17.1 INTRODUCTION

Today Mars is a cold, dry, desert planet. The atmosphere is thin and liquid water is not stable at the surface. But there is evidence that very early in its history it was warmer and wetter. Since Mariner 9 first detected fluvial features on its ancient terrains, researchers have been trying to understand what climatic conditions could have permitted liquid water to flow on the surface. Though the evidence is compelling, the problem is not yet solved.

The main issue is coping with the faint young Sun. During the period when warmer conditions prevailed (~3.7–4.1 Ga), the Sun's luminosity was ~25% less than it is today. How can we explain the presence of liquid water on the surface of Mars under such conditions? A similar problem exists for Earth, which would have frozen over under a faint Sun even though the evidence suggests otherwise.

Attempts to solve the “faint young Sun paradox” rely on greenhouse warming from an atmosphere with a different mass and composition than we see today. This is true for both Mars and Earth. However, it is not a straightforward solution. Any greenhouse theory must (a) produce the warming and rainfall needed, (b) have a plausible source for the gases required, (c) be sustainable, and (d) explain how the atmosphere evolved to its present state. These are challenging requirements, and judging from the literature they have yet to be met.

In this chapter we review the large and growing body of work on the early Mars climate system. We take a holistic approach that involves many disciplines, since our goal is to present an integrated view that touches on each of the requirements listed in the preceding paragraph. We begin with a brief discussion of geological and mineralogical eras in Section 17.2.1, and the major events in early Mars’ history (Section 17.2.2). We then step through the observational evidence of the early climate system, which comes from the geology (Section 17.3.1), mineralogy (Section 17.3.2), and isotopic (Section 17.3.3) data. Each of the datasets presents a consistent picture of a warmer and wetter past with a thicker atmosphere. How much warmer and wetter are matters of debate, but conditions then were certainly different than they are today.

We then discuss the origin and evolution of the early atmosphere (Section 17.4), from accretion and core formation to the end of the late heavy bombardment, including estimates of the volatile inventory, outgassing history, and potential loss mechanisms. This sets the stage for a comprehensive look at the climate system of early Mars (Section 17.5) and the attempts to solve the faint young Sun problem. We take some time to review the basic physics involved and then discuss the different ideas highlighting their strengths and weaknesses. We conclude in Section 17.6 with a summary and a discussion of potentially promising avenues of future research.

17.2 OVERVIEW OF EARLY MARS

17.2.1 Geological and Mineralogical Eras

The geological history of Mars has been grouped into three epochs: Noachian, Hesperian, and Amazonian. These eras are mainly classified on the basis of crater populations linked to a model that converts them to absolute ages (e.g. Hartman and Neukum, 2001). The oldest recognizable terrains, those from the Noachian era, preserve geologic features that date from the period of heavy bombardment, which ended ~3.7 Ga. How far back in time the geologic record extends is unclear because of uncertainties in the early cratering history, but the oldest Noachian surfaces visible in imaging data are probably ~4.1 Ga in age (Werner, 2008). Some workers therefore identify a pre-Noachian period that extends from the time of formation of the crust to ~4.1 Ga (Frey, 2003). Vague circular structures, remnants of large impact basins, probably date from this earliest era. Noachian terrains have crater densities and fluvial features similar to those in the Noachis Terra region of the southern highlands. Hesperian terrains, named after those in Hesperia Planum, are of intermediate age and date from the end of the Noachian to ~2.9–3.3 Ga. Volcanic plains dominate Hesperian terrains, but this does not necessarily imply an increase in volcanic activity, since much of the evidence for earlier volcanism would have been destroyed by the high impact rates (Greeley and Schneid, 1991). Amazonian terrains are the youngest and date from the end of the Hesperian to the present. These terrains were shaped during a period of declining geologic and volcanic activity.

Bibring et al. (2006) suggested a parallel timeline classification based on the detection of phyllosilicates and sulfates by Mars Express. Phyllosilicates require liquid water to form, and sulfates are found in sedimentary environments. Phyllosilicates require alkaline conditions and are present mostly on the oldest terrains. Sulfates, on the other hand, require acidic environments and appear to have formed after the phyllosilicates. However, mapping on smaller spatial scales shows that phyllosilicates

---

1 We use the abbreviations Ma (mega-annum) and Ga (giga-annum) to denote millions and billions of years ago, respectively.
and sulfates are found together in Meridiani Planum (Wiseman et al., 2008; Wray et al., 2009a), Gusev Crater (Wang et al., 2006), Gale Crater (Milliken et al., 2010), Cross Crater (Swayze et al., 2008), Mawrth Vallis (Farrand et al., 2009), and are interbedded in Columbus Crater in Terra Sirenum (Wray et al., 2009b). This may indicate that aqueous environments during the Noachian varied locally with different pH, water/rock ratios, salinity, and temperature. Today, the surface is dominated by anhydrous iron oxides, which can form in the absence of liquid water. Thus, Bibring et al. (2006) suggested three epochs: the Phyllosian, when neutral waters weathered basalts, the Theilkian, when a transition occurred to more acidic environments, and the Siderikian, marked by a final transition to a dry environment. These epochs roughly correspond to the Nouchian, Hesperian, and Amazonian epochs.

### 17.2.2 Timeline of Major Events

Figure 17.1, which is based on material drawn from Fassett and Head (2010), Carr and Head (2010), and Werner and Tanaka (2011), is a summary of the major events that occurred on early Mars. Note that the timing is somewhat uncertain, and the relative chronologies, though plausible, are also subject to revision. The intent here is to provide a context for the discussions in later sections where details and references are provided.

**Mars Accretes (4.57–4.55 Ga).** Mars accreted to its present size within the first 20 Ma of Solar System history, perhaps even sooner. During this period, the planet differentiated into a core, mantle, and crust. During accretion the devolatilization of impactors probably created an atmosphere mostly composed of water (steam). Accretion may have been fast enough to produce a magma ocean. Hydrodynamic escape powered by the high extreme ultraviolet (EUV) fluxes of the young Sun probably resulted in massive loss of volatiles.

**Pre-Noachian (4.55–4.1 Ga).** As the planet cooled, convection within the core generated a dynamo and a magnetic field that held off the solar wind, thereby suppressing the loss of atmospheric constituents by solar wind stripping. The global hemispheric dichotomy was established. Weathering of surface materials may have begun during this period, but the alteration products are poorly constrained. The Tharsis volcanic province began forming and many major impact basins formed. This period ends with the Hellas impact event, which we take to occur at 4.1 Ga, but which may have been as late as 3.9 Ga, depending on crater modeling assumptions. By this time the dynamo ceased, the magnetic field collapsed, and the atmosphere was exposed to solar wind erosion. If there was a period of hydrodynamic escape, it likely ended late in this period.

**Noachian (4.1–3.7 Ga).** This is the period where the ancient geologic record is best preserved. And it is for this period that warm and wet conditions are thought to have prevailed. During this period, the valley networks formed, global-scale layered terrains developed, erosion and cratering rates were high, and oceans, lakes, and/or seas may have been present. The atmosphere at that time may have been thick enough to provide enough greenhouse warming to drive an active hydrological cycle with rainfall and runoff. Impacts may have created episodic warm environments. By the end of this period, surface waters became more acidic, sulfates began forming, the atmosphere thinned, and erosion rates declined. Note that the time evolution of surface minerals is a general trend and that local exceptions can be found (e.g. phyllosilicates above and below sulfates). Also note that valley network formation may have continued into the early Hesperian (Fassett and Head, 2010, 2011).

### 17.3 Evidence for a Different Atmosphere and Climate System

#### 17.3.1 Geological Evidence

The morphology of the Martian surface strongly supports the supposition that surface conditions during the Noachian were, at least episodically, both warmer and wetter than they were.
subsequently. Valley networks and associated features, such as deltas and lakes, common in terrains that survive from this early era, almost certainly resulted from precipitation and the subsequent flow of water across the surface. Erosion rates were much higher during this early era than they were later, and the mineralogical evidence discussed in the next section supports moist surface conditions. Despite these conclusions, considerable uncertainty persists as to how frequent, intense, and sustained the fluvial episodes were, what enabled the fluvial episodes, and whether there was ever a stable, global hydrologic system in which precipitation, runoff, infiltration, and groundwater flow were in quasi-equilibrium with evaporation and sublimation from large bodies of water and ice.

The valley networks have long been viewed as strong evidence of precipitation on early Mars (e.g. Carr, 1981; Baker, 1982) and, as we acquire better imagery and altimetry, the evidence becomes more compelling. The global distribution of valley networks is shown in Figure 17.2. In planimetric form the valleys resemble terrestrial river valleys. They are almost exclusively incised into Noachian or Lower Hesperian terrains, although there are rare younger exceptions (e.g. Mangold et al., 2004). They branch upstream and are typically 1–4 km wide and 50–200 m deep. They tend to have V-shaped cross-sections in their upper reaches and rectangular cross-sections downstream. While a few are over 2000 km long, the vast majority are less than 200 km long and drain into local lows, rather than regional lows such as the northern plains and the Hellas Basin. Drainage densities (stream length per unit area) vary considerably with location and the resolution of the observation from extremely low values up to 1 km\(^{-1}\), which is well within the terrestrial range (Craddock and Howard, 2002; Hynek and Phillips, 2003; Hynek et al., 2010; Ansan and Mangold, 2013). The apparent low drainage densities observed in early imagery, amphitheater heads of tributaries, common in prominent valleys such as Nirgal Vallis, and rectangular cross-sections suggested to many early workers that groundwater sapping had played a major role in the formation of many of the valleys (Pieri, 1980; Carr and Clow, 1981; Baker et al., 1990; Gulick, 1998), although all acknowledged that precipitation and/or hydrothermal circulation were needed to recharge the groundwater system to enable sustained or episodic flow. However, better imaging and altimetry now show that dense, area-filling networks indicative of surface runoff, such as shown in Figure 17.3, are common throughout the Noachian terrains (Hynek et al., 2010). Precipitation followed by surface runoff, coupled with infiltration and groundwater seepage, must have occurred at least episodically in the Noachian (Craddock and Howard, 2002; Irwin and Howard, 2002; Hynek and Phillips, 2003; Stepinski and O’Hara, 2003; Howard et al., 2005; Carr, 2006). Dimensions of channels within some valleys, meander lengths, and boulder sizes suggest peak discharges of \(10^2–10^4\) m\(^3\) s\(^{-1}\) (Moore et al., 2003; Jerolmack et al., 2004; Howard et al., 2005; Irwin et al., 2005), comparable to discharges of typical terrestrial rivers and substantially smaller than the discharges of the large flood features that formed mostly later in Mars’ history.

A distinction must be made between the general, pervasive degradation of the ancient highlands and the formation of most of the valley networks. During the Noachian there was widespread erosion of crater rims and other high ground, and partial infilling of low areas such as craters (Figure 17.4). Several impact craters as large as several hundred kilometers across, with original rim heights of hundreds of meters and depths of a few kilometers, that are superimposed on Hellas ejecta, for example, have almost completely filled and had their rims almost completely eroded away so that they are barely visible. However, the observed valley networks appear to have contributed little to that degradation. They were incised into the degraded landscape, but were not its cause. Poorly dissected
Noachian areas such as northwest Arabia and the region between Hellas and Argyre are just as degraded as the dissected areas. The pristine-appearing valleys incised into the highland terrains appear to result from a late Noachian to early Hesperian episode of intense incision (Howard et al., 2005; Irwin et al., 2005). Degraded, difficult-to-discern valleys throughout the highlands (e.g. Baker and Partridge, 1986) may be traces of fluvial activity prior to the formation of the crisply defined valleys.

The Noachian terrains are clearly much more eroded than younger terrains. While Hesperian craters as small as a few kilometers across generally preserve all their primary impact features, even delicate textures on their ejecta, Noachian impact craters several hundreds of kilometers across mostly have highly eroded rims and partly filled interiors. The contrast implies a precipitous drop in erosion rates at the end of the Noachian. The number of fresh-appearing craters with well-preserved ejecta patterns on Noachian terrains is comparable to the number on the oldest Hesperian terrains, which also suggests that high erosion rates persisted until the end of the Noachian and then rapidly declined (Craddock and Maxwell, 1993). Golombek and Bridges (2000) and Golombek et al. (2006), in summarizing all the data available on Martian erosion rates, concluded that, although Noachian erosion rates were 2–5 orders of magnitude higher than they were subsequently, they still appear to have been lower than average terrestrial rates, being comparable to those of old flat cratons on Earth, or about 5 \( \mu \text{m yr}^{-1} \).

The valleys form an immature drainage system with low basin concavities, undissected interfluves between drainage basins, and poor correlation of basin circularity with elevation within the basins (Stepinski and O’Hara, 2003). Stream profiles are poorly graded and closely follow the regional slopes (Howard et al., 2005). Either erosion by the incised valleys was not sufficiently sustained to allow regional integration of the valleys, or disruption by impacts and other events prevented it. The terminal episode during which most of the valleys formed appears to have ended abruptly because associated fans and deltas are undissected (Figure 17.5). During the late incision phase and during the period of degradation that preceded it, eroded debris was mostly deposited in local lows where layered deposits are common. Large drainage basins comparable to the Amazon and Mississippi that would have resulted in transport of erosional debris to regional lows did not develop. Only modest amounts...
of erosional debris ($\leq 3 \times 10^4$ km$^3$) have, for example, accumulated in Hellas despite a drainage area of $1.5 \times 10^7$ km$^2$, several kilometers of relief, and pervasive erosion of its rim.

Lakes were likely common throughout Noachian terrain while the valleys were forming (Cabrol and Grin, 1999, 2010), as expected of a poorly graded landscape undergoing fluvial erosion. Most valleys terminate in closed local depressions such as craters or inter-craterr lows, where at least transient closed lakes likely formed. Seemingly fine-grained, horizontally layered, easily erodible sediments underlie most of these low areas (Figure 17.6). Chlorine-rich deposits found in local lows within the Noachian uplands may be the result of evaporation of lakes (Osterloo et al., 2008). Temporary, late Noachian to early Hesperian lakes are also suggested by the MER findings in Meridiane. Some of the sulfate-rich deposits found there may have been deposited in transient inter-dune lakes, and subsequently altered as a result of oscillations in the local groundwater table (Grotzinger et al., 2005). Intermittent lakes in Gale Crater possibly lasting for thousands to millions of years (Grotzinger et al., 2015) further support the notion that lakes were common in the early Hesperian. Almost 300 lows have been identified within the highlands that have both inlet and outlet valleys indicating the former presence of an open basin lake, i.e. one that overflowed (Fassett and Head, 2008). Balancing estimated stream discharges against measured basin area and estimated evaporation rates, Matsubara et al. (2011) inferred that some of the larger open lakes would have taken thousands of years to overflow. However, most basins are closed. From the fact that most local basins did not overflow, coupled with estimates of peak discharges from channel dimensions and computer simulations of fluvial erosion of cratered landscapes, Howard (2007), Barnhart et al. (2009), and Matsubara et al. (2011) concluded that the valley network system did not result from a few deluge-type events, such as might be caused by large impacts, but rather from modest fluvial events over extended periods of time. The model times required to simulate the observed landscapes depend on the frequency and magnitude of the fluvial events, but times could be as long as millions of years (Howard, 2007). Similarly, hydrologic modeling suggests that the Meridiane and other layered lacustrine deposits in Arabia formed by ground-water upwelling over millions of years. The fine-scale regular layering of many sediments in putative lake deposits suggests that inflow into the lakes was controlled by regularly changing conditions such as might be induced by obliquity changes.

Many deltas have been observed where valleys intersect local lows. A particularly striking example in Eberswalde Crater has many of the characteristics of terrestrial deltas and dimensions of its component parts that suggest discharge of several hundred cubic meters per second, comparable to terrestrial rivers (Malin and Edgett, 2003; Moore et al., 2003). The time taken to build the delta is not narrowly constrained but is estimated to be in the range of thousands to millions of years (Moore et al., 2003). Over 50 deltas situated along the plains’ upland boundary, close to the $-1.8$ km contour, have been attributed to sedimentation along the edge of a northern ocean (Di Achille and Hynek, 2010). In most cases, however, neither fans nor deltas are observed where a valley enters a low, possibly because the sediment load is generally too fine-grained (Howard, 2007).

Whether there were ever oceans on Mars is one of the planet’s most controversial issues (Parker et al., 1989, 1993; Baker et al., 1991; Head et al., 1999; Baker, 2001; Clifford and Parker, 2001; Carr and Head, 2003). Discussion has focused mainly on the possibility of post-Noachian oceans mainly because they could have resulted from the large post-Noachian floods, and because any evidence for oceans would be better preserved for the post-Noachian era than for the Noachian. However, the Noachian is the time for which we have the best evidence for warm conditions under which oceans might be present. Moreover, large bodies of water were likely needed to provide the precipitation that eroded the valley networks. Clifford and Parker (2001) argue from estimates of the global inventory of water and the thermal conditions implied by the valley networks that possibly one-third of the planet was covered by oceans during parts of the Noachian, and have tentatively identified shorelines that could be Noachian in age. The presence of deltas along a possible shoreline has already been mentioned. In addition, Moore and Wilhelms (2001) suggested that two well-defined breaks in slope within the Hellas Basin could be
17.3 Evidence for a Different Atmosphere and Climate

17.3.1 Summary

The geologic evidence for precipitation in the Noachian and flow of water across the surface is compelling. The dimensions of the valleys, lakes, and associated features such as deltas suggest rivers with discharges comparable to terrestrial rivers. The erosion rates, although poorly constrained, are estimated to have been similar to arid and semi-arid regions on Earth where water is the erosive agent. But not knowing the flow frequency and the time over which conditions were favorable to fluvial activity, the amount of water that flowed across the surface is poorly constrained. Nevertheless, simulations of the evolution of the Noachian landscape suggest that times of millions of years and precipitation rates equivalent to arid to semi-arid regions on Earth were required to achieve the observed degraded landscape (Howard, 2007). If true, then evaporation from large bodies of water must have occurred to provide precipitation and drive a global hydrologic cycle.

17.3.2 Mineralogical Evidence

The surface of Mars is predominantly basalt. When basaltic minerals react with water, the resulting solution can produce phyllosilicates, silica, sulfates, carbonates, iron oxides, and other alteration products (e.g. Burns, 1987; Catling, 1999; Christensen et al., 2001). Consequently, the water that flowed on the surface of early Mars, as demonstrated in Section 17.2.1, should have produced such minerals. Indeed, they have been detected, as summarized in Table 17.1. The question, however, is: Were they formed in a sustained warmer and wetter climate with an active hydrological cycle, or were they formed by other means, such as impact events or subsurface hydrothermal alteration, which are independent of the climate system?

17.3.2.1 Carbonates

Carbonates have been of particular interest since they are expected to form in warm wet conditions produced by thick CO₂ atmospheres (Kahn, 1985; Pollack et al., 1987). Such atmospheres were thought to produce the greenhouse effect needed to explain the fluvial features (see Section 17.5.2). Carbonates form when CO₂-rich water reacts with basaltic materials to release various cations (principally Fe²⁺, Mg²⁺, Ca²⁺, K⁺, and Na⁺) and bicarbonate anions (HCO₃⁻). When the resulting solution concentrates, bicarbonate combines with the cations to produce minerals such as siderite, calcite, or magnesite. Used up in this way, for example, 1 bar of CO₂ would produce a planet-wide cover of ~20 m of pure calcite².

Carbonates are present on Mars, but they do not outcrop globally (Niles et al., 2013). Occurrences include small volumes (<1 vol.%) in Martian meteorites (Bridges et al., 2001), layered carbonates overlying olivine-rich units in Nili Fossae (Ehlmann et al., 2008a; Ehlmann and Edwards, 2014), layered carbonates excavated by an impact in Syrtis Major (Michalski and Niles, 2010), magnesite outcrops in Gusev Crater (Morris et al., 2010), a 2–5 wt.% magnesite component of the global Martian dust (Bandfield et al., 2003), and 3–5 wt.% carbonate in soil at the Phoenix landing site of mixed Ca/Mg/Fe content (Boynton et al., 2009; Sutter et al., 2012). The area of carbonate outcrops detected in the near-infrared is ~10⁷ km², assuming that they extend under associated olivine units. Given an average thickness of ~100 m and assuming 100% calcite, the CO₂ they contain would be only ~3 mbar.

