CHAPTER

1 INTRODUCTION TO SYNTHETIC SOLAR IRRADIANCE

Jamie M. Bright

Bright, J. M. (2021), "Introduction to synthetic solar irradiance," in Bright, J. M. (ed.), *Synthetic Solar Irradiance: Modeling Solar Data*, Melville, New York: AIP Publishing, pp. 1-1–1-32

1.1 INTRODUCTION

The target audience of this book is anyone with an interest in solar energy engineering or specifically to the emerging field of synthetic solar irradiance. Synthetic solar irradiance is, in essence, generating fake solar data that are representative of measured solar data. Generating fake data is often necessary as we do not measure high-resolution solar data with any great regularity, yet it is often a requirement for studies and applications surrounding solar engineering. If this is a topic that interests you, then you are reading the correct book.

The field of solar energy engineering has—like many other research fields—transcended any single discipline of education. It boasts contributions from a myriad of core competencies and specialties ranging from statistics, through computer science, to atmospheric chemistry. Substantial advancement of the solar energy field has been unlocked through interdisciplinary collaboration, and so it is unfair to assume that you, the reader, has all the prerequisite knowledge specific to solar energy. One hopes, in fact, that this book welcomes new minds with new ideas to the field so that you can apply your knowledge—if this applies to you, dear reader, welcome! If you do have a strong background in solar energy modeling, my congratulations on an excellent choice of research field; I hope you find this chapter interesting and, more important, accurate. That said, in this chapter, I target those readers in need of a refresher or those who need an introduction to solar energy modeling concepts that are needed to get the most out of this book. If you are confident with fundamental terms of solar irradiance, then you are encouraged to jump to the section setting out what exactly synthetic solar irradiance *is*. Some of the later chapters are quite mathematically intense; readers are expected to have a reasonable grasp of statistics.

Synthetic Solar Irradiance, Jamie M. Bright

This introductory chapter really does go back to the beginning in Sec. 1.2 where a fun and brief history of solar energy is presented; a fair warning must be issued, as I am no historian and I embellish, speculate, and skip vast swathes of development. The brief history is mainly to emphasize the importance of solar energy in society. Section 1.3 introduces the core concepts and definitions of the earth–sun geometry and how we calculate where the sun is in the sky. Section 1.4 introduces the reader to the fundamentals of irradiance where key terms, principles, and relationships are discussed. Section 1.5 extends this knowledge further and introduces some of the most common indices used to characterize and leverage solar irradiance data. Next, Sec. 1.6 delivers the context of synthetic solar irradiance, what it is and why it is important. Some requisite skills are discussed in Sec. 1.7, misconceptions in Sec. 1.8, and the book objectives in Sec. 1.9.

The main objective of this book is to enable new researchers to enter the domain of synthetic solar irradiance, and to inform possible industrial stakeholders of the benefits and limitations of the techniques. By the end, one should be able to understand the key concepts and techniques of synthetic solar irradiance and know how to enter the research field themselves through either direct contribution of new techniques or perhaps adoption of existing methodologies. On behalf of the many contributing authors to this book, we thank you for your interest and hope you find our collection of ideas useful, insightful, and empowering.

1.2 A BRIEF HISTORY OF SOLAR ENERGY

The basis of all life on earth[1] is a direct consequence of radiation from the sun (Mertens, 2013). Each hour the earth receives 1.73×10^{12} kWh of solar energy. This energy is responsible for driving weather patterns and for providing heat and nourishment for all life on the planet. Of this vast quantity of received solar energy, only a very tiny portion is harnessed through land-based solar energy technologies and converted to heat or electricity (Neville, 1995). It is estimated that a "small" portion of the energy received would more than satisfy the annual energy demand of the human race (Solanki, 2009). This comes with the caveat that it can be delivered to where it is needed of course, but the potential is striking.

Solar energy can be used—and by this we mean leveraged to perform a function in our lives—in a number of ways through either passive or active means. An example of passive solar energy usage is the deliberate architectural design of buildings to allow for space heating from the sun, whilst active solar energy is the use of solar energy as an input for manipulation, such as a photovoltaic (PV) solar cell (Thorpe, 2013). PV is an acronym that is used regularly throughout this book; it is intended to be understood as an electricity-generating solar panel or equivalent. It is typically used with the adjectives small-scale PV or residential PV to mean private solar installations on a home; it can also

[1] I am conveniently ignoring the fact that deep sea vents can support life without sunlight. This statement is true for the vast majority of life on earth.

be called a farm, plant, utility-scale, or large-scale to mean a significant installation with the purpose of generating and selling high quantities of electricity into the power system.

The sun provides society with heat, light, and edible vegetation, and so it is perhaps unsurprising that throughout history there are many examples of solar deities—a god that represents the sun. The Ancient Egyptian Pharaoh Akhenaten (1375–1358 B.C.) recognized the critical importance of the sun and held the sun god Aten in the highest of regard. He instated the earliest known single deity religion called Atenism when he raised Aten as the only god, excluding all others (of which there were many at the time) (Hornung and Lorton, 2001). More famously, there are myths and legends that describe the rising and setting sun as the role of a deity. Helios pulls the sun across the sky every day using a horse-drawn chariot; Ra ferried the sun in a barge. These are perhaps the earliest examples in recorded history that, in essence, describe a prediction of the daily variability of solar energy. Humans observed that the sun "traveled across the sky" every day, and to describe this observation, they used gods as an explanation.

There is evidence from early documented history that humans have passively utilized the sun's energy. We are not simply talking about warming up through the day, but to have deliberately planned to leverage solar energy to benefit oneself or society. Leveraging the knowledge of solar predictability, ancient Greek philosophers Socrates and Aristotle developed and taught that homes should be constructed in such a way that they could be heated by the sun. The Greeks further developed the use of glass in windows to enhance the sun's thermal effects (Forbes, 1957) for heating the home. They described what we now call passive solar heating, and ancient Greek towns were built accordingly (Morris, 2006). Not only were architectural practices followed, but the law itself was modified to enable access to solar energy. This passive solar heating practice was legally enforced in ancient Rome to ensure that homes had access to sunlight. This practice exists still in modern legal forms, for example, the "Rights to Light" is a form of easement in modern English law giving property owners the right to maintain their level of illumination without new infringement upon it (Kerr, 1865; and Law Commission). Passive solar heating and illumination is found in all modern developments, such as in architectural design coupled with the intricate detailing and calculation of thermal transfer and illumination processes in design stage (Behling and Behling, 2000; Bainbridge and Haggard, 2011; and Chwieduk, 2014).

Examples of scientific study into the active manipulation of sunlight have been documented from potentially as early as 212 B.C., when Archimedes purportedly protected the city of Syracuse when it was besieged by the Romans using parabolic mirrors that ignited advancing Roman ships—note that this is widely disputed (Simms, 1977). Depictions of this event (Fig. 1.1) show the active use of concentrated solar power redirected to target the ships. Whilst we may take with a large pinch of salt that these polished metal solar receivers did anything other than provide a serious glare at these ships, it isn't too much of a stretch to think that this equipment must have existed to provoke the story, and they were probably used to illuminate dark areas during daytime or perhaps as signals across the city.

FIG. 1.1
Depiction of the Siege of Syracuse, where Archimedes purportedly set fire to advancing Roman ships using parabolic mirrors and the sun rays. (Image taken from Wikimedia Commons.)

Concrete examples of active solar technology begin in 1515 with none other than Leonardo da Vinci, who designed a large curved mirror to concentrate energy for heating water in the cloth dying industry (Morris, 2006). Elements of this design are still leveraged in concentrating solar power applications around the world.

The first documented event of using solar energy in conjunction with electricity came much later. The photovoltaic effect was discovered in 1839 by 19-year-old French scientist Alexandre-Edmond Becquerel, who placed two platinum-coated electrodes in a container with an electrolyte (Wolfe, 2013). He observed that the strength of the measurable current changed with exposure to light (Mertens, 2013). Later, in 1887, Heinrich Hertz accidentally discovered the photoelectric effect while experimenting with electromagnetic waves between two conductors (Ostdiek and Bord, 2012). It was not until 1905 that Becquerel's and Hertz's discoveries were theoretically explained by Albert Einstein, who later earned himself the Nobel Prize in 1921 for services to theoretical physics, with particular contribution to the discovery of the fundamental laws of the photoelectric effect (Wolfe, 2013). Over time, laboratory-based materials were developed that leveraged the photoelectric effect converting sunlight to electricity, though the efficiencies of such prototypes were incredibly small. It was not until 1954 that the efficiencies of solar converting materials gained any substantial momentum.

For an entertaining and insightful read on the history of the modern photovoltaic age, the reader is referred to a broader perspectives paper by Joseph J. Loferski (1993), who was witness to the birth of the modern solar PV cell. Loferski points to three research papers (Chapin *et al.*, 1954; Rappaport, 1954; and Reynolds *et al.*, 1954) all published in 1954 that each describe positive-negative junction semiconductor devices capable of converting solar energy radiation into electricity with significantly improved efficiency on original designs. These three publications first demonstrated PV cells had the potential for large-scale power generation. RCA and AT&T's Bell Laboratory developed working PV solar cells with an 8% efficiency and were granted a patent for a "Solar Energy Converting Apparatus" in 1954, though commercial application at the time was entirely limited due to the high cost of manufacture (Wolfe, 2013). It was the space exploration program that noticed the excellent synergy between PV technology and space flight. With no atmosphere or weather in space to impede access to solar energy, using solar panels in the vacuum of space was a clear application. During the 1960s, significant research and development resources were committed to the technology's growth. PVs remained at substantial cost and were only considered plausible for the heavily funded space industry (Tiwari and Dubey, 2010). Loferski (1993) reports that the OPEC oil embargo of 1973 resulted in substantial investment into reducing the USA's dependence on oil imports. Of the identified opportunities, solar PV was highlighted and funding for its development came in abundance. Oil companies provided one of the earliest terrestrial commercial use cases for solar PV when they identified its usefulness in powering uninterruptible warning systems on offshore oil rigs. Electricity generation from solar power was identified as a key player (Breeze, 2014) and it remains so to this day. It is perhaps a little ironic that oil companies are the ones that funded and commercialized the fastest growing and most popular renewable energy technology that is a huge threat to their own business operations.

