The case for rainfall on a warm, wet early Mars

Robert A. Craddock
Center for Earth and Planetary Studies, National Air and Space Museum, Smithsonian Institution, Washington, District of Columbia, USA

Alan D. Howard
Department of Environmental Sciences, University of Virginia, Charlottesville, Virginia, USA

Received 11 April 2001; revised 10 April 2002; accepted 10 June 2002; published 23 November 2002.

[1] Valley networks provide compelling evidence that past geologic processes on Mars were different than those seen today. The generally accepted paradigm is that these features formed from groundwater circulation, which may have been driven by differential heating induced by magmatic intrusions, impact melt, or a higher primordial heat flux. Although such mechanisms may not require climatic conditions any different than today’s, they fail to explain the large amount of recharge necessary for maintaining valley network systems, the spatial patterns of erosion, or how water became initially situated in the Martian regolith. In addition, there are no clear surface manifestations of any geothermal systems (e.g., mineral deposits or phreatic explosion craters). Finally, these models do not explain the style and amount of crater degradation. To the contrary, analyses of degraded crater morphometry indicate modification occurred from creep induced by rain splash combined with surface runoff and erosion; the former process appears to have continued late into Martian history. A critical analysis of the morphology and drainage density of valley networks based on Mars Global Surveyor data shows that these features are, in fact, entirely consistent with rainfall and surface runoff. The necessity for a cold, dry early Mars has been predicated on debatable astronomical and climatic arguments. A warm, wet early climate capable of supporting rainfall and surface runoff is the most plausible scenario for explaining the entire suite of geologic features in the Martian cratered highlands.

INDEX TERMS: 5407 Planetology: Solid Surface Planets: Atmospheres—evolution; 5415 Planetology: Solid Surface Planets: Erosion and weathering; 1815 Hydrology: Erosion and sedimentation; 1824 Hydrology: Geomorphology (1625); Keywords: Mars, Martian, erosion, climate, hydrology, fluvial


1. Introduction

[2] Most of the geologic and climatic history of Mars is recorded in the cratered highlands. These materials are the oldest on the Martian surface and cover ~60% of the planet [Tanaka et al., 1988]. However, the geology of the highlands is complicated, and reconstructing the processes that have operated through time on Mars has proven difficult. The valley networks, which are contained almost exclusively in the highlands, are the most obvious features indicating that ancient geologic processes, and the intensity at which they operated, were different than today’s [e.g., Carr, 1996, p. 71]. As a result, much attention over the last several decades has focused on their origin. Such analyses have led to the conclusion that valley networks were formed primarily or exclusively by groundwater sapping [Pieri, 1976, 1980; Carr and Clow, 1981; Baker and Partridge, 1986; Brakenridge, 1990; Baker et al., 1992; Goldspiel et al., 1993; Carr and Chuang, 1997]. Often this interpretation has been influenced by the paradigm that Mars has always been cold and dry such that liquid water was never stable on the surface. To resolve this paradox, several models propose that groundwater sapping would have been driven by geothermal heating [Gulick and Baker, 1990; Clifford, 1993; Squyres and Kasting, 1994; Gulick, 1998; Tanaka et al., 1998; Malin and Carr, 1999; Carr and Malin, 2000; Goldspiel and Squyres, 2000]. We find this explanation as the primary cause of valley networks to be untenable as it ignores the regional geologic context. Specifically, hydrothermal circulation does not explain erosion of highland impact craters, which occurred contemporaneously with erosion of the valley networks. We maintain that an equally compelling case can be made that early Mars was warm and wet, so that hydrothermal circulation driven by geothermal heating is no longer a necessary process.

[3] Geochemical, climatic, and geologic analyses provide the basis for understanding early Martian environmental conditions and processes. In this paper we review the available evidence, reaching the conclusion that a preponderance of climatic and geochemical evidence either sup-
ports a warm, wet early Mars or is inconclusive. Basically, the geochemical data remains unclear. Likewise, climatic analyses depend upon many variables, most of which are uncertain, thus yielding ambiguous results. Consequently, it is the geology that provides the best record for determining the early history of Mars.

While there is some evidence indicating that groundwater sapping did occur, our analysis of the geologic record indicates that rainfall and surface runoff were the dominant processes responsible for eroding the valley networks and modifying highland impact craters. Rainfall and surface runoff occurred throughout most of the Noachian and possibly through the early Amazonian, but the intensity at which these processes operated may have waned with time. As on the Earth, groundwater sapping on Mars appears to have been maintained by precipitation, not geothermal heating.

The geologic evidence in the cratered highlands is most easily explained by a thicker, warmer primordial atmosphere. It is probable that this atmosphere gradually condensed and collapsed into the regolith while simultaneously being lost to space. Eventually, warmer, wetter surface conditions waned into a drier, periglacial climate that existed prior to the onset of conditions observed today.

2. Geochemical Evidence

In a warm, thick atmosphere, the weak acid-base reaction of CO₂ gas and water should produce carbonate deposits on Mars [Kahn, 1985; Catling, 1999; Morse and Marion, 1999]. Such deposits, however, have eluded spectral identification, and this has often been cited as evidence against a warm, wet early climate [Blaney and McCord, 1989; Christensen et al., 2000a]. Because valley networks and outflow channels indicate that liquid water was present on the surface at least periodically in the past, the missing carbonate deposits have been called one of the “major conundrums” of Martian science [Carr, 1998]. There are, however, several plausible reasons why remote sensing instruments have not identified carbonates on Mars.

1. Early environmental conditions on Mars were reducing, thus never allowing abundant limestone or dolomite to synthesize (R. G. Burns, personal communication, 1992). Similar to early Earth [Kasting, 1993], the early Martian atmosphere and any released juvenile water was probably reducing. Outgassing of water vapor and sulfates may have combined to form acid rain, which could have destroyed any carbonates that did form. Limestone does not synthesize under current Martian conditions UV irradiation causes sulfates to rapidly displace any carbonate species, releasing CO₂ back into the atmosphere. This hypothesis is further supported by laboratory simulations [Mukhin et al., 1996]. Large deposits of carbonates may be present on Mars, but because remote sensing data are only sensitive to the upper few centimeters, they could go undetected.

2. Carbonates may have undergone secondary alteration. In the current oxidizing environment on Mars, SO₂ gas released from volcanism and magmatic outgassing would be quickly converted to SO₃, which reacts with any H₂O molecule present to form H₂SO₄ [Clark, 1999]. Sulfuric acid has a very low freezing point and would remain stable until it came into contact with a mineral grain to react with; thus, the upper layer of rock and soil on Mars could be scrubbed free of any carbonates [Gooding, 1978; Clark, 1999; Catling, 1999]. UV breakdown could also remove carbonate from the upper layer or rock and soil. Clark [1979] showed that under current Martian conditions UV irradiation causes sulfates to rapidly displace any carbonate species, releasing CO₂ back into the atmosphere. This hypothesis is further supported by laboratory simulations [Mukhin et al., 1996]. Large deposits of carbonates may be present on Mars, but because remote sensing data are only sensitive to the upper few centimeters, they could go undetected.

3. Soils forming under current climatic conditions may simply be volumetrically more abundant than those formed during the early history of Mars [Blaney, 1999]. The mineralogy and formation of present-day soils is contentious, but most mechanisms do not require a warm, wet climate [Bell, 1999]. If a bulk of the Martian soils formed over the last billion years or more when environmental conditions were the same as today’s, then the spectral signature of any carbonate deposits may be washed out. To resolve this issue, Blaney [1999] recommends that soil formation models also consider the rate and volume of material produced and not just chemistry alone.

4. Particle size and surface texture greatly influence the spectral signature. Although calcite exhibits strong adsorption bands at 6.5, 11.25, and 35 μm in the thermal infrared, investigations by Kirkland et al. [2000a] show that surfaces covered by ~50% to 100% calcrete can exhibit virtually no spectral signatures due to multiple scattering by surface cavities [Kirkland et al., 2000b], which greatly increase the emissivity and reduce the spectral band contrast. Thus, large outcrops of carbonates may be present on Mars, but have eluded spectral identification because of a rough, pitted surface [Kirkland et al., 2000a, 2000b]. Blocky carbonate surfaces can also remain undetected due to the cavity effect. In addition, small carbonate particles are optically thin with rough edges, so such material also looses spectral reflectance due to volume scattering [Salisbury and Wald, 1992]. Large deposits of carbonates may be present as rough or blocky outcrops or as fine-grained dust on Mars, but they would elude detection from instruments such as TES.

Positive identification of carbonate deposits through remote sensing data is not as straightforward as it would seem. To be visible to remote sensing instruments, the surface of any carbonate deposit would have to remain fairly smooth, stay free of dust, resist secondary alteration, and stay volumetrically abundant (>10%) over billions of years. Given these caveats, it may not be surprising that carbonates have eluded detection. Although it is a weak argument to suggest that carbonates may be present on Mars, yet they simply cannot be seen, their absence does not preclude an early warm, wet climate. Analyses of the SNC meteorites indicate that carbonates are present on Mars, so it is likely that there is a crustal sink for these materials [e.g., McSween, 1994]. Suggesting that some process is “hiding” carbonates on Mars is, therefore, not ad hoc.

Ultimately, however, the presence or absence of carbonates says nothing definitive about past climates or geologic conditions on Mars. Synthesis of carbonate is
possible even under current Martian conditions [Booth and Kieffer, 1978]. Thus, a cold, dry early Mars does not explain the lack of spectral evidence either. Likewise, should we eventually identify carbonates, it does not mean a priori that early Mars was warm and wet. A future lander, for example, may be capable of accessing fresh bedrock or unaltered regolith and thus find carbonates; however, the challenges will be in determining the age of such materials and the environment in which they formed. The missing carbonates are a "conundrum," but what this actually means is unclear.

Finally, identification of olivine from TES spectra has been presented as an additional geochemical argument against a warm, wet early Mars [Clark and Hoefen, 2000]. Basically, olivine is easily weathered in above freezing conditions when abundant water is present, so the interpretation is that such conditions never existed on Mars. However, spectral olivine deposits are located in dark albedo materials, which result from exposed bedrock or locally derived coarse debris typically associated with Amazonian volcanic flows [Soderblom, 1992]. Since these materials are <1 billion years old, they provide no evidence to the early climatic history of Mars.

3. Climatic Evidence

Three general issues influence our perception of the early Martian climate. These include the thickness and composition of the primordial atmosphere, the luminosity of the early Sun, and conclusions from climate models that assess how the early atmosphere may have responded to these other parameters. Increasingly, the accepted paradigm within the planetary science community is that early Mars was cold and dry, primarily because of imposing cases presented by some climate models. However, it is important to remember that climate models are generally based in part on largely untenable assumptions and unconstrained variables. In every instance, an opposing, argument has been made supporting warm surface conditions on early Mars. Because these arguments are often integral to interpreting the geologic history of Mars, summaries of the competing scenarios are presented below.

3.1. A Doomed Climate?

Similar to the Earth, Mars probably had a thicker atmosphere early in its history. This atmosphere would have formed in part as volatiles were released from impacts during accretion [Boslough et al., 1980]. However, a majority of the volatiles composing this atmosphere, such as water and CO₂, would have been reincorporated into the growing planet because of their solubility [Jakosky and Aherns, 1979]. Mars would have taken longer to accrete than the Earth [Weidenschilling, 1976] and should have remained generally colder. For this reason it is more likely that Mars reincorporated volatiles as it continued to accrete [Jakosky and Aherns, 1979]. A more important mechanism for creating an early, thicker atmosphere on Mars may have been degassing from large-scale melting following accretion and subsequent planetary differentiation as suggested for the Earth [Fanale, 1971]. However, Holland [1984, p. 76] states that the behavior of volatiles during such an episode is difficult to analyze. The comparative importance of both mechanisms remains controversial for the Earth (let alone Mars) and the amount and relative proportions of volatiles both within the crust and atmosphere following accretion is largely unknown. Whatever original, thicker atmosphere was present, it may have been especially susceptible to impact erosion because of the low mass and surface gravity of Mars [Melosh and Vickery, 1989]. In addition, hydrodynamic escape of hydrogen may have also stripped the early atmosphere of heavier gases, which is evidenced by the mass-fraction of noble gases in the current Martian atmosphere compared to the telluric value [Hunten et al., 1987; Hunten, 1993].

A tenuous, primordial atmosphere becomes more problematic for supporting warmer surface conditions on Mars when astronomical observations of the Sun are considered. Because of its greater distance from the Sun, Mars receives only 43% as much sunlight as the Earth. This alone makes it difficult for climatic conditions to generate above freezing surface temperatures. Gough [1981], however, suggested that compositional changes in the Sun through time have caused it to burn brighter, so its original luminosity may have only been 70% its current value. The "faint, young Sun" is the principal reason why it is so difficult to model conditions where early Mars was not freezing [e.g., Haberle et al., 1994], apparently dooming it to a perpetual cold, dry climate.

To counteract the effects of a "faint, young Sun," Mars would need some mechanism to push surface temperatures higher, and the simplest solution is an early atmosphere containing several bars of CO₂ [Pollack et al., 1987]. If this were the case, however, the argument is that hundreds of meters of carbonates should be sequestered in the Martian regolith, and, as discussed previously, there is no spectral evidence supporting this [Blaney and McCord, 1989; Christensen et al., 2000a]. Alternatively, perhaps there was an efficient greenhouse gas in the early atmosphere (e.g., ammonia [Kuhn and Atreya, 1979]), but such gases are thought to be short-lived and are hard to defend in the absence of a viable recharge mechanism.

Although a variety of climate models have been proposed, Kasting's [1991] analysis of the surface effects resulting from an early CO₂-rich atmosphere has become a principal tenet for a cold and dry early Mars. Simply stated, his results suggest that any CO₂ ice clouds that formed in the primordial atmosphere would reduce the convective lapse rate of the troposphere and thus reduce surface heating. Accepting Kasting's [1991] results together with a "faint, young Sun," the suggestion that the early atmosphere was tenuous, and the potential problems with viable greenhouse gases, make a compelling case that the Martian climate has always been cold and dry. However, such an argument is one-sided, as we elaborate below.

3.2. A Warmer Climate?

3.2.1. A Thicker Atmosphere

Although it is possible that most of the volatiles that were outgassed during accretion may have been reincorporated into the mantle during differentiation [Jakosky and Aherns, 1979], it is unclear how much may have been subsequently re-released. In contradistinction to Lewis' [1972] cosmochemical model, analyses of nonradiogenic volatiles measured by Viking suggest that Mars is depleted in volatile elements relative to the Earth. In addition, radiogenically
produced isotopes indicate that outgassing on Mars was 4 times less complete than on Earth [Anders and Owen, 1977]. Despite these estimates, Anders and Owen [1977] calculate that early Mars outgassed as much as 525 millibars of CO₂ alone, which they suggest was enough to permit the existence of liquid water on the surface.

[16] Substantial impact erosion also supports the presence of an early, thick atmosphere. Although this process is often cited as mechanism for removing the early atmosphere and driving a colder climate [e.g., Squyres and Kasting, 1994], calculations by Melosh and Vickery [1989] suggest that the primordial atmosphere must have been at least ~1 bar. Without this early, thick atmosphere to start with, impact erosion would have reduced pressures well below their current value of 0.06 bar. More recent analyses, however, suggest that large impacts may not have played as an important role in removing planetary atmospheres [Newman et al., 1999]. Notwithstanding, smaller, more numerous impactors were adding to the early atmosphere. This “late-accreting veneer” is becoming accepted as the main contributor of volatiles on the inner planets [Anders and Owen, 1977; Owen, 1992]. Comets in particular may have added substantial amounts of volatiles to the planetary reservoirs and may be responsible for creating the oceans on Earth and Mars [Chyba, 1987, 1990]. Because the surface of Mars probably stayed cooler during accretion [Weidenschilling, 1976], a higher percentage of volatiles may have been incorporated into the crust [Jakosky and Aherns, 1979]. C. Bertka (personal communication, 1999) suggests that a high percentage of water in the crust may also explain the basaltic andesite composition measured by Mars Pathfinder [Morris et al., 2000]. Thus, the amount of volatiles released during volcanic outgassing may have been proportionally higher on Mars than the Earth. Volatiles released during late-stage cometary impacts and volcanic outgassing could have played an important role in maintaining a dense atmosphere on early Mars.

[17] As a minimum estimate, it is likely that the early Martian atmosphere was ~0.5 bar thick, and pressures greater than 1 bar do not seem unreasonable. Because both impact cratering and outgassing from volcanism are stochastic processes, it is likely that the primordial atmosphere did not simply monotonically decline into the current one. Rather, between impact and volcanic events there may have been punctuated episodes on Mars where the atmosphere was 0.5 bar or thicker, but the frequency and amplitude of these episodes should have declined through time.