The absence of global outcropping has been attributed to the presence of sulfuric or sulfurous acids that would have suppressed their formation (Fairén et al., 2004; Bullock and Moore, 2007; Halevy et al., 2007). On early Mars, a slow rain of microscopic sulfuric and sulfuric acid droplets could have been generated from atmospheric oxidation of volcanic SO₂ hydrated by water vapor (Settle, 1979; Smith et al., 2014). While carbonate suppression by acidic waters is valid for steady-state conditions, in reality weathering fluids in contact with basalts become more alkaline in time. So carbonate formation is unlikely to have been suppressed everywhere at all times.

Do the known carbonates tell us anything about past climates? Notably, the Nili Fossae and Gusev Crater carbonates are magnesite-dominated. In Nili Fossae, magnesite is associated with the probable alteration of olivine-rich ejecta around Isidis (Mustard et al., 2008). On Earth, magnesite commonly forms through hydrothermal alteration of Mg-rich igneous rocks. Consequently, the association of magnesite with olivine probably tells us more about past geothermal activity than past climate.

Unlike the carbonates in Nili Fossae, however, those excavated from the deep crust have spectral signatures of siderite or possibly calcite (Michalski and Niles, 2010). This may be a glimpse of very ancient, deeply buried sedimentary carbonates from a warmer climate or it could simply be evidence of ancient subsurface hydrothermal alteration. In either case, ferrous iron in siderite implies a reducing environment, unlike the modern oxidizing surface environment, which can favor Mg-rich carbonates (Hausbrath and Olsen, 2013).

The carbonates in global dust and Phoenix soils may not have any connection to the Noachian climate. While the magnesite in global dust could have a contribution from the erosion of deposits such as those in Nili Fossae or in Gusev (Ehlmann et al., 2008a),

\[ \text{Depth} = \left( \frac{P_{\text{sat}}}{P_{\text{sat}}(g \rho_{\text{CO}_2})} \right) \left( \frac{\rho_{\text{CO}_2}}{g M_{\text{CO}_2}} \right), \]

where surface pressure \( P_s = 10^7 \) Pa, \( M_{\text{CO}_2} = 0.1 \) kg mol⁻¹, \( M_{\text{CO}_2} = 0.044 \) kg mol⁻¹, \( g \) is gravity, and density \( \rho_{\text{CO}_2} = 2710 \) kg m⁻³.
Table 17.1 Alteration minerals reported from orbital or in situ measurements on Mars. (Adapted and updated from Ehlmann (2010) and Murchie et al. (2009a).)

<table>
<thead>
<tr>
<th>Class</th>
<th>Minerals</th>
<th>Chemistry</th>
<th>Locations on Mars</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-silicates</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carbonates</td>
<td>Calcium carbonate</td>
<td>CaCO$_3$</td>
<td>Soil at Phoenix site; large impact crater east of Syrtis Major</td>
</tr>
<tr>
<td>Sulfates</td>
<td>Magnesite</td>
<td>MgCO$_3$</td>
<td>Nili Fossae; Gusev Crater; global dust</td>
</tr>
<tr>
<td></td>
<td>Siderite</td>
<td>FeCO$_3$</td>
<td>Large impact crater east of Syrtis Major</td>
</tr>
<tr>
<td></td>
<td>Gypsum</td>
<td>CaSO$_4$·2H$_2$O</td>
<td>Juventae Chasma$^a$; Iani Chaos; north circumpolar dunes</td>
</tr>
<tr>
<td></td>
<td>Bassanite</td>
<td>CaSO$_4$·0.5H$_2$O</td>
<td>Mawrth Vallis</td>
</tr>
<tr>
<td></td>
<td>Mono- and poly-hydrated sulfates</td>
<td>(Fe,Mg)SO$_4$·nH$_2$O, e.g. kieserite, MgSO$_4$·H$_2$O</td>
<td>Valles Marineris region: Ius, Hebes, Capri, Candor</td>
</tr>
<tr>
<td></td>
<td>Jarosite</td>
<td>KFe$_{3+}$(SO$_4$)$_2$(OH)$_6$ (H$_2$O)</td>
<td>Mawrth Vallis, Meridians</td>
</tr>
<tr>
<td></td>
<td>Alunite</td>
<td>KAl$_3$(SO$_4$)$_2$(OH)$_6$</td>
<td>Terra Sirenum craters</td>
</tr>
<tr>
<td></td>
<td>Metal chlorides</td>
<td>type not identified</td>
<td>Terra Sirenum plains</td>
</tr>
<tr>
<td></td>
<td>No mineral names exist</td>
<td>Mg$_2$ClO$_4$, (Na,K)ClO$_4$</td>
<td>Soil at Phoenix site; north polar cap</td>
</tr>
<tr>
<td>Iron oxides</td>
<td>Hematite</td>
<td>Fe$_2$O$_3$</td>
<td>Meridians; Aram Chaos; Candor; Melas, Juventae, Tithonium, and Eos Chasmata</td>
</tr>
<tr>
<td></td>
<td>Goethite</td>
<td>FeOOH</td>
<td>Gusev Crater rocks</td>
</tr>
<tr>
<td><strong>Sheet silicates</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phyllosilicates</td>
<td>Fe/Mg smectites</td>
<td>(Ca$<em>{0.3}$Fe$</em>{2.5}$Mg$<em>{0.2}$Al$</em>{0.1}$)$<em>{2}$(OH)$</em>{2}$, e.g. Fe-rich nontronite Na$<em>0.5$Fe$</em>{2+}$(Si,Al)$<em>2$O$</em>{10}$(OH)$<em>2$ and Mg-rich saponite Ca$</em>{1+6}$ Mg$_2$(Al,Si)$<em>2$O$</em>{10}$(OH)$_2$</td>
<td>Mawrth Vallis; Nili Fossae; Holden, Eberswalde, Jezero, Terby, and Gale Craters; walls of Valles Marineris; Terra Sirenum; Meridians; southern highlands</td>
</tr>
<tr>
<td></td>
<td>Kaolinite group</td>
<td>Al$_2$Si$_2$O$_5$(OH)$_4$ (kaolinite), e.g. halloysite Al$_2$Si$_2$O$_5$(OH)$_4$·(0–2)H$_2$O</td>
<td>Mawrth Vallis, Nili Fossae</td>
</tr>
<tr>
<td></td>
<td>Montmorillonite</td>
<td>(Na$<em>{0.3}$Ca$</em>{0.7}$Al$_{2}$Si$<em>3$O$</em>{10}$)(OH)$_2$</td>
<td>Mawrth Vallis</td>
</tr>
<tr>
<td></td>
<td>Al,K phyllosilicates</td>
<td>KAl$_3$Si$_2$O$_8$(OH)$_2$</td>
<td>Nili Fossae</td>
</tr>
<tr>
<td></td>
<td>Chlorite</td>
<td>(Mg,Fe$_{2+}$)$_2$Al$_2$(Si$_3$Al$<em>1$)O$</em>{10}$(OH)$_2$</td>
<td>Large northern plains or southern highland crackers</td>
</tr>
<tr>
<td></td>
<td>Serpentine</td>
<td>(Mg,Fe)$_2$Si$_2$O$_5$(OH)$_4$</td>
<td>Claritas Rise; Nili Fossae; highlands impact craters; regional olivine-rich unit near Isidis</td>
</tr>
<tr>
<td></td>
<td>Prehnite</td>
<td>Ca$_2$Al$_2$(Si$_3$Al$<em>1$)O$</em>{10}$(OH)$_2$</td>
<td>Large northern plains craters; southern highlands</td>
</tr>
<tr>
<td><strong>Framework silicates</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silica</td>
<td>Opaline silica</td>
<td>SiO$_2$·2H$_2$O</td>
<td>Gusev Crater; Valles Marineris Hesperian plains; Ius/Melas Chasmata; Noctis Labyrinthus; Mawrth Vallis</td>
</tr>
<tr>
<td>Analcime (analcite)</td>
<td>Analcime/analcite</td>
<td>NaAlSi$<em>3$O$</em>{10}$·H$_2$O</td>
<td>Highlands bordering Isidis or Hellas; Terra Cimmeria</td>
</tr>
</tbody>
</table>

$^a$ Gypsum is reported by Gendrin et al. (2005) in Juventae Chasma but disputed by others (Bishop et al., 2009; Kuzmin et al., 2009).
it could also have been produced relatively recently. Laboratory experiments show that a few weight percent of carbonate could form on the surface of basaltic fines in “wet” conditions of only \(10^2 \text{–} 10^3\) monolayers of water over about 1 Ga, although the exact timescale depends on the specific surface area \(\text{m}^2 \text{g}^{-1}\) of the fines (Stephens, 1995a,b). Such conditions could have been produced during the Amazonian at times of high obliquity (see Chapter 16). Furthermore, the soil at the Phoenix site is likely derived from the ejecta material of the 600 Ma old Heimdall Crater, which has also been exposed to periodic wet conditions (Boytont et al., 2009).

### 17.3.2.2 Sulfates

Sulfates are abundant on Mars. Near-infrared spectroscopy shows that hydrated sulfates exist in light-toned layered deposits (LLDs) (Gendrin et al., 2005; Bibring et al., 2006; Murchie et al., 2009a), while in situ analyses demonstrate that sulfate is ubiquitous in the soil (Yen et al., 2005; Kounaves et al., 2010). Astonishingly, sulfates outcrop as mountain-sized LLDs inside low-latitude chasms and canyons, often in association with chaotic terrain (Catling et al., 2006; Glotch and Rogers, 2007; Bishop et al., 2009; Lichtenberg et al., 2010; Roach et al., 2010a,b). On the ground, the Opportunity Rover provided a close-up view of LLDs in Meridiani Planum (Squyres et al., 2006), where millimeter-scale spherules of hematite (Weitz et al., 2006) are embedded in sulfate-rich sandstones.

The Meridiani LLD sandstones date to the Noachian–Hesperian boundary \(\sim 3.7\) Ga (Hynek and Phillips, 2008) and are composed of a mixture of sulfates, hydrated silicates, iron oxides, and basaltic debris (Clark et al., 2005). Regionally, they are several hundred meters thick and cover \(>500\) 000 km\(^2\). The meter-scale cross-bedding seen in the Burns formation and Victoria Crater (Squyres et al., 2009) indicate that these sedimentary stacks formerly consisted of migrating dunes (Grotzinger et al., 2005).

Squyres et al. (2006) proposed the following scenario for the formation of the Meridiani LLDs. Playa deposits of sulfates were wind-eroded and redeposited as sand dunes (McLennan and Grotzinger, 2008). Groundwater subsequently penetrated the sandstone. Hematite concretions formed when the dissolved iron in the groundwater precipitated on timescales as short as \(\sim 10^3\) years, constrained by the typical size and volume fraction of concretions (Sefton-Nash and Catling, 2008). Hydrothermal conditions may also have existed (Golden et al., 2008). The centimeter-scale cross-laminations in the upper meter of the Burns formation, the so-called “festoons”, suggest gentle subaqueous flow (Grotzinger et al., 2005). Thus, groundwater upwelled to the surface and ponded between dunes. Inter-dune ponding requires a rise in the water table but not direct rainfall (McLennan and Grotzinger, 2008). However, if such hydrology is responsible, the late Noachian or Lower Hesperian climate might require significant rainfall to charge groundwater aquifers.

\[\text{17.3.2.3 Phylosilicates}\]

Phylosilicates\(^3\) are products of basaltic weathering and require high water activity in near-neutral or alkaline conditions to form (Ehlmann et al., 2013). They were originally thought to be present in airborne dust from Mariner 9 observations (Toon et al., 1977) and later more definitively identified in Noachian surface materials by Mars Express (Bibring et al., 2005, 2006; Poulet et al., 2005; Ehlmann et al., 2011). However, only \(\sim 3\)% of Noachian surfaces have hydrous minerals such as phylosilicates or hydrous carbonates (Carter et al., 2012). Phylosilicates constitute a large group of minerals such as kaolinites or

---

\(^3\) The term “clay” alone refers to particulate materials with grain size less than 1/256 mm or 4 μm. The term “clay mineral” is generally taken to mean phylosilicates, including aluminosilicate minerals with a crystal structure that derives from phylosilicates.
smectites that are defined by their sheet arrangements4. Phyllosilicates have also been identified in Martian meteorites (e.g., Gooding, 1992) and confirmed by in situ measurements of mudstone material by MSL.

There are three stratigraphic classes of phyllosilicates: (i) deep phyllosilicates, (ii) layered phyllosilicates, and (iii) phyllosilicates in sedimentary basins (Murchie et al., 2009a). The deep phyllosilicates are excavated by impact craters with diameters of tens of kilometers in the Noachian southern highlands (Ehlmann et al., 2009; Wray et al., 2009b) or northern lowlands (Michalski and Niles, 2010). They are also found in layers outcropping from underneath Hesperian units (Mustard et al., 2009). The minerals in these phyllosilicates include chlorite, prehnite, and analcime, which are typically formed under hydrothermal conditions. Prehnite, for example, forms under temperatures of 200–350°C and lithostatic pressures below ~2 kbar. However, the presence of prehnite in impact ejecta and its thermal stability suggests that prehnite existed at depth before excavation (Fairén et al., 2010).

Layered phyllosilicates are laterally extensive in Mawrth Vallis and Nili Fossae and consist of Al phyllosilicates overlying kaolinite and Fe/Mg smectites in light-toned outcrops (Ehlmann et al., 2009). In Mawrth Vallis, the Al phyllosilicate layer can also include opaline silica, montmorillonite, and kaolinite (Bishop et al., 2008). Because smectite units drape underlying topography, possible precursors could be layers of volcanic ash that were altered by water. Fe–Mg smectites are also exposed by small ~100 m scale craters in the southern highland craters, and western Arabia – serpentine is detected. Whether phyllosilicates have implications for the climate depends on their formation environment. The coexistence of Mg-rich smectites and carbonates suggests an equilibrium \( pCO_2 \) of 1–10 mbar, which is comparable to the present surface pressure (Chevrier et al., 2007). Consequently, if the phyllosilicates formed in equilibrium with the atmosphere under warmer, wetter conditions, Chevrier et al. (2007) argue that non-CO2 greenhouse gases would be necessary. However, in the subsurface, under hydrothermal conditions, magnesium carbonates and serpentine can form independently of climate. It has also been suggested that many phyllosilicates on Mars could form from water-rich magma-derived fluids in the subsurface (Meunier et al., 2012).

The geomorphology and chemistry of particular deposits can be examined to determine whether groundwater alteration of ashfall, sedimentary placement of eroded clays, or percolating rainfall through soils (pedogenesis) was the likely mechanism for forming the phyllosilicates. In Mawrth Vallis, alteration of volcanic ash is favored because smectites drape over topography (Bishop et al., 2008; McKeown et al., 2009). Subsequent alteration of the phyllosilicates is also possible. Some sulfates may be found alongside phyllosilicates because of sulfuric acid weathering of previously deposited phyllosilicates, which produces different phyllosilicates along with silica and Al or Fe sulfates (Altheide et al., 2010). However, it is doubtful that this can explain layers of sulfates that are sandwiched between phyllosilicate layers (Wray et al., 2009b).

Phyllosilicates in fan and delta deposits could be particularly convincing evidence of liquid water if they formed in the lakes where these features arose, thereby combining mineralogical with geomorphic evidence. CRISM has detected phyllosilicates in Eberswalde, Holden and Jezero Craters, which contain fan and delta deposits (Ehlmann et al., 2008b; Mustard et al., 2008). However, the phyllosilicates are the same type as those in surrounding sediment source regions and so are probably detrital, i.e. washed in.

17.3.2.4 Silica

Opaline silica also potentially sets environmental constraints. Its coexistence with jarosite on the Hesperian plateau surrounding Valles Marineris, indicates low pH (Milliken et al., 2008). The silica can derive directly from weathered basalts or from phyllosilicates. Low-temperature to hydrothermal conditions can produce silica (McAdam et al., 2008) but high concentrations of silica (such as up to ~91% detected by the Spirit Rover) generally favor a hydrothermal genesis (Squyres et al., 2008).

17.3.2.5 Chlorides

These have been spectroscopically detected in association with Fe–Mg smectites in Terra Sirenum, in the Noachian southern highlands (Osterloo et al., 2008). Chlorides are found in intercrater plains and occasionally some crater floors. The chlorides are more recent and embay ancient phyllosilicate-rich knobs (Glotch et al., 2010). One hypothesis for formation of chlorides is groundwater discharge and evaporation, leaving chloride crusts (Jensen and Glotch, 2011).

17.3.2.6 Iron Oxides

High-albedo regions show absorption features characteristic of oxidized iron both from orbit (Bibring et al., 2006) and from landers (Bell et al., 2000), while orbital near-infrared spectra lack absorption signatures for hydration (Bibring et al., 2006). Mars’ red surface (Morris et al., 2000) and the butterscotch sky color (Huck et al., 1977) are both due to ferric iron oxides in finely crystalline or poorly crystalline form on the surface and in airborne dust. How such iron oxides formed is debated, but...
deposition of atmospheric peroxides (accompanied by escape of hydrogen to space) should produce surface oxidation (Bullock et al., 1994; Zahnle et al., 2008), even in the Amazonian. When this oxidation process started is relevant for whether the early atmosphere was reducing or oxidizing.

### 17.3.2.7 Perchlorates

The Phoenix Lander detected $\sim 0.6$ wt.% $\text{ClO}_4^-$, perchlorate, in soil samples through wet chemistry, which was corroborated by the evolved oxygen from heated samples (Hecht et al., 2009). Spectral data indicate that perchlorate has been concentrated into patches, possibly in brines (Cull et al., 2010). Pyrolysis of soil on MSL has also released oxygen consistent with perchlorates (Glavin et al., 2013). Perchlorate is also tentatively detected on the north polar cap (Massé et al., 2010). Initially, Mg, Na and K perchlorates were favored forms at the Phoenix site (Kounaves et al., 2010; Marion et al., 2010) but reanalysis suggests the possible presence of Ca perchlorate (Kounaves et al., 2014). The eutectic temperatures of Mg and Ca perchlorates are very low, $\sim 57^\circ C$ (Stillman and Grimm, 2011) and $\sim 77^\circ C$ (Pestova et al., 2005), respectively, which means that liquids will exist on Mars above these temperatures wherever these perchlorates contact ice. Moreover, metastable solutions can supercool tens of degrees below eutectic temperatures (Toner et al., 2013). The origin of the perchlorate may be similar to that on Earth, which is oxidation of chlorine volatiles in the air to perchloric acid, which dry-deposits (Catling et al., 2010), although presently unknown gas–solid reactions are required for Mars (Smith et al., 2014). This would have occurred in the past on Mars when chlorine gases were released by volcanism or impact volatilization.

### 17.3.2.8 Summary

Many alteration minerals have been detected in diverse areas on Mars. Phyllosilicates and magnesium carbonate are Noachian in age, while chlorides, layered sulfates, and opaline silica are found in Noachian and Hesperian units. The coexistence of Noachian phyllosilicates and carbonates suggests alkaline or weakly acidic conditions, while some sulfates, such as alunite or jarosite, imply acidic environments. The link between aqueous chemistry and climate is indirect, however. Climate implications depend on the extent to which rainfall was required for the aqueous alteration versus hydrothermal systems or low-temperature brines. Nonetheless, it is reasonably secure to conclude that groundwater systems, likely charged by rainfall, were necessary to form Noachian–Hesperian hematite concretions, and that hydrothermal waters were needed to form some Noachian phyllosilicate minerals such as serpentine. The current inventory of carbonate outcrops is tiny and taken at face value (i.e. assuming deeply buried carbonates are not significant in volume) would imply considerable escape of the atmosphere if the past inventory of $\text{CO}_2$ was large. In the present environment, a relatively large concentration of Mg or Ca perchlorate in the soil is permissive of brines at temperatures on Mars today. Consequently, low-temperature brines should exist today and such brines may have been important in the past.

### 17.3.3 Isotope Evidence

Strong evidence for atmospheric escape is preserved in the isotopes of Ar, H, Xe, and N. Modest evidence for escape is seen in the isotopes of C and O. Before examining the evidence, we briefly review the kinds of escape processes pertinent to Mars. (The reader is referred to Chapter 15 for details.) These provide a framework for interpreting the evidence. We then present the evidence and see how it fits the framework element by element, beginning with the noble gases.