Today, solar energy technologies are increasingly popular for both large-scale investors or for private residential purposes. They are cheap, effective, simple to install, safe, they scale easily, and have an excellent public perception. Sunlight, the raw resource, is distributed freely to all countries with only climate and geography governing total availability. It is distributed without societal or political bias (if we are allowed to ignore any connection between society and climate), which is perhaps the most significant advantage of solar energy as an electricity generation resource. The current energy resource market is heavily dependent on the global trade of raw and processed resources. Solar radiation cannot be simply collected and sold for pre-dispatch, it must first be converted to electricity or heat. Heat can be stored for long periods of time, and electricity can be stored in various states; solar energy (by which I am generally referring to electricity throughout this book) is mostly exported and used instantaneously, though storage in batteries or in some alternative potential like pumped hydropower is constantly growing in interest. By contrast, fossil fuels and nuclear material resources are not evenly distributed geographically around the world, and so access to the raw resource is either fortuitously dependent on the natural availability, or price dependent through import and resource processing. This can isolate lesser economically developed countries that struggle to compete financially (Cherp *et al.*, 2013). This trade structure and captive dependency on fossil fuels can result in political instability (Lefèvre, 2010). Unsurprisingly, this then leads to energy security being a crucial target for energy policy makers around the world (Winzer, 2012). As the raw solar resource is not held at a price—it does not require any extraction, processing, or transportation—it is conceivable that solar energy generation, combined with other renewable energy generation technologies, could largely remove this political and economic barrier of resource availability. This is, of course, dependent on fair access to solar energy technologies. Alas, this is a common barrier to all energy generation techniques; nevertheless, we see a marked improvement on the status quo.

Little needs to be said about the meteoric rise in solar energy technologies worldwide, other than it is unsurprising and certainly credible. However, the technology is not without its downsides. In order for it to remain cheap to developers and private installers, the electricity generated must be instantaneously consumed. This means that the electricity is either sold to the grid as soon as it is generated or consumed locally (for example in the home). However, the generated electricity is not always immediately required; in contrast, it is not always generated when it is in fact needed. So without the capacity to store the generated electricity, which can be sold at a more convenient time, solar energy technologies are at the whims of the weather and whether or not the sun is shining. Of course, we could store the electricity in batteries, heat or as some alternative potential, and often these approaches mitigate a lot of the variability, but forecasting and understanding fluctuations in available solar irradiance still have a role to play in these storage technologies.

The average annual solar resource is reasonably predictable around the world. We intuitively understand this in the most basic sense. For example, we intuitively know that we would have a better chance of a sunny holiday next year were we to go to the South of France than anywhere in England. What we have really just done is intrinsically underpin the solar potential and typical meteorology of lower latitudes in a temperate climate. We understand the seasons and have done so for time immemorial; this is perhaps a sweeping statement, but we observe this understanding in plants and animals and so it is safe to say that we instinctively understand them too. Consider an example. Were we to ask a child from each region of the world to describe the weather throughout the year where they live, we could build a global model showing the interannual variability of weather and infer from this a solar energy potential map. This is because humans instinctively understand solar energy, even if not consciously.

It will be shown throughout this chapter how simple it is to model these relationships of solar potential with regard to the distance to the sun, the tilt of the earth, and the geometry to the point on the globe. These periodic changes are known as seasonality. It is the benchmark to which we can begin to understand how much potential solar energy is available at a location. However, once we take the estimation further—and by this, I mean to high resolutions in both space and time—accounting for the atmosphere and its constituents and the dynamic nature of the clouds, the whole process becomes increasingly complex and harder to underpin.

For many decades now, solar energy has been modeled mathematically. Long-term seasonality estimates of solar energy are satisfactorily accurate. However, the chaos of clouds and the acquisition of data are still to this day an incredibly challenging hurdle for solar applications. To describe the nature of solar energy research today, it can broadly be categorized into the following basic themes: solar technologies, solar materials, solar system design, solar monitoring, solar physics, and solar energy meteorology. For the purposes of this book, we are mainly interested in energy meteorology and, to some extent, monitoring. Today, it is possible to have highly accurate solar monitoring equipment as well as reasonable estimates of resource history from satellite or other gridded methodology (Yang and Bright, 2020). However, one of the main challenges is to have a long history of high-resolution data in both space and time for the location that you desire. The chances of having excellent data retrieval at the location of economic interest is slim to none. This desire for data is where the field of synthetic solar irradiance was born.

1.3 SOLAR GEOMETRY

This book deals with the topic of "modeling the sun," however, it is not intended to be a textbook such that you walk away with all the mathematical equations needed. The book is instead supposed to inform you of the state of synthetic solar irradiance and how one might engage with this research field. However, we must be introduced to the definitions of the more important solar geometries, irradiance definitions, and indices. Hence, we introduce these core concepts here. Note that this section is far from comprehensive. In fact, there are many more suitable books and websites that teach a complete understanding of these more common features of solar irradiance modeling. For example, Duffie and Beckman (2013) wrote a very detailed book covering these facets, and Brownson (2014) dedicated an entire chapter to deriving just the solar geometry. We simply wish to cover the critical components that feature throughout this book.

The definitions provided are purely derivations that have had considerable usage over the years. To obtain high accuracy solar angles, it is important to leverage the best available techniques. Fortunately, you would not need to—and should not have to—code these expressions from scratch, except if you desire for your own curiosity. There are prominent coding packages across various scripting languages that have already done this hard work for you, most notably the `pvlib` package (https://pvlib-python.readthedocs.io/) that is in both Python and MATLAB. The solar geometry calculations most commonly used today are called pvlib-python/spa.py, which is the National Renewable Energy Laboratory's (NREL, USA) solar position algorithm (SPA). At the time of writing, there are 42 functions leveraged in the NREL SPA needed to calculate all the solar angles. Hence, this introductory chapter will only give a shallow overview with simple equations.

Solar modeling has the objective of deriving the amount of irradiance received by an observer on the earth's surface. The term *observer* is used here, but you can instead imagine this as a solar farm in a field. To calculate this quantity, we must know the path between the sun and the earth, and to what angle sunlight is approaching the observer. These angles are time variant, as the angles all depend on the continuous orbit and rotation of earth.

One must assume that, to calculate solar geometry, you have a location in mind that is our reference point on the earth's surface. This reference point is defined using the latitude Θ and longitude Φ, expressed in degrees in what follows.

The first angle for us to consider is called the solar declination. Solar declination is the angle made between (i) a line drawn between the center of the earth to the center of the sun, and (ii) a line drawn from the center of the earth that passes through the equator. Because the earth rotates on a tilt, the equator does not always point directly at the sun (though it does twice a year at times known as the equinox). The solar declination angle is independent of our position on earth, as it is explained only by the tilt of the earth and the time of year. The declination explains the length of a day at the location being modeled. The declination is a function of n, which is the day position in the year, e.g., February 15 would be the 46th day of the year, hence in this case $n = 46$. The declination angle δ is calculated as

$$\delta = 23.45^\circ \sin\left(\frac{360}{365}(284 + n)\right). \tag{1.1}$$

The next quantity to understand is known as the "equation of time." The earth wobbles on its axis. Were we were to take a photograph of the sun from an identical vantage at the exact same time of every day (by a wrist watch, not by solar time), we would observe the sun increasing in elevation in the sky and decreasing again in a line that is exactly explained by the declination. However, we would also observe sideways variation in position such that we actually see more of a figure-of-eight; this is known as an analemma and is illustrated in Fig. 1.2. The analemma is due to the fact that the earth is not on a perfectly circular orbit, but a more elliptical orbit. As the earth moves closer

FIG. 1.2
Afternoon analemma photo taken in 1998–99 in Murray Hill, New Jersey, USA, by Jack Fishburn. The Bell Laboratories building is in the foreground.

to the sun, the speed that the earth orbits increases. As the earth travels faster, the sun would have advanced along its east-to-west location sooner than the previous day, yet our clocks are perfectly spaced throughout the year; hence, we would see displacement. The opposite effect happens as the earth slows down at the farthest point away from the sun. For solar modeling, we wish to correct for this variation by adjusting the time of day with which we take the image to force the sun to appear on the same line each day. Fortunately, the wobble can be described mathematically to a high level of accuracy (or at least high enough for solar modeling). The number of minutes to be added or removed is calculated and is known as the equation of time (EoT). EoT can be calculated with the following (Duffie and Beckman, 2013) (there are, of course, alternatives):

$$B = (n-1)\frac{360}{365}\frac{180}{\pi}, \tag{1.2}$$

where n again is the day of year. This is then used as so:

$$\begin{aligned}\mathrm{EoT} = {} & 229.2 \cdot (0.000075) \\ & + 229.2 \cdot (0.001868 \cos B - 0.032077 \sin B) \\ & - 229.2 \cdot (0.014615 \cos 2B + 0.04089 \sin 2B).\end{aligned} \tag{1.3}$$

Both the solar declination and the equation of time are displayed in Fig. 1.3.

