

1

Overview

This chapter has several goals. The first is to provide some general historical background on how Mars was perceived before spacecraft exploration started with the launch of Mariner 4 in 1964. The second goal is to provide an overview of what conditions are like on Mars today. Most of the book concerns the record of past events as preserved in the landscape and in the rocks at the surface. Although conditions may have been different in the past, those that prevail today provide strong constraints on how we interpret that past record. A third purpose is to give a brief overview of topics that are important for our understanding of the planet, but which are a little off the main theme of the book, which is to describe the major geological features of the planet and their origin. A brief description of the present atmosphere is included here, for example. We also include a section on martian meteorites. These are both huge topics with a vast literature, and no attempt is made in the book to treat them comprehensively. A fourth aim of the chapter is to provide a short geological overview so that the subsequent, more detailed chapters can be read in light of a general knowledge of the planet's geology. Most geological topics are just touched upon here and referenced to later chapters.

Telescopic observations

Mars is the fourth planet from the Sun. With a mean radius of 3389.5 km, it is intermediate in size between the Earth (6378 km) and the Moon (1738 km). As Earth and Mars move in their orbits around the Sun, telescopic viewing conditions change. When Earth and Mars are on opposite sides of the Solar System, they are close to 400 million km apart and Mars subtends an angle of only 3.5 arcsec. At closest approach (opposition) the distance between the two planets may be as small as 55 million km, and the planet subtends an angle of 25 arcsec. Telescopic viewing is thus best at opposition when features as small as 150 km across can be distinguished with the best ground-based telescopes. Oppositions are spaced roughly 780 days apart. The exact spacing varies because Mars' orbit, unlike the Earth's, is distinctly eccentric. This results in an orbital velocity that

changes according to where the planet is in its orbit. The spacing between oppositions therefore changes according to where Mars is at opposition. Eccentricity also affects the quality of the oppositions, the best being when Mars is at perihelion. This has caused a bias of telescopic observations toward the southern hemisphere, since at perihelion Mars' southern hemisphere is tilted toward the Sun and hence toward Earth.

No topography can be seen from Earth-based telescopes (Figure 1.1). What are seen are variations in the reflectivity (albedo) of the surface, including the polar caps, and changes in the opacity of the atmosphere. Although the surface markings may change in detail from opposition to opposition or over decadal time scales, the gross pattern has remained constant for the entire period of telescopic observation. The most prominent features outside the poles are dark markings in the 0–40°S latitude belt, although the most prominent dark feature on the planet, Syrtis Major, is outside this belt. The dark areas were originally thought to be seas and so were called maria. They are mostly areas that have been swept partly clean of the bright dust that covers much of the surface. Most dark markings do not correspond to topographic features, although some do. Some bright markings, such as Hellas and Nix Olympica, noted on some early maps, also correspond to topographic features, probably because of persistent clouds in these areas. The most famous features of the planet from the telescopic era, the canals, were portrayed on almost all twentieth-century maps until the mid 1960s. They are largely illusions on the part of observers straining to see markings at or below the limits of telescopic observation.

Transient brightenings of part, or all, of the telescopic image of the planet were correctly attributed to clouds, of which two types were identified: yellow clouds interpreted as dust storms, and white clouds interpreted as condensate clouds. The yellow dust clouds were observed to occur mostly in the southern hemisphere in southern spring and summer. In some years such as 1956 and 1971 the dust storms became truly global. The classical markings disappeared and

