Chapter 2 Solar Radiation at the Top of the Atmosphere

2.1 The Sun as an Energy Source

Once upon a time about 4.6 billion years ago, the sun condensed out of the center of a thin, hot, spinning disk of interstellar material, according to a theory proposed by Laplace (1796). The sun is a typical second-generation G2 star in the cosmic hierarchy. Among the billions of stars in the universe, the sun is about average in mass but below average in size. The sun has one unique feature in that it is 300,000 times closer to the earth than the next nearest star. With a mean distance of about 1.5×10^8 km between the earth and the sun, virtually all of the energy that the earth receives and that sets the earth's atmosphere and oceans in motion comes from the sun.

The sun is a gaseous sphere with a visible radius of about 6.96×10^5 km and a mass of approximately 1.99×10^{30} kg. Its main ingredients are primordial hydrogen (H) and helium (He), plus a small amount of heavier elements including oxygen (O), carbon (C), nitrogen (N), neon (Ne), iron (Fe), silicon (Si), magnesium (Mg), sulfur (S), and calcium (Ca). Hydrogen makes up roughly 90% of the mass, and the remaining 10% or so is helium. The temperature of the sun decreases from a central value of about 5×10^6 K to about 5800 K at the surface. The density within the sun falls off very rapidly with increasing distance from the center. The central density is about 150 g cm^{-3}, and at the surface, it is about 10^{-7} g cm^{-3}. The average density is about 1.4 g cm^{-3}. Approximately 90% of the sun's mass is contained in the inner half of its radius.

Solar energy is believed to be generated by the steady conversion of four hydrogen atoms to one helium atom in fusion reactions, which take place in the deep interior of the sun with temperatures up to many millions of degrees, as shown in Fig. 2.1. The amount of energy released in nuclear fusion causes a reduction of the sun's mass. According to Einstein's law relating mass and energy, $E = mc^2$, and converting the energy radiated by the sun, we find that almost 5 million tons of mass per second are radiated by the sun in the form of electromagnetic energy. In a billion years, it is believed that the sun will radiate into space about 10^{26} kg, which is less than one part in 10^4 of its total mass. Thus, only an insignificant fraction of the sun's substance has been lost by electromagnetic radiation. It is estimated that only 5% of the sun's total mass has been converted from hydrogen to helium in its lifetime thus far.

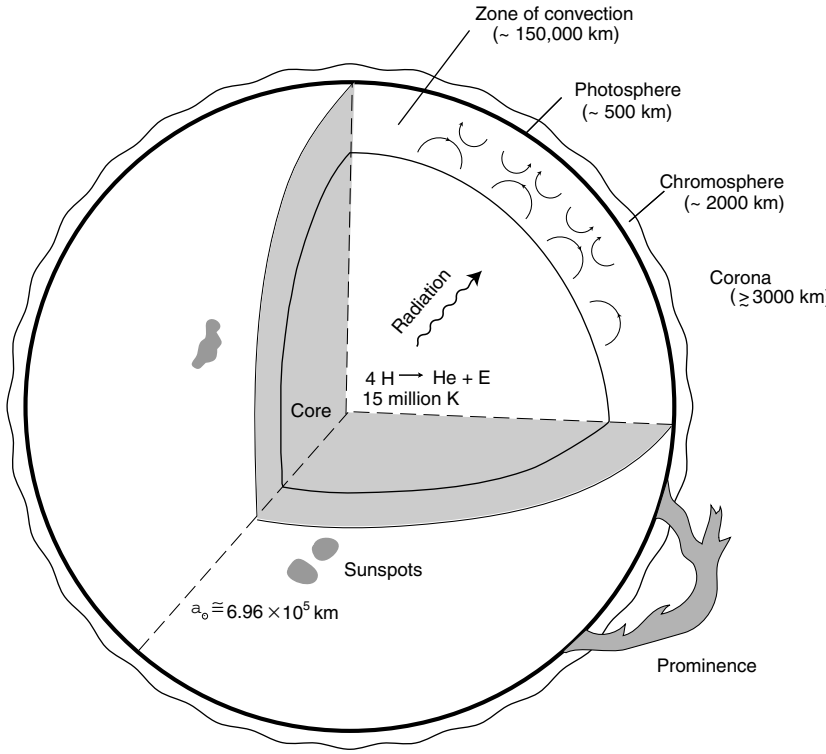

Figure 2.1 A cross section of the sun illustrating the solar interior and atmosphere. The solar interior includes the core with a temperature of about 1.5×10^7 K, the radiation zone, and the convective zone. The solar atmosphere includes the photosphere, the chromosphere, and the corona. The former two layers are exaggerated for illustration purposes.

As a result of the extremely high temperatures in the deep interior of the sun, collisions between atoms are sufficiently violent to eject many electrons from their orbits. Only the tightly bound inner electrons of heavy atoms will be retained. The energy emitted by nuclear fusion in the form of photons can pass through the inner part of the sun without being absorbed by the electrons. However, closer to the sun's surface, the temperature decreases and the heavier atoms such as iron begin to recapture their outer electrons. These outer electrons are bound to the nucleus by relatively small forces and can be easily separated from the nucleus by the absorption of photons. It follows that the flow of photons coming from the interior is blocked by the appearance of the absorbing atoms. The blocking of these photons will cause the temperature to drop sharply at some depth below the surface. Thus, the outer region of the sun consists of a layer of relatively cool gas resting on top of a hotter interior. As a consequence, the gas at the bottom of the cool outer layer is heated by the hot gas in the interior. It undergoes expansion and rises toward the surface. Once it reaches the surface, the hot gas loses its heat to space, cools, and descends into the interior. The entire outer

Table 2.1
Magnitude and Variability of the Solar Sources of Terrestrial Energy[a]

Source	Energy ($W\,m^{-2}$)	Solar cycle change ($W\,m^{-2}$)	Terrestrial deposition altitude
Solar radiation			
Total irradiance	1366	1.3	Surface, troposphere
UV 2000–3000 Å	15.4	0.16	0–50 km
UV 0–2000 Å	0.1	0.02	50–500 km
Particles			
Solar protons	0.002		30–90 km
Galactic cosmic rays	0.000007		0–90 km
Solar wind	0.0003		About 500 km

[a]Data taken from National Research Council (1994).

layer breaks up into ascending columns of heated gas and descending columns of cooler gas. The region in which this large-scale upward and downward movement of gases occurs is called the *zone of convection* (Fig. 2.1), which extends from a depth of about 150,000 km to the surface of the sun. Below this depth, it is believed that energy is transported within the sun by means of electromagnetic radiation, i.e., by the flow of photons. Near the surface, however, because of the substantial blocking of radiant energy by the absorption of heavier elements, energy is transferred partly by convection and partly by electromagnetic radiation. Above the surface, energy transport is again by means of electromagnetic radiation. The sun provides electromagnetic, particle, and plasma energy to the earth as summarized in Table 2.1. It is clear that electromagnetic radiation characterized by wavelengths from gamma rays to radio waves (see Fig. 1.1) is by far the largest solar energy source for the earth and the most important for its weather and climate processes.

2.1.1 The Structure of the Sun

The visible region of the sun is called the *photosphere,* where most of the electromagnetic energy reaching the earth originates. Although the sun is a gaseous body, the photosphere is referred to as the *surface* of the sun. The photosphere is marked by relatively bright *granules* about 1500 km in diameter, which are separated by dark regions and variable features called *sunspots*. The bright granules are fairly uniformly distributed over the solar disk and are believed to be associated with ascending hot gases in the uppermost layer of the zone of convection discussed previously.

The photosphere is a comparatively thin layer about 500 km thick that constitutes the source of the sun's visible radiation, as illustrated in Fig. 2.1. The temperature in this layer varies from 8000 K in the lower layer to 4000 K in the upper layer. Matching the theoretical Planck curve versus wavelength (see Fig. 2.9) with the measured spectral radiant energy emitted by the sun, the best agreement was found for a temperature of approximately 5800 K. This temperature is an average over the

temperature range of the photosphere. The *effective temperature* of the photosphere may also be obtained by measuring the luminosity of the sun. On the basis of the Stefan–Boltzmann Law, and with the knowledge that the sun's radius assumes a sharply defined surface, the effective temperature also yields a value close to 5800 K. Radiation emitted from the photosphere is essentially continuous and will be discussed further in Section 2.3.1.

The region above the photosphere is called the *solar atmosphere*. It is characterized by tenuous and transparent solar gases. The solar atmosphere is divided into two regions called the *chromosphere* and the *corona* (see Fig. 2.1). The chromosphere lies above the photosphere to a height of approximately 2000 km. Its temperature increases from a minimum of about 4000 K and stays between 4000 and 6000 K up to about 2000 km. Above this height, the temperature rises drastically, reaching about 10^5–10^6 K. The temperature structure of the solar atmosphere is shown in Fig. 2.2, in

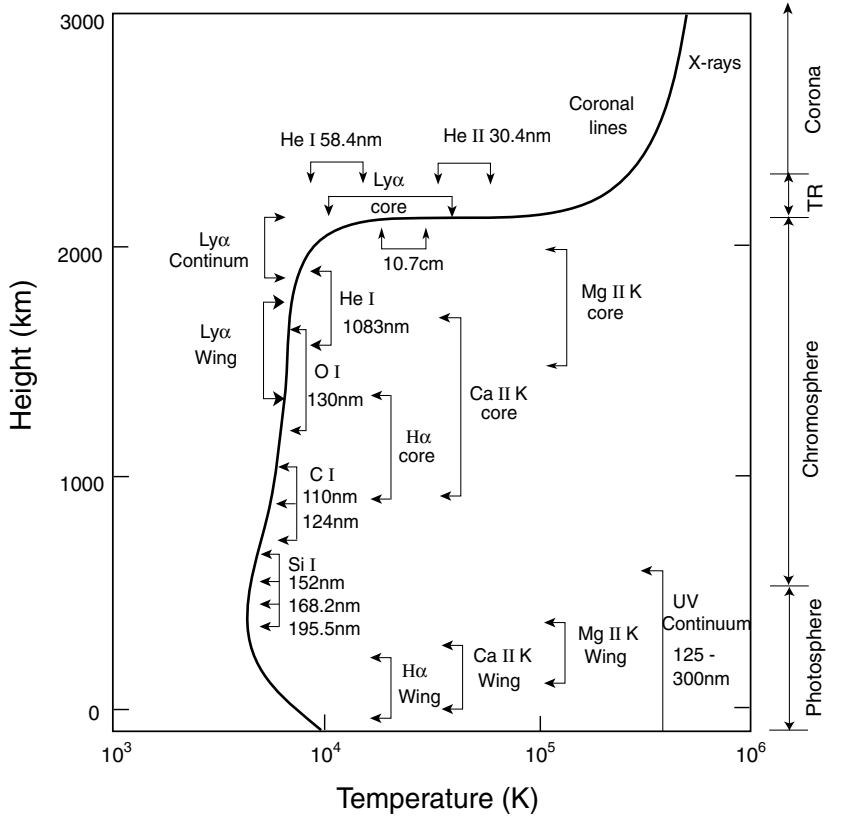

Figure 2.2 The temperature structure of the solar atmosphere and the regions where absorption lines are formed. Data constructed from the results published in Lean (1987) and Livingston *et al.* (1988). The symbols I and II denote a neutral atom and a singly ionized atom, respectively, and TR refers to the transition region.

