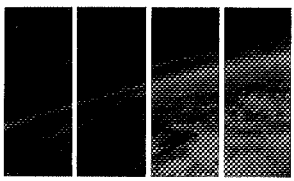
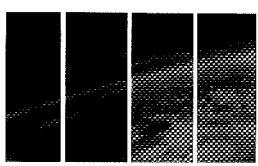


MSATT

Mars Surface and Atmosphere Through Time

Reprinted from the *Journal of Geophysical Research*

The MSATT logo incorporates Viking images of the southern hemisphere limb of Mars, depicting the surface and atmosphere in early and late winter (NASA Photos 53A65 and 79B06). Design by Ronna Hurd, Lunar and Planetary Institute.



MSATT

Mars Surface and Atmosphere Through Time

Reprinted from the *Journal of Geophysical Research*

Volume 100, Number E3, March 25, 1995

Published by THE AMERICAN GEOPHYSICAL UNION

Martian Surfaces and Atmosphere Through Time

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5233–5234, MARCH 25, 1995

Introduction to the MSATT special section

Robert M. Haberle and James F. Bell III

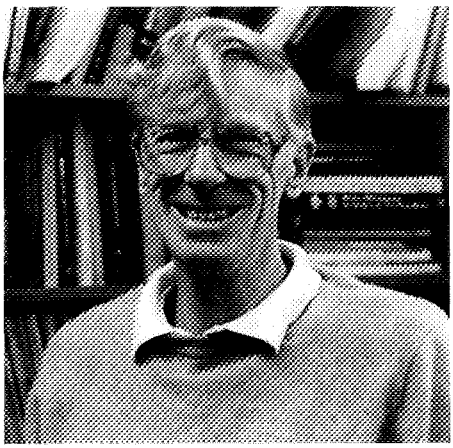
Space Sciences Division, NASA Ames Research Center, Moffett Field, California

The papers published here are based on a workshop entitled "Mars: Past, Present, and Future: Results from the MSATT Program." MSATT (Mars Surface and Atmosphere Through Time) was the last of the Mars data analysis programs and functioned mainly through a series of focused workshops, the final one being held at the Lunar and Planetary Institute in Houston, Texas on November 15-17, 1993. The program began and ended with workshops that brought the entire MSATT community together. At the kickoff workshop (held in Boulder, Colorado, September 23-25, 1991), research areas were identified and collaborations were formed. At the final workshop, progress was assessed and results were presented.

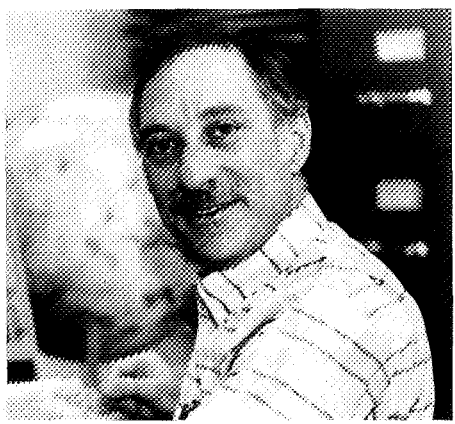
This special MSATT section of *JGR-Planets* is dedicated to the memories of Roger Burns and Jim Pollack. Both men were MSATT investigators, and Roger served on the steering committee. Their dedication to the program, and to planetary science in general, has served the community well, and they will be sorely missed.

Though ill and undergoing chemotherapy, Roger Burns presented a paper at the final workshop and chaired the session on Mineralogy and Future Observations. In his usual gentle, patient style, Roger lead a spirited discussion of ongoing controversies in Mars surface science. His presence was inspirational and illustrative of his devotion to the field. Roger's main planetary science interests were on the mineralogy and weathering history of the Martian surface. Along with Amos Banin, Roger organized a lively and successful workshop on Mars chemical weathering, held at Cocoa Beach, Florida, in September 1992. At that meeting, Roger presented some of his first results on the presence of dehydroxylated clay silicates on Mars, as well as more advanced work on the oxidation of Martian basaltic minerals. Roger was a champion of the search for exotic ferric- and sulfate-bearing phases on Mars because he, more than anyone else, was keenly aware of the potential of such materials for revealing substantial information on the style and rate of Martian weathering and climatic variations.

Jim Pollack was unable to attend the final workshop, as he was then convalescing from his second surgery. But, like Roger, Jim did not let his illness prevent him from participating in the program. He attended three workshops prior to the final one, and continued his research at home. Jim's main MSATT interests were characterizing the optical properties of airborne dust and quantifying the various factors that control the seasonal CO_2 cycle. He realized the importance of suspended dust on the thermal structure and heating of the Martian atmosphere. So, with Maureen Ockert-Bell, he used Viking images to better define the size distribution and single scattering properties of this ubiquitous aerosol. With Francois Forget, he carried out a reanalysis of the Mariner 9 IRIS and Viking IRTM data sets with a particular focus on determining the presence and radiative effects of CO_2 ice clouds in the Martian polar regions. The results of the Ockert-Bell collaboration are reported in this issue; the CO_2 ice cloud studies continue and will be reported on later.



Roger Burns.



Jim Pollack.

Paper number 95JE00223.

Of course, both Roger and Jim were interested in all MSATT research, and the papers published here, as well as those from the kickoff workshop (see the special sections of *JGR-Planets*, February and June 1993), illustrate how broad the program really was. Here you will find papers that address the geology, mineralogy, and meteorology of Mars in an effort to assess how the surface and atmosphere of this fascinating planet have evolved over time. Could early Mars have been warmed by a brighter young sun instead of a massive greenhouse effect? Were glaciers and hydrological cycles part of Mars' relatively recent past, or was aeolian activity responsible for the putative glacial features? Do the SNCs come from a single source region, or is more than one site involved? And what really are the properties of Martian soils and what do they tell us about the weathering environment? Clearly, these are difficult questions, but progress toward answers can be found in this issue. Also contained in this issue are a mix of theoretical and observational papers that deal with the general circulation of the current atmosphere, the factors that drive it (dust properties), and the role it plays in controlling the current climate system.

Looking back on the MSATT program, which sponsored nine workshops in two years, it is safe to say that it was an incredibly productive program. Those of us who participated will miss the frequent focused workshops that stimulated our thinking and guided our research. We can only hope that soon we will obtain new data pertaining to the Martian surface and atmosphere system, and that soon we will have another data analysis program as successful as MSATT.

J.F. Bell III and R.M. Haberle, Space Sciences Division, NASA Ames Research Center, Moffett Field, CA 94035.

(Received January 13, 1995; accepted January 13, 1995.)

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5235–5250, MARCH 25, 1995

Viking Lander image analysis of Martian atmospheric dust

James B. Pollack[1]
Space Sciences Division, NASA Ames Research Center, Moffett Field, California

Maureen E. Ockert-Bell[2]
Meteorology Department, San Jose State University, San Jose, California

Michael K. Shepard[3]
Department of Earth and Planetary Sciences, Washington University, St. Louis, Missouri

Abstract. We have reanalyzed three sets of Viking Lander 1 and 2 (VL1 and VL2) images of the Martian atmosphere to better evaluate the radiative properties of the atmospheric dust particles. The properties of interest are the first two moments of the size distribution, the single-scattering albedo, the dust single-scattering phase function, and the imaginary index of refraction. These properties provide a good definition of the influence that the atmospheric dust has on heating of the atmosphere. Our analysis represents a significant improvement over past analyses (Pollack et al. 1977, 1979) by deriving more accurate brightnesses closer to the sun, by carrying out more precise analyses of the data to acquire the quantities of interest, and by using a better representation of scattering by nonspherical particles. The improvements allow us to better define the diffraction peak and hence the size distribution of the particles. For a lognormal particle size distribution, the first two moments of the size distribution, weighted by the geometric cross section, are found. The geometric cross-section weighted mean radius (r_{eff}) is found to be 1.85 ± 0.3 μm at VL2 during northern summer when dust loading was low and 1.52 ± 0.3 μm at VL1 during the first dust storm. In both cases the best cross-section weighted mean variance (v_{eff}) of the size distribution is equal to 0.5 ± 0.2 μm. The changes in size distribution, and thus radiative properties, do not represent a substantial change in solar energy deposition in the atmosphere over the Pollack et al. (1977, 1979) estimates.

1. Introduction

Solar heating of the Martian atmosphere and surface plays a fundamental role in determining the temperature structure and dynamics. As with Earth, the heating of the atmosphere provides the energy to drive the winds and general circulation pattern. Unlike Earth, a majority of the absorbed solar and infrared energy comes from absorption by suspended dust particles [*Gierasch and Goody*, 1972; *Pollack et al.*, 1979; *Haberle et al.*, 1982]. Understanding the interactions between the atmospheric dust and solar energy enhances our ability to understand and predict the seasonal and longer-term climate cycles.

Over the past 20 years, scientists have analyzed data obtained by various spacecraft to define the key radiative properties of airborne dust, summarized in Table 1 of this paper and by *Kahn et al.* [1992] and *Murphy et al.* [1993]. Mariner 9 and the Viking orbiters and Landers have provided a wealth of information. The investigation of atmospheric dust particles has been centered around several key descriptive parameters: the particle size distribution (characterized by the geometric cross-section weighted mean radius and variance (r_{eff} and v_{eff})), the shape of the particles, the single-scattering albedo ω_0, and the imaginary index of refraction n_i. Given this information, one can calculate the single-scattering phase function $p(\theta)$, and the extinction coefficient Q_{ext}. Each of these parameters is important in determining the extent of energy absorption and scattering by the dust particles.

The single-scattering properties of the airborne dust particles can be derived, in part, from their size distribution, such as the modified-gamma or lognormal size distributions. *Toon et al.* [1977] fit brightness temperature models to the Mariner 9 infrared interferometer spectrometer (IRIS) observations in the thermal infrared to obtain the dust composition and size. A good fit to the spectra was found in montmorillonite. Using the optical properties of montmorillonite and the modified-gamma distribution, they extracted the mean infrared properties of the dust. Their best fit size distribution for α=2 and γ=0.5 was r_g = 0.4 μm. Using these values, it is possible to calculate the first two moments of the size distribution: r_{eff} = 2.70 μm and v_{eff} = 0.38 μm.

Pollack et al. [1979] also assumed a modified-gamma size distribution to analyze Viking Lander sky images. They found the Toon et al. size parameters to be consistent with Viking Lander image results. *Pollack* [1982] presented a number of single-scattering parameters that were determined from these

[1]Deceased, June 13, 1994.
[2]Also at Space Sciences Division, NASA Ames Research Center, Moffett Field, California.
[3]Now at National Air and Space Museum, Smithsonian Insititution, Washington, D. C.

Paper number 94JE02640.
0148-0227/95/94JE-02640$05.00

Table 1. Comparison of the Martian Dust Size Parameters Inferred from Spacecraft Observations

r_{eff}, μm	v_{eff}, μm	L_s and Year	λ, μm	Other Parameters	Source
			Ultraviolet		
1.3	0.2	ds, 1971	0.268, 0.305	spherical particles n_r=1.8, n_i=0.02 - 0.01	Mariner 9 [*Pang et al.*, 1976]
0.2	0.9	ds, 1971	0.268, 0.305	n_r=1.6, n_i=0.004-0.015	Mariner 9 [*Chylek and Grams*, 1978]
			Visible		
2.70	0.38	143.5,1977	0.5-0.9	ω_0= 0.86, g=0.79	Viking lander [*Pollack et al.*, 1979]
≤0.4		169, 1977 - 185, 1979	0.3-3.0	ω_0= 0.92, g=0.55 spherical particles	Viking orbiter IRTM EPF [*Clancy and Lee*, 1991]
1.85	0.5	143.5, 1977	0.5-0.9	ω_0= 0.9, g=0.6, n_r=1.5, n_i=0.0025	this work
1.52	0.5	208.1, 1977	0.5-0.9	ω_0= 0.9, g=0.6, n_r=1.5, n_i=0.004	this work
			Infrared		
2.75	0.42	143.5	5-50	spherical particles	Mariner 9 [*Toon et al.*, 1977]
0.53	0.023	343-348, 1972	5-50	spherical particles	Mariner 9 IRIS [*Santee and Crisp*, 1993]
1.26	0.25	2-18, 1989	0.76-3.14		Phobos ISM [*Drossart et al.*, 1991]
1.26	0.2	2-18, 1989	1.9, 3.7	12 - 35 km altitude	Phobos [*Korablev et al.*, 1993]
1.8	0.4	2-18, 1989	1.9, 3.7	12 km altitude	Phobos [*Korablev et al.*, 1993, Figure 7]
1.8	0.8	295-350, 1971	5-50	spherical particles	Mariner 9 IRIS [*Clancy et al.*, this issue]

Here, ds refers to the dust storm of November 1971.
Unless otherwise noted, all work assumes nonspherical particles.

observations. The single-scattering albedo (ω_0 = 0.67 - 0.99) and asymmetry parameter (g = 0.74-0.90) presented in Table 1 of *Pollack* [1982] are significantly different from the many values at the visible wavelengths contained in the review by *Murphy et al.* [1993].

Clancy and Lee [1991] used Viking orbiter infrared thermal mapper (IRTM) emission-phase-function (EPF) measurements of several regions to better constrain the particle single-scattering behavior. To the first order, their phase function results correspond well with those of *Pollack et al.* [1977, 1979]. However, the asymmetry factor was lower than that of *Pollack et al.* [1979] (g = 0.55 versus g=0.79), the single-scattering albedo was higher (ω_0 = 0.92 versus ω_0 = 0.86). The cross-section weighted mean particle size differed by nearly a factor of 10 ($r_{eff} \leq 0.4$ μm versus r_{eff} = 2.5 μm).

The Phobos mission provided another chance to collect data that could supply estimates of the first two moments of the dust particle size distribution. *Chassefiere et al.* [1992] and *Korablev et al.* [1993] analyzed solar occultations measured with the IR spectrometer to obtain the volume extinction coefficient at 3.7 and 1.9 μm in the altitude range of 12-35 km. The modified-gamma size distribution (with α=2 and γ varied) and the optical indices of a basalt-type mineral were used to extract the dust particle size distribution information. Nine observation sequences were analyzed. Korablev et al. concluded that r_{eff} = 1.26 ± 0.2 μm for v_{eff} = 0.2 ± 0.1 μm was the best fit to the data averaged over all altitudes. The discrepancies between these results and earlier results were attributed to the atmospheric conditions during the observations (clear (Phobos) versus dusty (Viking)) and to the difference in altitude (limited altitude region (Phobos) versus column integrated (Viking)). Yet in trying to rectify the size values to the Toon et al. sizes, Korablev et al. found that the number density profiles were nonphysically constant and would require an enormous flux to maintain.

