CHAPTER 5

Economics of Power Generation

Manfred Hafner and Giacomo Luciani

Electricity can be generated either chemically (as in photovoltaic panels) or, more frequently, mechanically, through the rotary movement of a generator (a magnet moving within a net of cables). The needed rotary movement can be obtained by the force of steam expanding at high temperature, water flowing, or wind blowing in a turbine; or even by using a regular internal combustion engine. The high temperature needed to raise steam can be derived from burning coal, oil, gas, waste and biomass; from controlled fission in a nuclear reactor; by concentrating solar radiation; or by extracting heat from the earth crust. The bottom line is that there are numerous solutions to generate electricity, and each of them has specific characteristics that render it more adapted to the specific conditions and circumstances where and when electricity is required.

In order to provide a satisfactory treatment of power generation technology and economics, a single chapter would have expanded beyond a practical dimension: accordingly the discussion has been divided into a general

---

M. Hafner (✉)
SciencesPo—Paris School of International Affairs, Paris, France

Johns Hopkins University SAIS Europe, Bologna, Italy

Fondazione Eni Enrico Mattei (FEEM), Milan, Italy
e-mail: Manfred.hafner@sciencespo.fr

G. Luciani
SciencesPo—Paris School of International Affairs, Paris, France

Graduate Institute of International and Development Studies,
Geneva, Switzerland
e-mail: giacomo.luciani@sciencespo.fr

M. Hafner, G. Luciani (eds.), *The Palgrave Handbook of International Energy Economics*, https://doi.org/10.1007/978-3-030-86884-0_5

introduction and a sequence of specific chapters each devoted to a different generation solution: thermal power based on fossil fuels (coal, oil, and gas)—Chap. 6; thermal power based on nuclear fission—Chap. 7; hydroelectricity—Chap. 8; solar power—Chap. 9; wind power—Chap. 10; geothermal power—Chap. 11; and power from tides and waves—Chap. 12. In this introductory chapter, we touch upon the major economic differences between these multiple solutions, highlighting the comparative advantages and disadvantages of each. In the end, a well-functioning electricity system will always necessitate a combination of different technologies assorted in an appropriate way to satisfy a range of situations that are expected to arise in time.

When discussing electricity and comparing different power generation technologies and their properties, the reader must first and foremost keep in mind the distinction between *capacity* (or power) and *energy (or electricity generated* or consumed). Capacity (or power) is the electricity that a generation plant produces (or an electricity device consumes) instantaneously. It is measured in watts, kilowatts (1 kW = 1000 watts), megawatts (MW: 1 MW = 1000 kW), and gigawatts (1 GW = 1000 MW). The installed (or nominal) capacity of a power plant is (generally) the maximum capacity of a power plant. The amount of electricity a power plant produces (or an electricity device consumes) over a given time is measured in kilowatt-hours (kWh). Kilowatt-hours are determined by multiplying the number of kW produced (required) by the number of hours of production (use). Energy is thus the amount of electricity generated (consumed) over time and is measured in watt-hour, kilowatt-hour (kWh), megawatt-hour (MWh), gigawatt-hour (GWh), and terawatt-hour (TWh).

Unlike coal, oil, or gas, electricity cannot be stored easily. It must thus be generated and delivered at the precise moment it is needed. The most important element to be considered when addressing power generation is the *demand load curve*. A load curve shows the variation of load (in kW or MW) over time (in hours). The load curve can be plotted for 24 hours a day, it is then called a daily load curve; if one year is considered, it is called annual load curve. The load curve is important because the electricity capacity demanded by consumers (industry, residential, and commercial) varies over time. Typically, industrial activities are the highest during the day, commercial activities are high during the day and the early evening hours, and residential activities are high mainly in the evening when everybody is at home and turns on the lights, watches television, and uses other electric devices.

The resulting daily load curve of a country is one with a low level during the night and a higher level during the day with some peaks either during the day or in the evening hours. Moreover, the load curve differs from day to day (on weekends and festivities when industrial activities are reduced, the load curve is generally lower) and across seasons (in cold climates, electricity load is higher during the winter months due to heating, while in hot climates, it may be highest in summer months due to cooling). Some high-income countries with a relatively temperate climate may nowadays have two seasonal peaks: a winter peak due to heating and a summer peak due to cooling. The load curve thus differs from country to country due to cultural and meteorological differences. The integral (surface) below the load curve represents the electricity demand (electricity demand = capacity × time = kW × h = kWh).

All of this is of utmost importance because the load curve will define the amount of the electricity demand which is *base load* (load needed all year), *peak load* (load needed only a few hours a day), and *intermediate load (or mid-load)* for operating hours between base load and peak load. Different power plants with their different repartition of capital cost and operating cost will be used to satisfy different load segments.

