Energy Primer

A Handbook of Energy Market Basics

A staff report of The Division of Energy Market Oversight

Office of Enforcement | Federal Energy Regulatory Commission

July 2012

3 Wholesale Electric Markets

Overview

Electricity is a physical product – the flow of electrons. It is a secondary energy source in that it results from the conversion of other energy forms such as natural gas, coal or uranium, or the energy inherent in wind, sunshine or the flow of water in a river. It may not be visible, but it can be turned on and off and measured.

> **Quick Facts: Measuring Electricity**
>
> Electricity is measured in terms of watts, typically in kilowatts (1,000 watts) or megawatts (1,000 kilowatts).
>
> A kilowatt (or watt or megawatt) is the amount of energy used, generated or transmitted at a point in time. The aggregation of kilowatts possible at a point in time for a power plant, for example, is its capacity. The aggregation of kilowatts used at a point of time is the demand at that point.
>
> The number of kilowatts used in an hour (kilowatt-hour or kWh) is the amount of electricity a customer uses or a power plant generates over a period of time. Kilowatt-hours are used to bill customers.

Electric markets have retail and wholesale components. Retail markets involve the sales of electricity to consumers; wholesale markets typically involve the sales of electricity among electric utilities and electricity traders before it is eventually sold to consumers. This paper focuses on wholesale markets, although it addresses retail demand and other instances where retail markets strongly influence wholesale markets.

Much of the wholesale market and certain retail markets are competitive, with prices set competitively. Other prices are set based on the service provider's cost of service. For wholesale markets, FERC either authorizes jurisdictional entities to sell at market-based rates or reviews and authorizes cost-based rates.

In competitive markets, prices reflect the factors driving supply and demand – the physical fundamentals. In markets where rates are set based on costs, these fundamentals matter as well. Supply incorporates generation and transmission, which must be adequate to meet all customers demand simultaneously, instantaneously and reliably.

Consequently, key supply factors affecting prices include fuel prices, capital costs, transmission capacity and constraints and the operating characteristics of power plants. Sharp changes in demand, as well as extremely high levels of demand, affect prices as well, especially if less-efficient, more-expensive power plants must be turned on to serve load.

Background

Electricity on Demand

In the United States and other developed countries, consumers expect electricity to be available whenever they need it. Electricity use has grown enormously as consumers now consider not only refrigerators, TVs and hair dryers but also computers, iPods and other electronic devices as necessities. Consumers also expect to pay reasonable prices for the electricity they use.

Meeting these customer expectations is challenging. With few exceptions, electricity cannot be stored, in any appreciable quantities, and thus must be produced as needed. Further, unlike most other markets, electricity's historical inelastic demand does not move with prices. To provide electricity on demand, electric system operations have to be planned and conducted with that goal in mind. Lacking storage and responsive demand, operators must plan and operate power plants and the transmission grid so that demand and supply exactly match, every moment of the day, every day of the year, in every location.

The Drive for Enhanced Value

The electric industry has met this growing demand with increasing efficiency. Between 1929 and 1967, the national average cost of electricity for residential customers plummeted from about 60¢/kWh to 10¢/kWh (in 2005 dollars), and remains around there today. How did the industry achieve such tremendous cost savings and then keep the real price of electricity flat over the past 40 years? Part can be explained by greater efficiency – power plants use less fuel, and new techniques make it cheaper to extract the coal and natural gas that fuels generators. Another part of the answer, though, stems from changes in the way the industry is organized and operated.

Economies of Scale

Electric power is one of the most capital intensive industries. Generation alone can account for roughly 65% of a customer's electric bill. Spreading these relatively fixed costs over more customers helps bring down the cost that each customer pays.

Thomas Edison's first street lighting project in the 1880s grew to electrifying whole neighborhoods, towns and cities. Providing service over larger areas allowed utilities economies of scale in generating technology. The cost per unit of production dropped as power plants grew larger and larger. The companies building these facilities were basically self-contained – they owned and operated the generation, transmission and distribution facilities. Power lines were built from their generation to their population, or load, centers. These companies were vertically integrated.

One downside of larger generating units is that they are difficult to replace if they experience unexpected shutdowns. For a single utility building a new and larger unit, the only way to ensure reliable service is to build two units – creating a capacity reserve. When coal and nuclear unit sizes grew to 500 or 1,000 MW, building two units became very expensive for any individual company.

Reserve Sharing, Interconnection and Power Pools

The solution to high reserve costs was to share reserves with adjacent utilities. Instead of building two large units, utilities could buy from their neighbors in times of need, and cut their costs significantly. To facilitate reserve sharing, utilities built major interconnecting transmission lines large enough to deliver power in case of a major generator outage. Today's bulk power grid began as a way to maintain reliable service while lowering costs.

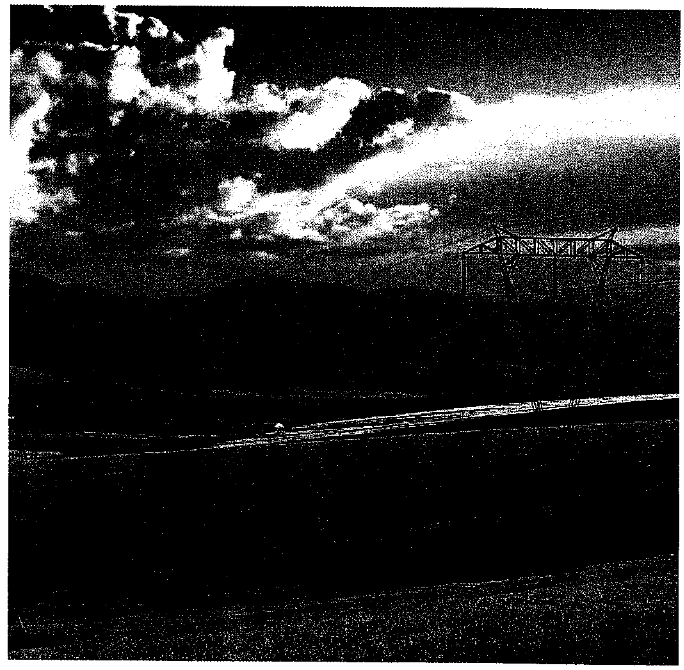
As more utilities share reserves, the smaller the reserves each must carry, and the lower the costs. The value of reserve-sharing agreements led to the formation of power pools, the forerunners of today's regional transmission organizations (RTOs).

Coordinating reserves also led to closer coordination of other utility functions, such as the process of determining which generating units to use, called unit commitment. Operators want to commit just enough capacity for the next day to ensure reliability but no more than needed, to minimize costs. This began a new phase of using economies of scale in system operations encompassing whole regions of the country.

Regional coordination also was spurred by special circumstances, particularly in the West. Large federally owned dams on the Columbia and Colorado rivers generate power from the spring runoff of melting mountain snow. When the reservoirs are full and the turbines are spinning, there is not enough local demand to use the power. Since the hydropower was cheaper than any alternative, long distance transmission lines were built to deliver the excess power from the Northwest and Southwest to load centers in California.

With the transmission interconnections in place, northwestern utilities found that they could get cheaper power from southern power generation at other times of the year. These seasonal and regional disparities in availability and price provide for a lively bilateral trading market.