Clancy et al. [this issue] combined analysis of the Mariner 9 IRIS data and Viking IRTM EPF, and assumed a "palagonitelike" composition for their analysis. They found two good fits to the data: (1) r_{eff} = 1.2 μm for v_{eff} = 0.4 μm and (2) r_{eff} = 1.8 μm for v_{eff} = 0.8 μm. Clancy et al. favor the second size distribution because this distribution fits the 20 to 30 μm dust opacity as well as the visible to 9-μm dust opacity ratio found by *Martin* [1986].

We have carried out a new analysis of Viking Lander 1 and 2 (VL1 and VL2) images of the Martian atmosphere to evaluate the radiative properties of the atmospheric dust particles. The properties of interest were the first two moments of the particle size distribution, the imaginary index of refraction, and other single-scattering properties, such as the single-scattering albedo and phase function. The method of analysis was improved by taking advantage of vignetting (i.e., partial shadowing of the camera) to allow us to obtain data closer to the sun and thus refine previous estimates of the particle size parameters, especially r_{eff}. We also included a better representation of light scattered by randomly oriented nonspherical particles. The refined estimate of particle size and shape has improved the estimates of the other single-scattering properties.

In section 2 we present the data and the methods used in this analysis. Section 3 presents the results, including some

Table 2. Viking Lander Image Information

Image ID	Sol	L_s	τ_{dust}†	Solar Elevation	Azimuth Angles
12A002/S	0	97.0	0.50	37.0	0-220
21B220/BGR	97	143.5	0.36	27.7	10-85
21B221/IR1	97	143.5	0.36	26.4	10-85
12C118/BGR	211	208.1	1.93	35.5	5-60
12C119/IR1	211	208.1	1.93	34.4	5-60

Image ID is given, where the first digit represents lander number, the second is camera number, and the subsequent characters are the picture and filter identification. The effective wavelength of the filters are as follows: blue=0.49 μm, green=0.55 μm, red=0.67 μm, IR1=0.86 μm, and survey=0.69 μm.

Sol is given as the number of days since VL1 touchdown. All images were taken in the afternoon.

†The optical depth of the dust as given by a sundiode measurement taken close in time.

sensitivity testing of the various parameters. In section 4 we compare the results of this study to those presented above and discuss the implications for the heating of the Martian atmosphere.

2. Data and Methods

The data that were used for this study include several images taken by VL1 and VL2 (Table 2). The images were converted to brightness values using the ground-based calibration files of *Huck et al.* [1975]. Brightness was expressed as the ratio of the observed brightness to that of a perfectly reflecting Lambert surface, normally illuminated by sunlight at Mars distance from the sun (*I/F* or the radiance factor, as described by *Hapke* [1981]). The images were corrected for vignetting (described below) and were translated from the camera coordinate system (CCS) to the local horizon coordinate system (LHCS) [*Tucker*, 1978]. The data were scanned at a constant observer elevation, so that azimuthal distance from the sun would closely match the scattering angle of the observation.

Above 20° in elevation CCS, vignetting due to the camera contamination cover was important. When the contamination cover was closed, it partially shadowed the camera lens, with the shadow fraction monotonically increasing with increasing elevation angle. Thus vignetting partially offset the tendency of the sky to brighten when observations were made close to the Sun. This helped to reduce the saturation at angles close to the Sun. Here, we carefully selected scans that were unsaturated at the smallest angular distance from the Sun to get as far into the particles' diffraction pattern as possible. Note that brightness values at an azimuthal distance of less than 10° from the sun are suspect [*Pollack et al.*, 1977] and thus were not used. We used calibration files for the fractional reduction in brightness due to vignetting to make a good correction for it in scans of interest [*Wolf*, 1976].

Figure 1 illustrates the importance of vignetting to the work that we are presenting. This comparison shows the difference between a real scan taken from an image where there was vignetting (solid line) and what the scan would have looked like had the image not had vignetting (dotted line). The scans are in pseudo-data number (DN) with saturation equal to a value of 63. Had the contamination cover not shadowed the camera, saturation would have occurred at 15° scattering angle. With vignetting, we are able to extract meaningful data down to 10° scattering angle. Accounting for vignetting also allows for a greater dynamic range in the data and thus portrays a better representation of the particles' diffraction pattern.

Equally important to our data selection was the choice of scanning elevations and azimuths. Although the phase function relies on all of the particle-scattering properties, certain areas can be used to optimize some of the parameters. Figure 2 illustrates this idea.

Figure 2, zone 1. The shape of the curve at small scattering angles (θ) was used to find the particle size distribution information (r_{eff} and v_{eff}). An elevation as close as possible to that of the Sun was used to minimize the angular distance of the last unsaturated data point.

At small scattering angles the shape of the observed intensity curve is dominated by single scattering. This area of the single-scattering phase function (zone 1) is highly diagnostic of particle size, but it is not sensitive to particle shape [e.g., *Hansen and Travis*, 1974; *Pollack et al.*, 1977; *Pollack and Cuzzi*, 1980]. To illustrate the sensitivity of the shape of the intensity curve versus azimuth in zone 1, consider a limiting case (Figure 2b, r_{eff} = 0.1 μm). The single-scattering phase curve of very small particles closely resembles that of Rayleigh scatterers [*Hansen and Travis*, 1974]. For the geometry of our images, this means that we expect at most a 5% increase in *I/F* between 20° and 10° azimuth angle. The observed *I/F* (cf. Figures 3-7) actually increases by a factor of 3-5 over this same range of angles. Thus the mean particle size of Martian dust must have a size parameter, $x = 2\pi r_{eff} / \lambda$ >>0.1 or r_{eff} >> 0.01 μm. For large particles (e.g., Figure 2b, r_{eff} = 2.5 μm), the slope of the phase function is much steeper in zone 1 with as much as a factor of 10 or more increase. For these particles, $x > 1$, the phase function is controlled by light diffracted around the particle with the amplitude of the phase function increasing and its width decreasing as x increases [*Pollack and Cuzzi*, 1980]. Both criteria imply that $x << 100$ or $r_{eff} << 10$ μm.

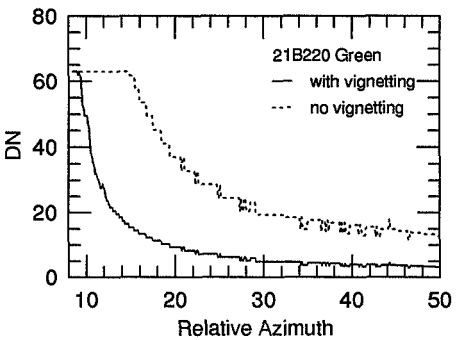


Figure 1. Scans of image 21B220 (red filter) taken at the solar elevation angle. The data are converted to pseudo-DN to better illustrate the effect of vignetting on image scans. The solid line is from the calibrated image. The dotted line is a representation of what the same image scan would have looked like had the vignetting not occurred.

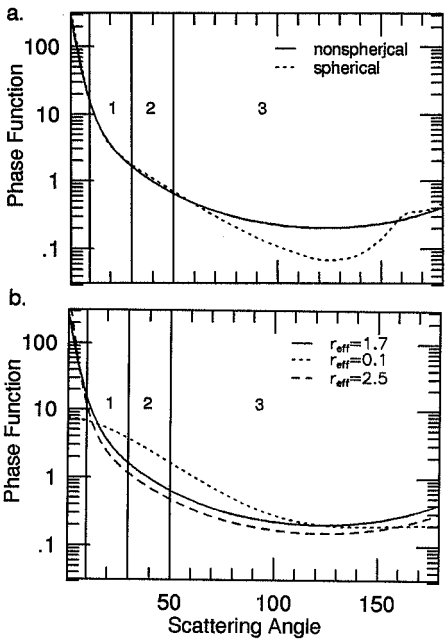


Figure 2. Comparison of five single-scattering phase functions, demonstrating the effect of particle (a) shape and (b) size on the scattering behavior. The numbers denote the specific areas that were chosen for optimizing the phase function parameters: zone 1, particle size distribution; zone 2, imaginary index of refraction and single scattering albedo; and zone 3, shape.

Figure 2, zone 2. The absolute value of the brightness at scattering angles of 30°-50° was used to determine the value of the single-scattering albedo, which in conjunction with the size parameters found from zone 1, yielded estimates of the particles' imaginary index of refraction (for an assumed value

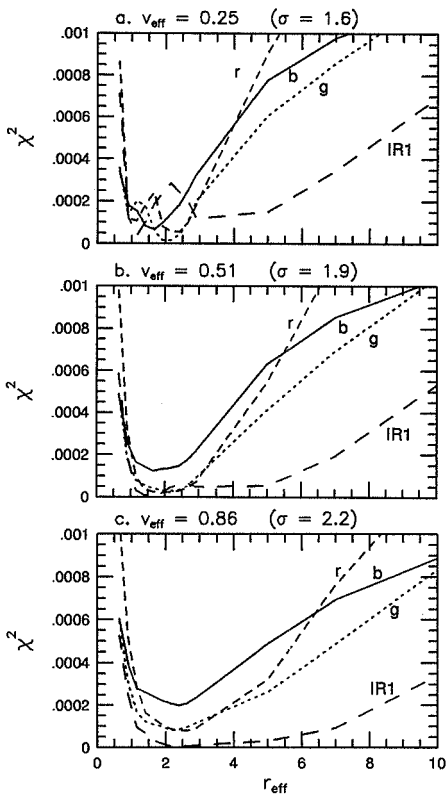


Figure 4. Plot of χ^2 for several values of v_{eff} and r_{eff} for image 21B220 (sol 97). Each panel shows the results of calculations for different values of v_{eff}. The best fit filter values for r_{eff} were averaged to find the best value for this sol.

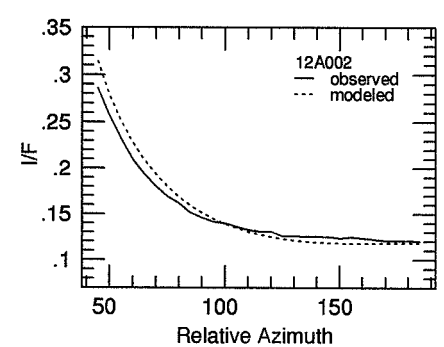


Figure 3. Scan of image 12A002 (sol 0) taken at the 15° elevation angle (solid line) plotted with the best fit model (dotted line), where SLP = 0.013, ThMin = 118.5.

of their real index). The scattering angles of zone 2 were chosen because there is significant multiple scattering at these angles [e.g. *Hansen and Travis*, 1974], which enhances particle absorption, and because the deduced single-scattering albedo depends least on knowledge of the particles' single-scattering phase function, a corollary to Russell's rule [see *Veverka*, 1971]. An elevation below the area of vignetting was chosen.

Figure 2, zone 3. The data at intermediate and larger scattering angles determined several of the particle shape-dependent parameters. This point is illustrated by the divergent behavior of the phase function curves of Figure 2a for scattering angles greater than 50° (also see *Pollack and Cuzzi* [1980]). Fits in this zone yield the shape parameters related to the slope in this zone (FTB or SLP). They are fully defined later in this section in the discussion of the procedure used to evaluate single-scattering properties. An elevation below the area of vignetting was chosen for these scans as well.

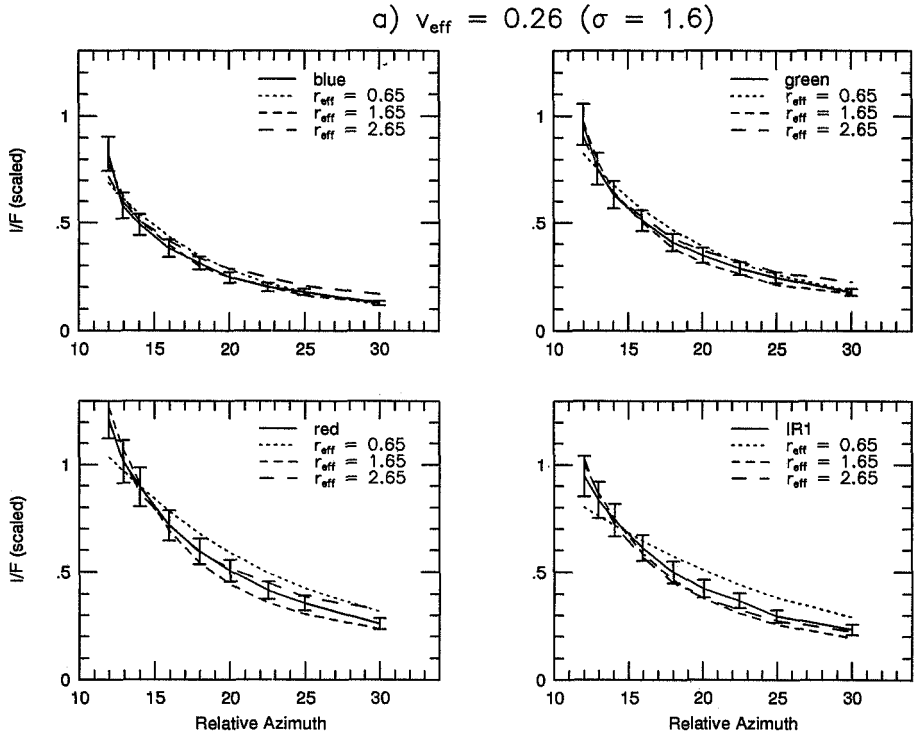


Figure 5. Scan of image 21B220, 21B221 (sol 97) taken at the solar elevation angle (solid line) plotted with several models. Each panel shows the results of calculations for a specific value of v_{eff}. Within each panel, all four filters are plotted with three models of r_{eff}. The median value for r_{eff} in each plot is close to the best fit value.

A number of constraints were placed on the selection of the Viking Lander images. (1) They needed to be afternoon images to minimize the interfering effect of diurnal water ice clouds [*Colburn et al.*, 1989]. (2) They needed to be close in time to a sundiode image to give a value for the atmospheric dust optical depth (see below). (3) The solar elevation angle and azimuth angles close to the Sun had to be in the image to provide the zone 1 information, or azimuth angles opposite the Sun had to be in the image to provide zone 3 information.