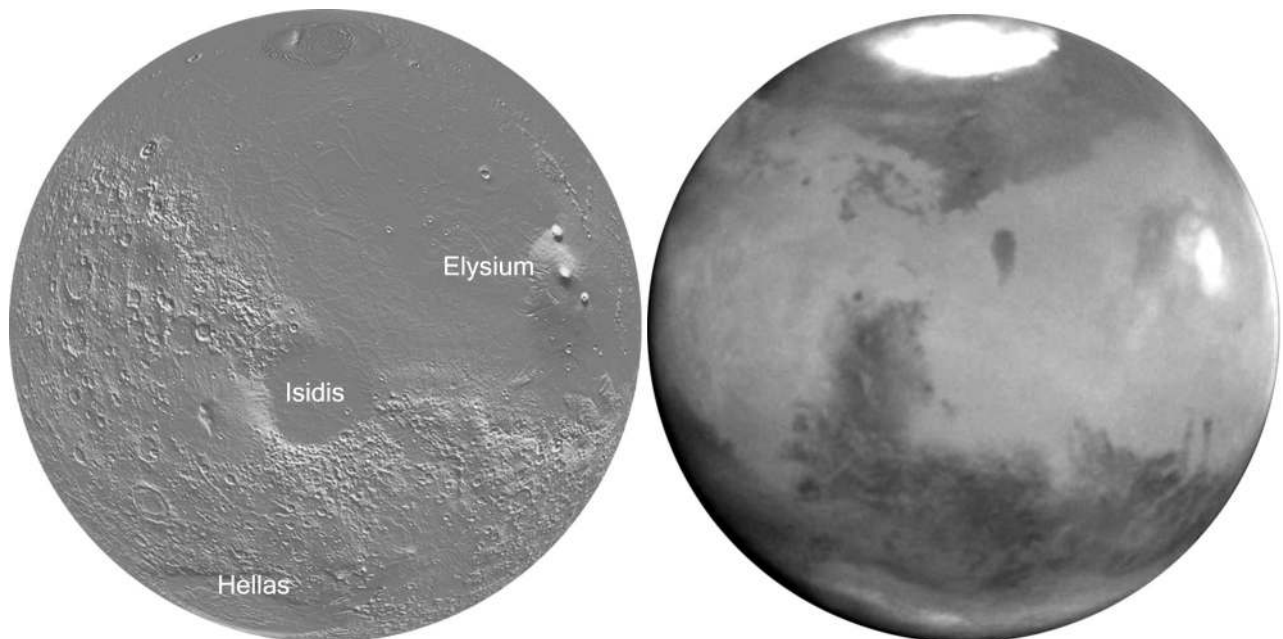


Figure 1.1. Comparison of martian terrain with what is seen in a telescopic image. The view on the left is a MOLA reconstruction of martian terrain from the same perspective as the telescopic image on the right. The MOLA image is of surface relief. The telescopic image was taken by Hubble Space Telescope in late northern spring when the planet was 103×10^6 km from Earth. It shows variations in the reflectivity of the ground and the atmosphere. Bright clouds are present in Hellas and Elysium, and the seasonal CO_2 cap in the north has almost completely dissipated. The dark and light surface markings are only poorly correlated with the relief.

became re-established only after several months. White clouds were observed at places, such as Olympus Mons, Alba, Patera, and Sinai Planum, where we now know from spacecraft observation that water ice clouds are common. Two other phenomena deserve mention. The first is the wave of darkening, a progressive darkening of the dark markings that proceeded from pole to equator as the polar caps receded. It was variously interpreted as growing vegetation, release of water from the receding cap, and sweeping of dust from the area around the receding cap by strong off-pole winds. The second phenomenon was the appearance of blue clearings, times when the dark markings appear particularly crisp and clear. Neither phenomenon has been confirmed by spacecraft observations. For a comprehensive summary of Mars as viewed from the telescope see Martin *et al.* (1992).

Orbital and rotational motions

The orbital and rotational motions of the planet (Figure 1.2) affect how much insolation falls on the planet and how the amount changes with time of year and latitude. The motions, therefore, affect surface temperatures, atmospheric circulation, and climatic

conditions in general. Variations in obliquity (the angle between the spin axis and the orbit normal) are particularly important for Mars since the changes are large and can cause significant variations in atmospheric pressure and transfer of water between the poles and lower latitudes

The Mars day is 24 hr 39.6 min and the year is 687 Earth days or 669 Mars days (sols). Instead of months, the areocentric longitude of the Sun (L_s) is used to denote time of year. This is the equivalent of the Sun-centered angle between the position of Mars in its orbit and the position of the northern spring equinox. At the start of northern spring $L_s = 0^\circ$, at northern summer solstice $L_s = 90^\circ$ and so on. Mars' rotation axis is tilted 25° with respect to the orbit plane so that the planet has seasons like the Earth. The orbit of Mars is, however, distinctly elliptical (eccentricity of 0.093), in contrast to the near-circular orbit of the Earth (eccentricity of 0.017), and this affects the length and intensity of the seasons. At closest approach to the Sun (perihelion), the Mars–Sun distance is 1.381 AU (One Astronomical Unit or AU is the mean Earth–Sun distance or 149.5×10^6 km.) At its furthest distance from the Sun aphelion, the Mars–Sun distance is 1.666 AU. Since the solar flux varies with the