3.2.2. Warmer Arguments

[18] A thick primordial atmosphere is generally not held in contention, but other climatic issues are controversial. Although the “faint, young Sun” [Gough, 1981] is an assumption in many climatic models [Kasting, 1991; Haberle et al., 1994], this hypothesis is not without controversy. In fact, some astronomers have found it a bit disconcerting that climate modelers would take the nature of the early Sun as a known quantity, as the whole question of the early, faint Sun is still very much open (L. R. Doyle, personal communication, 2000).

[19] To explain the observed depletion of lithium in the solar atmosphere, computational models of solar evolution and convection indicate that the young Sun was more massive (~1.04 – 1.2 solar masses) than the present [Weyman and Sears, 1965; Hobbs et al., 1989; Boothroyd et al., 1991; Graedel et al., 1991]. This hypothesis has been supported with some observational data [Soderblom et al., 1990; Whitmire et al., 1995; Doyle et al., 1996], and it agrees with observed mass-loss rates in solar-type dwarfs [Mullan et al., 1989; Whitmire et al., 1995]. These observations imply that the young Sun was brighter than it is today, which could help create warmer surface temperatures on early Mars [Whitmire et al., 1995]. In the absence of a “faint, young Sun,” climatic models indicate that above freezing conditions would have been possible under early, high atmospheric pressures [Haberle et al., 1994].

[20] The effectiveness of CO₂ clouds at surface cooling [Kasting, 1991] has also been questioned. Forget and Pierrehumbert [1997] recalculated Kasting’s [1991] findings while including the scattering effects of CO₂ ice crystals. Because CO₂ ice clouds reflect outgoing thermal radiation back to the surface, a condensing, CO₂-rich atmosphere might be extremely effective at raising Martian surface temperatures above freezing. In fact, Forget and Pierrehumbert [1997] calculate that if the paleoatmosphere contained more than just a few tenths of bars of CO₂, then surface temperatures would have been high enough to permit liquid water on the surface even in faint, young Sun conditions. The entire issue of CO₂ clouds and their effectiveness at raising surface temperatures is currently being reexamined. Using a similar model, Mischna et al. [2000] treated thermal infrared scattering more consistently. Their findings generally concur with Forget and Pierrehumbert’s [1997]; however, they also found that low, optically thick clouds could actually reduce surface temperatures. Colaprete and Toon [2000] found that CO₂ ice clouds may cause temperatures in the surrounding air to increase, causing the clouds to dissipate within a few days. Thus, warming of the surface and atmosphere on Mars by clouds may be self-limiting. The disparity in these types of ongoing analyses emphasizes how sensitive such models are to input parameters. At the very least, properly assessing the effects of CO₂ clouds on raising surface temperatures will require three-dimensional models that consider atmospheric pressure, changes in atmospheric temperatures, cloud location, cloud height, and optical depth.

[21] Finally, arguments for early greenhouse gases also continue. Sagan and Chyba [1997], for example, suggested that any ammonia gases contained in the early terrestrial atmosphere may have been shielded from photodissociation by a high-altitude layer of organic aerosols produced from methane photolysis. Depending on the composition of the atmosphere, atmospheric shocks from impacts as well as interplanetary dust particles would also contribute to this layer [Chyba and Sagan, 1992]. They suggest that it may be possible to warm the surface without invoking several bars of atmospheric CO₂ even if there was a “faint, young Sun.” Although their calculations were made primarily with the Earth in mind, the implication is that there may be means for increasing the lifetimes of greenhouse gases in the atmospheres of the terrestrial planets that do not require ad hoc resupply mechanisms.

3.3. Our Current Understanding

[22] Determining the nature of the early climate on Mars is extremely difficult. Currently the composition and pres-
sire of the early atmosphere are not understood with any reasonable degree of certainty, and the nature of the early Sun is controversial. These problems not only make it difficult to explain how liquid water may have existed on the surface of early Mars [e.g., Haberle et al., 1994], but they also make it difficult to explain how life originated on Earth [e.g., Cloud, 1988, p. 124]. Clearly the latter implies a weakness in the “faint, young Sun” hypothesis or, at the very least, that some type of mechanism existed for creating a greenhouse on the early Earth. Determining what the most important or sensitive variables were in creating the early climate on Mars (and Earth) is ultimately the value of performing climatic models. However, it would be unwise to force the interpretation of any geologic feature on Mars to fit a given result or scenario. To offer one perspective, climate models basically attempt to predict the weather on an alien world almost four billion years ago; yet, even the daily forecasts for the Earth are frequently inaccurate. As outlined in following sections, the geology of the highlands offers the best record for understanding the climate of early Mars when geomorphic features are examined collectively.

4. Geologic Evidence From Valley Networks

[23] A variety of geologic features located in the Martian highlands attest to the presence of water or ice. There is evidence that Hadriaca and Tyrhena Paterae erupted explosively during their early development as rising magma interacted with volatiles in the crust [Crown and Greeley, 1993]. The morphology of “softened terrain” at midlatitudes is consistent with landscapes modified by ice-induced creep [Squyres et al., 1992]. Putative deltas, terraces, and shorelines suggest that water ponded in a number of impact craters [Cabrol and Grin, 1999]. Similarly, Ma’adim Vallis appears to have formed from catastrophic releases of water from a large, intercrater lake [Irwin et al., 2002]. On a smaller scale, small gullies found on the sunward facing slopes of impact craters and valleys suggest seepage of water that may have occurred geologically recently [Malin and Edgett, 2000b]. MOC images also indicate that materials contained in some impact craters or which drape the landscape in certain locations display characteristics consistent with sedimentary rocks, suggesting fluvial processes [Malin and Edgett, 2000a]. In addition, the fluidized form of many fresh impact craters suggests that water may be present at the near surface [e.g., Costard, 1989]. To date, however, the most intensely studied geologic features in the Martian highlands are the valley networks [Pieri, 1976, 1980; Masursky et al., 1977; Carr and Clow, 1981; Brakenridge et al., 1985; Brakenridge, 1990; Baker et al., 1992; Goldspiel et al., 1993; Carr, 1995; Carr and Chuang, 1997; Gulick, 1998; Malin and Carr, 1999]. Because so much information has been collected about valley networks that can provide clues about ancient climatic conditions and past geologic processes, it is important to review and evaluate this evidence.

[24] It is important to first point out that the term “valley network” may be a bit of a misnomer. By definition a valley is any low-lying land bounded by higher ground on either side that is usually traversed by a stream or river [e.g., Bates and Jackson, 1984, p. 550]. A channel, on the other hand, is simply the lowest portion of a stream, bay or strait [e.g., Bates and Jackson, 1984, p. 82]. The distinction is that water may be contained in a channel or stream, and a channel or stream may be contained in a valley. It is unclear whether Martian valley networks are actually channels or true valleys as the literature appears to suggest both. For example, Malin and Carr [1999] present evidence for a channel at the bottom of Nanedi Vallis, but such features are rare. As we discussed below, the lack of channels at the bottom of most valley networks may be a preservation issue. However, many valley networks may also be true channels. We try to make this distinction clear in our discussion.

[25] In general, past studies have provided information regarding valley network drainage densities, morphology, and age. Almost always investigators have concluded that valley networks formed primarily from groundwater sapping. The most frequently cited observation used to support this interpretation is the apparent low drainage densities of many valley network systems [Pieri, 1976, 1980; Carr and Clow, 1981; Brakenridge et al., 1985; Gulick and Baker, 1989; Wilhelms and Baldwin, 1989; Brakenridge, 1990; Baker et al., 1992; Clifford, 1993; Gulick et al., 1997; Carr and Chuang, 1997; Gulick, 1998]. Carr and Chuang [1997, for example, measured valley networks in the equatorial highlands that were visible at Viking orbiter mapping resolution (~100–200 m/pixel). They compared their measurements to terrestrial drainage densities derived from Landsat images degraded to Viking image resolutions and found that the densities of the valley networks are consistently lower than terrestrial runoff channels. Gulick [1998, 2001] also suggested that valley networks occur as either isolated features or in clusters, which is consistent with groundwater sapping. If valley networks resulted from rainfall and surface runoff, Gulick [1998] argues they should uniformly dissect a surface with apparently similar age, compositions and topography.

[26] Morphological evidence for groundwater sapping includes the abrupt headward termination of valleys [Pieri, 1976], absence of fine dissection in uplands between valleys [Malin and Carr, 1999], relatively constant valley width downstream [Goldspiel et al., 1993], and, in some cases, short, stubby tributaries [Baker and Partridge, 1986]. These characteristics are well exhibited in terrestrial groundwater sapping systems that occur in layered sedimentary rocks [Lalit and Malin, 1985; Howard et al., 1988; Luo et al., 1997].

[27] Finally, Carr [1996, p. 90] states that the Amazonian-age valley networks found on the Tharsis volcanoes “represent the strongest evidence against the simple climate change model in which early Mars was warm and wet but changed early to conditions that resemble the present and that persisted for most of Mars’ history.” The interpretation is that these valley networks formed after the period of fluvial activity that occurred in the older cratered highlands. Without the benefit of an early warm, wet climate, some other mechanism is needed to explain such valley networks. Gulick and Baker [1990] proposed geothermal heating from the Tharsis volcanoes as a source for melting ground ice. In their model, warm, less-dense water rises, creating a buoyancy-driven flow. The resulting groundwater circulation is envisioned to partially discharge to the surface, carving the
valley networks through sapping. With some climatic scenarios suggesting that Mars has always been cold and dry, hydrothermal circulation models have also been proposed to explain valley network formation in the Martian highlands [Squyres and Kasting, 1994; Gulick, 1998, 2001]. Hydrothermal circulation seemingly explains the lack of observed carbonates, how fluvial erosion could occur in a cold and dry climate (e.g., modeling by Goldspiel and Squyres [2000]), as well as valley network morphology, density, and age. Hydrothermal circulation seems to have become a paradigm for understanding the geology of the highlands that concords with the cold, dry interpretation of the early Martian climate. Because of the wide acceptance this hypothesis has gained and the implications it has for interpreting past climatic conditions and geologic processes, it deserves a critical review.

4.1. Groundwater Sapping

[28] Groundwater sapping has been assumed to be an almost inevitable accompaniment to hydrothermal circulation. As a result, it is often invoked for the valley network formation without much attention being paid to the requisite processes and hydrological conditions. Groundwater sapping can be loosely defined as erosion due to water that has reemerged from subsurface flow. Such erosion is often contrasted to erosion due to overland flow fed directly by precipitation. This water source distinction is, however, not easy to make in natural landscapes. Classic “Hortonian overland flow” occurs when the rainfall intensity exceeds the infiltration capacity of the soil or were rock is exposed, but surface runoff also occurs when shallow subsurface throughflow reemerges to the surface or when groundwater saturates the regolith and seeps out to combine with precipitation (“saturation overland flow”). Such saturated areas often occur in valleys and lowlands, where their extent depends upon prior precipitation, leading to the “variable source area concept” proposed by Dunne [1978]. In shale badlands, most runoff on slopes occurs as shallow throughflow in shrinkage cracks and small eroded pipes (see review by Howard [1994a]). Channel heads in many, if not most, terrestrial landscapes may be maintained and extended by erosion due to reemergent subsurface flow [DeVries, 1976; Dunne, 1980, 1990; Vaughn, 1989; Dietrich and Dunne, 1993; Onda, 1994; Uchapi and Oldale, 1994; Schumm et al., 1995; Nash, 1996]. In this sense, groundwater sapping loses significance as a process distinct from runoff erosion fed by precipitation.

[29] Erosion by runoff from emergent groundwater flow can occur through several mechanisms depending upon the material properties and the hydrologic setting. Slopes in sandy cohesionless soils are readily eroded by emerging groundwater. Such slopes are undercut by mass wasting due to upward seepage, and the emergent water transports the sand downstream from the headcut [Howard and McLane, 1988]. In planform, valleys are stubby and poorly branched [Baker et al., 1988]. Examples can be found on beaches due to water emerging during low tide and along rivers with banks of cohesionless alluvium following rapid lowering of river stage (e.g., along the Colorado River in the Grand Canyon [Budhu and Gobin, 1994]). In cohesive (but not indurated) soils, intergranular groundwater flow is inhibited by small pore sizes and groundwater erosion occurs primarily by hydraulic detachment by flow through discrete fractures and macropores, often enlarging them into networks of subsurface pipes [Jones, 1990; Parker and Higgins, 1990], which may eventually collapse to form gully systems [Higgins, 1990a, 1990b]. Seepage in concave headwater hollows can also cause landslides and debris avalanches [Dietrich and Dunne, 1993; Dietrich et al., 1995], helping to form and maintain the headwater tips of channel systems. Bank erosion by groundwater seepage may control channel width [Hagerty, 1991a, 1991b].

[30] It is important to note that most of the above sapping processes occur at a smaller spatial scale than the valley systems attributed to groundwater sapping on Mars. Physical and chemical weathering by emerging groundwater in indurated soils and rock, however, can produce large valley systems. Two types of terrestrial valleys attributed to groundwater sapping have been suggested to be analogs to Martian valleys. The first are valleys developed in layered sedimentary rocks, particularly sandstones, such as the valleys in Navajo Sandstone on the Colorado Plateau, western United States [Laity and Malin, 1985; Howard et al., 1988], alluvial materials of the Canterbury Plain in New Zealand [Schumm and Phillips, 1986], and in limestone scarps in the Sahara desert [Luo et al., 1997]. The second are deep, U-shaped valleys on the flanks of the Hawaiian Island volcanoes [Kochel and Piper, 1986; Baker et al., 1988; Gulick, 1998].

[31] Given their size, morphology, and evidence for widespread layering in the Martian highlands [Binder et al., 1977; Malin and Edgett, 1999], the best terrestrial analog for the Martian valley networks may be terrestrial sapping channels that occur in layered sedimentary rocks. Formation and retreat of box canyons in sandstone on the Colorado Plateau due to groundwater sapping processes has been clearly established [Laity and Malin, 1985; Howard et al., 1988], and many of these valleys show the classic morphology of amphitheater valley heads, weak, stubby canyon development, and nearly uniform canyon width downstream. Valley morphology is, however, strongly controlled by lithology and stratigraphy. Groundwater flow is contained within the Navajo Sandstone, which is overlain by the Carmel Formation and underlain by the Kayenta Formation, both of which are impermeable. Canyon development does not take place until the Carmel Formation has been stripped by surface runoff [Laity and Malin, 1985]. On the east side of the Escalante River, valleys are wider, terminate with amphitheater heads, and occupy a large portion of the drainage basin. The units are oriented so that groundwater flow through the Navajo Sandstone occurs toward the valley headwalls. Simulation modeling indicates that such morphology is characteristic of valley wall retreat that is a positive function of groundwater seepage rates [Howard, 1995]. Even in this environment, however, surface flow remains an important agent as higher discharge rates accompanying storms are needed to remove the larger materials that are eroded from seepage and by mass wasting from canyon walls [e.g., Schumm and Phillips, 1986]. Discharge from seeps with ~20 km² drainage area are about 0.01 m³/s [Laity and Malin, 1985]. Such discharges are sufficient to entrain fine sand, but grains of 1-cm diameter would require a gradient of 0.1, and gravel could
not be moved at reasonable fluvial gradients. On the opposite side of the Escalante River, the Navajo Sandstone is oriented so that groundwater flow occurs away from the valley headwalls. Broad canyons do not form, the terminations of the valleys are tapered, and erosion is controlled primarily by surficial flow [Laita and Malin, 1985].

The case for a groundwater sapping origin for the deep valleys on the Hawaiian Island volcanoes is highly circumstantial. The valleys are suggested to be more strongly eroded than adjacent V-shaped valleys because they tap groundwater reserves trapped by dikes intruded along the major rift zones beneath the volcanoes [Kochel and Piper, 1986]. Inspection of several headward ends of these valleys by one of the authors (Howard) did not reveal conspicuous signs of enhanced weathering or erosion at the bottom of the waterfalls at these locations (such as cavernous weathering or emergent springs). Rather, valley walls are generally smooth and end abruptly at waterfall plunge pools. Howard et al. [1994] suggest that such valleys may erode headward primarily by downward ‘‘drilling’’ by waterfalls. Of related importance is the fact that Kohala, Hawaii receives over 175 inches of rain per year, which makes Hawaii a poor analog to valley network formation in a cold, dry early Mars [e.g., Gulick, 1998].

The principal point to this discussion is that groundwater flow, recharge, and erosion are closely related to precipitations and surface runoff on the Earth. Although the importance of surface runoff and groundwater flow can vary as a function of lithology, stratigraphy, and climate, they do not occur independently of one another. This relation is often disregarded when groundwater sapping is invoked for Mars.