which the regions of formation of different absorption lines discussed below are also displayed. The layer with a minimum temperature of 4000 K extends to a few thousand kilometers, consisting of relatively cool gases lying over the hotter gases. These cool gases absorb continuous radiation emitted from the photosphere at wavelengths characteristic of the atoms in the sun, and generate the solar absorption spectrum. As was discussed in Section 1.3.1, when an atom absorbs radiant energy, it is excited to a new energy level. The excited atom then makes a transition to a lower excited state or to the ground state, during which a quantum of energy is emitted. Consequently, the emission spectrum of the chromosphere is formed. Since the absorption spectrum is produced by the initial transition of atoms from a low-energy to a high-energy state, while the emission spectrum results from the subsequent transition of the same atoms in the reverse direction, it is clear that the lines in the sun's emission spectrum are the same as those in its absorption spectrum. When the photosphere is eclipsed by the moon or by instrument, line emissions, mostly from hydrogen, helium, and calcium, can be observed. Because a bright line emission spectrum flashes into view briefly at the beginning and the end of a total eclipse, it is called the *flash spectrum*. The Hα line at 6563 Å is one of the strongest absorption lines in the solar spectrum. Because of the large amount of energy emitted in this line, the chromosphere becomes visible and has a characteristic reddish appearance during an eclipse.

Above the chromosphere lies the region of the solar atmosphere called the *corona*. The corona layer extends out from the edge of the solar disk many millions of kilometers. It is visible as a faint white halo during total eclipses. Figure 2.3 illustrates the solar corona during the total eclipses of March 1970 and July 1991. It is generally believed that the corona has no outer boundary. A stream of gas called *solar wind* (see Table 2.1) flows out of the corona and into the solar system continuously. An instrument called a *coronagraph* has been used in the past to study both the chromosphere and the corona in the absence of a natural eclipse. Strong emission lines of hydrogen and helium originating within the chromosphere disappear with increasing altitude and are replaced by the continuous spectrum of white light characteristic of the corona. The spectrum of the corona contains a number of weak emission lines, of which the most intense is the green line of ionized iron. The generation of this emission line requires an enormous amount of energy, and it is believed that the temperature in large regions of the corona is close to 10^6 K.

2.1.2 Solar Surface Activity: Sunspots

Several observable features of the sun are particularly interesting and important because of their transient nature. The best known and most frequently observed of these variable features are *sunspots,* which are relatively dark regions on the photosphere—the surface of the sun. Sunspots have an average size of about 10,000 km but range from barely visible to areas that cover more than 150,000 km on the sun's surface. The spots usually occur in pairs, or in complex groups, which follow a leader spot in the direction of the sun's rotation. Small sunspots persist for several days or a week, while the largest spots may last for several weeks, long enough for these spots to

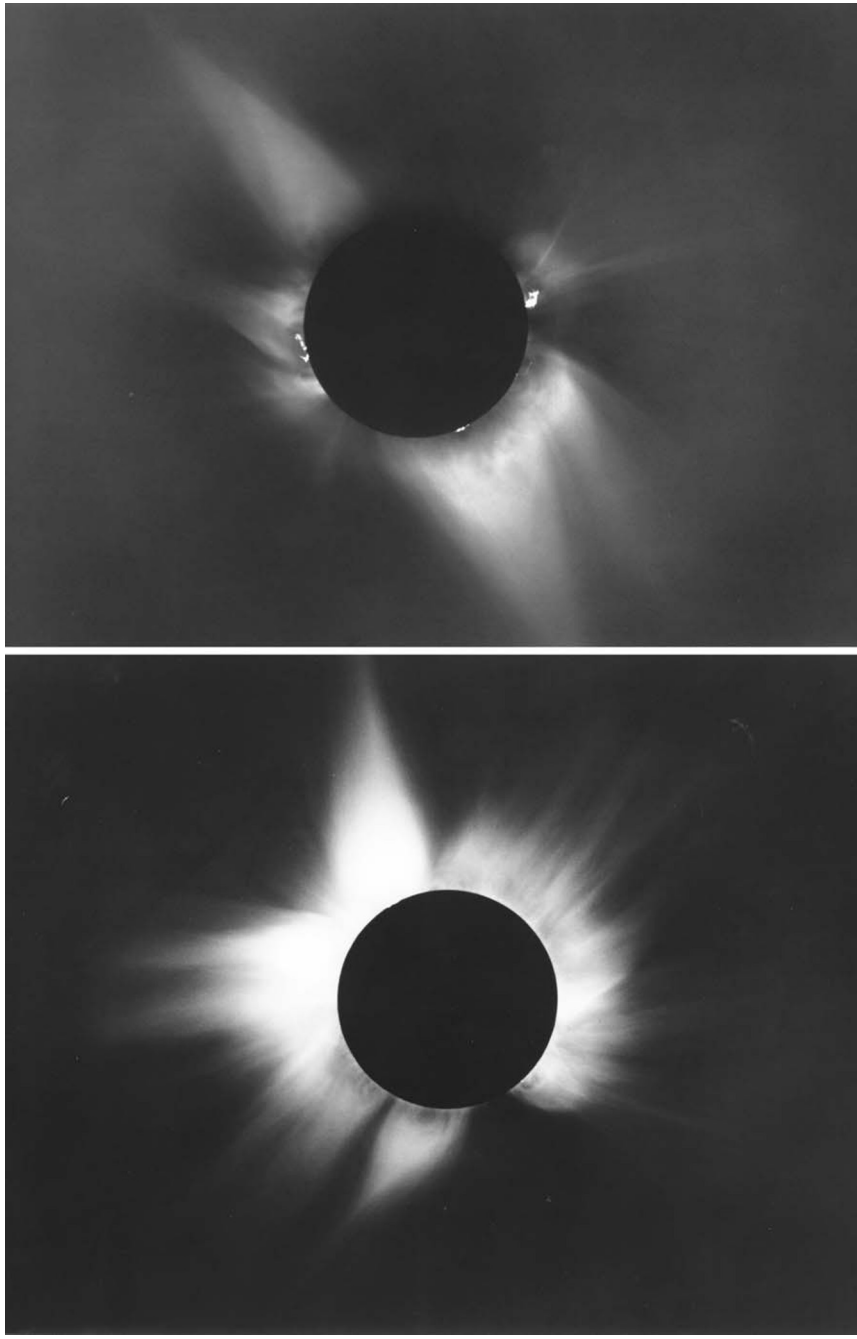

Figure 2.3 The top picture is the total solar eclipse of July 11, 1991, photographed from Mauna Loa, Hawaii. The bottom picture is the solar corona during the total eclipse of March 7, 1970. Features are visible at a distance of about 4.5 solar radii or 3 million kilometers (courtesy of Rhodes College and High Altitude Observatory, Boulder, Colorado).

reappear during the course of the sun's 27-day rotation. Sunspots are almost entirely confined to the zone of latitude between 40° and the equator, and they never appear near the poles. Just after their minimum occurrence, spots first appear near latitude 27° in both hemispheres. As the cycle proceeds, they drift toward the equator and disappear close to latitude 8°.

Sunspots are cooler regions with an average temperature of about 4000 K, compared to an average temperature of 6000 K for the photosphere. Because of their relatively low temperature, sunspots appear black. Sunspot activity has been observed with the aid of high-power telescopes. The number of sunspots that appear on the solar disk averaged over a period of time is highly variable. There are periods of time when spots are relatively numerous, whereas a few years later spots occur hardly at all. These periods are called *sunspot maxima* and *sunspot minima,* respectively. The periodic change in the sunspot number is referred to as the *sunspot cycle*. For more than 200 years, the number of spots appearing every day and the position of these spots on the face of the sun have been recorded continuously. The average length of time between sunspot maxima is about 11 years; the so-called *11-year cycle*. Figure 2.4

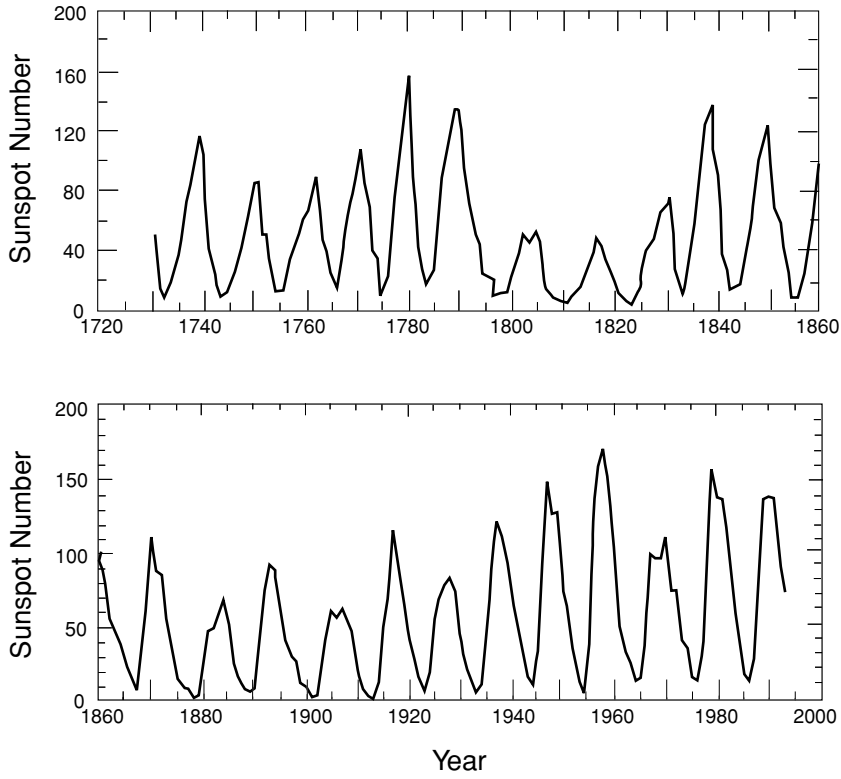

Figure 2.4 Variability of the sunspot number as function of year. Results from 1730 to 1870 are taken from Eddy (1977); those after 1860 are from Lean and Rind (1998).

depicts the variation in the number of sunspots since about 1730. The number ranges from a few to a maximum of about 150. In the years of sunspot maxima, the sun's surface is violently disturbed, and outbursts of particles and radiation are commonly observed. During sunspot minima periods, however, outbursts are much less frequent.