Current escape rates are low. Escape can be thermal (Jeans escape) or non-thermal (e.g. sputtering, dissociative recombination). Low levels of escape are often highly fractionating, either because the lighter isotope extends higher above Mars and thus is more prone to escape, or because the lighter isotope is more likely to acquire escape velocity by chance (Jeans escape). Such escape can leave a strong signal in what is left behind, but because the escape rate is low, the signal can only be seen if the reservoir left behind is quite small. Large fractionations in Ar, H, and N have been attributed to such processes.

Although sputtering and dissociative recombination can cause escape from Mars today, these mechanisms would have been more effective on ancient Mars, because the young Sun would have been a stronger source of ionizing radiation and would have had a more intense solar wind (Luhmann et al., 1992). Some models have predicted ancient escape rates by sputtering that are several orders of magnitude higher than today (Luhmann et al., 1992). The exception to this is that, when young, Mars had a significant intrinsic magnetic field, which may have protected its atmosphere against sputtering.

Ancient escape processes to consider are impact erosion and hydrodynamic hydrogen escape. The roughly uniform hundredfold depletion of Mars’ noble gases with respect to Earth suggests that Mars lost much of its atmosphere by a relatively efficient, non-fractionating process. Impact erosion – the expulsion of atmosphere by impacts – is a leading candidate (Melosh and Vickery, 1989; Zahnle, 1993, Brain and Jakosky, 1998). Impact erosion is expected to have been efficient on Mars because its escape velocity is small compared to typical impact velocities. Impact erosion discriminates strongly between volatiles in the atmosphere and those that are condensed on or under the surface. Thus impact erosion preferentially removes the noble gases and nitrogen, while favoring retention of water, carbonate rock, chlorine, and sulfur. Isotopic fractionations are possible if the condensed state favors one isotope over another. Elements that could have been fractionated in this way on Mars are C, O, and to some extent H.

Impact erosion is restricted to early in Mars’ history when impact rates were high. We can set an upper bound on impact erosion by adopting the Melosh and Vickery (1989) “tangent plane” model for all impacts. In this model it takes roughly 500 craters bigger than 100 km diameter to erode a 1 bar atmosphere. Crater counts restrict this to times before 4 Ga (Hartmann, 2004). By the same model, it takes roughly 500 craters bigger than 30 km diameter to erode a 0.01 bar atmosphere. This corresponds to 3.5 Ga (Hartmann, 2004). If Mars had a thick (0.5–1 bar) atmosphere in the late Noachian or later, the atmosphere could not have been removed by impact erosion.
Hydrodynamic hydrogen escape is more speculative, but is probably required to explain fractionation in Xe. Hydrodynamic escape refers to escape driven by a pressure gradient. The gas flows into space; there is no clearly defined edge to the atmosphere. By definition, escape rates are much higher than today. If the escape rate is high enough, the hydrogen carries other gases with it and the gases left behind become mass-fractionated.

Hydrodynamic escape has two requirements. First, hydrogen must be a significant constituent of the atmosphere. This is likely shortly after Mars formed (Dreibus and Wänke, 1987), less likely at later times. Second, if hydrodynamic escape is to fractionate heavier gases like Ar and Xe, the escape flux must be high and this requires a solar EUV flux 10–50 times what it is today (Pepin, 1991). Given that the Sun’s EUV luminosity declined inversely with time (Zahnle and Walker, 1982), hydrodynamic escape is restricted to earlier than ~4 Ga. It should be noted that hydrodynamic escape is not likely to be greatly hindered by a planetary magnetic field – the outflowing planetary wind is too strong.

17.3.3.1 Data

Data sources are: in situ measurements by Viking, Phoenix, and MSL; trapped gases in Martian meteorites; and telescopic observations. The in situ observations have not always been self-consistent nor always consistent with the gases found trapped in glasses of the youngest of meteorites, EETA 79001 in particular (Bogard et al., 2001). With one notable exception, the available MSL measurements resemble the meteorite data more closely than they resemble Viking, Phoenix, or telescopic data. As the meteorites are the chief source of historical information, we concentrate on them.

Nyquist et al. (2001) and Bogard et al. (2001) comprehensively reviewed Martian meteorites and their volatiles. Martian meteorites formed at different times. The shergottites come from surface flows that might sample the atmosphere at the time of their crystallization ages (~170 Ma for some, ~450 Ma for others; Nyquist et al., 2001). Others, the nakhlites and Chassigny in particular, are older, were more deeply buried, and may sample the atmosphere at 1.4 Ga. Collectively, the shergottites, nakhlites, and chassignites are referred to as the SNC meteorites. ALH 84001 (which is not classified as an SNC) crystallized 4.1 Ga (Lapen et al., 2010). If it truly samples the Martian atmosphere at that time, it is the most relevant to early Mars (Leshin et al., 1996; Marti and Mathew, 2000; Miura and Sugiura, 2000; Sugiura and Hoshino, 2000; Mathew and Marti, 2001; Greenwood et al., 2008, 2010). NWA 7034, “Black Beauty”, is a 4.4 Ga impact melt breccia, but its trapped gases appear to be Amazonian in age (Cartwright et al., 2014). The meteorites also provide some glimpses into non-atmospheric (Mars interior) volatile reservoirs, although the latter are not always interpretable.

17.3.3.2 Noble Gases

First, it bears repeating that noble gases in the Martian atmosphere are a hundred times rarer than on Earth (Anders and Owen, 1977) and 10,000 times rarer than they are in carbonaceous chondrites or gas-rich enstatite chondrites (Pepin, 1991). Put another way, if Mars accreted 0.01% of its mass in a volatile-rich late veneer (roughly equal to the sum total of all the visible impact basins on Mars, excluding the hemispheric dichotomy), the atmosphere would carry the isotopic signature of that late veneer in its noble gases (Zahnle, 1993). However, no such signature is seen. This implies that a conventional late veneer was not the source of current atmospheric gases.

We now address the isotopic evidence in noble gases. Argon and xenon show strong isotopic evidence for fractionating escape. By contrast, neon (lighter than Ar) and krypton (lighter than Xe) show little and none, respectively. The noble gases are therefore naturally grouped in pairs, with Ne complicating the interpretation of Ar, and Kr doing the same for Xe. Table 17.2 summarizes the discussion.

(a) Argon Argon has three stable isotopes. The heavy isotope \(^{40}\text{Ar}\) would be exceedingly rare were it not for radioactive decay of \(^{40}\text{K}\) (half-life 1.25 Gyr). If \(^{40}\text{Ar}\) has not escaped, the amount of \(^{40}\text{Ar}\) in the atmosphere implies that Mars is only 1% degassed (Dreibus and Wänke, 1987). This could be accomplished by degassing a kilometer-thick basaltic crust over all of Mars’ history, or it could be accomplished by degassing Mars to the core during its first 20 Ma, or by any of many other scenarios.

The lighter isotopes \(^{36}\text{Ar}\) and \(^{38}\text{Ar}\) are non-radiogenic. The \(^{36}\text{Ar}/^{38}\text{Ar}\) ratio on Earth is 5.32 and the chondritic ratio is 5.30 (Pepin, 1991); the solar wind ratio is 5.50 (Vogel et al., 2011). Against these, the Martian atmospheric ratio of 4.2±0.1 (Atreya et al., 2013) is strikingly heavy. The MSL measurement confirms what had been deduced from the SNC meteorites (Wiens et al., 1986; for a review see Bogard et al., 2001). The argon fractionation is unique to the Martian atmosphere and therefore requires substantial mass-fractionating escape of Ar from Mars.

The argon escape mechanism that can operate throughout Mars’ history is sputtering (Jakosky et al., 1994; Hutchins and Jakosky, 1996; Hutchins et al., 1997). Broadly put, sputtering escape occurs when the solar wind picks off atoms that rise above the exobase. Because low-mass atoms rise highest, the process discriminates by mass. Jakosky et al. (1994) showed that sputtering implies the loss of at least 50–75% of \(^{36}\text{Ar}\) to generate the observed \(^{40}\text{Ar}/^{36}\text{Ar}\). Hutchins and Jakosky (1996) developed a more complete model that included radiogenic \(^{36}\text{Ar}; they estimated loss of 85–95% of \(^{36}\text{Ar}\) and loss of 70–88% of outgassed \(^{40}\text{Ar}. Hutchins et al. (1997) revised the model to include shielding by an ancient magnetic field, which increased the uncertainties (e.g. 75–99% loss of \(^{36}\text{Ar}\).

Recent sputtering poses several problems. Krypton is not subject to sputtering escape, and thus to prevent the Ar/Kr ratio from shrinking too much the model implicitly requires that Ar be replenished preferentially. Neon is more susceptible to sputtering than Ar, and therefore it too needs to be replenished. Hutchins and Jakosky (1996) argue that recent volcanic degassing would fall one or two orders of magnitude short; the problems with neon will be addressed with that element.

It therefore appears that modern sputtering losses have been overestimated and that the \(^{40}\text{Ar}/^{36}\text{Ar}\) fractionation may be an ancient feature of the atmosphere. This is consistent with sputtering expected to have been much more effective early in Mars’ history (Luhmann et al., 1992). But an ancient origin also opens the door to hydrodynamic escape. In an \(\text{H}_2–\text{CO}_2\) atmosphere, the maximum rate that hydrogen can flow through...
Evidence for a Different Atmosphere and Climate

a CO₂ atmosphere that does not escape is the diffusion limit (Hunten and Donahue, 1976). What this means for Ar, which is slightly less massive than CO₂, is that Ar can escape with the hydrogen, but that escape is inefficient and more strongly mass-fractionating than sputtering (Zahnle et al., 1990). In this model, a 36Ar/38Ar of 4.2 requires only a third of 36Ar to escape, and none of the 40Ar. Both sputtering and diffusion-limited hydrodynamic escape predict significant neon fractionation (e.g. Zahnle et al. (1990) predicted 20Ne/22Ne ~ 7), and thus both models require replenishing the atmosphere with a small amount of isotopically light neon.

Radiogenic Ar also sets useful constraints that are broadly consistent with Ar fractionation being ancient rather than recent. The 40Ar/36Ar ratio of ~1800 ± 100 deduced from trapped gases in SNC meteorites (Bogard et al., 2001) and confirmed by MSL (1900 ± 300; Mahaffy et al., 2014) is much higher than on Earth (294). There is evidence in trapped gases in SNCs that the 40Ar/36Ar ratio has not changed appreciably in the past 1.4 Ga (Bogard and Garrison, 2006; Cassata et al., 2012). This suggests that neither 40Ar degassing nor 36Ar sputtering losses have been important in the past 1.4 Ga, as either process would raise the 40Ar/36Ar ratio. For comparison, Hutchins and Jakosky (1996) predicted that the 40Ar content in the atmosphere should have increased by a factor of 2.5 in the past 1.4 Ga.

A very ancient atmospheric 40Ar/36Ar ratio of 626 ± 100 has been reported in a 4.16 ± 0.04 Ga mineral in ALH 84001 (Cassata et al., 2012). This ratio is surprisingly high given that only 20% of Mars’ 40K had decayed and only a fraction of the 40Ar could have reached the atmosphere. The implication is that the atmospheric inventory of 36Ar was already small by Earth’s standards. Cassata et al. (2012) developed a model around this datum to set an upper bound of 0.4 bar on atmospheric CO₂ at 4.16 Ga. Their argument is too model-dependent to stand as a true upper bound, but the overall sense that by 4.16 Ga Mars had lost more than 90% of the 36Ar it began with seems inescapable.

(b) Neon Neon has three stable isotopes that to within (ample) error reveal little by themselves. The 20Ne/22Ne ratio in SNC impact glasses is ~10 (Bogard et al., 2001), quite comparable to Earth’s. But when viewed against the backdrop of argon escape, the relatively modest fractionation of neon isotopes is telling. All proposed models to account for 36Ar/38Ar fractionation predict that the 20Ne/22Ne ratio should be lower than it is because Ne escapes more easily than Ar (e.g. Zahnle et al. (1990) predict 20Ne/22Ne ~ 7). Hence all models require that new neon be added to the atmosphere after argon escape (Hutchins and Jakosky, 1996). The source of fresh neon would constrain when Ar escape took place if the source could be quantified. However, sources with the required properties – a high Ne/Ar ratio and a high 20Ne/22Ne ratio – are solar. Very few meteorites have these properties. Pesyanoe, an enstatite chondrite, is an exception. If all post-Noachian impacts on Mars were Pesyanoe-like, it is possible for Ar fractionation to post-date the Noachian. Otherwise an exogenous source can work only if Ar fractionation is ancient. Volcanic sources are also problematic. The interior component in Chassigny is probably the best guide we have to the composition of Martian volcanic gases (Ott, 1988; Bogard et al., 2001). If the 36Ar content of Chassigny is representative, and if the Ne/O ratio of the gases was solar, it would take Tharsis-like degassing to reset the atmosphere’s neon isotopes. This is unlikely in the post-Noachian period.

Thus, three lines of evidence suggest that Ar fractionation was established early. The constant 40Ar/36Ar ratio in SNCs precludes significant Ar escape and significant Ar degassing after 1.4 Ga. The high 40Ar/36Ar ratio in ALH 84001 shows that most

---

**Table 17.2. Isotope data and their interpretation.**

<table>
<thead>
<tr>
<th>Isotopic ratio</th>
<th>Mars atmosphere</th>
<th>Earth atmosphere</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>36Ar/38Ar</td>
<td>&lt;3.9 (SNCs)</td>
<td>5.3</td>
<td>Mass-fractionating escape (sputtering, hydrodynamic escape)</td>
</tr>
<tr>
<td></td>
<td>4.01 (MSL)</td>
<td></td>
<td>Early escape of 36Ar (hydrodynamic escape, impact erosion)</td>
</tr>
<tr>
<td>40Ar/36Ar</td>
<td>3000 (Viking)</td>
<td>300</td>
<td>Early escape of 36Ar (hydrodynamic escape, impact erosion)</td>
</tr>
<tr>
<td></td>
<td>1800 (shergottites)</td>
<td></td>
<td>Early escape of 36Ar (hydrodynamic escape, impact erosion)</td>
</tr>
<tr>
<td>20Ne/22Ne</td>
<td>~10 (shergottites)</td>
<td>9.8</td>
<td>Possible escape</td>
</tr>
<tr>
<td>Xe isotopes (Figure 17.7)</td>
<td>Fractionated</td>
<td>Fractionated</td>
<td>Hydrodynamic escape</td>
</tr>
<tr>
<td>129Xe/132Xe</td>
<td>2.5</td>
<td>0.97</td>
<td>Early escape of 132Xe (impact erosion)</td>
</tr>
<tr>
<td>14N/15N</td>
<td>170</td>
<td>272</td>
<td>Mass-fractionating non-thermal escape</td>
</tr>
<tr>
<td>D/H</td>
<td>Variable, (1−6.7)×SMOW a</td>
<td>SMOW</td>
<td>Mass-fractionating thermal escape</td>
</tr>
<tr>
<td>δ¹⁸O in CO₂</td>
<td>48±5‰ (MSL)</td>
<td>0</td>
<td>Possible escape</td>
</tr>
<tr>
<td>δ¹³C in CO₂</td>
<td>46±4‰ (MSL)</td>
<td>0</td>
<td>Possible escape</td>
</tr>
</tbody>
</table>

* Standard mean ocean water.
The Early Mars Climate System

$^{36}$Ar escape took place before 4.16 Ga. The relatively unfracti-
nated Ne isotopes today require injecting new neon, so that
Ar escape was complete before volcanic activity or impact bom-
bardment subsided. Although any of these arguments can be
questioned, they are independent, so each must be contradicted
independently. Links between Ar and CO$_2$ are model-
dependent, but at least one such model concludes that the atmosphere was
no thicker than 0.4 bar at 3.8 Ga (Cassata et al., 2012).

(c) Krypton Krypton has six stable isotopes. Martian Kr is
indistinguishable from solar Kr and clearly distinct from chon-
dritic Kr and terrestrial Kr. Krypton has not been affected by
fractionating escape. But Kr has escaped: its overall abundance
is a little high but not grossly out of line with the abundances of
Ar and Xe. This supports the hypothesis that the bulk of noble
gas escape – including Ar – was by impact erosion, and not by
a fractionating process like hydrodynamic escape. It also means
that the total amounts of Ar and Xe lost during fractionating
escape have been modest, no more than a factor 2 or 3.

(d) Xenon Xenon has nine stable isotopes, several of which
are partially radiogenic. As shown in Figure 17.7, Mars’
atmospheric Xe (trapped in SNC glasses) is strongly mass-
fractionated with respect to solar or meteoritic Xe (Swindle et al., 1986; Wiens, 1988; Bogard et al., 2001). This observation
suggests that some Xe was lost by a mass-fractionating process
that did not affect Kr.

Atmospheric radiogenic xenon is very depleted. Best esti-
mates (Swindle and Jones, 1997; Mathew et al., 1998) are that
the atmosphere retains only ~0.1% of its possible radiogenic
$^{129}$Xe (from decay of $^{129}$I, half-life 15.7 Myr) and ~0.3% of its
possible radiogenic $^{136}$Xe (from spontaneous fission of $^{244}$Pu,

\[ \text{Relative to Nonradiogenic Solar Wind} \]

\[ \begin{array}{c|c|c|c|c|c}
\hline
\text{Atomic Mass} & 124 & 126 & 128 & 130 & 132 & 134 & 136 \\
\hline
\text{CI chondrites (AVCC)} & 1.1 & 0.9 & 0.7 & 0.5 & 0.3 & 0.1 & 0.0 \\
\text{Fractionated AVCC @ 2.4%/amu} & 1.1 & 0.9 & 0.7 & 0.5 & 0.3 & 0.1 & 0.0 \\
\text{Mars (Swindle et al 1986)} & 1.1 & 0.9 & 0.7 & 0.5 & 0.3 & 0.1 & 0.0 \\
\text{Fractionated SW@ 4.0%/amu} & 1.1 & 0.9 & 0.7 & 0.5 & 0.3 & 0.1 & 0.0 \\
\text{Chassigny (Ott 1988)} & 1.1 & 0.9 & 0.7 & 0.5 & 0.3 & 0.1 & 0.0 \\
\text{Frac. SW @ 2.9% per amu + PU} & 1.1 & 0.9 & 0.7 & 0.5 & 0.3 & 0.1 & 0.0 \\
\hline
\end{array} \]

Figure 17.7. Xenon in Mars’ atmosphere is well explained by mass fractionation of carbonaceous chondritic Xe with no fission products
(Swindle et al., 1986), or it can be fit by mass fractionation of solar wind xenon with a small addition of fission Xe (Swindle and Jones, 1997;
Mathew et al., 1998). The latter is preferred because Xe in Chassigny appears solar (Ott, 1988). Martian Kr (not shown) is identical to solar Kr.
Martian $^{129}$Xe plots at 2.5 (beyond the top of the diagram). These relationships superficially resemble what is seen on Earth, but differ in every
detail, which implies that parallel processes have operated on different materials on the two planets.

Martian internal xenon as seen in Chassigny differs strikingly
from the atmosphere (Ott, 1988; Mathew et al., 1998; Bogard et al., 2001). Isotopically, it closely resembles solar xenon (Figure 17.7).
Martian atmospheric xenon can be derived by mass-fractionating Chassigny xenon and adding radiogenic Xe from $^{129}$I decay and
fissiogenic xenon from $^{244}$Pu decay (Swindle and Jones, 1997; Mathew et al., 1998). Chassigny is also rather Xe-rich. Degassing
of 800 m of Chassigny-like basalt (~10$^8$ km$^3$) would double the
xenon reservoir in the atmosphere (Zahnle, 1993). It takes less
volcanic degassing to reset xenon’s isotopes than neon’s.

The ancient ALH 84001 contains several different xenons.
The most interesting of these is something Mathew and Marti
(2001) call “early atmosphere”, which is radiogenically
enriched ($^{129}$Xe/$^{132}$Xe $= 2.16$; 2.5 is the signature of the mod-
ern Martian atmosphere) and correlated with a rather Earth-like
nitrogen ($\delta^{15}$N $= +7\%$), but which is not as mass-fractionated as
xenon in the atmosphere. If this really is a sample of the atmos-
phere 4.1 Ga, it means that most of the mass-fractionating Xe
escape took place after 4.1 Ga.