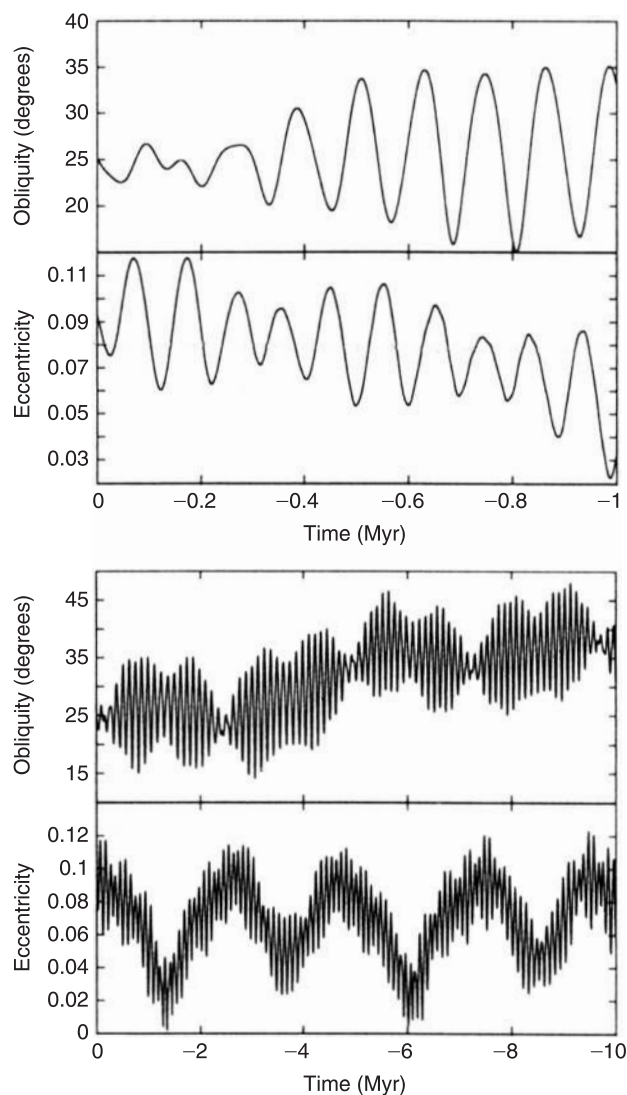


Figure 1.3. Changes in eccentricity and obliquity projected back in time. The upper two panels show projections back 1 Myr. The lower two panels show projections back 10 Myr. Modulation of the 1.2×10^5 yr obliquity cycle has been modest for the last 0.4 Myr. From 0.4 to 4 Myr ago, obliquities ranged from 15° to 35° about a mean of 25° . From 4 to 10 Myr ago the mean was close to 35° . Prior to 10 Myr obliquities are chaotic and cannot be definitively predicted. Obliquity may affect a wide range of phenomena such as atmospheric pressure, the stability of water ice at the surface and in the ground, and the incidence of dust storms. The effects are most marked at the poles (Chapter 10). (From Laskar *et al.*, 2002, copyright © Nature Publishing Group.)

Precession causes a slow rotation of the line of equinoxes – the intersection of the equatorial plane and the orbit plane – with a 175,000 yr period and a rotation of the line of apsides – the line joining perihelion and aphelion – with a period of 72,000 yrs.

The net result is that the longitude of perihelion, the angle between equinox and perihelion passage, changes with a period of 51,000 yrs. Thus, while today perihelion occurs in southern spring causing southern springs and summers to be short and hot, 25,000 yrs from now it will be the northern springs and summers that are short and hot.

Changes in obliquity are likely to have much larger climatic effects than changes in eccentricity and the timing of perihelion. The present obliquity is 25.19° but it undergoes large changes (Figure 1.3). During the current epoch, it is thought to oscillate between 15° and 35° , about a mean of 24° (Laskar *et al.*, 2004). The oscillations have a period of 1.2×10^5 yrs with an amplitude that is modulated on a 2 Myr cycle. Variations in obliquity have a particularly strong effect in the polar regions. At obliquities higher than 54° the average solar flux is higher at the poles than at the equator. Moreover, during polar summers at high obliquities, the pole is constantly illuminated, leading to high sublimation rates of any ice that may be present and deep penetration of a large annual thermal wave. During these periods water ice may be driven from the poles and accumulate at low to mid latitudes (Chapters 8 and 10).

There are considerable uncertainties as to what past obliquities were (Ward, 1992; Laskar and Robutel, 1993; Touma and Wisdom, 1993; Laskar *et al.*, 2004). Minute differences in the starting values for the calculations of past motions lead to large differences in the solutions when projected backward (or forward) in time, such that projections larger than 10 Myr are uncertain. Part of the problem concerns resonances. If the period of precession of the spin is commensurate with one of the periods of variation of the orbit, then spin-orbit resonances can occur. Excursions in obliquity significantly larger than are suspected from the current oscillations are then possible. These variations cause the obliquity to be chaotic, at least on time scales greater than 10 Myr. Laskar *et al.* (2004) ran a large number of simulations in order to estimate the distribution of obliquities over geological time. They found that the average obliquity is close to 40° , that there is a 63 percent probability of reaching 60° in the next 1 Gyr and >5 percent probability of exceeding 70° in 3 Gyr. In this respect Mars differs from the other terrestrial planets. The obliquities of Mercury and Venus have been stabilized by dissipation of solar tides, and that of the Earth by the presence of the Moon. Although the obliquity variations are chaotic on time scales longer than 10 Myr, calculations on the time scale of fewer than 10 Myr are reproducible (Laskar *et al.*, 2004).

They indicate that obliquities were significantly higher prior to 3 Myr ago. Between 3 Myr and 10 Myr ago they oscillated between 25° and 46°, instead of the present 15–35°. Possible geological and climatic effects of the obliquity cycle are discussed in Chapters 8 and 10.

Global structure and topography

Mars, like the Earth, is differentiated into a crust, mantle, and core (Chapter 4). Because we have no seismic data, the size of the core is poorly defined but the radius is estimated to be between 1300 and 1500 km. From the partitioning and depletion of core-forming elements in the mantle, as indicated by the composition of martian meteorites, the core appears to be more sulfur-rich than the Earth's (Treiman *et al.*, 1986; Wänke and Dreibus, 1988). Present-day Mars has no magnetic field so that the core is probably solid, but large remanent crustal magnetic anomalies indicate that the core was molten early in the planet's history (Acuna *et al.*, 1999). From relations between gravity and topography, the crust is estimated to range in thickness from 5 to 100 km, with a thicker crust in the southern hemisphere than in the north (Chapter 4). The crust is basaltic in composition. No crust analogous to terrestrial "granitic" continental crust has been detected. Two crustal compositions have been identified from Thermal Emission Spectrometer (TES) data (Bandfield, 2002). At low latitudes (<30°), where not dust covered, the surface has a basaltic spectrum. Higher latitudes have a different spectrum that was initially interpreted as that of basaltic andesite. Subsequently, Wyatt *et al.* (2004) suggested that the spectrum was more likely that of weathered basalt, the weathering having preferentially occurred at high latitudes because ice is stable at these latitudes.

