressure–enthalpy diagram for refrigerant R-22.

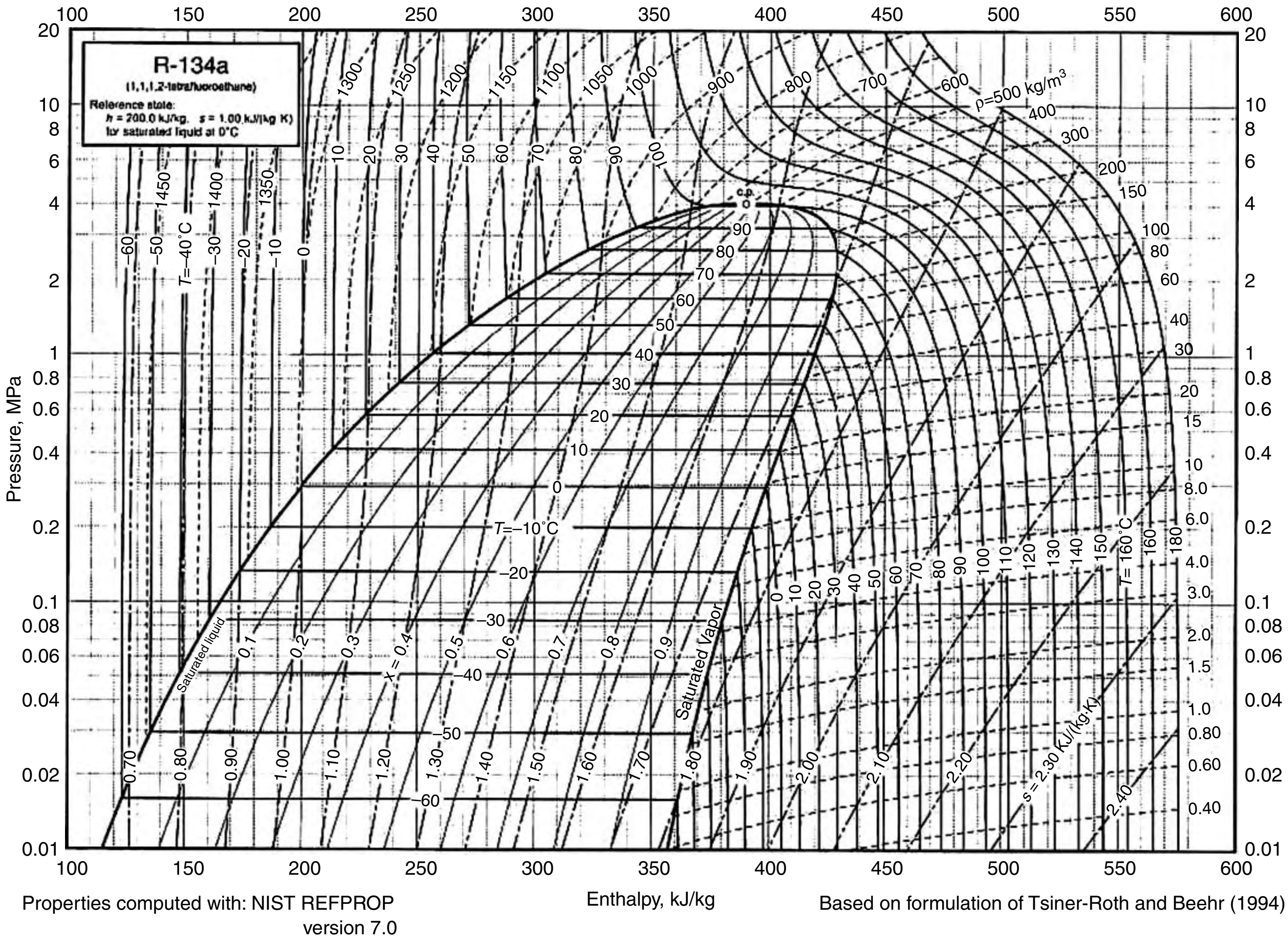


FIGURE A5.2 Pressure–enthalpy diagram for refrigerant R-134a.