With EoT, we can now arrive at an expression of the hour angle ω. The hour angle is a representation of time in terms of hourly deviation from solar noon, expressed in degrees. First, then, we must arrive at a representation of local solar noon. Before that, a quick message on time: it is far simpler to avoid complexities of daylight savings and other time-zone-related issues by operating instead in

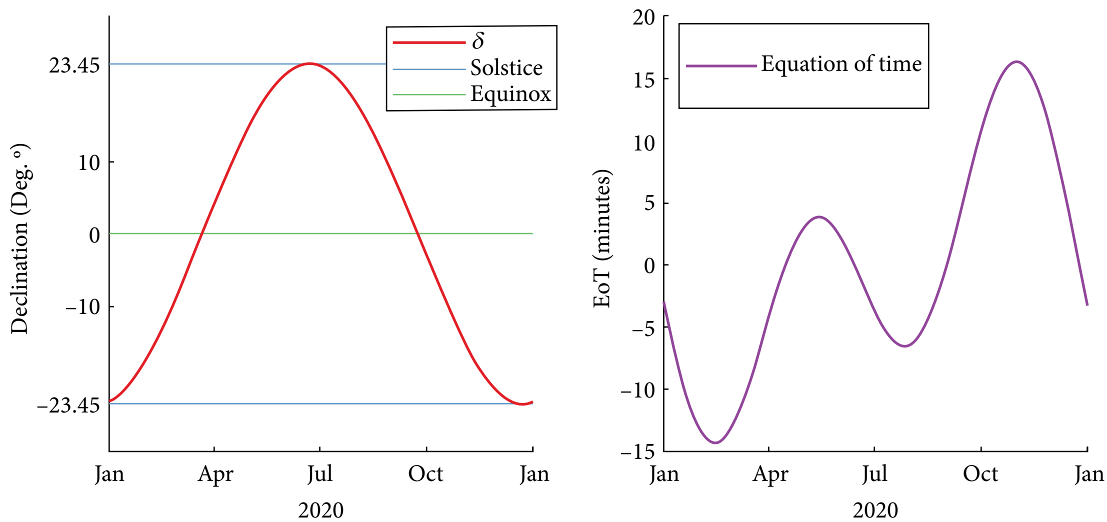

FIG. 1.3
(Left) Solar declination throughout the year with equinox and solstices indicated. (Right) Equation of time throughout the year.

coordinated universal time (UTC), which is the time at the prime meridian ($\Phi = 0°$). A useful quantity is to represent time in decimal time t_d, whereby time 11:30 A.M. is 11.5, and 5:10 P.M. is 17.167. This can be simply calculated as

$$t_d = H + \frac{M}{60} + \frac{S}{3600}, \tag{1.4}$$

where H, M, and S are the hours, minutes and seconds expressed in 24-hour clock hourly notation and in UTC time zone. For example, 1:30 pm in Singapore time (+8 UTC) would become $H = 13$, $M = 30$ and $S = 0$.

To calculate local solar time t_n, we leverage the longitude and the fact that the 360° of the globe can be divided into 24 hours in increments of 15°:

$$t_n = 12 - \frac{\Phi}{15} - \frac{\mathrm{EoT}}{60}. \tag{1.5}$$

With t_n, we can now derive ω:

$$\omega = 15(t_d - t_n), \tag{1.6}$$

where ω ranges between $-180°$ and $180°$ (mod 2π).

Finally, we may now derive the most important angle in solar modeling, the solar zenith angle (SZA, θ_z). SZA is simply the angle from the *zenith* to the sun—zenith is an imaginary line directly above a point on a sphere (sometimes helpfully thought of as the opposite direction of gravity). If stood on a flat surface with one hand pointing vertically above you (zenith) and the other pointing at the sun, you would be representing SZA as the angle between your arms. So when the sun is exactly at the horizon (sunrise/sunset), SZA is exactly 90°. At locations about the equator (depending on δ), the sun achieves a SZA of 0° around noon, though higher latitudes do not experience this and the sun always stays lower in the sky. It is also common to see this same angle expressed instead as the solar elevation angle, which is the complement of SZA in respect to the horizon, e.g., if the sun is 15° above the horizon it would have an elevation angle of 15° and a SZA of $90° - 15° = 75°$. To derive SZA, we require all the previously defined angles and positional angles:

$$\cos(\theta_z) = \sin(\Theta)\sin(\delta) + \cos(\Theta)\cos(\delta)\cos(\omega). \tag{1.7}$$

A representation of SZA and the hour angle are provided in Fig. 1.4. SZA is used with incredible frequency within solar modeling and synthetic solar irradiance models. This is because it enables us to group data to be representative of the position of the sun in the sky, which effectively removes any geographic location locking. For example, $\theta = 70°$ at a latitude of 60° or 0° may appear at different solar times; however, the angle of sunlight through the atmosphere is the same, and so we can apply generalized calculations to both. Using only the zenith angle, it is possible to arrive at an estimate of

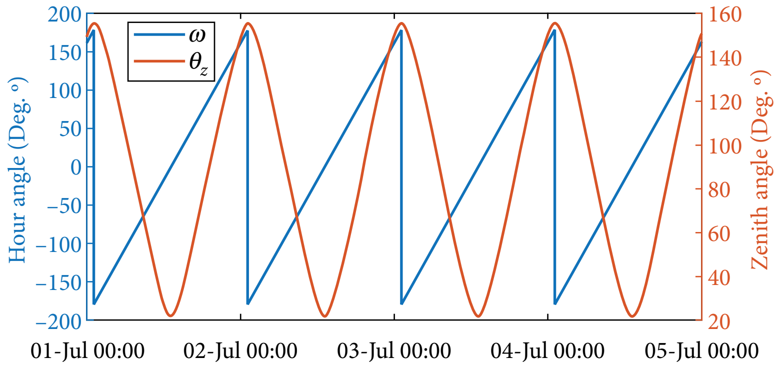

FIG. 1.4
Four-day progression of the hour angle and solar zenith angle. The hour angle jumps from 180° to −180° at solar midnight when SZA is at its peak. SZA is 90° at sunrise and sunset.

the potential available irradiance that would be received at the earth surface in the absence of clouds. It is now time to discuss irradiance.

1.4 TERMS OF IRRADIANCE

By this point in the chapter, I have already mentioned the word *irradiance* a few times, and it appears many more times moving forward; it is also in the title. This cannot be a comprehensive introduction to the fundamentals of solar irradiance, as that would require an entire book—a good example would be Iqbal (2012). Instead, the most important terms of irradiance are introduced so that researchers new to this field can have a rudimentary starting point.

The sun emits electromagnetic radiation. It behaves approximately as a black body[2] with an effective temperature of 5777 K, in that the total energy emitted from the sun and a black body of 5777 K is equal, though it does differ slightly at different wavelengths. The amount of energy within the electromagnetic radiation is called the radiant energy (in joules). When radiant energy is measured over a unit of time, we arrive at the radiant flux or radiant power (in watts = joule/second). In solar engineering, we are most interested in the density of radiant flux upon a surface—a solar panel for instance. This is called the radiant flux density (in watts per square meter) and is also known more simply as the irradiance. Hence, irradiance is the energy from the sun arriving on a surface every second, or more simply, the power from the sun incident upon a surface. As irradiance is emitted across a spectrum of wavelengths, each with a different density, irradiance can be reported per wavelength as spectral irradiance (Gueymard, 2004). A demonstration of spectral irradiance in space and at the surface of earth are visualized in Fig. 1.5; additionally, the spectral irradiance emitted by a black body of thermal equilibrium at 5777 K following Plank's law is plotted for reference.

Irradiance is typically synonymous with the broadband irradiance covering wavelengths from 0.2 to 4 micrometers (μm). The device we use to measure irradiance at the surface is called a pyranometer,

[2] A black body is an idealized physical surface that absorbs radiation from all wavelengths regardless of angle of incidence. A black body at equilibrium also emits thermal electromagnetic radiation.

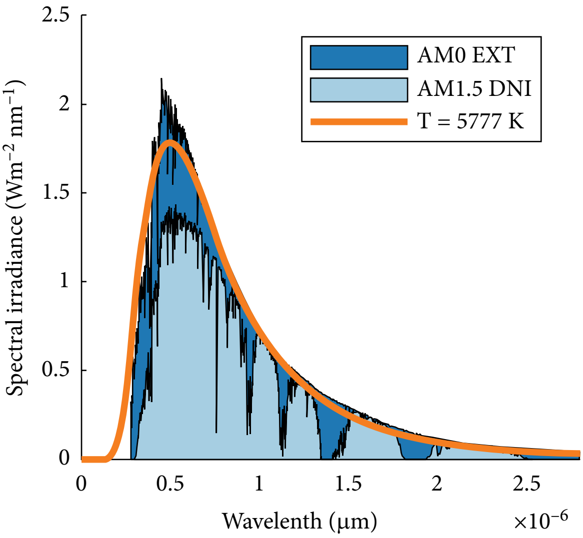

FIG. 1.5
Spectral extraterrestrial normal irradiance at an air mass (AM) of 0 in darker blue, and the spectral direct normal irradiance at AM1.5 in pale blue. The orange line plot is the spectral irradiance emitted by a black body at thermal equilibrium at 5777 K following Plank's law. [The data for this figure were developed by the American Society for Testing and Materials (AST) G173-03 standards and provided courtesy of The National Renewable Energy Laboratory (NREL) operating for the U.S. Department of Energy by the Alliance for Sustainable Energy, LLC ("Alliance")].

which is a sensor that converts the received solar radiation to a measurable electrical signal. Typically, pyranometers operate in the 0.285 to 2.8 μm spectral range. From Fig. 1.5, we observe that this encompasses the majority of significant spectral irradiance contributions.

The total solar irradiance (TSI) is a measurement of the sum of all solar power from every wavelength per square meter. TSI is measured normal/perpendicular to the incoming solar radiation (i.e., points directly toward the sun). Taking a long-term mean of TSI at the average distance between the sun and the earth (one astronomical unit, AU), we arrive at the solar constant (SC, E_{sc}). Gueymard (2018) provides a very detailed exploration of the history and an update of SC. Whilst called a constant, it does in fact fluctuate depending on solar activity. SC is the general starting point for most solar resource assessment applications as it is used to assess the potential solar energy. SC was derived from TSI most recently (Gueymard, 2018) to be a value of 1361.1 Wm^{-2}.