Although Mars has only 28 percent of the surface area of the Earth, it has much larger variations in surface relief. The range is 29.429 km, from –8.200 km in the floor of Hellas to 21.229 km at the summit of Olympus Mons (Figures 1.4, 1.5). Since Mars has no sea level, elevations have to be referenced to some artificial datum. During the Mariner 9 mission it was decided to use the elevation at which the atmospheric pressure is 6.1 mbar, the triple point of water. This surface was approximated by a triaxial ellipsoid with radii defined by occultation measurements and shape defined by a fourth-order representation of the gravity field (Wu, 1978). With the acquisition of direct measurements of radii by the Mars Orbiter Laser Altimeter (MOLA) (Smith *et al.*, 2001) and a vastly improved gravity field from Mars

Global Surveyor (Lemoine *et al.*, 2001), the elevations are now referenced to an equipotential surface whose average radius at the equator is 3396 km. The precision of the elevation measurements is close to 1 m.

A fundamental feature of Mars' topography is the so-called global dichotomy. Much of the northern hemisphere is at elevations well below the reference surface, while much of the southern hemisphere stands above the reference. One result is a south polar radius of 3382.5 km as compared with 3376.2 for the north, a 6.3 km difference. Other results are a center of mass/center of figure offset of 2.99 km, and a distinctly bimodal distribution of elevations with maxima at 1.5 km above the mean and 4 km below the mean (Smith *et al.*, 2001). The north–south dichotomy is also expressed as differences in crater density and differences in the thickness of the crust (Chapter 4).

Other than the dichotomy the largest positive topographic feature on the planet is the Tharsis bulge, 10 km high and 5000 km across centered on the equator at 265°E. The bulge formed very early in the history of the planet and has been a focus of volcanic activity ever since. A much smaller bulge is centered in Elysium at 25°N, 147°E. The largest negative topographic feature is the impact basin Hellas at 47°S, 67°E, which has a floor that is mostly 9 km below the rim. The rim itself forms a broad annulus around the basin that includes most of the highest terrain in the eastern part of the southern hemisphere. A second large impact basin, Argyre, at 50°S, 318°E is much shallower, with a floor that is mostly only 1–2 km below the datum.

Atmosphere

The Mars atmosphere is thin and composed largely of CO₂ (Tables 1.2 and 1.3). At the Viking 1 landing site, at an elevation of –2 km, the pressure ranged from 6.9 to 9 mbar (Figure 1.6), less than one-hundredth of the atmospheric pressure at sea level on Earth. The pressure varies with the seasons as up to 25 percent of the atmosphere condenses on the winter pole. The cycle is dominated by growth and dissipation of the south polar cap, it being more extensive than that in the north because of the longer, colder winters in the south. The pressure cycle is highly repeatable from year to year. The scale height of the atmosphere (the height over which the pressure declines to 1/e or 0.3678 of its original value) is roughly 10 km. This implies that for a pressure of 7.5 mbar at the –2 km elevation of the Viking 1 landing site, the pressure at the surface ranges from 0.7 mbar at the summit of Olympus Mons to 14 mbar in the deepest part of Hellas.