In the 1960s, the electric industry created an informal, voluntary organization of operating staff to aid in coordinating the bulk power system. Then, in 1965, the largest power blackout until that time hit the northeastern United States – including New York – and southeastern Ontario, Canada, affecting 30 million people. The blackout led to the development in 1968 of the National Electric Reliability Council (NERC), shortly thereafter renamed the North American Electric Reliability Council. and nine regional reliability councils. Rather than serving as a pool or other entity for sharing resources, NERC focused on reliability. In 2006, using authority granted in the U.S. Energy Policy Act of 2005, FERC certified NERC as the electric reliability organization for the United States, and reliability standards became mandatory and enforceable.



Optimizing Unit Commitment and Economic Dispatch

The industry also reduced costs by using computers and communication technology to optimize system operations. Utilities use algorithms for optimizing the commitment of their generating units, while RTOs' day-ahead market software does this for suppliers bidding into their markets.

In real time, demand is changing all the time. Without storage and responsive demand, the output of some generators must change to follow constantly changing demand. This is known as load following. Utilities use economic dispatch to optimize the use of these units and minimize real-time costs.

Economy Energy Trade

Since transmission interconnections were built primarily for the rare need to deliver reserves in emergencies, the industry had excess transmission capacity. This allowed utilities to use the lines to trade power. Major utilities generally owned sufficient capacity to meet their own peak power needs. However, sometimes the cost of operating their marginal generation was higher or lower than that of their neighbors. Transmission availability provided opportunities for utilities to save money by buying energy when it was cheaper than generating and selling energy to utilities with higher costs. This is called economy energy trading.

Evolving Public Policies

Different public policy theories have shaped the electric power industry over its history. All of these public policies are still in play to some extent today. Five concepts that helped shape the electricity industry and markets are outlined below.

Not-for-Profit Utilities

One of the first approaches to ensuring customer value was to depend on nonprofit electric providers. In the early days of the industry, electrification started in towns and cities. In many places, this utility service was provided by the municipal government. The federal government stepped in to develop and market the nation's significant hydroelectric resources. The Depression-era rural electrification program relied on customer-owned rural electric cooperatives and low-interest government loans. There are currently more than 1,700 municipal and almost 900 cooperative utilities in the United States.

Regulated Natural Monopolies

A second model for operating power systems was investor-owned regulated monopolies. In the early days of the industry, while many cities went the municipal route, many investor-owned utilities were also starting up. These private utilities are regulated, typically by a state agency. Initially, they agreed to be regulated to overcome a lack of retail competition, and were granted exclusive service territories (franchise). Today, regulation focuses on mitigating market power, among other things, because many utility functions are seen as natural monopolies.

State regulators approve a utility's investments in generation and distribution facilities, either in advance of construction or afterwards when the utility seeks to include a facility's costs in retail rates. Some states eventually developed elaborate integrated resource planning (IRP) processes to determine what facilities should be built.

Power Pools

Power pools are multilateral arrangements with members ceding operational control over their generating units and transmission facilities to a common operator. Members provided incremental cost data about their units and system status data to the operator. The operator ran an energy management system that used the unit cost data to optimize on a multilateral basis unit commitment and economic dispatch.

PJM began in 1927 for utilities to share their generating resources, forming the world's first power pool. The New York Power Pool was formed in 1966 and the New England Power Pool in 1971 in response to the 1965 Northeast blackout. The Electric Reliability Council of Texas (ERCOT) and the Southwest Power Pool (SPP) formed in 1941 to pool resources for the war effort.

Competition, Part 1: Competitive Generation and Open Access

The environmental movement and initiatives to open the airline and trucking industries to competition also helped shape the energy industry in the 1970s. A provision in President Carter's energy plan led to passage of the Public Utility Regulatory Policies Act of 1978 (PURPA), which ushered in the next era.

PURPA established a program implemented by states and overseen by the FERC to encourage the use of efficient cogeneration (using the heat from industrial or other processes to generate electricity) and small scale renewable generation. FERC's role was to issue regulations for the program and certify that qualifying facilities (QFs) met statutory requirements. States administratively set the price to be paid to these generators at the cost the utilities would avoid by purchasing the power rather than generating it themselves.

Most states set their avoided cost rate so low that they got little QF capacity. However, California, Texas and Massachusetts set very generous avoided cost rates and were overwhelmed with QF capacity, much of which received prices that turned out to be higher than the actual costs avoided by the purchasing utility. The rapid growth and size of the QF industry surprised many policymakers and entrepreneurs, and got them thinking about the viability of generation independent of regulated monopolies.

In 1988, FERC proposed rules to allow states to set their avoided-cost rate based on an auction. Instead of taking all capacity at a set rate, states could set the rate based on bids to supply a certain amount of needed capacity. The Commission also proposed to open the avoided-cost auction up to independent power producers (IPPs) that did not qualify as QFs. In this way, a regulatory program was transformed into a competitive initiative.

Under the regulated monopoly model, utilities owning and operating transmission lines had no obligation to allow others to use them. This posed a significant barrier to the development of an independent power industry. The Commission started conditioning approval in merger cases with the voluntary provision of open transmission access. The Energy Policy Act of 1992 gave the Commission authority to grant transmission access on request. These approaches to open access resulted in patchwork transmission access.

By the mid-1990s, support for opening the transmission grid to all users encouraged the Commission to pursue a generic solution. Order 888 required mandatory open transmission access by all transmitting utilities and a reciprocity provision successfully extended open access to nonjurisdictional entities (municipal, cooperative and federal utilities).

Order 889 addressed matters needed to implement open access. The rule established the Internet-based Open Access Same-Time Information System (OASIS) for posting available transmission capacity and reserving transmission capacity. These rules required significant changes to utility control room operations and limited the ability of companies to share transmission-related information with their own power marketing operating units.

Competition, Part 2: Integrating Markets and Operations – RTOs

Order Nos. 888 and 889 were designed for an industry of bilateral energy markets, in which parties negotiated transactions among themselves. The open-access transmission tariff and rules did not work well for multilateral power pools open to independent power producers. This led to the development of independent system operators (ISO) and, subsequently, regional transmission organizations (RTO). This primer uses RTO to stand for both RTOs and ISOs.

RTOs did more than operate the transmission system and dispatch generation, however. They developed markets in which buyers and sellers could bid for or offer generation. The RTOs used the bid-based markets to determine economic dispatch.

Major parts of the country operate under more traditional market structures, notably the West (excluding California) and the Southeast. Notably, two-thirds of the nation's electricity load is served in RTO regions.

Electricity Demand

Americans use electricity for heat and light, to run machinery and to power a growing number of products such as televisions, radios, computers, hair dryers, cell phones and iPods. This use has been increasing, reaching 3,865 gigawatt-hours (GWh) of electricity in 2008. Demand dropped in 2009 with the recession, but has since regained its upward trend.

The bulk of the electricity generated is sold to consumers, known as end-users or retail customers. Some consumers generate some or all of the power they consume. Some of the electricity sold to retail consumers is generated by integrated investor-owned utilities, federal entities, municipally owned and co-operatively owned utilities that sell the power directly to consumers. The rest of the electricity ultimately consumed by retail customers is bought and sold through wholesale electric markets.

This primer focuses on wholesale markets, which generally involve the sale of electricity to entities that resell the power to retail customers. However, retail consumers' electric use shapes demand and, therefore, the wholesale markets.

Quick Facts: Measuring Electricity

Electric use is described in terms of quantity and time.