Multispectral imaging on the Viking landers was accomplished by alternately selecting three diodes (blue, green, red, or IR1, IR2, IR3) for three vertical scans at each azimuth position [*Tucker*, 1978]. Images 21B220 (color triplet) and 21B221 (IR1 from the IR triplet) supplied information on zone 1 unknowns (r_{eff} and v_{eff}) and zone 2 unknowns (ω_0 and hence n_i). These images provided multifilter (wavelength) coverage at a time well before major dust storms. Additionally, these were the same images analyzed in our earlier studies [*Pollack et al.*, 1977, 1979] and thus provided ready comparisons with our earlier work. Images 12C118 and 12C119 provided comparable information to 21B220 and 21B221, but near the peak of the first of two global dust storms that occurred during the time that Viking was operational. We selected image 12A002 for the analysis of zone 3 unknowns because it covered a wide range of azimuth angles at intermediate and large scattering angles.

The process of simulating the *I/F* observed by the Viking Landers employed a number of theoretical models. We had to determine the scattering properties of the airborne species, the scattering behavior of the surface, and combine the results in a realistic manner. Below we describe the models used to determine the light-scattering properties of the airborne dust, the surface scattering, the radiative transfer model used to combine these scattering models, and the method used to determine a best fit to the free parameters.

The shape of a particle is particularly important when considering the light-scattering behavior of airborne dust particles at intermediate and large scattering angles. We used the semi-empirical theory for scattering and absorption by randomly oriented nonspherical particles of *Pollack and Cuzzi* [1980] to define the single-scattering properties of Martian airborne dust particles. Here, we briefly summarize key aspects of this theory. Mie theory was used for particles where the ratio of the circumference $2\pi r$ to the wavelength λ is small.

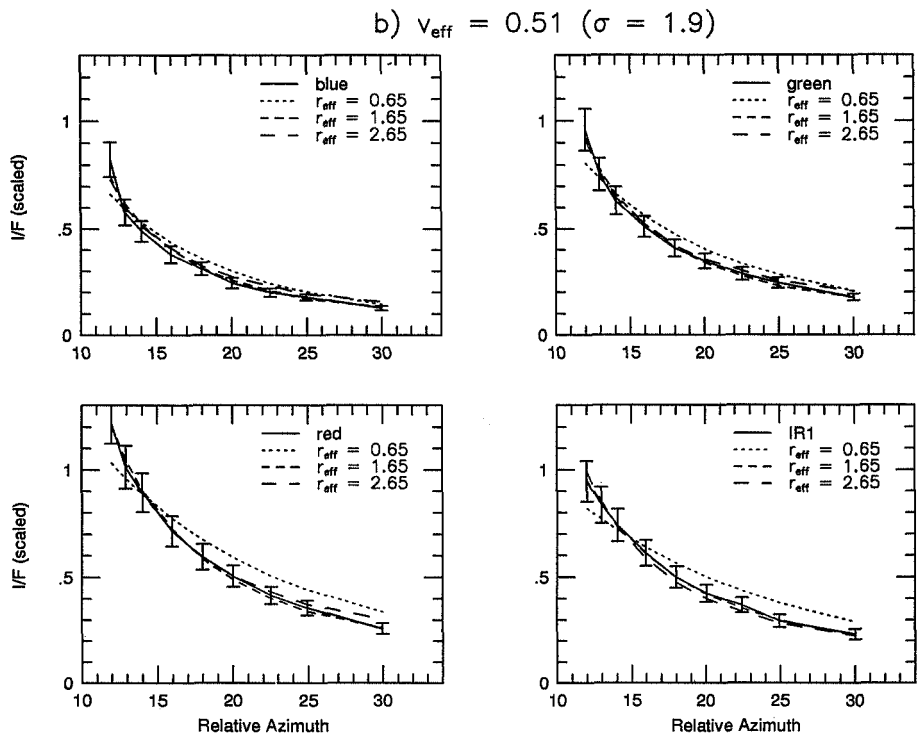


Figure 5. (continued)

The scattering characteristics of the large, nonspherical particles are divided into three components: a diffracted component, an externally reflected component, and internally refracted and transmitted components. Nonspherical theory was most important for the internal components of the scattered light, the other two were equivalent to their counterparts for spheres. *Pollack and Cuzzi* [1980] empirically fit laboratory measurements and found that the data had a linear behavior, ln $I/F \sim \theta$, which can be described by the slope (FTB) of the empirical function in the high scattering angle region. *Showalter et al.* [1992] added a quadratic term to this theory, based on the results of *Liou et al.* [1983], to include a point (ThMin) at which the slope (SLP, which is evaluated at 30° scattering angle) may change, in order to accommodate a modest backscattering peak. Finally, the scattering portion, but not the absorption portion, of the interaction cross section was enhanced by the ratio of the irregular particle's surface area to that of an equal-volume sphere (SAR).

The remaining free parameters in the nonspherical theory are X1 and X2. Mie theory was used to define single-scattering properties for the smallest particles, i.e., those with a size parameter $x < X1$. The nonspherical particle theory was used for the larger particles, i.e., those with a size parameter $x > X2$. A weighted average of the two methods was employed at intermediate values of x to provide a smooth transition from one regime to the next, as opposed to the abrupt transition between the two regimes (X0) originally suggested by *Pollack and Cuzzi* [1980]. Thus in the current version of this nonspherical particle theory there are a total of at least four free parameters: SAR, X1, X2, and FTB. In the quadratic theory, FTB is substituted with SLP and ThMin, making a total of five free parameters. Realistically, we can determine only two of these parameters, while assuming values for the remainder (reasonable values were found by *Pollack and Cuzzi* [1980] for a number of particle shapes). Here, we let SLP and ThMin be the two free parameters because of their direct relationship to the single-scattering phase function at intermediate and large angles of scatter.

In our analysis we assumed a lognormal size distribution for the dust. This distribution function was chosen because it provides a useful representation for many aerosol species on Earth and because there are only two free parameters, r_g and σ:

$$n(r) = \frac{c}{r} \exp\left(-0.5 \frac{(\ln r - \ln r_g)^2}{(\ln \sigma)^2} \right) \quad (1)$$

$$\ln r_g = \int_0^\infty \ln r \; n(r) \; dr \quad (2)$$

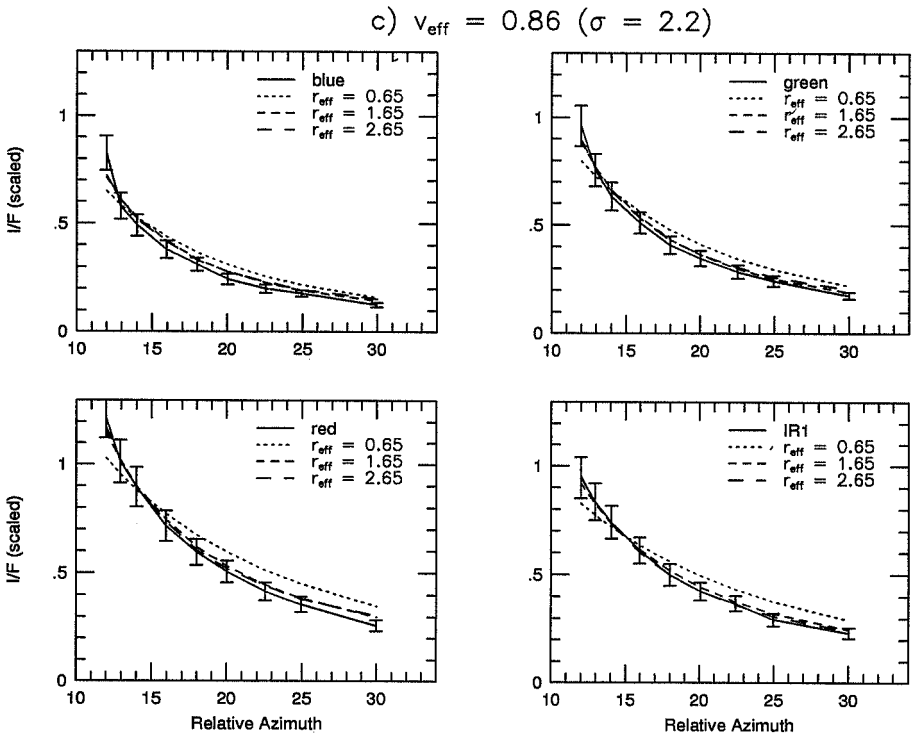


Figure 5. (continued)

$$(\ln \sigma)^2 = \int_0^\infty (\ln r - \ln r_g)^2 \, n(r) \, dr) \qquad (3)$$

where c is a constant, r_g is the modal radius, and σ is a parameter giving the width of the distribution. Note that our parameter σ is equivalent to $\exp(\sigma_g)$ of *Hansen and Travis* [1974]. The two σ are not the same because of slightly different definitions of the lognormal size distribution. The two free parameters (r_g and σ) are associated with the first two moments of the size distribution, r_{eff} and v_{eff}:

$$r_{eff} = \frac{r_g}{\exp\left(-2.5 (\ln \sigma)^2\right)} \qquad (4)$$

$$v_{eff} = \exp\left((\ln \sigma)^2\right) - 1 \qquad (5)$$

It is important to note that v_{eff} can be found from the value of σ, but r_{eff} is dependent on both the values of r_g and σ.

Hansen and Travis [1974] have shown that in many applications, many different particle size distribution functions having the same first two moments have nearly identical single-scattering properties. Conversely, it is difficult to derive much more size distribution information than the first two moments using brightness data alone.

Table 3 summarizes the parameters used to determine the light-scattering properties of the dust particles. Some parameters were fixed to the "nominal" values shown in this table. The range of values used for the free parameters are also listed in Table 3. In section 3 we present the results of a series of sensitivity calculations, in which the impact of varying each of these parameters is tested.

Multiple scattering from the surface can account for less than 1% or as much as 10% of the total intensity. The photometric properties of the surface were modeled using the *Hapke* [1981, 1986] photometric function. The Hapke function has at least four independent variables related to the surface particle characteristics: the single-scattering albedo, the average macroscopic surface roughness, and the width and magnitude of the opposition effect. *Arvidson et al.* [1989] analyzed several Viking Lander images of the surface using the color filters to better quantify the surface-scattering parameters. The values presented in their Table 1 were used in this study.

A single-scattering description of the atmosphere was inadequate for our purposes [*Pollack et al.*, 1977]. Hence we used a model that employs multiple scattering in a plane-parallel atmosphere based on the doubling method [*Hansen*, 1969]. We assumed that the single-scattering properties of

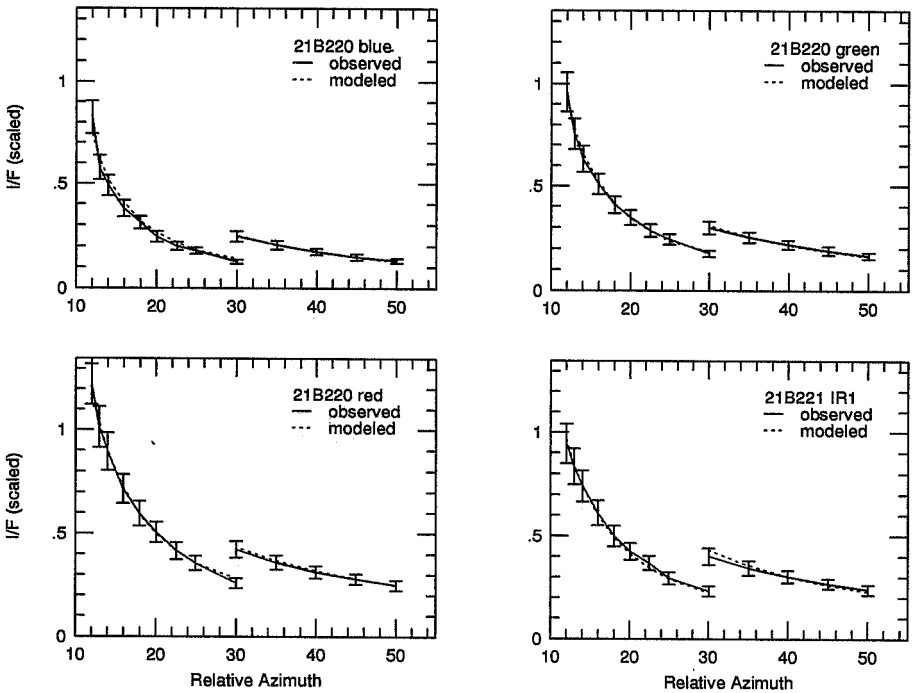


Figure 6. Scans of image 21B220, 21B221 (sol 97) (solid line) taken at the solar elevation angle for relative azimuth angle <30° (with vignetting) and at 15° elevation for angles between 30° and 50° (without vignetting). The scans are plotted with the best fit model (dotted line) v_{eff} = 0.51 μm, r_{eff} = 1.85 μm. The discontinuity at 30° most likely is due to the uncertainty in the absolute calibration of the vignetting correction.

the particles were invariant with altitude. The surface was included via the adding method described by *Hansen* [1969]. The specification of atmospheric optical depth was the only input to this program aside from the single-scattering particle properties. For each of the sky images the optical depth was obtained directly from a sundiode image (λ_{eff} = 0.67 μm) taken as close in time as possible to the image exposure, usually a few hours or at most a few days [*Colburn et al.*, 1989]. At other wavelengths, the optical depth scaled linearly with extinction coefficient.

We used an iterative method to fit the modeled *I/F* to the observed data of a given zone. We solved for the best fit parameter values and their uncertainties by finding the values that minimized χ^2, the weighted sum of squares of the residuals:

$$\chi^2 = \frac{1}{N} \frac{\sum_{j=1}^{N} h \left[I/F(\text{model})_j - I/F(\text{observed})_j \right]^2 w_j}{\sum_{j=1}^{N} w_j} \quad (6)$$

where N is the total number of observations modeled. The weighting factor w estimates variations among the different filters in the total measurement analysis error. A weighting scheme was devised to account for leaks in the detectors, assumed to be the greatest error in the observations. The scaling factor h is chosen to minimize χ^2 for each model parameter and thus fits the modeled data to the shape of the curve. This protocol was followed whenever feasible to remove the influence of errors in absolute calibration, almost always much larger than the relative errors of the derived model parameters. When the absolute value of the data was modeled, h was given a value of 1.

We sequentially analyzed zones 1, 2, and 3 to derive the optimum values of the free parameters of one regime, while using the best current estimates of the free parameters of the other two regimes. Each parameter for a zone was systematically varied, one at a time, and χ^2 was minimized. This procedure was iterated until the free parameters of the zone of current interest changed minimally. We then proceeded to the next zone using these newly determined values from the previous zone. In general, this process had to be repeated several times for the best fit for all of the parameters to be found.