With the SC, we can estimate the solar irradiance at the top of the atmosphere (in space just before it interacts with anything). This irradiance is called the extraterrestrial irradiance (EXT) and is very predictable, as there has been no interference to the radiation traveling through the vacuum of space; it is the irradiance we would receive if there were no atmosphere. EXT is a function of SC and the ratio between the mean sun-to-earth distance (to which SC is derived) and the apparent sun-to-earth distance. The apparent sun-to-earth distance depends on what point in the earth's orbital cycle we are calculating; hence, time and SC are the governing inputs to arrive at EXT.

There are three types of planar projection that are important to introduce at this point: horizontal, normal, and tilted. As irradiance is measured per unit area (i.e., on a surface or plane), it is important to acknowledge that the orientation of the surface has a significant influence on the irradiance received. The two most common projections are normal and horizontal. Normal refers to a plane that is perpendicular to the sunlight, e.g., pointing directly at the sun. Horizontal refers to a plane that is perpendicular to gravity. For example, if the earth was perfectly smooth, then horizontal irradiance would be the power incident upon anywhere on the surface. Occasionally, the horizontal irradiance is the same as the normal when the horizontal surface happens to be pointing directly at the sun; this happens around the equator for short periods of time when the sun is perfectly above (at zenith). The final projection is called tilted, meaning a plane of arbitrary orientation—it could be facing in any possible direction. The tilted projection is used when understanding the irradiance incident upon a solar panel itself. For notation, we tend to use E to refer to irradiance and the subscripts inform which type of irradiance and the projection. EXT as the starting place for solar modeling is represented as E_0, but because it is represented on a normal projection, it is additionally denoted with a subscript n; EXT is E_{0n}. EXT on the horizontal would be E_{0h}. The notation varies among authors.

Converting between irradiance on the normal projection to the horizontal is straightforward and only requires SZA to be known. By multiplying EXT by the cosine of SZA, we can convert any normal projection of irradiance to horizontal. Hence, $E_{0h} = E_{0n} \cos \theta_z$. EXT and SC are displayed in Fig. 1.6. On the left panel, the solar constant is shown on the blue horizontal line; as suggested, it is constant throughout the whole year, whereas the received normal EXT fluctuates throughout the year relative to the sun–earth distance. The right panel illustrates horizontal EXT in green. This

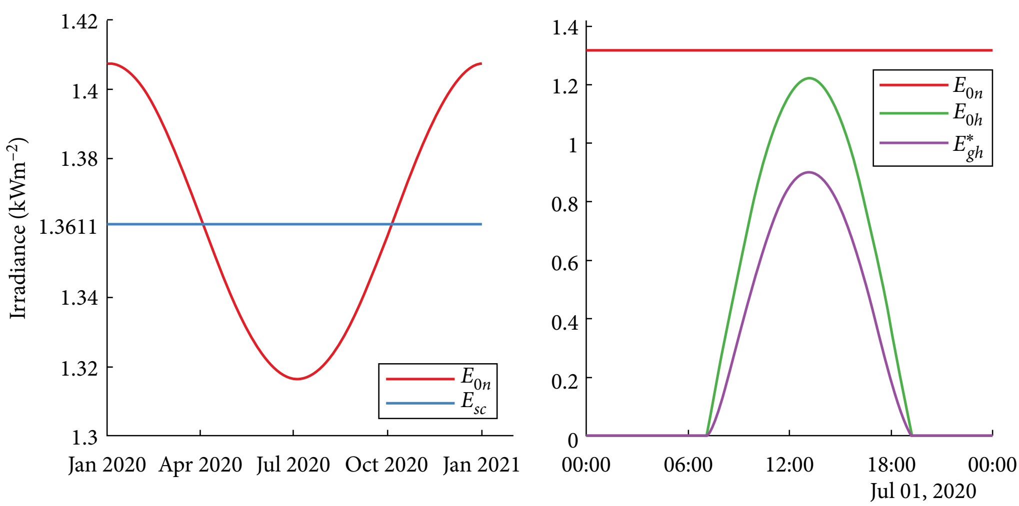

FIG. 1.6
(Left) Extraterrestrial normal irradiance (EXT) and the solar constant (SC) throughout the year. (Right) EXT on the normal and horizontal and an example of global horizontal clear-sky irradiance (GHIcs) for the location of the SERIS, NUS, Singapore, in Singapore time.

significantly depends on latitude and time. In the example, the horizontal plots are for the location of SERIS, NUS in Singapore (1.30°, 103.77°).

From EXT, we can make an approximation of the irradiance at the earth's surface in the absence of clouds. This is known as the clear-sky irradiance. From Fig. 1.5, we saw that spectral irradiance at the surface was diminished when compared to EXT; this is strictly due to interaction of radiation with the atmosphere. And so to arrive at an estimate of the earth's surface, we must now account for atmosphere. At this stage it is important to introduce three new terms: global, direct, and diffuse irradiance.

Irradiance is composed of direct and diffuse components. Global irradiance is the sum of both irradiance components incident upon a plane from any direction. As sunlight is scattered and reflected from the atmosphere and the ground, the term *global* is a catchall term for all aspects of irradiance. Global irradiance is most commonly reported on the horizontal projection and as such is referred to as the global horizontal irradiance (GHI) and is commonly denoted as either E_{gh} or G_h. GHI can also be expressed in terms of clear-sky irradiance—the irradiance received in the complete absence of clouds, GHIcs. It is commonly denoted as E^*_{gh} or $G_{h,cs}$. A representation of GHIcs is on the right panel of Fig. 1.6 in purple. The difference between GHIcs and EXT is, therefore, only due to the atmospheric interaction with radiation.

The direct normal irradiance (DNI, E_{bn}; under clear-sky is DNIcs, E^*_{bn}) is the irradiance received within a 5° radius of the sun, ignoring all other directions (e.g., irradiance from reflections and scattering). DNI is regularly known as beam irradiance, hence the subscript *b*. It is also regularly known as beam normal irradiance BNI or B_n. DNI is the high energy component that is essential for concentrating solar plants and solar thermal technology.

The diffuse horizontal irradiance (DIF, E_{dh}; under clear-sky is DIFcs, E^*_{dh}) is all other radiation received at the surface excluding DNI. DIF is also regularly expressed as D_h or DHI. DIF is composed of radiation that is scattered or reflected by the atmosphere. Radiation is absorbed by the molecules in the atmosphere. These molecules convert some of that energy to heat, kinetic, or activation energy before re-emitting again in all directions. Hence, we receive solar radiation in the form of DIF from all directions (except the 5° around the sun which is considered DNI). DIF is the reason that we can see when it is fully overcast. DNI is negligible when a cloud is blocking the sun, and so any energy received by a pyranometer during these periods is entirely DIF. There are additional complexities regarding ground-reflected irradiance, etc., however, they are not discussed further here as these complexities are largely redundant for the current state of synthetic solar irradiance.

Examples of GHI, DNI, and DIF are presented in Fig. 1.7. The first day of the example is so clear that the measured quantities of GHI, DNI, and DIF are equivalent to their clear-sky counterparts GHIcs, DNIcs, and DIFcs, respectively. The second day illustrates how the DNI drops with cloud cover, which is mirrored by a rise in DIF due to the brightening effect of clouds. Notice how GHI does not fall below DIF, and that DNI can drop to zero.

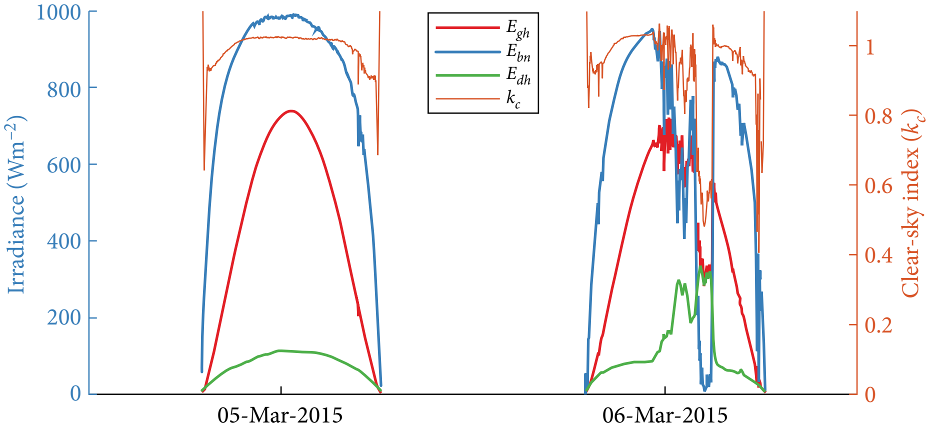

FIG. 1.7
Example of global, direct, and diffuse irradiance and the clear-sky index for two days at the Baseline Surface Radiation Network station of Sioux Falls, South Dakota, USA. The day of 05-MAR-2015 was so clear that the measured irradiance components are equivalent to their clear-sky counterparts.

A particularly interesting thing to know about DNI, DIF, and GHI is that, were we to measure these quantities for, say, ten years consecutively, the annual total received power would most likely differ each time. This might not be news to you; we all know that we can have good or bad summers that would strongly influence the total, but it is important to highlight, as it plays a relevant role for synthetic solar irradiance. The same is also said about irradiance over any averaged time scale; monthly and daily totals vary considerably too—each day is unique. Perhaps you have heard of catastrophic events described as a 1-in-1000-year event? It is regularly used to describe volcanic eruptions, tsunamis, floods, earthquakes, hurricanes, wild fires, etc. It means that a particular event was so extraordinary in magnitude that the probability of it occurring again within the next 1000 years is slim indeed. It is probability, so of course it could happen again in the next year, though that would be incredibly unlucky due to the low probability of it actually happening. This phrase can also be applied less excitingly to solar irradiance. Each year is different due to the complexity of weather systems. For example, we could have had a particularly cloudy summer, as such the total annual irradiance would be much lower.