The unit of measure of the quantity used is the kilowatt (kW), or 1,000 watts. The maximum number of all the kilowatts used by consumers on an electrical system at a point of time is peak demand.

The amount of electricity a consumer uses over a period of time is described as the number of the kilowatt-hour (kWh) - 1,000 watts working for one hour. Consumers pay based on the number of kWh they consume in a billing period, typically a month.

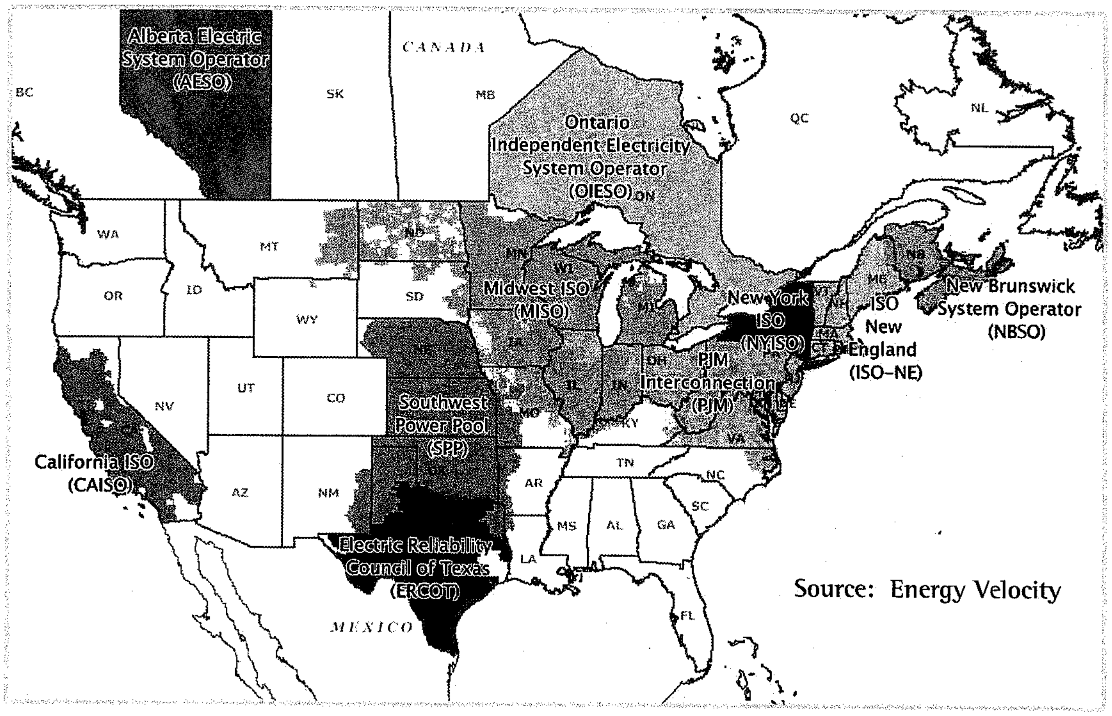
Source: EIA

Demand Characteristics

Demand is often characterized as baseload or peak. Baseload is demand that occurs throughout the day or throughout the year. Refrigerators, for example, may create baseload demand. Peak load is demand that shows up during part of the day or year, all at the same time – heating or air conditioning, for example.

Demand for electricity follows cycles, throughout the day and year. Regionally, electric demand may peak in either the summer or the winter. Spring and fall are typically shoulder months, with lower peak demand. Seasonal peaks vary regionally, although the highest levels of power load in almost all regions of the United States occur during heat waves and are most acute during the daily peak load hours reached in the late afternoon. However, a minority of regions reach their peak load when the weather is extremely cold. These are primarily areas with significant space-heating requirements and little summer air conditioning load. A majority of these systems are in the far northern areas of

North American Regional Transmission Organizations



the United States, where air conditioning load is not significant. South Florida's seasonal peak also occurs during the winter, when the population and tourism surges and uses more power than native Floridians do in the summer.

Daily demand typically peaks in the late afternoon, as commercial and domestic activities peak, and, in the winter, when lighting needs grow.

Electricity use also varies between weekdays and weekends. Commercial and industrial activities are lower on weekends and peoples' noncommercial activities change with their personal schedules. The load on different weekdays also can behave differently. For example, Mondays and Fridays, being adjacent to weekends, may have different loads than Tuesday through Thursday. This is particularly true in the summer.

Because demand historically has not varied with price and because storage options are limited, generation must rise and fall to provide exactly the amount of electricity customers need. The cost of providing power typically rises as demand grows, and falls as demand declines, so wholesale power prices are typically highest during peaks. Consequently, system planners, power marketers and traders all carefully track weather trends, economic growth and other factors to forecast power demand.

Demand Drivers

In general, the amount of electricity demanded is relatively insensitive to the price of electricity in the short-term (inelastic). One reason for this is that many customers – especially smaller customers – do not get price signals to which they can respond. Most residential customers are billed monthly on a preset rate structure. Large industrial customers, on the other hand, may receive real-time price signals.

Further, electricity is a necessity to most people and businesses. While they may be able to reduce their demand in the short-term – by turning down the thermostat or turning off lights, for example – electric consumers find it difficult to do without electricity altogether. There is little storage for electricity now and few realistic substitutes. Consequently, demand tends to drive price, especially when the system is stressed.

In the longer-term, options for reducing electricity use include switching to natural gas, installing insulation and implementing other energy efficiency measures. Larger consumers may consider building their own generation facilities.

Governments and businesses are also developing demand-response programs, which provide plans in which customers agree to reduce load in exchange for compensation.

Factors driving demand include demographics, climate and weather, economic activity and policies and regulations.

Demographics

Population levels affect demand, with greater population levels tending to increase electric consumption. Shifts in population also affect regional demand. Population flight in the 1980s from northern industrial regions – the Rust Belt – to warmer climates in the South affected residential consumption patterns. In the 1990s, consumption in the South surpassed that in the Midwest, making it the region with the greatest electric use.

Climate and Weather

Weather is the biggest factor driving demand. General climatic trends drive consumption patterns and therefore the infrastructure needed to ensure reliable service. Cold weather and short days drive winter demand in northern regions. Southern regions rely more on electric space heating, and, thus, see demand rise in the winter, although demand typically peaks in the summer with air conditioning load. In the winter, lighting contributes to the occurrence of peaks during the seasonally dark early morning and early evening hours.

Weather also can have extreme short-term effects on electricity usage. A sudden cold snap can drive heating use up quickly and a heat wave can push air conditioning loads. Other, less obvious weather patterns affect demand – rain and wind, for example, may result in sudden cooling, affecting heating or air conditioning.

Economic Activity

The pattern of socioeconomic life affects the cycle of electric use, with weekends and holidays showing a different pattern than weekdays. Demand typically rises as people wake up and go to work, peaking in the afternoon.

The health of the United States and regional economies also affects power demand. During periods of robust ac-

Quick Facts:
Heating and Cooling Degree Days

In the United States, engineers developed the concept of heating and cooling degree days to measure the effects of temperature on demand. Average daily temperatures are compared to a 65° F standard – those in excess of 65° yield cooling degree days; those below 65° yield heating degree days. A day with an average temperature of 66° would yield one cooling degree day.

tivity, loads increase. Similarly, loads drop during recessions. These changes are most evident in the industrial sector, where business and plants may close, downsize or eliminate factory shifts. In addition to reducing overall demand, these changes may affect the pattern of demand; for example, a factory may eliminate a night shift, cutting baseload use but continuing its use during peak hours. In some cases these effects can be significant.