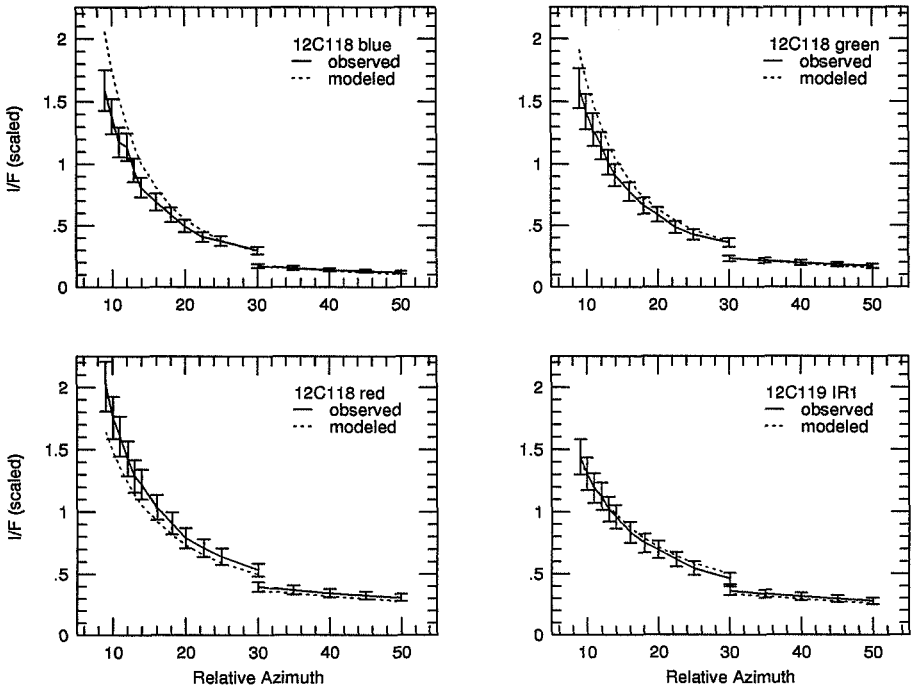


Figure 7. Scans of image 12C118,12C119 (sol 211) (solid line) taken at the solar elevation angle for relative azimuth angle <30° and at 15° elevation for angles between 30° and 50°, plotted with the best fit model (dotted line) $v_{eff} = 0.51$ μm, $r_{eff} = 1.52$ μm. Note that in this case the model value of r_{eff} is too large for the blue and green filter and too small for the red filter. This is true for all cases of v_{eff}.

3. Results

In this section we apply the protocol of section 2 to determine the size, shape, and radiative properties of Martian airborne dust during two very different dust-loading conditions. The first refers to the summer season in the northern hemisphere far in time from the season of global dust storms, when dust loading (or equivalently, optical depth) is close to its minimum value of a few tenths. The second period occurs during midfall in the northern hemisphere at a time close to the peak of the first global dust storm of Viking year 1 (the 1977a storm). In each case, a four-color image sequence (blue, green, red, IR1) was used to determine r_{eff}, v_{eff}, and ω_0 (and hence n_i) from analyses of zones 1 and 2. The optical depth of these afternoon images at an effective wavelength of 0.67 μm was obtained from sundiode images taken in the afternoon of the same Martian day (sol) [*Colburn et al.*, 1989].

We simulated zone 3 data from the panorama image obtained shortly after touchdown of VL1. The dust optical depth was obtained from the closest relevant data points of *Colburn et al.* [1989]. This image was obtained during early northern summer at a time of low optical depth (cf. Table 2) as were the northern midsummer images taken at VL2. We analyzed the two sets of images concurrently to provide the best overall fits for zones 1, 2, and 3. Few images met the necessary constraints for images used in analyzing zone 3 parameters. Thus the shape parameters found for image 12A002 were applied directly and without change to the analyses of the dust storm images. The best fit values of SLP and ThMin are summarized in Table 5. The simulation of the panorama data is shown in Figure 3. A very good, if not perfect, fit is obtained over a large range of scattering angles. The value of ThMin (<180°) implies the presence of a backscattering peak in the dust particles' single-scattering phase function.

3.1. Particle Properties

As alluded to earlier, it is possible to describe a size distribution well using the first two moments, r_{eff} and v_{eff}, when one's interest centers on such radiative transfer quantities as the single-scattering phase function. With the lognormal size distribution, we attempted to find the first two moments that could best account for the scattering behavior of the observed data between 10° and 30° relative azimuth (zone 1). For the purposes of this work, relative azimuth (sun azimuth minus observed azimuth) is roughly equal to

Table 3. Model Parameters and Their Values

Parameter	Type	Nominal Value	Range of Values
r_{eff}	free		0.01 - 10
v_{eff}	free		0.25 - 1.0
n_r	fixed	1.5	
n_i	free		0.0 - 0.1
ω_0	free		0.5 - 1.0
SAR	fixed	1.3	
SLP	free		0.0 - 0.5
ThMin	free		110 - 180
X1	fixed	3.0	
X2	fixed	5.0	

scattering angle, since the data were scanned close to the solar elevation angle. The best values of r_{eff} and v_{eff} were determined by modeling the *I/F* and comparing the shape of the curve in zone 1 to that of the observed data.

The method of determining the best values of r_{eff} and v_{eff} involved systematically altering each variable and finding the minimum value of χ^2 (Figure 4). Three plots are shown, corresponding to alternative choices of σ (or v_{eff}). Line scans giving zone 1 information of the four filter images taken during northern summer were modeled in order to construct these plots. As expected from the discussion in section 2, both small and large values of r_{eff} lead to poor fits to the data (i.e., χ^2 has a high value). Optimum values of r_{eff} lie between about 1 and 3 μm (more exact values will be given later). The double-valued nature of some of the χ^2 curves, as first alluded to by *Pollack et al.* [1979], illustrates the need to use multifilter images. The occurrence of a broad minimum in χ^2 for r_{eff} values lying between about 1 and 3 μm implies that, while rough estimates of r_{eff} can readily be obtained from the Viking sky images, it is much harder to get "exact" values. Our use of four filter images enhances our ability to pinpoint r_{eff} for a given v_{eff}, but we also need to find v_{eff} from the same data.

The best fit value of r_{eff} was found separately for three alternative choices of v_{eff}. The nature of these fits to the northern summer images (sol 97) is illustrated in Figure 5. Table 4 summarizes these best fits by giving the best fit values of r_{eff} and r_g, along with the specified values of v_{eff} and σ. Rather than find the best overall value of r_{eff} by simply averaging the four separate determinations, we searched for one that was able to provide an acceptable fit to all four filters within zone 1 (Figure 6).

Table 4. Size Parameters for a Lognormal Size Distribution Composed of Nonspherical Particles

		Sol 97		Sol 211	
v_{eff}	σ	r_g, μm	r_{eff}, μm	r_g, μm	r_{eff}, μm
0.25	1.6	0.69 ± 0.24	1.20 ± 0.42	0.67 ± 0.10	1.17 ± 0.17
0.51	1.9	0.66 ± 0.12	1.85 ± 0.34	0.54 ± 0.11	1.52 ± 0.32
0.86	2.2	0.50 ± 0.01	2.34 ± 0.07	0.45 ± 0.11	2.14 ± 0.53

It is clear from Figure 5 that only a crude estimate of v_{eff} can be made when the primary goal is to get a fine estimate of r_{eff}. The largest values of v_{eff}, which have size distributions skewed toward numerous big particles, provide noticeably poorer fits to the data than do calculations made for smaller values of v_{eff} (compare Figure 5c with Figures 5a and b). However, airborne dust in the Earth's troposphere lofted by dust storms tends to have broad size distributions [*Westphal et al.*, 1987]. For these reasons, we selected a median value as the best fit, v_{eff} = 0.51 μm ± 0.25 μm (see Table 5).

Estimates of the best fit value of the single-scattering albedo ω_0 were obtained from simulations of zone 2. In modeling zone 2 data, the imaginary index of refraction was systematically varied. This required a fit to the absolute value of the sky brightness. Our best fit parameters, such as ω_0, fit the observed data very well (Figure 6). Values of ω_0 presented in Table 5 vary strongly with wavelength in the visible and near IR due to a strong wavelength dependence of the absorption coefficient of ferric iron contained in the dust [*Soderblom*, 1992]. The efficiency of single scattering by the dust particles in the red wavelengths could account for the reddish color of the Martian sky [*Pollack et al.*, 1977] and the similarity in color between optically thick dust storms and bright surface features [*McCord et al,*. 1977]. The results given in Table 5 are in qualitative agreement with these expectations with one notable exception: ω_0 for the IR1 filter in the non-dust storm time period was found to be slightly smaller than that of the red filter. We expected that ω_0 for the IR1 filter would have been at least as large as ω_0 for the red filter. We advocate that caution be used when examining the value of ω_0 for the IR1 filter because this filter was harder to calibrate. It was partially transparent, so multiple scattering of photons off of the back interface was an important source of photons. Also, this filter was susceptible to degradation by neutrons originating from the RTG power sources. The blue, green, and red filters were not. Although calibration measurements of this effect were taken on the ground prior to launch, they were limited in scope. We therefore believe that the IR1 filter has a ω_0 that is at least as large as that of the red filter, as is indicated in ground-based measurements of bright surface features [*Mustard and Bell*, 1994].

We also obtained estimates of the free parameters of zones 1 and 2 for the dust storm images of Table 2 (sol 211). The

Table 5. Radiative Parameters for a Lognormal Size Distribution Composed of Nonspherical Particles

Sol	Effective Wavelength	ω_0	g	Q_{ext}	n_i
97	0.49	0.79	0.68	2.82	0.010
	0.55	0.84	0.66	2.89	0.007
	0.66	0.94	0.63	3.04	0.0025
	0.86	0.89	0.65	3.11	0.0065
211	0.49	0.73	0.71	2.83	0.018
	0.55	0.81	0.67	2.94	0.010
	0.66	0.93	0.63	3.13	0.004
	0.86	0.92	0.63	3.23	0.0065

For sol 97, r_{eff}=1.85 μm and v_{eff} = 0.51 μm.
For sol 211, r_{eff}=1.52 μm and v_{eff} = 0.51 μm.
Here ω_0, g, and Q_{ext} are good to within 5%; n_i is good to within 10%.

ability of the derived best fit values to fit the observed sky brightness is shown in Figure 7. Note that the model fits to the data of each filter are not nearly as good as for the low dust optical depth case (Figure 6). In this case it was difficult to find an average value of r_{eff} that fit each filter well. It is likely that this is due to the increased importance of multiple scattering in the high optical depth case, which would make the results much more sensitive to slight differences in n_i and r_{eff}. Table 4 gives the numerical values and associated error bars of r_{eff} for each value of v_{eff}. Table 5 summarizes the best fit values for all free parameters and the single-scattering properties. A most surprising finding is that the best fit values of r_{eff} at times of low and large dust optical depth are very similar. This interesting result will be discussed in the next section of the paper.

3.2. Sensitivity Testing

In carrying out the simulations described above, we fixed the values of a number of parameters (cf. Table 3). To determine whether the inferred best fit values of the "free parameters" were robust, we systematically varied, one at a time, the values of the fixed and free parameters. The results of this sensitivity study are summarized in Table 6. We find that the nonspherical parameters are mostly robust (changes of 10% in one parameter affect the others by less than 5%) and they do not have a large effect on the size distribution and imaginary index. SAR was the exception; changes in SAR increased the values of the phase function at the large scattering angles and required compensation by changes in either the value of the imaginary index or the mean size of the particle. This is because SAR provides a weight to the relative contribution by small (x < X1) particles and large (x > X2) particles to the total scattering properties [see *Pollack and Cuzzi*, 1980]. The free parameters (size distribution parameters and the imaginary index) are connected by the transition from zone 1 data to zone 2. That is, the shape of the curve in zone 1 (size parameters) and the absolute value of the data in zone 2 (optical constants) are bound. Decreases in one optical constant (e.g., n_r) can be compensated by increases of the other (e.g., n_i) or by decreases in the mean particle size. The size parameters can also compensate for changes within their own zone (an increase in v_{eff} leads to an increase in r_{eff}) or by a change in one of the optical constants. We believe that we have iterated enough times to distinguish the connections of these parameters and to determine the best fit given the information content of the data.

Table 6. Sensitivity Testing of the Model Parameters

Parameter	Effect
r_{eff}	n_i 10% or v_{eff} 10%
v_{eff}	n_i 10% or r_{eff} 10%
n_r	n_i inversely correlated 10%
n_i, ω_0	either 10% change in n_r or in r_{eff}
SAR	5-10% change in n_i or n_r
SLP	robust
ThMin	robust
X1, X2	2% change in r_{eff} or 10% in SLP

Each parameter was varied by 10% from their best fit values, and the impact on the other variables was assessed.

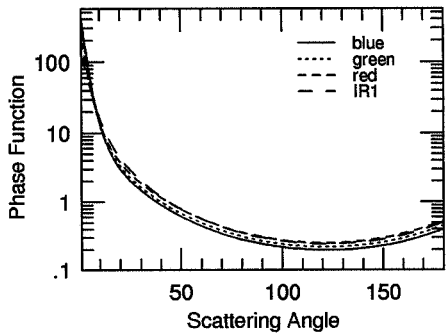


Figure 8. The best fit phase function for images 12A002 (sol 0) and 21B220, 21B221 (sol 97). The phase functions were calculated using a lognormal size distribution with r_{eff} = 1.85 μm and v_{eff} = 0.51μm.

The shape of the single-scattering phase function at the effective wavelengths of the four filters is illustrated in Figure 8. We used the best fit parameters found from the northern summer images to calculate these phase functions. Very similar curves pertain to the best fit parameters of the northern spring dust storm. At scattering angles less than about 30°, with the exact value of this boundary depending on r_{eff} and λ, the phase function increases progressively more sharply with decreasing scattering angle due to diffracted light [*Hansen and Travis*, 1974; *Pollack and Cuzzi*, 1980]. As expected, the phase function $p(\theta)$ has its largest amplitude and narrowest width at the smallest scattering angles for the blue filter due to the size parameter x increasing as λ decreases.

Double refraction events into and out of the particles, and to a lesser degree external reflection, lead to a more gradual fall of $p(\theta)$ at scattering angles lying between about 30° and 50°. Both these types of interactions depend only weakly on the particles' shape. Note that the value of $p(\theta)$ is close to 1 in zone 2, a result anticipated in an earlier section of this paper. It results from the phase function normalization condition that demands that the solid angle average of $p(\theta)$ equal 1. Finally, shape-dependent behavior occurs at scattering angles greater than about 50° due to more complicated paths that are followed by some photons.