The most important take home message here is that, were we to measure 1 year of solar irradiance data at a location, how can we know whether or not it was a 1-in-1000 year event that we just recorded? There is a distribution of potential annual totals from the worst case to the best case with all possibilities in-between. Our year of measured data must fall within that distribution somewhere. If we design a solar farm based on this one year of measurement and we assumed that all years of operations would be the same, then if our measured year was a very high percentile estimate, we would dramatically overestimate the yield of the solar farm. Likewise, if it was a very low percentile, we would receive much more than expected and may have sized the equipment sub-optimally. Hence, a longer measurement history is needed to start to build that distribution. With this distribution, it would then be suitable to make an assessment of the solar potential of a location.

Clear-sky irradiance (GHIcs, DNIcs, and DIFcs) is a heavily researched field; there are numerous methodologies to produce an estimate that accounts for a cloud-free atmosphere. In fact, the most recent validation on the topic evaluated 95 different clear-sky irradiance approaches for GHIcs, DNIcs, and DIFcs (Sun *et al.*, 2019; and Sun *et al.*, 2021). We have briefly mentioned that the clear-sky irradiance is what would be received under perfectly cloud-free conditions. However, this hides certain nuances that are also noteworthy. Clouds are not the only influential factor when it comes to solar irradiance losses from EXT down to the clear-sky irradiance at the surface. Let us make an extreme example by considering a significant dust storm event. A dust storm is not technically cloud, and so a dust storm that happens on a clear day is technically a clear-sky day, yet there would be a considerable reduction in power from EXT to GHIcs. Hence, the constitution of the atmosphere between the observer (point on the earth's surface) and the sun is critical for accurate clear-sky irradiance modeling. Of the 95 clear-sky irradiance models evaluated by Sun *et al.* (2021), the list of variables that are recorded with enough regularity to be used in model design are barometric pressure, ambient temperature, total column water vapor, total column ozone, total column nitrogen dioxide, and information about the airborne aerosols (represented as an optical depth at various wavelengths, as well as scattering parameters). The clear-sky models vary in complexity from including none or all of the above. Naturally, those that consider more variables in their physical modeling parameterization tend to do better in global validation. The most widely used clear-sky model is either the REST2 (Gueymard, 2008; and Bright *et al.*, 2020), though often advocated for is the McClear (Lefèvre *et al.*, 2013) model, as it is easily accessible online.

I like to consider the clear-sky irradiance as the "potential" irradiance that could be received on a plane at the earth's surface. It represents the best-case scenario for a solar farm, as it is the maximum that could be received that day resulting in uninterrupted generation. However, it does not necessarily mean it is the possible upper limit of what is received. In fact, the clear-sky irradiance is regularly exceeded. Sometimes clear-sky models are simply just underestimating. However, there are other times when the clear-sky irradiance is exceeded, and it may surprise you to learn that it is when there are clouds in the sky. Whilst I did say that clouds are responsible for the majority of irradiance reduction, they can also *contribute* to the received irradiance in instances when the sunlight reflects off the side of a cloud momentarily enhancing the amount of irradiance received. This is called an overirradiance event or a cloud edge enhancement event. Clear-sky irradiance is still one of the most useful variables to calculate when working with solar data. This becomes more apparent in the following section where it is used to derive features and indices of irradiance.

1.5 COMMON INDICES OF IRRADIANCE

Now that we have discussed many of the different terms of measured irradiance, as well as derived the solar zenith angle SZA and introduced the clear-sky irradiance, we can begin to talk about specific features of irradiance. Note that this is far from a comprehensive list (in fact I don't believe a complete list exists yet); these are simply the ones most commonly seen in literature.

A feature, in data science, can be considered as a quantifiable property of a phenomenon under observation. For our purpose with irradiance, a feature broadly can be translated to anything that can be derived or deduced from the measured time series. We set limits as to what features are permissible to only those that can be derived from the basic variables that we would have available to us when modeling: location, time, and any commonly measured irradiance data. We learned from the previous chapter that with only location and time, we can derive EXT and SZA. We can also use some of the more simpler clear-sky irradiance estimation tools to obtain GHIcs, DNIcs, and DIFcs. This is a remarkable amount of useful features from such a simple collection of input data.

Once we incorporate the measured quantity of irradiance (usually only GHI) we are then able to derive additional features. Extraterrestrial and clear-sky irradiances represent versions of the solar potential at a time and location, so we can use them to detrend a time series of measured irradiance. The clearness index (denoted as k_T here, but widely varies in literature) is an expression of the measured irradiance normalized by the extraterrestrial irradiance:

$$k_T = \frac{E_{gh}}{E_{0h}}. \tag{1.8}$$

As GHI is normalized by EXT, we express the received irradiance in terms of the maximum potential ignoring all influences presented by atmosphere and weather. That is to say that a solar panel in space (at the edge of the earth's atmosphere) on the horizontal plane would have a constant k_T of 1. This means that the total attenuation of sunlight from the edge of space to the ground is entirely encapsulated by the clearness index. The clearness index is particularly useful, as it inherently captures *all* possible reductions in irradiance from the edge of space to the ground without errors propagating from the clear-sky irradiance model. It is also popular to use, on occasion, because it is incredibly simple to calculate.

The other popular index of irradiance is the clear-sky index (denoted k_c here, but again varies massively in the literature; κ is also very common). Much like the clearness index, the clear-sky index is the measured irradiance normalized by the clear-sky irradiance:

$$k_c = \frac{E_{gh}}{E^*_{gh}}. \tag{1.9}$$

From the equation, we see that the clear-sky index is calculated using GHI and GHIcs. One can also derive the clear-sky beam or diffuse index, normalizing by their respective clear-sky component. The clear-sky index is an incredibly powerful tool in solar energy engineering. The main benefit is that, assuming an accurate clear-sky model, the losses due to atmosphere are already accounted for as they are encompassed within the calculation of GHIcs already. This means that the losses described by the clear-sky index can be exclusively isolated to the influence of clouds. Of course, any other systematic loss is captured within k_c, such as shading or fouling of equipment. For now, we can assume that these issues are extraneous. The clear-sky index, therefore, explains the optical losses associated to clouds.

The other major perk to expressing irradiance in terms of the k_T or k_c is that they are the closest approximation of stationarity that we can have with solar irradiance time series. A stationary time series has the same mean, variance, and autocorrelation over time. As the presence of clouds is not stationary (except over coarse temporal averages) and the indices represent losses due to clouds, a time series of k_c or k_T is not perfectly stationary. That said, we effectively remove all influence of seasons and the daily curve of irradiance (from sunrise and sunset) from the time series.

Example times series of k_c and k_T are shown in Fig. 1.8, which uses identical data to Fig. 1.7. It is immediately clear that k_c is higher than k_T, naturally so because the former is normalized by GHIcs and the latter by EXT. For the extremely clear-sky day on 05-Mar-2015, we observe that k_c achieves a more consistent level of stationarity than k_T. Again, this is because the atmosphere is almost entirely accounted for by E^*_{gh}, whereas the length of the path that the sunlight makes through the atmosphere is constantly changing with SZA. As such, it hits a minimum rate of change as a function of SZA. An interesting

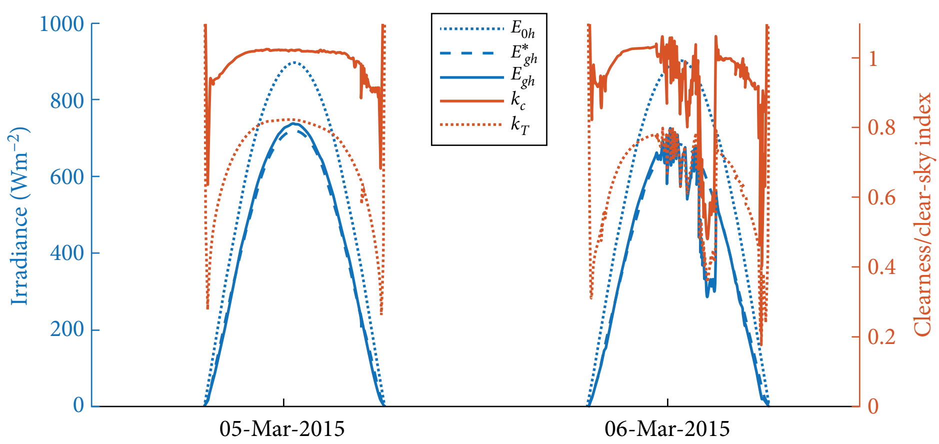

FIG. 1.8
Example of clearness and clear-sky indices.

observation, one that is a regular feature of solar modeling, is how the indices behave at sunrise and sunset. As the values of irradiance at these periods are in the order of 0–10 Wm^{-2}, the indices become very sensitive to discrepancies between the clear-sky estimate and measurement, furthermore, irradiance measurement equipment suffers considerable accuracy issues with such low solar angles. It is common in solar energy engineering to ignore the first 5° of solar elevation (or $\theta_z > 85°$).

Other useful features that regularly appear in solar modeling are the diffuse fraction k_d and the beam fraction k_b. The diffuse fraction indicates the share of GHI that is attributed by DIF:

$$k_d = \frac{E_{dh}}{E_{gh}}. \tag{1.10}$$

The diffuse fraction is equal to 1 when there is no DNI present, meaning that GHI is entirely equal to DIF. Under clear-sky conditions, the DIF can be much higher than usual if there is more scattering in the atmosphere, such as that produced by particulate scattering (e.g., a larger portion of the direct sunlight is scattered by the molecules present in the atmosphere). The beam index is calculated as

$$k_b = \frac{E_{bn} \cos \theta_z}{E_{gh}}. \tag{1.11}$$

Note that the DNI must be converted onto a horizontal projection first. The beam fraction can never be 1, as this would imply no scattering.