Energy Policies and Regulations

State regulatory agencies set prices and policies affecting retail customer service. Some states are considering changes that would enable customers to receive more accurate price signals. They include, among other things, changing rate structures so that the rate varies with the time of day, or is even linked to the cost of providing electricity.

Efforts to reduce overall demand by improving energy efficiency are underway through several governmental and utility venues.

Retail Customer Mix

Most electric utilities serve different types of customers: residential, commercial and industrial. Each class uses electricity differently, resulting in a differing load profile, or the amount that each customer class uses and the daily shape of the load. If a consumer uses electricity consistently throughout the day and seasons, his load shape is flat, and the load will be baseload. Another consumer may use more at some times than others, resulting in baseload and peaks. Greater variability in demand is typically more expensive to serve, especially if the peak occurs at the same time other customers' use peaks. Consequently, the mix of customer types affects a region's overall demand.

Residential consumers form the largest customer segment in the United States at approximately 38 percent of electricty demand. Residential consumers use electricity for air conditioning, refrigerators, space and water hearing, lighting, washers and dryers, computers, televisions, cell phones and other appliances. Prices for residential service are typically highest, reflecting both their variable load shape and their service from lower-voltage distribution facilities, meaning that more power lines are needed to provide service to them.

Commercial use is the next largest customer segment at approximately 36 percent, and includes hotels and motels, restaurants, street lighting, retail stores and wholesale businesses and medical religious, educational and social facilities. More than half of commercial consumers' electric use is for heating and lighting.

Industrial consumers use about 26 percent of of the nation's electricity. This sector includes, for example, manufacturing, construction, mining, agriculture and forestry operations. Industrial customers often see the lowest rates, reflecting their relatively flat load structure and their ability to take service at higher voltage levels.

Transportation demand for electricity stems primarily from trains and urban transportation systems. This is less then 1 percent of electricty demand.

Load Forecasting

Demand is constantly changing, challenging grid operators and suppliers responsible for ensuring that supply will meet demand. Consequently, they expend considerable resources to forecast demand. Missed forecasts, where actual demand differs significantly from the forecast, can cause wholesale prices to be higher than they otherwise might have been.

Forecasts are necessary as well for the variety of actions that must occur if sufficient supply is to be available in the immediate or long term: planning the long-term infrastructure needs of the system, purchasing fuel and other supplies and staffing, for example. Load forecasts are also extremely important for suppliers, financial institutions and other participants in electric energy generation, transmission, distribution and trading.

Load forecasting uses mathematical models to predict demand across a region, such as a utility service territory or RTO footprint. Forecasts can be divided into three categories: short-term forecasts, which range from one hour to one week ahead; medium forecasts, usually a week to a year ahead; and long-term forecasts, which are longer than a year. It is possible to predict the next-day load with an accuracy of approximately 1%-3% of what actually happens. The accuracy of these forecasts is limited by the accuracy of the weather forecasts used in their preparation and the uncertainties of human behavior. Similarly, it is impossible to predict the next year peak load with the similar accuracy because accurate long-term weather forecasts are not available.

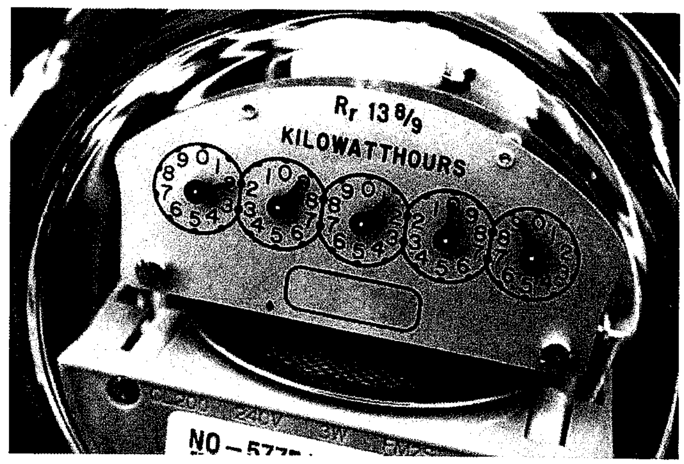
The forecasts for different time horizons are important for different operations within a utility company. Short-term load forecasting can help to estimate transmission system power flows and to make decisions that can prevent overloading of transmission systems. Timely implementation of such decisions leads to the improvement of network reliability and to the reduced occurrences of equipment failures and blackouts. Forecasted weather parameters are the most important factors in short-term load forecasts; temperature and humidity are the most commonly used load predictors.

The medium- and long-term forecasts, while not precise, take into account historical load and weather data, the number of customers in different customer classes, appliances used in the area and their characteristics, economic and demographic data, and other factors. For the next-year peak forecast, it is possible to provide an estimated load based on historical weather observations. Long-term forecasts are used for system infrastructure planning and are meant to ensure that there are sufficient resources available to meet the needs of the expected future peak demand. These forecasts are made for periods extending 10 to 20 years into the future.

Demand Response

Electric demand is generally insensitive to price, meaning that demand does not typically fall significantly when wholesale prices rise. However, some utilities and grid operators are developing ways to stimulate a response from consumers through demand-response programs.

Demand response (DR) is the ability of customers to respond to either reliability or price triggers by forgoing



power use for short periods, by shifting some high energy use activities to other times or by using onsite generation. The programs may use price signals or incentives to prompt customers to reduce their loads. The signals to respond to electric power system needs or high market prices may come from a utility or other load-serving entity, a regional transmission organization (RTO) or an independent DR provider. These programs are administered by both retail and wholesale entities.

DR has the potential to lower systemwide power costs and assist in maintaining reliability. It can be used instead of running power plants or to relieve transmission congestion There can also be environmental benefits because peaking units tend to be costly - and dirty - to run.

Demand response rewards consumers for reducing load during certain market conditions and at specific times. However, it is difficult to measure and quantify this reduction. Measuring and verifying the reduction requires development of consumers' baseline usage, against which their actual use is measured to determine the reduction in the event they are called to lessen their load. An accurate measure of their typical usage is important to prevent (or detect) gaming by participants.

Demand-Response Programs

Programs generally fall into three categories: curtailing, shifting or on-site generation.

Curtailing, or forgoing, involves reducing power use (load) during times of high prices or threats to reliability without making up the use later. For example, residential customers might turn off lights or raise thermostats during hot weather. Commercial facilities may turn off office equipment, lower building lighting or change thermostat settings by a few degrees.

Shifting involves moving or rescheduling high energy-use activities in response to high prices or DR program events to off-peak periods – evenings, nights or weekends. Industrial customers might reschedule batch production processes to evening hours or the next day. Commercial establishments may delay high-energy operations. Residential customers may wait until evening or night to use high-energy consuming appliances, such as clothes dryers or dishwashers. In shifting, the lost amenity or service is made up at a subsequent time.

On-site generation is when some customers may respond by turning on an on-site or backup emergency generator to supply some or all of their electricity needs. Although customers may have little or no interruption to their electrical usage, their net load and requirements from the power system are reduced. The ability to use on-site generation is most common for institutional customers, such as hospitals, large schools or data centers.