4.2. Valley Network Drainage Density

4.2.1. Factors Influencing Drainage Density

There is great morphologic diversity in what are called valley networks (Figure 1), and a number of classification schemes have been proposed [Pieri, 1979, pp. 39–68; Masursky et al., 1977; Baker, 1982, pp. 34–55]. However, most valley network drainage densities studies have collectively grouped these features together [Pieri, 1976, 1980; Carr and Clow, 1981; Carr and Chuang, 1997]. As a result, drainage densities that have been quantified represent a rough global or, at best, regional average, and these can often be biased by both the resolution of the images and the investigator. At the very least, the differences in valley network morphologies indicate that the geologic conditions that existed while they were forming were not always the same, yet these conditions are not well understood. As explained below, drainage density is a function of many environmental factors.

1. Slope [e.g., Horton, 1945]. On most surfaces, there is a critical slope above which surface runoff concentrates sufficiently that a critical shear stress is exceeded and incision or sapping occurs. This critical slope is a function of flow volume (discharge) and surface material erodibility.

2. Contributing area, slope length, and relief [e.g., Marcus, 1980]. Runoff discharge, whether contributed by runoff or groundwater discharge to the surface, generally correlates strongly with total contributing area. Area, in turn, generally increases with slope length and relief. The roles of contributing area and gradient are explicitly considered in process rules for fluvial incision in terrestrial landscape evolution models [Seidl and Dietrich, 1992; Howard, 1994b, 1995, 1998; Howard et al., 1994; Stock and Montgomery, 1999; Whipple and Tucker, 1999]. Each semiempirical models generally assume that channel erosion rate $\frac{dz}{dt}$ is a function of some measure of flow intensity, $\vartheta$:

$$\frac{dz}{dt} = -K_s(\vartheta - \vartheta_c)^\zeta$$

where $\vartheta_c$ is a critical flow intensity that must be exceeded for erosion to occur, $K_s$ is an erodibility that depends upon bedrock properties, and $\zeta$ is an exponent (most models have assumed that $\zeta$ is unity). The most common assumed measures of flow intensity are the bed shear stress, $\tau$, or the stream power per unit area of channel bed, $\omega = \tau V$, where $V$ is mean flow velocity. Simple equations of continuity, flow

![Figure 1.](image)

Although studies of valley networks tend to treat these features the same, there is a wide variety in morphology. (A) Branching dendritic network northeast of Argyre (Viking image 084A47). (B) ‘‘Fluviatile’’ [Masursky et al., 1977] valley networks on Alba Patera (Viking image 004B57). (C) Small dendritic valley networks (‘‘Filamentous’’ [Masursky et al., 1977]) typified by Warrego Valles (Viking image 063A08). (D) Tributaries (arrows) emptying into a large sinuous valley network (Viking image 084A43). (E) Parallel gullies (left) and a trellis valley network system (right) in southeast Aeolis (Viking image 629A30). (F) Simple dendritic valley network in Libya Montes (Viking image 377S77). These variations indicate that there were differences in how these valley networks formed, suggesting that groundwater sapping was not the only process.
resistance, and downstream hydraulic geometry are usually also assumed [e.g., Howard, 1994b].

\[ \tau = \gamma RS, \]
\[ V = K_\rho R^{2/3} S^{1/2} / N, \]
\[ Q = K_\rho RWV, \]
\[ Q = K_\rho A^e, \]
\[ W = K_\rho Q^b, \]

where \( \gamma \) is the unit weight of water, \( R \) is hydraulic radius, \( S \) is channel gradient, \( V \) is mean velocity, \( N \) is Manning's resistance coefficient, \( Q \) is an effective discharge, \( A \) is drainage area, and \( K_\rho, K_p, K_a, K_w \) are coefficients. The coefficients and exponents are generally assumed temporally and spatially invariant. These, when substituted into (1) allow the erosion rate to be expressed as a function of drainage area and local gradient:

\[ \frac{\partial x}{\partial t} = -K_v(A^e S^b - \theta), \]

where the various coefficients are incorporated into \( K_v \). The exponents have the values \( b = 0.6(1 - b) \) and \( h = 0.7 \) for \( \theta = \tau \), whereas \( g = e(1 - b) \) and \( h = 1.0 \) for \( \theta = \omega \). Measurements of bedrock channel geometry [Montgomery and Gran, 2001] suggest that a value for the width exponent \( b \) of about 0.5 is reasonable, which is identical to the observed exponent for terrestrial alluvial channels. If the water source is from precipitation, channel discharge should reasonably correlate with drainage area, as suggested by (5). In terrestrial landscapes the discharge exponent, \( e \), is about 0.6 to 0.8, although it can be lower in arid landscapes [Benson, 1962, 1964]. Therefore, if dissection results from discharge derived from precipitation, whether directly from runoff or indirectly from groundwater seepage, the intensity of dissection could be expected to depend upon both the steepness and length of regional slopes (the contributing drainage area would increase with the length of the regional slope). The exponent for gradient, \( h \), in (7) is at least two times greater than the exponent for area, \( g \), indicating the dominance of gradient in determining the degree of fluvial incision.

3. Climate [e.g., Gregory and Gardiner, 1975]. Mars has approximately as much land surface area as the Earth, and even today Mars has seasonal and regional climatic variations. Climatic variability on early Mars may have been large. Although Gulick [1998] suggests that all valley networks were formed under the same conditions, the area presented in her Figure 1a is approximately 700,000 km\(^2\) (~260,000 mi\(^2\)), or roughly twice the size as the state of California. San Diego, San Francisco, Lone Pine, and Amboy all have very different climates, which argues that morphology differences and spatial variations exhibited by Martian valley networks may also be related to differences in past climates.

4. Scale. Abrahams [1981] distinguishes between three scales that influence quantitative assessments of drainage densities. Macroscale studies reflect the influence of climate on drainage densities. Mesoscale studies occur within a single climatic regime so that other environmental factors are reflected in the measured drainage densities. Microscale studies occur within a single drainage basin. Because much of the emphasis has been placed on the length of flowing streams [e.g., Gregory and Gardiner, 1975], Abrahams [1984] argues that such studies are inconsistent with the definition of a channel as a trenchlike feature, which is used to calculate drainage densities at the other scales. By comparison, most Martian studies have been essentially macroscale with the assumption that they were probably mesoscale. However, most of the terrestrial data are from mesoscale and microscale studies, and few comparable studies have been performed on Mars. In two notable exceptions, Grant [2000] delimited individual drainage basins in the Margaritifer Sinus region and compare these to terrestrial drainage basins, a microscale study consistent to both planets. He found that Martian drainage densities can in fact reach values comparable to terrestrial drainage basins incised by rainfall and surface runoff. A similar conclusion has been reached by Irwin and Howard [2002] for the Cimmeria region of Mars.

5. Lithology [e.g., Johnson, 1933]. Surface composition, permeability, and stratigraphy all influence drainage density development. While there has been some Thermal Emission Spectrometer data from the MGS mission indicating that the Martian highlands are predominately basaltic [Bandfield et al., 2000], and there is some MOC data suggesting that these materials are layered [Malin and Edgett, 1999], the permeability and stratigraphy of highland materials is largely unknown. Suggestions that Martian volcanoes may be systematically layered by permeable and impermeable layers [Gulick, 1998] is speculative.

6. Preservation. The valley networks are often billions of years old [Carr and Clow, 1981; Dohm and Scott, 1993]. Since their formation they have been subjected to not only mass wasting and eolian infilling [Carr and Malin, 2000], but also impact cratering. Potentially, the smaller valleys have been eradicated while the larger ones have become more subdued. In contrast, most terrestrial runoff and sapping channels systems are still active. Thus, studies such as those presented by Carr and Chuang [1997] should typically result in drainage densities that are lower on Mars than they are on the Earth. It is also likely that Carr and Chuang's [1997] terrestrial estimates were influenced by the fact that plants concentrate around water. Because they used Landsat data to calculate terrestrial drainage densities, and Landsat spectral bandpasses were purposely designed to be sensitive to vegetation, smaller channels should be easier to recognize even if the data were degraded to a resolution comparable to Viking.

7. Time or Stage of Development. Glock [1931] showed that the drainage density of an area will change as the system evolves through time. Currently, it is unclear how mature any drainage basins are on Mars. Of particular importance is the fact that Martian valley networks were forming simultaneously with impact cratering [Baker and Partridge, 1986; Irwin and Howard, 2002]. Simulation modeling by A. D. Howard (in preparation) shows that impact events could have frustrated valley network development as the drainage area was constantly evolving. Under
such conditions, the resulting drainage densities from surface runoff would be lower because of the constantly changing slopes and the competition for surface area from impact craters.

[35] In summary, there are many factors that influence drainage densities. Unless these factors are identical, comparisons between Martian valley networks and terrestrial runoff or sapping channels cannot be made reliably. To date even the data sets used for such analyses have not been entirely comparable. It is also important to note that there are no absolute drainage densities that can be used to distinguish between channels formed by runoff and those formed by groundwater sapping. For example, sapping channels developed in sandy sediments on terrestrial beaches and riverbanks exhibit very high drainage densities. The intended purpose of evaluating terrestrial drainage densities is not to determine whether channels formed by runoff or sapping. Rather, it is used to measure the topographic texture or linear scale of landscapes and predict runoff characteristics [Abrahams, 1984]. As such, it is odd that so much emphasis has been given to the qualitative assessment of "low" Martian drainage densities and that many investigators have attributed it a priori to groundwater sapping. Few efforts have been made to quantify these densities and evaluate what they would imply about actual valley network runoff characteristics. The tacit assumption in the cold, dry early Mars scenario is that somehow hydrothermal circulation would allow groundwater sapping to provide these unquantified characteristics. On the Earth, precipitation is an agent in any drainage basin. Given the right environmental factors the resulting surface runoff and groundwater sapping produce a wide range of drainage densities.

4.2.2. The Spatial Distribution of Valley Networks

[36] Comparing factors controlling the distribution of channel incision on Earth and the Martian highlands has been handicapped by low resolution image data (Viking coverage was generally lower than 200 m/pixel) and the lack of detailed topographic information. The lack of topographic data has been particularly crippling because of the overwhelming importance of gradient in determining the rate of fluvial incision (Factor 1, above). With the advent of Mars Global Surveyor data better, more accurate comparisons between valley networks and terrestrial drainage densities can be made. It is now possible to perform a simple, straightforward test to determine the importance of precipitation versus hydrothermal circulation by investigating the relationship between topography (i.e., slope and contributing area) and the location of valley network systems. Gulick [1998, 2001] stated that valley networks occur as isolated features or in clusters, suggesting localized water sources such as what might be expected from hydrothermal activity. If precipitation were involved, however, the spatial variation in the degree of dissection (whether resulting from surface runoff or groundwater sapping) should correlate with both the steepness and length of regional slopes. MOLA data were utilized to produce a topographic map for a portion of the Margaritifer Sinus/Sinus Sabaeus region (Figure 2a). We then independently distinguished between two qualitative degrees of valley network drainage densities using Viking images at resolutions ranging from 50m/pixel to 220m/pixel (Figure 2b). Heavy incision consists of

Figure 2. With the advent of MOLA data, it is now possible to determine the relationship between dissection and topography. On a portion of the cratered highlands of Mars (A) MOLA topography of an area in Margaritifer Sinus/Sinus Sabaeus centered on $-22.5^\circ$ lat., $357^\circ$ long. The contour interval is 100 m, and the scale is in kilometers. (B) Photomosaic for the same region. (C) Degree of dissection of the region shown above based upon Viking images with resolution of 50–230 m/pixel. Unshaded areas have few obvious valleys, lightly shaded areas are moderately dissected, and dark areas appear to be completely dissected into valley and ridge systems. The major valley systems are shown in dash-dot lines. Note the strong correlation between drainage density and slope, which is consistent with rainfall.
closely spaced channels (drainage density exceeding 0.1 km/km² and locally approaching 1.0 km/km²), which on MOLA profiles typically indicates valley depths of about 20–150 m. Moderate incision includes regions with recognizable valleys of lower drainage density. The major trunk valleys in this area are depicted with dash-dot lines on Figure 2c. The dissected interior crater rims were not mapped except for the two larger, >100 km basins (Newcomb and an unnamed one).

The extent of dissection correlates strongly with the steepness and length of regional slopes. The heavily dissected terrain occurs on the steepest parts of interior or exterior crater rims and on the steeper scarp bordering flat highlands (e.g., at −21 to −24° lat., 1.5° long., and −24° lat., 355° long.). Moderate dissection also largely corresponds to long, relatively steep regional slopes as well as a zone surrounding the north-trending unnamed trunk valley bisecting the region at about 356° long. Areas free of apparent dissection correspond to basin floors as well as the flat part of major basin rims and flat uplands, such as those centered at −24 to −26° lat., 352–355° long., and −27° lat., 359° to 3° long.

From this analysis we conclude that the spatial variation in the degree of dissection of the study area (Figure 2) correlates strongly with both the steepness and length of regional slopes, as would be expected from models of fluvial incision resulting from distributed precipitation. If dissection resulted from local subterranean sources of water, such as thermal anomalies from igneous intrusions or crater hydrothermal systems, as suggested by Gulick [1998, 2001], then such a correlation should not be expected. In particular, it would be likely that only some of the steep regional slopes would be strongly dissected [Gulick, 1998]. This is not the case. Similarly, Tanaka et al. [1998] found the degree of incision in the Thaumasia region correlates with closeness to large impact basins. Although they attribute this correlation to the greater potential for hydrothermal activity in large basins, an alternative explanation is that greater relief associated with larger impact basins produces greater potential for precipitation-induced fluvial incision. This is demonstrated by the strong dissection associated with the 250-km-diameter Newcomb impact basin centered at −24° lat., 1° long. in our study area (Figure 2).

If the source of flow that carved the Martian valley systems were dominated by hydrothermal sources, then mainstreams deriving their water from distant, localized sources would commonly erode isolated valleys with sharp, steep valley walls and few tributary valleys. If main trunk valleys did have local sources of hydrothermal water creating tributary valleys, such sources would be unlikely to have been distributed symmetrically on opposite sides of the trunk channel or uniformly along the trunk valley. However, dissection along trunk valleys commonly occurs along nearly the entire length of incised portions of the valley (e.g., Figure 2c). Similarly, tributary valleys are symmetrically developed along a 400 km incised segment of Loire Valles in the Margaritifer Sinus region (Figure 3a). In this section, Loire Valles is incised about 400–500 m below a modestly dissected plateau (Figure 3b). Tributary valley systems occur on both sides of Loire Valles in a broad valley system about 30–50 km wide. Nearly uniformly distributed sources of surface flow resulting from precipitation-fed runoff or groundwater seepage are the most plausible agent of fluvial erosion, as suggested by Grant [2000]. Again, the extent of fluvial incision in the Loire Valles region (Figure 3a) correlates strongly with the steepness of the regional slope (Figure 3b). Although the incision of Loire itself has provided most of the local relief, strong incision also occurs along the flanks of valleys tributary to Loire Valles, on the steep south-facing slope of the basin at the southeast corner of Figure 3 (−20° to −21° lat., 12–14° long.), as well as on every other surface with a long, steep, regional slope. We conclude that the spatial distribution of valley networks in the Martian highlands reflects local topography, which is indicative of a rainfall source for runoff or groundwater sapping.

4.3. Valley Network Morphology

Many valley networks have rectangular cross-sections and flat floors with amphitheater heads, which is often characteristic of valleys developed by groundwater sapping [Pieri, 1980; Carr and Clow, 1981; Baker et al., 1992; Carr, 1996, p. 72; Gulick, 1998, 2001]. Unfortunately, it is often assumed that all valley networks share these characteristics, which implies that all valley networks formed by sapping,
which is then used as evidence to invoke a cold, dry early Mars. However, some valley networks are clearly V-shaped and do not have an amphitheater head, which implies that they were developed by surface runoff [Goldspiel et al., 1993; Williams and Phillips, 2001]. This cannot be explained in a cold, dry early Mars scenario where groundwater sapping is invoked as the only working geologic process. Furthermore, a rectangular cross-section or amphitheater head is not always indicative of groundwater sapping as runoff valleys and channels can often share these same characteristics. Thus, we suggest that the ubiquity of groundwater sapping as an erosive agent on early Mars may have been overinterpreted. Finally, younger processes appear to have modified many valley networks [Goldspiel et al., 1993; Carr and Malin, 2000; Williams and Phillips, 2001]. V-shaped valleys would become more rectangular in cross-section as they are degraded and infilled, further causing groundwater sapping to be overinterpreted. The evidence for such processes may imply that the early geologic and climatic conditions shifted gradually into those present today.