There are many other features that can be extracted from an irradiance time series, especially when one starts to look at features evaluated over a period of time, such as the size of a ramping event, or, for example, evaluating a feature over a 10-min window. There are ways to quantify variability (discussed in depth in Chap. 4); there are even attempts to classify the type of clouds from just the observation data. There has yet to be a comprehensive study of all the features possible to derive from time and location, and any additional ones once the ground data are incorporated. This would actually be a very welcomed piece of work, mainly because, for the most part, solar modeling is achieved largely by leveraging these features.

As an example, the clear-sky index k_c was shown to have distinct distributions when grouped by other meteorological variables, such as the total cloud cover amount and SZA (Smith *et al.*, 2017). The total cloud cover amount is an expression of what percentage of sky is covered by clouds (where 0% is clear and 100% is overcast). We can intuitively theorize how these distributions may look. We know that when there are no clouds, the ground GHI should be equivalent to the GHIcs (assuming we have an accurate clear-sky model). Hence, we would have a large peak at $k_c = 1$. We also know that clouds cause a reduction in k_c, and that clouds generally look similar depending on the type of cloud, with a few rare exceptions. We also discussed breaches of $k_c > 1$ under cloud enhancement events. So we can reason that each type of cloud will have a particular distribution and there will be a long tail beyond 1. Combining the distributions for clear and cloudy periods into a single distribution, we arrive at a bimodal distribution. However, this has previously been investigated in more detail and shown to have a trimodal or even more complex than that with each 1/10 fraction of cloud cover having a distinct distribution that even changes with SZA.

Does this then mean that knowing the cloud cover amount and SZA that we can then arrive at an estimate of the solar irradiance? Sort of. Knowing the percentage of cloud in the sky could, in theory, let us map the clouds and see what would be received at the ground. However, we do not have an understanding of where the clouds are, how large they are, how thick they are, which direction they are moving, etc. So we could not generate the actual ground irradiance (not unless we had something like a skyward-facing camera that could give us this information). We could, however, create a time series that would be entirely realistic under that scenario. These realistic data are an example of what is known as synthetic solar irradiance.

1.6 WHAT IS SYNTHETIC SOLAR IRRADIANCE AND WHY MIGHT WE USE IT?

As we now understand what all the terms of irradiance are and how features of irradiance can be deduced, we have arrived at the main topic of this book: synthetic solar irradiance. But what is it?

Synthetic solar irradiance is realistic but fake irradiance data. We favor the term *synthetic* as it captures many of the facets of real solar irradiance, but is instead computer generated. The *Oxford English Diction*ary defines the adjective *synthetic* as "(of a substance) made by chemical synthesis, especially to imitate a natural product." We abstract the chemical synthesis to mean instead computer-generated synthesis. In our case, we are using computational modeling to imitate GHI.

Perhaps the simplest method of describing what exactly *is* synthetic irradiance is to provide an example of when it might be needed. Pretend that we are solar farm developers. We intend to commission a large-scale PV farm in the north of England. Let us simplify this by assuming all political inertia, demand requirements, equipment suppliers, and grid-based connection issues are identical regardless of location. An incredibly important variable for decision-making is, therefore, where we position the potential solar farm. The answer is wherever it is most profitable. So where is most profitable? PV farm performance between two prospective sites (assuming identical layout, demand, and maintenance) is a function of the solar resource availability. We could extend this to know about possible limitations set by the electricity network operators that could potentially be violated due to solar variability. Hence, we must discover (1) how much PV generation is there likely to be and (2) what are the largest possible ramping events we may experience. To answer these questions, we require solar data from our prospective locations. We discussed in Sec. 1.4 that we would need an extended history of ground data in order to make an informed decision about plausible yield. Hence, we only have a few options.

The first and most concrete option is to have scoped the site many years before this stage and had a long history of observation data. This isn't particularly ideal if we did not have the foresight or the need to move quickly. It also isn't ideal because mapping and measuring data at all possible locations is costly and time-consuming. Hence, we would next consider a solar resource assessment, most likely from a private consultation contract. It is quite likely you would obtain some gridded reanalysis estimates or satellite-derived estimates of irradiance (Yang and Bright, 2020). Now, this is an excellent option, though the resolution over Europe and Africa is at best 15 min, over Asia at 10 min, and only recently at 5 min over the continental USA. Also, the estimation accuracy is often questionable and could lead to biases. It would also not resolve issues of variability in space and time, as the spatial resolution of these devices is coarse, such that a solar farm would be lost within a single pixel of a satellite image. An alternative option, then, is to leverage synthetic solar irradiance.

Synthetic solar irradiance is one or many realistic time series that are statistically representative of real measured data. They are computer generated using a collection of data-driven mathematical modeling techniques. The intention is that synthetic solar irradiance data behaves in the exact same manner as the real data, but the moment to moment magnitudes will not be the same except by chance. All the facets and features of solar irradiance (as discussed in the previous sections) should be present, including the temporal and spatial correlation. The major advantage of using synthetic solar irradiance data is that many years of data can be generated for the same location with no necessary requisite for preexisting data (though it is always useful to have). Using solar irradiance features and other variables that are more commonly measured (such as the historical cloud cover amount), synthetic solar irradiance models can leverage the statistical relationships observed at the locality to generate plausible time series. An additional powerful aspect of synthetic solar irradiance is that the time series can be generated on a grid such that each PV panel in the design of a solar farm can have its own bespoke time series that correlates to the others in space. Due to the nature of any training or relationships to variables that synthetic solar irradiance models rely upon, the output should automatically be appropriate to the geography modeled. Another possibility of synthetic solar irradiance would be to interpolate coarser time series into higher resolution time series in a process known as downscaling, such as taking the satellite-derived irradiance at 15 min resolution and returning 1 min resolution data. The downscaled versions should retain all the properties of the coarser data, but introduce realistic variability as estimated from the features within the data (such as leveraging the diffuse fraction as an indication of cloud presence; Frimane *et al.*, 2020).

As part of our thought exercise, we can immediately see how synthetic solar irradiance can complement all the other datasets that are conventionally obtained in a solar farm siting investigation. Say that we are now armed with a year of ground data, a good few years of historical gridded data, and synthetic solar irradiance time series that have been generated for as many years needed to create a good distribution of annual yield. With the many years of synthetic data, we can get an idea for the plausible distribution of yields that the sites receive. When we analyze the ground measured sample, we see that it falls within the higher end of the annual yield distribution. Furthermore, we see that for that year, the more westerly site had a marginally higher annual yield than the easterly site. If we were to have only leveraged the measured time series and designed all financial planning on this higher than usual sample, we would have overestimated our expected revenue for the majority of years. Looking at the solar resource estimates from gridded data, we compare it to the ground data and find that it overestimated by 2% at one site and underestimated by 2% at the other; we are uncertain whether this is a consistent bias such that we can modify the whole history by these values. Validations at other independent sites show a general scattering of biases that largely are balanced at 0% globally. However, are the estimates for our sites accurate?

Once we consider all the data, we might learn that the more easterly site has higher annual yield distribution over the many years than the westerly site. This is in contrast to what the ground measurements showed. However, we can now see that the measurements were within the distribution but at different percentiles for each site. We also see that the spatial variability on the easterly site is much higher than the westerly site. It was only possible to derive the spatial variability from the synthetic solar irradiance, as we only had one ground measurement device. From these two analyses, we can anticipate maximum ramping possibilities and probabilistic yields with which to finalize the design of the plant. Without synthetic data, we would have had to rely on typical meteorological years (TMYs). TMYs are a collection of the most standard months that have been observed (we delve much deeper into this in Chaps. 2 and 5). However, they do not represent the full range of yield, hence they induce more risk than strictly necessary.

From this one example we can see simply how synthetic data could be useful in PV farm siting and planning. However, there are many more applications where it can pay dividends. In the next chapter, when defining synthetic solar irradiance more formally, we present a few different applications where synthetic irradiance is useful and practical. In the final chapter, we look forward to more possible applications of synthetic solar irradiance. Chapter 3 shows all the different techniques tried to actually produce synthetic solar irradiance, and Chap. 4 shows how we validate the output time series.

In summary, synthetic solar irradiance at present has two key styles: downscaling and entirely synthetic. Downscaling is a form of representative interpolation, whereby a coarse temporal resolution time series (such as hourly irradiance) is increased to a higher temporal resolution (such as 1 min) by introducing realistic solar irradiance variability. Entirely synthetic solar irradiance does not necessarily start with a coarse time series (though it could for training purposes), and instead generates synthetic solar irradiance based on stochastic elements or any other statistical inference. With entirely synthetic irradiance, the longer-term averages are not required to match any observation, for example, entirely synthetic solar irradiance could be used to generate 1000 irradiance time series for a year in the future.

1.7 REQUISITE SKILLS

We should briefly touch upon the topic of what kind of mathematics and experience is needed to enter the domain of synthetic solar irradiance. Whilst we have presented the basic approaches to modeling the sun, in that we can now arrive at estimates for solar potential anywhere in the world, we must now focus on specifics for synthetic solar irradiance.

Synthetic solar irradiance models are almost always data driven and in the form of statistical models. There is no simple formula that can be used to estimate synthetic solar irradiance, as such synthetic models are not analytic. Analytic models require certain observational inputs and leverage a physically trained model to arrive at an estimated variable, such as clear-sky irradiance models that accept variables such as aerosol scattering and use the physical relationship between the concentration and size of the aerosols to estimate the direct and diffuse irradiance. Instead, synthetic irradiance models tend to always derive statistical properties from data and apply them with various time series analysis techniques. There is no single factor responsible for the randomness we observe in ground GHI data; there are in fact many, some of which we know about. It is broadly impossible to determine them all and isolate them into empirical relationships. This difficulty is what prevents analytical methods in the field of solar meteorology in general. By using statistical methods, we are able to separate what is known from what is unknown. What is known is represented in the deterministic components (EXT, SZA, GHIcs, etc.). What is not known is represented by a random component implemented by distributions extracted from data (plus any relationships the modeler believes are true). Hence, a solid background in statistics, time series analysis, and data analysis are essential to enter this field.