Hydrodynamic hydrogen escape can fractionate xenon
because a heavier isotope is dragged to space less easily than
a lighter isotope (Sekiya et al., 1981; Hunten et al., 1987;
Sasaki and Nakazawa, 1988; Pepin, 1991). This leads to the
same problem with sputtering escape of argon and neon: if Xe
is fractionated by hydrodynamic escape, then Kr should be as
well. But it is not.

There are several ways to address this. One approach is to
presume that the other noble gases are entirely lost, so that only
fractionated Xe remains. Then fresh Ne, Ar, and Kr are added to the atmosphere, but fresh Xe is not. If this is done endogenously, Xe must remain inside Mars when Ne, Ar, and Kr degas (Pepin, 1991). Why Xe does not degas is unspecified. If the noble gases are replenished exogenously, a class of impactor is presumed in which Kr is much more abundant than Xe. This requirement might be met by some comets (Dauphas, 2003): experiments show that Kr is preferentially adsorbed by amorphous ice at ~50 K compared to Xe (Owen and Bar-Nun, 2001).

It is interesting to point out that the comet hypothesis does not actually require any Xe escape at all. The same story can be told if the mass-fractionated Xe were the result of gravitational settling inside porous planetesimals (Ozima and Nakazawa, 1980; Zahnle et al., 1990). This highly fractionated Xe is accompanied by a very low Kr/Xe ratio (also attributable to gravitational settling). Gravitational settling of atmospheric gases in deep snow is a well-known phenomenon in glaciology. If comets have very low Xe/Kr ratios and are the source of Kr, there is no need for hydrodynamic escape.

Another speculative possibility is that Xe escapes from planetary atmospheres by a different mechanism than Kr. This can be the case in hydrodynamic escape if Xe, the only noble gas more easily ionized than hydrogen, escapes as an ion in an ionized hydrogen wind. The strong Coulomb interaction between ions couples the Xe+ to the flow. An ionized wind is possible along open (polar) magnetic field lines or in the absence of a magnetic field. In principle this could take place with relatively modest (but still significant) hydrogen escape rates and relatively recently, which makes it the only hypothesis that can accommodate fractionating Xe after 4.1 Ga.

17.3.3.3 Other Gases

(a) Nitrogen Nitrogen in the Martian atmosphere is isotopically heavy. Viking reported $^{14}\text{N}/^{15}\text{N} = 168\pm 17$, a measurement confirmed by MSL, which reported 173±11 (Wong et al., 2013; Earth is 272). In delta notation, $^{15}\text{N} = 570\pm 100_{\text{oo}}$. SNC impact glasses also contain heavy nitrogen, but in no case is the nitrogen as heavy as atmospheric nitrogen. Several SNCs also carry isotopically light nitrogen ($^{15}\text{N} = -30_{\text{oo}}$; air is 0$_{\text{oo}}$ by definition) that resembles nitrogen in enstatite chondrites. This appears to be an internal Martian reservoir and a plausible starting point for nitrogen evolution. The SNCs are therefore best interpreted as trapping a mixture of atmospheric and interior nitrogen (Bogard et al., 2001).

McElroy (1972) predicted that 99% of Mars’ initial nitrogen would have been lost to space by dissociative recombination. This occurs when an N$_2$+ ion (ionized by solar EUV radiation) recombines with an electron to form two high-velocity neutral atoms. The energy released is enough for the upward-moving atom to escape from Mars if it does not collide. Therefore, if dissociative recombination takes place above the exobase, a nitrogen atom is lost to space. Viking’s discovery that Martian nitrogen is heavy was taken to confirm the theory (Nier and McElroy, 1977).

Nitrogen escape from Mars has been revisited many times since, with increasing sophistication (McElroy et al., 1977; Wallis, 1989; Fox and Haé, 1997a; Fox, 2007). Nitrogen can be lost by sputtering as well as by dissociative recombination. Both mechanisms favor $^{14}\text{N}$ escape because the $^{28}\text{N}_2/^{30}\text{N}_2$ ratio is higher at the exobase than in the lower atmosphere. Another effect is that, after dissociative recombination, the $^{14}\text{N}$ atoms have higher velocities than the $^{15}\text{N}$ atoms. Most calculations suggest that N escape is fast and rather too easily accounts for the $^{15}\text{N}$ enrichment in the remaining N$_2$. Indeed, the mechanism is often computed to be so efficient that it requires a substantial non-atmospheric nitrogen reservoir to damp the fractionation (McElroy et al., 1977; Bogard et al., 2001). In the limit, Wallis (1989) and Manning et al. (2009) suggested that the atmospheric $^{14}\text{N}/^{15}\text{N}$ ratio might be in steady state between a terrestrial-like source of N and escape. The range of elevated $^{14}\text{N}/^{15}\text{N}$ values in SNCs might then be taken as snapshots that chronicle the “steady-state” response of the atmosphere to rapid escape and stochastic resupply.

Nitrogen in ALH 84001 is paradoxical. Mathew and Marti (2001) report a correlation between $^{15}\text{N}$ and $^{129}\text{Xe}/^{132}\text{Xe}$ that they use to define what they call the “early atmosphere”. The nitrogen isotopic value corresponding to the highest $^{129}\text{Xe}/^{132}\text{Xe}$ (= 2.16) is $+7_{\text{oo}}$. Mathew and Marti therefore suggest that, at 4 Ga, Martian N had evolved from an initial value of $-30_{\text{oo}}$ to $+7_{\text{oo}}$. However, Miura and Sugiura (2000) report that ALH 84001 contains the heaviest N found in a Martian meteorite, with $^{15}\text{N} = 420_{\text{oo}}$ ($^{14}\text{N}/^{15}\text{N} = 190$). Miura and Sugiura regard the heavy N as representative of the atmosphere at 4 Ga. There is thus a paradox: evidence has been reported that $^{14}\text{N}$ has evolved over the past 4 Ga and that it has not evolved over the past 4 Ga.

(b) Hydrogen Hydrogen is the most strongly fractionated element on Mars and, with Ar, the most certain to have escaped. The D/H ratio in water vapor measured by ground-based telescopic observations is 5–7 times higher than it is on Earth or in most meteorites (5.5±2×SMOW; Owen et al., 1988; Owen, 1992; Krasnopolsky et al., 1997, 1998; Krasnopolsky, 2000; Villanueva et al., 2008, 2015; Novak et al., 2011). MSL has confirmed this (Webster et al., 2013), both for water vapor (5.95±1.08) and for the surprisingly abundant adsorbed water in the Rocknest fines (6.88±0.06). The other reported telescopic measurement is an inference that D/H = 2.4±1×SMOW in H$_2$ (Krasnopolsky, 2002), a measurement that is directly relevant to escape.

Heavy hydrogen is seen in SNCs (Watson et al., 1994; Leshin et al., 1996; Leshin, 2000; Greenwood et al., 2008, 2010). D/H ratios ranging from terrestrial to Martian water vapor are reported. The young (~170 Ma; Nyquist et al., 2001) SNC basals Shergotty, Zagami, and Los Angeles have D/H ratios in apatites of 5.6×SMOW, 5.4×SMOW, and 4.5–5.1×SMOW, respectively (Greenwood et al., 2008). These are thought to represent atmospheric water. The extent to which the more Earth-like D/H ratios are attributable to terrestrial contamination versus indigenous water of Mars has been debated (Leshin et al., 1996; Doctor et al., 2003; Greenwood et al., 2008, 2010), but at least some of the low-D/H water is Martian. Leshin et al. (1996) report that a number of nakhlites (1.4 Ga) have D/H ~ 1.6×SMOW and a few shergottites (0.17 Ga) also have D/H ~ 2.2×SMOW in hydrous minerals. These are plausibly interpreted as sampling meteoric crustal water (Carr, 1996; Leshin et al., 1996). Bogard et al. (2001) list D/H = 1.09±SMOW for Martian interior water. Kurokawa et al. (2014) argue that Mars’

5 Standard mean ocean water, i.e. Earth.
original D/H ratio was 1.275×SMOW. It seems, therefore, that the starting point for D/H evolution on Mars was SMOW-like.

ALH 84001 also contains both heavy (4.0±0.1×SMOW; Greenwood et al., 2008) and light water (1.08×SMOW; Bogard et al., 2001; see also Boctor et al., 2003). As with the SNCs, the simple interpretation is that the rock records both internal undegassed water (low D/H) and environmental water (high D/H). This interpretation would imply that D/H fractionation and hydrogen escape were early processes that pre-date 4.1 Ga (Greenwood et al., 2008).

A recent *in situ* finding by MSL reports that D/H = 3.0±0.1×SMOW in water (hydroxide) that is chemically bound in clay minerals extracted from the mudstones of Yellowknife Bay (Mahaffy et al., 2014). Although the age of these minerals is uncertain, they appear to have been subjected to fluvial activity as late as 3.0–3.5 Ga (Grant et al., 2014; Grotzinger et al., 2015). It is likely that this surface water was exchanging with the atmosphere so that the D/H in these minerals would reflect the composition of the atmosphere at that time.

Interpretation of the D/H data is complicated by the fact that they vary so much (Villanueva et al., 2015), and that it is difficult to ascertain the source of the water (e.g. Boctor et al., 2003). The simplest approach to interpreting the data, illustrated in Figure 17.8, is through the use of Rayleigh distillation, which relates changes in fractionation to changes in the size of reservoirs (e.g. Yung et al., 1988; Carr, 1990; Yung and Kass, 1998; Kass and Yung, 1999; Villanueva et al., 2008, 2015). The only assumptions are (i) there is no resupply and (ii) the fractionation factor $F$ that describes the relative efficiency of D escape to H escape is constant. A range of fractionation factors has been reported in the literature ($F = 0.32$, Yung et al., 1988; $F = 0.016$, Krasnopolsky et al., 1998). Since hydrogen appears to be escaping at close to the diffusion limit (Zahnle et al., 2008), we use $F = 0.4$, which is the ratio of D/H in H$_2$ (= 2.4, Krasnopolsky, 2002) to D/H in H$_2$O (= 6, see above references).

While the exact value of $F$ in this range has an effect on reservoir sizes, the effect is modest and it does not change the basic conclusions about the need for significant escape over time.

In Rayleigh distillation, the remnant reservoir $R_{\text{now}}$ is related to the initial reservoir $R_{\text{init}}$ by

$$\frac{R_{\text{init}}}{R_{\text{now}}} = (\frac{D/H}_{\text{now}}) (\frac{D/H}_{\text{init}})^{-F}$$

(17.1)

With $F = 0.4$, and a D/H now of 6×SMOW, only 5% of an initial SMOW water reservoir would remain. To convert relative reservoir sizes into absolute reservoir sizes requires another piece of information, either a known reservoir size or the hydrogen escape flux. If our best estimate of the current escape rate has been constant through time, Mars loses ~8 m of water in 4 Ga (Zahnle et al., 2008), leaving it with ~40 cm of water today. Thus, this model predicts that we are looking at Mars just before it dries up. These updated estimates differ only in details from what Yung et al. (1988) found when they first presented this argument. It is an inevitable consequence of starting D/H evolution from water initially at SMOW.

On the other hand, Rayleigh distillation works much better if evolution begins at ~3 Ga (a rough age for the Gale clays) with D/H = 3.0×SMOW. For D/H = 6.0 currently, $F = 0.4$, and today’s H escape rate, the 6 m of water to escape over 3 Ga is two-thirds of an initial 9 m inventory, which leaves 3 m today. This model is also consistent with the younger SNCs, because it predicts that 170 million years ago D/H = 5.6×SMOW.

A feature of this simplest of possible models is that water inventories scale with the hydrogen escape rate. Carr and Head (2014) find that the total amount of near-surface unbound water on Mars today is ~50 m. If this represents the exchangeable reservoir, it implies that 150 m of exchangeable water was present when the Gale clays formed, and that 100 m of water have escaped since. This means that the average hydrogen escape rate for the past 3 Ga was 17 times...
bigger than it is today. The requirement for more H escape on a 3 Ga timescale is reasonable, although it needs to be emphasized that the rate of hydrogen escape is controlled by the oxygen sink (Hunten and Donahue, 1976; Zahnle et al., 2008). Escape and crustal oxidation are the main known sinks for oxygen. Since oxygen escape models cannot account for 100 m of water loss (Lammer et al., 2003), a sizable fraction of the excess oxygen must have been taken up by the crust. Of course, these numbers scale directly with the present reservoir size. A more likely scenario is that 20 m of water is presently exchangeable. This represents the size of the polar ice caps, which are thought to come and go with orbital variations (see Chapter 16). If only 20 m of the near-surface water is currently exchangeable, then 60 m of water was exchanging when the Gale clays formed, the average escape rate was only seven times the present value, and much less crustal oxidation is required.

One thing Rayleigh distillation models cannot generate is a lower D/H ratio at 3 Ga from the higher D/H ratio at 4.1 Ga in ALH 84001. If the D/H = 4 number truly represents atmospheric D/H at 4.1 Ga as Greenwood et al. (2008) claim, then a great deal of fresh isotopically light water must have been released to the surface after 4.1 Ga. The most attractive candidate is the liberation of old low-D/H crustal water by outflow events, which, as described by Carr (1996), are of the right order of magnitude. It is tempting to identify the old water with the D/H = 1.6 water in nakhlite and the D/H = 2.2 water in the (shallower) shergottites (Leshin et al., 1996). It is then reasonable to think that the lakes in Gale sample an unusual moment when an exceptionally large amount of old water had been recently mobilized. Exogenous water and juvenile volcanic water cannot be ruled out, but neither mechanism is quantitatively attractive. The amount of low-D/H exogenous water required is at the high end of what could be delivered after 4 Ga (cf. Chyba, 1990), while juvenile low-D/H volcanic water works only if the erupting matter is wet (>1% H2O) or the volume of volcanism is in excess of the Tharsis construct.

To get to a D/H = 4 at 4.1 Ga requires significant early escape. Hydrodynamic escape, which best explains the xenon data, is weakly fractionating in a CO2 atmosphere such that $F = 0.8$ (Zahnle et al., 1990). To get D/H = 4×SMOW at 4.1 Ga, Rayleigh distillation requires the loss of 99.999% of the initial reservoir, which Dreibus and Wänke (1987) estimate to be 150 km. Thus, 150 m would remain at 4.1 Ga. While the loss of this much water is possible because the process includes escape during accretion, it seems quite extreme. What is more likely is that this much water is possible because the process includes escape, but these would be associated with episodes of lower D/H. A generally similar story was explored by Carr (1996).

In summary, fitting the D/H data to a Rayleigh distillation model implies the loss of significant amounts of water over Mars’ history. If the current exchangeable reservoir is ~20 m, and hydrogen has escaped at the diffusion limit ($F = 0.4$), then the exchangeable reservoir at 3 Ga was ~60 m, the average H escape rate over the past 3 Ga was ~7 times higher than it is presently, and only a modest surface sink for oxygen is needed. An early epoch of hydrodynamic escape would have escaped many kilometers of water but fractionation of SMOW-like initial water to 4×SMOW at 4.1 Ga was likely accomplished after hydrodynamic escape ended and most of Mars’ residual water was more or less permanently locked up in the crust. The subsequent lowering of D/H to 3×SMOW at 3 Ga could have been accomplished by mixing with low-D/H crustal waters brought to the surface by outflow channels.

(c) Oxygen Oxygen escape has been of interest since McElroy (1972) predicted that C and O could escape by dissociative recombination of CO2. Oxygen escape by this and many other possible non-thermal mechanisms has been extensively studied (Hunten and Donahue, 1976; Liu and Donahue, 1976; McElroy et al., 1977; Luhmann et al., 1992; Fox, 1993; Jakosky et al., 1994; Fox and Haé, 1997b, 2010; Luhmann, 1997; Lammer et al., 2003). Many of these studies were based on the expectation that O should escape at half the rate that H does (assuming no surface sinks), so that it is H2O that on net escapes. The newer studies compute O escape rates that are much too small to balance H escape (Lammer et al., 2003). Whatever the details, in general the lighter isotopes escape more easily, so that, if escape has been important, mass-dependent fractionation is expected (Jakosky et al., 1994).

Oxygen in atmospheric CO2 today is somewhat heavier than the terrestrial standard (SMOW), with δ18O = 48±5‰ from MSL (Mahaffy et al., 2013). This can be compared to ground-based telescopic observations (δ18O = +18±18‰, Krasnopolsky et al., 2007) and in situ measurements by the Phoenix Lander (δ18O = +31±5.7‰; Niles et al., 2010). H2O in the Rocknest fines is quite heavy, with δ18O = +84±10‰ (Webster et al., 2013). Oxygen in water vapor is not reported (Webster et al., 2013). Oxygen in carbonate in SNCs and ALH 84001 is typically between +10‰ and +20‰ (Niles et al., 2010; Halevy and Eiler, 2011), with enough scatter to mask any obvious trends (Niles et al., 2010).

The fractionation in O is rather modest, and there is no sign that δ18O has evolved from ALH 84001 and Shergotty in the carbonate isotopes. The fractionation is not outside the bounds of what geochemistry can do, and thus not so big that only O escape can explain it. The usual interpretation for a modest, slowly evolving fractionation is that either H2O or CO2 (or both) is escaping with a much larger reservoir. An alternative explanation is that current and historic non-thermal O escape rates have been greatly overestimated.

Oxygen isotopes in meteorites also show multiple levels of strong, distinctive mass-independent fractionations (MIF) in silicates, water, carbonates, and sulfates in SNCs.Bulk Mars, as represented by the silicates, is distinct from Earth with Δ17O = 0.3‰. Water and carbonates (in Lafayette and Nakhla) are
more enriched in $^{17}$O than the silicates with $0.5\% < \Delta^{17}$O < 1.0$\%$ (Karlsson et al., 1992; Farquhar and Johnston, 2008), and sulfate in Nakhla is still more enriched with $\Delta^{17}$O = 1.2$\%$ (Farquhar and Johnston, 2008). A similar MIF is seen in carbonate in ALH 84001 (Farquhar and Johnston, 2008).

On Earth the chief source of $^{17}$O enrichments is ozone. Most highly oxidized species in Earth’s atmosphere (nitrates, sulfates, perchlorates, and peroxides) have at least some $^{17}$O enrichment that can be traced back to ozone (Thiemens, 2006). Ozone is also an important photochemical species on Mars; it is plausible that the water and carbonate acquired their $^{17}$O enrichment from atmospheric oxidants like hydrogen peroxide or ozone reacting with the soil (Farquhar and Thiemens, 2000).

Oxygen isotopes in Lafayette and Nakhla indicate that water was equilibrated with carbonate – i.e. that carbonate formed at temperate conditions with the aid of water – but neither is equilibrated with silicate or sulfate (Karlsson et al., 1992; Farquhar and Thiemens, 2000; Farquhar and Johnston, 2008). Thus, the SNCs appear to derive from at least three separate oxygen reservoirs (Carr, 1996). The silicates represent bulk Mars. The water and carbonate carry excess $^{17}$O. This lack of interaction between the atmosphere and hydrosphere on the one hand and the rocky planet on the other is something that dates back at least as far as ALH 84001 (Farquhar and Johnston, 2008). It is an argument for an extremely dry planet (Carr, 1996). The sulfate carries an even greater excess of $^{17}$O. Farquhar and Johnston (2008) point out that the lack of isotopic interchange between sulfates and water is characteristic of abiotic systems.