DR programs can be further distinguished by whether they are controlled by the system operator (dispatchable) or the customer (nondispatchable). Dispatchable demand response refers to programs that reduce customer energy use, such as direct load control of residential appliances or directed reductions to industrial customers. Dispatchable DR is used for reliability or economic reasons. Nondispatchable demand response lets the retail customer decide whether and when to reduce consumption in response to the price of power. It includes time-sensitive pricing programs based on rates that charge higher prices during high-demand hours and lower prices at other times.

As a result of technology innovations and policy directions, new types and applications of DR are emerging that encompass the use of smart appliances that respond in near real-time to price or other signals. These models may allow customers to respond more easily as they require little customer monitoring or interaction.

Demand Response in Retail Markets

Many states require utilities to use energy efficiency, DR or renewable resources. Energy Efficiency Resource Standards (EERS) in more than half of the states (plus Washington, D.C.) require utilities to achieve electric energy savings; many of these standards include peak load reduction targets. These mandates provide incentives for utilities to reduce customers' energy consumption, such as mechanisms that decouple profits from the amount of electricity sold, or performance bonuses for utilities that meet or exceed reduction targets.

Some states are implementing dynamic pricing, in which retail rates change frequently to better reflect system costs. Time-based rates depend on advanced meters at customer premises that can record usage. In time-of-use programs, customers are charged different prices at different times, with hours of peak demand costing more than off-peak hours.

In real-time pricing (RTP) programs, customers are charged prices reflecting the immediate cost of power. Industrial or very large commercial customers are often on RTP tariffs.

Critical peak pricing (CPP) uses real-time prices at times of extreme system peak, and is restricted to a small number of hours annually but features prices higher than time-of-use prices during the critical peak. Consumers do not know in advance when a critical peak might be called. A CPP program for residential customers uses a carrot without the stick: critical-peak rebates. Participating customers get rebates on their bills for responding to utility price-signals, but are not penalized if they do not lower use in those hours.

Wholesale Market Programs

Retail programs may aid RTOs, although the RTO may not be able to invoke them or even see specifically the amount of response that occurs. Wholesale-level DR occurs in the RTOs, which differ in how demand-response resources (DRR) may participate in their markets. Some RTOs permit DRR to participate in their markets as voluntary reliability resources. For example, NYISO has an emergency demand-response program, which permits DRR to participate through an aggregator or other interface party, and receive energy payments for providing curtailments when called.

DRR also can participate in wholesale markets as capacity resources and receive advance reservation payments in return for their commitment to participate when called. Resources that fail to perform when called are penalized.

Finally, DRR can bid into RTO day-ahead (DA) markets as energy resources, specifying the hours, number of megawatts and price at which they are willing to curtail. ISOs set minimum bid values – NYISO's program has a $75/MWh floor. In New York, a resource scheduled in the DA market is obligated to curtail, and failure to perform results in a penalty.

Some of the RTO DR comes from individual entities; the rest is accumulated through third-party aggregators, or curtailment service providers (CSPs), who recruit customers too small to participate on their own, such as schools, commercial chains or groups of residential customers. In aggregating small customers, CSPs have increased customer participation in many wholesale reliability and emergency programs. In NYISO's two incentive-based programs, CSPs increased their share of subscribed DRR to 77 percent in 2008 from 44 percent in 2003. CSPs were responsible for more than 60 percent of total DRR capacity, and 70 percent of new DRR in ISO-NE.

Demand-Response Use in Planning and Operations

Different DR programs can be used at various times to support planning and operations (see graphic, page 50). Energy efficiency programs that reduce baseload or peak demand over the long-term are incorporated into system planning. Dispatchable programs that are quickly implemented and targeted for short-term peak reductions – such as direct load control – lie on the other end of the spectrum, and are used in the moment of operation.

Electricity Supply and Delivery

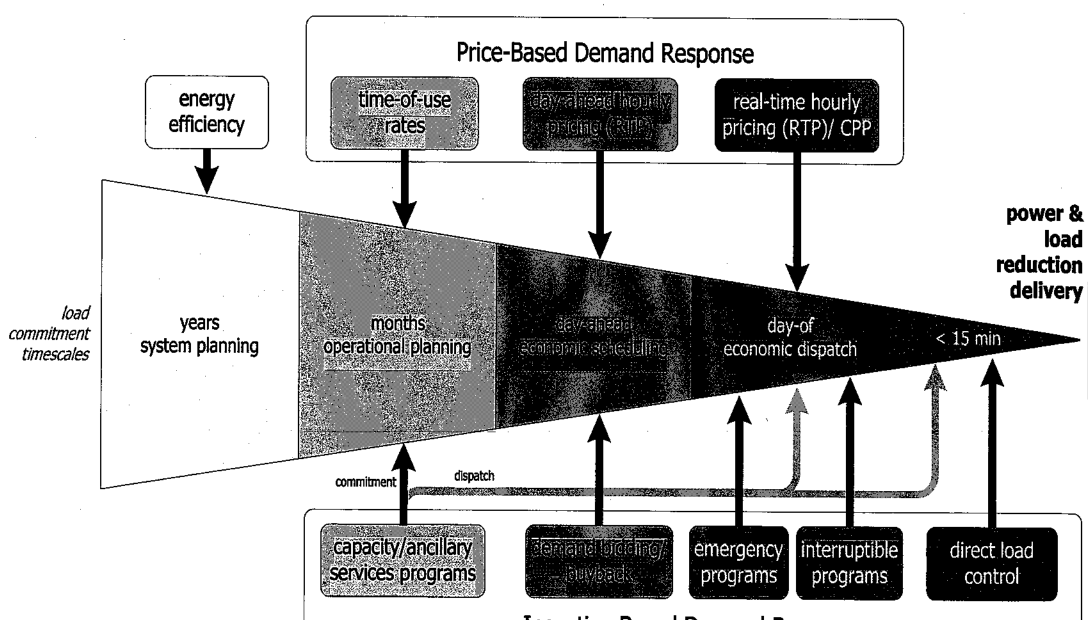
Unlike many other products, electricity cannot be stored in any appreciable quantities. Further, electricity is a necessity for most consumers, whose use responds little to price changes. Finally, electric equipment and appliances are tuned to a very specific standard of power, measured as voltage. Deviations in voltage can cause devices to operate poorly or may even damage them. Consequently, the supply side of the electric market must provide and deliver exactly the amount of power customers want at all times, at all locations. This requires constant monitoring of the grid and close coordination among industry participants.

Electricity service relies on a complex system of infrastructure that falls into two general categories: generation and the delivery services of transmission and distribution. Together, the power generation and high-voltage transmission lines that deliver power to distribution facilities constitute the bulk power system. Transmission and distribution facilities are also referred to as the power grid. These are coordinated and at times operated by a grid coordinator.

Nationally, the grid is split into three main sections – the Western, Eastern and Texas Interconnections. These sections operate independently and have limited interconnections between them.