It is believed that sunspots are associated with the very strong magnetic fields that exist in their interiors. Magnetic field measurements show that pairs of sunspots often have opposite magnetic polarities. For a given sunspot cycle, the polarity of the leader spot is always the same for a given hemisphere. With each new sunspot cycle, the polarities reverse. The cycle of sunspot maxima having the same polarity is referred to as the *22-year cycle*. Sunspot activity has been found to have a profound influence on many geophysical phenomena and atmospheric processes.

Sunspots are not the only source of solar variability. When a sunspot is near to the limb, it can be seen to be surrounded by networks of enhanced emission, called *faculae,* which appear before and disappear after the sunspots. Although faculae have less magnetic flux than sunspots, they extend over considerably more of the sun's disk and persist longer. When the sun is viewed by monochromatic light from a single element such as the hydrogen Hα line, the sunspots are visible but are surrounded by bright areas, known as *plages*. These prominent outburst features also occur at high latitudes, but are usually observed in the vicinity of large, complex sunspot groups. They are known as *solar flares* and are associated with great increases of hydrogen Lyman α at 1216 Å and other ultraviolet radiation. The burst of radiation and energy particles from a large flare may produce interference with radio communications and cause substantial variations in the earth's magnetic field.

The sun's other transient features are the *prominences* produced by photospheric eruptions. They extend into the chromosphere and can be observed on the limb of the sun. A typical prominence may be 30,000 km high and 200,000 km long with a temperature of about 5000 K. Because prominences are cooler than the photosphere, they may be seen in Hα light as dark filaments on a bright background. Disturbances in the corona are closely related to the sunspot cycle and changes in sunspot number, based on solar radio emission observations. Outbursts are accompanied by large increases in the far-ultraviolet and x-ray emission from the sun.

All of the preceding variabilities are associated with magnetic activity. Variations in the magnetic field are produced by the interactions among the convective motion, the solar rotation, and the general magnetic field of the sun. Sunspots contain the strongest magnetic fields. As noted earlier, the polarity in sunspot pairs reverses in successive sunspot cycles.

2.2 The Earth's Orbit about the Sun and Solar Insolation

2.2.1 Orbital Geometry

The earth is one of the nine planets in the solar system. The four planets closest to the sun (i.e., Mercury, Venus, Earth, and Mars), are referred to as the terrestrial

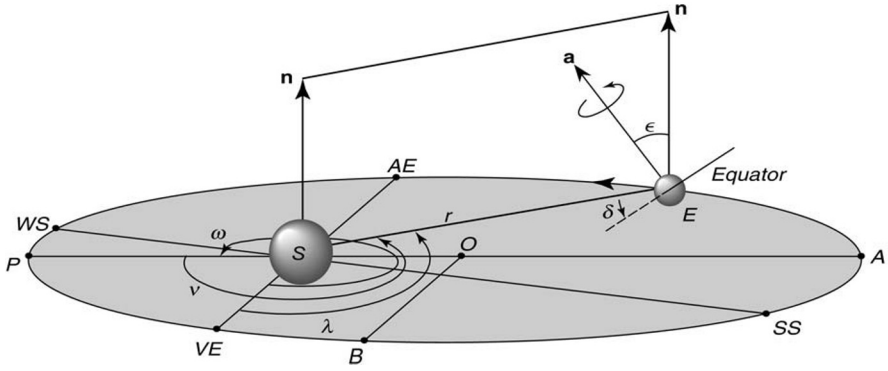

Figure 2.5 The earth–sun geometry. P denotes the perihelion, A the aphelion, AE the autumnal equinox, VE the vernal equinox, WS the winter solstice, and SS the summer solstice, **n** is normal to the ecliptic plane, **a** is parallel to the earth's axis, δ is the declination of the sun, ϵ the oblique angle of the earth's axis, ω the longitude of the perihelion relative to the vernal equinox, ν the true anomaly of the earth at a given time, λ the true longitude of the earth, O the center of the ellipse, OA (or $OP = a$) the semimajor axis, OB $(= b)$ the semiminor axis, S the position of the sun, E the position of the earth, and ES $(= r)$ the distance between the earth and the sun.

planets, and the remaining planets (i.e., Jupiter, Saturn, Uranus, Neptune, and Pluto), are called the major planets. All of the planets revolve around the sun in the same direction, and with the exception of Uranus, they also rotate in the same direction about their axes. In addition, all the planetary orbits except Mercury and Pluto lie in almost the same plane.

Once every 24 hours the earth makes a steady rotation eastward around the axis of its poles. This rotation is the cause of the most obvious of all time periods, involving the alternation of day and night, which comes about as the sun shines on the different parts of the earth exposed to it. At the same time, the orbital motion of the earth, with a mass of 6×10^{24} kg, in an ellipse with the sun at one focus, takes approximately 365 days to complete. The distance between the earth and the sun varies. The earth's orbit about the sun and the earth's rotation about its axis, which is tilted as shown in Fig. 2.5, are the most important factors determining the amount of solar radiant energy reaching the earth, and hence, the climate and climatic changes of the earth–atmosphere system. Because of the rotation of the earth about its axis, the earth assumes the shape of an oblate spheroid, having equatorial and polar radii of 6378.17 and 6356.79 km, respectively.

The position of the sun is defined by the *solar zenith angle,* which is determined from other known angles. In reference to Fig. 2.6 let P be the point of observation and OZ the zenith through this point. Assume that the sun is in the direction of OS or PS and let D be the point directly under the sun. Then the plane of OZ and OS will intersect the surface of the earth in a great circle. The angle ZOS, measured by the arc PD of this circle, is equal to the sun's zenith distance θ_0. In the spherical triangle NPD, the arc ND is equal to 90° minus the solar inclination δ, which is the angular distance

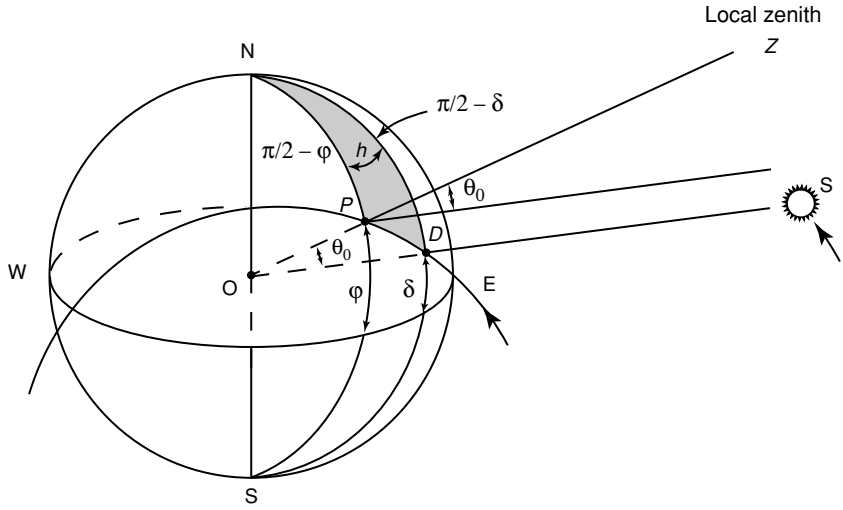

Figure 2.6 Relationship of the solar zenith angle θ_0 to the latitude φ, the solar inclination angle δ, and the hour angle h. P and D are the point of observation and the point directly under the sun, respectively (see text for further explanation).

of the sun north (positive) or south (negative) of the equator. The arc NP is equal to 90° minus the latitude ϕ of the observation point. The angle h is the hour angle, or the angle through which the earth must turn to bring the meridian of P directly under the sun, and is given by $h = 2\pi t/\Delta t_\odot$, where $\Delta t_\odot = 86{,}400$ sec. From spherical trigonometry, the cosine of the solar zenith angle as given in Fig. 2.6 and Appendix C is

$$\cos\theta_0 = \sin\varphi \sin\delta + \cos\varphi \cos\delta \cos h. \tag{2.2.1}$$

At the poles, $\varphi = \pm\pi/2$, so that $\cos\theta_0 = \sin\delta$ or $\theta_0 = \pi/2 - \delta$. The solar elevation angle $(\pi/2 - \theta_0)$ is equal to the declination of the sun. During the 6 months of daylight, the sun simply circles around the horizon and never rises more than about 23.5°. At solar noon at any latitude, the hour angle $h = 0$. Thus, $\cos\theta_0 = \cos(\varphi - \delta)$ or $\theta_0 = \varphi - \delta$. Except at the poles, $\theta_0 = \pi/2$ at sunrise and sunset. From Eq. (2.2.1), the half-day, H (i.e., from sunrise to noon or from noon to sunset), is defined by

$$\cos H = -\tan\varphi \tan\delta. \tag{2.2.2}$$

If $\varphi = 0$ (equator) or $\delta = 0$ (equinoxes), then $\cos H = 0$. The length of the solar day is 12 hours. The latitude of the polar night ($H = 0$) is determined by $\varphi = 90^\circ - |\delta|$ in the winter hemisphere. The solar inclination is a function of the day of the year only and is independent of the location of the observation point. It varies from $23^\circ 27'$ on June 21 to $-23^\circ 27'$ on December 22. The hour angle is zero at solar noon and increases by 15° for every hour before or after solar noon.

The plane of the earth's orbit is called the *plane of the ecliptic,* as shown in Fig. 2.5. The position of the sun is located at one of the foci of an ellipse. Let a and b denote the semimajor and semiminor axes of the ellipse. The eccentricity of the earth's orbit

is defined as the ratio of the distance between the two foci to the major axis of the ellipse and is given by

$$e = (a^2 - b^2)^{1/2}/a. \tag{2.2.3}$$

The tilt of the earth's axis with respect to normal to the ecliptic plane is defined by the oblique angle ε. The longitude of the perihelion, i.e., the closest point of the earth to the sun, relative to the vernal equinox, is defined by the angle ω. For a given time, the position of the earth is defined by the true anomaly ν in reference to the perihelion.