(d) Carbon Oxygen and carbon are usually considered together because they are tightly entangled in CO$_2$. Unfortunately, the reported $\delta^{13}$C data vary widely. Ground-based telescopic observations find $\delta^{13}$C = $-22\pm20\%$ in atmospheric CO$_2$ today (Krasnopolsky et al., 2007). Viking measured $\delta^{13}$C = 0$\pm50\%$ (Bogard et al., 2001). The situation is exacerbated by differences in the composition of atmospheric CO$_2$ reported by Phoenix, $\delta^{13}$C = $-2.5\pm4.3\%$ (Niles et al., 2010), and MSL, $\delta^{13}$C = 46$\pm4\%$ (Mahaffy et al., 2013). Reported $\delta^{13}$C values in Martian meteorite carbonates also span a huge range, from $-20\%$ to +60$\%$ (Niles et al., 2010). Carbonates in the 4.1 Ga ALH 84001 and 1.4 Ga Nakhla carry heavy carbon with $\delta^{13}$C = +40$\%$e (Romanek, et al., 1994; Jull et al., 1999; Halevy and Eiler, 2011), while the younger Zagami (170 Ma) has $\delta^{13}$C = $-20\%$e (Jull et al., 1997; all fractionating escape and its timing and duration remain uncertain, but they are all fractionating to some degree and have played some role in producing the observed patterns. The implication for early Mars is that the atmosphere was thicker than it is today. How much thicker is difficult to ascertain from the isotope data alone, though one estimate based puts an upper limit of ~400 mbar during the Noachian (Cassata et al., 2012). In the next section we review the origin and evolution of the Martian atmosphere and place some constraints on its mass and composition.

17.4 ORIGIN OF THE EARLY MARS ATMOSPHERE

17.4.1 Accretion and Core Formation

The origin of the Martian atmosphere is intimately linked to the materials from which it formed, and how they were assembled to construct the planet. The standard model of Solar System formation has planets forming from condensed materials within the solar nebula (e.g. Lin, 1986). Thus, the primitive atmosphere might consist of gases contained in the solar nebula. However, as mentioned above, noble gases on Mars are significantly depleted with respect to solar abundances (Owen, 1992), so a nebular origin of its atmosphere is not likely. Instead, the early Martian atmosphere must have formed by some process that extracted volatiles from the solid materials from which the planet formed, from volatiles contained in the asteroids and comets delivered after the planet formed, or some combination of both.

The materials from which the planet formed likely had compositions similar to the primitive chondritic meteorites, which constitute the bulk of known meteorites (see Scott and Krot (2005) for an overview of their origins and compositions). They are thought to be among the oldest objects in the Solar System and are rich in the silicate minerals olivine and pyroxene, but also contain significant volatile materials such as water (in the form of hydrated minerals) and carbon (as both inorganic and organic compounds).

In the carbonaceous chondrites, there are refractory inclusions of Ca and Al that are believed to be the first materials to condense in the solar nebula. These calcium aluminum inclusions (CAIs) have been precisely dated to 4.5672 Ga$\pm0.6$ Ma (Amelin et al., 2002) and thus provide an accurate time for when the Solar System began forming. One class of carbonaceous chondrites (CI) has a composition nearly identical to the Sun’s photosphere, and therefore represents the unfractinated primitive material from which Solar System objects formed. The CI carbonaceous chondrites are often used to assess the degree of processing that Solar System objects have experienced.
Terrestrial planet formation models show three distinct phases of evolution beginning with the formation of kilometer-sized planetesimals, to lunar-sized planetary embryos, and finally to full-scale planets themselves (for a review see Lunine et al., 2011). Mars appears to have formed primarily through the accretion of planetesimals. Its small size implies that it escaped late-stage bombardment by lunar-sized objects, unlike the situation for Earth and Venus. This could have occurred if it accreted in a low-density region of the solar nebula (Chambers, 2001), suffered no collisions with planetary embryos by chance (Lunine et al., 2003), or migrated away from the embryo impact zone early in its history (Minton and Levison, 2011). Each of these scenarios implies that Mars would have a smaller inventory of volatiles compared to Earth, unless it formed locally from volatile-rich planetesimals (Drake and Righter, 2002).

As Mars grew, the energy released during accretion and the decay of short-lived radiisotopes (most notably $^{26}$Al) eventually became large enough to melt the planet, causing it to differentiate into a metallic core and silicate mantle. From chemical analyses of the SNC meteorites, which provide information about the Martian mantle (McSween, 1994), we now know that accretion and differentiation were complete within 20 Ma after the CAIs first condensed (Lee and Halliday, 1997; Kleine et al., 2002). The main evidence for this comes from the measured excess of $^{182}$W in the SNCs. The tungsten, which has an affinity for iron, is produced from the radioactive decay of $^{182}$Hf, which has an affinity for oxygen. The excess of $^{182}$W in the mantle implies that the core must have formed before most of it was produced. Otherwise, it would have followed iron into the core if differentiation occurred after it was produced. The half-life of $^{182}$Hf is 9 Ma. Using precise estimates of the mantle Hf/W ratio, Dauphas and Pourmand (2011) find that Mars reached half its present size in less than several million years. There also is evidence that the crust formed as early as 4.5 Myr. Theoretical predictions of rapid cooling from a possible magma ocean (Elkins-Tanton et al., 2005) and the breccia-like nature of the 4.4 Ga meteorite NWA 7533/7034 suggest that the bulk of the Martian crust formed within 100 Ma of Martian history (Humayun et al., 2013). Thus the core, mantle, and crust were largely in place very early in the planet’s history.

17.4.2 Primary Atmosphere

17.4.2.1 Magma Ocean and Steam Atmosphere

The rapid establishment of these three geochemical reservoirs has several implications for the early atmosphere and climate system. The first is that the short accretion time may have led to the formation of a magma ocean. While there is no observational requirement for a magma ocean, some melting of the planet must have occurred to permit differentiation. Indeed, some have argued that core formation is not possible without complete melting (e.g. Terasaki et al., 2005). The energy released by accreting planetesimals and the decay of short-lived radioactive isotopes is more than enough to melt the entire planet (see Elkins-Tanton et al., 2005). However, if a magma ocean did form, its energy must be retained against radiative loss to space. The most plausible way to do this is through the insulating effect of an optically thick atmosphere.

During accretion, volatiles will be released from the impacting planetesimals when their kinetic energy reaches a critical value. Water is the most common volatile in planetesimals of chondritic composition and it absorbs in many parts of the infrared spectrum (see Section 17.5.1). Thus a vapor atmosphere begins to develop. Whether it remains in the atmosphere or condenses on the surface depends on the accretion time since this determines the main energy input to the surface, and the infrared opacity of the atmosphere, since this determines the efficiency of heat retention. The accretion model of Matsui and Abe (1987) includes these processes. They found that impact devolatilization began on Mars when its radius grew to about 0.4 of its final value, and that the thermal blanketing effect of a steam atmosphere becomes sufficient to melt the surface and maintain a magma ocean if the accretion time is less than 5 Ma. At the time their paper was published, these times were deemed improbably short. However, as we have seen, these short accretion times are consistent with the Hf/W data.

17.4.2.2 Hydrodynamic Escape

Whether or not Mars had a steam atmosphere and a magma ocean, it is likely that its primitive atmosphere was hydrogen-rich. Impact devolatilization and outgassing driven by high early heat flows would have created an atmosphere with abundant water, which would have been photodissociated by the higher EUV output of the young Sun and/or used to oxidize reduced minerals in the surface to produce a hydrogen-rich atmosphere. Such an atmosphere would have been prone to hydrodynamic escape whose outflow could drag off other volatiles of relevance to the climate system such as CO$_2$ and N$_2$. Hydrodynamic escape has been invoked to explain the noble gas isotopic patterns on Mars and Earth (see Section 17.3.3), as well as the high D/H ratio on Venus (Kasting and Pollack, 1983). Thus, the intensity and duration of an early epoch of hydrodynamic Mars will have a direct influence on the inventory of climatically significant volatiles that remain when it ends.

There is evidence that hydrodynamic escape did occur on early Mars. The hydrogen blow-off will preferentially remove the lighter elements and thus fractionate isotopic patterns. As discussed above, xenon is strongly fractionated and this is consistent with an early epoch of hydrodynamic escape. Furthermore, the high $^{129}$Xe/$^{132}$Xe ratio implies that primordial $^{129}$Xe was stripped off the planet by an epoch of hydrodynamic escape that ended within several half-lives (15.7 Myr) of $^{129}$I (Zahnle, 1993). A longer period is possible if radiogenic contribution to the abundance of $^{129}$Xe by fission of $^{244}$Pu, which has a half-life of 82 Ma, is factored in (Jakosky and Jones, 1997).

---

8 Elements with an affinity for iron are referred to as siderophiles, those with an affinity for oxygen are lithophiles, while calcophiles are paired with sulfur.
17.4.3 Secondary Atmosphere

There is no direct evidence for the mass and composition of the atmosphere after the original atmosphere escaped. However, it was likely derived from magmatic outgassing and its composition therefore would depend on the chemistry and oxidation state of the mantle. One line of thinking (Figure 17.9) is that, because the core formed early, iron quickly separated from the volatiles, leaving the mantle mildly oxidized (Beaty et al., 2005). In this case the key volatiles (H, C, and S) would have come out mainly as H$_2$O, CO$_2$, and SO$_2$, which are excellent greenhouse gases. However, the SNC meteorites suggest more reducing conditions (Wadhwa, 2008), where H$_2$, CO, CH$_4$, and H$_2$S would also be in the mix. These too are good greenhouse gases. We shall return to the question of the redox state of the mantle and its outgassing products in Section 17.5.2.

Historically speaking, CO$_2$ and H$_2$O were thought to be the main constituents of the early Mars atmosphere. Yet their evolution is difficult to assess, since we do not really know the degassing history of early Mars, and both are subject to removal from the atmosphere by escape out the top and chemical reactions at the surface. The best we can do is to provide an upper limit on how much was available, assess potential loss rates, and then speculate on various evolutionary scenarios.

17.4.3.1 Inventory of CO$_2$ and H$_2$O

Estimates for the original amount of CO$_2$ and water have been made on the basis of scaling arguments, geochemical data, and geological landforms. Venus, for example, has 90 bar of CO$_2$ in its atmosphere, while Earth has 60–90 bar of CO$_2$ locked up mostly in the rock reservoir as carbonates. Thus, Earth and Venus have similar amounts of CO$_2$. If we assume that Venus, Earth, and Mars accreted from materials with similar volatile contents, then scaling by mass and accounting for the difference in gravity would give Mars an initial inventory of ~15 bar of CO$_2$.

The C/$^{84}$Kr ratio also suggests a large initial inventory of CO$_2$. Krypton is not subject to the escape processes acting since the end of the late heavy bombardment and has therefore probably been retained in the atmosphere since then. The C/$^{84}$Kr ratio is ~4×10$^{-7}$ for Venus and Earth, but is ~4.4–6×10$^{-6}$ for Mars. This implies that Mars had 10 times the amount of CO$_2$ in its atmosphere at the end of the late heavy bombardment than it does today. Today, the mean annual surface pressure on Mars is about 6 mbar. If we assume this represents the bulk of the CO$_2$ presently available on the planet today, then at the end of the late heavy bombardment the minimum amount of CO$_2$ would have been 60 mbar. Furthermore, $^{84}$Kr on Mars per unit mass is depleted by a factor of 100 with respect to Earth and Venus. If we assume that escape during and before the end of the late heavy bombardment accounts for the depletion, then it implies that Mars had an initial inventory of ~6 bar of CO$_2$, which is 2.5 times less but roughly consistent with the scaling argument made in the paragraph above.

The initial inventory of water has been estimated from geochemical arguments. Dreibus and Wänke (1987) proposed a two-component model of accretion that began with 150 km of water but ended with only 130 m. Unlike the Earth, these two components, one volatile-poor and the other volatile-rich, accreted homogeneously, which allowed the water to react with iron, generating vast amounts of hydrogen that escaped to space, leaving the mantle dry. Of the 130 m left in the mantle, Dreibus and Wänke estimated that only 10 m was outgassed to the atmosphere. This amount is far less than needed to account for the Noachian fluvial features (Carr and Wänke, 1992), and smaller than the ~20 m in the polar caps (Carr and Head, 2014). As we have seen, early core formation would reduce the time that volatiles had to mix with iron, so that the Dreibus and Wänke (1987) assumption of complete mixing may be overstated.

The amount of water remaining after an early epoch of hydrodynamic escape and impact erosion has been estimated from the volume of water needed to carve various fluvial landforms on the surface, estimates of the volume of water present in possible ocean basins, and estimates of the volume of water outgassed from volcanoes. Carr (1986) estimated that 425–475 m was present from the volume of water needed to cut the large outflow channels and from geomorphic indicators of near-surface ground ice. As discussed in Section 17.3.1, the former presence of oceans on Mars has been suggested by a number of investigators and several shorelines have been tentatively identified (Lucchitta et al., 1986; Parker et al., 1989; Baker et al., 1991; Clifford and Parker, 2001). If these are true shorelines, then a near-surface inventory of ~1 km is implied. Finally, Baker et al. (1991) proposed an ocean model to explain Amazonian glacial features and post-Noachian valley networks. In this model, further elaborated upon in Baker et al. (2000) and Baker (2001), post-Noachian Mars was mostly cold and dry, but experienced episodes of warm and wet conditions brought about by ocean-forming events linked to magmatic activity in Tharsis that changed the climate system. They estimate that the volume of water involved in these events could have been as high as ~2.5 km.

---

**Figure 17.9.** The oxidation state of outgassed materials transitions from reduced conditions (left) to mildly oxidizing conditions (right) after the core forms. From Catling and Claire (2005).
The existence of former oceans on Mars is still being debated. Carr and Head (2003) argued against the shoreline interpretation, citing their large range in elevations, the fact that some sections have a clear volcanic rather than sedimentary origin, and that the volumes of water corresponding to the larger oceans are implausible since they are comparable to the size of the Earth’s oceans. However, more recent developments favor the ocean hypothesis. The large-amplitude fluctuations in the elevations of the shorelines could be explained by true polar wander (Perron et al., 2007), and the enhanced potassium, thorium, and iron abundances within shoreline boundaries detected by the Gamma Ray Spectrometer (GRS) on Mars Odyssey are consistent with leaching and transport of these elements from the highlands to the lowlands (Dohm et al., 2009).

Whether or not oceans existed in the past, some ponding from the outflow flood events must have occurred. Thus to some extent the debate is more about the size of the ponded water, rather than the existence of an ocean per se. Certainly lakes and seas were likely in the past (see Section 17.3), and possibly even oceans. While most of the discussion has focused on post-Noachian times, it is relevant for our purposes in that it helps provide an estimate of how much water was available to the early Mars climate system. The volumes of water contained in the oceans proposed by the above investigators ranges from hundreds to thousands of meters global equivalent. As this represents only surface waters, it is a lower limit since it does not account for groundwater. Thus the inventory of water based on estimated ocean volumes is at least consistent with the geological estimates.

Thus the estimated initial inventory of CO₂ and water accreted by the planet is 6–15 bar of CO₂ and 150 km of water. How much of this initial inventory remained after the first 400 Ma is uncertain, but hundreds of meters to several kilometers of water must have been available during the Noachian to explain the fluvial landforms. This implies the loss of significant amounts of accreted water, which is consistent with the D/H data. For CO₂, much of its initial inventory must have been lost as well. The ⁴₀Ar/³⁶Ar in ALH 84001 suggests a Noachian atmosphere of <400 mbar of CO₂ (Cassata et al., 2012), and secondary crater statistics put an upper limit of ~1–2 bar depending on surface properties (Kite et al., 2014). As discussed below, there are a variety of loss mechanisms that can significantly reduce the initial inventory of CO₂.

17.4.3.2 Outgassing History

The outgassing history of Mars has been estimated from the volume of volcanic lavas that have erupted onto the surface, their ages, and their presumed volatile compositions. Greeley and Schneid (1991) published estimates of the volume and ages of volcanic material since the onset of the geological record at about 3.8 Ga ago, and Craddock and Greeley (2009) used these data to estimate that ~0.8 bar of CO₂ has been outgassed to the atmosphere through the Amazonian. The synthesis of these data suggest that outgassing peaked during the early Hesperian ~3.3 Ga, and was minimal during the mid- to late Noachian. However, erosional modification of Noachian terrains makes it impossible to estimate the true volume of volcanic material for that epoch. Consequently, while this approach is useful for the post-Noachian eras, it may greatly underestimate outgassing rates in the Noachian.

Phillips et al. (2001) estimated that 1.5 bar of CO₂ was outgassed during the Noachian. They based this estimate on the volume of magma that produced the Tharsis rise and assumed it had a composition similar to Hawaiian basalts. Follow-up work by Hirschmann and Withers (2008) and Grott et al. (2011) suggest much smaller outgassing totals. In particular, Grott et al. (2011) combined the solubility model of Hirschmann and Withers (2008) with a thermo-chemical evolution model to self-consistently calculate the amount of CO₂ dissolved in Martian magmas. Their calculations show that most of the outgassing occurs during the first few hundred million years and eventually ceased between 3.5–2 Ga depending on mantle oxidation state. They estimate that a total of ~1 bar of CO₂ was delivered to the atmosphere during this period, and that only ~250 mbar was outgassed between 4.1 and 3.7 Ga. These low outgassing volumes are contradicted by Wetzel et al. (2013), whose laboratory studies show that as much as 2.3 bar of carbon would have outgassed due to its greater solubility in the melt than previously assumed. Thus, there is a wide range of estimates for the outgassing volume.

However, none of these published estimates approach the 6–15 bar of CO₂ that outgassing volume. The reservoirs we can quantify include the atmosphere (~6 mbar, Haberle et al., 2008), the seasonal caps (~1 mbar, Kelly et al., 2006), the south polar residual cap (~0.1 mbar, Thomas et al., 2009), and the ice buried beneath it (5–6 mbar, Phillips et al., 2011). Together, these constitute less than 13 mbar of CO₂. The reservoirs whose volumes are difficult to quantify are the regolith and rock reservoirs. Zent and Quinn (1995) put an upper limit of ~40 mbar of CO₂ adsorbed in the regolith reservoir, though the presence of substantial volumes of near-surface ground ice at middle and high latitudes discovered later (e.g. Boynton et al., 2002) makes this seem generous. Carbonates have been detected in Martian dust and soil, the SNC meteorites, and regionally restricted outcrops, which could contain up to ~10 mbar of CO₂ (see Section 17.3.3, Ehlmann and Edwards, 2014). However, if the volume fractions of these detections (0.5–5%) are representative of the Martian crust, then 0.25–2.5 bar of CO₂ could be sequestered in a 1 km thick layer. Taken together, we can account for no more than several bars of CO₂, which implies that if the 6–15 bar estimate for the initial inventory is correct, then either outgassing was incomplete and the missing CO₂ is still in the magma, or outgassing was complete and substantial amounts of CO₂ have escaped to space.

17.4.3.3 Escape Mechanisms

Atmospheric escape is easier on Mars compared to Earth because of its smaller size. Also, the absence of an intrinsic magnetic field for most of Mars’ history facilitates loss mechanisms not possible on the Earth. We have already mentioned that Mars may have experienced an early episode of hydrodynamic escape and that the bulk of its volatiles could have been carried off in the subsequent outflow. However, other escape processes will operate when hydrodynamic escape ends. For the pre-Noachian, the most important are thermal escape and impact...
eroded; non-thermal escape rates are slow by comparison, and sputtering would not have begun until the magnetic field shut down at about 4.1 Ga (Hutchins et al., 1997). However, after 4.1 Ga, sputtering can remove as much as 90% of the post-Noachian atmosphere (Jakosky and Jones, 1997), and non-thermal escape can account for the observed enrichment in the $^{15}$N/$^{14}$N ratio (McElroy, 1972). A goal of the MAVEN mission now operating at Mars is to quantify present-day escape rates for these various mechanisms. (See Jakosky et al. (2015) for initial results, and Chapter 15 for details.)

Thermal escape (i.e., Jeans escape) occurs in the exosphere when molecules in the high-energy tail of the Maxwell distribution achieve velocities that exceed the escape velocities. Since molecular motions are proportional to temperature, the energy fluxes controlling those temperatures are important. For present-day Mars, thermal escape of molecules heavier than helium is negligible. However, on early Mars the higher EUV flux from the Sun could dramatically alter the temperature and altitude of the exobase. Using a one-dimensional thermosphere–ionosphere model that self-consistently calculates the ionization and excitation states of thermospheric constituents, Tian et al. (2009) found that the escape flux of carbon under the high EUV fluxes estimated by Ribas et al. (2005) ranged from $10^{12}$ molecule s$^{-1}$ at 4.5 Ga to $10^{11}$ molecule s$^{-1}$ at 4.1 Ga. Sustained fluxes of this magnitude would remove 1 bar of CO$_2$ in 1 Ma and 10 Ma, respectively. Though this model did not include H$_2$ and may therefore overestimate carbon escape (H$_2$ would take up some of the energy for escape), thermal escape of a thick CO$_2$ atmosphere during the period of high EUV fluxes on early Mars is likely to have been fast enough that even the most optimistic outgassing scenarios would have difficulty keeping pace with the loss to space. Under these circumstances, the only way to generate a thick CO$_2$ atmosphere before the end of the Noachian (3.7 Ga) is to outgas most of it after 4.1 Ga when the EUV fluxes declined enough for CO$_2$ to accumulate.