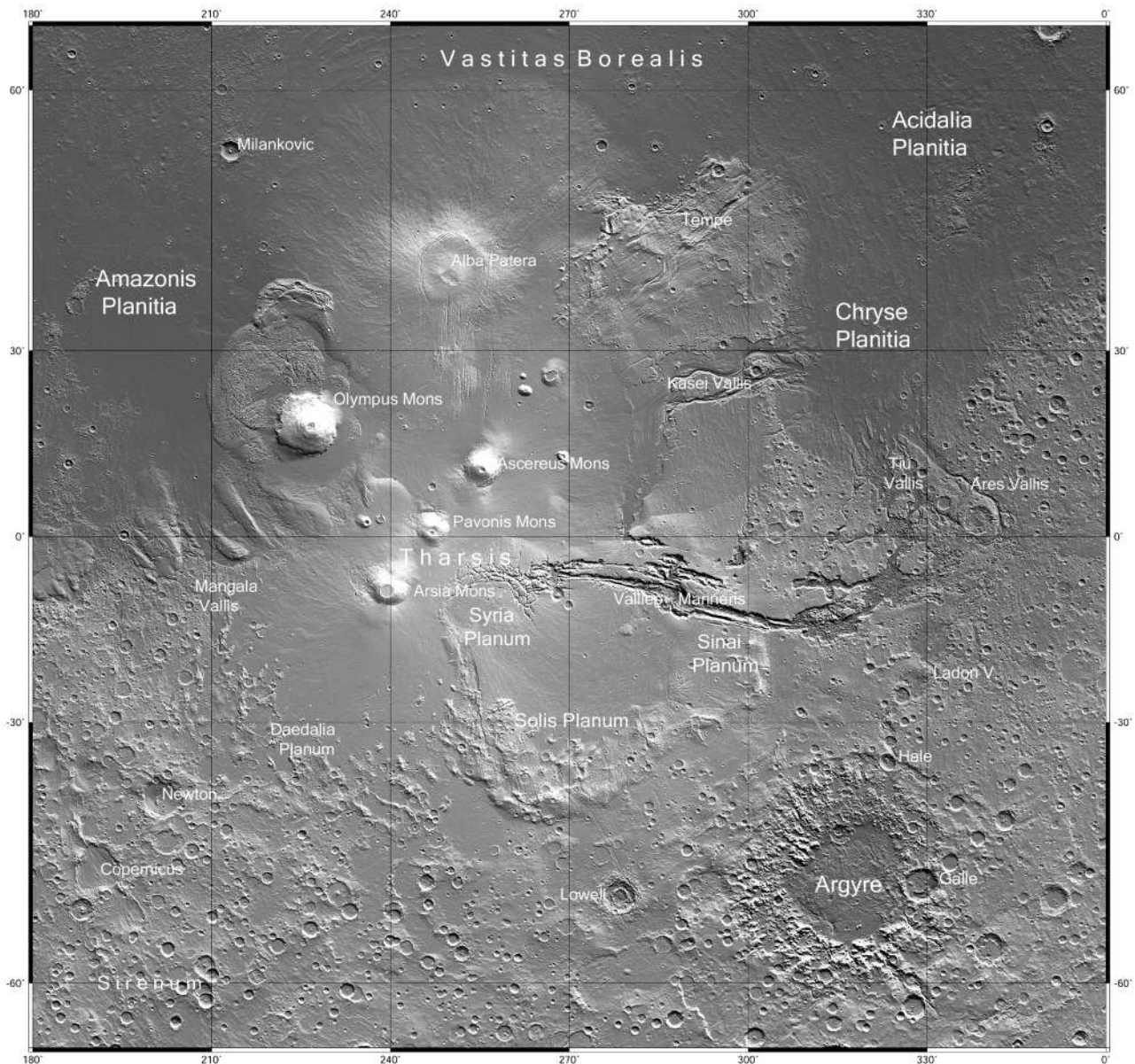


Figure 1.4. The western hemisphere of Mars. Higher areas have lighter tones. The hemisphere is dominated by the Tharsis bulge, centered on the equator at 260° E. The bulge straddles the boundary between the low-lying, sparsely cratered plains to the north and the high-standing cratered terrain to the south (MOLA).

The vertical temperature profile of the martian atmosphere is quite different from the Earth's, partly because of the lack of ozone to create a stable stratosphere, and partly because of the effects of dust (Zurek, 1992). In the lowest part of the Earth's atmosphere, the troposphere, temperatures decline with elevation and are controlled largely by radiative and conductive heat exchange with the surface, and release of latent heat from the condensation of water vapor. In the stratosphere, above the

tropopause, temperatures increase with elevation as a consequence of adsorption of ultraviolet radiation by ozone. Because of the reversed temperature gradient, there is very little vertical mixing. Above the Earth's stratosphere, in the mesosphere, the temperature gradient reverses again and temperatures decline with elevation, being controlled by radiative emission and absorption by CO₂. Finally, at the mesopause, the temperatures again start to increase with elevation, as heating is by conduction from above