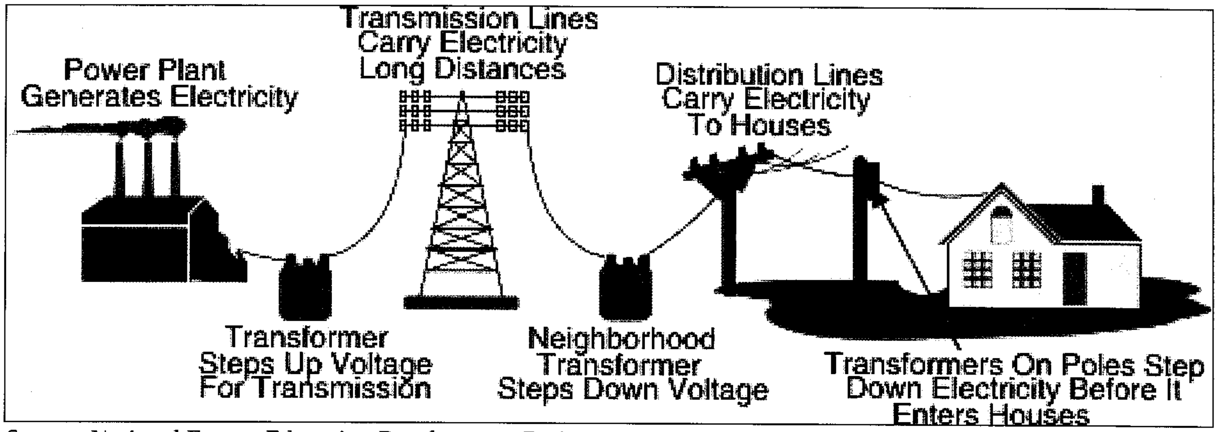
The nation, along with Canada and a small part of Mexico, is also divided into regional entities, (see map, page 51). The regional reliability entities fall under the purview of the North American Electric Reliability Corp. (NERC), which was designated by FERC as the nation's energy reliability organization and which develops standards, among other things, to ensure the grid's reliability. The standards, once issued by FERC, must be met by all industry participants – the standards are mandatory and enforceable. Consequently, the grid is designed and operated to meet these standards.

Demand-Response Program Use in Electric System Planning and Operations



Source: DOE

Electricity Service Schematic



Source: National Energy Education Development Project

NERC's regions include:

- Florida Reliability Coordinating Council (FRCC),
- Midwest Reliability Organization (MRO),
- Northeast Power Coordinating Council (NPCC),
- Reliability First Corporation (RFC),
- SERC Reliability Corp. (SERC),
- Southwest Power Pool (SPP),
- Texas Reliability Entity (TRE) and
- Western Electricity Coordinating Council (WECC)

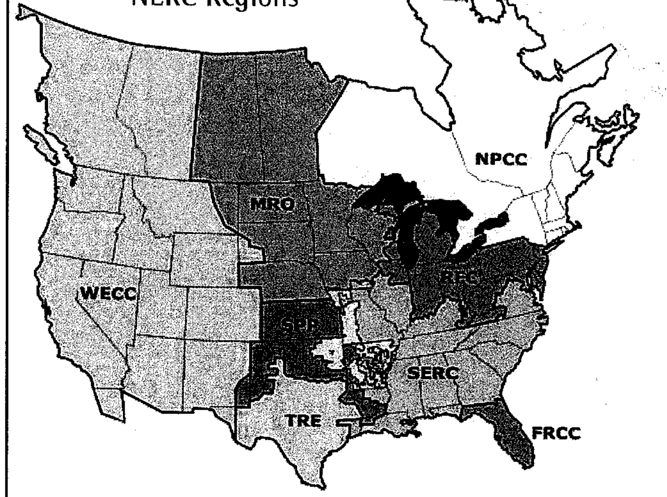
Generation

Power generators are typically categorized by the fuel they use and subcategorized by their specific operating technology. The United States has more than 1,000 gigawatts (GW) of total generating capacity. Coal, natural gas and nuclear dominate the power generation market.

Power plants each have differing costs and operational characteristics, both of which determine when, where and how plants will be built and operated.

Plant costs fall into two general categories: capital investment costs, which are amounts spent to build the plant, and operational costs, the amounts spent to maintain and run the plant. In general, there is a trade-off between these expenses: the most capital intensive plants are the cheapest to run – they have the lowest variable costs – and, conversely, the least capital intensive are more expensive to run – they have the highest variable cost. For example, nuclear plants produce vast amounts of power at low variable costs, but are quite expensive to build. Natural gas-fired combustion turbines are far less expensive to build, but are more expensive to run.

NERC Regions



Grid operators dispatch plants – or, call them into service – with the simultaneous goals of providing reliable power at the lowest reasonable cost. Because various generation technologies have differing variable costs, plants are dispatched only when they are part of the most economic combination of plants needed to supply the customers on the grid. For plants operating in RTOs, this cost is determined by the price that generators offer. In other areas, it is determined by the marginal cost of the available generating plants.

Construction of different generating technologies is subject to a number of issues, including community concerns, regional emission restrictions and the availability of fuels or other necessary resources:

- Wind plants are generally built in areas with the appropriate meteorological conditions. In most cases, these sites are located in rural areas with limited transmission access. For example, in West Texas, the transmission lines connecting wind farms with consumer centers in Dallas and Houston can become overloaded, requiring generators to curtail production.
- Coal plants have environmental characteristics that limit both their siting and operations. Specifically, they emit NO_x, SO_x, particulates, mercury and substantially higher levels of CO_2 than gas-fired plants. This has made financing these plants and siting them near urban centers difficult.
- There have been virtually no new nuclear plants built in the United States in the past 30 years. The technology of older plant designs became a source of concern following the accident at the Three Mile Island plant in the United States in 1977, the Chernobyl plant meltdown in Ukraine in 1986 and the Japanese earthquake, tsunami and nuclear plant destruction in 2011. New plant designs have been put forward over the past few years and are expected to be very expensive and controversial to build. Further, the disposition of high-level radioactive waste remains an unresolved problem, and the waste remains at plant locations.

Conventional Generation

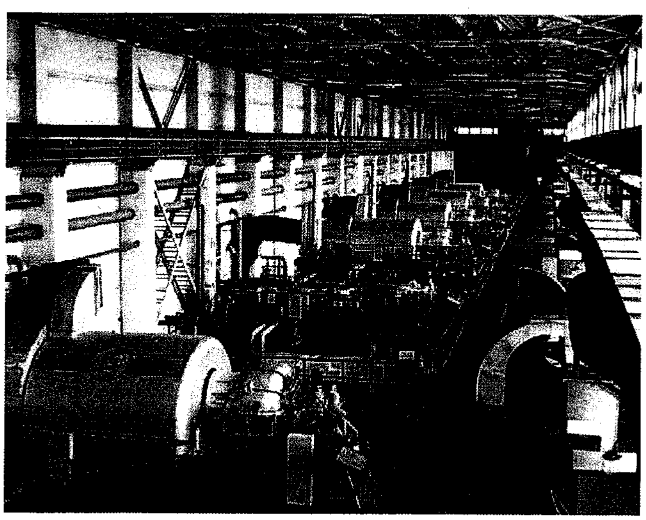
Natural gas power plants: These feature three major technologies, each with its distinct set of market advantages and limitations. They are steam boilers, gas turbines and combined cycle generators. Natural gas fuels nearly a third of electricity generation.

Steam boiler technology is an older design that burns gas in a large boiler furnace to generate steam at both high pressure and a high temperature. The steam is then run through a turbine that is attached to a generator, which spins and produces electricity. Typical plant size ranges from 300 MW to 1,000 MW. Because of their size and the limited flexibility that is inherent in the centralized boiler design, these plants require fairly long start-up times to become operational and are limited in their flexibility to produce power output beyond a certain range. Furthermore, these plants are not as economical or easy to site as newer designs – which explains why none has been built in recent years.