Impact erosion is another mechanism that makes it difficult to sustain a thick early atmosphere. Melosh and Vickery (1989) showed that large impactors could remove substantial atmospheric mass in the expanding vapor plume. If the impactor is large enough (~3 km for the current atmosphere) and fast enough (~14 km s$^{-1}$), it can remove the atmosphere above the tangent plane at the impact site. From the observed cratering record, they determined the flux of impactors capable of eroding the atmosphere and integrated it back in time to obtain estimates of the total amount of CO$_2$ that could be removed by this process. In this manner they estimated that the surface pressure on Mars 4.5 Ga was 60–100 times greater than it is today. Thus, ~350–600 mbar of CO$_2$ could have been eroded since the beginning of Martian history.

Brain and Jakosky (1998) elaborated on this model by focusing on the observed crater densities of the oldest terrains rather than the difficult-to-determine cratering rate as a function of time. They pointed out that it is the amount of erosion since the onset of the geological record during the late Noachian that is of interest rather than for all of the planet’s history, since it was at that time that the observed fluvial features formed. Based on the observed range of crater densities, they estimate between about 50% and 90% of the late-Noachian atmosphere has been removed by impacts, which implies an atmosphere with surface pressures of ~12–60 mbar. However, when they factored in sputtering losses, which they took to be 90%, the combination could remove 95–99% of the late-Noachian atmosphere, leaving therefore 1–5% in the present atmosphere. The 120–600 mbar surface pressures implied by this calculation represent a lower limit, since CO$_2$ could have migrated into non-atmospheric reservoirs (see above). These higher early surface pressures are supported by the modeling work of Hu et al. (2015), who match the MSL $\delta^{13}$C values through fractionation by photodissociation and sputtering modulated by carbonate formation.

### 17.4.3.4 Surface Sinks

As mentioned above, a potentially important surface sink for atmospheric CO$_2$ on Mars is conversion to carbonate rocks. On Earth, the long-term CO$_2$ content of the atmosphere is governed by the carbonate–silicate cycle schematically illustrated in Figure 17.10. CO$_2$ dissolves in rainwater producing a mildly acidic carbonic acid solution that leaches cations from silicate rocks on land surfaces and transports them into the ocean where they are eventually precipitated as carbonates onto the ocean floor by shell-forming organisms in the sea. Some of these carbonates are buried in sediments, subducted to great depths as the sea floor spreads, and eventually thermally decompose and release CO$_2$ gas back into the atmosphere through volcanoes. If oceans existed on early Mars that were maintained by an active hydrological cycle, carbonate precipitation should have occurred$^{10}$, and recycling through plate tectonics would not have been effective. Instead, hot spot volcanism (Pollack et al., 1987) or impact gardening (Carr, 1989) would have been the main recycling mechanisms. However, both would have declined at the end of the Noachian, at which point conversion of atmospheric CO$_2$ to carbonate would have been irreversible.

### 17.4.3.5 Summary

Mars formed quickly with accretion and core formation largely complete within the first several million years. Noble gas data rule out a primary atmosphere captured from the solar nebula. Instead, a steam atmosphere and possibly a magma ocean likely dominated the early environment. Much of this water was probably lost during a brief episode of hydrodynamic escape powered by enhanced EUV fluxes from the young Sun. A secondary atmosphere subsequently developed from volcanic outgassing. The mass and composition of this atmosphere are uncertain. Rapid core formation favors an atmosphere dominated by CO$_2$ and H$_2$O, though the meteorite data suggest more reduced gases (H$_2$, CO, CH$_4$). Estimates of their initial accreted volatile abundances range from 6 to 15 bar of CO$_2$ and up to 150 km for water. Although all indications are that most of the initial water escaped to space, enough (hundreds to a thousand meters) must have been present during the Noachian to form the observed fluvial features. And while there are a variety of loss mechanisms that can also limit the buildup of thick CO$_2$ atmospheres prior to 4.1 Ga, it is entirely plausible that, because outgassing rates

$^{10}$ Living organisms are not required to precipitate carbonates.
were higher, the Noachian atmosphere was thicker than it is at the present time. How much thicker is difficult to determine. Outgassing models put the number between 250 and 1500 mbar; a theoretical interpretation of the $^{40}$Ar/$^{36}$Ar ratio in ALH 84001 suggests up to 400 mbar; secondary crater statistics can accommodate 1–2 bar depending on surface properties; and a reducing mantle could contribute up to 2.3 bar. Thus, published estimates of the size of a CO$_2$ Noachian atmosphere range from 0.25 to 2.3 bar.

### 17.5 THE CLIMATE OF EARLY MARS

The processes affecting the early Mars’ climate are summarized in Figure 17.11. Some of these have already been discussed. In this section we discuss the potential mechanisms for warming early Mars. The principal challenge for greenhouse models is coping with the faint young Sun. Standard solar evolution models suggest that, when the Solar System formed at 4.56 Ga, the Sun was less luminous than it is today (Newman and Rood, 1977; Gough, 1981). The luminosity of Sun-like stars increases with time. The fusion of hydrogen into helium increases the mean molecular weight of the core. To maintain the balance between the pressure gradient force and gravity, the core contracts and warms. The increased densities and temperatures increase the rate of fusion and hence the star’s luminosity increases with time. A good fit to this increase with time for standard solar models is given by (Gough, 1981)

$$L(t) = \frac{L(t_0)}{1 + \frac{1}{2}(t/t_0)}$$  \hspace{1cm} (17.2)

where $L$ is luminosity, $t$ is time, and $t_0$ is the age of the Solar System (4.56 Ga).

Figure 17.12 shows the solar luminosity and the planet’s effective temperature, $T_e$, as functions of time. The latter is defined as the temperature at which a black body radiates away the total energy it absorbs. If $A_p$ is the planetary albedo and $S_o$ is the annual mean solar flux at Mars’ orbit, then $T_e$ is given by

$$T_e = \frac{(1 - A_p)S_o}{4\sigma}^{1/4}$$  \hspace{1cm} (17.3)

where $\sigma$ is the Stefan–Boltzmann constant ($5.67 \times 10^{-8}$ W m$^{-2}$ K$^{-4}$). In Figure 17.12 we assume that $A_p = 0.25$ and is constant in time, and that $S_o$ for Mars today is 590 W m$^{-2}$. Thus, today the planet’s effective temperature is 210 K, but during the Noachian epoch (3.7–4.1 Ga), its effective temperature under these assumptions would have been 195 K. To raise global mean annual surface temperatures to the melting point of water thus requires the atmosphere to produce 78 K of greenhouse warming. Considering that the Earth’s atmosphere provides only 33 K of greenhouse warming (Table 17.3), this presents a challenge. However, it is not impossible, as Venus’ atmosphere produces over 500 K of warming (though with 90 bar of CO$_2$)!

Attempts to resolve the faint young Sun problem for Mars depend on what the required surface temperatures actually are. Unfortunately, the geological and mineralogical data are ambiguous on this point. Traditionally, a global and annual averaged surface temperature, $T_s$, of at least 273 K has been the assumed requirement. While this makes some sense, there is little discussion in the literature about the rationale. If, instead, the requirement is that mean annual temperatures at specific locations be above freezing, then this constraint could be relaxed. It could be further relaxed if these surface temperatures need only be above freezing seasonally or daily, and further yet if brines are permitted. Thus, it is not clear what requirements we are demanding from an early Mars greenhouse effect. What is clear, however, is that surface temperatures must be above freezing when liquid water is present.

A conceptual argument can be made that $T_s$ must be close to or above 273 K if the valley networks, eroded terrains, and hydrated mineral deposits require a long-lived continuously warm and wet Noachian climate with an active hydrological cycle involving rainfall and runoff. If this is the case, then large bodies of open water comparable to seas or oceans must exist. Scattered lakes or ponds could not sustain such a climate system (Soto et al., 2010). These seas/oceans would exist in the northern lowlands, as has been suggested (see Section 17.3.1), which means that mean annual temperatures in most
Table 17.3. Approximate surface and effective temperatures of the terrestrial planets.

<table>
<thead>
<tr>
<th>Planet</th>
<th>$S_0/4$ (W m$^{-2}$)</th>
<th>$A_p$</th>
<th>$T_w$ (K)</th>
<th>$T_e$ (K)</th>
<th>$T_w - T_e$ (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Venus</td>
<td>657</td>
<td>0.75</td>
<td>740</td>
<td>232</td>
<td>508</td>
</tr>
<tr>
<td>Earth</td>
<td>342</td>
<td>0.30</td>
<td>288</td>
<td>255</td>
<td>33</td>
</tr>
<tr>
<td>Mars</td>
<td>148</td>
<td>0.25</td>
<td>215</td>
<td>210</td>
<td>5</td>
</tr>
</tbody>
</table>

Figure 17.11. Cartoon summary of the processes acting on early Mars. Graphic courtesy of Christina Olivas. A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.

Figure 17.12. Solar luminosity (solid line, left axis) and effective temperature (dashed line, right axis) as functions of time.
of the northern hemisphere would have to be above freezing, otherwise a frozen surface would raise the albedo, lower surface temperatures, reduce evaporation rates, and slow the hydrological cycle. Similarly, mean annual temperatures in the southern hemisphere, where most of the valley networks are found, would also have to be above freezing to permit rainfall and runoff. Thus, if long-lived continuously warm and wet conditions are required, then the annual and globally averaged surface temperature \( T_m \) must be close to or above freezing. Somewhat lower temperatures are possible if the oceans become salty and/or only partially freeze over. How much lower is difficult to quantify. But if rainfall is truly required, mean annual global average surface temperatures cannot be too much below 273 K.

Much of the present debate about early Mars hinges on this issue. If long-lived warm wet conditions are required throughout the Noachian, then large bodies of open water are required, and the atmosphere must produce and sustain a greenhouse capable of raising global mean annual surface temperatures close to the melting point of water. Alternatively, the Noachian climate could have been mostly cold and dry with \( T_m < 273 \) K, but punctuated with transient periods of globally warm and wet conditions brought about by external forcings such as impacts or volcanism. Finally, it is possible that global mean annual surface temperatures were always less than 273 K and that warm wet conditions were only needed at certain places and certain times. Thus, after consideration of climate processes in Section 17.5.1, we divide this part of our review into three categories: long-lived greenhouse atmospheres (Section 17.5.2), transient greenhouse atmospheres (Section 17.5.3), and cold climates with locally wet conditions (Section 17.5.4). A fourth possibility, which we also consider, is that the luminosity of the early Sun did not follow the standard solar model and the Sun was actually bright enough to warm early Mars with only modest changes to the mass and composition of the atmosphere. This idea is more speculative, but since it has been explored in the literature we include it here for completeness (Section 17.5.5).

### 17.5.1 Climate Processes

#### 17.5.1.1 Radiative Transfer

Because of its high temperature, the Sun radiates most of its energy at visible wavelengths, while planets, being much cooler, emit mainly in the thermal infrared. This provides a clean separation for assessing planetary energy budgets. Planets gain energy by absorbing the sunlight intercepted by their cross-sectional area, and lose energy by emitting in the infrared over their entire surface. In the absence of an atmosphere, \( T_e \) as defined by (17.3) is equivalent to the global and annual mean effective surface temperature\(^{11} \) \( T_{se} \) and the total outgoing long-wave radiation (OLR) is \( q_{se} \) (Haberle, 2013). However, if we introduce to such a planet an atmosphere that is transparent at solar wavelengths but opaque in the infrared, the OLR will be reduced by an amount depending on the concentration, distribution, and absorption properties of those gases. Consequently, it will be absorbing more energy than it is emitting and the temperature of its atmosphere and surface will increase until a new balance is established. Though there are many subtle features that must be considered, this is the essence of the greenhouse effect. It explains why the surface temperature can be higher than the effective temperature, and it gives us a convenient measure of the greenhouse power of a given atmosphere (Haberle, 2013).

Table 17.3 lists the difference between the effective temperature, and it gives us a convenient measure of the greenhouse power of a given atmosphere (Haberle, 2013). Table 17.3 lists the difference between the effective temperature, and it gives us a convenient measure of the greenhouse power of a given atmosphere (Haberle, 2013). Table 17.3 lists the difference between the effective temperature, and it gives us a convenient measure of the greenhouse power of a given atmosphere (Haberle, 2013). Table 17.3 lists the difference between the effective temperature, and it gives us a convenient measure of the greenhouse power of a given atmosphere (Haberle, 2013). Table 17.3 lists the difference between the effective temperature, and it gives us a convenient measure of the greenhouse power of a given atmosphere (Haberle, 2013). Table 17.3 lists the difference between the effective temperature, and it gives us a convenient measure of the greenhouse power of a given atmosphere (Haberle, 2013). Table 17.3 lists the difference between the effective temperature, and it gives us a convenient measure of the greenhouse power of a given atmosphere (Haberle, 2013).

### Table 17.4. Fundamental vibrational frequencies of common greenhouse gases (cm\(^{-1}\)).

<table>
<thead>
<tr>
<th>Gas</th>
<th>( \nu_1 )</th>
<th>( \nu_2 )</th>
<th>( \nu_3 )</th>
<th>( \nu_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO(_2)</td>
<td>1388(^a)</td>
<td>667</td>
<td>2349</td>
<td></td>
</tr>
<tr>
<td>H(_2)O</td>
<td>3657</td>
<td>1595</td>
<td>3756</td>
<td></td>
</tr>
<tr>
<td>SO(_2)</td>
<td>1152</td>
<td>518</td>
<td>1362</td>
<td></td>
</tr>
<tr>
<td>H(_2)S</td>
<td>2615</td>
<td>1183</td>
<td>2626</td>
<td></td>
</tr>
<tr>
<td>CH(_4)</td>
<td>2917(^a)</td>
<td>1534(^a)</td>
<td>3019</td>
<td>1306</td>
</tr>
<tr>
<td>N(_2)O</td>
<td>2224</td>
<td>589</td>
<td>1285</td>
<td></td>
</tr>
<tr>
<td>NH(_3)</td>
<td>3337</td>
<td>950</td>
<td>3444</td>
<td>1627</td>
</tr>
</tbody>
</table>

\(^a\) These transitions do not change the dipole moment and are therefore infrared-inactive.

The main challenge is finding plausible greenhouse gases that can plug up the gaps in the OLR as a function of wavelength. The most common greenhouse gases that have been considered for early Mars include carbon dioxide (CO\(_2\)), water (H\(_2\)O), sulfur dioxide (SO\(_2\)), hydrogen sulfide (H\(_2\)S), methane (CH\(_4\)), nitrous oxide (N\(_2\)O), and ammonia (NH\(_3\)). In the infrared, these molecules absorb photons that mainly change their vibrational–rotational states. These changes occur at discrete frequencies that are listed in Table 17.4. Figure 17.13 shows their infrared absorption cross-sections as a function of wavelength. The top panel is the Planck function for a surface temperature of 273 K. If an early Mars greenhouse produced global mean surface temperatures of 273 K, this plot shows the spectral regions in the infrared where greenhouse gases need to be absorbing. Ideally, the absorption should be strongest in the 20–1200 cm\(^{-1}\) region where about 85% of the energy is radiated to the atmosphere by the ground.

The second panel shows the absorption cross-section for CO\(_2\) and H\(_2\)O gases that could have been present in the early Mars atmosphere in climatically significant amounts. CO\(_2\) is a linear molecule and has a much richer set of absorption features throughout the infrared. Its vibrational–rotational modes allow it to contribute significant absorption near 200 cm\(^{-1}\) and 1500 cm\(^{-1}\). The combination of these two gases can provide considerable opacity throughout the infrared, except in the so-called “window” region between 800 and 1200 cm\(^{-1}\). Since roughly 25% of the energy radiated by a 273 K surface is in the window, gases that can provide additional opacity in this

\(^{11} T_{se} \) is the temperature at which a black body emits the equivalent amount of radiation emitted by the entire surface over a year.
The Early Mars Climate System

region could significantly boost the greenhouse effect of a CO₂–H₂O atmosphere.

The bottom panel of Figure 17.13 shows absorption cross-sections for other potential greenhouse gases. By far, NH₃ has the greatest potential to plug up the window region, as it exhibits significant absorption from 700 to 1300 cm⁻¹. Other gases that can contribute opacity in the window are N₂O, SO₂, and H₂S. Outside the window, water vapor dominates the absorption spectra, though SO₂, N₂O, and CH₄ can contribute if present in sufficient quantities. Based on the Planck-weighted mean cross-sections listed in Table 17.5, on a per molecule basis water vapor is the most effective greenhouse gas and H₂S is the least effective; NH₃, SO₂, and CO₂ are major contributors, while N₂O and CH₄ have only minor effects. Of course, the actual greenhouse potential of these greenhouse gases will depend on the thermal structure of the atmosphere, their abundance and vertical distribution, and their lifetimes (i.e. sources and sinks). More sophisticated calculations are therefore needed to determine their real greenhouse potential.

It is important to note that the ability of these gases to absorb at the listed frequencies is enhanced by a number of processes that broaden their lines. The most familiar are “pressure” and “Doppler” broadening. The former is due to collisions between molecules, while the latter is the result of molecular motion. Thus, the enhanced absorptions produced by these mechanisms depend on pressure and temperature. However, at high enough pressures, collisions can also induce absorption. This collision-induced absorption (CIA)

---

Figure 17.13. (Top panel) The Planck function for a 273 K surface. (Middle and bottom panels) The infrared absorption cross-sections for the gases listed in Table 17.4. A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.
Table 17.5. Planck-weighted (273 K) cross-sections.

| Gas     | Mean cross-section  
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(10^{-20} cm^2 molecule^{-1})</td>
</tr>
<tr>
<td>H2O</td>
<td>2.48</td>
</tr>
<tr>
<td>NH3</td>
<td>2.06</td>
</tr>
<tr>
<td>SO2</td>
<td>1.35</td>
</tr>
<tr>
<td>CO2</td>
<td>0.99</td>
</tr>
<tr>
<td>N2O</td>
<td>0.43</td>
</tr>
<tr>
<td>CH4</td>
<td>0.14</td>
</tr>
<tr>
<td>H2S</td>
<td>0.13</td>
</tr>
</tbody>
</table>

is complicated and not well understood, but generally results from temporary dipoles, forbidden transitions, or short-lived dimers. It should be emphasized that CIA is distinctly different from pressure broadening in that new lines appear, which absorb at different frequencies. For pure CO2 atmospheres, CIA becomes important for surface pressures exceeding several hundred millibars with absorptions occurring in the 0–250 cm\(^{-1}\) and 1200–1500 cm\(^{-1}\) regions (Wordsworth et al., 2010). For H2-rich atmospheres, CIA may also be important. Sagan (1977) first proposed the idea as a solution to the faint young Sun problem for both Earth and Mars, and it has been recently revived by Wordsworth and Pierrehumbert (2013) for the Earth, and by Ramirez et al. (2014) for Mars (see below).

In addition to the discrete line absorptions described above, gases often exhibit “continuum” absorption. Water vapor has this property. Its continuum absorption has been widely studied because of its relevance to the Earth’s climate system. Laboratory measurements clearly show that there is more absorption in the window region (800–1200 cm\(^{-1}\)) than can be accounted for from nearby lines. The cause for this excess absorption is uncertain. Super-Lorentzian behavior (meaning that the line shapes differ from the shape commonly assumed and produce more absorption) in the far wings of the lines is the prevailing view, but the theoretical and observational basis for this is weak. The approach to including water vapor continuum absorption has therefore been empirical and is based on laboratory measurements and direct observations.