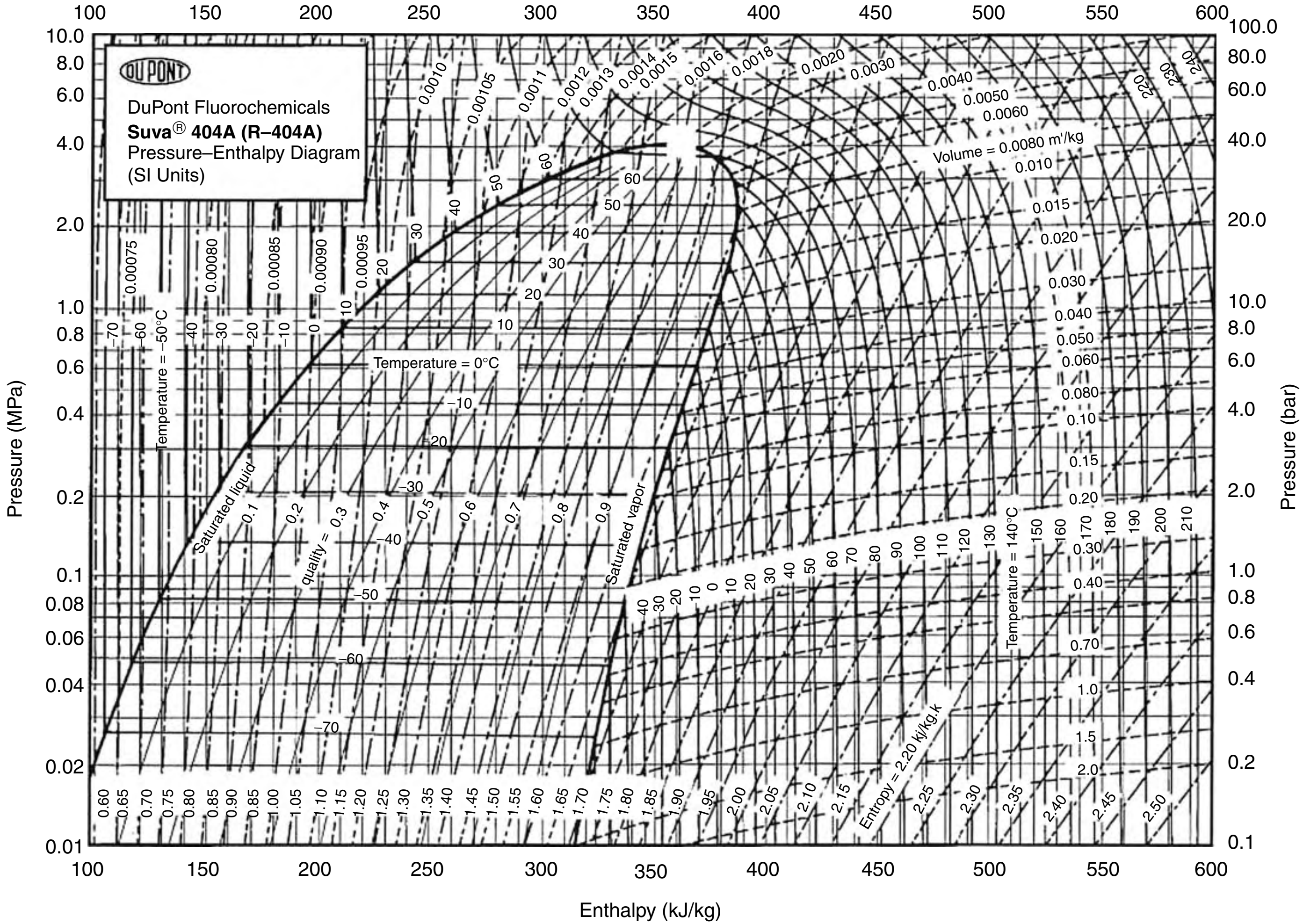


FIGURE A5.3 Pressure–enthalpy diagram for refrigerant R-404A.