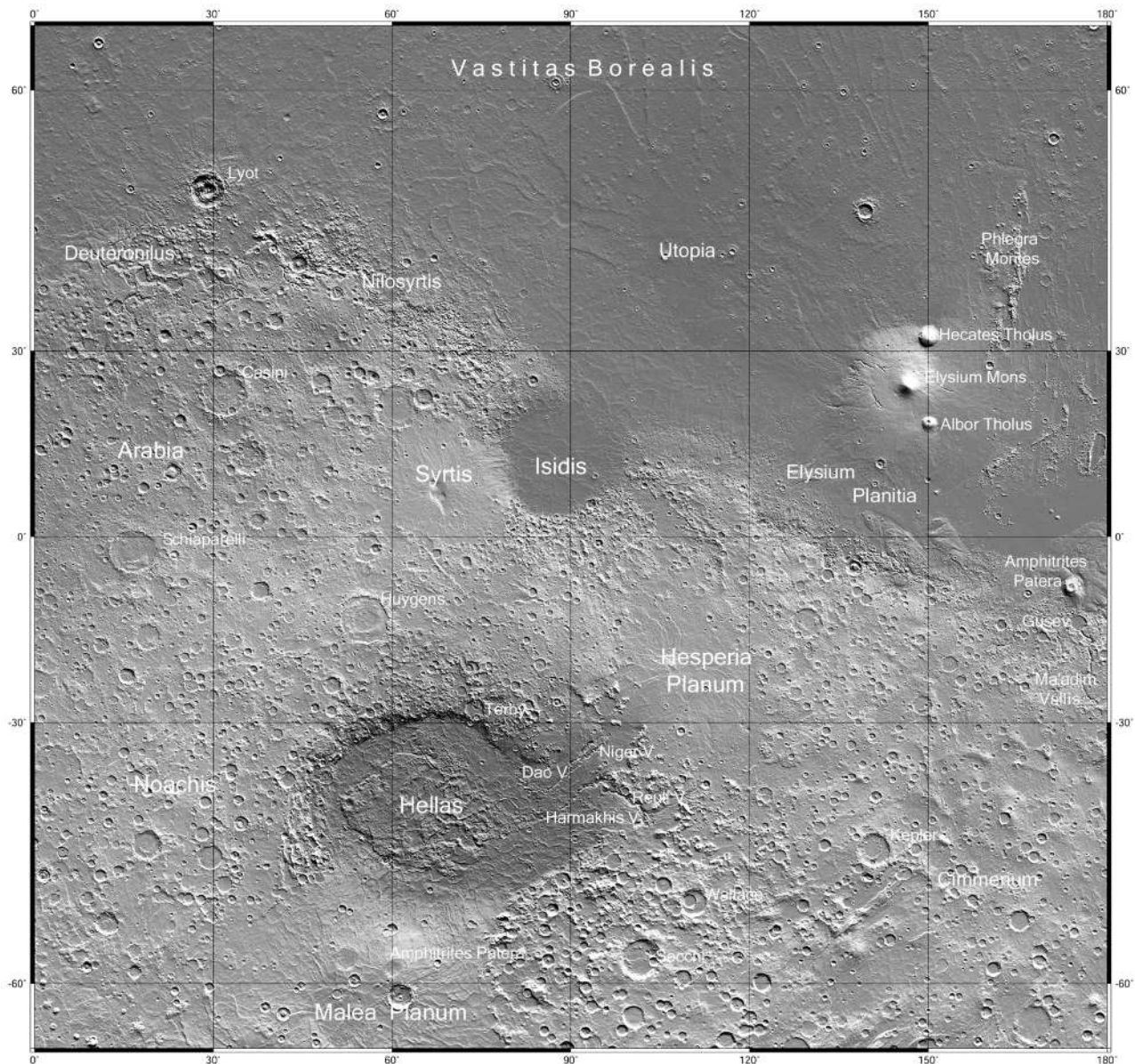


Figure 1.5. The eastern hemisphere of Mars. The most prominent features of this hemisphere are the two impact basins Hellas and Isidis and the volcanic province of Elysium. The north–south dichotomy boundary is more obvious in this hemisphere than in the west where it is buried by Tharsis (MOLA).

where the Sun's extreme ultraviolet radiation is absorbed.

When clear, temperatures in the lower part of the martian atmosphere decline with elevation as in the Earth's troposphere, but there is no gradient reversal due to ozone as in the Earth's atmosphere. Up to an elevation of about 45 km, temperatures are controlled largely by exchange of heat with the ground. With the small amounts of water present, latent heating is negligible. From 45 to 110 km elevations, temperatures continue to fall but radiative emission and absorption

by CO₂ dominate. Above the mesopause, at 110 km, the temperature gradient reverses, like on Earth, as the effects of absorption of extreme ultraviolet radiation high in the atmosphere becomes important. Just above the mesopause, at 125 km, is the homopause, above which atmospheric gases begin to separate diffusively. At still higher elevations is the exosphere where the atmosphere is so thin that atoms and molecules are on ballistic trajectories and can escape. Diffusive separation of atmospheric components above the homopause and escape from the exosphere

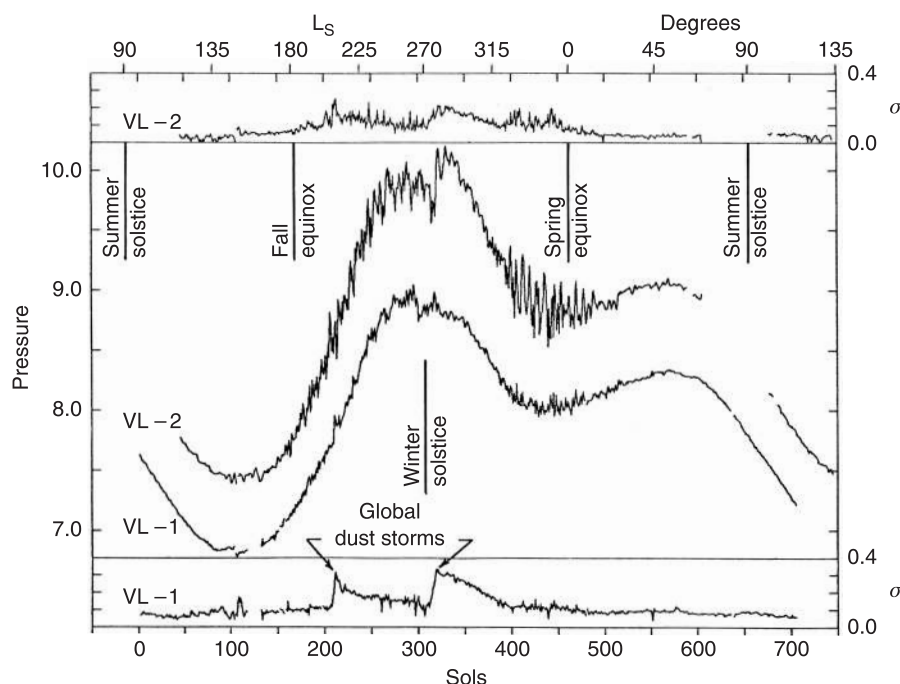


Figure 1.6. Variation in surface pressure at the two Viking landing sites. The origin is at the northern summer solstice when the southern seasonal cap is approaching its maximum extent, and the atmospheric pressure is near its lowest. The second minimum, at the end of northern winter, is much shallower than the first because of the smaller seasonal cap in the north. The upper and lower panels show the standard deviation of the pressure within a sol. These increase substantially when the atmosphere is dusty because of heating throughout the atmosphere rather than from the surface (Hess *et al.*, 1980, copyright 1980, American Geophysical Union, reproduced by permission of the American Geophysical Union).

have resulted in a substantial enrichment of the atmosphere in deuterium and other, heavier, isotopes.

The conditions just described are for a clear atmosphere. The martian atmosphere holds at least a small fraction of dust at all times. Because the dust directly absorbs the Sun's radiation, heat transfer from the surface becomes less important in controlling the temperature. The more dust that is present, the more isothermal is the vertical profile. Diurnal temperature variations at the surface are also suppressed.