From the geometry shown in Fig. 2.5, the declination of the sun can be expressed in terms of the oblique angle of the earth's axis, the longitude of the perihelion relative to the vernal equinox, and the true anomaly of the earth at a given time. Also, letting the true longitude of the earth, counted counterclockwise from the vernal equinox, be λ, from three-dimensional geometry one can derive the following relationships:

$$\sin\delta = \sin\varepsilon \sin(\nu + \omega) = \sin\varepsilon \sin\lambda. \tag{2.2.4}$$

Having defined the relevant geometric parameters, we shall now introduce the three basic laws governing the motion of the planet. The planet's orbital ellipse can be described by Kepler's first law (the law of orbits), in which the distance is related to the true anomaly and eccentricity in the form

$$r = \frac{a(1 - e^2)}{1 + e\cos\nu}. \tag{2.2.5}$$

This equation can be derived from the geometry of an ellipse (Exercise 2.2). To compute the solar flux over a certain time period, Kepler's second law (the law of areas) is required. This law is a statement of the conservation of angular momentum that the radius vector, drawn from the sun to the planet, sweeps out equal areas in equal times (Exercise 2.3). Letting $\widetilde{T}$ denote the tropical year (i.e., the time between successive arrivals of the sun at the vernal equinox), which is equal to 365.2422 mean solar days, and noting that the area of an ellipse is πab, we have

$$r^2 \frac{d\nu}{dt} = \frac{2\pi a^2}{\widetilde{T}}(1 - e^2)^{1/2}. \tag{2.2.6}$$

In order to compute solar insolation, which will be defined in Subsection 2.2.3, the mean distance between the earth and the sun must be defined. From Kepler's second law, we may define a mean distance based on the conservation of angular momentum such that

$$r_0^2 = \frac{1}{2\pi}\int_0^{2\pi} r^2 d\nu = a^2(1 - e^2)^{1/2} \cong a^2. \tag{2.2.7}$$

Kepler's third law (the law of periods) states that

$$a^3/\widetilde{T}^{\,2} = k, \tag{2.2.8}$$

where $\widetilde{T}$ is the planet's period of revolution around the sun, and k has the same value for all planets. This law is a consequence of the balance between the gravitational

and centrifugal forces governing a planet in orbit. The semimajor axis of the earth's orbit is invariant. Because the factor $(1 - e^2)^{1/2}$ is very close to 1, the mean distance between the earth and the sun may be set as the invariant semimajor axis (i.e., $r_0 = a$).

Solar insolation as a function of latitude and the time of year, which will be defined in Eq. (2.2.21), requires the values of $(a/r)^2$ and δ, which in turn are computed from the eccentricity e, the oblique angle ε, and the longitude of the perihelion relative to the vernal equinox ω. Based on celestial mechanics, the secular variations in these three parameters are associated with the perturbations that other principal planets exert on the earth's orbit. Milankovitch (1941) has provided mathematical expressions for the computation of solar insolation including orbital parameters. Berger (1978) developed simplified trigonometric expansions for the efficient computation of the aforementioned three parameters. Figure 2.7 shows their values for the past 200,000 years. The eccentricity varies from about 0.01 to 0.04 with a mean value of about 0.017 and has a characteristic period of about 100,000 years. The oblique angle varies

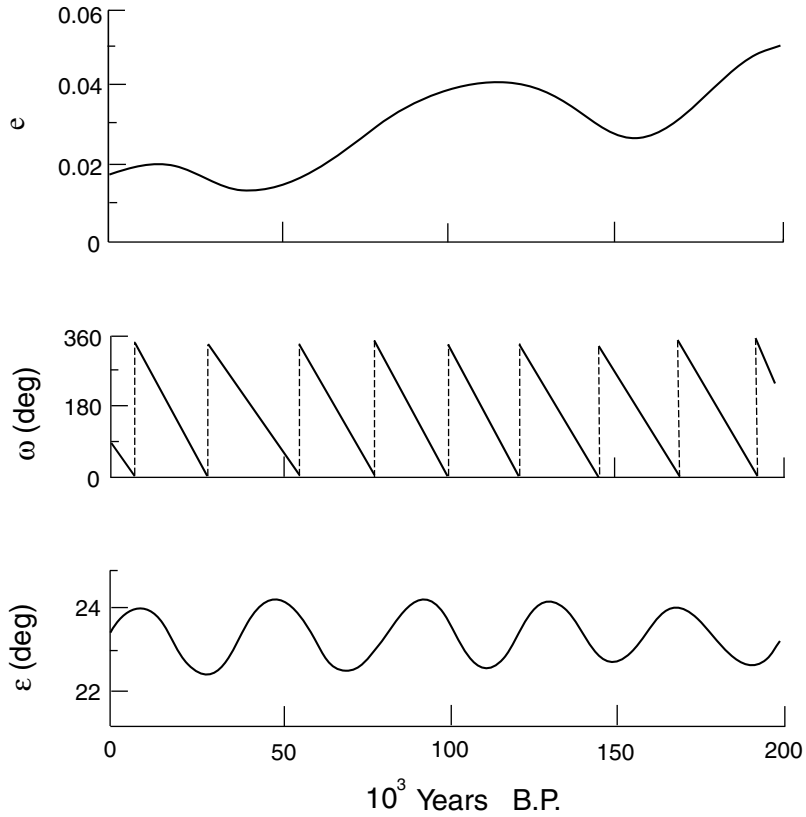

Figure 2.7 The eccentricity e, the obliquity of the ecliptic ε, and the longitude of the perihelion ω of the earth as functions of year before the present.

Table 2.2
Coefficients for the Calculation of the Sun–Earth Distance and the Declination Angle

n	a_n	b_n	c_n	d_n
0	1.000110	0	0.006918	0
1	0.034221	0.001280	−0.399912	0.070257
2	0.000719	0.000077	−0.006758	0.000907
3	—	—	−0.002697	0.000148

from about 22° to 24.5° with a dominant period of about 41,000 years. The longitude of the perihelion has a periodicity of about 21,000 years due to the advance of the perihelion by about 25 minutes each year, referred to as the periodic precession index.

The sun–earth distance can be approximated with an accuracy of about 10^{-4} as follows:

$$\left(\frac{a}{r}\right)^2 = \sum_{n=0}^{2}(a_n \cos nt + b_n \sin nt), \tag{2.2.9}$$

where $t = 2\pi d/365$, with $d = 0$ for January 1 and $d = 364$ for December 31. Also, the declination angle can be evaluated from

$$\delta = \sum_{n=0}^{3}(c_n \cos nt + d_n \sin nt), \tag{2.2.10}$$

with an accuracy of 0.0006 radians. The coefficients a_n, b_n, c_n, and d_n are listed in Table 2.2.

The most distinguishable feature of climatic change is the seasons. The revolution of the earth about the sun and the tilt of the earth's axis cause this seasonal variation. At the time of the summer solstice, which occurs about June 22, the sun appears directly overhead at noon at latitude 23.5°N, called the *Tropic of Cancer*. The elevation of the sun above the horizon and the length of the day reach their maximum values in the northern hemisphere at the summer solstice, and everywhere north of the Arctic Circle (latitude 66.5°N), the sun remains above the horizon all day. In the southern hemisphere on the June solstice, the sun's elevation is at a minimum, the days are shortest, and everywhere south of the Antarctic Circle (latitude 66.5°S), the sun does not rise above the horizon. This is the beginning of the northern hemisphere summer, whereas the southern hemisphere summer begins with the winter solstice on about December 22. Having reached the southernmost point in its annual migration, the sun then stands directly overhead at noon at latitude 23.5°S, called the *Tropic of Capricorn*. Both the elevation of the sun above the horizon and the length of the day are then at their minimum values in the northern hemisphere and their maximum values in the southern hemisphere, and the sun does not rise within the Arctic Circle or set within the Antarctic Circle. At the vernal (spring) and autumnal equinoxes, the days and nights everywhere are equal (12 hours), and the sun appears directly overhead on the equator at noon. The sun crosses the equator from north to south at

the autumnal equinox, and from south to north at the vernal equinox. The distances between the centers of the sun and the earth vary between the extreme values of 147×10^6 km at about winter solstice, and 153×10^6 km at about summer solstice. The mean distance is about 150×10^6 km, as noted earlier in Section 2.1.

2.2.2 Definition of the Solar Constant

The distribution of electromagnetic radiation emitted by the sun as a function of the wavelength incident on the top of the earth's atmosphere is called the *solar spectrum*. The *solar constant* S is a quantity denoting the amount of total solar energy (i.e., covering the entire solar spectrum) reaching the top of the atmosphere. It is defined as the flux of solar energy (energy per unit time) across a surface of unit area normal to the solar beam at the mean distance between the sun and the earth.

The sun emits energy at the rate of 6.2×10^7 W m^{-2}. On the basis of the energy conservation principle and if there is no intervening medium present, the energy emitted from the sun must remain the same at some distance away. Thus,

$$F_{\odot}^{*} 4\pi a_{\odot}^2 = S 4\pi r_0^2, \tag{2.2.11}$$

where $F_{\odot}^{*}$ denotes the solar emittance, $a_{\odot}$ the radius of the sun, and r_0 the mean distance between the sun and the earth. Hence, the solar constant may be expressed by

$$S = F_{\odot}^{*}(a_{\odot}/r_0)^2. \tag{2.2.12a}$$

Since the sun may be considered as an isotropic emitter, the solar intensity (or brightness) I is then given by $F_{\odot}^{*}/\pi$ [see Eq. (1.1.10)]. Thus, Eq. (2.2.12a) can be rewritten in the form

$$S = I \cdot \pi a_{\odot}^2 / r_0^2 = I \cdot \Omega, \tag{2.2.12b}$$

where the solid angle Ω is the angle from which the earth sees the sun. The intensity, i.e., the energy contained within the solid angle, is invariant: it is the same at the position of the sun as it is at the position of the earth. If the intensity varies within the solid angle, then we must use Eq. (1.1.9) to obtain the flux density. A number of exercises at the end of the chapter require the use of the intensity concept.

The total energy intercepted by the earth whose radius is a_e is given by $S\pi a_e^2$. If this energy is spread uniformly over the full surface of the earth, then the amount received per unit area and unit time at the top of the atmosphere is given by

$$\bar{Q}_s = S\pi a_e^2 \big/ \left(4\pi a_e^2\right) = S/4. \tag{2.2.13}$$

To estimate the equilibrium temperature $T_{\odot}$ of the sun, we use the blackbody assumption. From the Stefan–Boltzmann law, i.e., $F_{\odot}^{*} = \sigma T_{\odot}^4$, we find

$$T_{\odot}^4 = (r_0/a_{\odot})^2 (S/\sigma). \tag{2.2.14}$$

Inserting values of S, σ, r_0, and $a_{\odot}$ into Eq. (2.2.14), we obtain an equilibrium temperature of about 5800 K for the sun. Thus, once the solar constant is measured, the effective temperature of the sun can be computed.