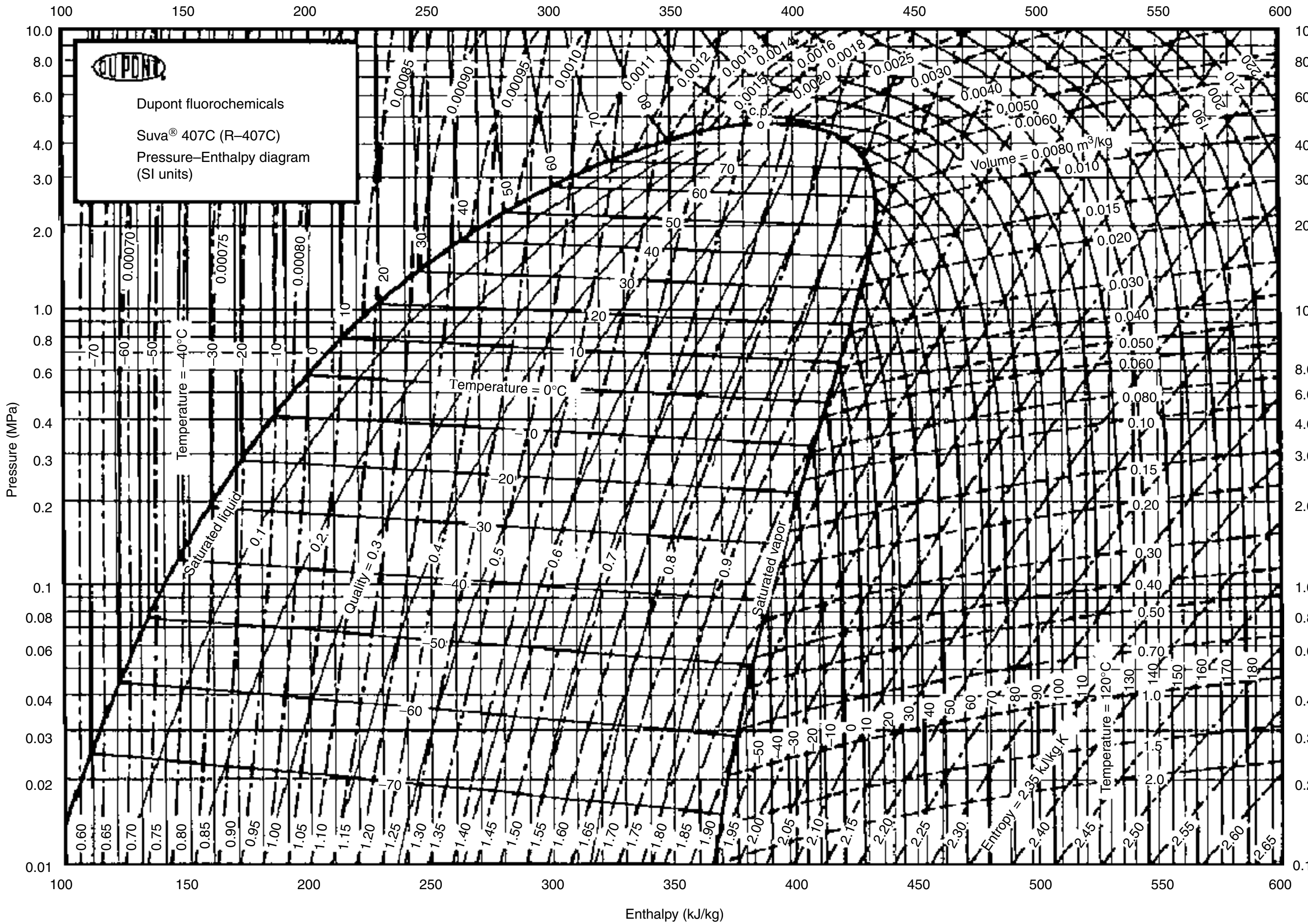


FIGURE A5.4 Pressure–enthalpy diagram for refrigerant R-407C.