The research field is extensively performed in computer code—normally open-source scripting language options. There is an element of computational development needed in order to engage with the ideas presented in this book. This is because validation and reproducibility are vital components to research, hence, the models must be tested against ample amounts of data. One does not need to be a computer developer to appreciate or learn about this field; however, in order to develop and test models, it is expected that the reader knows how to code in a scripting language.

1.8 A COMMON MISCONCEPTION

There is a common misconception that arises regularly for those entering the field of synthetic solar irradiance or with those who have a passing interest. The misconception is that synthetic solar irradiance models are predictive. By this, I mean that some people incorrectly believe that synthetic solar irradiance models attempt to derive the actual measurement at that exact time and location, such that it would be accurate to measured data were it to have actually been measured. This misconception often leads people to believe that synthetic solar irradiance models are equivalent to forecasting models or that they are versions of solar resource estimation techniques.

A solar forecasting model attempts to estimate the exact magnitude of irradiance at a point in the future. A solar resource methodology attempts to estimate the exact magnitude of irradiance that occurred at some point in history (up until the immediate present). What these two techniques have in common is that they both have the objective of arriving at the exact expected magnitude of irradiance. Both models use as much information as is feasibly available to them. Synthetic solar irradiance is not constrained by this feature; it is not required to find the actual magnitudes, just as long as the collection of magnitudes are appropriately representative of the location in question.

Of course, it would be a desirable trait if synthetic solar irradiance could produce accurate interpolated data. However, this is impossible without additional information. Attempting to downscale an hourly average into what truly happened at the 1-minute scale is a form of inverse problem. A forward problem is one where we calculate the likely outcome from a set of data. An example would be that we see an elephant walking across a muddy plain. If we know the size and weight of the animal, we can anticipate that there would be footprints of a certain depth and frequency. The inverse of this problem is that we arrive at the muddy plain to observe the footprints left behind. With only this information, we must infer the weight, size, and type of the animal that made the footprints. This might seem trivial, because not all inverse problems are impossible if they are well-posed problems. However, if there is insufficient detail, the problem becomes ill-posed. Jacques Salomon Hadamard defined a well-posed problem as having a unique solution. We know that synthetic downscaling can have numerous versions that satisfy the end result.

Consider a digital picture of the Mona Lisa as shown in Fig. 1.9. The more pixels used to build the image, the clearer the image is to us. Each pixel is simply the average color of the data contained within that small part of the painting. However, we lose detail with fewer pixels in a process known as pixelation. It is very easy for us to apply pixelation to the starting image. However, how do we invert the problem to arrive at the original version from the pixelated version? This is an example of an "inverse problem," as we have the result (the pixelated image) and we wish to arrive at the start (the original image). Some machine-learning approaches may be able to take the whole pixelated image and identify it as the Mona Lisa and then simply replace the pixelated version with the copy of the original from a database (this would be an effective method of solving this inverse problem). However, what if we never had the original version in the first place? Even if we were able to identify the pixelated version as the Mona Lisa, could we really recreate all the detail of da Vinci's brush strokes and color selection with any accuracy? No. This is because pixelation is a destructive process whereby the underlying data are destroyed. We could use certain photo editing software tricks to blur and sharpen and arrive at a reasonable representation of the Mona Lisa, but never the same as the original.

Let us use a modern example. Have you seen blurry photography taken from CCTV footage released by the police to try to identify a suspect? What the police are doing, in essence, is expanding their training database by accessing the public's memories for matching faces. The police probably know many common suspects of an area; however, if the new suspect does not remind them of anyone,

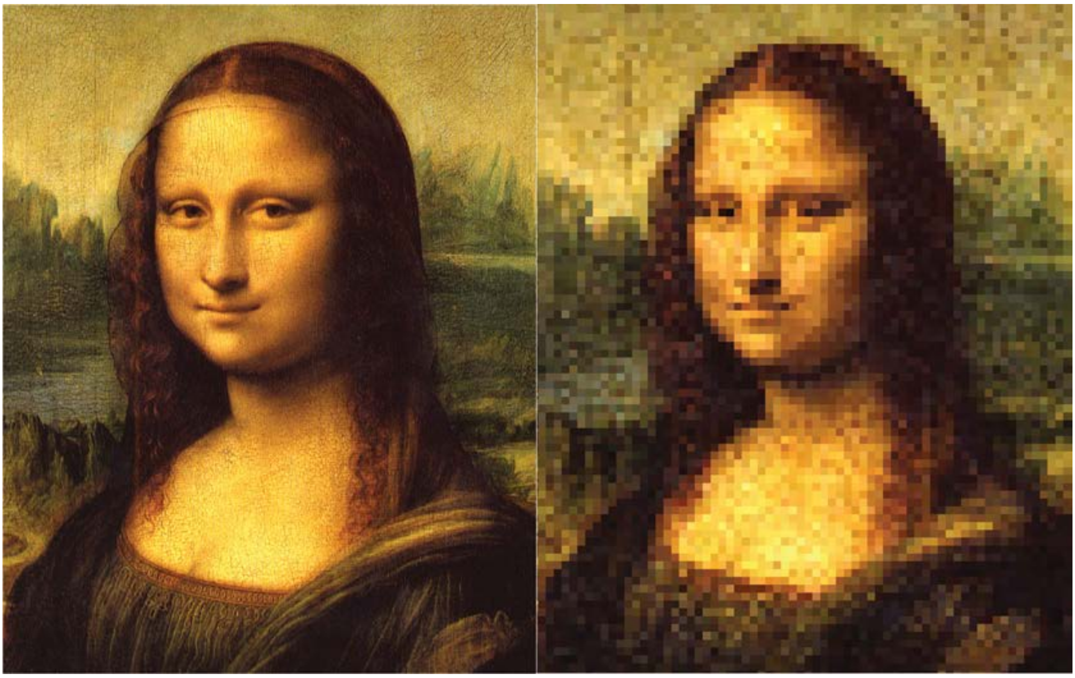

FIG. 1.9
Pixelation of the Mona Lisa.

then they will need help with identification. They hope that someone will see the blurry image and be able to recognize the suspect based on their personal "database." This is because it is not possible to accurately infer more information of what is behind each pixel without some additional source of data or access to the original. We would define this inverse problem as ill-posed as it has no unique solution. There are an extraordinarily large number of solutions to unpixelating the Mona Lisa. Consider just a single pixel that we wish to convert into a 10×10 grid. As each new pixel can be potentially one of 16 000 000 colors (24-bit), the number of permutations become incredibly large. The number of solutions is clearly not unique. This same principle applies to a solar irradiance time series measured at 1-hour resolution when we are trying to obtain 1-minute resolution. Fortunately in this case we have 60 new data points to construct that must obey certain constraints that we introduce.

Considering that synthetic irradiance downscaling is an inverse problem that attempts to leverage only the resulting data, we know that it will never arrive at the true answer except by rare chance. Any attempts to improve the predictive accuracy of synthetic solar irradiance will require the provision of supporting data, such as sky imagers, additional satellite imagery, or nearby sensors. This is, of course, doable. However, we would then have left the research field of synthetic solar irradiance and instead be attempting to perform solar resource estimation.

Certain synthetic solar irradiance techniques leverage transition matrix probabilities of particular meteorological variables or directly on the clear-sky index. As these probabilities are trained on historical data, there is some probabilistic forecasting potential for synthetic solar irradiance, whereby synthetic irradiance is generated for many future scenarios, and the distribution of expected outcomes forms the basis of a forecast. This, however, is an application of synthetic solar irradiance to the field of solar forecasting. The synthetic time series would be able to generate a representative solar scenario for the next year, not just the next 30 minutes.

1.9 KEY OBJECTIVES OF SYNTHETIC SOLAR IRRADIANCE

This section aims to make explicit the objective of synthetic solar irradiance. This chapter has already touched on the some of the key aspects within the field. We have categorized synthetic solar irradiance into two classes: either entirely synthetic data or downscaled data. Either type can be spatial, temporal, or spatiotemporal in nature.

The main aim of synthetic solar irradiance is to provide solar irradiance data in cases where there are no ground-measured data or where the global gridded irradiance data are too coarse in spatial or temporal resolution. This extends to scenario generation whereby we may have ground measured data but wish to use them to produce many more scenarios for the same place.

Fundamentally, all types of synthetic solar irradiance attempt to achieve the same objectives:

- To produce data that statistically resemble real data.
- To produce data that capture equivalent variability to real data.
- To produce data that are plausible, yet can capture all extents of possibility.
- To produce data that are representative of observation data for the geography of the target location, both temporally and spatially (if applicable).

These targets all have objective measures that can be demonstrated and should be validated.

There are stretch goals that would benefit synthetic solar irradiance models too. They mainly come as a wish list for future activity in the field:

- New models should be coded and made available under a license that encourages research.
- The code should be computationally efficient.
- The model should not require obscure input/training data.
- Results and models should be easily reproducible/accessible, such that the code is very well documented and provided with a reproducible example.

- Models should be trained on well-known datasets that are publicly available.
- Models should not be site-dependent and should be tested worldwide if possible.

1.10 SUMMARY

This book aims to set out the current state of synthetic solar irradiance. The research field has a relatively small core group of researchers actively participating in the field; of those, most have contributed to a chapter in this book. We do not pretend to be an authority on this topic. What we do hope is that you feel welcomed and invited to the research field, and that this book serves as a good starting point for your journey into this research space.

I hope that, by the end of this chapter, you now have a clearer understanding of synthetic solar irradiance. Additionally, I hope that you now understand and appreciate the context as to where synthetic solar irradiance may be useful, the objectives to which researchers strive toward, and the fundamental definitions of solar energy modeling so that the rest of this book can be enjoyed.