The circulation of the atmosphere has several components (Zurek *et al.*, 1992). A north–south (meridional) flow results from seasonal exchange of CO₂ between the two poles, as 10–15 percent of the atmosphere condenses on the northern polar cap in northern winter and 25 percent on the southern polar cap in southern winter. Heating of the atmosphere at low to mid latitudes in the summer hemisphere causes air to rise there, promoting a seasonal meridional overturning (Hadley cell) that extends across the equator, thereby facilitating exchange of water vapor between the two hemispheres. The north–south flow and strong latitudinal thermal gradients in the winter hemisphere cause instabilities (eddies) at mid latitudes.

As a result, eastward-propagating planetary waves develop, accompanied by strong westerly winds, high-altitude jet streams and traveling storm systems. While the Viking landers were on the surface, the storms passed regularly on a roughly 3-day cycle. Modeling suggest that the northern storms are more intense than those in the southern mid latitudes. Other elements of the circulation include low-latitude, westward-propagating thermal tides, driven by the diurnal heating cycle, and quasi-stationary waves caused by the large-scale topography and large-scale variations in albedo.

At the Viking landing sites, outside the dust storm season, winds were typically a few meters per second with daily maxima of 8–10 m s⁻¹. During the dust storm period, and at the times of the winter storms in the north, where the landing sites were, winds at the Viking sites were in excess of 10 m s⁻¹ 10 percent of the time and gusts reached almost 40 m s⁻¹. The wind sensor was 1.6 m above the ground. At this elevation, winds of 20–60 m s⁻¹ are needed to cause saltation of surface grains, the exact value depending on the size of the grains and the surface roughness (Greeley *et al.*, 1992). Dust is commonly

raised by dust devils which have left criss-crossing tracks in many areas of the planet (Chapter 9).

When the atmosphere is clear, it is heated mostly from below, with the result that a convective boundary layer expands to a few kilometers thick during the day then collapses at night. Most of the daily temperature variations damp out within 2 km of elevation above the surface. Because the atmosphere is so dry, latent heat effects are negligible and the vertical profile is close to the dry adiabat. The thermal stability of the base of the atmosphere at night, caused by the extremely cold surface temperatures, decouples the atmosphere from surface friction and a strong nocturnal jet may develop, particularly at those places and at those times when the general circulation has a strong north–south component.

During southern spring and summer, dust storms tend to start at low latitudes wherever there are large slopes and/or large gradients in surface albedo or thermal inertia. They also may start along the edge of the seasonal caps. Areas where dust storms have historically been initiated are the northwest rim of Hellas, the Claritas Fossae region, and low-lying parts of Isidis Planitia (Gierasch, 1974). Dust storms may be local, or they may grow to global proportions as they did in 1971, at the start of the Mariner 9 mission and in 1977 during the Viking mission. Global dust storms can spread to encompass almost the whole planet in several weeks. The storms are most common in southern spring and summer, close to perihelion, when summer temperatures are at their highest. During the 1977 dust storm the optical depth at the Viking 1 landing site, far from the initiation site of the dust storm, rose from a value of 0.5 just before the storm to a high of 5 during the peak of the storm. The total amount of dust elevated into the atmosphere in the global storms is small, equivalent to a few micrometers spread over the whole planet (Kahn *et al.*, 1992). At present, the storms appear to be causing a net transfer of dust from the southern hemisphere to bright low-thermal inertia regions at low northern latitudes.

Surface temperatures

Because the atmosphere is dry and thin, it has a low heat capacity and absorbs little of the Sun's incoming short-wave radiation or the outgoing long-wave radiation, at least when it is clear. As a result, for most of the year, surface temperatures have a wide range and are close to those expected from a simple balance between the solar radiation adsorbed at the surface, the emitted infrared radiation, and heat conducted into or out of the ground (Figure 1.7).

Mean diurnal temperatures range from close to 150 K (−123°C) at the poles to 240 K (−33°C) at the warmest locations at mid summer in the southern hemisphere (Kieffer *et al.*, 1977). Daily maxima can reach 300 K (27°C) during summer at mid southern latitudes, but this is somewhat deceiving since the high temperatures are reached only where the thermal inertia (see below) is low. At such locations, the daily fluctuations damp out rapidly with depth (Figure 1.8) so that the above freezing temperatures are reached only within the upper centimeter of the soil.

Surface temperatures depend on latitude and season, and on the albedo and thermal inertia of the surface. They are also affected by the slope of the ground and the configuration of the surrounding terrain. The radiometric albedo is the fraction of the total incident solar radiation not adsorbed by the surface. Albedos of unfrosted ground range from 0.095 to 0.415, with preferred values at 0.135 and 0.275 (Kieffer *et al.*, 1977). The thermal inertia (I) is a measure of the responsiveness of a material to changes in the thermal regime. It is defined as $(K\rho c)^{1/2}$ where K is the thermal conductivity, ρ is the density, and c is the specific heat of the material. Because the density and specific heat of rock materials do not vary greatly, most of the variations in thermal inertia are caused by variations in the thermal conductivity. Solid rocks have high thermal inertias; loose granular materials with abundant void spaces have low thermal inertia, conduction of heat through them being largely restricted to the contact points between grains. Thermal inertias have historically been measured in units of $\text{cal cm}^{-2} \text{s}^{-1/2}$, but are normally referred to in units (sometimes informally referred to as Kieffers) where the actual values have been multiplied by 10^3 . One thermal inertia unit = $10^{-3} \text{cal cm}^{-2} \text{s}^{1/2} = 41.84 \text{J m}^{-2} \text{s}^{-1/2}$. Thermal inertias of the martian surface range from 1 ($41.84 \text{J m}^{-2} \text{s}^{-1/2}$) to 15 ($627.6 \text{J m}^{-2} \text{s}^{-1/2}$) (Kieffer *et al.*, 1977; Palluconi and Kieffer, 1981). Bare rocks typically have thermal inertias in excess of 30 ($1255 \text{J m}^{-2} \text{s}^{-1/2}$), so all the martian surface is at least partly covered with loose materials. The thermal inertias of the martian surface cluster around two values, 6 and 2.5, corresponding respectively to the preferred albedo values of 0.135 and 0.275 (Kieffer *et al.*, 1977). Figure 1.7 shows how surface temperatures change during the day as a function of latitude for typical martian values for albedo and thermal inertia.