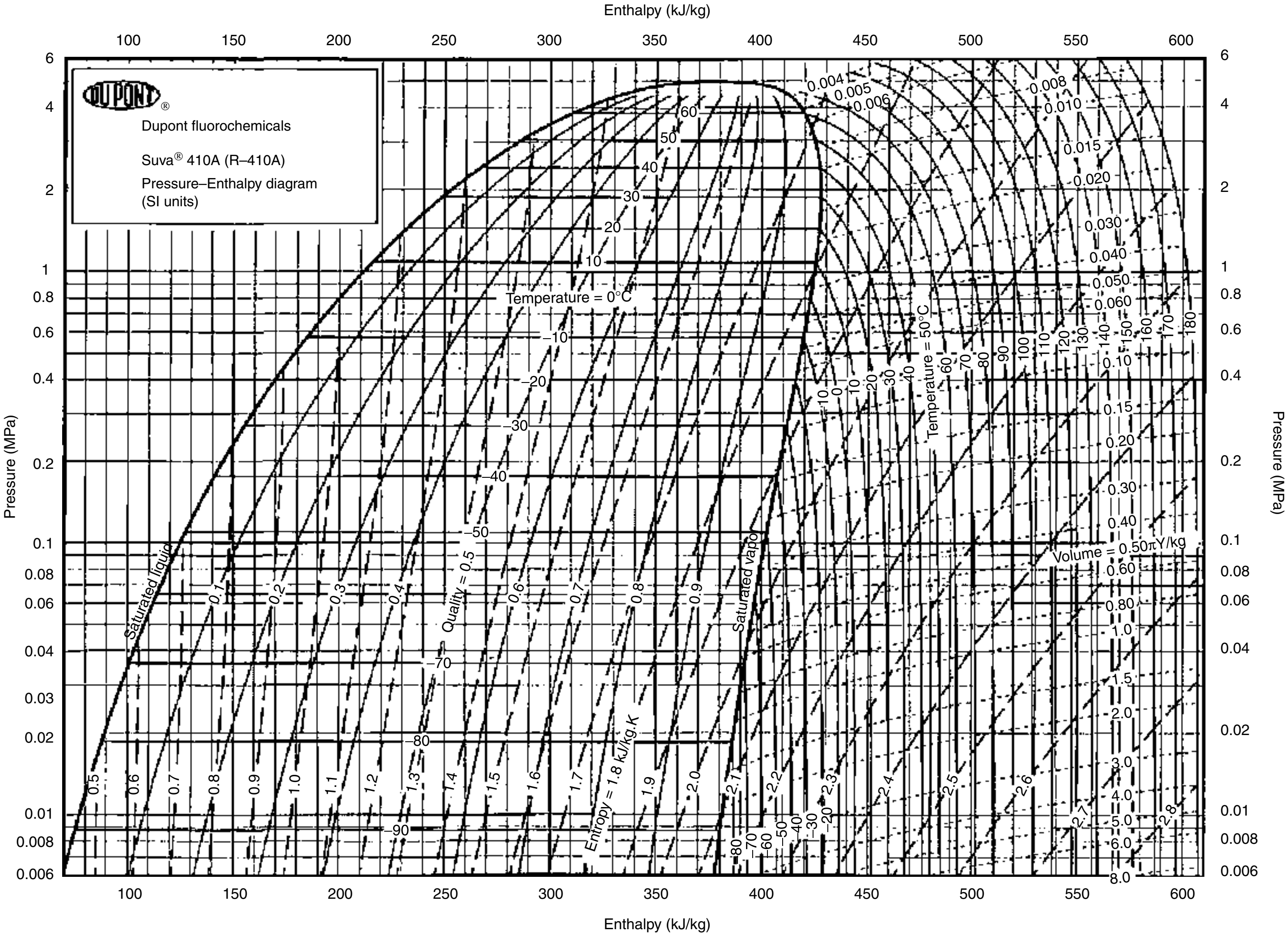


FIGURE A5.5 Pressure–enthalpy diagram for refrigerant R-410A.