4.3.1. Valley Network Cross-Sections

[41] In simplest terms a fluvial channel is either eroding or transporting sediments. When a channel is at equilibrium (i.e., “graded” [Mackin, 1948]), erosion occurs primarily in the headwater area and the upper reaches of the tributaries. However, should the climate change, the terrain be uplifted, or the base level decreased, the entire length of the channel will adjust by downcutting into the landscape. When a channel or portion of a channel is actively eroding, the valley it occupies typically assumes a V-shaped profile. This type of valley morphology is unique to erosion by surface runoff in that it is the most efficient shape for focusing the energy from the surface downward. A V-shaped valley profile is, therefore, almost always indicative of surface runoff. While a rectangular cross-section is a morphologic characteristic associated with many valley networks, there are clear instances of valleys with V-shaped profiles. From topographic data, Goldspiel et al. [1993] have shown that Maja Valles are V-shaped, while Williams and Phillips [2001] demonstrated that V-shaped profiles are fairly common. These observations support our assertion that incision from surface runoff occurred on Mars.

[42] Conversely, erosion by groundwater sapping occurs along the boundary between the aquifer and the underlying aquitard, so the resulting valley morphology is predominately rectangular. However, a valley with a rectangular cross-sectional morphology is not strictly diagnostic of groundwater sapping. Surface runoff can generate canyon-like valleys when headward erosion occurs into scarp bounding a flat upland or incision takes place through a resistant layer, such as duricrust or a layer of volcanic material. Both types of materials may be widespread on Mars [Binder et al., 1977; Malin and Edgett, 1999]. More commonly, valleys assume a rectangular shape when a channel is graded because downcutting rates diminish and the valley is permitted to become wider due to channel meandering and erosion of the valley walls.

[43] Channels also frequently have rectangular profiles. In perennial streams where sediment can be moved from both the bed and the banks, Henderson [1961] states that two conditions must be met simultaneously: the shape of the channel must be able to transmit the flow, but the banks must also remain stable. As water flows through the channel, it exerts a shear stress along the bed and banks that is a function of the roughness of the channel boundary, the slope, and the depth of the flow (e.g., equations (2) and (3)). In an ideal environment where the physical characteristics of the material are uniform and channel maintains a constant flow depth and slope, Koechlin [1924] calculated that the exerted stresses would produce a channel with a sinusoidal shape. In nature, however, the amount of discharge in a channel is often highly variable. Because the banks of the channel are unable to stand the stresses exerted by deeper flows, they give way and the central portion of the channel simply enlarges, resulting in a channel with a rectangular profile [Leopold et al., 1964, p. 201].

[44] Although a rectangular cross-section is a common morphologic attribute of perennial streams [e.g., Leopold and Maddock, 1953], a discussion of ephemeral streams may help emphasize our point better because they are fed by primarily by rainfall and surface runoff following storms and the contribution by groundwater to channel flow is extremely small [Mabbutt, 1977, p. 151]. Because they are storm fed, flow in ephemeral streams is short-lived, peak discharges occur as pulses, drainage networks are poorly connected, and flooding is localized and highly variability within the drainage system [Knighton and Nanson, 1997]. Concentrations of sediments in arid streamflows is usually very high due to the sparse vegetation of the surrounding uplands [Mabbutt, 1977, p. 158]. Typically most of the sediment carried in these floods is sand, which is supplied by partial weathering and the effectiveness of slope wash to carry such particles [Schumm, 1961]. Consequently, >35% of the sediment is often transported as bed load [Schumm, 1961]. Often, however, material is transported only for short distances because transmission losses are usually high due to high evaporation rates [Leopold et al., 1966]. To adjust for the high amounts of bed load, the stream conforms to a rectangular cross-section where shear stresses are high over the bed and flow velocities are low along the erodible banks (Figure 4). The width and depth ratio is balanced by the resistance of the bank to decreasing lateral scour [Mabbutt, 1977, p. 160]. Channels become deeper if more silt and clay is available for transport because these sediments can be transported in suspension (i.e., “mixed load” [Schumm, 1961]). Lateral erosion will then cease, and the high shear stresses will cause scouring into the channel bed. The result is always a channel with a rectangular or trapezoidal cross-section.

[45] The principal point in this discussion is that there is nothing about the morphology of valley network morphology that is inconsistent with precipitation and surface runoff. While V-shaped valley networks may provide the best evidence that incision from surface runoff occurred on Mars, rectangular valley networks may also reflect this. In fact, the spatial relationship between V-shaped and rectangular valleys observed by Williams and Phillips [2001] provides strong support for such an argument. Using topographic data from MOLA, they found that V-shaped valleys occur upstream of rectangular (i.e., “U-shaped”) valleys. They interpreted these morphologies as representing two separate phases of valley network development: a runoff phase followed by a groundwater sapping phase. Baker and
counterparts. The opposite should be true if significant
often shallower than their rectangular (i.e., "U-shaped")
[Williams and Phillips, 1999] that small craters >10 m deep have not been obli-
This result agrees with observations by
Hartmann et al.
incision and downcutting typically occur in the upper
reaches of a valley, resulting in a V-shaped profile, while
further downstream the valley assumes a rectangular shape
due to channel meandering and continuing erosion of the
valley walls. The valley network morphology described by
Williams and Phillips [2001] is what one would typically see
if water were suddenly removed from the rivers on the Earth.

4.3.2. Modification of Valley Network Morphology
[46] We argued earlier that smaller order tributaries may
not have been preserved due to degradation by mass wasting,
eolian infilling, and impact cratering, but these processes
would also modify the main trunk channels as well. How-
ever, the extent to which such modification may have
occurred is not well known. Carr and Malin [2000] sug-
gested that the flat floor morphology of many valley net-
cworks could be the result of eolian infilling, but this does not
seem practical. Assuming these valleys began with a V-
shaped profile, the morphometric measurements from Gold-
spiel et al. [1993] indicate that it would often take tens (if not
hundreds) of meters of infilling to create a significantly
broad flat floor. Analysis of degraded impact craters indi-
cates that only ~10 m of eolian material was deposited in
topographic lows in the highlands [Craddock et al., 1997].
This result agrees with observations by Hartmann et al.
[1999] that small craters >10 m deep have not been obli-
terated on Mars. There is simply no independent evidence
supporting the necessary amount of infilling. In addition,
Williams and Phillips [2001] found that V-shaped valleys are
often shallower than their rectangular (i.e., "U-shaped")
counterparts. The opposite should be true if significant

infilling had occurred. This implies that many of the dunes
observed at the bottom of valley networks [Carr and Malin,
2000] may contain a significant fraction of reworked fluvial sediments.

[47] It is apparent that valley networks have experienced
some modification, but how much is unclear. Morphometric
measurements made from photoloclinometric profiles indicate
that the walls of most valley networks are not steep, but
rounded, which Goldspiel et al. [1993] interpret to be the
result of extensive modification by mass wasting. However,
MOLA data suggest that the amount of modification that
has taken place is minimal [Williams and Phillips, 2001].
Currently it is unclear how much modification has taken
place, and this may be important. “Extensive” mass wast-
ing resulting in backwasting and infilling would cause the
valley networks to become shallower and more rectangular,
perhaps causing the importance of groundwater sapping to
be overinterpreted. Equally unclear are the conditions driv-
ing the modification. What are the processes involved and
could they be supported under current climatic conditions?

[48] On Earth, mass wasting can influence channel mor-
phology. For example, terrestrial canyons eroded by surface
runoff typically have valleys that gradually narrow head-
ward [Laity and Malin, 1985; Howard, 1995], but subse-
quent mass wasting tends to widen and shallow these
valleys, creating rounded valley heads that are difficult to
distinguish from sapping valleys [Howard, 1995]. However,
the mechanics of this process is poorly understood. Laity
[1983] performed a petrographic study of rock samples
from seepage faces in Navajo Sandstone canyon heads,
and noted that intergranular deposition of calcite disrupted
the sandstone fabric. Winter freeze-thaw may also contrib-
ute to weathering [Laity, 1983], but any weathering pro-
cesses would also clearly depend strongly on the physical
and chemical composition of the host rocks and the ground-
water chemistry. Whether similar processes occurred on
Mars is unknown.

[49] Small diameter impactors can also induce modifica-
tion of the surface. Such processes are extremely important
for degrading small-scale features on the lunar surface [e.g.,
Craddock and Howard, 2000], but they are extremely slow.
Such “meteorite creep” is also driven by the dust particles
shed by comets as they approach the Earth [Shoemaker, 1983],
but it is unclear whether such small particles could
survive passage through the Martian atmosphere, particu-
larly one that may have been denser in the past. It is also
important to note that “meteorite creep” on Mars is unlikely
to have been of any great importance at larger scales, since
similar modification is only discernable in lunar impact
craters that are <5-km in diameter [Craddock and Howard,
2000].

4.4. Valley Network Age
[50] Several investigators have stated that the principal
inconsistency between a warm, wet early Mars and the
geologic record is that valley networks continued to form
in localized regions long after this climate was supposed have
90]. In particular, valley networks on the young Tharsis
volcanoes appear to record an episode of fluvial activity that
occurred after activity ceased in the highlands [Carr and
Clow, 1981; Gulick and Baker, 1990]. Such an apparent late
episode of fluvial activity recorded on the Tharsis volcanoes has been explained by a climate change caused by Oceanus Borealis, which Baker et al. [1991, 2000] suggest was the result of outflow channel formation and the release of a putative CO$_2$ clathrate from buried permafrost. Evidence for an ocean in the northern plains was first presented by Parker et al. [1989], and it has been recently supported by topography derived from MOLA data [Head et al., 1999].

[51] The belief that there were two separate episodes of valley network formation is based on the presumption that all valley networks in the Martian are ancient because the highlands are old [Carr and Clow, 1981; Gulick and Baker, 1990]. In fact, “ancient” is an adjective commonly used to describe all valley networks in the Martian highlands. However, detailed crater-age dates reveal that valley networks formation in the highlands continued up to the Amazonian [Dohm and Scott, 1993]. Some investigators have also suggested that valley networks have been preserved in various stages of modification [Baker and Partridge, 1986; Goldspiel et al., 1993; Carr and Malin, 2000], which would further suggest that they are not all the same age; however, such interpretations are arguable [Williams and Phillips, 2001]. Perhaps a stronger argument can be made that many impact craters in the highlands are in various stages of modification. Fresh crater populations, which record the time at which surfaces became stable, indicate that highland degradation was also operating continuously up until the early Amazonian [Craddock and Maxwell, 1993].

As we explain below, such processes were coupled to valley network formation. Essentially, these analyses reveal that there is no discontinuity between the ages of highland valley networks and those found on the Tharsis volcanoes. The geologic record does, in fact, support continued erosion and modification in the Martian highlands throughout the early history of Mars, implying long-lived episode of valley network formation whose rate diminished through a gradual climatic change. This gradual decrease in erosion rates would have been accompanied by short-term variability due to major impacts, episodes of volcanic activity, and quasi-cyclical climate variations.

4.5. Hydrothermal Models for Valley Network Formation

[52] The possibility that valley networks on the Tharsis volcanoes were temporally separate from those found in the highlands [Carr and Clow, 1981] as well as the warmer, wetter climate that may have created them needed an explanation. This inspired Gulick and Baker [1990] to first suggest a hydrothermal mechanism for valley network formation. In their model, magmatic intrusions heat the local groundwater which creates a density contrast, which subsequently initiates a buoyancy-driven flow. Eventually, warm groundwater was released to the surface to carve the valley networks on the volcanoes.

[53] As acceptance of the cold, dry early Mars paradigm increased, the highland valley networks also faced a similar paradox. How could they form if early climatic conditions did not support liquid water at the surface? Squyres and Kasting [1994] proposed that geothermal heating following accretion coupled with the heat released from impact cratering may have also driven hydrothermal convection and groundwater seepage in the highlands. Both heat mechanisms may have been naturally more pronounced during the early history of Mars. Because accretion occurs much more rapidly than conductive heat loss, there should have been substantial melting of the Martian interior [Coradini et al., 1983]. Also, analysis of some shergottite-nakhlite-chassignite (SNC) meteorites indicates that Mars probably differentiated very early [Shih et al., 1982]. The heat released from accretion and core formation is much higher than the heat released slowly through radiogenic decay, and Squyres and Kasting [1994] estimate that the global mean geothermal gradient during the first 500 million years on Mars may have been a factor of 12 times greater than today’s. Under current climatic conditions, they suggest that early higher heat gradients may have allowed liquid water to exist within 150 to 300 m of the surface [Squyres and Kasting, 1994], a depth comparable to many valley networks [Goldspiel et al., 1993]. Potentially these depths would have been shallower if the groundwater was saline or if there were local areas where magmatic activity was concentrated [Squyres and Kasting, 1994; Gulick, 1998]. They also suggest that the early, higher geothermal gradient may have also been aided by the generation of impact melt because the flux was greater during the period of heavy bombardment. Newsome [1980], Brakenridge et al. [1985], and Newsome et al. [1986] suggest that these mechanisms may have caused ground ice to melt, circulate through the regolith, and subsequently carve the valley networks through sapping. Besides excluding the important roles of rainfall and surface runoff in forming terrestrial sapping channels, these models face a number of other difficulties that are frequently overlooked.

4.5.1. Recharging the Aquifer

[54] Recharge of the aquifers feeding the valley networks is necessary to maintain their development. Without recharge, aquifers would be emptied quickly before much erosion could take place. On the Earth, aquifer recharge is supplied by rainfall. In a cold, dry Martian environment, however, rainfall is not supported, and the greatest difficulty for hydrothermal models of valley formation is providing a mechanism for aquifer recharge.

[55] In one proposed mechanism water evaporated from Oceanus Borealis and eventually precipitated as snow on the Tharsis volcanoes [Baker et al., 1991]. Gulick et al. [1997] suggest local heating from underlying magmatic intrusions subsequently caused the snow to melt and infiltrate into the ground. There are a number of problems with such a recharge mechanism, however. (1) Baker et al. [1991] suggested that Oceanus Borealis resulted from discharge from the large outflow channels, and was thus present only during the Amazonian. If so, there needs to be an alternative mechanism for recharging the majority of highland valley networks that formed in the Noachian and Hesperian. (2) Snowfall would not concentrate on top of a magmatic intrusion; rather, it would be widely distributed across the planet. Snowfall would eventually accumulate in areas not heated by magmatic intrusions, eventually cutting off the supply. (3) It is ad hoc to suggest that CO$_2$ heating allowed liquid water to accumulate into an ocean [Baker et al., 1991, 2000], yet these conditions did not support liquid precipitation. If rainfall did occur, it is no longer necessary to melt snow through magmatic intrusions. Hydrothermal circulation is also no longer necessary.
Another proposed recharge mechanism is that groundwater released to the surface from geothermal buoyancy-driven flows simply infiltrated back into the system downstream [Gulick and Baker, 1989; Squyres and Kasting, 1994; Gulick, 1998; Goldspiel and Squyres, 2000]. Water may also be drawn into the system from the surrounding area [Gulick and Baker, 1989, 1990; Gulick, 1998]. While diagrams in these models show water being released on the flanks of steep volcanoes or recharge occurring from broad areas, topography is not considered in the actual numerical models. It seems unlikely that water would be very efficient at traveling hundreds of kilometers underground while rising several kilometers in elevation. In addition, the ambient surface pressure on Mars must have been 1 bar to prevent boiling during the valley network development [Gulick, 1998]. With a denser atmosphere in place, there are a variety of mechanisms that could also support warmer surface conditions resulting in rainfall.

Clifford [1993] has suggested an additional method of recharge from below that meets some of the inherent difficulties in simple hydrothermal circulation. If surface temperatures average below freezing, upward vapor flow driven by the geothermal heat gradient might form an icy permafrost near-surface layer. Melting of this permafrost could supply limited quantities of water for hydrothermal circulation, but only in the initial stages of any hydrothermal source. If the surface temperatures oscillated repeatedly above and below freezing, alternate formation (from the top down) and melting (also from the top down) of the permafrost layer could episodically release groundwater to the surface, possibly at fairly high elevations in the landscape. Mars throughout its history has probably had extreme variations of obliquity and eccentricity [Ward, 1979]. Although the permafrost recharge model seems theoretically plausible, we presently have no direct or indirect observations that can test the model. It also remains to be seen whether melting of any permafrost layer could occur rapidly enough to supply the discharges necessary to carve the observed valleys. Additionally, such a model could also be troubled by the exponential decline in the global heat flux, suggesting that the required model parameters would be short-lived (described below), which is not supported by the long interval during which valley network formation and crater degradation took place.