Gas turbines (GT) are small, quick-start units similar to an aircraft jet engine. These plants are also called simple cycle turbines or combustion turbines (CT). GTs are relatively inexpensive to build, but are expensive to operate because they are relatively inefficient, providing low power output for the amount of gas burned, and have high maintenance costs. They are not designed to run on a continuous basis and are used to serve the highest demand during peak periods, such as hot summer afternoons. GTs also run when there are systemwide shortages, such as when a power line or generator trips offline. GTs typically have a short operational life due to the wear-and-tear caused by cycling. The typical capacity of a GT is 10-50 MW and they are usually installed in banks of multiple units.

Combined cycle power plants (CCPPs) are a hybrid of the GT and steam boiler technologies. Specifically, this design incorporates a gas-combustion turbine unit along with an associated generator, and a heat recovery steam generator along with its own steam turbine. The result is a highly efficient power plant. They produce negligible amounts of SO_2 and particulate emissions and their NO_x and CO_2 emissions are significantly lower than a conventional coal plant. CCPPs, on average, require 80 percent less land than a coal-fired plant, typically 100 acres for a CCPP versus 500 acres for comparable coal plant, and CCPPs also use modest amounts of water, compared to other technologies.

Coal plants: These generate more than one-third of the electricity in the United States. These facilities tend to be large, baseload units that run continuously. They have high initial capital costs and are also somewhat complex in their design and operations. However, coal plants have low marginal costs and can produce substantial amounts of power. Most of the coal-fired plants in the United States are owned by traditional utility companies and located in the Southeast and Midwest.



Oil-fired plants: These play a minor role in U.S. power markets. These facilities are expensive to run and also emit more pollutants than gas plants. These plants are frequently uneconomic and typically run at low capacity factors. Like gas-fired generators, there are several types of units that burn oil; primarily, these are steam boilers and combustion turbines.

Generally, two types of oil are used for power generation: number 2 and number 6 (bunker) fuel oil. Number 2 is a lighter and cleaner fuel. It is more expensive, but because it produces fewer pollutants when burned, it is better for locations with stringent environmental regulations such as major metropolitan areas. Conversely, number 6 fuel oil is cheaper, but considered dirty because of its higher emissions. It is highly viscous (thick and heavy) and it comes from the bottom of the barrel in the refining process.

Nuclear plants. These provide roughly 20 percent of the nation's electricity; there are 104 operating plants with a total capacity of 100 GW. These plants are used as baseload units, meaning that they run continuously and are not especially flexible in raising or lowering their power output. Nuclear plants have high capital and fixed costs, but low variable costs, which includes fuel cost. They typically run at full power for 18 month, which is the duration of a unit's fuel cycle. At that point, they are taken off-line for refueling and maintenance. Outages typically last from 20 days to significantly longer, depending on the work needed.

Following the Three Mile Island plant accident in 1977, there was a cessation in the development of new plants. Most projects under construction in 1977 were finished, albeit with tremendous cost overruns. The last unit built in the United States came online in 1996.

Renewable Generation

Renewable resources use fuels that are not reduced or used up in the process of making electricity. They generally include biomass, geothermal, hydropower, solar, onshore and offshore wind, hydrokinetic projects, fuel cells using renewables and biogas.

Renewable generation provides a small percent of total U.S. capacity and generation. Even in 2009, when total U.S. electric output fell, average renewable generation grew 12% and wind output grew 28%, spurred by state regulations and federal tax credits. As renewable generation becomes a larger percentage of generation resources, integrating them into the operating power grid has presented challenges.

Capacity: Wind and solar capacity have grown faster than other renewable resources in recent years. Geothermal has more installed capacity than solar, but is growing more slowly. Wind added the second highest amount of capacity after gas-fired generation in 2009 and 2010.

Additions are usually reported in megawatts of nameplate capacity. Actual capability varies from the nameplate for any unit type due to age, wear, maintenance or ambient conditions. But as renewable resources are often weather-dependent, their capacity factors – the ratio of average generation to the nameplate capacity for a specific period – have been much lower (as low as 30 percent) than for fossil-fuel-fired generation. Markets care about the difference between nameplate and capacity factor values when they evaluate capacity available to cover expected load. Prior to sufficient operating experience with a renewable technology, markets usually estimate capacity value conservatively.

Average capacity factors for new renewable resources in early 2010 were 34-35 percent for large wind projects, 74 percent for geothermal, 15-21 percent for commercial solar photovoltaic (PV), 35 percent for concentrating solar power (CSP) without storage (or 43-45 percent with storage technology) and 63 percent for biomass. Capacity factors have risen with technological innovation and improved manufacturing processes.

Characteristics: Wind power is the fastest-growing renewable resource, in part due to earlier cost declines and technology improvements as well as earlier receipt of federal tax credits. A 1.5-MW wind turbine was the most frequently installed size in 2009, although 2.3-MW turbines later became more common. Wind is largely pollution-free and can be located on farms and ranches.

Because the best wind resources are often far from load centers, insufficient transmission presents a challenge to delivering its output. Other market challenges for future wind development include its variable output, which is often inversely correlated to demand (seasonally and daily); system operators' inability to dispatch wind resources to meet load increases; difficulties related to accurately forecasting its ramping; and the need for companion generation (usually fossil-fueled) to be available to balance wind generation when the wind is not blowing.

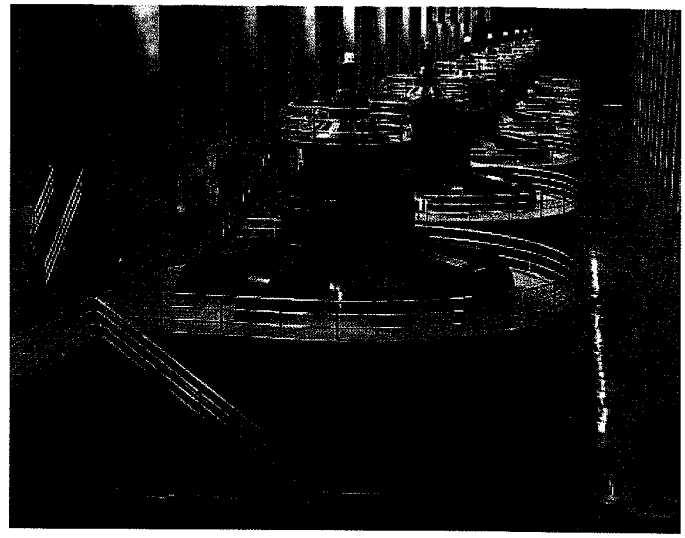
Geothermal energy taps into reservoirs of steam and hot water deep beneath the earth's surface to produce power. The best resources are in the intermountain West. Geothermal potential is determined by thermal conductivity, thickness of sedimentary rock, geothermal gradient, heat flow and surface temperature. While geothermal power was less than 0.4 percent of U.S. generation in 2009, it was 11 percent of nonhydro renewable output in April 2010. California hosts more than 80 percent of U.S. operating capacity. The five states with the most geothermal capacity in development are California, Nevada, Oregon, Utah and Idaho.