A modest backscatter peak at θ=180° (Figure 8) gives additional clues about particle shape, suggesting one or more internal reflections per scattering event [*Pollack and Cuzzi*, 1980]. The amount of light scattered into the large scattering angles is enhanced by partial or total internal reflection and by particle internal geometry, especially equidimensionality. Internal reflections are abetted by sharp corners, such as occur in a cube or parallelepiped. They are also strongly affected by the real index of refraction through the Brewster angle condition for total internal reflection. *Zerull et al.* [1977] note that a backscattering enhancement can be seen in "fluffy" aggregate particles as well.

4. Discussion

4.1. Comparisons With Other Work

Using a number of data sets from past spacecraft missions, as well as Earth-based platforms, several investigators have

derived mean values for the dust particles' single-scattering properties and their size and shape (see Table 1 of this paper, *Kahn et al.*, [1992], and *Murphy et al.* [1993, Table 1]). Most of the work summarized here suggests that the particle single-scattering albedo (ω_0) is quite high (greater than 0.7) and the asymmetry factor (g) is approximately 0.6. This research agrees with those results. However, the many different studies do not agree well on the mean particle radius. The use of the different size distributions among the various research projects has very little effect on the deduced value of the mean size, as long as it is equated to the cross-section weighted mean radius (r_{eff}) and the cross-section weighted mean variance (v_{eff}). Here we make further comparisons with other investigators' values of r_{eff} and v_{eff} to see how well our values correspond. We attempt to understand the reasons for some major differences in the results and give our opinions and reasoning for an optimum choice.

Even though the present work examined some of the same images, the mean particle size values (r_{eff}) are smaller than those presented by *Pollack et al.* [1979]. This can be attributed to the addition of the vignetting correction to the data processing algorithm, to the improved method of data selection at small azimuth angles, to the use of four concurrent narrow filter images, and to the improved computer codes: the χ^2 fitting, the particle properties, and the radiative transfer programs. While *Pollack et al.* [1977, 1979] were aware that vignetting occurred, they could not carefully correct for it and so used data below the area of vignetting (15° elevation angle). Also *Pollack et al.* [1979] only used data from the blue and IR1 filters to obtain their results. Both of these filters have calibration problems, as shown by *Huck et al.* [1975]. The single-scattering properties do not agree well with those presented by *Pollack* [1982], particularly the asymmetry factor (g) and ω_0. The smaller value of g and higher value of ω_0 in this work are consistent with the smaller particle size and smaller values of imaginary index found in this work.

The new values for g obtained in the present work agree well with other estimates, such as those given by *Clancy and Lee* [1991]. The latter result was found by fitting a variety of Henyey-Greenstein phase functions to intensity measurements made at intermediate and large scattering angles with the Viking orbiter IRTM experiment's solar channel. The IRTM solar channel (λ_{eff} = 0.67 μm), which is close to the λ_{eff} of the visible channels of the Viking Lander cameras, has a much broader bandwidth (0.3-3.0 μm). Since g varies slowly with wavelength for our deduced value of r_{eff}, this difference in the width of the bands for the two data sets should not produce major differences in values deduced for g. Thus the value for g obtained by *Clancy and Lee* [1991] (g = 0.55) is viable; it was found with a more empirical approach than ours and was anchored to data other than Lander images at intermediate and large angles of scatter. We prefer the g values given in this paper (g = 0.63 at λ = 0.66 μm) because they are based on fits to one data set over almost a complete suite of scattering angles.

As discussed in the introduction to this paper, a wide range of values has been given for r_{eff}. Here we compare our values with these earlier investigations to see the degree to which these differences stem from time variations, differences in protocol, and the adequacy of the data set for estimating r_{eff}. Our new values are smaller than the value ~2.5 μm given by *Pollack et al.* [1979], who also analyzed our northern summer data set of images. We attribute this difference to the factors enumerated earlier in this subsection, namely, our more thorough data analysis and more careful attempt to get useful *I/F* values at the smallest possible scattering angles.

Significant differences are found when our results are compared with the r_{eff} value (2.5 μm) found by *Toon et al.* [1977], who simulated the Mariner 9 IRIS spectra during the decay of the great global dust storm of 1971. We believe that these differences, if real, may reflect differences in location of the two data sets. The IRIS spectra were all situated near the subsolar point in the southern subtropics within the latitude region of the dust storm source material and were taken at a local time close to the expected diurnal maximum for the ground temperature. Thus the dust in this region may have the largest value for r_{eff}, since some of the biggest particles lofted into the atmosphere from their initial source region may not have had enough time to settle back to the ground. Conversely, all of the data used in this paper were taken at the Viking Lander sites in the northern hemisphere. Assuming the source region is in the southern hemisphere, all of the dust seen in the Viking Lander images would have been advected from the southern hemisphere by the storm, allowing for considerable sedimentation time. Indeed, it is surprising that these two estimates do not diverge more than they do. Interannual variations of dust particles due to local dust storms and ice hazes are other possible sources of the particle size difference.

The values of r_{eff} in this work and in *Clancy and Lee*'s [1991] analysis of the IRTM emission phase function (EPF) sequences do not agree. When investigating the Viking EPF measurements, *Clancy and Lee* [1991] found $r_{eff} \leq 0.4$ μm (Table 1), assuming a montmorillonite composition and modeling the visible/9-μm dust opacity ratio rather than the single-scattering phase function. The Viking IRTM data are of special value, since they come from specially designed sequences that view the same area for a variety of emission and phase angles. Since they encompass only scattering angles that are greater than 60°, they have a much greater sensitivity to particle shape than to particle size. In their analysis they used only measurements made at high scattering angles, and they note that this could cause problems in determining particle size. We believe and R. T. Clancy agrees that this is likely (personal communication, 1993).

In their more recent work, *Clancy et al.* [this issue] use the Mariner 9 IRIS spectra of the 1971 global dust storm in the southern subtropics (as do *Toon et al.* [1977]) to model the particle size distribution assuming a "palagonitelike" composition. The cross-section weighted mean particle radius and variance presented in that study (r_{eff} = 1.8 μm, v_{eff} = 0.8 μm) correspond very well to the work presented here, possibly too well, considering the factors described above for the Toon et al. work.

The r_{eff} values presented by *Korablev et al.* [1993] also agree well with our values (cf. Tables 1 and 4). Korablev et al. present (their Figure 7) altitude profiles of particle effective radius for one of the sequences. It is interesting to note that for the lowest altitudes shown in their plot (12 km) and for a value of v_{eff} close to the value that we quote (0.4), the particle effective radius is approximately 1.8 μm. The Viking data refer to the vertically averaged column of dust (effective altitude 11 km). It would appear from their altitude profiles, then, that we are close to perfect agreement on the mean particle size. As mentioned in the above, we would expect to see large size differences between these two data sets due to the

time and space differences of the observations. Additionally, the method of the Phobos analysis was constrained by the limited wavelength baseline and by the use of Mie theory to extract extinction efficiencies. By contrast, we use data at four wavelengths and model the shape of the diffraction peak, which is nearly particle-shape independent. However, there is a remarkable correlation between the two data sets.

4.2. Dust Particle Properties

Table 5 summarizes the best fit values of the single-scattering properties and size, and shape parameters of Martian airborne dust particles above the Viking Landers at a time of low optical depth (northern summer) and near the peak of the first global dust storm (northern fall), as deduced from the analyses of this paper. They may also represent optimum values of these key properties, as argued in section 4.1. Figure 8 shows the dust particles' single-scattering phase function at the effective wavelengths of the blue, green, red, and IR1 filters of the Viking Lander cameras. Here, we describe and try to understand several striking aspects of these results.

First, to within the error bars of the current analyses, the r_{eff} values derived at times of two very different dust optical depths and activity are close to the same value. This similarity is much greater than might be expected if one postulates greatly reduced dust raising between global dust storms and loss of dust to the surface due to a size-dependent sedimentation rate, with the biggest particles being lost the quickest. One possible reason for the similarity in r_{eff} might be the presence of eddy mixing, a mechanism that is size independent, as a more important dust removal process than sedimentation during most seasons of the year. The key region where eddy mixing must dominate over sedimentation in transporting dust downward is the planetary boundary layer. This mechanism has been invoked to explain the surprising lack of size evolution during global dust storms [*Conrath*, 1975; *Toon et al.*, 1977]. Eddy mixing could be enhanced by local dust storms continually supplying fresh material to the atmosphere. Many local dust storms occur between times of global ones [*Murphy et al.*, 1993]. We suspect that both processes play roles in explaining the similarity in r_{eff}. The combination of these two mechanisms does not allow the dust sufficient time to undergo much size evolution.

Local dust storms as a mechanism makes particular sense if there is a "universal dust," i.e., one having the same composition and size distribution, at source regions on the ground. Some support for there being a "universal" dust on Mars is provided by the similarity in values of the imaginary index of refraction (n_i) and r_{eff} for the two cases of dust loading analyzed here (cf. Tables 4 and 5). Additional support is given by the similarity in geometric albedos and visible and near IR spectra of optically thick airborne dust [*McCord et al.*, 1977] and various bright surface regions [*McCord and Westphal*, 1971; *Bell et al.*, 1990]. The assumption that the dust originates in bright surface source regions is limited by the spatial resolution of the two ground-based data sets. *Mustard and Bell* [1994] note in their comparison of Phobos images and ground-based telescopic images of bright areas that there are many more features in the high-resolution spacecraft data than previously seen in the low-resolution telescopic data.

Finally, we can make a crude estimate of the amount of visible/near IR light absorbing mineral(s) in the airborne dust. Analyses of high resolution spectra of the wavelength region from 0.4 to 0.9 μm strongly suggest that the light absorbing mineral(s) are chiefly iron oxides [*Soderblom*, 1992]. *Huffman and Stapp* [1973] measured the optical constants of magnetite, one of the more promising candidate iron oxides. Assuming that this mineral is one of the light absorbing minerals in Martian dust, we obtain a very rough estimate of its fractional volume by dividing the n_i of airborne dust that was found here by the imaginary index of magnetite at the various wavelengths of interest. In this way we obtain a fractional volume of 1.2% magnetite in the Martian airborne dust. It would be interesting to compare the imaginary index found in our work with those of a number of other rocks and minerals. As yet, such a library of optical constants for the iron oxide minerals and rocks that are likely to be abundant on Mars does not exist.

4.3. Solar Heating

The revised values of some of the radiative properties of the airborne dust particles may engender significant changes in the amount of solar radiation that is absorbed within the atmosphere and at the surface, as compared with that found by *Pollack* [1982]. Here, we make crude estimates of the magnitude of these changes to obtain a preliminary estimate of their significance. A more complete and quantitative analysis of this issue will be made in paper 2 of this series on airborne dust properties, at which point we will have extended the definition of these monochromatic values across essentially the entire solar spectrum.

In the present paper, we have obtained modest to major changes in the values of three key properties of airborne dust in the Martian atmosphere that could alter the solar energy deposition profile, as deduced from the properties given by *Pollack* [1982]. These are the effective particle radius r_{eff}, the single-scattering asymmetry parameter g, and the single-scattering albedo ω_0 for the IR1 filter.

Our preferred values for r_{eff} of 1.5-1.9 μm are somewhat smaller than values of 2.5-2.7 μm obtained in earlier work. The particle size affects the solar energy deposition profile chiefly by playing a major role in determining the wavelength dependence of the dust optical depth (Figure 9). The phase

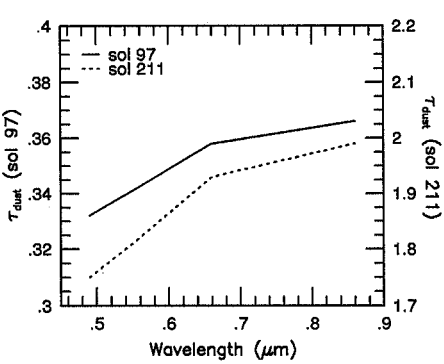


Figure 9. The wavelength dependence of dust optical depth for the two images sol 97 and 211. The optical depth was originally measured by the sundiode instrument at 0.67 μm, equivalent to the red filter. The apparent optical depth scales linearly with Q_{ext}.

shift parameter provides a measure of the phase shift experienced by an axial ray passing through the center of the particle,

$$\rho = \frac{4\pi r_{eff}\,(n_r - 1)}{\lambda} = 2x\,(n_r - 1) \qquad (7)$$

For both the old and revised values of r_{eff}, ρ remains above 2.5 throughout the visible and near IR [*Pollack and Cuzzi*, 1980]. For such values of ρ, Q_{ext} is expected to lie near 2, which is the sum of the geometric optics and diffraction extinction efficiencies. In this case the wavelength dependence of the dust optical depth, which scales linearly with Q_{ext}, changes little throughout the solar-dominated part of the spectrum. Therefore the altered value of r_{eff} should not lead to major changes in the solar energy deposition profile.

The new value of $g = 0.67$ at a wavelength of 0.55 μm, taken here as a representative solar wavelength, is significantly smaller than the value of 0.82 at this same wavelength that was given by *Pollack* [1982]. The new value of g implies more backscattering. This change may be important in the alteration of the solar energy deposition profile. To illustrate, we consider its effect on the albedo A of a global average dust layer with optical depth τ when $\omega_0 = 1$ and the surface albedo is 0. This can be approximated by [*Pollack and Sagan*, 1967]

$$A = \frac{\tau\,(1 - g)}{1 + \tau\,(1 - g)} \qquad (8)$$

When $\tau\,(1 - g) < 1$, a condition met for almost all of the cases of interest, A is proportional to $(1 - g)$. For such a situation the revised value of g leads to a decrease of 65% in the bolometric albedo of an optically thin dust layer of fixed optical depth. The surface absorbs considerably more sunlight when the new value of g is used. By implication, the amount of sunlight absorbed by an optically thin dust layer having a large, but not unit, single-scattering albedo could increase by several tens of percent when the new value of g is used because sunlight penetrates, on average, significantly deeper when g decreases. Hence A decreases and, by conservation of energy, the absorption and transmission increase.