A related problem to recharge is the large amount of water that must be released to the surface from hydrothermal circulation. To produce the volume of erosion observed in the valley network systems, large quantities of water must be discharged. Typically, groundwater sapping of cohesionless sand may require as little as 10 times the volume of sand eroded to form the valleys [Howard and McLane, 1988], but erosion of cohesive soils and indurated rocks generally requires a water to rock ratio of at least 10^4 to 10^5 [Howard, 1988; Goldspiel and Squyres, 1991]. This requires regional groundwater levels to be very close to the surface as well as a subsurface permeability exceeding 1 darcy [Gulick, 1998], which is not typical of most geologic materials. The problem is that most of the cratered highlands lie 5 to 10 km above the Martian lowlands (the northern plains and the floor of Hellas). Simulations of global groundwater flow by Howard [1991] indicate that more than 75% of groundwater in the highlands would migrate to the northern plains within 10 to 50 million years given an average permeability of 1 darcy, and water levels would drop significantly below the surface in much less time. If lateral flow of groundwater was restricted by permeability barriers, such as unfractured rock or permafrost, hydrothermal activity would quickly exhaust nearby surface supplies of water also leading to the quantitative supply issue discussed above.

The Depth of Melting

Hydrothermal models sometimes place an ad hoc control on where melting occurs in the Martian regolith. Squyres and Kasting [1994], for example, suggest that geothermal heating caused ground ice to melt within a few hundred meters of the surface, thus carving the valley networks. However, the exponential decline in the global heat flux they describe means that subsurface melting occurred at increasingly greater depths with time. Valley network formation would be quickly shut off. Conversely, the heat flux would have been greater in early Martian history and above freezing temperatures would have been even closer to the surface. Basically, there is only a brief period when geothermal heating could provide above freezing temperatures precisely at the depths of the valley networks. Thus, it cannot explain the range in ages of the valley network systems [Dohm and Scott, 1993].

The Missing Erosion Mechanism

Terrestrial sapping channels are often initiated by surface runoff, and sapping begins once incision intersects the water table [e.g., DeVries, 1976]. Alternatively, seepage occurs at the base of scarp and headward erosion creates a sapping channel [e.g., Onda, 1994]. It is unclear which mechanism would work in the proposed hydrothermal models. It seems unlikely that surface erosion could intersect the water table without precipitation. Likewise, there is scant evidence for scarps at the base of the valley networks. A related point is that hydrothermal seeps generally deposit more than they erode because of the heavy dissolved load and the decrease in temperature and pressure when water erupts at the surface (e.g., the Yellowstone hydrothermal district), which further undermines the effectiveness of such a process eroding the Martian valley networks.

The Lack of Surface Manifestations Related to Geothermal Systems

On Earth, excavation of valley systems by hydrothermal discharge is extremely rare. Yet, there are a large number of common hydrothermal features that have not been identified on Mars. Depending on the reservoir temperature and the discharge rate of heated water, typical surface evidence of terrestrial geothermal systems includes seeps, fumaroles, hot springs, boiling springs, geysers, phreatic explosion craters, zones of acid alteration, deposits of silica sinter, travertine, and bedded breccias surrounding phreatic craters [Wohletz and Heiken, 1992, pp. 119–140]. Although the geochemical signature from hydrothermal activity could be masked, its absence undermines such models. Because pyroclastic deposits resulting from hydrothermal activity could be a precursor to many of the highland paterae [e.g., Crown and Greeley, 1993], there should at least be a few distinct examples of phreatic explosion craters from hydrothermal activity, but such features are not recognized. Surface manifestations of geothermal processes should be especially pronounced near the
source areas of the valley networks in the Gulick [1998] model, but there is no such evidence. Although there is indication from the SNC meteorites indicating hydrothermal flow did take place at depth on Mars [McSween, 1994], the lack of mineral deposits and geomorphic features analogous to those on the Earth indicates that this process never worked with any efficiency near the surface. If Martian hydrothermal systems were as widespread as implied by the abundance of valley networks in the highlands, then there should be a number of mineral deposits and additional surface features reflecting this phenomenon.

4.5.5. The Origin of Ground Ice
[62] There is abundant evidence for ground ice on Mars [Squyres et al., 1992]. The tacit assumption in the hydrothermal models is that ground ice (or water) is already present in the crust to melt and carve the valley networks or be released in catastrophic outbursts [Baker et al., 1991, 2000]. However, no explanation has been presented as to where this water originally came from. Any volatiles brought into Mars as it was accreting would be vaporized during impact. Volatiles released during subsequent outgasing would also be in the form of vapor. The easiest way to place water into the Martian crust is to have it condense and then precipitate into the regolith, which necessitates warmer surface temperatures early in Martian history.

4.5.6. Impact Melt and Hydrothermal Systems
[63] Brakenridge et al. [1985] proposed that melt sheets emplaced during impact crater formation interacted with ground ice and drove local hydrothermal systems. In their model, permeable outcrops allowed hot springs to release warm water to the Martian surface. Valley networks resulted through a combination of headward sapping and down-valley water and ice flow. Again, Squyres and Kasting [1994] also suggested that heat from impact melt may have been an additional mechanism for raising the global heat flux because of the high impact rate during the period of heavy bombardment. Gulick [1998] notes that because melt sheets blanket the surface, they probably have a different cooling history than typical magmatic intrusions, and they are probably less efficient at cycling water to the surface. Gulick et al. [1988] calculated that only craters larger than ~100 km

Figure 5. (opposite) A number of observations severely undermine the validity of models that suggest groundwater circulation was driven by heating from impact melt. (A) There is no clear relation between crater diameter and the onset of dissection. This 50-km-diameter crater, for example, is highly dissected by valley networks, yet according to existing models [Gulick et al., 1988; Pope et al., 2000] is too small to have generated enough impact melt to drive hydrothermal systems. (MOC image M0303829.) (B) Valley networks appear to originate from this impact crater, yet the ejecta and associate impact melt deposits have been eroded. There is no apparent heat source supporting these systems. (MOC images M0020004, M0801747, M0900475, M1103094.) (C) Young valley networks in Bakhuysen have incised the crater up to the rim crest. It would be difficult for geothermal heating to release enough groundwater at this elevation to cause channel incision. (MOC images M0304865-6.) Such valley networks also attest to the fact that not all are “ancient.” (D) If climate is not a factor as suggested by Brakenridge et al. [1985], then impact melt should cause valley network to form on all impact craters, including those forming today. However, fresh, unaffected impact craters and basins form adjacent to similar diameter impact structures that have been incised by valley network systems. The largest fresh basin in this area is marked with an “F” while smaller basins containing valley networks systems (arrows) are marked with an “M.” The area shown is in Aeolis between −22.5° to −27.5° lat., 192.0° to 200.0° long.
in diameter would be capable of sustaining a hydrothermal system long enough to produce an integrated valley network system. However, because of the lower impact velocity and effects of gravity scaling on Mars, Pope et al. [2000] calculate that \( \sim 100 \) km is the threshold diameter for generating a melt sheet free of any breccia, but such a melt sheet would be too thin and cool too quickly to provide an effective hydrothermal system. Their results suggest that impact basins with diameters >600 km are needed.

[65] Similar to magmatic intrusions, the impact melt model is appealing because it does not require atmospheric conditions any different than those that exist today [Brakenridge et al., 1985]. However, there are several observations that make large impact melt driven hydrothermal systems implausible. (1) Craters much smaller than 100 km diameter are frequently associated with integrated valley network systems (Figure 5a). Essentially, there is no ubiquitous onset diameter. (2) Valley network systems are typically associated with heavily degraded impact craters where the impact melt sheet (along with the crater rim) has been completely eroded (Figure 5b). There would no longer be an impact melt deposit driving hydrothermal circulation in such instances. (3) It would be difficult to place enough water at the crater rim crest through either hydrothermal circulation or vapor diffusion [Clifford, 1991] to induce valley network incision (Figure 5c). Basically, there is a finite amount of water that could be released at the crest of a crater rim from groundwater seepage, but an infinite amount of water could be precipitated on this surface. (4) If different climatic conditions are not required, then valley networks should form on all impact craters (with the caveat on size), including those forming today. However, there is an entire population of fresh, unaffected impact craters in the Martian highlands, many which have formed next to craters incised by valley network systems (Figure 5d). Something has to be different to have allowed these craters to remain undissected. (5) Kieffer and Simonds [1980] predict that impacts into volatile rich targets (such as the Martian highlands) would result in impact melt sheets dispersed as fine particles. This fine dispersal of particles caused by the expansion of shocked volatiles during crater formation would cool quickly and form glass, which would further reduce the effectiveness of any hydrothermal circulation. Recent assessments of TES data reveal that the surface materials on Mars do consist of a glass component [Bandfield et al., 2000]. While this signature is strongest in the northern plains, significant concentrations (\( > \sim 30\% \)) are distributed across the highlands as well.

5. Geologic Evidence From Modified Highland Impact Craters

[65] In addition to valley networks, the Martian highlands contain a large population of impact craters that have undergone extensive modification. Because crater modification was occurring simultaneously with valley network formation (Figure 6), the geologic processes responsible for creating one set of features should be consistent with the other. Previous analyses of crater modification have been both qualitative [e.g., Craddock and Maxwell, 1990, 1993] as well as quantitative [Craddock et al., 1997]. The morphology of these craters indicates that they were eroded as opposed to being buried or aggraded. Because craters are preserved in a variety of degradational stages, modification must have taken place over a long period of time as the craters were forming. Topographic data from both fresh and modified crater populations were analyzed to determine the extent to which erosion took place. Using a computer simulation model [Howard, 1994b], we explored the effects a variety of geologic processes would have on crater morphology from known rate laws. These results indicate that rainfall and surface runoff were the most important geologic processes in the early history of Mars. There is also evidence that the intensity at which these processes operated waned with time.

5.1. Evaluation of Crater Morphology

5.1.1. Evidence for Widespread Erosion on Early Mars

[66] One of the first observations made from early Mariner data was the Martian craters did not look like those seen on the Moon [Leighton et al., 1965]. To explain the unique morphology of the Martian craters, many subsequent investigators proposed burial by air fall deposits [Hartmann, 1971; Wilhelms and Baldwin, 1989; Grizzaffi and Schultz, 1989; Grant and Schultz, 1990; Moore, 1990]. This interpretation may have been influenced by the prevalence of ongoing eolian processes on Mars, but also by the presence
of large air fall deposits in the regions where many of these studies took place (e.g., Arabia [Christensen, 1982, 1986]). The problem with a uniform blanket of material, such as an air fall deposit, lava flow, and even many sedimentary deposits, is that such materials do not explain the size range of modified craters. Typically, these deposits obtain a fairly uniform thickness, and would thus obliterate craters with depths less than the deposit, obscure all but the rims of craters similar in-depth to the deposit, and would only modestly affect deeper, larger craters. This would result in a narrow range of crater diameters with the characteristic flat floor, rimless morphology. Erosional processes, on the other hand, would affect a wider range of crater diameters similarly. A small diameter crater would be eroded to the same general morphology as a larger diameter crater only faster. This scale dependency means that smaller diameter craters may be completely eradicated, which has been argued by Jones [1974] and Chapman and Jones [1977] to explain the deficit in the number of smaller craters observed in size-frequency distribution curves of highland crater populations.

[67] The emplacement of a contiguous, uniform deposit is also inconsistent with the timing of highland erosion as well as the temporal relation between valley networks and modified impact craters. Although Wilhelms and Baldwin [1989] suggested that valley incision took place after a volatile-rich air fall deposit was emplaced, crater modification was a long-lived process that continued to erode impact craters as they were forming [Craddock and Maxwell, 1993; Craddock et al., 1997]. An air fall deposit emplaced quickly does not explain the stages in crater modification. If it were emplaced slowly over time, then the deposit would quickly bury the smaller valley networks before affecting the larger impact craters; however, both valley networks were clearly forming simultaneously while craters were being modified (Figure 6). Alternatively, Clifford [1997] proposed that highlands craters were modified by liquefaction induced by the seismicity following large impact events. However, such an erosional mechanism would simultaneously destroy the smaller valley networks, and this is clearly not the case. More recently, Malin and Edgett [2001] proposed that erosion, deposition, and impact cratering occurred simultaneously, resulting in a complex interbedding of craters and previously exposed surfaces. However, they do not identify the erosion mechanisms or the scale that the resulting “heavily cratered volume” would become important. Their model appears to explain large layered deposits in the highlands that are fairly isolated [Malin and Edgett, 2000a] as well as exhumed craters tens of meters in diameter, but not the modified morphology of larger diameter craters that are prevalent throughout the highlands. From the analysis of crater populations in the Martian highlands, we proposed that only an erosional process, most likely fluvial, could explain the morphology and the size-range of affected craters (2 km to basin size impacts) [Craddock and Maxwell, 1990, 1993; Craddock et al., 1997].

5.1.2. Timing and Stages of Crater Degradation

[68] Detailed analyses of crater modification in the Martian highlands have been presented for the cratered (Npl1) and dissected (Npld) highland geologic units defined at the 1:15M-scale [Scott and Tanaka, 1986; Greeley and Guest, 1987] and located between ±30° latitude [Craddock and Maxwell, 1990, 1993; Craddock et al., 1997]. Fresh craters in these materials were identified by their well-defined raised rims, hummocky floors with central peaks or pits, and obvious ejecta deposits (Figure 7a). Because fresh impact crater populations can be useful in estimating the time at which crater modification ceased and the highland

Figure 7. Stages in crater modification can be seen in these images. (A) Fresh craters have sharp central peaks with hummocky ejecta and rim deposits. (MOC image M0900776.) (B) During the initial stages of modification, the central peak is quickly eroded and buried. Often the ejecta becomes incised by valley networks and the interior walls of the crater develop gullies. (MOC image M0901754.) (C) Although the central peaks and ejecta have been removed, some craters appear to have been “smoothed” or “softened.” Viking-based analyses suggest that such craters represent a style of modification that occurred later in Martian history. (MOC image M1700974.) (D) During advanced stages of degradation, the rim of the crater is nearly completely removed and the craters become progressively infilled. This leads to the classic flat-floored, rimless morphology (arrows). (MOC image M0807597.) (E) Frequently flat-floored, rimless craters become breached by valley networks. (F) Eventually the craters become heavily eroded and buried, producing a “ghost” crater (arrows). (MOC image M0700832.) Analyses of such modified impact craters indicate that rainfall occurred early in Martian history.
surface became stable, these data were binned by geologic unit, latitude, and the elevation data available at the time and, by convention, normalized to 1 million km² so that N[X] = number of craters > [X] km diameter per 10⁶ km². Although there was no relation between cessation and latitude, fresh crater populations suggest that highland crater modification operated from the end of heavy bombardment (N(5) = 359 ± 26) through the early Amazonian (N(5) = 35 ± 8), which correlates well with the timing of valley network formation [Dohm and Scott, 1993].

Many modified highland craters lack obvious ejecta deposits, suggesting that such materials were easily eroded. The remaining crater ejecta appears etched or hummocky and is frequently dissected by valley networks (Figure 7b) [Masursky et al., 1977; Gurnis, 1981; Craddock and Maxwell, 1993]. Valley networks occur frequently on the continuous ejecta of larger impact craters (>40 km) and often form extensive radial drainage patterns. However, not all modified craters have incised ejecta (Figure 7c). As explained below, such craters may represent a temporal difference in the degradational process. The smooth floors also associated with most craters in the early stage of degradation indicate that central peaks were eroded and/or buried quickly. Crater interiors are often dissected by numerous parallel valleys that terminate at the crater floor. These features closely resemble the debris chutes associated with

Figure 8. Massive erosion and infilling are evidenced by this group of craters on the border of the Sinus Sabaeus quadrangle centered at −21°, 317°. (A) The two largest craters are both ~55 km in diameter and have undergone appreciable modification. (B) However, MOLA data show that the western crater has been infilled by perhaps as much as ~1 km of material. Either it has experience extensive backwasting or sediments were transported from the surrounding terrain. It is unlikely volcanic infilling has take place since such deposits should fill surrounding craters to the same level [Howard, 1999; Howard and Craddock, 2000].
terrestrial impact structures in the early stages of dissection (e.g., Meteor Crater [Grant and Schultz, 1991c]). Crater rims also became lower and generally more rounded.

[76] As degradation progressed the crater rims were removed by both continued downward erosion and interior backwasting. This results in the characteristic flat-floored, rimless crater morphology (Figure 7d). By balancing the mass loss from backwasting and infilling at a variety of crater diameters and stages of degradation, crater diameters were estimated to have been enlarged by ~10% initially and by as much as ~30% during the terminal stage of degradation [Craddock et al., 1997]. Similar amounts of enlargement have been estimated from analyses of terrestrial craters with similar degraded morphology [Grant and Schultz, 1993]. Only craters <~60 km in diameter seem to have been modified extensively by such extreme erosion and backwasting [Arvidson, 1974; Craddock and Maxwell, 1993].