Solar energy transforms sunlight into electricity using one of two technologies: photovoltaic (PV) or concentrating solar power (CSP). PV modules, or panels, transform sunlight directly into power using silicon wafers or nonsilicon thin-film technologies. They can be installed on roofs of buildings or at ground-level PV farms. CSP plants use a two-step process to transform the sun's energy. First, mirrors direct sunlight towards a receiver that captures the heat. CSP then employs a thermal process to create steam, driving an engine or turbine to produce electricity. CSP plants, which are dispatchable, can include low-cost energy storage that extends their availability later in peak hours.

PV growth has been highly concentrated as a result of state policy incentives: 10 states have 95% of PV capacity; California alone has more than 50 percent. Annual PV additions rose to more than 100 MW beginning in 2006, spurred by tax incentives in the Energy Policy Act of 2005. More recent growth was spurred by falling costs, technology innovation, expanded federal tax benefits and an increase in state policies promoting investment.

In 2010, 432 MW of CSP was operational and 81 MW was under construction. Seven western and southwestern states have extensive CSP potential: Utah, New Mexico, Arizona, Nevada, Texas, California and Colorado. Developing that potential will require overcoming challenges of siting, transmission and the need for extensive water supplies to clean mirrors.

Hydroelectric power is created when the kinetic energy of falling water drives turbine generators, which convert the energy into electricity. There are two types of hydroelectric projects: conventional and pumped storage. Conventional projects, which use a dam in a waterway, can operate in a run-of-river mode, in which water outflow from the project approximates inflow, or in a peaking mode, in which the reservoir is mostly drained to generate power during peak periods when energy is more valuable. Pumped storage



projects use bodies of water at two different elevations. Water is pumped into elevated storage reservoirs during off-peak periods when pumping energy is cheaper; the water is then used to generate power during peak periods as it flows back to the lower elevation reservoir. Pumped storage is the only significant commercially deployed electricity storage technology available today.

Biomass includes many waste byproducts, such as agricultural residues, landfill gas, municipal solid waste and wood resources. The largest biomass category is wood waste, burned for heat and power in the lumber, pulp and paper industries. Challenges to biomass production include impacts on food supplies (for example, converting corn into ethanol), conserving natural resources and minimizing water pollution. State policies on renewable generation differ on eligibility of biomass technologies.

Biogas energy is created through the anaerobic (without oxygen) bacterial decomposition of manure, which is turned into a gas containing 60-70 percent methane. Biogas recovery can be installed at farms anywhere, used to run farm operations and reduce methane emissions from natural manure decomposition.

Renewable Energy Policies

Renewable development is frequently tied to policies promoting their use because of their higher cost relative to other technologies. Financial incentives include tax credits, low-cost loans, rebates or production incentives. Federal funding of research and development (R&D) has played an important role in lowering costs or reducing the time it takes for renewable technologies to become commercially viable.

Congress has passed tax incentives to spur renewable resource investments. Wind, biomass, geothermal, marine and hydrokinetic project developers can receive federal production tax credits (PTC) based on a facility's production. It is an inflation-adjusted credit that runs for 10 years from the date a facility goes online. Initially set at 1.5¢/kilowatt hour (kWh), its value in 2010 was 2.1¢/kWh. To qualify, a facility has to be operational before the PTC expires.

Unlike other renewables projects, solar projects are eligible for a federal investment tax credit (ITC), worth 30 percent of a project's equipment and construction costs. The ITC begins the year a project starts commercial operation and depreciates over five years. The Emergency Economic Stabilization Act (2008), the legislation enacted to help shore up the U.S. financial system, extended the solar ITC for eight years, through 2016. It also expanded both the PTC and ITC to include utilities, which were previously ineligible for these credits. This change created the impetus for a model of utility owned and operated renewable generation, and led to a flurry of project announcements.

Provisions of the American Reinvestment and Recovery Act (ARRA) of 2009 extended the PTC and gave developers new options. It extended the credit for wind to 2012 and for other eligible technologies to 2013, and gave PTC-qualified developers the option to claim the 30 percent ITC on a project-by-project basis for the PTC's current duration.

Due to the economic crisis, ARRA gave developers another option for projects that began construction by the end of 2010 – they could apply for Treasury-administered cash grants, which monetized the ITC's value up front. ARRA funds helped support renewable energy research and development and aided capacity growth in 2009, despite the economic downturn.

State renewable portfolio standards (RPS) and renewable energy standards (RES) have been significant drivers in the growth of investment in renewable generation. An RPS requires a certain percentage of energy sales (MWh) to come from renewable resources. Percentages usually increase incrementally from a base year to an ultimate target. Currently, 29 states plus Washington, D.C., have an RPS and six states have renewable goals without financial penalties for nonachievement. As utilities build more renewable-powered generation, the markets in which they participate continue to address the integration of renewable output into their day-ahead and real-time operations and model expected growth as part of their long-term transmission-planning processes.

To encourage the development of distributed generation (DG), or the production of electricity at the site of consumption, and solar power, 16 states plus Washington, D.C., created RPS carve-outs or set-asides to give an extra boost to these resources, which are not yet cost-competitive with other renewables.

Renewable energy certificates (RECs) allow state regulators to track compliance with mandatory RPS targets or verify progress in voluntary state renewable programs. They also allow compliance entities to purchase credits – subject to state imposed limits on amount and price - if they have not generated or bought enough renewable energy to meet their annual requirements. Each reported megawatt-hour (MWh) of eligible generation results in a system-issued REC with a unique identification number to prevent double-counting. Each REC includes attributes such as generator location, capacity, fuel-type and source, owner and the date when operations began.

States and local utilities offer a variety of financial incentives for renewable energy to complement policy mandates. These include tax credits for in-state manufacture of renewable energy equipment, consumer rebates for purchase and installation of renewable generation or production incentives. Production incentives include extra credits for solar output based on RPS solar set-asides and feed-in tariffs.

Five states, two municipal utilities and Ontario, Canada, adopted or expanded feed-in tariffs (FITs) in 2009 to support their energy and environmental goals. Also called feed-in rates or advanced renewable incentives, these programs typically are designed to encourage development of new small- and medium-sized renewable generation projects by residential and independent commercial developers.

FITs require utilities to buy the renewable generation at a fixed rate that is higher than that provided to other generators, under multiyear contracts. This enables smaller distributed renewable generators to avoid having to participate in renewable portfolio standard (RPS) auctions or other competitive procurements and compensates them for more expensive technologies. The utility passes the costs of the program to its customers.

Transmission

The alternating current (AC) power grid operates like an interconnected web, where, with a few exceptions, the flow of power is not specifically controlled by the operators on a line-by-line basis. Instead, power flows from sources of generation to consumers across any number of lines simultaneously, following the path of least resistance. There are a limited number of direct current (DC) lines, which are set up as specific paths with definite beginning and end points for scheduling and moving power. These lines are controllable by operators and have other characteristics that make them attractive to grid planners and operators, such as providing greater grid stability and lower line losses. However, DC lines cost significantly more than AC lines to construct. Consequently, DC lines are typically built for certain specialized applications involving moving large amounts of power over long distances, such as the Pacific Intertie, which extends between the Northwest and California.

Transmission lines provide a certain amount of resistance to the flow of power as electricity travels through them. This resistance is not unlike the wind resistance that a car must overcome as it travels along a highway. The resistance in power lines creates losses: the amount of power injected