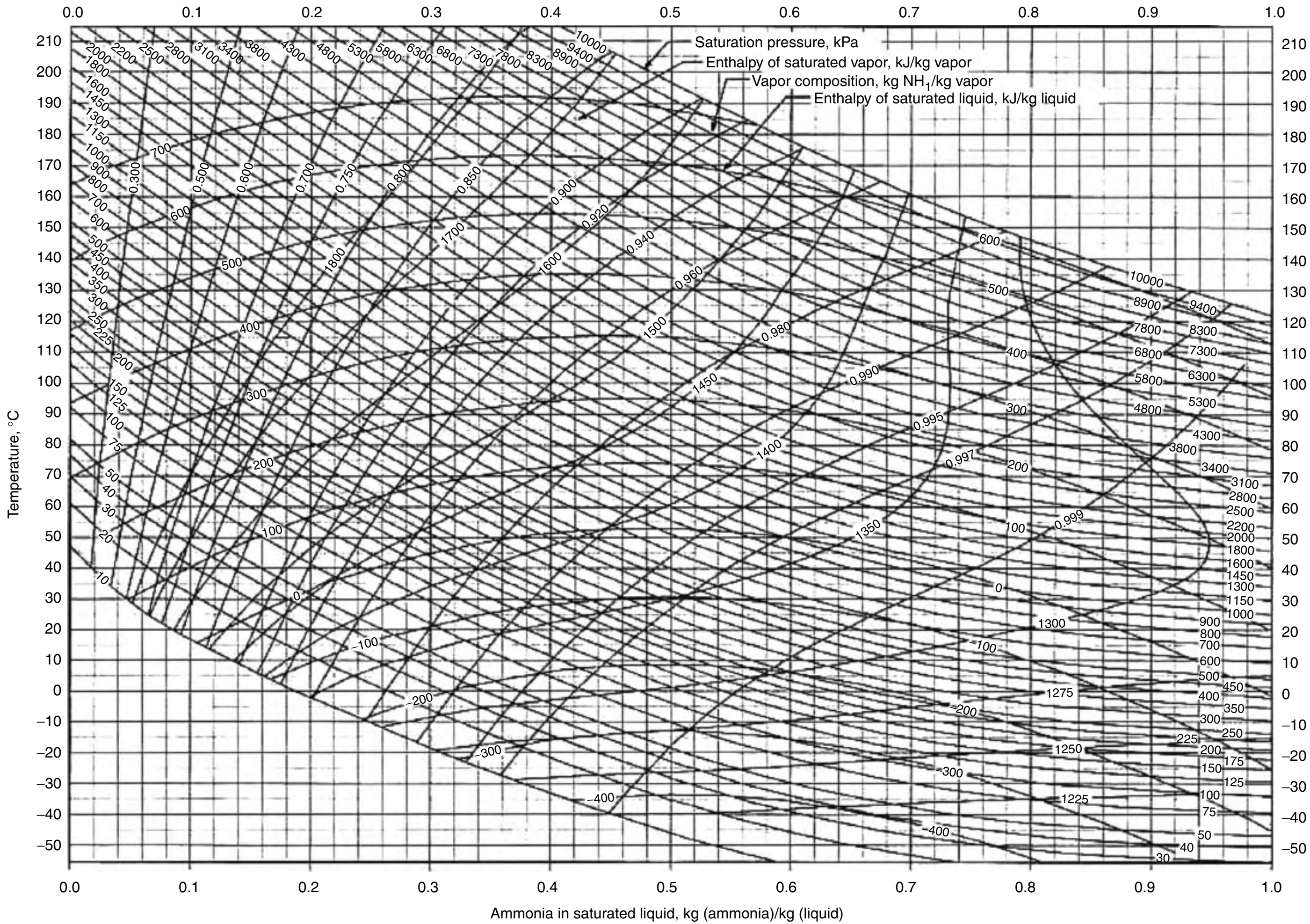