2.2.3 Distribution of Solar Insolation

Solar insolation is defined as the flux of solar radiation per unit of horizontal area for a given locality. It depends primarily on the solar zenith angle and to some extent on the variable distance of the earth from the sun. The flux density at the top of the atmosphere may be expressed by

$$F = F_{\odot} \cos\theta_0, \tag{2.2.15}$$

where $F_{\odot}$ represents the solar flux density at the top of the atmosphere when the instantaneous distance between the earth and sun is r, and θ_0 denotes the solar zenith angle. The definition of the solar constant is $S = F_{\odot}(r/r_0)^2$. Thus, we have

$$F(t) = S(r_0/r)^2 \cos\theta_0, \tag{2.2.16}$$

where S is the solar constant corresponding to the mean earth–sun distance r_0 defined previously. If we define the solar heating received at the top of the atmosphere per unit area as Q, then the solar flux density may be written as

$$F = \frac{dQ}{dt}. \tag{2.2.17}$$

Thus, the insolation for a given period of time is

$$Q = \int_t F(t)\, dt. \tag{2.2.18}$$

The total solar energy received by a unit of area per one day may be calculated by integrating total insolation over the daylight hours. Upon substituting Eq. (2.2.16) into (2.2.18) and noting that variation of the distance r in one day can be neglected, we can define the daily insolation as follows:

$$Q \cong S\left(\frac{r_0}{r}\right)^2 \int_{\text{sunrise}}^{\text{sunset}} \cos\theta_0(t)\, dt. \tag{2.2.19}$$

Inserting Eq. (2.2.1) into Eq. (2.2.19) and denoting the angular velocity of the earth ω by dh/dt ($= 2\pi$ rad/day), we obtain

$$Q = S\left(\frac{r_0}{r}\right)^2 \int_{-H}^{H} (\sin\varphi \sin\delta + \cos\varphi \cos\delta \cos h)\frac{dh}{\omega}, \tag{2.2.20}$$

where H represents a half-day, i.e., from sunrise or sunset to solar noon. After performing this simple integration, daily solar insolation is given by

$$Q \cong \frac{S}{\pi}\left(\frac{a}{r}\right)^2 (\sin\varphi \sin\delta H + \cos\varphi \cos\delta \sin H), \tag{2.2.21}$$

where we have set $r_0 = a$ as shown in Eq. (2.2.7), and H in the first term on the right-hand side is expressed in units of radians ($180^\circ = \pi$ rad). Note that the factor $(a/r)^2$ never departs from unity by more than 3.5%. It ranges from 1.0344 on January 3 to 0.9674 on July 5. The computations of (a/r) and δ have been discussed previously.

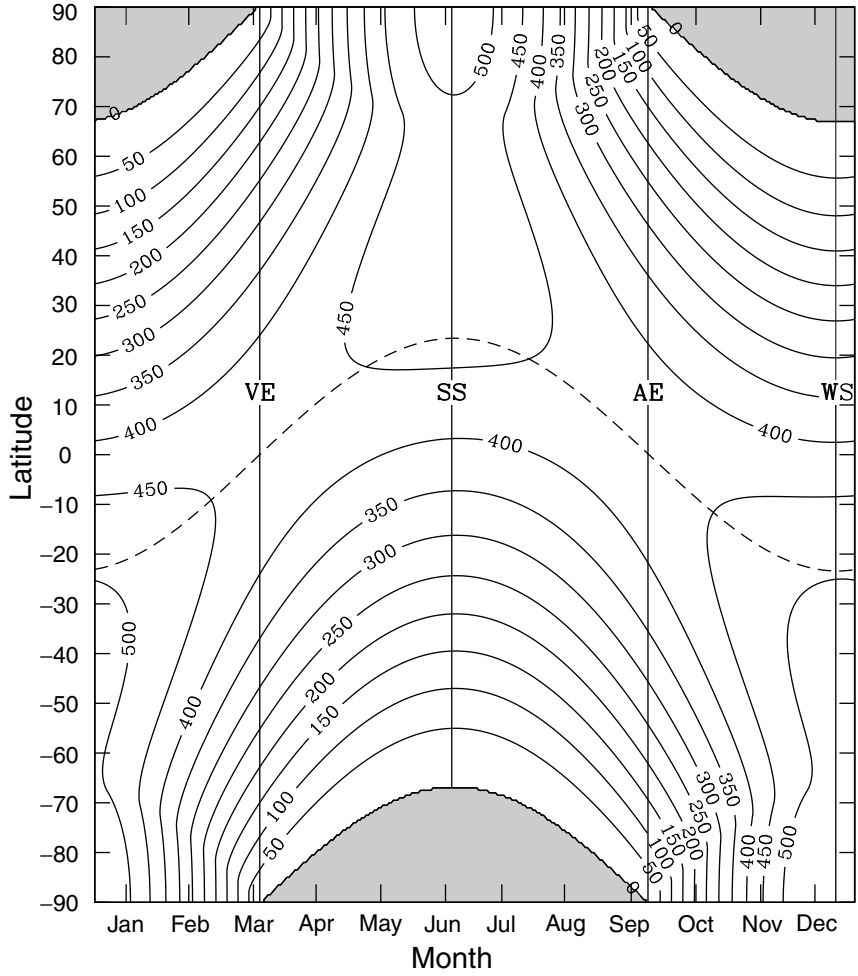

Figure 2.8 Daily mean solar insolation ($Q/24$ hr) as a function of latitude and day of year in units of W m^{-2} based on a solar constant of 1366 W m^{-2}. The shaded areas denote zero insolation. The position of vernal equinox (VE), summer solstice (SS), autumnal equinox (AE), and winter solstice (WS) are indicated with solid vertical lines. Solar declination is shown with a dashed line.

Daily solar insolation is shown in Fig. 2.8. The distribution of solar insolation is independent of longitude and is slightly asymmetric between the northern and southern hemispheres. The sun is closest to the earth in January (winter in the northern hemisphere), so that the maximum solar insolation received in the southern hemisphere is greater than that received in the northern hemisphere. At the equinoxes, solar insolation is at a maximum at the equator and is zero at the poles. At the summer solstice of the northern hemisphere, daily insolation reaches a maximum at the North Pole because of the 24-hour-long solar day. At the winter solstice, the

sun does not rise above the horizon north of about 66.5°, where solar insolation is zero.

The calculation of seasonal and annual insolations is rather involved and will not be detailed here. However, in the domain of the true longitude of the earth as shown in Fig. 2.5, the year can be divided into astronomical spring $(0 - \pi/2)$, summer $(\pi/2 - \pi)$, autumn $(\pi - 3\pi/2)$, and winter $(3\pi/2 - 2\pi)$. Seasonal solar insolation can be evaluated using these divisions. Because of the cosine property of the true longitude of the earth, λ, solar insolation is the same for spring and summer, and for autumn and winter. Consequently, it suffices to consider the summer half-year (spring plus summer) and winter half-year (autumn plus winter) for the calculation of seasonal solar insolation. Considering the total solar insolation for the winter $(\pi, 2\pi)$ half-year and expressing this insolation in terms of that for the summer $(0, \pi)$ half-year, and performing a lengthy algebraic analysis, we find

$$Q_{s,w} = \frac{S\widetilde{T}}{2\pi(1 - e^2)^{1/2}}[\widetilde{S}(\varphi, \varepsilon) \pm \sin\varphi \sin\varepsilon], \tag{2.2.22}$$

where the insolation function is defined by

$$\widetilde{S}(\varphi, \varepsilon) = \frac{\sin\varphi \sin\varepsilon}{2\pi} \int_0^{2\pi} (H - \tan H) \sin\lambda \, d\lambda, \tag{2.2.23}$$

and the positive and negative signs are for the summer and winter solar insolations, respectively (see Exercise 2.15 for the reason for this difference). The half-day defined in Eq. (2.2.2) is given by

$$\cos H = -\frac{\tan\varphi \sin\varepsilon \sin\lambda}{(1 - \sin^2\varepsilon \sin^2\lambda)^{1/2}}, \tag{2.2.24}$$

where the definition of the declination angle defined in Eq. (2.2.4) is used. The annual insolation for a given latitude is the sum of the summer and winter insolations and is given by

$$Q_a = \frac{S\widetilde{T}\widetilde{S}(\varphi, \varepsilon)}{\pi(1 - e^2)^{1/2}}. \tag{2.2.25}$$

Because the insolation function is the same for the northern and southern hemispheres, i.e., $\widetilde{S}(\varphi, \varepsilon) = \widetilde{S}(-\varphi, -\varepsilon)$, annual solar insolation is the same for corresponding latitudes in each hemisphere.

Finally, annual global solar insolation can be evaluated by using the instantaneous solar insolation for the entire earth, which is given by $S(a/r)^2\pi a_e^2$. Distributing this energy over the surface area of the earth, $4\pi a_e^2$, the mean solar insolation for one day is given by $\Delta t_\odot (a/r)^2 S/4$. We can perform an integration over a year via Kepler's second law to obtain

$$Q_{at} = \int_0^{\widetilde{T}} \frac{S\Delta t_\odot}{4} \left(\frac{a}{r}\right)^2 \frac{dt}{\Delta t_\odot} = \frac{S}{4}\widetilde{T}(1 - e^2)^{-1/2} \cong \frac{S}{4}\widetilde{T}(1 + e^2/2). \tag{2.2.26}$$

The annual global insolation is proportional to $(1 + e^2/2)$, but is independent of the declination of the sun δ and the true anomaly ν.

2.3 Solar Spectrum and Solar Constant Determination

2.3.1 Solar Spectrum

The solar spectrum covers wavelengths ranging from gamma rays to radio waves, as shown in Fig. 1.1. Because of the nonquantized electronic transitions, most solar energy is carried by the continuum, i.e., radiation is continuous rather than selective. The single most important contributor is hydrogen, both in its neutral state and as negative ions. A radiation transition from one level to another is characterized by an absorption or an emission line whose frequency is governed by Planck's relation. However, in the ionization process the atom (or molecules) may absorb more than the minimum energy required to remove the electron. This additional energy may be thought of as supplying kinetic energy to the freed electron and is not quantized. As a consequence, absorption is not selective but rather continuous. The ionization continuum occurs on the high-frequency (shorter wavelength) side of the ionization frequency. Neutral hydrogen has ionization continua associated with lines, some of which were defined in Fig. 1.9. Metallic atoms also contribute to the continuum in the ultraviolet spectrum. The continuum absorption in the visible and infrared spectrum, however, is produced by negative hydrogen ions.

Electromagnetic radiation emerging from within the sun is continuously emitted and absorbed by atoms. As shown in Fig. 2.2, the radiative temperature first drops off to a minimum value of about 4500 K just above the photosphere, and then levels off and slowly rises in the chromosphere, followed by a rapid rise in the transition region to several million degrees in the corona. At each temperature, probabilities of the electronic transition exist that any atom will achieve a particular excited state, leading to the formation of absorption lines at different levels in the solar atmosphere. The core of a line forms at the temperature where the maximum transition probabilities of an electron moving from one orbital level to another occur (see Fig. 1.8). The wings of a line form at different temperature levels because of the required transition probabilities. Each absorption line has a preferred formation region in the solar atmosphere. Those lines that absorb very little radiation are known as weak lines, which can form in narrow layers of the solar atmosphere. Some of the absorption lines in the solar atmosphere were displayed in Fig. 2.2.