Infilling dominated in the late stages, perhaps depositing as much as ~60 m of sedimentary material in the interior of rimless craters in the Sinus Sabaeus and Margaritifer Sinus area [Craddock et al., 1997]. Extreme deposition in other areas produced very shallow "ghost" craters which appear to be almost completely buried (Figure 7e). Crater walls that remain are steep-sided and contain interior gullies. In particular, smaller diameter craters (<20 km) appear to have been completely modified or eradicated [McGill and Wise, 1972; Craddock and Maxwell, 1990, 1993].

5.1.3. Crater Infilling

[77] The recent availability of detailed topographic information from the Mars Orbiter Laser Altimeter (MOLA) permits precise measurements of crater morphology and a reexamination of modification processes. Determining the types of processes responsible for crater infilling is of particular interest because of the geomorphic evidence suggesting that many contained paleolakes [Cabrol and Grin, 1999] and morphometric analyses suggesting they acted as sediment sinks [Craddock et al., 1997]. Typically, the processes responsible for crater degradation caused a significant decrease in overall crater depth simply by eroding the crater rim, predominately through backwasting [Craddock et al., 1997]. Once the crater rim has been completely eroded, there is no longer a topographic obstruction preventing material eroded from the surrounding landscape from being transported into the crater. Thus, contrary to suggestions by Malin and Edgett [2001], we have remained conscious of the possibility that material within modified impact craters may have resulted from erosion of neighboring terrain [Craddock and Maxwell, 1993; Craddock et al., 1997]. As an example, Figures 8a and 8b compare a modified crater with one at a similar diameter but at the terminal stage of degradation. MOLA data reveal that the degree of infilling during the terminal stages of degradation is often dramatic.

[78] Using MOLA data, we quantified the amount of infilling by measuring the depths of both fresh and degraded craters >10 km in diameter in the parts of the Iapygia and Margaritifer Sinus regions of Mars (Figure 9a). Only MOLA tracks crossing close to the center of craters were utilized. Fresh craters were defined as those exhibiting primary texture on the ejecta blankets and little or no presence of smooth interior crater deposits. Degraded craters were conversely defined to have no obvious primary ejecta morphology and to have smooth interior deposits. To better evaluate the processes infilling degraded craters, we did not include any craters that were breached by channels. Calculation of crater depth utilized the average of the highest points on the two crossings of the MOLA track. For comparison, Figure 9a also shows summary relationships for fresh crater depth previously derived using photoclinoity [Craddock et al., 1997] and global analysis of MOLA Digital Elevation Maps [Garvin et al., 2000], as well as the relationship for craters in increasingly advanced stages of degradation [Craddock et al., 1997]. The greater average depth of fresh craters in the present study may be influenced by the fact that most MOLA profiles do not precisely pass along the crater center, and thus underestimate the crater diameter. Degraded craters are generally at least 500 m to as much as 3000 m shallower than fresh craters at equivalent diameters (Figure 9a). A small amount of shallowing can be attributed to erosional lowering of the crater rim. In addition, extensive backwasting often makes it difficult to determine the diameter of the crater prior to modification (i.e., its initial, fresh diameter). A smaller diameter crater is inherently shallower prior to backwasting. Nevertheless, much of the change in-depth must be the result of crater infilling [Craddock and Chuang, 1996; Craddock et al., 1997].

[79] Appreciable amounts (hundreds of meters) of post-Noachian eolian infilling are unlikely. MOC images show that eolian deposition and scour in craters occurs in spatial patches considerably smaller than the size of large basin floors [Edgett and Malin, 2000]. Air fall deposition during the Noachian may have included planet-wide deposition of ejecta and volcanic ash, as well as eolian deposition [Grant and Schultz, 1990, 1991a, 1991b]. However, the preservation of channels and persistence of crater rims and unburied intercrater uplands indicates that if extensive air fall deposition occurred, such deposits would have had to be subsequently eroded and transported by other processes to contribute appreciably to crater infilling. Also, Edgett and Malin [2000] noticed that dark regions (e.g., Syrtis Major) do not appear to be capable of "cleaning themselves" of bright dust following a global dust storm as suggested by Christensen [1988], which may indicate that there is no homogenous, planet-wide fall-out of eolian material today.

[80] MOLA data indicate that a number of craters and basins with diameters of ~100 km or greater have flat floors (e.g., Dawes at -9° lat., 322° long, Figure 10). Occasionally these basins also contain wrinkle ridges (e.g., Flaugergues at -17° lat., 341° long.). Lava infilling is a plausible explanation, either from eruptions from within the crater or from flows breaching the crater rim from elsewhere. However, because sources for lava flows are generally isolated, lava infilling is unlikely to account for the nearly symmetrical profiles of the majority of degraded craters <60 km in diameter (Figures 9b and 11b–11e). It is also important to note that wrinkle ridge formation can occur in sedimentary deposits [Watters, 1988], so such features are not diagnostic of lithology. Also, the highly dissected interior crater rims indicate that at least some of the interior crater deposits are of fluvial origin. Lava flows may have contributed to large basin infilling, but the final stages of infilling appear to be fluvial, as discussed below. Because of the discrete number of lava sources, volcanic
infilling would be a plausible candidate if future work indicates that depths of infilling do not vary systematically with crater size.

[75] Besides a common flat floor, most degraded craters have nearly symmetrical, inward-sloping floors with gradients of 0.005 to 0.02 (Figures 9b and 11b–11e). Fluvial erosion and sedimentation appear to be the most likely processes to have produced these fan-like deposits. The channeling of inner crater walls (e.g., Figure 10a) also requires an advective process, such as fluvial incision or rapid, presumably wet, debris avalanches [Howard, 1994a, 1994b]. In addition, the sediment eroded from crater walls had to be carried several tens of kilometers to produce the smoothly graded crater floors, and again fluvial transport is the most likely mechanism. MOLA data allow this fluvial interpretation to be evaluated both by comparing basin slopes and profile shapes with terrestrial basin deposits and through simulation modeling of fluvial crater degradation.

[76] Two properties of the crater floor morphometry can be compared with terrestrial basin deposits (pediments, fans, bajadas, and playas). These are the gradient and the profile concavity. The average basin gradient is measured from the base of the crater wall, which is usually an abrupt break in slope at the inner edge of the lower gradient crater floor, to the center of the crater. On terrestrial fans and basins gradients were measured from the basin fill-mountain contact to the basin center, the inner edge of playas, or through-flowing streams, whichever was closest to the mountain. These data were derived from topographic maps (at scale ranging from 1:24,000 to 1:250,000) and ~90 m DEMs at 1:250,000-scale. Locations were selected where the mountain front was nearly linear and flow into the basins was neither strongly convergent nor divergent. For Mars craters, the basin source area was defined as extending from rim crest to the bottom of the crater wall. For terrestrial basins the source area was defined from the drainage divide to the head of fluvial fans/basin fills.

[77] Average gradients of Martian degraded crater floors are lower than alluvial fans in the Basin and Range and Mojave Desert region (“Earth Fan”, Figure 9b), but are equivalent to low gradient fans and basin fills in humid environments (“Earth Basin”, e.g., the Rocky Mountain-Great Plains transition, fans in the Shenandoah Valley), and alluvial fans in low relief basins in Arizona. Low gradient terrestrial basin fills can result from high water to sediment ratios (e.g., humid climate or low sediment delivery rates) or

Figure 9. (opposite) Geometric properties of Martian craters and terrestrial fan and basin deposits. (A) Plot of crater depth (from rim to crater floor) versus crater diameter for fresh and degraded craters. Measurements made from MOLA data in the Margaritifer Sinus and Iapygia regions of Mars. Summary depth/diameter relationships for relatively unmodified craters are from Craddock et al. [1997] “Craddock Fresh,” based upon photoclinometric techniques and from MOLA DEM data “Garvin Fresh” from Garvin et al. [2000]. Also shown are summary relationships for moderately degraded craters of class B and C “[b/c]” and for highly degraded craters “[e]” from Craddock et al. [1997]. A plot of the temporal change of crater depth and crater diameter is shown for simulated dissection of a 50-km crater (see Figure 11). (B) Plot showing the relationship between the average gradient of the inward-sloping surfaces of degraded crater floors (e.g., Figure 10) and the average gradient of the corresponding interior crater rim, from measurements made from MOLA profiles. Also shown are the gradients of alluvial fans in the Southwestern United States plotted against the average gradient of the source mountains (from mountain crest to fan head), as well as a similar relationship for lower-gradient alluvial basins in the United States. The terrestrial data is from topographic maps and 1:250,000 DEM data. (C) The relationship between the concavity of Martian degraded crater floors and the gradient of the crater floor, as well as corresponding data from terrestrial fans and basins. (D) The total relief of Martian degraded crater floors (from the base of the interior crater wall to the center of the crater) is plotted versus crater diameter.
fine-textured sediment (sand and finer sediment rather than gravel). In addition, the lower Martian gravity can result in a 0 to 40% decrease in gradient for equivalent water and sediment supply, depending upon whether threshold, bed load, or suspended load conditions apply (A. D. Howard, in preparation). Furthermore, water flow into the interior of craters is strongly convergent, which may result in greater discharges and lower gradients as compared to nonconvergent flow. The average gradient of both terrestrial fans/basins and Martian crater floors exhibits a moderate positive correlation with the average gradient of the source region (interior crater wall for craters, the source mountains for terrestrial fans/basins) (Figure 9b).

The long profile concavity of fans and crater floors was measured by fitting an exponential downstream decrease in gradient

\[ S = Ke^{-bx} \]  

where \( S \) is gradient, \( x \) is the distance from the crater rim (or mountain crest), \( K \) is an estimated constant, and \( b \) is the estimated concavity. Both terrestrial basin deposits and Martian degraded crater floors exhibit similar concavity (Figure 9c), again suggesting a similarity in emplacement mechanisms. Crater floor concavity could also be the result of differential compaction or dissolution of interior fill deposits, however. The average relief from crater floor edge to crater center for 28 measured craters (15–90 km in diameter) is 140 m. Based upon estimated crater fill depths (Figure 9a), >10% decrease in volume would be required to result in the observed concavity, which does not seem reasonable. Concentric fractures commonly occur at the margins of basin fills which have undergone appreciable downwarping (e.g., in the lunar mare [Maxwell, 1978]). None are apparent in Martian crater fills. With the notable exception of the occasional basin containing wrinkle ridges, there is little evidence for interior compression that should also occur with appreciable downwarping [e.g., Muehlberger, 1974].

![Figure 10](image1.png)

**Figure 10.** The 180-km-diameter crater Dawes. (A) A portion of MDIM quadrangle M10S322 showing the track of MOLA orbit 10922. (B) The MOLA topographic data showing the flat-floored crater floor with less than 100 m of relief. Vertical exaggeration is about 44X.

![Figure 11](image2.png)

**Figure 11.** The results of a 50-km-diameter crater that has been eroded using the computer model of Howard [1994b]. In these simulations the erosion rate is limited by a finite rate of weathering of the host rock. (A) A simulated 50-km fresh crater. (B) Degraded crater corresponding to modification stage in Figure 7b. (C–D) Advanced stages of degradation. Note the increase in crater diameter due to backwasting. (E) Topographic profiles through the center of the crater showing advanced stages of degradation. Vertical exaggeration is about 18.3X. Compare with Figure 12.
The inward sloping fan-like surfaces forming the floor of degraded craters <60 km in diameter seldom exhibit flat central regions that might correspond to the playa lake beds in interior-drained desert basins of the southwestern United States. This could indicate that the sediment eroded from crater walls had little material in the silt and clay size range that would have been carried in suspension to be deposited in temporary lakes. Another possibility is that precipitation levels were sufficiently high that lakes of appreciable (but variable) depth occupied crater interiors during the epoch of crater degradation. Channels debouching into such lakes would extend a variety of distances into the crater interior depending upon existing lake levels. The effect that lakes varying in depth would have upon the morphology of crater floors is uncertain. Shorelines and deltas might not survive subsequent eolian reworking and small impact gardening if lake levels were highly variable; Pleistocene lake basins in the Basin and Range Province of the United States that lack an outflow sill controlling lake level have numerous poorly developed shorelines over a wide range of lake depths. A few Martian highland craters exhibit symmetrical interior terraces, and a lacustrine origin for such terraces has been suggested [Forsythe and Zimbelman, 1995; Cabrol and Grin, 1999]. Lacustrine processes may be an alternative to lava inundation as an explanation for large flat-floored craters, such as Dawes (Figure 10). Because of the great initial depth of large crater basins, they would have preferred sites for intermittent or perennial lakes.

To evaluate the affects of infilling on crater morphology and morphometry, a spatially explicit computer model has been developed to simulate landform evolution on the ancient cratered terrain of Mars [Howard, 1994a, 1994b; Craddock et al., 1997; A. D. Howard, in preparation]. The model includes the processes of cratering, weathering, mass wasting, fluvial erosion and deposition, and eolian erosion and deposition. For the present discussion, only weathering, mass-wasting, and fluvial erosion and deposition are utilized to simulate modification of a 50 km fresh crater (Figure 11). For this simulation rainfall and surface runoff are assumed to be areally uniform, and the crater is assumed to be composed of bedrock that weathers at a finite rate into regolith that is 10 times more erodible. Fluvial erosion concentrated at the base of the crater wall causes backwasting of the crater walls and deposition of a crater basin fill with concave profile. Additional simulations with other process assumptions are presented by A. D. Howard (in preparation), but the general pattern of strong interior crater wall erosion and radial infilling by deposited sediment is characteristic of all simulations, and the simulated morphology bears a striking similarity to degraded Martian craters.

Figure 9a shows the path of change of crater diameter and depth for the simulation shown in Figure 11. The curve is jagged because measured crater diameters are quantized by the 0.4 km grid spacing used in the simulation. Because of the bowl-like shape of fresh crater interiors, crater depth diminishes more rapidly in early stages of degradation than during later stages (Figure 9a). For a 50 km crater the modest B/C stage of degradation corresponds to about 0.9 km reduction in-depth and about a 10% expansion of diameter. The highly degraded E stage corresponds to about 1.3 km reduction in-depth and a 20% increase in diameter. These figures are close to the estimates reported by Craddock et al. [1997]. The simulations indicate that crater diameter increases nearly linearly with time, assuming temporally constant intensity of fluvial erosion, so that a crater near the terminal stage of degradation (stage E) has undergone about twice the duration of modification as a crater with modest amounts of degradation (stage B/C). The scattering of very shallow degraded craters lying well below the line denoting terminal degradation (stage E) have extreme amounts of sediment infilling, possibly from channels debouching into the crater interior, or extensive, local accumulations of eolian or volcanic materials.

5.2. Evidence for Rainfall

To determine the extent of erosion as well as the types of geologic processes responsible for crater degradation, we made morphometric measurements of fresh and degraded impact craters using both Viking-based photoclinometry [Craddock et al., 1997] and MOLA topography. Beginning with the morphometric data from fresh impact craters, we ran a variety of computer simulations using Howard’s [1994b] model to explore the effects different geologic processes might have on crater morphology from known rate laws. Results of these simulations were compared to measurements of actual degraded craters, thus allowing us to directly evaluate the efficiency of any process or combination of processes at producing the observed crater morphology. The tacit assumption was, of course, that fresh craters preserved on the Martian surface today are morphologically similar to those that formed during the early history of Mars prior to erosion. Our results indicate that a combination of rainfall and surface runoff best explain the observed crater morphology. There is also evidence that the intensity at which these processes operated waned with time.

5.2.1. Computer Simulations

At a minimum, erosional processes are both diffusional (i.e., “dispersive”) and advective (i.e., “concentrative”) [Howard, 1994b]. Diffusional processes become less efficient as the contributing area increases or, as Howard [1994b] noted, they are scale-inefficient. In steady state conditions where all parts of the landscape being eroded at the same rate diffusional processes cause the gradient to increase downslope. Although there are exceptions (e.g., rock avalanches on steep, bedrock slopes), diffusional processes are commonly equated with mass wasting. Mass wasting can be expressed by a linear rate law, where the mass wasting flux increases linearly with increasing slope gradient, and a failure rate law, where the mass wasting flux increases without limit as slope gradient approaches a critical value [Howard, 1994b; Roering et al., 1999]. A terrestrial example where the dominant erosional process is diffusional are debris mantled slopes affected by solifuction, which result in slope profiles that are convex-to-concave [e.g., French, 1976, pp. 150–155].

Craddock et al. [1997] used the photoclinometric algorithm presented by Davis and Soderblom [1984] to analyze the morphometry of 264 fresh impact craters and 186 degraded impact craters ~3 to 76 km in diameter in the Sinus Sabaeus and Margaritifer Sinus regions of Mars. Crater modification was modeled in cross-section and the rate laws expressed in cylindrical coordinates. Simulations
were performed on both representative 20-km and 40-km-diameter fresh craters to investigate scale effects on crater modification. The simplest process modeled was linear diffusional creep (Figures 12a and 12b). As a mass wasting process linear diffusional creep describes the effects of rain splash, solifluction or, in the case of the Moon [Craddock and Howard, 2000], micrometeorite bombardment. Creep is assumed to occur at shallow depths relative to the depth of the crater and is essentially a surficial process. The original crater shape is depicted by black squares in Figure 12 while progressive stages of modification are depicted by other symbols. Linear diffusional creep clearly destroys crater rims and fills in the crater bottoms. As expected, the degree of modification of the 40-km crater (Figure 12b) for equivalent duration of degradation is less because creep is scale-inefficient, meaning that its efficiency at erosion decreases with scale due to longer transport distances.