An extensive body of literature exists that describes the theoretical and experimental basis for these mechanisms and how they can be applied to radiative transfer calculations. Textbook discussions can be found in Petty (2006) and Pierrehumbert (2010), while more Mars-specific applications are presented in Halevy et al. (2009) and Mischna et al. (2012). The reader is referred to these works for further details.

### 17.5.1.2 Convection

In addition to greenhouse warming, convection in the lower atmosphere plays an important role in regulating surface temperatures. Convection sets the lapse rate, which is the rate of decrease of temperature with altitude, and therefore controls the thermal structure of the troposphere. Radiative processes and surface heat exchange tend to destabilize the lower atmosphere and produce lapse rates that exceed the adiabatic lapse rate, the maximum lapse rate in a stable hydrostatic atmosphere. Convection acts to restore the adiabatic lapse rate by transporting heat from the lower atmosphere to the upper atmosphere through turbulent motions. Thus, convection cools the near-surface atmosphere and warms the upper atmosphere, with the amount of cooling or warming depending on the moisture content of the atmosphere. For the same initially unstable temperature profile, convection in dry atmospheres produces less surface cooling than convection in moist atmospheres because the latent heat released in rising air parcels containing enough water vapor to condense provides an additional heat source. Consequently, the moist adiabatic lapse rate is not as steep as the dry adiabatic lapse rate, i.e. temperatures fall off more slowly with height. This is also the case for dry atmospheres, where the main constituent can condense. In this instance, the lapse rate follows the frost-point temperature. This is the situation for the Martian atmosphere, which today is 95% CO2, and for early Mars may have been an even higher percentage.

### 17.5.1.3 Horizontal Heat Transport

Horizontal heat transport also plays a role in regulating surface temperatures, though in this case it is the spatial distribution of temperatures that is mostly affected rather than global mean temperatures. However, horizontal heat transport can affect global mean temperatures through feedbacks such as the ice-albedo feedback. For planets like Mars, where more energy is absorbed at low latitudes compared to high latitudes, horizontal wind systems will transport heat toward the poles, thereby lowering tropical temperatures and raising polar temperatures. The nature of the wind systems so produced and the strength of the transport they generate depends on a variety of factors, including the solar forcing, mass and composition of the atmosphere, and orbit properties (obliquity and rotation rate in particular).

For early Mars the most important factors for determining the degree to which transport can warm polar temperatures are the mass and composition of its atmosphere, since these will determine the partitioning of condensables between the atmosphere and polar caps. Mars’ atmosphere today, for example, is too thin to transport enough heat into the polar regions to prevent CO2 from condensing onto the surface (Leovy and Mintz (1969) were the first to show this). This weak transport is part of the reason that CO2 ice at the south pole survives all year long. The existence of a permanent cap on Mars is critical since its heat balance will determine the mean annual surface pressure (see Chapter 12 for a review). However, as the mass of the atmosphere increases, more and more heat will be transported into the polar regions and the greenhouse effect will be strengthened. At some point, permanent polar caps will not be possible and all the CO2 in the system will exist in the atmosphere.

This is a critical issue for early Mars since it means there is a minimum surface pressure, below which greenhouse warming and heat transport are insufficient to prevent the atmosphere from collapsing into a permanent polar cap. The value of that minimum surface pressure depends on the solar luminosity, the albedo of the cap, and the planet’s orbit properties. Haberle
et al. (1994), Kreslavsky and Head (2005), Manning et al. (2006), and Soto et al. (2011) have studied this issue. It is difficult to state with any certainty, since it depends on what other absorbers might be present, but based on these studies surface pressures on the order of at least several hundred millibars and possibly much higher are needed to stabilize the atmosphere of early Mars with present-day orbital conditions.

### 17.5.1.4 Climate Models

Climate models do not use the line-by-line approach to radiative transfer because of the complexity of the line distribution, which would require an overwhelming amount of computational time to find solutions from it. Consequently, alternative, more efficient, methods are used to calculate spectrally integrated radiative fluxes in the bands of interest, such as the 15 µm CO₂ band. There are a variety of approaches to radiative transfer in climate models, ranging from simple gray models, to band models, to exponential sums and correlated-k techniques that offer fast and reasonably accurate methods to compute temperature profiles.

Climate models also have a range of spatial resolution. One-dimensional globally averaged models were the main tool for the initial studies during the last several decades of the 20th century (e.g. Pollack, 1979). These models sought equilibrium solutions to a given radiative forcing and included convective adjustment in the lower atmosphere. Recent versions of these one-dimensional models use many spectral bands and include gaseous and aerosol scattering. With the increase in computing technology, most notably the speed and memory of modern machines, full three-dimensional general circulation models with quite sophisticated physics packages are now being utilized to study this problem (e.g. Wordsworth et al., 2013).

### 17.5.2 Long-Lived Greenhouse Atmospheres

#### 17.5.2.1 CO₂–H₂O Atmospheres

Carbon dioxide and water vapor are the most extensively studied greenhouse gases for early Mars (see Haberle (1998) for a review of the earlier work). These gases are plentiful, relatively stable, and form the basis of greenhouse models of Earth, runaway Venus, and early Mars. As shown in Figure 17.13 they can provide significant infrared opacity if present in large enough quantities. The first modeling studies of these atmospheres indicated that between 5 and 10 bar of CO₂ was needed to raise temperatures to the melting point of water in the presence of the faint young Sun (Pollack, 1979; Pollack et al., 1987). The main problem with these atmospheres is that at some point the atmosphere saturates and CO₂ condenses. Kasting (1991) pointed out this problem and illustrated the consequences with a one-dimensional radiative–convective model. As surface pressures increase, CO₂ begins to condense in the atmosphere. The associated release of latent heat pins temperatures to the frost point and the lapse rate begins to follow a moist CO₂ adiabat instead of a dry adiabat. As a consequence, the upper atmosphere warms and the surface must cool to maintain energy balance. As surface pressures continue increasing, more and more of the atmosphere saturates and eventually the entire atmosphere follows the moist CO₂ adiabat. At this point, further increases in CO₂ simply result in precipitation to the surface. For a solar luminosity appropriate to conditions 3.8 Ga (i.e. 75% of today’s value), this ultimate limit was reached in Kasting’s model at ~2 bar of CO₂ and the corresponding surface temperature was ~220 K.

Another setback for CO₂–H₂O atmospheres is the finding that CIA appears to be weaker than previously calculated (Figure 17.14). The standard parameterization for CIA comes from the work of Kasting et al. (1984) and this has been used in a number of studies of early Mars’ greenhouse potential (Pollack et al., 1987; Kasting, 1991; Forget and Pierrehumbert, 1997; Mischna et al., 2000). However, this parameterization was found to overestimate CIA in pure CO₂ atmospheres (Halevy et al., 2009; Wordsworth et al., 2010). As shown in Figure 17.14, absorption due to the induced-dipole transitions in the 250–500 cm⁻¹ range was greatly overestimated. Models using this parameterization have therefore overestimated the greenhouse effect. Wordsworth et al. (2010) showed that global mean surface temperatures may have been overestimated by as much as 20–30 K at high surface pressures with the original parameterization.

The Kasting (1991) model has since been updated (see Kopparapu et al., 2013). The most recent update includes new absorption coefficients for CO₂ and H₂O as well as the CIA parameterization suggested by Wordsworth et al. (2010). The model now predicts a maximum surface temperature of 230 K at 3.7 Ga for a pure CO₂–H₂O atmosphere (Ramirez et al., 2014). Figure 17.15 shows the latest results for four different values of the solar luminosity. For solar luminosities less than 80% of the present-day value, which corresponds in time to ~2.85 Ga, Ramirez et al.’s new model cannot raise surface temperatures to the melting point of water. The reduced temperatures at high pressures predicted by Kasting’s latest model are due to the new CIA parameterization, which provides substantially less absorption than the old parameterization. This makes it even more difficult to warm early Mars with a pure CO₂–H₂O atmosphere.

However, the potential for CO₂ condensation complicates the picture. CO₂ clouds not only modify air temperatures through latent heat release, but also interact with the solar and infrared radiation fields in ways that depend on their optical properties. Kasting (1991) speculated that, because CO₂ ice clouds are much more transparent in the infrared than water ice clouds, they would have an additional net cooling effect since the clouds should be highly reflective in the visible. Thus, the clouds would not only reduce the greenhouse effect by decreasing the lapse rates, but also increase the planetary albedo and provide little infrared opacity to cloud upwelling surface emission. Implicit in this line of reasoning is the assumption that the cloud particles would be too small to interact with infrared radiation. However, one unique aspect of CO₂ ice clouds on Mars is that during condensation the flow of gas to growing ice crystals will be hydrodynamic rather than diffusive (Rossow, 1978). This is likely to result in the rapid growth of particles large enough (10–100 µm) to efficiently scatter infrared radiation before they fall to the surface (Pierrehumbert and Erlick, 1998).

To assess this possibility, Forget and Pierrehumbert (1997) modified Kasting’s (1991) model to include reflective CO₂ ice clouds and found that optically thick clouds containing 10
17.5 The Climate of Early Mars

Figure 17.14. (a) Extinction for a 1 bar CO$_2$ atmosphere at 273 K as a function of wavenumber. (b) Comparison of the updated CIA extinction (solid lines) with the older parameterization of Kasting et al. (1984) (crosses) for several different temperatures. Note the much stronger absorption in the old parameterization in the 250–500 cm$^{-1}$ range. From Wordsworth et al. (2010).

Figure 17.15. Global mean surface temperature as a function of surface pressure. The different lines correspond to different values of the solar luminosity $S/S_0$, where $S_0$ is the present value. From Ramirez et al. (2014).
The viability of the scattering greenhouse depends on the concentration and particle size of the ice particles. Poorly understood CO$_2$ microphysical processes control these. While models have begun simulating the behavior of CO$_2$ ice clouds (e.g. Forget et al., 1998; 2013; Mischna et al., 2000; Colaprete and Toon, 2003; Colaprete et al., 2003; Wordsworth et al., 2013), very few laboratory studies have been conducted to measure their nucleation and growth properties (Glandorf et al., 2002). Thus, there is uncertainty in the net radiative effect of CO$_2$ ice clouds. Furthermore, if the opacity and sizes do produce surface warming, then the atmosphere will also warm and this may limit cloud concentrations and surface warming. In their fully coupled self-consistent simulations, Colaprete and Toon (2003) obtained a modest 5–10 K surface warming from the scattering greenhouse effect because of this effect.

Spatial coverage is another issue for CO$_2$ ice clouds. Forget and Pierrehumbert (1997) assumed 100% cloud coverage, reasoning that clear skies might be more difficult to achieve when the major constituent of the atmosphere is condensing. However, they recognized the importance of fractional cloud cover, as simulations with 75% cloud coverage were 20–30 K colder than with complete coverage. This issue is best studied with full three-dimensional global circulation models, as dynamical processes will have a strong influence on cloud coverage. To date, only a few such studies have appeared in the peer-reviewed literature (Forget et al., 2013; Wordsworth et al., 2013). While it does appear that CO$_2$ ice clouds cover a major part of the planet in these models, their optical depths are not high enough to raise surface temperatures by more than 15 K even under the most favorable circumstances (Forget et al., 2013). Thus, CO$_2$ ice clouds do not provide as much warming as originally thought.

Thus, for a variety of reasons, CO$_2$–H$_2$O atmospheres are not as promising as they once were. To date, no model is able to generate a steady-state climate system with a pure CO$_2$–H$_2$O atmosphere capable of producing global mean annual surface temperatures at or above the melting point. However, radiatively active water ice clouds in such atmospheres could, in principle, enhance the greenhouse effect. Urata and Toon (2013) show this using an Earth global climate model modified for Mars conditions. Under the right conditions – large cloud particles, reduced precipitation, and complete cloud coverage in grid cells – mean surface temperatures near the melting point can be achieved. Segura et al. (2008) first recognized the potential for a cloud greenhouse effect on Mars in their studies of post-impact environments on early Mars. Also it can significantly raise surface temperatures at times of high obliquity (Haberle et al., 2012; Madeleine et al., 2013). However, the LMD (Laboratoire de Météorologie Dynamique) model of Wordsworth et al. (2013), which also includes radiatively active clouds, does not find a significant role for a water ice cloud greenhouse. Thus, while the potential for a significant cloud greenhouse on early Mars is intriguing, it is model-dependent and therefore needs further study.

### 17.5.2.2 Supplemental Greenhouse Gases

The difficulty of raising surface temperatures to the melting point of water with pure CO$_2$–H$_2$O atmospheres has focused attention on finding additional trace gases to boost the greenhouse effect. Sulfur dioxide (SO$_2$), methane (CH$_4$), ammonia (NH$_3$), and nitrous oxide (N$_2$O) have all been mentioned in the literature, but only for SO$_2$ has there been significant published work on this topic.

Sulfur dioxide absorbs in parts of the spectrum not covered by CO$_2$ and H$_2$O (Figure 17.13). Its strongest absorption feature is due to the v$_1$ vibrational transitions near 7 μm. Postawko and Kuhn (1986) considered high levels of SO$_2$ (up to 1000 ppmv) in combination with a thick CO$_2$–H$_2$O atmosphere but were unable find solutions with global mean surface temperatures at 273 K. This negative result, and the fact the SO$_2$ is photochemically unstable, soluble in water, and readily oxidized to sulfate, did not motivate much further research at the time. One paper suggested that small amounts of SO$_2$ (0.1 ppmv) might stave off the CO$_2$ condensation problem (Yung et al., 1997). However, with the discovery of Noachian/Hesperian-aged sulfate deposits by the MER Rovers, and later mapped by OMEGA and CRISM, interest in SO$_2$ has been renewed, particularly since SO$_2$ could have acidified surface waters and inhibited carbonate formation (see Section 17.3.2).

While there are ways to form sulfates that do not involve the atmosphere, volcanic emission of reduced sulfur species (mainly SO$_2$ and H$_2$S) must certainly occur, and since Mars was much more volcanically active during the late Noachian and early Hesperian, an atmospheric delivery of sulfur to the surface is inevitable (e.g. Settle, 1979). The early Martian mantle is thought to be rich in sulfur (Wänke and Dreibus, 1994) such that outgassing during intense periods of volcanism could have emplaced enough SO$_2$ in the atmosphere to raise surface temperatures at least to levels where brines could exist (~250 K), or even temporarily well above the freezing point. The three-dimensional greenhouse model of Johnson et al. (2008) suggests this possibility. Their 500 mbar CO$_2$–H$_2$O atmosphere simulation, for example, gives global mean surface temperatures of 283 K and 315 K for SO$_2$ mixing ratios of 6.14 and 245 ppmv, respectively. Unfortunately, the warming estimated by this model appears to be overestimated (Mischna et al., 2013). And there are other reasons, discussed below, that warming global mean temperatures to the melting point is not likely to be correct for early Mars.

An SO$_2$ greenhouse strong enough to sustain liquid water on the surface in a cloud- and aerosol-free CO$_2$-dominated atmosphere requires high surface pressures and elevated SO$_2$ levels. The values required are model-dependent, but are in the ranges 0.5–4 bar of CO$_2$ and 1–100 ppmv of SO$_2$ (Halevy et al., 2007; Johnson et al., 2008; Tian et al., 2010). Whether SO$_2$ concentrations at this level are sustainable depends on the details of an early Martian sulfur cycle. Volcanic outgassing rates must be high enough to balance atmospheric losses. The main sinks
are rainout, gas-phase reactions, and photolysis. Halevy et al. (2007) proposed a sulfur cycle that operated analogously to the CO$_2$ cycle on Earth. They focused particularly on the late Noachian when Tharsis volcanism could have sustained sulfur outgassing rates twice those on present-day Earth. In their model, the generally reducing nature of the outgassed sulfur (as SO$_2$ and H$_2$S) would have exhausted the supply of oxidants and therefore limited the removal of SO$_2$ by gas-phase oxidation reactions. Oxidants on Mars are created when hydrogen escapes to space, and they reasoned that this was a relatively slow process compared to the faster supply of reductants from volcanoes. They further argued that Rayleigh scattering limited loss by photolysis in a thick CO$_2$ atmosphere. Rainout would not be significant until temperatures approached the melting point. Thus, with the main sinks of atmospheric SO$_2$ suppressed, its concentration could build until the greenhouse warming produced liquid water, thereby accelerating its removal and limiting further increases. This feedback on atmospheric SO$_2$ is similar to how the Earth’s carbonate–silicate cycle regulates the long-term concentration of CO$_2$ in the atmosphere.

Some elements of this scenario are supported by the one-dimensional photochemical studies of Johnson et al. (2009) and Tian et al. (2010). Photolysis rates at low levels do significantly decline in thick CO$_2$ atmospheres (because of Rayleigh scattering) and this does stabilize SO$_2$ in the lower atmosphere. The abundance of oxidants also declines, which further limits the sink of SO$_2$ due to gas-phase reactions. However, these are not the main sinks for SO$_2$ in these models. Rainout is the main removal mechanism and this limits the lifetime of SO$_2$ to hundreds, perhaps thousands, of years. This is a much longer lifetime than that estimated by Wong et al. (2004) for the present Martian atmosphere (~0.5 year). Johnson et al. (2009) concluded that the enhanced lifetime of SO$_2$ on early Mars meant that volcanic outgassing events would be followed by transient periods of warm wet conditions that could help carve the observed fluvial features.

In practice, however, an SO$_2$ greenhouse still faces significant challenges. The main challenge is avoiding the production of sulfur-bearing aerosols that inevitably cool the planet (Tian et al., 2010). These were not considered by Halevy et al. (2007) and Johnson et al. (2008, 2009). On Earth, SO$_2$ is converted to sulfurous acid (H$_2$SO$_3$) by the oxidation of SO$_2$ to SO$_3$ (by reaction with OH), which then reacts with water to form H$_2$SO$_4$. The sulfuric acid then condenses to form sulfate aerosols, which are highly reflective at visible wavelengths. The cooling effect of sulfate aerosols is a well-known consequence of volcanic eruptions on Earth (e.g. Hansen et al., 2002), and the sulfuric acid clouds on Venus (e.g. Bullock and Grinspoon, 2001). On early Mars, OH might be limited by the supply of reductants from volcanic emissions as Halevy et al. (2007) suggest. However, there is an additional supply of oxidants on Mars that is not present on Earth. Carbon dioxide, the assumed main constituent of the early Mars atmosphere, photoyzes to produce CO and O. Thus, the oxidation of SO$_2$ to SO$_3$ can also proceed by reaction with the atomic oxygen so produced. Furthermore, even in highly reducing conditions, the photolysis of SO$_2$ (and H$_2$S) can produce elemental sulfur S$_2$ that can polymerize to make stable particles of S$_8$ (Kasting et al., 1989). These particles have optical properties that also cool (Tian et al., 2010). Thus, regardless of the oxidation state of the atmosphere, elevated SO$_2$ levels will produce sulfur aerosols that offset the greenhouse warming.

There are other problems. The timescale for producing sulfur-bearing aerosols is on the order of months (Tian et al., 2010), which is much less than the timescale for removal of SO$_2$ by rainout. Thus, the duration of a transient greenhouse would be much less than envisioned by Johnson et al. (2009). And the injection of the minimum amount of SO$_2$ needed for a transient greenhouse (~10 ppmv) apparently requires volcanic eruptions much more powerful than those on Earth (Tian et al., 2010). Finally, there is still the high solubility of SO$_2$ in water to deal with. The rainout rates in the above models were parameterized and varied in sensitivity studies with results showing a lifetime of hundreds of years. However, a more recent study with a detailed cloud microphysical model indicates that in warm wet conditions when liquid water is present, the lifetime is closer to several months because precipitation rates are very high (40 cm yr$^{-1}$) under such conditions (McGouldrick et al., 2011). While it is possible that surface waters will saturate with SO$_2$ and the rainout could be balanced by return fluxes (Halevy et al., 2007), the resulting elevated SO$_2$ levels will inevitably lead to a cooling haze layer. Given all these issues, a sustainable powerful SO$_2$ greenhouse does not seem plausible for early Mars.