In view of the preceding discussion, the solar spectrum consists of a continuous emission with a superimposed line structure. The visible and infrared spectrum of the photosphere shows absorption lines, known as the *Fraunhofer spectrum*. The strongest of these lines are produced by H, Mg, Fe, Ca, and Si, as well as singly ionized Ca and Mg. Most of the lines shorter than 1850 Å produced from the photosphere exhibit in emission. Light from the chromosphere and the corona has emission lines at all observed wavelengths.

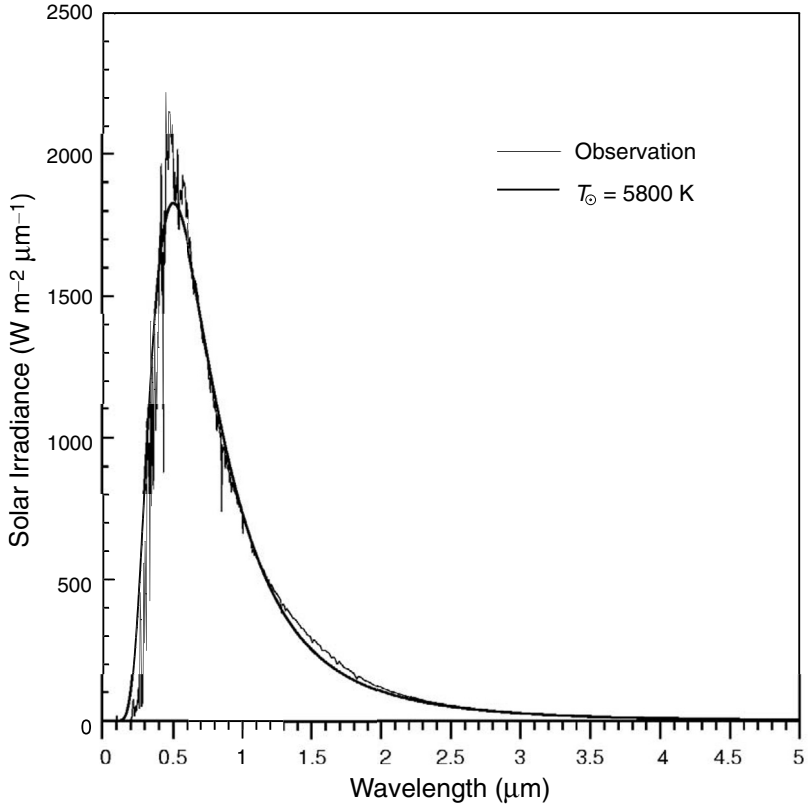

Figure 2.9 Solar irradiance for a 50 cm^{-1} spectral interval at the top of the atmosphere based on the results presented in the MODTRAN 3.7 program. Also shown is the Planck flux with a temperature of 5800 K accounting for the mean distance between the earth and the sun.

Figure 2.9 shows the spectral solar irradiance observation at the top of the atmosphere averaged over a 50 cm^{-1} spectral interval as a function of wavelength up to 5 μm, based on the results presented in the MODTRAN 3.7 program (Anderson *et al.,* 1995). Although the total solar irradiance derived from this program is 1373 W m^{-2}, the spectral solar irradiance curve presented here is scaled with respect to the recently proposed solar constant of 1366 W m^{-2} (see Section 2.3.3 for further discussion). A 50 cm^{-1} spectral average has been performed to smooth out the rapid fluctuations produced by the absorption/emission line structure. However, some variabilities can still be seen, particularly in the ultraviolet spectrum. Also shown is the Planck curve with an emitting temperature of 5800 K, taking into account the mean distance between the sun and the earth. This temperature appears to fit closely with the visible and infrared spectrum characteristic of radiation emitted from the photosphere. For atmospheric applications, it is critically important to have reliable spectral solar irradiances for use in radiative transfer models. Table 2.3 gives tabulated data from 0.2 to 5 μm with

Table 2.3

Distribution of Solar Spectral Irradiance S_λ from 0.2 to 100 μm in Terms of the Accumulated Energy and Percentage Based on the Values Listed in the MODTRAN 3.7 Program[a]

λ (μm)	S_λ (W m^{-2} μm^{-1})	$S_{0-\lambda}$ (W m^{-2})	$S_{0-\lambda}$ (%)	λ (μm)	S_λ (W m^{-2} μm^{-1})	$S_{0-\lambda}$ (W m^{-2})	$S_{0-\lambda}$ (%)
0.20	2.0832E+01	2.08317E+00	0.15250	3.8	1.0564E+01	1.35239E+03	99.00390
0.30	5.4765E+02	5.68479E+01	4.16163	3.9	9.6162E+00	1.35335E+03	99.07430
0.40	1.4042E+03	1.97272E+02	14.44155	4.0	8.6980E+00	1.35422E+03	99.13797
0.50	1.9619E+03	3.93464E+02	28.80410	4.1	7.9180E+00	1.35502E+03	99.19593
0.60	1.7632E+03	5.69780E+02	41.71153	4.2	7.2072E+00	1.35574E+03	99.24870
0.70	1.4300E+03	7.12778E+02	52.17994	4.3	6.5062E+00	1.35639E+03	99.29633
0.80	1.1257E+03	8.25347E+02	60.42075	4.4	5.7954E+00	1.35697E+03	99.33875
0.90	8.8835E+02	9.14182E+02	66.92404	4.5	5.2622E+00	1.35749E+03	99.37727
1.00	7.2943E+02	9.87125E+02	72.26392	4.6	4.8180E+00	1.35798E+03	99.41255
1.10	5.8743E+02	1.04587E+03	76.56425	4.7	4.4724E+00	1.35842E+03	99.44529
1.20	4.8921E+02	1.09479E+03	80.14558	4.8	4.1565E+00	1.35884E+03	99.47573
1.30	4.0851E+02	1.13564E+03	83.13614	4.9	3.8504E+00	1.35922E+03	99.50391
1.40	3.4450E+02	1.17009E+03	85.65813	5.0	3.5740E+00	1.35958E+03	99.53008
1.50	2.9066E+02	1.19916E+03	87.78592	6.0	1.8385E+00	1.36303E+03	99.78240
1.60	2.4644E+02	1.22380E+03	89.58999	7.0	1.0108E+00	1.36404E+03	99.85639
1.70	2.0453E+02	1.24425E+03	91.08726	8.0	5.9672E−01	1.36464E+03	99.90007
1.80	1.6829E+02	1.26108E+03	92.31927	9.0	3.7458E−01	1.36501E+03	99.92751
1.90	1.3725E+02	1.27481E+03	93.32404	10.0	2.4702E−01	1.36526E+03	99.94559
2.00	1.1624E+02	1.28643E+03	94.17501	11.0	1.6932E−01	1.36543E+03	99.95798
2.10	9.7416E+01	1.29617E+03	94.88816	12.0	1.2005E−01	1.36555E+03	99.96677
2.20	8.2132E+01	1.30439E+03	95.48942	13.0	8.7276E−02	1.36563E+03	99.97315
2.30	6.9594E+01	1.31134E+03	95.99889	14.0	6.5062E−02	1.36570E+03	99.97792
2.40	5.9198E+01	1.31726E+03	96.43226	15.0	4.9463E−02	1.36575E+03	99.98154
2.50	5.1023E+01	1.32237E+03	96.80577	16.0	3.8307E−02	1.36579E+03	99.98434
2.60	4.4280E+01	1.32679E+03	97.12994	17.0	3.0112E−02	1.36582E+03	99.98655
2.70	3.8672E+01	1.33066E+03	97.41305	18.0	2.3991E−02	1.36584E+03	99.98831
2.80	3.3815E+01	1.33404E+03	97.66058	19.0	1.9351E−02	1.36586E+03	99.98973
2.90	2.9589E+01	1.33700E+03	97.87720	20.0	1.5797E−02	1.36588E+03	99.99088
3.00	2.6133E+01	1.33962E+03	98.06850	30.0	3.4388E−03	1.36598E+03	99.99860
3.10	2.3093E+01	1.34193E+03	98.23756	40.0	1.0465E−03	1.36599E+03	99.99937
3.20	2.0476E+01	1.34397E+03	98.38746	50.0	4.2098E−04	1.36600E+03	99.99968
3.30	1.8186E+01	1.34579E+03	98.52059	60.0	2.0151E−04	1.36600E+03	99.99983
3.40	1.6191E+01	1.34741E+03	98.63913	70.0	1.0860E−04	1.36600E+03	99.99991
3.50	1.4562E+01	1.34887E+03	98.74574	80.0	6.3779E−05	1.36600E+03	99.99995
3.60	1.3032E+01	1.35017E+03	98.84114	90.0	3.9985E−05	1.36600E+03	99.99998
3.70	1.1670E+01	1.35134E+03	98.92657	100.0	2.6459E−05	1.36600E+03	100.0000

[a] The solar constant is taken to be 1366 W m^{-2}.

a 0.1-μm spectral interval. From 5 to 100 μm, solar irradiance accounts for about 6 W m^{-2}. Based on these values, about 50% of the total solar irradiance lies in wavelengths longer than the visible, about 40% in the visible region, and about 10% in wavelengths shorter than the visible. Note that from 3.5 to 5 μm, the emitted thermal infrared radiation from the earth and the atmosphere system becomes significant.