**Figure 12.** The effects of different geologic processes on crater morphometry are depicted in these computer simulations. In each instance the original crater shape is represented by black squares. 20-km-diameter (A) and 40-km-diameter (B) craters modified exclusively by linear diffusional creep have rounded rims and interior slopes. Although the shapes of such craters are similar to lunar craters modified by micrometeorite bombardment [Craddock and Howard, 2000], modified Martian craters are much larger by a factor of 10 or more, further emphasizing the enormous amounts of erosion that took place on early Mars. By comparison, 20-km-diameter (C) and 40-km-diameter craters modified by fluvial erosion and deposition have steep interior slopes as well as flat floors. The general difference is that diffusional processes tend to soften relief while advective processes, such as fluvial erosion, accent slopes. A majority of modified craters on Mars can be explained by a combination of processes (E and F). However, a separate class of craters appears to have been modified by diffusional processes alone (Figure 8). The effects on preexisting modified craters (G and H) would be minimal while preserving evidence for earlier advective processes.
scale-efficient, more rapid erosion occurs for larger contributing areas at the same gradient. This results in approximately the same degree of erosion for both the 20-km and 40-km-diameter craters. For this reason too, erosion of the crater rim is small while erosion of the outer ejecta blanket and lower crater wall is rapid. The steeper crater walls themselves are predicted to be bedrock channels that are detachment-limited. Aggradation occurs on the shallow gradient crater floors and portions of the ejecta. Crater interior walls become steeper, and a sharp break in slope develops between the crater wall and floor.

The best fit to profiles of actual modified impact craters at a variety of diameters resulted from a combination of linear diffusional creep and advective processes (Figures 12e and 12f). However, because erosion and backwasting increase the apparent diameter of a crater during degradation by as much as 30% [Craddock et al., 1997], it is not always possible to precisely match a simulated degraded crater profile to an actual measured one. A 34-km-diameter degraded crater, for example, may have originally been a 26-km-diameter fresh crater. Because erosion modeling was only originally performed for fresh 20-, 30-, 40-, 50-km-diameter craters, matches were frequently imprecise. As a result, we have expanded our library of simulated degraded crater profiles by modeling erosion of 20- to 60-km-diameter in 5-km intervals, and we have increased the number of degraded profiles as described below. Figure 13 does a better job, perhaps, of illustrating the importance of both diffusional and advective processes in modifying Martian highland impact craters.

The mass balance calculations were made to determine where eroded materials were being transported. These calculations were preformed at different crater diameters as the stages of modification progressed to understand the
influence of both time and scale. Initially, these calculations implied that a greater amount of backwasting had taken place at larger crater diameters [Craddock et al., 1997]. Basically, the simulations had to be run longer at larger diameters to explain the volume of eroded materials calculated to be inside heavily modified impact craters, which is opposite to what should have been observed from fluvial erosion and deposition. Smaller diameter craters should erode faster. However, these calculations assumed that the volume of infilling resulted from backwasting the interior crater wall alone. If \( \sim 60 \text{ m} \) of additional sedimentary materials were added from the surrounding terrain once the crater rim had been breached, then the mass balance calculations make sense. The effect of these additional materials would be more pronounced at smaller crater diameters, thus making it appear as though they had experienced less backwasting. The mass balance calculations were fairly straightforward when less modified impact craters were analyzed, indicating that only breached or rimless craters have been infilled any appreciable amount. This suggests that fluvial transport may have taken place over great distances (tens to hundreds of kilometers).

5.2.2. Groundwater Sapping Considered

[88] Because of the implied role groundwater sapping played in creating the valley networks, and the suggestions from hydrothermal models that this process operated exclusively, it is important to evaluate the effects this process would have on crater degradation [Howard, in press]. Simulations of crater dissection by groundwater sapping indicate that erosion, even with recharge from precipitation, is limited to the outer portions of the ejecta blanket and the lower parts of interior crater walls. The problem of erosion on upland summits is illustrated by considering an isolated hill at about \(-32^\circ\) lat., \(9^\circ\) long. (Figure 14a) partially imaged by the MOC at a resolution of about 15 m per pixel (Figure 14b). Topography from representative MOLA tracks (Figure 14c) indicate that the hill summit is about 300–400 m above surrounding plains. The hill is strongly gullied by channels spaced about 500–1000 m apart. These channels locally extend to within 1 km of the summit laterally and within about 50–100 m vertically. Supply of water to these channels from upward recharge is unlikely, since upward moving water would exit close to the base of the hill. The quantity of water that could be stored as an initial reservoir of water above the heads of the channels is clearly insufficient to account for the depth and extent of incision. Recharge or runoff from precipitation is a more realistic interpretation. The summits and divides of this hill are strongly rounded, indicating strong diffusional mass wasting during and/or after the period of channel incision. The southern flank of the hill is bordered by a probable pediment/fan surface with a gradient of about 0.5°, similar to those of the fills in the interior of degraded craters (as discussed above).

5.2.3. Rain or Snow?

[89] The simplest way of explaining our observations is that the diffusional component of modified crater morphology was generated by rain splash and the advective component resulted from surface runoff. These interpretations, however, are not necessarily unique. Alternatively, it is possible that solifluction generated the diffusional component while any surface runoff was the result of snowmelt. Such an alternative may be appealing to some since rain or even a “warm and wet” climate often invokes images of tropical conditions on the Earth. However, it is important to remember that solifluction and snowmelt both require above-freezing temperatures at some point. Even in terrestrial periglacial environments, rainfall occurs when the conditions are warmer and plays an important role in the overall denudation of the landscape [e.g., French, 1976, pp. 141–143]. Finally, unlike midlatitude regions where there is evidence for “terrain softening” [e.g., Squyres et al., 1992] there is no clear indication that solifluction occurred in the equatorial regions of Mars. We examined the interior walls of modified impact craters using high-resolution MOC images and found no evidence for solifluction lobes. To the contrary, interior walls are often gullied. However, surface runoff from seasonal slope wash occurs even on slopes where solifluction is the dominate process, and it can often result in rilling across solifluction sheets and lobes [French, 1976, pp. 141–145]. This would tend to conceal the evidence of solifluction, but it should not hide it completely. The simplest explanation is that rainfall and corresponding surface runoff eroded impact craters on Mars. Whether
above-freezing temperatures were only seasonal or periodic on early Mars, erosion would only be possible during these warmer intervals. For additional clues as to whether early Mars had a more temperate or periglacial environment, the rate of highland erosion was also estimated.

5.3. Erosion Rate Estimates

Climate can have a strong influence on the rate at which a landscape is eroded [e.g., Fournier, 1960]. Estimating the erosion rates in the Martian highlands offers an additional clue to past climatic conditions. At the very least it may be possible to infer the intensity of past erosional processes. Early efforts to estimate the amount and rates of erosion in the Martian highlands [Craddock and Maxwell, 1990, 1993] were crude. Theoretical crater production curves were subtracted from actual production curves, and the “missing” craters were assumed to have been completely eroded from the population. The rim and ejecta volumes of these “missing” craters were estimated, and these values were subsequently divided by the surface area of the region investigated to determine the average thickness of eroded material. This value was then divided by the relative time erosion occurred, which is provided by the fresh crater populations. Obviously, such estimates are model dependent since multiple theoretical production curves have been proposed for Mars [Hartmann et al., 1981; Neukum and Hiller, 1981; Neukum, 1983]. In addition, the volume of material removed from partially eroded craters was not considered.

To obtain a more accurate measurement, Craddock et al. [1997] divided degraded highland craters into four classes based on morphologic differences. Photoclinometric measurements of both fresh and degraded craters were used to calculate the amount of material eroded and redistributed at each stage of crater degradation as a function of diameter. These equations were then applied to every degraded crater observed in the Sinus Sabaeus and Margaritifer Sinus study area. Based on this, the total volume of material eroded was determined to be $2.2 \times 10^5$ km$^3$, which is equivalent to a layer ~200 m thick. Assuming that liquid water was stable on Mars approximately the same time it was on the Earth (~4 billion years ago [Holland, 1984, p. 110]), and taking into account the differences in model chronologies of the absolute ages of the Martian periods [Tanaka, 1986], degradation was estimated to have been active over a 400- to 600-million-year period. This suggests that the rate of erosion in the highlands was ~0.0003 to 0.0005 mm/yr (0.3–0.5 Bubnoff units), which is comparable to those seen in modern day periglacial environments such as those in Alaska and northern Canada or the Australian Outback [Fournier, 1960]. Considering that the highlands are an unvegetated surface, these rates are very low by terrestrial standards, but they are much higher than those occurring on Mars today [Golombek and Bridges, 2000]. Although the erosion rate estimates suggest an early periglacial or arid environment, there are several caveats to consider. (1) It is impossible to determine precisely when liquid water became stable on the Martian surface. Potentially the erosion rates were higher if, for some reason, water became stable later on Mars than the Earth. (2) The erosion rate estimates are averages based on the entire period highland degradation was occurring. Potentially rates were different in the past and changed with time. (3) The amount of erosion attributable to valley network development is not considered, and MOLA data suggests that this volume of material could be extensive. Evidence is presented below suggesting that erosion of the cratered highlands during the Noachian was substantial, averaging perhaps 10s to 100s of meters of net erosion on uplands and even more substantial depths of accumulation in local basins (such as crater floors) [Malin and Edgett, 2000a]. As such, ~0.0003 to 0.0005 mm/yr should be considered a lower estimate. Currently, erosion rates suggest that the ancient climate of Mars was arid, but more humid conditions cannot be ruled out.

5.4. New Evidence for Extensive Erosion

MOLA topography has revealed that locally valley networks have incised up to 1 km into steep scarps. The areal extent of such steep scarps is limited, but more extensive regions with gentler slopes are also strongly dissected. An example is a ~10,000 km$^2$ area in Tyrrhenum centered at approximately $12^\circ$ lat., $244^\circ$ long, with a gradient of about 0.75° toward the northwest (Figure 15a). This slope has been nearly uniformly dissected by valleys spaced about 1–3 km apart with relative relief of about 30–100 m between valley bottoms and intervening divides (Figures 15a and 15b). Divides between adjacent valleys are rounded to sharply pointed (Figure 15c). The divides show no remnants of an original undissected upland surface. Thus, a minimal estimate of the average depth of incision on this regional slope is about half the local relief, or about 14 to 50 m. In addition, images and MOLA data suggest that incision was deep enough to create a hierarchy of divides with a relative relief of ~100–200 m, suggesting an average erosion depth of 50–100 m. Depths of erosion may actually have been considerably greater because few remnants of Noachian craters smaller than 5–10 km diameter occur on this regional slope (or elsewhere on the Noachian uplands for that matter). This suggests erosion amounts upwards to the equivalent depths of such craters, or 500–1000 m [Craddock et al., 1997].

High amounts of erosion also occurred on the isolated mountainous uplands (massifs) found locally on the cratered highlands (for example, on the Hellas and Isidis Basin rims). An example occurs at ~0°N 262.5°W (Figure 16) in the Libya Montes. MOLA profiles indicate that the Libya Montes massif crests extend about 3000 m above the surrounding plains (Figure 16c). The divides are sharply pointed, rock outcrops are common on the ridge crests (Figure 16b), and sideslopes appear to be talus covered with gradients at the angle of repose (~35°). The mountains are indented by several deep reentrants and numerous small, steep basins. These massifs are probably remnants of ring structures of the Isidis basin [Wilhelms, 1973]. Lunar massifs are similar in height and exhibit local steep slopes; however, they do not show the pronounced radial alignment of reentrants or the hierarchical arrangement of smaller basins tributary to larger. Furthermore, sideslope gradients on the Martian massifs are commonly close to the angle of repose on slopes of all orientations and positions on the massif. The massifs strongly resemble isolated high mountains of the Basin and Range province of the southwestern United States both in overall form (Figure 9b) and in detailed morphology (Figure 9c). An even stronger resem-
balance emerges when the morphology of the lowlands surrounding the massifs is considered. MOLA profiles reveal that the massifs are bordered by fan-like surfaces sloping away from the massifs in all directions (Figures 16a and 16c). Gradients were measured on these apparent fans along four MOLA profiles (heavy lines in Figures 16a and 16c). Gradients ranged from 1.5° to 5.1° and averaged 2.2°. Because MOLA profiles are seldom oriented directly along the line of steepest descent, these gradients are probably slight underestimates of actual values. Such gradients are typical of terrestrial fans formed of debris shed from adjacent mountains (Figure 9b). The original shape and size of the massifs is uncertain. However, the massifs have clearly been steepened and backwasted by the development of steep, hierarchical valleys. Debris from these massifs has apparently been spread onto alluvial fans radiating from the massifs. Preferential incision and slope steepening along hierarchical valley systems requires advective processes in the sense discussed above [Howard, 1994b]. As an advective process, fluvial erosion tend to steepen slopes, whereas most diffusive or mass wasting processes (creep, landslides, etc.) tend to smooth the affected topography. On steep slopes, some rapid mass wasting processes (avalanches) can affect terrain similar to advective processes producing spur and gully topography [Howard, 1998], but dry avalanches generally deposit their debris at the base of slopes at the angle of repose. Wet debris avalanches and mudflows can erode steep slopes and transport debris onto fan surfaces at gradients close to those observed surrounding the Martian massifs [e.g., Whipple and Dunne, 1992]. That is, they can be advective. Whether by wet debris avalanches or less debris-laden fluvial flows, erosion of these massifs and construction of their surrounding fans is best explained by water derived from precipitation and involving erosion and deposition of large quantities of material.

[94] Additional evidence for large amounts of erosion and deposition in the Martian highlands is the very existence of long, integrated valley network systems. The Samara Valles system, for example, exceeds 1400 km in length [Grant, 1987, 2000]. If affected by rainfall and runoff, a surface produced solely by cratering (e.g., the lunar highlands) would contain flow path lengths averaging only a few kilometers unless sufficient runoff occurred to fill craters to overflowing. Computer simulations (A. D. Howard, in preparation) indicate that combined erosion of divides (generally crater rims) and infilling of basins (generally crater floors) gradually results in integrated drainage networks, but the amount of erosion required to produce long channel systems is great. Simulated erosion of an initially saturated cratered surface as a function of time (Figure 17) shows the evolution of an average flow path length (the initial and final simulation states are shown in Figure 18, discussed below). The final average path length for this simulation is close to the maximum that can occur for the doubly periodic boundary conditions, or about 1/4 of the maximum boundary length of 102.4 km. If sufficient water were available to
cause deep crater lakes (as suggested by Cabrol and Grin [1999]), overflow of crater rims could greatly reduce the time required to integrate drainage networks. Infilling of crater and basin floors by fluvial sediments, air fall deposits, or lava flows/intrusions could also reduce the amount of required erosion. Finally, superimposed long regional slopes (for example, the topographic slopes radiating away from Hellas [Smith et al., 1998] and the highlands-lowlands boundary), and structural depressions can make network integration more rapid, but development of the long, integrated networks that are common on the Martian highlands has probably required extensive reworking of the highlands surface by fluvial erosion and deposition.

The regional topographic character of the cratered highlands is another indicator of extensive past erosion. Much of the highlands have a character similar to the region shown in Figure 2a in that there are extensive low relief plains and plateaus at varying elevations separated by isolated ridges and scarps. Of course, many of the basins occupy the floors of large impact basins and many of the scarps are interior crater rims. Many of the intercrater plains, however, are irregular in shape. Some even occur as high plateaus, such as those at the southeast and southwest corners of Figure 2a. Often the exterior rims of major basins are nearly as steep and high as the interior crater rims, including parts of Newcomb, a 200 km basin centered at −24°, 359°. Locally the rims of large craters have flat, nearly undissected surfaces bordered by dissected inner rims on one side and a steep outer scarp on the other, for instance

**Figure 16.** Morphology and topography of a massif complex centered at about 0° lat., 262.7° long. (A) MOC context image M0202858 showing selected MOLA tracks identified by orbit number and the location of the high-resolution image shown in (B). Heavy portions of MOLA tracks were used in measurement of gradients of possible alluvial fans at the base of the massifs. (B) MOC high-resolution image M0202857 centered on the crest of a massif. Image width is about 2.95 km. (C) MOLA profiles crossing the massif complex (orbit numbers increment on the equator for north-to-south MOLA profiles). Profile 11926 is displaced vertically by ~1500 m, profile 11806,7 by >6000 m and profile 11479,80 by >4000 m. Vertical exaggeration is about 11.5X.