Finally, the new study raises serious questions about the reliability of the ω_0 values for any of the IR filters, including IR1. The single-scattering albedo was derived from the absolute value of the I/F, which may have had much larger error bars for the IR1 filter than for the other filters. In the current studies, ω_0, as deduced for the IR1 filter having an effective wavelength of 0.86 μm, was found to be 0.89-0.92. These values are comparable to those inferred for the red filter at an effective wavelength of 0.66 μm (0.93-0.94). For reasons discussed earlier, we consider the value of ω_0 for the IR filter to be suspect and consider a value larger than ω_0 for the red filter to be more likely. As about 45% of the sunlight at the top of the Martian atmosphere lies within this longer wavelength domain ($\lambda > 0.75$ μm), the use of the higher value of ω_0 for the IR1 filter could significantly decrease the amount of sunlight deposited in the airborne dust. Correspondingly, it would increase the amount of sunlight deposited at the surface by a similar amount (it will not have exactly the same magnitude because the atmosphere/surface system now has a higher albedo). For optically thin situations, the amount of sunlight deposited within the atmospheric dust D_d scales as $(1 - \omega_0)$, implying as much as a 10% reduction in D_d with the hypothetically larger value of ω_0.

In summary, major changes of several tens of percent may result from the revised values of g and ω_0 in the near IR. The revised value of g leads to an enhancement in the amount of sunlight deposited within the atmosphere. Alteration of ω_0 in the near IR has the opposite effect. These two changes, and even the change in particle size, should not greatly disturb current estimates of solar energy deposition in the Martian atmosphere.

5. Summary

In this paper we have presented results of radiative transfer modeling of Viking Lander images. The data analysis and modeling have resulted in significant improvements over past Viking Lander image analyses [*Pollack et al.*, 1977, 1979] to values of a number of key properties of the Martian airborne dust. We also explored some of the implications of these new results.

1. We have analyzed three sets of Viking Lander images to determine the size, shape, and radiative properties of Martian airborne dust during two very different dust-loading conditions as summarized in Table 5 and Figures 3-8. One image was taken by VL2 during the summer season in the northern hemisphere, when dust optical depth was close to its minimum value of a few tenths. The second set of images was taken at VL1 during midfall in the northern hemisphere at a time close to the peak of the first global dust storm of Viking year 1. The third image was a panorama taken by VL1 just after touchdown, when the dust optical depth was quite low. The first two sets of images were used to find the particle cross-section weighted mean radius (r_{eff}) and variance (v_{eff}) and their wavelength dependent single-scattering albedos (ω_0). The best fit v_{eff} for both images was found to be 0.5 ± 0.2 μm. The values of r_{eff} varied from 1.85 ± 0.3 μm for the low dust optical depth VL2 image to 1.52 ± 0.3 μm for the high dust optical depth VL1 image. The panorama image was used to determine the degree of irregularity of the dust particles. The particles are nonspherical. They scatter light fairly efficiently into the backward hemisphere, which implies equidimensionality and possibly very sharp corners.

2. These values of r_{eff} and v_{eff} are surprisingly consistent between the two dust loadings and with the latest research done with Phobos and Mariner 9 IRIS data (see Table 1). One would expect a much greater discrepancy due to the uplifting of dust during the global dust storm. Eddy mixing can be invoked as a more important dust removal mechanism than sedimentation of large particles, since eddy mixing is size independent. This in combination with frequent local dust storms would account for the lack of change of mean particle size with time and location. A second explanation may be the presence of dust source regions where the dust is the same composition and size, a "universal" dust.

3. The single-scattering properties (summarized in Table 5) are consistent with other investigators' results. Over the Viking wavelength range (0.49 to 0.86 μm), the single-scattering albedo ranges from 0.6 to 0.9, the asymmetry factor from 0.6 to 0.7, and the extinction coefficient from 2.8 to 3.2. Also shown in Table 5 are the values of the imaginary index of refraction for each image, assuming that the real index of refraction is 1.5. The imaginary index ranges from 0.01 to 0.0025 for the low dust-loading example (sol 97) and from 0.018 to 0.004 for the dust storm case (sol 211). Due to ground-based calibration and possible RTG damage, absolute

I/F values for the IR1 filter have large errors and thus so do ω_0 and n_i. However, the trends shown in Table 5 (such as the increasing value of imaginary index) appear to be consistent with ground-based results [*Pollack et al.*, 1993].

4. Even though the revised values of r_{eff} are smaller than the values found by *Toon et al.* [1977] and *Pollack et al.* [1979], the extinction values are not largely different. Therefore the solar energy deposition within the atmosphere and at the surface of the planet due to the different values of the mean particle size would change little. Major changes of several tens of percent may result from the revised values of g. The lower value of g leads to an enhancement in the amount of sunlight absorbed by the atmospheric dust. A reduction of absorbed light of corresponding size and opposite magnitude may exist due to a higher ω_0 in the near IR than quoted in Table 5.

5. It is possible to determine the fractional composition of the dust particles from the imaginary index of refraction, given the n_i of other minerals or rocks. An example of 1.2% composition of magnetite is found. It would also be possible to compare the shape of the imaginary index curve to those of iron oxide rocks and minerals and find one or more good candidates for the dust composition. Unfortunately, an extensive library of the complex index of refraction for iron oxide minerals and rocks in the visible wavelengths does not exist yet.

6. Notation

c,α,γ	size distribution adjustable constants.
CCS	Viking Lander camera coordinate system.
χ^2	sum of the squares of the residuals.
FTB	ratio of integral forward light transmission to backward.
g	scattering asymmetry factor.
I/F	ratio of observed brightness to that of a Lambert surface.
λ	wavelength of light in micrometers, μm.
L_s	areocentric longitude of Mars.
n	differential number of particles with radius r.
n_i	imaginary index of refraction.
n_r	real index of refraction.
$p(\theta)$	normalized phase function.
Q_{ext}	particle extinction efficiency.
r	particle radius.
r_{eff}	geometric cross-section weighted mean particle radius.
r_g	modal radius.
σ	size distribution width parameter.
SAR	ratio of irregular particle surface area to equal volume sphere.
SLP	slope of phase function at 30°.
sol	number of Martian days since VL1 touchdown.
θ	scattering angle.
ThMin	location of phase function minimum.
v_{eff}	geometric cross-section weighted variance.
ω_0	single-scattering albedo.
x	particle size parameter, equal to $2\pi r / \lambda$.
X1	maximum value of x for Mie particles.
X2	minimum value of x for nonspherical particles.

Acknowledgments. The authors wish to thank R. E. Arvidson, J. Bell, A. F. C. Bridger, R. Haberle, J. Murphy, and T. Roush for helpful discussions and insights. The authors are indebted to M. Showalter for versions of his Mie and nonspherical computer code, radiative transfer code, and for his help in implementing changes to these codes. The time and effort put into the review of this manuscript by T. Z. Martin and S. W. Lee are greatly appreciated. Funds for the support of this study have been allocated to the San Jose State University Foundation by NASA Ames Research Center under interchange NAS2-740. This work was supported in part by a grant from NASA's Planetary Atmospheres program, RTOP 154-10-80-01.

References

Arvidson, R. E., E. A. Guiness, M. A. Dale-Bannister, J. Adams, M. Smith, P. R. Christensen, and R. B. Singer, Nature and distribution of surficial deposits in Chryse Planitia and vicinity, Mars, *J. Geophys. Res., 94*, 1573-1587, 1989.

Bell, J.F., III, T.B. McCord, and P. D. Owensby, Observational evidence of crystalline iron oxides on Mars, *J. Geophys. Res., 95*, 14,447-14,461, 1990.

Chassefiere, E., J. E. Blamont, V. A. Krasnopolsky, O. I. Korablev, S. K. Atreya, and R. A. West, Vertical structure and size distributions of Martian aerosols from solar occultation measurements, *Icarus, 97*, 46-69, 1992.

Chylek, P., and G. W. Grams, Scattering by nonspherical particles and optical properties of Martian dust, *Icarus, 36*, 198-203, 1978.

Clancy, R.T., and S.W. Lee, A new look at dust clouds in the Martian atmosphere: Analysis of emission-phase-function sequences from global Viking IRTM observations, *Icarus, 93*, 135-158, 1991.

Clancy, R.T., S.W. Lee, G. R. Galdstone, W. W. McMillan, and T. Roush, A new model for Mars atmospheric dust based upon analysis of ultraviolet through infrared observations from Mariner 9, Viking, and Phobos, *J. Geophys. Res*, this issue.

Colburn, D. J., J. B. Pollack, and R. Haberle, Diurnal variations in optical depth at Mars, *Icarus, 79*, 159-189, 1989.

Conrath, B. J., Thermal structure of the Martian atmosphere during the dissipation of the dust storm of 1971, *Icarus, 24*, 36-46, 1975.

Drossart, P., J. Rosenquist, S. Erard, Y. Langevin, J.-P. Bibring, and M. Coombes, Martian aerosol properties from the Phobos ISM experiment, *Ann. Geophys., 9*, 754-760, 1991.

Gierasch, P.J., and R.M. Goody, The effect of dust on the temperature structure of the Martian atmosphere, *J. Atmos. Sci., 29*, 400-402, 1972.

Haberle, R.M., C. B. Leovy, and J. B. Pollack, Some effects of dust storms on the global circulation of Mars, *J. Atmos. Sci., 50*, 322-367, 1982.

Hansen, J.E., Radiative transfer by doubling very thin layers, *Astrophys. J., 155*, 565, 1969.

Hansen, J.E., and L.D. Travis, Light scattering in planetary atmospheres, *Space Sci. Rev., 16*, 527-587, 1974.

Hapke, B., Bidirectional reflectance spectroscopy, 1 Theory, *J. Geophys. Res., 86*, 3039-3054, 1981.

Hapke, B., Bidirectional reflectance spectroscopy, 4, The extinction coefficient and the opposition effect, *Icarus, 67*, 264-280, 1986.

Huck, F. O., E. E. Burcher, E. J. Taylor, and S. D. Wall, Radiometric behavior of the Viking Mars lander cameras, *NASA Tech. Rep., TMX-72692*, 1975.

Huffman, D. K., and J. L. Stapp, Optical measurements on solids of possible interstellar importance, in *Interstellar Dust and Related Topics*, edited by M. Greenburg and H. C. van de Hulst, p. 297, D. Reidel, Norwell, Mass., 1973.

Kahn, R. A., T. Z. Martin, R. W. Zurek, and S. W. Lee, The Martian dust cycle, in *Mars*, edited by H. H. Kieffer et al., p. 1017-1053, University of Arizona Press, Tucson, 1992.

Korablev, O. I., V. A. Krasnopolsky, A. V. Rodin, and E. Chassefiere, Vertical structure of Martian dust measured by solar infrared occultations from the Phobos spacecraft, *Icarus, 102*, 76-87, 1993.

Liou, K. N., Q. Cai, J. B. Pollack, and J. N. Cuzzi, Light scattering by randomly oriented cubes and parallelepipeds, *Appl. Opt., 22*, 3001-3008, 1983.

Martin, T. Z., Thermal infrared opacity of the Mars atmosphere, *Icarus, 66*, 2-21, 1986.

McCord T.B. and J. A. Westphal, Mars: Narrowband photometry, from 0.3 to 2.5 microns, of surface regions during the 1969 apparition, *Astrophys. J., 168,* 141-153, 1971.

McCord, T. B., R. L. Huguenin, D. Mink, and C. Pieters, Spectral reflectance of Martian areas during the 1973 opposition: Photoelectric filter photometry 0.33-1.10 μm, *Icarus, 31,* 25-39, 1977.

Murphy, J.R., R. M. Haberle, O. B. Toon, and J. B. Pollack, Martian global dust storms: Zonally symmetric numerical simulations including size-dependent particle transport, *J. Geophys. Res., 98,* 3197-3220, 1993.

Mustard, J.F. and J.F. Bell III, New composite reflectance spectra of Mars from 0.4 to 3.14 μm, *Geophys. Res. Lett., 21,* 353-356, 1994.

Pang, K., J. M. Ajello, C. W. Hord, and W. G. Egan, Complex refractive index of Martian dust: Mariner 9 ultraviolet observations, *Icarus, 27,* 55-67, 1976.

Pollack, J.B., Properties of dust in the Martian atmosphere and its effect on temperature structure, *Adv. Space Res., 2,* 45-56, 1982.

Pollack, J.B., and J.N. Cuzzi, Scattering by nonspherical particles of size comparable to a wavelength: A new semi-empirical theory and its application to tropospheric aerosols, *J. Atmos. Sci., 37,* 868-881, 1980.

Pollack, J. B., and C. Sagan, An analysis of Martian photometry and polarimetry, *Space Sci. Rev., 9,* 243-299, 1969.

Pollack, J.B., R. Kahn, J. Hunter, W. Van Camp, C. E. Carlson, and D. C. Pidek, Properties and effects of dust suspended in the Martian atmosphere, *J. Geophys. Res., 82,* 4479-4496, 1977.

Pollack, J.B., D. S. Colburn, F. M. Flaser, R. Kahn, C. E. Carlson, and D. C. Pidek, Properties and effects of dust suspended in the Martian atmosphere, *J. Geophys. Res., 84,* 2929-2945, 1979.

Pollack, J. B., M. E. Ockert-Bell, R. Arvidson, and M. Shepard, The wavelength dependence of Martian atmospheric dust radiative properties, *LPI Tech. Rep., 93-06,* 44, 1993.

Santee, M., and D. Crisp, Thermal structure and dust loading of the Martian atmosphere during late southern summer: Mariner 9 revisited, *J. Geophys. Res., 98,* 3261-3279, 1993.

Showalter, M. R., J. B. Pollack, M. E. Ockert, L. R. Doyle, and J. B. Dalton, A photometric study of Saturn's F ring, *Icarus, 100,* 394-411, 1992.

Soderblom, L.A., The composition and mineralogy of the Martian surface from spectroscopic observations: 0.3-50 μm, in *Mars,* edited by H. H. Kieffer et al., pp. 557-593, University of Arizona Press, Tucson, 1992.

Toon, O. B., J. B. Pollack, and C. Sagan, Physical properties of particles composing the Martian dust storm of 1971-1972, *Icarus, 30,* 663-696, 1977.

Tucker, R. B., Viking lander imaging investigation picture catalog of primary mission experiment data record, *NASA Ref. Publ., 1007,* 1978.

Veverka, J., The meaning of Russell's law, *Icarus, 14,* 284-285, 1971.

Westphal, D. L., O. B. Toon, and T. N. Carlson, A two-dimensional numerical investigation of the dynamics and microphysics of Saharan dust storms, *J. Geophys. Res., 92,* 3027-3049, 1987.

Wolf, M., Viking lander camera vignetting, *JPL Interoffice Memo. 76-295,* Jet Propul. Lab., Pasadena, Calif., 1976.

Zerull, R. H., R. H. Giese, and K. Weiss, Scattering measurements of irregular particles vs. Mie theory, *Proc. Soc. Photo-Opt. Instrum. Eng., 112,* 191-199, 1977.