According to solar flux observations, the ultraviolet region ($<0.4\,\mu$m) of the solar spectrum deviates greatly from the visible and infrared regions in terms of

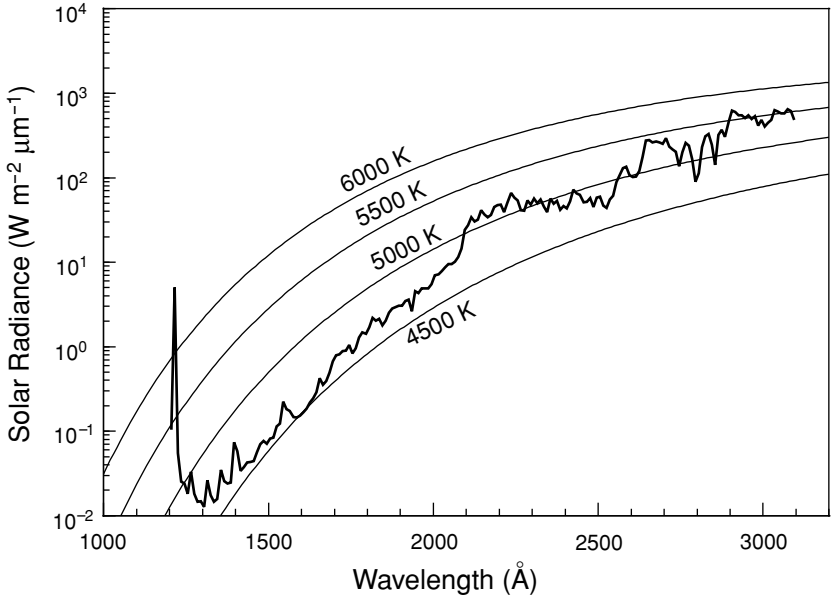

Figure 2.10 Observed irradiance outside the earth's atmosphere in the ultraviolet region (data taken from Brasseur and Simon, 1981) and comparison with the Planck curves for temperatures ranging from 4500 K to 6000 K.

the equivalent blackbody temperature of the sun. Figure 2.10 illustrates a detailed observed solar spectrum from about 1000 to 3000 Å, along with blackbody temperatures of 4500, 5000, 5500, and 6000 K. In the interval 2100–3000 Å, the equivalent blackbody temperature of the sun lies somewhat above 5000 K. It falls gradually to a minimum level of about 4700 K at about 1400 Å. From there toward shorter wavelengths, a larger amount of energy flux is observed at the Lyman α emission line of 1216 Å associated with the transition of the first excited and ground states of hydrogen atoms. The ultraviolet portion of the solar spectrum below 3000 Å contains a relatively small amount of energy. However, because the ozone and the molecular and atomic oxygen and nitrogen in the upper atmosphere absorb all this energy, it represents the prime source of the energy in the atmosphere above 10 km.

2.3.2 Determination of the Solar Constant: Ground-Based Method

For historical reasons, we shall first introduce the ground-based method for the determination of the solar constant. Ground-based observations of solar irradiance for the purpose of determining the solar constant require three primary instruments. These are the *pyrheliometer,* the *pyranometer,* and the *spectrobolometer*. The pyrheliometer is used to measure the direct, plus some diffuse, solar radiation, while the pyranometer, utilizing a suitable shield to block the direct solar radiation from striking the instrument, measures only the diffuse solar radiation for arriving at a pyrheliometer

correction. The amount of direct sunlight can then be calculated by subtracting the flux density measured by the pyranometer from that measured by the pyrheliometer. The spectrobolometer is a combination of a spectrograph and a coelostat. A coelostat is a mirror that follows the sun and focuses its rays continuously on the entrance slit of the spectrograph, which disperses the solar radiation into different wavelengths by means of a prism or diffraction grating. In the Smithsonian solar constant measurements, about 40 standard wavelengths between 0.34 and 2.5 μm are measured nearly simultaneously from the record of the spectrograph. The instrument for these measurements is called a *bologram*. There are two techniques of measuring the solar constant from the ground-based radiometer, called the *long* and *short* methods of the Smithsonian Institution. The long method is more fundamental and establishes the basis for the short method. The long method uses the Beer–Bouguer–Lambert law and is introduced in the following.

Consider an atmosphere consisting of plane-parallel layers. At a given position of the sun, which is denoted by the solar zenith angle θ_0, the effective path length of the air mass is $u \sec\theta_0$, where

$$u = \int_{z_1}^{z_\infty} \rho dz. \tag{2.3.1}$$

In this equation z_1 is the height of the station and z_∞ denotes the top of the atmosphere. On the basis of the Beer–Bouguer–Lambert law, the irradiance F of the direct solar radiation of wavelength λ observed at the surface level is given by

$$F_\lambda = F_{\lambda 0} \exp(-k_\lambda u \sec\theta_0) = F_{\lambda 0} T_\lambda^m, \tag{2.3.2}$$

where $F_{\lambda 0}$ is the monochromatic solar irradiance at the top of the atmosphere, k_λ denotes the monochromatic mass extinction cross section, T_λ is the monochromatic transmissivity defined in Eq. (1.4.10), and m ($= \sec\theta_0$) represents the ratio of the air mass between the sun and the observer to the air mass with respect to the local zenith distance. Upon taking the logarithm, we find

$$\ln F_\lambda = \ln F_{\lambda 0} + m \ln T_\lambda. \tag{2.3.3}$$

Observations of F_λ may be made for several zenith angles during a single day. If the atmospheric properties do not change during the observation period, then the transmissivity T_λ is constant. A plot of $\ln F_\lambda$ versus m shown in Fig. 2.11 may be extrapolated to the zero point, which represents the top of the atmosphere ($m = 0$). This is referred to as the *Langley plot*. If observations of the monochromatic irradiance are carried out for wavelengths covering the entire solar spectrum, then we have

$$F_\odot = \int_0^\infty F_{\lambda 0} d\lambda \approx \sum_{i=1}^{N} F_{\lambda_i 0} \Delta\lambda_i, \tag{2.3.4}$$

where N is the total number of the monochromatic irradiances measured. The irradiance $F_\odot$ corresponds to the actual distance between the earth and the sun, r. By using

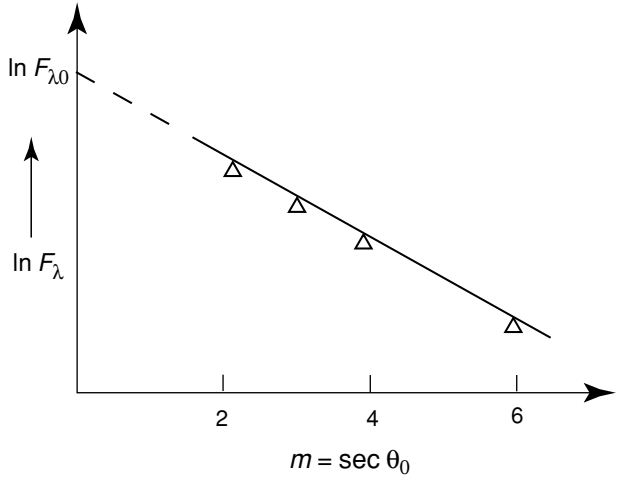

Figure 2.11 Hypothetical observed monochromatic solar irradiances F_λ as a function of the effective path length from which the solar irradiance at the top of the atmosphere can be graphically determined, referred to as the Langley plot.

the mean distance $r_0 = a$, the solar constant is defined by

$$S = F_\odot (r/a)^2. \tag{2.3.5}$$

The foregoing outlines the theoretical procedures of the Smithsonian long method for the determination of the solar constant. However, the atmosphere is essentially opaque for wavelengths shorter and longer than about 0.34 μm and 2.5 μm, respectively. Consequently, flux density observations cannot be made in these regions. Therefore, empirical corrections are needed for the omitted ranges, which account for about 8% of the solar flux.

There are other sources of error inherent in the Smithsonian long method caused by (1) empirical corrections for the absorption of ultraviolet by ozone, and the absorption of infrared by water vapor and carbon dioxide in the wings of the solar spectrum; (2) an unknown amount of diffuse radiation entering the aperture of the observing instrument; (3) variations of k_λ and the possible effects of aerosols during a series of measurements; and (4) measurement errors. Therefore, in spite of careful evaluation and observation, a certain amount of error is inevitable.

Employing the Smithsonian long method, each determination requires about 2 to 3 hours of observation time, plus twice that much time for the data reduction. In addition, there is no assurance that atmospheric properties and solar conditions will remain unchanged during the observation period. Because of this uncertainty and the burdensome, time-consuming work involved, a short method was devised to determine the solar constant.

In the short method, the diffuse component of solar radiation (the sky brightness) is measured for a given locality over a long period of time, so that a mean diffuse intensity

can be determined. Thus, a pyranometer reading of diffuse solar radiation will differ from the mean by an amount ε, called the *pyranometer excess*. In reference to Section 1.1.4 and Fig. 3.9, the attenuation of solar radiation on a clear day is due to scattering by molecules and aerosol particles, and absorption by various gases, primarily water vapor. If total precipitable water is given by w, an empirical relationship between the attenuation of direct solar irradiance and scattering and absorption effects may be expressed in the form $F_\lambda = w + q_\lambda \varepsilon$, where q_λ is a constant empirically determined for each wavelength for a given locality. With q_λ known, the spectral value of the solar irradiance can be found from the observed precipitable water and a pyranometer reading.

On the basis of a long series of previous observations of F_λ, m, and T_λ at a given location where the solar constant measurement has been made, a graph of F_λ versus air mass m can be constructed for a set value of T_λ. Thus, for a particular measurement of F_λ with a known air mass m, the corresponding transmissivity T_λ can be found from the graph. Once T_λ has been determined, solar irradiance at the top of the atmosphere $F_{0\lambda}$ can be evaluated through Eq. (2.3.2). After this point, evaluation of the solar constant proceeds in the same manner as in the long method. In the short method, the required measurements include a bologram of the sun, an observation of sky brightness by the pyranometer, and air mass determined by the position of the sun from a theodolite. These three measurements take only about 10 to 15 minutes. From the thousands of observations at various locations around the world during a period of more than half a century, the best value of the solar constant determined by the Smithsonian methods is 1353 W m^{-2}.

The presence of aerosols in the atmosphere imposes limitations on the accuracy of ground-based radiometric measurements of the solar constant (see, e.g., Reagan *et al.,* 1986). To minimize atmospheric effects, a number of measurements have also been made in the upper atmosphere and outer space. These have included observations made from balloons floating in the 27- to 35-km altitude range, jet aircraft at about 12 km, the X-15 rocket aircraft at 82 km, and the Mars Mariner VI and VII spacecrafts entirely outside the atmosphere. The solar constant derived from these experiments varies. Based on a series of measurements from high-altitude platforms, a standard solar constant of 1353 ($\pm$21) W m^{-2} was issued in 1976 by the National Aeronautics and Space Administration (Thekaekara, 1976).

2.3.3 Satellite Measurements of the Solar Constant

Measurements of incoming solar irradiance have been routinely made from satellite platforms since the mid-1970s. However, the high-accuracy, high-stability satellite-borne radiometer was only developed and incorporated in the Nimbus 7 satellite in 1978. This radiometer was an electrically calibrated cavity radiometer. The basic concept of blackbody cavity radiation was shown in Fig. 1.6. Radiometers of this design for use in satellites had a black painted cavity that absorbed nearly all the solar radiation impinging on it. The absorbed radiation raised the temperature of the cavity so that a radiant power could be measured corresponding to the increase in temperature. Such a cavity can also be heated by an electrical element in a manner equivalent to the incident sunlight. Because the input electrical power can be measured

accurately, the temperature response of the radiometer can be calibrated, and it is thus referred to as the *self-calibrating radiometer*.