At high latitudes in winter, surface temperatures are controlled mainly by condensation and sublimation of CO_2 . In the fall, temperatures fall until they reach 150 K, the frost point for CO_2 , at which point CO_2 starts to condense. The 150 K temperatures are

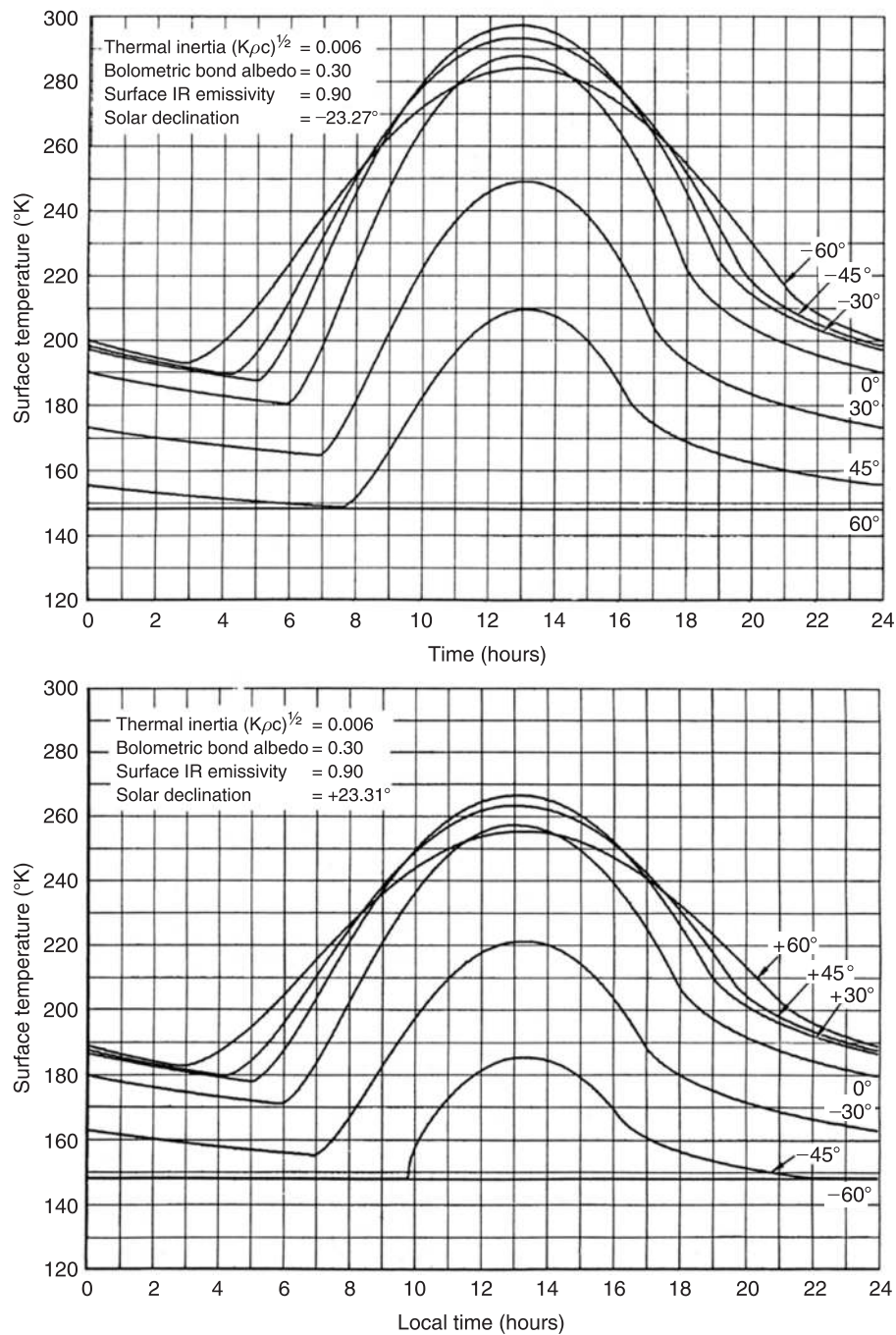


Figure 1.7. Models of the daily temperature fluctuations at the martian surface as a function of latitude. The upper diagram is for perihelion, when it is summer in the south. The lower diagram is for aphelion, when it is summer in the north. Because of the effect of eccentricity, peak summer temperatures in the south are significantly higher than those in the north. Temperatures rise rapidly at dawn, peak just after noon, then decline steadily to their pre-dawn lows (Michaux and Newburn, 1972, NASA/JPL).