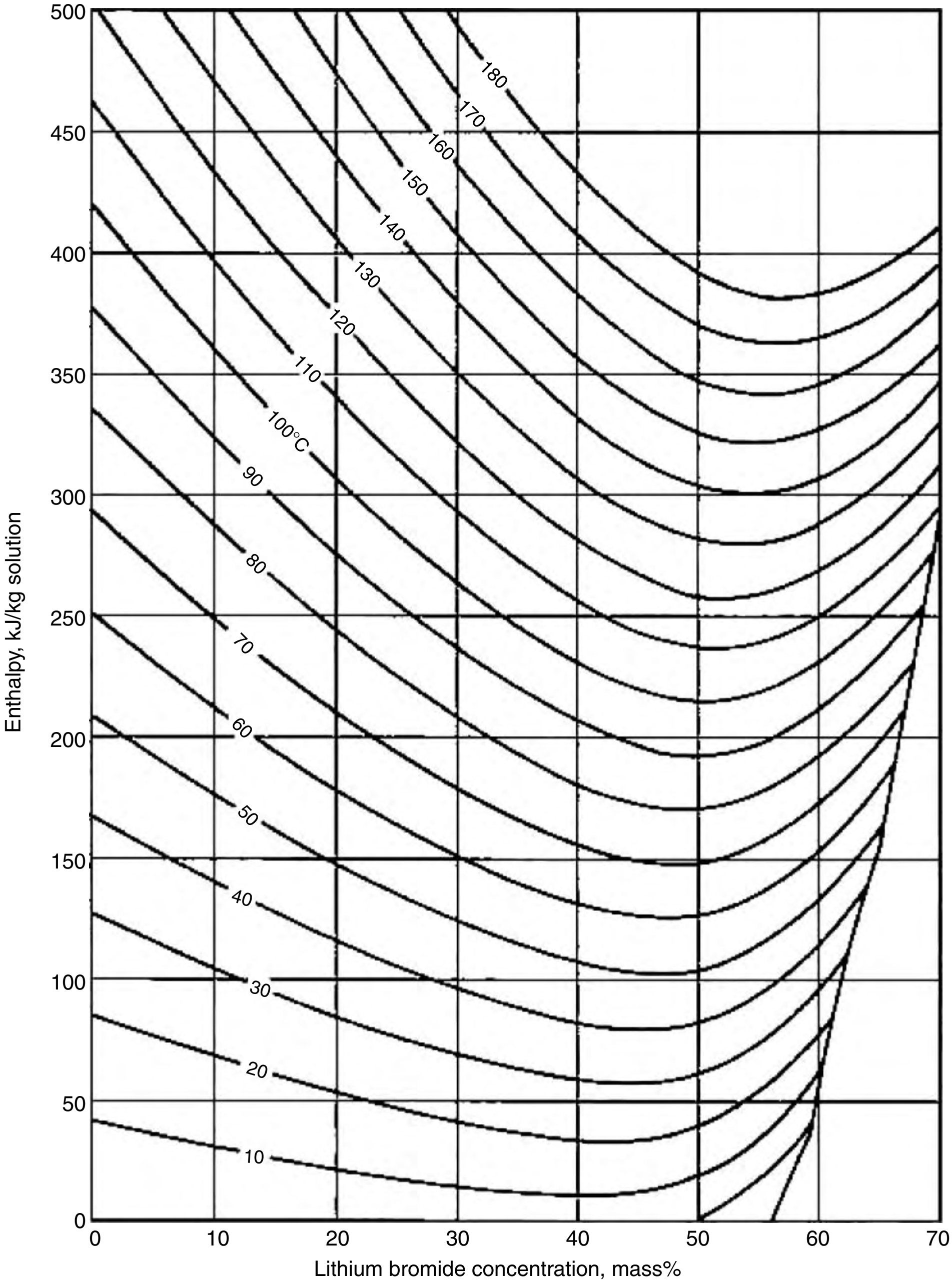