Solar constant data have been derived from total solar irradiance measurements made by self-calibrating radiometers aboard a number of satellites since 1978. These include the Nimbus 7 Earth Radiation Budget (ERB) mission in 1978; the Solar Maximum Mission (SMM) Active Cavity Radiometer Irradiance Monitor 1 (ACRIM I) in 1980; the Earth Radiation Budget Experiment (ERBE) on board the NASA Earth Radiation Budget Satellite (ERBS, 1984), NOAA 9 (1984), and NOAA 10 (1986); and the ACRIM II measurements on board the Upper Atmosphere Research Satellite (UARS, 1991). Figure 2.12 shows the daily measurements of the solar constant from these satellites from 1979 to 1996. The solid lines are 81-day running means of the daily data. The absolute radiance scale of the ACRIM II data has been adjusted to

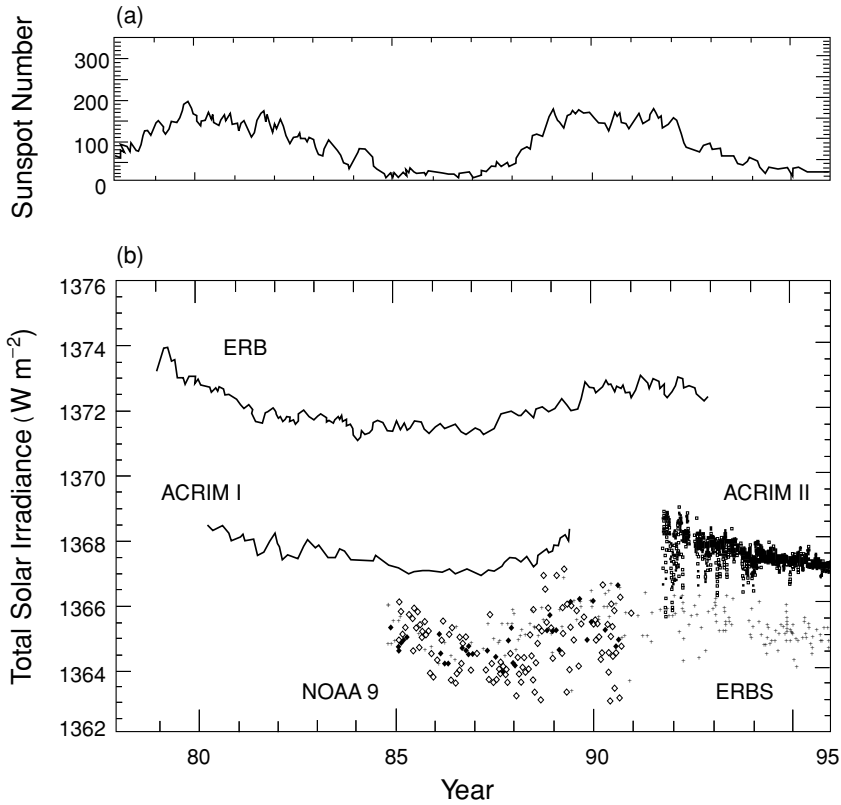

Figure 2.12 Solar activity variations from 1978 to 1996 illustrated by (a) the sunspot number and (b) changes in total solar irradiance. The results were obtained from the ERB radiometer on the Nimbus-7 satellite, ACRIM I on the Solar Maximum Mission (SMM) satellite, ACRIM II on the UARS, and the ERBE program (NOAA-9 and ERBS). The solid lines are 81-day running means of the daily data. Total solar irradiance increases during times of maximum solar activity relative to its levels in the intervening activity minimum. The differences in absolute irradiance levels among various measurements are of instrumental origin (data taken from Lean and Rind, 1998).

match the results of ACRIM I. The differences in absolute irradiance levels among various measurements depicted in this figure are attributed to the instrument sensitivity changes related to temperature or aspect drifts. In particular, ERB and ERBS data differed by about 10 watts per square meter. The top panel shows the sunspot numbers over the same period. It is quite clear that the data displayed in Fig. 2.12 provide irrefutable evidence of the 11-year solar constant cycle. When solar activity is high, as indicated by the sunspot number, the total and UV radiative outputs from the sun increase. Dark sunspots on the solar disk reduce total radiative output because their emission is less than that of the surrounding disk. However, after sunspots develop, magnetic regions involving faculae and plages where emission is enhanced also increase. These regions are evident as complexes of bright emission. The sun's irradiance fluctuates because radiation sources are not homogeneously distributed on its disk. Magnetic fields erupting from the solar convection zone (Section 2.1) into the overlying solar atmosphere generate active regions and complexes in which the local radiation is altered relative to the background solar disk. Magnetic activity erupts, evolves, and decays at different rates throughout the 11-year cycle, generating sunspots, plages, and faculae that modulate total and spectral solar radiative outputs. Finally, it should be noted that our knowledge of the 11-year irradiance cycle is imperfect because of uncertainties arising from the limited duration of space-borne solar monitoring that barely exceeds one 11-year cycle, as well as instrumental uncertainties that cause variable signals in individual satellite solar radiometers.

A number of analyses of the mean total solar irradiance have been reported. Based on the analysis of the solar irradiance measurements taken by the cavity sensor in a number of the satellites depicted in Fig. 2.12, a mean value for the solar constant of 1366 W m^{-2} with a measurement uncertainty of ± 3 W m^{-2} has been suggested (Lean and Rind, 1998). The solar constant value is critical in the interpretation of measured solar absorption and heating rates in the atmosphere.

Exercises

2.1 Compute the solar elevation angle at solar noon at the poles, 60° N(S), 30° N(S), and the equator. Also compute the length of the day (in terms of hours) at the equator and at 45° N at the equinox and solstice.

2.2 From the geometry of an ellipse and the equation defining it, derive Kepler's first law denoted in Eq. (2.2.5).

2.3 Based on the conservation of angular momentum that the radius vector drawn from the sun to the planet sweeps out equal areas in equal time, derive Kepler's second law denoted in Eq. (2.2.6).

2.4 (a) Derive Kepler's third law by equating Newton's law of universal gravitation and the centrifugal force required to keep the planet in a circular orbit. (b) Given that the NOAA polar satellites orbit at about 850 km above the earth's surface, what would be the period of these satellites? (c) Geostationary satellites have the same angular velocity as the earth. What would be the required height for these satellites?

2.5 Given the solar constant of 1366 $\mathrm{W\,m^{-2}}$, the mean earth–sun distance of 150×10^6 km, and the sun's radius of 0.70×10^6 km, calculate the equilibrium temperature of the sun.

2.6 If the average output of the sun is 6.2×10^7 $\mathrm{W\,m^{-2}}$, and the radius of the earth is 6.37×10^3 km, what is the total amount of energy intercepted by the earth in one day?

2.7 Compute the fraction of the sun's emittance intercepted by the earth.

2.8 Consider a circular cloud whose diameter is 2 km and assume that it is an infinitely thin blackbody with a temperature of 10°C. How much energy does it emit toward the earth? How much energy from this cloud is detected on a square centimeter of the earth's surface when the center of the cloud is 1 km directly over the receiving surface?

2.9 Assume that $\bar{r}$ is the mean albedo of the earth (*albedo* is defined as the ratio of the amount of flux reflected to space to the incoming solar flux), and that the earth–atmosphere system is in equilibrium. Show that the equilibrium temperature of the earth–atmosphere system $T_e = [(1 - \bar{r})S/4\sigma]^{1/4}$.

2.10 The following table gives the distances of various planets from the sun and their albedos. Employing the result in Exercise 2.9, compute the equilibrium temperatures of these planets.

Planet	Distance from sun (relative to earth)	Albedo (%)
Mercury	0.39	6
Venus	0.72	78
Earth	1.00	30
Mars	1.52	17
Jupiter	5.20	45

2.11 The height of earth-synchronous (geostationary) orbiting satellites, such as GOES satellites, is about 35,000 km. Using the solid angle derived in Exercise 1.2, calculate the equilibrium temperature of such a satellite in the earth–satellite system, assuming an effective equilibrium temperature of 255 K for the earth and assuming that the satellite is a blackbody.

2.12 Show that the change in the earth's equilibrium temperature T_e in terms of the earth–sun distance r is given by $\delta T_e / T_e = \delta r / 2r$. The distance between the earth and the sun varies by about 3.3% with a maximum and minimum on January 3 and July 5, respectively. Compute the seasonal change in the earth's equilibrium temperature.

2.13 Calculate the daily insolation at the top of the atmosphere at (a) the south pole at the winter solstice; and (b) the equator at the vernal equinox. Use the mean earth–sun distance in your calculations and check your values with those shown in Fig. 2.8.

2.14 Prove that annual insolation is the same for corresponding latitudes in the two hemispheres [you may use the results in Eqs. (2.2.23) and (2.2.24) for analysis].

2.15 Show that the difference between the length of summer and that of winter is given by $\widetilde{T}4e\sin\omega/\pi$. In carrying out this exercise, first define the length using the astronomical season definition and then utilize Kepler's expressions by approximation.

2.16 Reproduce the daily solar insolation graph presented in Fig. 2.8 using Eqs. (2.2.21), (2.2.9), and (2.2.10).

2.17 Compute and plot the solar irradiance at the top of the earth's atmosphere emitted from temperatures of 5000, 5500, and 6000 K. Compare your results with those presented in Figs. 2.9 and 2.10.

2.18 On a clear day, measurements of the direct solar flux density F at the earth's surface in the 1.5- to 1.6-μm wavelength interval give the following values:

Zenith angle (degree):	40°	50°	60°	70°
F (W m^{-2}):	13.95	12.55	10.46	7.67

Find the solar flux density at the top of the atmosphere and the transmissivity of the atmosphere for normal incidence [see Eq. (1.4.10)] in this wavelength interval.

Suggested Reading

Berger, A. L. (1988). Milankovich theory and climate. *Rev. Geophys*. **26,** 624–657. This paper offers an authoritative overview of contemporary theory of the earth's orbit about the sun and its impact on climate.

Coulson, K. L. (1975). *Solar and Terrestrial Radiation*. Academic Press, New York. Chapters 3 and 4 give a comprehensive illustration of various kinds of pyrheliometers and pyranometers.

Hoyt, D. V., and Schatten, K. H. (1997). *The Role of the Sun in Climate Change*. Oxford University Press, New York. Chapters 2 and 3 contain a readable discussion of the composition of the sun and the solar constant from a historical perspective.

Jastrow, R., and Thompson, M. H. (1984). *Astronomy: Fundamentals and Frontiers,* 2nd ed. Wiley, New York. Chapter 12 provides an in-depth discussion and delightful photos of the structure and composition of the sun.

Lean, J., and Rind, D. (1998). Climate forcing by changing solar radiation. *J. Climate* **11,** 3069–3094. This paper includes a comprehensive and authoritative review of the solar constant and solar spectrum measurements and variabilities.