---

**Figure 17.** Relationship between the mean path length of flow and simulation time (arbitrary but linear units) for the simulated fluvial dissection of an cratered surface. Mean path length of flow is average distance of flow of runoff from all portions of the simulated landscape to the center of the nearest enclosed depression. The simulation domain is 102.4 × 102.4 km. Results from simulation modeling of A. D. Howard (in preparation).
surfaces produced solely by impact cratering (e.g., the lunar highlands) do not have extensive plains surfaces and exterior crater rims are generally not steep (e.g., the simulated cratered surface in Figure 18a). Extensive fluvial erosion and deposition can modify a cratered surface so that it displays the type of steep scarps, dissected outer crater rims, and plains at various levels that occur widely on the Martian cratered highlands.

Figure 18. Simulated erosion of a cratered landscape. (A) Dense cratering superimposed on a low dimension fractal surface. Contours are 0.1 km on a 102.2 × 102.4 km domain. (B) Simulated erosion of (A) by weathering, mass wasting (linear diffusional creep), and fluvial erosion and deposition.

5.5. Evidence for Climatic Change

Photoclinometric measurements of 117 degraded craters located in highland materials (Npl1 and Npld) between ±30° latitude were made to determine the spatial distribution of crater degradation processes. These data were compared to MOLA profiles as well as previous photoclinometric measurement (Figure 13). The morphometric characteristics of degraded craters throughout the highlands appear similar to those described in detail for the Sinus Sabaeus and Margaritifer Sinus region [Craddock et al., 1997]. Thus, crater modification throughout the Martian highlands can usually be explained by a combination of mass wasting (linear diffusional creep) and fluvial erosion and deposition (Figure 13a). Such craters are typified by a rounded interior rim, steeply dipping walls, and a sharp break in slope at the crater floor. However, approximately 20% of the modified craters we measured have well-rounded rims with shallow sloping walls that gradually transition into the crater floor (Figure 13b). This type of modified crater can be explained by mass wasting (linear diffusional creep) exclusively. Where superposition relations could be established (Figure 19), linear diffusional creep was the sole degradation mechanism on the younger craters. Such craters also indicate that the difference in erosional mechanisms was not the result of lithologic variations in the highlands or regional climatic differences; but rather, they reflect a temporal variation in the early Martian climate. Additional computer simulations also supported this assessment (Figures 12g and 12h). Unless the late-stage episode of diffusional processes was extremely intense, it would not have been capable of erasing the signature of earlier, advective processes. Although crater rims and other high-frequency topographic objects (e.g., central peaks)
solifluction in these craters.

images provided no evidence supporting modification by precipitation. As mentioned previously, MOC ceased, presumably during the waning stages of climatic wasting, and overland flow had stopped. These observations suggest that mass wasting by rain splash continued after overland flow had ceased, presumably during the waning stages of climatic support for precipitation. As mentioned previously, MOC images provided no evidence supporting modification by solifluction in these craters.

Further evidence for a gradual climatic change comes from crater morphology. Initially two early stages of crater degradation were identified based on their morphology [Craddock and Maxwell, 1993]. The first stage was characterized by ejecta that was hummocky and highly incised by valley networks (Figure 7b). Later it appeared that most of the continuous ejecta had been removed, primarily because it had a smoother appearance and valley network drainage density was not as great. Notably, however, such craters typically have numerous parallel gullies along the crater walls (Figure 7c). However, photoclinometric measurements did not support this interpretation. Both craters with and without valley networks appear to have been eroded to approximately the same degree, and this was not explained [Craddock et al., 1997]. It now appears likely that the style of degradation waned through time and that this explains the observed differences in crater morphology. Early in Martian history fluvial erosion was intense enough so that shear stresses on shallow slopes, such as discontinuous ejecta blankets, could initiate incision and valley network formation. Later, however, the intensity of fluvial erosion fell so that only steeper slopes, such as those found on the interior crater walls or on the slopes of the Tharsis volcanoes, were capable of driving shear stresses to critical. Eventually fluvial processes became less important and diffusional processes dominated crater erosion. The edge of the surrounding plains gently rolls over into the crater interior (Figure 13a). Even in such craters there is evidence for some parallel gullies on the crater walls (Figure 7c), indicating that surface runoff (albeit limited) continued late into Martian history.

Additionally, valley network morphology may also record evidence for a gradual climatic change. Whether severe [Goldspiel et al., 1993] or minimal [Williams and Phillips, 2001], something was responsible for modifying the Martian valley networks after their formation. Although it is possible that small amounts of modification could take place under current environmental conditions, it may be that valley network modification also records a late-stage climate that supported mass wasting with limited surface runoff.

Finally, there is additional evidence for a gradual climatic shift when data from higher latitudes are considered. Measurements presented by Williams and Phillips [2000, 2001] indicate that depth to width ratios of valley networks increase toward the equator. Assuming that ratios are indicative of valley network maturity, they conclude that the more mature valley network systems occur near the equatorial band. In their model, precipitation and surface runoff carved the valley networks and recharged the aquifers in the very earliest history of Mars. As this early, warmer climate waned into a cooler one, they suggest that valley network formation was maintained by water-lubricated sapping and later by ice-lubricated sapping. Because of solar insolation, these processes would systematically shut off at higher latitudes first as implied by Carr [1999]. Although Williams and Phillips’ [2000] observations and interpretations are in general agreement with those presented for degraded highland craters, it follows from their model that the most severe crater degradation should also occur at lower latitudes. Currently we do not have the data to evaluate a latitudinal relationship. Also, several investigations have observed that eolian deposits apparently increase in thickness at higher latitudes [Soderblom et al., 1973; Arvidson et al., 1976; Mouginis-Mark, 1979]. These deposits could decrease the depth of valley network systems and potentially produce the observed relations. Such studies are important, however, because they emphasize the syner-
gist relationship between degraded craters and valley network formation.

6. Summary and Conclusions

[101] We have presented a review of the geochemical, climatic, and geologic arguments for and against a warm, wet early Mars. Our principal points are summarized below.

1. Although multispectral data have not identified carbonates on the Martian surface, such observations do not preclude an early warm, wet climate. Because carbonates should be synthesizing on the surface even under today’s cold, dry conditions [Booth and Kieffer, 1978], there must be some plausible explanation why carbonates have not been detected with remote sensing instruments. Sulfuric acid produced from volcanic and magmatic outgassing and UV irradiation may scrub the surface free of carbonates [e.g., Clark, 1999], other soils may be volumetrically more abundant than carbonates [Blaney, 1999], or particle size and surface texture may reduce the spectral band contrast of any carbonate deposit [Kirkland et al., 2000a, 2000b]. Carbonates are present in the SNC meteorites, so there must be a crustal sink for these materials [e.g., McSween, 1994]. Regardless, the presence or absence of carbonates is a poor indicator of past climatic conditions. Likewise, the observation of olivine in TES spectra [Clark and Hoefer, 2000] can be attributed to young, surficial deposits and thus has little bearing on ancient climatic conditions.

2. Climatic models suggesting that Mars has always maintained a cold, dry climate have provided much of the impetus for advocating valley network formation from hydrothermal circulation of groundwater. However, such models are based on a number of controversial assumptions and complex variables, including the “faint, young Sun” and the nature of the early Martian atmosphere. Conflicting results by Kasting [1991], Forget and Pierrehumbert [1997], Mischna et al. [2000], and Colaprete and Toon [2000] emphasize how sensitive climatic models are to input variables. Related to this argument is the fact that early Earth would have been compromised by the same astronomical and climatic constraints, yet somehow warm, wet conditions prevailed on our planet or life would not have developed. It is important to remember that the onus is on the Martian geologic record to provide the best clues for understanding the early climate and the types and intensities of processes it would have supported. As such, geologic interpretations should not be forced to fit a given climatic model; it should be the other way around.

3. Based on drainage density and morphology, most investigators have concluded that Martian valley networks formed as a result of groundwater sapping. However, there are a number of complex variables that influence drainage density, including slope, slope length, contributing area, relief, climate, scale, lithology, and preservation. Previously, these variables were unknown for Mars, so direct comparisons to terrestrial channels have been problematic. Preliminary assessments of valley network morphometry from MGS MOLA data are consistent with valley network formation by rainfall and surface runoff. At the very least, the spatial distribution of valley networks reflects local topography, which is indicative of a rainfall source for runoff or groundwater sapping.

4. The occurrence of V-shaped valley network profiles is indicative of incision by surface runoff, but the rectangular morphology observed in many valley network systems is interpretative. While it may be consistent with groundwater sapping, it is equally consistent with terrestrial runoff channels in equilibrium with discharge, slope, and the channel boundary. Both V-shaped and rectangular shaped (i.e., U-shaped) profiles have been observed in the same systems [Williams and Phillips, 2001]. Although this has been interpreted to represent two stages of valley network development [Baker and Partridge, 1991; Williams and Phillips, 2001], it is also the morphometric relation one would expect to find in many terrestrial rivers if the water suddenly disappeared.

5. On the Earth, groundwater sapping is intimately coupled to rainfall and surface runoff and does not occur as an isolated process. Rainfall provides the means for recharging aquifers and the pulses in discharge necessary for removing coarser materials from channels. These facts are often overlooked when interpreting valley network origin and Martian highland geology.

6. The consistence of groundwater sapping hypotheses with some features and the apparent need for a cold, dry early Mars has made such hypotheses persuasive. However, they are, in fact, plagued with many problems.

(a) To produce the volume of erosion observed in the valley network systems, large quantities of water must be discharged to the surface, which does not seem possible from hydrothermal circulation. Recharge must take place over hundreds of kilometers laterally and several kilometers vertically. Slope, in particular, has not been considered in any of the numerical models.

(b) Global heat sources [e.g., Squyres and Kasting, 1994] would not be able to maintain a steady depth of melting for the length of time necessary to form the highland valley networks (hundreds of millions of years).

(c) The erosion mechanism is poorly defined in existing hydrothermal models. Terrestrial sapping channels are often initiated by surface runoff, and sapping begins once incision intersects the water table. Surface runoff also provides the increase in discharge, which is necessary for removing larger particles from the valley system [Schumm and Phillips, 1986].

(d) Valley systems formed by hydrothermal discharge are extremely rare on Earth, yet the more common hydrothermal features have not been identified on Mars. There should be a number of phreatic explosion craters near the source of the valley networks as well as evidence for seeps, fumaroles, hot springs, boiling springs, geysers, zones of acid alteration, deposits of silica sinter, travertine, and bedded breccias surrounding the phreatic craters.

(e) Hydrothermal models do not explain the presence of ground ice in the Martian regolith. Interstitial ice would not be a direct by-product of accretion. Accreting volatiles would vaporize, condense, and subsequently precipitate into the regolith.

(f) Hydrothermal models do not provide a mechanism for modifying impact craters, which was occurring simultaneously with valley network formation. Because both degraded craters and valley networks are erosional features that formed coevally, models explaining one set of features should also explain the other.
7. Valley networks continued to form in the Martian highlands as they were forming on the flanks of the large volcanoes. Thus valley network formation occurred throughout much of Mars’ history, although at varying and generally decreasing rates. As evidenced by degraded crater morphology, it is likely that the valley networks found on the Tharsis volcanoes [e.g., Gulick and Baker, 1989, 1990] formed later in Martian history when the intensity of fluvial erosion fell so that shear stresses necessary for valley incision occurred only on steeper slopes, such as those found on large volcanoes.

8. The suggestions that hydrothermal systems were supported by impact melt have a number of additional problems.

(a) There is no consistent onset diameter related to valley network incision. Many smaller diameter craters have been heavily incised, which is inconsistent with existing models.

(b) Valley networks frequently incise impact craters where the impact melt has already been eroded.

(c) There is not enough volume at a crater rim crest to allow sufficient amounts of groundwater to be released to induce valley network incision.

(d) If a different climate is not required for valley network formation [Brakenridge et al., 1985], then they should continue to form on all Martian impact craters even in today’s environment. However, there is a large population of fresh impact craters that have not been modified. Frequently, fresh craters are found next to similar diameter craters that have been incised by valley networks. This demonstrates a clear temporal change that is best attributed to climate.

(e) Kieffer and Simonds [1980] predicted that impacts into volatile rich targets such as the Martian highlands would result in impact melt sheets dispersed as fine particles. This is supported by the glass component identified in analyses of multispectral data from TES [Bandfield et al., 2000].

9. Groundwater sapping cannot explain the morphology of modified impact craters in the highlands, further undermining the importance of hydrothermal circulation as a surficial process. While it is likely that groundwater sapping did occur on Mars, simulation modeling indicates that rain splash and surface runoff were the most important processes early in Martian history. Such processes are entirely consistent with valley network age and morphology.

10. Younger, superposed craters appear to have been modified exclusively by mass wasting (i.e., linear diffusional creep). While this could potentially be the result of solifluction in a late-stage periglacial environment, we found no independent photographic evidence from MOC images supporting this explanation. Although it is possible that early Mars experienced frozen precipitation, such as snow or sleet, it would still be necessary to raise surface temperatures above freezing to permit runoff. It is ad hoc to suggest that rainfall could not have occurred during such intervals.

[102] Rainfall and surface runoff are the best explanations for the topographic and morphological characteristics of the Martian cratered highlands. Other processes, such as lava inundation and air fall deposition may have contributed to surface modification, but do not appear to have been the dominant agents of terrain modification. Fluvial deposition is also most consistent with the great lateral extent and morphology of the crater floor fills, supporting the idea that interior crater lakes may have been present (ephemeral or perennially) throughout early Martian history [Cabrol and Grin, 1999]. In addition, a number of lines of evidence suggest that the modification of the cratered highlands by rainfall and surface runoff was considerably greater than previously appreciated. Some regional slopes have been incised by an average of at least 50 m and average amounts of erosion on upland areas may have been 200 m or more, with comparable average depth of deposits on plateaus and in basins. Finally, the intensity of these geologic processes appears to have changed over time. A warm, wet early Mars waned into a drier climate that preceded today’s. Thus, the climatic conditions for supporting rainfall and surface runoff appear to be long-lived and did not fail catastrophically.

[103] The scenario supported by measurements and observations of highland geologic features is fairly simple. Like Venus and the Earth, Mars had a substantial primordial atmosphere. Like the Earth, this atmosphere condensed and precipitated onto the surface and into the regolith. In the process it carved the valley networks, eroded impact craters as they continued to form, and redistributed large volumes of sediments. Eventually, enough water was placed into the Martian crust to create the hydraulic head necessary for supporting outflow channel formation. The volume of water released and sediment eroded may have contributed to the putative ocean in the northern plains, especially if the outflow channels were truly “catastrophic.” Alternatively, released water may have been repeatedly cycled back into the regolith, eventually resulting in multiple flood events that were probably less than bank-full discharge. Regardless, water probably stood in the northern plains at least on occasions in the early history of Mars. Eventually, early warm, wet conditions waned, the importance of advective processes (i.e., surface runoff) declined, and erosion from mass wasting dominated. Unlike the Earth, Mars lost its early atmosphere and most of its surface water before the middle Amazonian for a variety of reasons, including a decrease in volcanic outgassing with time, oxidation, adsorption into the regolith, and enhanced loss to space due to a lower gravity and the lack of a permanent or appreciable magnetic field. Also, unlike the Earth, which lost its primordial crust to plate tectonics and continued weathering, the record of early Mars is preserved in the highlands. Quite possibly the Martian highlands look similar to what the surface of the Earth may have looked like following ocean formation. This enhances the prospects of finding evidence for primitive life on Mars. At the very least, because the geologic processes and environmental conditions were so similar between the two planets, Mars should contain the precursor chemical elements that went into creating life on Earth. Ultimately, such information may actually be more important for understanding how life originates. Future work should focus on deconvolving temporal and spatial variations in the geologic processes that operated in the Martian highlands as well as identifying key landing sites where hypotheses regarding the history of early Mars can be tested.

[104] Acknowledgments. The authors appreciate the valuable conversations shared with Ray Arvidson, Connie Bertka, the late Roger Burns, Wendy Calvin, Lawrence Doyle, Ken Edgett, John Grant, Bob Haberle, Ross Irwin, Ted Maxwell, Jeff Moore, Hap McSween, and Jim Zimbelman. This manuscript benefited greatly from reviews by Bill Dietrich and Jeff Moore. This work is supported by NASA grant NAG5-9817 (Smithsonian), Mars Data Analysis, as well as NASA grants NAG5-3931 (Univ. of
CRADDOCK AND HOWARD: RAINFALL ON MARS


Pieri, D., Martian valleys: Morphology, distribution, age, and origin, Science, 190, 1093, 1975.