The few published papers on the viability of NH$_3$ and/or CH$_4$ as potential greenhouse gases are also not encouraging. Sagan and Mullen (1972) proposed that NH$_3$ in Earth’s atmosphere at concentrations of 10–100 ppm could solve the faint young Sun paradox. Though it was subsequently shown to be photochemically unstable on very short timescales (Kuhn and Atreya, 1979; Kasting, 1982), the idea was revived by Sagan and Chyba (1997), who suggested that a UV-absorbing organic haze produced from the photolysis products of CH$_4$ could shield the ammonia from photolysis. Such a haze layer does exist in the atmosphere of Titan. However, the idea was discounted for Earth because the haze was also likely to strongly absorb at visible wavelengths, creating an offsetting anti-greenhouse effect (McKay et al., 1999; Pavlov et al., 2001). More recently, this picture has changed as the fractal nature of organic haze particles allows for the possibility of low visible opacities, thereby avoiding the anti-greenhouse effect, but very high UV opacities, thereby providing an effective shield (Wolf and Toon, 2010). So the possibility of an NH$_3$–CH$_4$ atmosphere for early Earth cannot be ruled out.

For early Mars, a CH$_4$–NH$_3$ atmosphere would face similar issues. These gases, though powerful infrared absorbers, need a strong source and an effective shield to survive. An NH$_3$ greenhouse would require a mixing ratio of 5×10$^{-4}$ in a 4–5 bar CO$_2$ atmosphere to raise surface temperatures to 273 K (Kasting et al., 1992). Without shielding from ultraviolet radiation, maintaining this concentration requires a source strength several orders of magnitude greater than that estimated for early Mars (Brown and Kasting, 1993). If the shielding is provided by CH$_4$ photolysis products, as suggested for ancient Earth, then a reducing atmosphere (>10% CH$_4$) is required (Domagal-Goldman et al., 2008). Otherwise CH$_4$ will oxidize to CO$_2$ rather than polymerize to form hydrocarbons.

Alternatively, CH$_4$ itself was thought capable of providing the needed greenhouse effect. Initial calculations indicated that a several bar CO$_2$ atmosphere containing 1% CH$_4$ could have
kept early Mars warm and wet (Kasting, 1997). Unfortunately, the discovery of a coding error in those calculations showed that CH₄ greenhouse warming was greatly overestimated (Haqq-Misra et al., 2008). This, combined with the downward revision of CIA and the fact that CH₄ absorbs solar radiation in the near-infrared, severely limits the greenhouse warming from CH₄ (Ramirez et al., 2014).

However, a reducing mantle could conceivably outgas H₂ as well as CH₄. Ramirez et al. (2014) use Kasting’s updated model to show that CIA from the interaction of H₂ molecules with CO₂ could provide significant greenhouse warming on early Mars with temperatures reaching the melting point of water in atmospheres containing 1.3–4 bar of CO₂ and 5–20% H₂. The case for more reduced outgassing products is bolstered by recent laboratory experiments indicating that, for oxygen fugacities at or below iron–wüstite (IW) laboratory experiments indicating that, for oxygen fugacities at or below iron–wüstite (IW) –0.55 log₁₀ units, carbon in the mantle is stored as iron pentacarbonyl (Fe(CO)₅ rather than as carbonate (Wetzel et al., 2013). Upon degassing, the iron carbonyl dissociates to CH₄ and CO₂ which is then photochemically oxidized to CO₂. However, there is a wide range of oxygen fugacities in the Martian meteorites, ranging from just below IW all the way up to the quartz–fayalite–magnetite (QFM) buffer (Wadhwa, 2008). If the mantle during the Noachian was at the very low end of this range, then the outgassing products would have had a greater fraction of H₂, CH₄, and CO₂ required. Batalha et al. (2015) find that reaching 5% H₂ levels is possible, but it requires active volatile recycling and either additional sources of H₂ (serpentinitization), or an escape rate below the diffusion limit. If such conditions were met, thick hydrogen-rich CO₂ atmospheres could provide a solution to the faint young Sun problem for Mars.¹³

¹³ We note, however, that a hydrogen-rich atmosphere would also imply the absence of life, since both laboratory experiments (Kral et al., 1998) and calculations (Kasting et al., 2001) show that microbes tend to lower H₂ abundances to an energetic limit of ~10⁻³ bar.

17.5.3 Transient Greenhouse Atmospheres

17.5.3.1 Impact-Generated Climate Change

An alternative explanation for the late Noachian valley networks is that they were formed from the cumulative effects of temporary warm rainy climates following the impact of large asteroids onto the surface. To our knowledge this idea first appeared in print in a 1977 Scientific American article by Conway Leovy – see the last paragraph in Leovy (1977). It was later taken up by others (e.g. Matsui et al., 1988), but was first quantified by Segura et al. (2002), who demonstrated its feasibility with large impactors (>100 km). Follow-on studies focused on smaller impactors (Segura et al., 2008) and the possibility of multiple climate states (Segura et al., 2012). The basic idea is that the energy released by large impacts creates a globally dispersed ejecta blanket and a steam atmosphere that lasts for hundreds to thousands of years. As the atmosphere cools, water rains onto the surface and, when integrated over many impacts, the total volume of rainfall involved is enough to carve the valley networks we see today. This hypothesis has several attractive features. First, it does not require special conditions. Mars obviously experienced a high impact rate early in its history and the larger impactors can certainly affect the climate system. Second, the forcing function was mainly effective early in Mars’ history when the valley networks formed. Finally, it gets around the problem of finding ways to generate and sustain an effective greenhouse on early Mars. This latter feature makes it particularly attractive.

However, there are several aspects of the impact hypothesis that have been questioned with respect to the observations. The first involves the relative timing of the large impacts and valley network formation. Jakosky and Mellon (2004) pointed out that the largest impact craters formed well before the epoch of valley network formation, which they took to be near the end of the Noachian around 3.8 Ga. Since then there have been attempts to improve the estimated ages of these events. Werner (2008) re-evaluated the ages of 15 impact basins larger than 250 km and four smaller ones and found them to range in age from 3.7 to 4.1 Ga, with a mean crater age of ~3.9 Ga. Fassett and Head (2008) examined valley networks in the cratered uplands, which they estimated formed very near the Noachian–Hesperian boundary sometime between 3.7 and 3.9 Ga. This puts an average difference of no more than 200 Ma between the large impact era and valley network formation era. Of course, precise dating of specific events on early Mars is difficult, as it relies on visible crater populations that may not be representative, and models of the early Mars impact rate at Mars that may not be accurate. Consequently, the uncertainties can be large. However, on the basis of these new age estimates, it does appear that the interval of time between large impacts and valley network formation is narrower than envisioned by Jakosky and Mellon (2004).

Another criticism of the impact hypothesis has been its ability to produce the observed erosion rates. In their follow-on study of the impact hypothesis, Segura et al. (2008) estimated
that, depending on surface pressure, between about 1 and 50 m of material could have been removed by the post-impact rainfall resulting from impactors larger than 10 km in diameter. This is considerably less than the 50–2500 m they estimated would have been removed during the late Noachian if the erosion rates of Craddock and Maxwell (1993) lasted for 300 Ma. Their rainfall totals (~650 m) are also much less than the estimated hundreds to thousands of kilometers required from various models (Irwin et al., 2005; Howard, 2007; Barnhart et al., 2009; Hoke et al., 2011). These are large discrepancies to overcome and would appear to require that either the surface is much easier to erode than assumed – this seems unlikely given that they used an erodibility factor of $K \sim 0.12$, representative of a loamy sand surface – or their rainfall totals are significantly underestimated.

Barnhart et al. (2009) used a sophisticated landscape evolution model to determine the rainfall and surface characteristics of the Parana Valles drainage basin. A key observational constraint in their study was the lack of morphological evidence for breached craters in this basin area. This implies that water flowing through the valleys did not accumulate in craters to the point of overflowing the rims. They concluded that the Parana Valles system likely formed in a dry and climate with sporadic but numerous (hundreds of thousands) flooding events of sufficient intensity to erode the surface, but not so intense as to breach the craters. They ruled out the impact hypothesis on the basis that they are fewer in number and would therefore require deluge-style flooding events that would have created numerous exit breaches on the crater rims. We note, however, that breached craters do exist on Mars and they could have been caused by short-lived events of heavy rainfall (see e.g. Fassett and Head, 2005).

17.5.3.2 Orbital Variations

Orbital variations are particularly large for Mars (see Chapter 16) and must have played some role in the early Mars climate system. Obliquity, eccentricity, and precession variations occur on $10^5$–$10^6$ year timescales and, while they have a very small effect on the total annual insolation received, they do alter the latitudinal and seasonal distribution in ways that can be significant. However, this subject has received less attention for the Noachian era than it has for the Amazonian. Only a handful of studies for Noachian Mars have been reported in the literature. High-obliquity simulations with global circulation models show that dry, dense CO$_2$ atmospheres preferentially raise mean annual surface temperatures in the northern hemisphere due to increased insolation and an elevation effect, but these are not the regions where the valley networks and sedimentary units are located (Forget et al., 2013). These same simulations with a water cycle, however, do show that obliquity variations can shift the location of seasonal melting events between hemispheres, though the volume of meltwater generated is always small (1–2 mm yr$^{-1}$) and not necessarily in the right places (Wordsworth et al., 2013). More recently Mischna et al. (2013), also using a three-dimensional global climate model, studied the potential feedbacks between obliquity variations and volcanic emissions of SO$_2$ and found that, while such interactions do not raise global mean annual surface temperatures above freezing, they can broaden the times and locations where surface temperatures are above freezing.

17.5.4 Cold Climates With Locally Wet Conditions

A relatively new line of thinking, motivated by the problems with greenhouse warming, holds that early Mars was both cold and wet. In this view the fluvial features could be the result of several different processes. One possibility is that water flows on the surface in response to seasonal warming (McKay, 2004). The dry valleys of Antarctica provide an example. There, mean annual temperatures are well below freezing (253 K), precipitation is minimal (1–2 cm yr$^{-1}$) and always in the form of snow, yet lakes and rivers exist with liquid water (e.g. Lake Vanda and the Onyx river). Another possibility is that water can erupt from confined subsurface aquifers and flow on the surface (Gaidos and Marion, 2003). A cold early Mars would produce a thick cryosphere that could force large volumes of deep warm primordial groundwater to the surface. More recently, Faiérén (2010) has proposed yet another possibility in which the existence of widespread liquid water on the surface – even oceans – could be made possible by the presence of solutes from the weathering of typical Martian basalts. Solutes depress the freezing point and calculations suggest that liquids could be stable against evaporation at temperatures between 245 and 255 K (Faiérén et al., 2009). A cold ocean on early Mars would help explain the apparent rare occurrence of phyllosilicates in exposed Noachian terrains of the northern hemisphere (Faiérén et al., 2011), as well as glacial features such as fretted terrain and moraines near the hemispheric dichotomy boundary (Davila et al., 2013).

The most recent idea for forming the valley networks in a cold environment comes from general circulation modeling simulations of the early Mars hydrological cycle. Wordsworth et al. (2013) find that, for surface pressures higher than several hundred millibars, Mars becomes more Earth-like in the sense that high-altitude terrains will be colder than low-altitude terrains. This is due to the tighter coupling between the surface and atmosphere at higher pressures. This “adiabatic” effect favors snowfall in the southern highlands, which forms thick ice sheets that can later melt from impacts, volcanic activity, or basal melting. This “icy highlands” hypothesis is attractive in that it explains the predominance of valley networks in the southern hemisphere (though we do not know if they did form in the northern hemisphere but were subsequently buried), it does not require a thick atmosphere, and it makes use of plausible transient non-solar energy sources for melting. However, it is not clear if these energy sources can generate the meltwater volumes needed for the required erosion. In this view, early Mars is mostly a cold planet in which erosion and fluvial activity result from the episodic melting of surface ice deposits.

A common feature of these ideas is that they require higher surface pressures than Mars has today. For the icy highlands idea, high surface pressures are needed for the adiabatic effect. For the seasonal melting, groundwater eruption, and solutes in the water ideas, boiling is an important consideration. The mean annual surface pressure on Mars today is very close to the triple-point pressure of water. Below the triple point, water cannot exist as liquid. But even at those locations and times...
where pressures and temperatures exceed the triple point, liquid water on Mars today would be very close to the boiling point (Haberle et al., 2001). Boiling occurs when the vapor pressure of water is comparable to the surface pressure, and this leads to very high evaporation rates (Ingersoll, 1970; Kahn, 1985; McKay, 2004). Since the latent heat of vaporization is almost two orders of magnitude higher than the specific heat, evaporation will rapidly cool the water. Under these circumstances, water quickly freezes unless it is gaining energy from an external source. Thus, without higher surface pressures, liquid water will not last long enough to erode the surface to the levels observed. The exact value needed is situation-dependent, but it should be well above the triple point. McKay (2004) suggests a minimum value of about 100 mbar, which is a plausible value for the Noachian epoch.

17.5.5 A Brighter Early Sun?

If early in Mars’ history the Sun was brighter than predicted by the standard solar model, it would make it easier to stabilize liquid water on the surface. This could be achieved if the Sun was more massive in its early main-sequence phase (Whitmire et al., 1995). Solar luminosity is proportional to the fourth power of solar mass $M_{\odot}$, and, since a planet’s orbital distance $r$ is inversely proportional to $M_{\odot}$, and the solar flux varies as $r^{-2}$, the flux at a planet scales as $F \sim M^2$. Hence, small changes in solar mass can result in large changes in the solar flux.

If the Sun was more massive when it joined the main sequence, then it must have lost the excess mass over time in order to arrive at its present mass. The amount of excess mass and its loss history have been the subjects of several observational (e.g. Morel et al., 1997; Wood et al., 2002, 2005) and theoretical (e.g. Willson et al., 1987; Boothroyd et al., 1991; Swenson and Faulkner, 1992; Guzik and Cox, 1995; Sackmann and Boothroyd, 2003; Minton and Malhotra, 2007; Turck-Chièze et al., 2011) studies. The main constraints on mass loss are the observed lithium depletion in the Sun’s photosphere (Boothroyd et al., 1991), helioseismic observations that infer the speed of sound through its interior, and the fact that the Earth still has most of its water, which it would have lost if the Sun was too bright (Kasting, 1988). These considerations generally imply that the early Sun could not have been more than 10% more massive than it is today, and that most of its excess mass was lost early in the main-sequence phase (<1 Ga after birth).

Sackmann and Boothroyd (2003) have published perhaps the most definitive study on how early mass loss might have affected Mars and Earth. They considered three different mass-loss scenarios and tested how well their solar evolution models compared with helioseismic observations. The key results are shown in Figure 17.16. Their preferred scenario is for an initial mass of $1.07 M_{\odot}$, which gives results that are slightly (but not significantly) more consistent with helioseismic observations than the standard solar model (SSM). In this case, the solar flux at Mars 4.5 Ga would have been ~5% higher than it is today. By 3.4 Ga it would have declined to a value ~16% lower than today’s. Thus, throughout the Noachian, the solar flux would have been significantly higher than predicted by the SSM. Under these circumstances, the surface pressure requirement for a sustained CO$_2$ greenhouse would be reduced in comparison with previous predictions (e.g. Kasting, 1991; Wordsworth et al., 2013).

Observational support for the early mass loss hypothesis is tenuous. To date, loss rates at the level favored by Sackmann and Boothroyd (2003) have not yet been confirmed or observed for stars similar to the Sun. At the present time, the solar wind removes ~3x10$^{-14} M_{\odot}$ yr$^{-1}$, which would amount to ~0.01% of its present mass if this loss rate were constant over time. Lunar and meteoritic data suggest the mean solar wind over the past 3 Ga was an order of magnitude higher than it is today (Kerridge et al., 1991). Thus a soft lower limit for the Sun’s mass loss is ~0.1% of present mass. Of course, these data do not constrain mass loss rates during the first billion years, so observations of Sun-like analogs might help in this regard. Wood et al. (2002, 2005) have inferred mass loss rates from several nearby young Sun-like stars using observations of Ly $\alpha$ absorption that result from interactions of stellar winds with the interstellar medium. Their most recent results indicate loss rates of ~0.1% of initial mass for these stars, which is consistent with the lunar and meteoritic data for our Sun, but far less than

![Figure 17.16. Solar flux at Earth or Mars relative to the present solar flux as a function of time. Solid and dashed lines refer to different initial solar masses ($M_i$) relative to the present mass ($M_{\odot}$). Each panel represents a different loss scenario, as indicated. Thick double arrows give the lower flux limit of Kasting (1991) for the presence of water on early Mars. Thin arrows give the same limit for a low surface albedo. From Sackmann and Boothroyd (2003).](image-url)
the 7% mass loss favored by Sackmann and Boothroyd (2003). The removal timescale is also problematic, as it may be too short to keep the Earth from freezing after the first billion years or so (Minton and Malhotra, 2007). Thus, the size of the needed mass loss and its removal time are inconsistent with the available data. However, the observed mass loss rates are somewhat model-dependent, the sample size is very small, and exact solar analogs are difficult to find. Furthermore, the recent downward revision in the abundances of heavy elements in the Sun’s interior (Asplund et al., 2009; Caffau et al., 2010) increases the discrepancy of SSMs with helioseismic data, which can be partly mitigated with even larger mass loss rates (10–30%) (Guzik and Mussack, 2010; Turck-Chièze et al., 2011). Thus, the early mass loss hypothesis cannot be completely ruled out.

17.6 DISCUSSION AND FUTURE PROSPECTS

In spite of the considerable research conducted since Mariner 9 images first revealed ancient fluvial systems, we still do not know for certain what environmental conditions prevailed during the late Noachian on Mars. There is compelling evidence and widespread agreement that water flowed on the surface, and that the atmosphere was thicker in those times. But the notion of a sustained warm and wet climate system has been difficult to justify.

The early optimism that the greenhouse effect of a CO₂–H₂O atmosphere could offset the diminished output of the young Sun by providing enough warming to raise global mean surface temperatures to the melting point has faded to skepticism as researchers have delved deeper into the details of such a climate system. Not only does it appear that the Noachian atmosphere was less massive than previously assumed (though this is somewhat model-dependent), it is also less capable of warming the surface because of limitations imposed by CO₂ condensation, a weaker CO₂ cloud scattering greenhouse effect, and greatly reduced collision-induced absorption. State-of-the-art global climate models with pure CO₂–H₂O atmospheres are unable to raise global mean surface temperatures above ~250 K for any surface pressure unless a strong cloud greenhouse can be generated.

Supplementing such atmospheres with additional greenhouse gases such as SO₂, H₂S, CH₄, NH₃, and H₂ could boost their greenhouse power if their concentrations can be maintained at the right levels. But this is difficult to do; each of these gases has sinks that limit their effectiveness. SO₂ converts to sulfate, H₂S and CH₄ are not very effective in the first place, NH₃ – the most powerful greenhouse gas – is photochemically unstable, and H₂ requires a reducing mantle and very large outgassing rates. This does not rule them out, but special circumstances that have yet to be demonstrated are required to sustain the right concentrations. More research is needed to assess the viability of supplemental greenhouse gases in solving the early Mars dilemma.

Similarly, more research is needed into the role of impact-generated climate change. The few studies done thus far show that 10 km impactors can alter the climate system and produce rainfall, but the rainfall and erosion they produce falls short of what is required. An interesting feature of these studies that offers a promising new direction for early Mars research is the potential for a cloud greenhouse effect. The one-dimensional model of Segura et al. (2008) showed that the (water) clouds warm the surface and that they have a significant role in maintaining liquid water on the surface. More recent separate studies confirm the potentially warming effect of clouds, though with diverse estimates of the quantitative effect. Whether such clouds, generated either by impacts or at times of high obliquity, play an important role in warming the early Mars climate system warrants further study.

Finally, the continued analysis of existing data and the future acquisition of new data will help refine our understanding of the true nature of the early Martian climate system. Image analysis that further quantifies the volume of eroded material and the timing and duration of precipitation will help distinguish between the need for continuous versus episodic climate change. A more precise determination of the mineralogy of the surface should help better define the environmental conditions that existed when they formed. And the continued refinement of isotopic data from the SNCs and MSL will be of great value in piecing together escape scenarios. This, along with measurements of the current escape rates to come from MAVEN, should go a long way towards estimating the mass and composition of the early atmosphere.

ACKNOWLEDGMENTS

We are grateful for helpful reviews from Brian Toon, Jim Kasting, and Bruce Jakosky.

REFERENCES


References