Chapter 2 aims to deliver to the reader further in-depth definitions and applications of synthetic solar irradiance. Chapter 3 will guide you through specifics of the methodologies attempted so far. Chapter 4 addresses the complex challenge of validating synthetic data. Chapter 5 explores a case study as to how synthetic solar irradiance can be applied to industry. Finally, Chap. 6 delivers our thoughts as to where the research field may move in the future.

ACKNOWLEDGMENTS

Jamie M. Bright is funded by the Energy Market Authority (EMA), Energy Programme–Solar Forecasting Grant (NRF2017EWT-EP002-004). SERIS is a research institute at the National University of Singapore (NUS). SERIS is supported by NUS, the National Research Foundation Singapore (NRF), EMA, and the Singapore Economic Development Board (EDB).

Thanks to Âzeddine Frimane for reading this chapter and making excellent suggestions for improvement. Thanks also to my mother, Karen Bright, who diligently checked my writing and pointed out the difference between bearing fruit and baring fruits, an important life lesson. She is now well versed on the basic concepts of synthetic solar irradiance. Thanks also to my father, Mark Bright, who during his read of the chapter pointed out the difference between a steakholder and a stakeholder. We must always check our work and be conscious of homonyms, lest they distract from the science. I would also like to thank my supportive wife, Anna Flanagan, who has heard too much about this book over the last year.

REFERENCES

Bainbridge, D. and Haggard, K., *Passive Solar Architecture: Heating, Cooling, Ventilation, Daylighting and More Using Natural Flows* (Chelsea Green Publishing, White River Junction, VT, USA, 2011). ISBN 9781603584203.

Behling, S. and Behling, S., *Solar Power: The Evolution of Sustainable Architecture*, Architecture Series (University of Virginia, USA, Prestel, 2000). ISBN 9783791324111.

Breeze, P., *Power Generation Technologies* (Elsevier Science, Burlington, MA, USA, 2014). ISBN 9780080983363.

Bright, J. M., Bai, X., Zhang, Y., Sun, X., Acord, B., and Wang, P., "irradpy: Python package for merra-2 download, extraction and usage for clear-sky irradiance modelling," Solar Energy **199**, 685–693 (2020). ISSN 0038-092X. http://www.sciencedirect.com/science/article/pii/S0038092X20301894

Brownson, J. R. S., "Sun-earth geometry," in *Solar Energy Conversion Systems*, edited by J. R. S. Brownson (Academic Press, Boston, 2014), Chap. 6, pp. 135–178. ISBN 978-0-12-397021-3. http://www.sciencedirect.com/science/article/pii/B9780123970213000065

Chapin, D. M., Fuller, C. S., and Pearson, G. L., "A New Silicon p-n Junction Photocell for Converting Solar Radiation into Electrical Power," J. Appl. Phys **25**, 676 (1954).

Cherp, A., Jewell, J., Vinichenko, V., Bauer, N., and De Cian, E., "Global energy security under different climate policies, GDP growth rates and fossil resource availabilities," Clim. Change **136**, 1–12 (2013). ISSN 0165-0009.

Chwieduk, D., "Energy balance of a building with regard to solar radiation exposure," in *Solar Energy in Buildings*, edited by D. Chwieduk (Academic Press, Oxford, 2014), Chap. 6, pp. 173–288. ISBN 978-0-12-410514-0.

Duffie, J. A. and Beckman, W. A., *Solar Engineering of Thermal Processes* (Wiley, Hoboken, NJ, 2013). ISBN 9781118415412.

Forbes, R. J., *Studies in Ancient Technology*, *Vol. V*. (Brill Archive, Leiden, 1957).

Frimane, Â., Bright, J. M., Yang, D., Ouhammou, B., and Aggour, M., "Dirichlet downscaling model for synthetic solar irradiance time series," J. Renew. Sustain. Energy **12**(6), 063702 (2020).

Gueymard, C. A., "The sun's total and spectral irradiance for solar energy applications and solar radiation models," Solar Energy **76**(4), 423–453 (2004). ISSN 0038-092X. http://www.sciencedirect.com/science/article/pii/S0038092×03003967

Gueymard, C. A., "Rest2: High-performance solar radiation model for cloudless-sky irradiance, illuminance, and photosynthetically active radiation–validation with a benchmark dataset," Solar Energy **82**(3), 272–285 (2008). ISSN 0038-092X. http://www.sciencedirect.com/science/article/pii/S0038092X07000990

Gueymard, C. A., "A reevaluation of the solar constant based on a 42-year total solar irradiance time series and a reconciliation of spaceborne observations," Solar Energy **168**, 2–9 (2018). ISSN

0038-092X. http://www.sciencedirect.com/science/article/pii/S0038092X18303463. Advances in Solar Resource Assessment and Forecasting.

Hornung, E. and Lorton, D., *Akhenaten and the Religion of Light*. G - Reference, Information and Interdisciplinary Subjects Series (Cornell University Press, Ithaca, NY, 2001). ISBN 9780801487255.

Iqbal, M., *An Introduction to Solar Radiation* (Academic Press, Don Mills, Ontario, Canada, 1983). ISBN 0-12-373750-8.

Kerr, R. M., *On Ancient Lights: And the Evidence of Surveyors Thereon: with Tables for the Measurement of Obstructions* (J. Murray, London, 1865).

Law Commission, "Rights to light: Law com no 356," Technical report, UK Government, December 2014. https://www.gov.uk/government/publications/rights-to-light. ISBN 9781474112581, Law Com. 356, HC 796 2014-15.

Lefèvre, N., "Measuring the energy security implications of fossil fuel resource concentration," Energy Policy **38**(4), 1635–1644 (2010). ISSN 0301-4215.

Lefèvre, M., Oumbe, A., Blanc, P., Espinar, B., Gschwind, B., Qu, Z., Wald, L., Schroedter Homscheidt, M., Hoyer-Klick, C., Arola, A., Benedetti, A., Kaiser, J. W., and Morcrette, J.-J., "McClear: A new model estimating downwelling solar radiation at ground level in clear-sky conditions," Atmos. Meas. Tech. **6**, 2403–2418 (2013). https://hal-mines-paristech.archives-ouvertes.fr/hal-00862906

Loferski, J. J., "The first forty years: A brief history of the modern photovoltaic age," Prog. Photovolt.: Res. Appl. **1**(1), 67–78 (1993). https://onlinelibrary.wiley.com/doi/abs/10.1002/pip.4670010109

Mertens, K., *Photovoltaics: Fundamentals, Technology and Practice* (Wiley, Chichester, 2013). ISBN 9781118703373.

Morris, N., *Solar Power*, *Energy Sources Series* (Smart Apple Media, North Mankato, MN, 2006). ISBN 9781583409084.

Neville, R. C., *Solar Energy Conversion: The Solar Cell* (Elsevier Science, Amsterdam, 1995). ISBN 9780080541426.

Ostdiek, V. and Bord, D., *Inquiry into Physics* (Cengage Learning, Boston, MA, USA, 2012). ISBN 9781133104681.

Rappaport, P., "The electron voltaic effect in p-n junctions induced by beta-particle bombardment," Phys. Rev **93**, 246 (1954).

Reynolds, D. C., Leies, G., Antes, L. L., and Marburger, R. E., "Photovoltaic effect in cadmium sulfide," Phys. Rev. **96**, 533 (1954).

Simms, D. L., "Archimedes and the burning mirrors of Syracuse," Technol. Cult. **18**(1), 1–24 (1977). ISSN 0040165X, 10973729.

Smith, C. J., Bright, J. M., and Crook, R., "Cloud cover effect of clear-sky index distributions and differences between human and automatic cloud observations," Solar Energy **144**, 10–21 (2017). ISSN 0038-092X. http://www.sciencedirect.com/science/article/pii/S0038092X16306624

Solanki, C. S., *Solar Photovoltaics: Fundamentals Technologies and Applications*, 3rd edition (Prentice-Hall of India Pvt. Limited, Delhi, 2009). ISBN 9788120351110.

Sun, X., Bright, J. M., Gueymard, C. A., Acord, B., Wang, P., and Engerer, N. A., "Worldwide performance assessment of 75 global clear-sky irradiance models using principal component analysis," Renew. Sustain. Energy Rev. **111**, 550–570 (2019). ISSN 1364-0321. http://www.sciencedirect.com/science/article/pii/S1364032119302187

Sun, X., Bright, J. M., Gueymard, C. A., Bai, X., Acord, B., and Wang, P., "Worldwide performance assessment of 95 direct and diffuse clear-sky irradiance models using principal component analysis," Renew. Sustain. Energy Rev. **135**, 110087 (2021). ISSN 1364-0321. http://www.sciencedirect.com/science/article/pii/S1364032120303786

Thorpe, D., *Solar Technology: The Earthscan Expert Guide to Using Solar Energy for Heating, Cooling and Electricity* (Taylor and Francis, London and New York, 2019). ISBN 9781135069148.

Tiwari, G. N. and S. Dubey, *Fundamentals of Photovoltaic Modules and Their Applications*, *RSC Energy Series* (Royal Society of Chemistry, Cambridge, UK, 2010). ISBN 9781849730204.

Winzer, C., "Conceptualizing energy security," Energy Policy **46**(0), 36–48 (2012). ISSN 0301-4215.

Wolfe, P., *Solar Photovoltaic Projects in the Mainstream Power Market* (Taylor & Francis, London, UK, 2013). ISBN 9781136201653.

Yang, D. and Bright, J. M., "Worldwide validation of 8 satellite-derived and reanalysis solar radiation products: A preliminary evaluation and overall metrics for hourly data over 27-years," Solar Energy **210**, 3–19, Special Issue on Grid Integration (2020). ISSN 0038-092X. http://www.sciencedirect.com/science/article/pii/S0038092X20303893