All power generation plants are relatively capital-intensive, in the sense that the initial investment costs are a significant and frequently dominant component of total cost; however, the *ratio of capital vs. operating costs* varies significantly: it is highest for nuclear, wind, solar, large coal-fired, and some hydropower plants and smallest for gas turbines or plants based on internal combustion engines. Nuclear, coal-fired, and hydropower plants with large reservoirs are available for many hours, and it is convenient to keep them in use for as many hours as possible, in order to amortize the very high capital cost over the maximum number of hours and reduce the unit cost of producing each kilowatt-hour. The same would be true for solar and wind, except that these plants are non-dispatchable; therefore, the operator cannot control the extent of their use. Consequently, nuclear, coal-fired, and some hydropower plants are optimal to meet base load demand. In contrast, gas turbine power plants or generators based on internal combustion engines are typically preferred for meeting demand peaks or dealing for emergency situations, for example, in islands or other isolated tourist destinations during the high season, or in hospitals in case electricity from the grid is no longer available. Hydroelectric plants with small storage will be used during peak hours due to the high opportunity cost of these plants.

A further distinction of importance is between *indirect* and *direct* operating costs. Indirect costs are related to the upkeep of the plant independently of how much the plant is being used and are typically incurred on a yearly basis. In contrast, direct costs are directly related to the utilization of the plant, for example, the cost of fuel in a coal- or oil-fired thermal plant or in a gas turbine plant. Indirect costs are fixed and fundamentally unavoidable, while direct costs are directly related to the production of power. Hence, both capital and indirect costs are not part of *marginal* cost, which is the cost of producing one additional KWh of power, and exclusively reflects direct costs. Hence, some technologies, notably solar, wind, and most hydropower plants, have zero marginal costs, and nuclear has low marginal costs because the cost of fuel per kWh produced is very small. To the opposite extreme, gas turbines or internal combustion engine-based plants have significant marginal costs and will only be started if demand justifies it.

A generation plant will not always generate at full capacity: there will be times when it generates at less than full capacity, and times when it is not in use and does not generate at all. This may be due to the load curve or intrinsic non-availability by some plants (the most obvious are solar and wind availability for plants relying on these energy sources).

How much electricity will be produced by a plant of a given installed (or nominal) capacity depends on the number of hours that the plant is available for production (*availability*) and the number of available hours that the plant is actually in use.

There are 8760 hours in a year, and no power plant can be available throughout the year. Some plants may be available most of the time in a year (coal-fired, nuclear, biomass, geothermal, run-of-river hydroelectric plants, or hydroelectric plants with a very large reservoir of water) and may be relied upon for close to 8000 hours (due to maintenance and other outages they cannot operate all hours of the year).

Other plants, in contrast, are necessarily limited in their availability: for example, solar photovoltaic panels only produce electricity during the day and will produce very little when the sun is low over the horizon or it is covered by clouds, meaning that even in the best imaginable conditions, a photovoltaic panel cannot possibly reach 3000 hours of availability, and in many locations may be available for as little as 1000 hours. Similarly, a hydropower plant with a small reservoir (e.g. in the Alps) may only be able to produce electricity at nominal capacity for 2–3000 hours in the year.

The difference between the last two cases is that in the case of a hydropower plant the operator may normally decide when to use the plant using an opportunity cost approach (i.e. to decide when to "spend" the limited plant's hours of availability in order to maximize revenues), while in the case of solar photovoltaic the operator has no control at all on the availability, and electricity may be produced when it is needed, but possibly also when it is not needed. Hence a further key difference between various technologies is *dispatchability*. Some technologies (notably thermal power plants, independently of the source of heat, and hydropower plants with large reservoirs) are fully dispatchable, in the sense that the operator decides when the plant is in operation. At the opposite extreme, some technologies are not dispatchable at all (wind, solar, and hydro run-of-the-river, i.e. with no reservoir), and electricity is produced when the appropriate natural conditions exist, and not at other times.

This distinction is important because, as already mentioned, the demand and supply of electricity in a grid must be balanced at all times in real time. The power producer (or the manager of the grid, called a Transmission System Operator or TSO) has little or no control over demand and must adjust supply to demand—a task made considerably more difficult if power is produced from non-dispatchable technologies. The details of this are discussed in Chap. 13 on the economics of networks, and the integration of non-dispatchable renewables is discussed in Chap. 15. The issue of non-coincidence of demand and supply highlights the importance of electricity storage, which is limited and expensive: this is discussed in Chap. 14 as far as battery storage is concerned and in Chap. 8 as far as pumped storage is concerned (so far the only way to store electricity by converting it to potential energy).

Producing peak load electricity is more expensive than producing base load electricity; in the first case, an equipment needs to be built which only runs a few hours a day, and in the latter case, investments can be amortized producing electricity almost all year. Even though in physics a kWh is equal to a kWh regardless when and where it is consumed, in economic terms a kWh is not equal to a kWh. The cost of producing a kWh depends on the moment when it is consumed and thus when it needs to be produced since it cannot be easily and cheaply stored.