M. E. Ockert-Bell, NASA Ames Research Center 245-3, Moffett Field, CA, 94035. (e-mail: mbell@anarchy.arc.nasa.gov)

M. K. Shepard, CEPS MRC 315, National Air and Space Museum, Smithsonian Institution, Washington, D. C., 20560. (e-mail: mshepard@ceps.nasm.edu)

(Received March 28, 1994; revised October 3, 1994; accepted October 5, 1994.)

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5251–5263, MARCH 25, 1995

A new model for Mars atmospheric dust based upon analysis of ultraviolet through infrared observations from Mariner 9, Viking, and Phobos

R. T. Clancy,[1] S. W. Lee,[2] G. R. Gladstone,[3] W. W. McMillan,[4] and T. Rousch[5]

Abstract. We propose key modifications to the Toon et al. (1977) model of the particle size distribution and composition of Mars atmospheric dust, based on a variety of spacecraft and wavelength observations of the dust. A much broader (r_{eff}variance~0.8 μm), smaller particle size (r_{mode}~0.02 μm) distribution coupled with a "palagonite-like" composition is argued to fit the complete ultraviolet-to-30-μm absorption properties of the dust better than the montmorillonite-basalt, r_{eff}variance= 0.4 μm, r_{mode}= 0.40 dust model of Toon et al. Mariner 9 (infrared interferometer spectrometer) IRIS spectra of high atmospheric dust opacities during the 1971-1972 Mars global dust storm are analyzed in terms of the Toon et al. dust model, and a Hawaiian palagonite sample (Roush et al., 1991) with two different size distribution models incorporating smaller dust particle sizes. Viking Infrared Thermal Mapper (IRTM) emission-phase-function (EPF) observations at 9 μm are analyzed to retrieve 9-μm dust opacities coincident with solar band dust opacities obtained from the same EPF sequences (Clancy and Lee, 1991). These EPF dust opacities provide an independent measurement of the visible/9-μm extinction opacity ratio (≥2) for Mars atmospheric dust, which is consistent with a previous measurement by Martin (1986). Model values for the visible/9-μm opacity ratio and the ultraviolet and visible single-scattering albedos are calculated for the palagonite model with the smaller particle size distributions and compared to the same properties for the Toon et al. model of dust. The montmorillonite model of the dust is found to fit the detailed shape of the dust 9-μm absorption well. However, it predicts structured, deep absorptions at 20 μm which are not observed and requires a separate ultraviolet-visible absorbing component to match the observed behavior of the dust in this wavelength region. The modeled palagonite does not match the 8- to 9-μm absorption presented by the dust in the IRIS spectra, probably due to its low SiO_2 content (31%). However, it does provide consistent levels of ultraviolet/visible absorption, 9- to 12-μm absorption, and a lack of structured absorption at 20 μm. The ratios of dust extinction opacities at visible, 9 μm, and 30 μm are strongly affected by the dust particle size distribution. The Toon et al. dust size distribution (r_{mode}= 0.40, r_{eff}variance= 0.4 μm, $r_{cw\mu}$ =2.7 μm) predicts the correct ratio of the 9- to 30-μm opacity, but underpredicts the visible/9-μm opacity ratio considerably (1 versus ≥2). A similar particle distribution width with smaller particle sizes (r_{mode}= 0.17, r_{eff}variance= 0.4 μm, $r_{cw\mu}$=1.2 μm) will fit the observed visible/9-μm opacity ratio, but overpredicts the observed 9-μm/30-μm opacity ratio. A smaller and much broader particle size distribution (r_{mode}= 0.02, r_{eff}variance= 0.8 μm, $r_{cw\mu}$ =1.8 μm) can fit both dust opacity ratios. Overall, the nanocrystalline structure of palagonite coupled with a smaller, broader distribution of dust particle sizes provides a more consistent fit than the Toon et al. model of the dust to the IRIS spectra, the observed visible/9-μm dust opacity ratio, the Phobos occultation measurements of dust particle sizes (Chassefiere et al., 1992), and the weakness of surface near IR absorptions expected for clay minerals (Clark, 1992; Bell and Crisp, 1993).

[1]Space Science Institute, Boulder, Colorado.
[2]Laboratory for Atmospheric and Space Physics, University of Colorado at Boulder.
[3]Southwest Research Institute, San Antonio, Texas.
[4]Department of Geosciences, San Francisco State University, San Francisco, California.
[5]NASA/Ames Research Center, Moffet Field, California.

Paper number 94JE01885.
0148-0227/95/94JE-01885$05.00

Introduction

The long-standing model of Mars atmospheric dust is a predominantly montmorillonite composition with a particle size distribution corresponding to a cross-section weighted mean radius ($r_{cw\mu}$) of 2.7 μm (r_{mode} = 0.4 μm). The basis for this model was an analysis of Mariner 9 infrared interferometer spectrometer (IRIS) observations (5-50 μm) of the 1971 global dust storm, performed by *Toon et al.* [1977]. The actual conclusions from the Toon et al. analysis regarding the dust composition were not as specific as the montmorillonite model generally adopted by the Mars science community. Toon et al. concluded that the dust was a mixture of weathering products from silicic rocks (SiO_2>60%) and a secondary component such

as basalt. The initial IRIS analysis by *Hanel et al.* [1972] estimated a silica content of 60±10% for the atmospheric dust, based on the wavelength center of the 9-μm band absorption. The clay mineral montmorillonite provided the best fit to the 9-μm band absorption in the IRIS spectra and was considered by Toon et al. as one possible candidate for the silicic weathering product. Subsequently, *Pollack et al.* [1977, 1979] analyzed Viking Lander observations of visible (0.4-1.1 μm) sky brightness and transmission during the 1977 global dust storms. These observations were interpreted as confirmation of the montmorillonite, $r_{cw\mu}$=2.7-μm model of Mars atmospheric dust, with two important modifications. A non spherical shape for the dust particles was modeled to explain the observed single-scattering phase function of the dust, and a separate visible/ultraviolet absorbing component was postulated to match these Viking Lander observations and earlier Mariner 9 ultraviolet spectrometer (UVS) observations [*Pang and Ajello*, 1977]. *Pollack et al.* [1979] suggested that a several percent component of magnetite in the Mars atmospheric dust could provide the observed visible and ultraviolet absorption by the dust, since neither montmorillonite nor basalt absorbs visible or ultraviolet light efficiently.

Since these key studies, a number of important measurements ranging from analysis of Viking 9 micron dust opacities [*Martin*, 1986] to Phobos near-IR extinction observations of Mars dust [*Chassefiere et al.*, 1992] have been obtained. *Hunt* [1979] pointed out that the observed ratios of 9- to 20-μm dust absorption in Viking infrared thermal mapper (IRTM) measurements during the 1977 dust storms were not consistent with a simple montmorillonite composition, a conclusion which *Toon et al.* [1977] also determined from their analysis of the Mariner 9 IRIS spectra. Both Toon et al. and Hunt suggested that an additional component, such as basalt, might account for the relative weakness of the 20-μm dust absorption. *Zurek* [1982] pointed out that the ratio of visible-to-IR extinction opacities for the dust, obtained from comparison of the Viking Lander and IRTM measurements (~2), was not consistent with the ratio predicted by the standard dust model of Toon et al. (~1). *Toon et al.* [1977] also examined the visible-to-IR opacity ratio of the dust, but relied upon Mariner 9 UVS determinations of dust opacity [*Pang and Hord*, 1973], which were apparently inaccurate for the determination of this ratio. More recent studies include an analysis of Viking solar-band IRTM emission-phase-function (EPF) sequences by *Clancy and Lee* [1991], which suggested smaller absorption (visible single-scattering albedo near 0.92 versus a value of 0.86 from *Pollack et al.* [1979]) and larger backscattering (single-scattering asymmetry parameter of 0.55 versus 0.79 from *Pollack et al.* [1979]) for Mars atmospheric dust at visible/near-IR wavelengths. Clancy and Lee also noted that smaller particle sizes for the dust would be consistent with the observed visible/IR opacity ratio. Phobos solar occultation measurements of dust extinction at wavelengths of 0.75-3.15 microns indicated a value for $r_{cw\mu}$ nearer to 1.0-1.5 microns (r_{mode} ~0.2 μm) at 15- to 25-km altitude [*Chassefiere et al.*, 1992]. With regard to the dust composition, *Clark* [1992] interpreted ground-based near-IR spectra of Mars surface reflectance to place a very low limit (<1%) on the amount of crystalline montmorillonite-like clay materials. *Bell and Crisp* [1993] identified weak absorption features in the 2.20- to 2.25-μm wavelength region of ground-based reflectance observations, which suggested a small but measurable clay component to the Mars bright soil. In either case, the dominant character of the Mars bright regions appears more consistent with amorphous or poorly crystalline materials such as palagonite [e.g., *Singer*, 1982; *Morris et al.*, 1990; *Orenberg and Handy*, 1992; *Banin et al.*, 1992; *Soderblom*, 1992; *Roush et al.*, 1993; *Bell et al.*, 1993; *Murchie et al.*, 1993], than with soil containing large fractions of crystalline clay minerals such as montmorillonite.

In the following analysis, we consider a "palagonite-like" dust composition and a smaller and considerably broader particle size distribution than the Toon et al. size distribution as an alternative model for Mars atmospheric dust. We do not attempt to model a wide range of materials and particle sizes and, hence, cannot claim a unique solution for the composition and particle sizes of Mars atmospheric dust. In fact, the term "palagonite" accommodates a wide range of compositions and component crystallinity (usually amorphous to noncrystalline), since it refers to the weathering product of basaltic glass. Rather, we attempt to demonstrate that a single, plausible component and particle size distribution can fit the full range of Mars dust observations to a much better degree than does the currently accepted model of Mars dust. Our demonstration is made with a particular sample of palagonite from the upper slopes of Mauna Kea [*Singer*, 1982; *Clark et al.*, 1990; *Roush et al.*, 1991], which has been found to be a good spectral analog for Mars dust [e.g., *Singer*, 1982; *Morris et al.*, 1990]. One of the key elements in conducting this analysis is the ability to compute 7- to 30-μm dust absorption spectra for comparison to the 1971 Mariner 9 IRIS spectra, as well as the Viking IRTM 9-μm band observations in 1977.

Infrared Spectroscopic Modeling of Mars Atmospheric Dust

In order to reanalyze the key IRIS spectral observations as well as Viking IRTM EPF 9-μm sequences, one of us (R. Gladstone) has modified a doubling-and-adding radiative transfer code developed for terrestrial atmosphere studies [*Gladstone et al.*, 1984]. The adapted code is capable of accurate dust emission-absorption-scattering radiative transfer calculations over the 5- to 30-micron wavelength region for variable dust composition and particle size inputs, and incorporates both the Viking IRTM channel weightings and the Mariner 9 IRIS wavelength resolution for direct comparisons to these data sets. Two-to-five kilometer vertical resolution between the surface and 50-km altitude and 4 quadrature zenith angles characterize the radiative transfer modelling. We adopt atmospheric temperature profiles according to the algorithm of *Martin* [1986] in the case of the Viking IRTM comparisons, and estimated Mariner 9 IRIS temperature profiles from the 15-micron CO_2 band from IRIS spectra of dusty periods, for the case of the IRIS comparisons. Atmospheric opacity from CO_2, which is the primary molecular absorber in the 7- to 30-μm wavelength region, is not modeled in the dust radiative transfer. Hence, the 12- to 18-μm region, where the CO_2 opacity is important, is excluded in our analysis, as it was in the Toon et al. analysis. We model the dust absorption/scattering properties with Mie single scattering theory and the optical constants for a particular specimen of Hawaiian palagonite as retrieved by Roush [*Roush et al.*, 1991] for the 5- to 25-μm wavelength region, and over the 0.3- to 5.0-μm region as described below. For comparison, we also model the infrared absorption spectrum of montmorillonite, employing the infrared optical constants of montmorillonite 219b from the analysis of *Toon et al.* [1977].

We consider palagonite as the primary alternative to the montmorillonite model of Mars atmospheric dust, based on several considerations. As a primary weathering product of

basalt, palagonite is a very plausible constituent of the bright and dark regions on Mars. Palagonite absorbs in the ultraviolet and visible wavelength region due to its Fe^{+3} content. Montmorillinite does not absorb at UV/visible wavelengths, although it can present such absorption if coated by a ferric oxide stain [*Hunt et al.*, 1973; *Toon et al.*, 1977]. Palagonite may be consistent with the observed lack or weakness of 2.2- to 2.25-μm clay absorption for the bright regions of the Mars surface, due to the amorphous to nanocrystalline structure of many palagonites such as the modeled Mauna Kea sample. Furthermore, as apparent in the palagonite sample we model below, palagonite does not necessarily display strong, structured absorption near 20 microns as does montmorillonite. The observed lack of this 20-μm absorption feature in Mariner 9 IRIS spectra of the 1971 dust storm was emphasized by *Toon et al.* [1977] as an indication that pure montmorillonite was an incomplete model of Mars atmospheric dust. In fact, all of the additional compositions considered by Toon et al. (quartz, basalt, granite, andesite, basaltic glass, and obsidian) exhibited distinctive 20-μm absorptions which were not apparent in the IRIS spectra of Mars atmospheric dust. Montmorillinite was notable primarily for its ability to match the observed 9-μm absorption of Mars atmospheric dust. Montmorillinite with 25-50% basalt by weight led to a better fit of the observed 20-μm absorption, but also seriously degraded the model fit to the observed 9-μm absorption [*Toon et al.*, 1977].

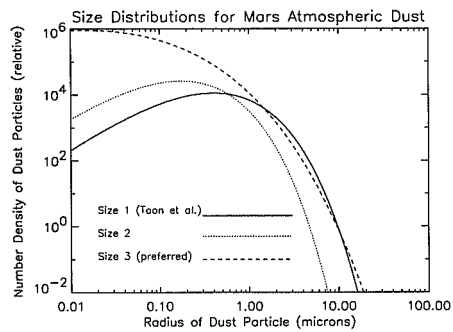


Figure 1b. Dust particle size distributions, based on the modified gamma distribution. All of the model calculations presented employ one of these size distributions, for which the parameters are provided in Table 2. Size distribution 1 corresponds to the Toon et al. model; size distribution 2 exhibits a size distribution width similar to the Toon et al. model, but with roughly one half the average particle sizes. Size distribution 3 exhibits a very broad size distribution width, and is found to obtain a dust opacity variation between visible and 30-μm wavelengths most appropriate to the observed variation.