FIGURE A5.6 Enthalpy–concentration diagram for ammonia/water solutions prepared by Kwang Kim and Keith Herold, Centre for Environmental Energy Engineering, University of Maryland at Colleage Park.



Equations Concentration range 40 < x < 70% LiBr Temperature range 15 < t < 165°C

$h = \Sigma_0^4 A_n X^7 + r\Sigma_0^4 B_n X^n + r^2 \Sigma_0^4 C_p X^n$ in kJ/kg, where t = °C and X = %LiBr

$A_0 = -2024.33$	$B_0 =$ 18.2829	$C_0 = -3.7008214$ E-2
$A_1 =$ 163.309	$B_1 = -1.1691757$	$C_1 =$ 2.8877666 E-3
$A_2 = -4.88161$	$B_2 =$ 3.248041E-2	$C_2 = -8.1313015$ E-5
$A_3 =$ 6.302948 E-2	$B_3 = -4.034184$ E-4	$C_3 =$ 9.9116628 E-7
$A_4 = -2.913705$ E-4	$B_4 =$ 1.8520569 E-6	$C_4 = -4.4441207$ E-9

FIGURE A5.7 Enthalpy–concentration diagram for water/lithium bromide solutions.

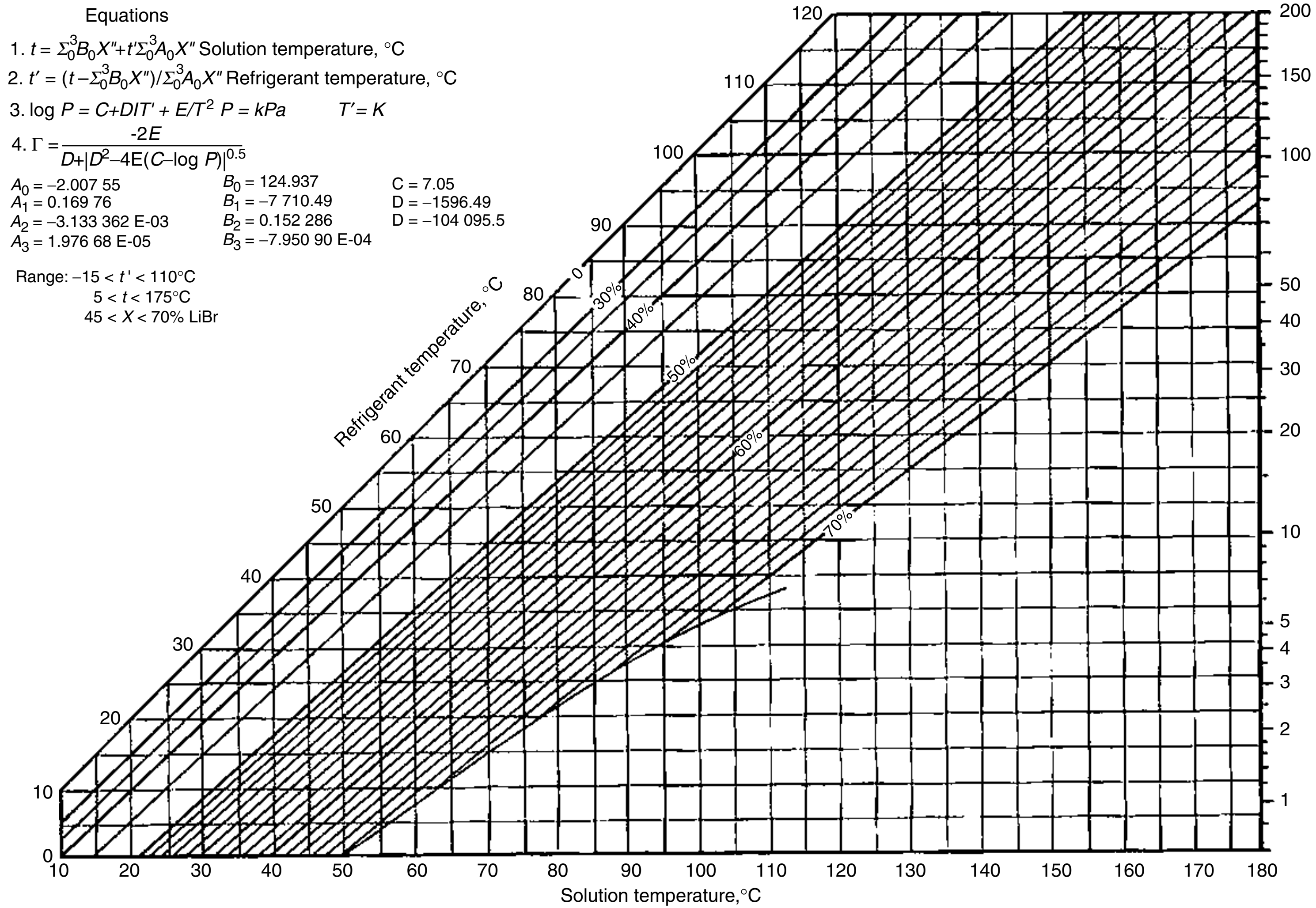


FIGURE A5.8 Equilibrium chart for aqueous lithium bromide solutions reprinted by permission of Carrier Corp.