This time element of demand (and thus production) is relevant not only for *power plant dispatching*, but also for *future capacity planning*. If overall demand of electricity in a country increases by a certain amount of TWh but most demand increase is expected to happen during peak hours, the required power plant investments will be fundamentally different compared to the case where the demand increase happens mainly during low load hours thus increasing base load.

A distinction needs to be made between capacity investment planning in order to satisfy future electricity demand evolutions and dispatching existing and available power plants for the hour or day ahead. For future capacity planning, a full cost (or *long-run marginal cost*) approach needs to be taken (including investment cost, operating and maintenance cost, fuel costs, and possibly the cost of carbon emissions—as well as possibly other costs aimed at internalizing environmental and other externalities), while for dispatching purposes, only the short-run marginal costs (fuel costs and other unit-based environmental costs) are taken into account. The choice of power generation technology (and thus energy) being used on a given moment of the day depends thus on the merit order (marginal costs) of the different power plants to satisfy demand. For dispatching purposes, all fixed costs are to be considered sunk cost.

With increasingly large shares of non-dispatchable power generation sources in electricity producing systems, flexibility mechanisms become of utmost importance. Non-dispatchable power generation means (e.g. wind and solar) are always first in the merit order, thanks to their zero short-run marginal cost, but they are largely not reliable in the sense that whenever the sun shines and the wind blows, you will use them, but whenever the sun does not shine and the wind does not blow, they are not available. In fact, dispatchable power plants no longer need to follow the "demand load curve" as defined by consumers, but the so-called net load curve, that is, the difference between the load curve as demanded by consumers and the electricity produced by non-dispatchable zero marginal cost electricity (mainly solar and wind). The net load curve is much less predictable and has much higher ramp up and ramp down requirements compared to the load curve of consumers. Needed *flexibility* mechanisms include (i) the capability of power plants to ramp up and ramp down quickly (storage hydroelectric and, to a slightly lesser extent, gas turbine power plants can ramp down/up very quickly, while steam turbine-based power plants [in particular large coal and even more so nuclear plants] are not well suited for fast ramp up/down of power output), (ii) interconnections to neighboring electricity systems, (iii) storage (so far mainly pump-storage, but in the future possibly to some extent also batteries), (iv) electricity demand side management (in particular demand response), and (v) sector coupling (e.g. power to heat, power to gas, power to vehicles).

A further differentiating characteristic is *size*, as measured by the plant's capacity. For some technologies, notably coal-fired and nuclear plants, economies of scale are potentially very important, favoring the construction of very large power plants (in excess of 1 GW of capacity). However, nuclear power plants can also be medium or small size (including less than 100 MW), and in fact, there is growing interest toward such smaller nuclear alternatives. Gas turbine-based plants can be small (gas turbines—GT) or medium size (combined cycle power plants—GTCC). Individual wind turbines are small (today up to 10 MW) and individual photovoltaic panels very small. Hydropower plants can be of all sizes: the largest power plants in the world are hydroelectric, but hydro solutions are available also for very small applications in locations where the grid does not reach.

Another relevant dimension of size is *space occupation* and the physical impact on the immediate environment. Hydropower plants with large reservoirs may entail the flooding of vast surfaces and the need to relocate large numbers of people, an obvious drawback. Large solar power plants also occupy very large surfaces for relatively limited capacity, an obstacle to their deployment in cultivated, forested, or inhabited spaces that are in demand for other purposes. This is one of the reasons why large solar power plants tend to be proposed for desert regions, where space has limited alternative potential use (the other reason is that in dry desert areas solar radiation is very high). To the opposite extreme, nuclear power plants are very small relatively to the very large capacity that they can reach, especially where several plants are grouped in a single location, as is frequently the case.

A final important characteristic is *locational constraints*. Some technologies are available only in specific locations, this being most evidently the case of hydro, but conditions for wind and solar are also greatly variable depending on latitude, meteorology, and orography. This is important because electricity is expensive and difficult to transport over long distances, and plants must be sized in view of the total demand that they can effectively reach and satisfy economically. Thus, some very promising locations for hydro, wind, and solar remain underexploited or unexploited because no demand is geographically close enough to justify creating transmission capacity and a generation plant.

In contrast, thermal power plants are extremely flexible from the point of view of their localization, as they basically only need proximity to a body of water for cooling purposes. Historically, this has allowed industry to be localized in the proximity of markets, or where other factors of production, notably labor, are present at low cost; while in past centuries, when energy was predominantly available in kinetic form, industry clustered in the proximity of energy sources (mostly flowing water). With increased reliance on renewable sources (solar, wind, and hydro) the pendulum may, at least to some extent, swing back to localizing industry close to the source of energy, with potentially momentous consequences on the international distribution of industrial production, especially in sectors that are highly energy-intensive.