Mauna Kea Palagonite

The palagonite sample was originally collected at several locations on the upper slopes of Mauna Kea, Hawaii [*Roush*, 1987]. Subsequent to collection, the sample whose optical constants are used here was separated from the bulk sample by liquid suspension in methanol. Initial characterization of this sample included X ray diffraction and determination of major elements via X ray fluorescence analyses. The sample was X ray amorphous and the major elements are provided by *Roush et al.* [1991]. Subsequent to the initial analyses, the sample was studied using transmission electron microscopy (TEM) [*Roush and Blake*, 1991]. The point-to-point resolution of the microscope in conventional imaging mode was about 0.5 nm, the lateral spatial resolution for elemental analysis was 30-50 nm, and selected area electron diffraction patterns could be recorded from particles as small as 100 nm laterally. Qualitative elemental abundances of 101 grains showed that Si, Al, O, Fe, and Ti were most abundant and lesser amounts of Ca, Cr, Mg, and P were present in some samples [*Roush and Blake*, 1991)] Grain-to-grain compositions were highly variable, but three major classes of grains having Si, Al, and Fe as the most abundant element were identified [*Roush and Blake*, 1991]. A limited number of grains from each major class were selected for characterization by electron diffraction. Most grains did not exhibit electron diffraction patterns indicating that they were amorphous. The diffraction patterns observed were consistent with layer lattice silicates [*Roush and Blake*, 1991]. Five of six grains that had Al as the most abundant element did not yield a diffraction pattern. A particle size distribution was determined by measuring the major (a) and minor (b) axes of 610 grains from micrographs calculating the projected elliptical area (πab), and the equivalent spherical particle diameter of this sample was determined to be 1.02±0.66 μm [*Roush and Blake* 1991].

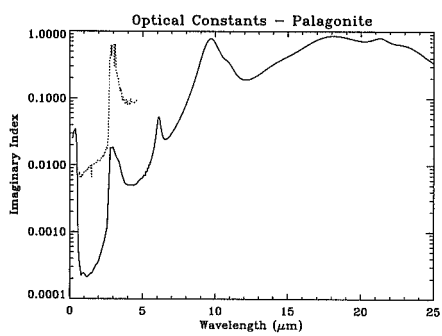


Figure 1a. The wavelength dependence of the imaginary index of refraction for the modeled Mauna Kea palagonite. The 0.3- to 0.7-μm dependence is adopted from *Clark et al.* [1990]. The 5- to 25- μm dependence is taken from *Roush et al.* [1991]. The calculated near infrared (0.7-5.0 μm) dependence (dotted line, see text) has been shifted to merge with the 5- to 25- μm behavior, in agreement with the results of *Clark et al.* [1990].

Over the infrared wavelength region (5-25 μm) we adopt the palagonite optical constants presented by *Roush et al.* [1991]. The imaginary index versus wavelength is presented in Figure 1a. In the near infrared (0.7 to 4.6 μm), the indices of refraction were derived using the techniques described by *Clark and Roush* [1984] and *Clark et al.* [1990]. This technique uses the measured reflectance of the sample, knowledge of the grain size, assumes the real index does not vary significantly, and iteratively applies *Hapke's* [1993] radiative transfer model to vary the imaginary index until the computed reflectance matches the measured reflectance. The initially derived values of the imaginary index using this technique are plotted, along with the infrared values

[*Roush et al.*, 1991], as the dotted line in Figure 1a. Clearly, there is a discrepancy in the imaginary index near 5 μm, between the long wavelength region of the near-infrared and the short wavelength region of the infrared values. To resolve this discrepancy, we note that there are fewer assumptions used in deriving the infrared values than the near-infrared values. As a result, we believe that the infrared imaginary indices are more accurate than the near-infrared values. Additionally, Clark *et al.* [1990] derived 0.3- to 5-μm absorption coefficients for a similar material, and when the absorption coefficients are converted to imaginary indices, they compare quite well with the values derived in the infrared for the sample studied here. Hence, we have arbitrarily scaled the near-infrared imaginary indices derived for the sample studied here to agree with the derived infrared values. These rescaled values are also shown in Figure 1a, as the solid line curve, which now merges with the mid-infrared values from *Roush et al.* [1991], and with the ultraviolet and visible values adopted from *Clark et al.* [1990].

In the visible and near-IR, the spectral properties of palagonites exhibit some variability (e.g., *Soderblom et al.*, 1978; *Evans et al.*, 1981; *Singer*, 1982; *Morris et al.*, 1990; *Bell et al.*, 1993; *Golden et al.*, 1993; *Bishop et al.*, 1993], so one would expect some variability in their infrared spectral behavior as well. *Roush* [1992] investigated the transmission properties of several palagonites collected from different rainfall regimes on the island of Hawaii. The transmission behavior of these samples generally fell into two distinct categories based on observed spectral features. One category exhibited a Si-O stretching fundamental that occurred at wavelengths ≤10 μm, a weak feature near 10.53 μm that can be associated with an Fe-OH stretching mode, and a doublet in the region Si-O bending fundamental, with maximum absorptions near 18.2 and 23.0 μm. The other exhibited a Si-O stretching fundamental that occurred at wavelengths ≥10 μm and only a single feature in the region Si-O bending fundamental, at approximately 16.1 μm. *Roush et al.* [1993] show the transmission spectra of two other palagonites that also illustrate this behavior.

Comparisons to Mariner 9 IRIS Spectra

We present model fits of palagonite and montmorillonite 219b to four individual IRIS spectra, obtained early during the Mariner 9 mission, when the atmospheric dust optical depths were large. The four Mars regions corresponding to the modeled IRIS observations are in Hellas Planitia (40°S, 306°W), Solis Planum (34°S, 10°W), Argyre Planitia (55°S, 51°W), and southern Chryse Planitia (4°N, 25°W). The solar longitudes of these IRIS spectral observations are L_S= 298° (orbit 17), 298° (orbit 18), 301° (orbit 28), and 305° (orbit 40), respectively. Table 1 presents a list of observational parameters for each of these IRIS spectral measurements. Because we present individual IRIS spectra for all four cases (i.e., no spatial or temporal averaging), the spectral noise is quite high in the 5- to 7-micron wavelength region. *Toon et al.* [1977] presented model comparisons to averages of three to 12 spectra within the Hesperia Planum region (20°S-40°S, 190°W-250°W) for orbits 8, 56, 80, 138, and 154. Toon et al. found the largest dust IR optical depths for orbits 8 ($\tau_{9\mu m}$~1.5) and 56 ($\tau_{9\mu m}$~0.8), and no clear indication of any change in the dust IR optical properties over the range of time studied (L_S~295-350°). As demonstrated below, we derive comparable dust opacities ($\tau_{9\mu m}$=1.1-0.7) and spectral line fits (with the montmorillonite 219b, $r_{cw\mu}$=2.7 μm model) for the three more southern regions (55, 40, and 34°S) we analyze here. Our more northern region (4°N) presents a smaller dust optical depth ($\tau_{9\mu m}$~0.5), suggesting a possible hemispheric gradient in the atmospheric dust loading at this time (L_S~300°). Perhaps more interesting, we find the highest dust opacity for the Solis Planum region, which has often been cited as a region of intense dust storm activity. This observation corresponds to a high-altitude region ($\tau_{9\mu m}$~1.1, altitude~7 km), and yet it indicates a significantly higher dust opacity than observed over the low-elevation Hellas region ($\tau_{9\mu m}$~0.7, altitude~-4 km) at the same time (i.e., orbit 17 versus orbit 18).

Temperature profiles used for interpretation of the Mariner 9 IRIS spectra were estimated directly from the spectra using the CO_2 15-μm band. The brightness temperature at a given wavenumber was assigned to the atmospheric temperature at a specific pressure level. At smaller pressures than the pressure level assigned to the band center, an isothermal temperature profile was assumed. The pressure level assigned to a given wavenumber corresponds to the peak of the respective contribution function for a retrieved temperature profile from a dust-free IRIS spectrum. Although the contribution functions for wavenumbers that peak below 1 mbar are affected by the surface, the atmospheric temperature profiles estimated from the spectra are close to a log P extrapolation from the profile above 1 mbar. The redundancy of the 15-μm band was utilized by averaging the brightness temperatures for pressure levels associated with wavenumbers on each side of the band. Surface pressures were calculated from altitudes given in the USGS I-2160 1991 map series using a seasonal correction and a constant scale height of 10 km. These derived atmospheric temperature profiles are a key input for the inference of the IR dust extinction opacities from the 9- and 20-μm absorptions in the IRIS spectra. The depth of the 9-μm dust absorption is roughly proportional to the difference between the surface brightness temperature and the atmospheric temperature ~20 km above the surface, as well as the dust extinction opacity. Retrievals of the 9-μm dust extinction opacities are not particularly sensitive to the details of the modeled dust composition and size distribution. This relative

Table 1. Mariner 9 IRIS Spectra

Region	Latitude, Longitude	Local Time, hours	Emission Angle, deg	DAS	Orbit	L_S, deg	Surface Pressure, mbar	Surface Temperature, K
Argyre Planitia	55.0°S 50.7°W	13.95	52.8	2605524	28	301.1	5.5	253
Hellas Planitia	39.8°S 306.4°W	12.2	46.7	2209744	17	297.8	8.7	246
Solis Planum	34.2°S 109.1°W	13.3	61.3	2245531	18	298.1	2.9	273
Southern Chryse Planitia	4.3°N 25.2°W	10.3	73.7	3032932	40	304.7	6.7	238

DAS, Data Automation Subsystem time (incremented 1 count per 1.2 seconds time)

insensitivity is due to the fact that absorption dominates over scattering within the dust 9-μm extinction over a wide range of dust parameters. As a result, the Viking IRTM 9-μm dust opacities derived by *Martin* [1986] on the basis of the Toon et al. model of Mars atmospheric dust are not overly dependent on this model, although their interpretation in terms of dust mass would be. More important, the detailed fit of the wavelength dependence of the 9-μm dust absorption and the relative depths of the visible, 9-μm, and 20-μm dust absorptions are very sensitive to the dust composition and particle size distribution.

Besides definition of the atmospheric temperature profile, it is also necessary to constrain the IR brightness temperature of the surface of Mars. It is possible to use the 5- to 7-μm brightness temperatures from the IRIS spectra to constrain the 5- to 7-μm surface brightness temperature, since the dust absorption is smaller (but not negligible) at these wavelengths. At longer IR wavelengths the atmospheric dust absorption is more significant. *Toon et al.* [1977] assumed a wavelength independent surface emissivity/brightness temperature, based on comparisons to an IRIS spectrum (orbit 206) taken well into the decay of the 1971-1972 dust storm. A recent analysis of IRIS emission angle sequences for orbits 174 and 176 also indicated no gradient in surface brightness temperatures between 5 and 30 μm [*McMillan et al.*, 1993].

In the following analysis, we consider the effects of a wavelength-dependent surface emissivity on the surface brightness temperatures, based on Viking IRTM observations of 7- to 11-μm surface brightness variations [*Christensen*, 1982, 1986] and IRIS observational limits on the 5- to 30-μm variation in surface emissivity. *Christensen* [1982] found a significant variation in the IR emissivity of lower albedo (albedo<0.28) regions on Mars versus a wavelength-independent IR surface emissivity for the brightest (albedo>0.28) regions on Mars. The maximum 7- to 20-μm variation in the IR emissivity of the Mars surface (Syrtis Major, albedo=0.10) is 0.98 to 0.92, corresponding to a decrease in surface brightness temperature of roughly 8 K between wavelengths of 7 and 20 μm. In fact, Christensen found that most of the decrease in the surface emissivity occurred at wavelengths between 7 and 11 μm, which would lead to considerable ambiguity in separating the 9-μm atmospheric dust absorption from the surface emissivity effects. Since the IRTM analysis of *Christensen* [1982] is broad band in nature and indicates considerable variability and/or uncertainty in the detailed wavelength dependence of the surface IR emissivity, we model only the effects of a simple linear decrease in the surface brightness temperature between 5 and 30 μm. The implications of 8- to 11-μm spectral structure in the surface emissivity are considered in the discussion of our modeling results. A separate wavelength dependence in the surface brightness temperatures of Mars results from differences in the thermal properties of the dust and rock components of the Mars surface. *Christensen* [1986] analyzed the nighttime differences in the 7- and 20-μm IRTM brightness temperatures to map rock abundance on the Mars surface. However, this effect is nearly zero during the daytime, when the above IRIS spectra were taken, and does not influence the current analysis.

We model three separate particle size distributions for the Mars atmospheric dust, all based on the modified gamma distribution (e.g., *Hansen and Travis*, 1974) for direct comparison to the Toon et al. model. The modeled size distributions, presented in Figure 1b and Table 2, include (size 1) the Toon et al. model (r_{mode}=0.40 microns), (size 2) a size distribution with roughly two times smaller particles (r_{mode}=0.17 microns) but the same size distribution width as the Toon et al. model (varreff~0.4 μm), and (size 3) a size distribution with much smaller particles (r_{mode}=0.02 microns) combined with a much broader size distribution (r_{eff} variance~0.8 μm) than the Toon et al. model. At visible wavelengths, the effective (cross-section weighted, $r_{cw\mu}$) radii for these three size distributions are 2.7, 1.2, and 1.8 μm, respectively. As demonstrated below, the particle size distribution is the primary influence on the relative variation in the dust optical depth at visible, 9-μm, and 30-μm wavelengths. The Toon et al. size distribution was derived primarily on the basis of the 7- to 30-μm absorption properties of the atmospheric dust. The second particle size distribution allows a better fit to the relative visible/9-μm variation in dust opacity and the Phobos near-IR observations of the dust extinction opacities, but requires a substantial 5- to 30-μm variation in surface emissivity to fit the IRIS brightness temperatures at 18-30 μm. The third particle size distribution is chosen to provide agreement with the optical depth variation of the dust from visible, near-IR, 9-μm, and 30-μm observations, for the case of minimal wavelength dependence of surface emissivity between 5 and 30 μm. As seen in Figure 1b, this particle size distribution exhibits a great deal of very fine dust (r < 0.1 μm), yet provides a significant amount of 1- to 10-μm sized particles due to the broadness of the size distribution.

Table 2. Modified Gamma Size Distribution Parameters

	a	b	c	r_{mode}	$r_{cw\mu}$	r_{eff} Variance
Size 1	2	6.32	0.5	0.40	2.7	0.38
Size 2	2	9.6	0.5	0.17	1.2	0.42
Size 3	1	12.0	0.3	0.014	1.8	0.79

Argyre Planitia

The bottom half of Figure 2a presents the brightness temperature calculations for the standard model of Toon et al. (montmorillonite 219b, size 1, $r_{cw\mu}$ = 2.7 μm) as compared to

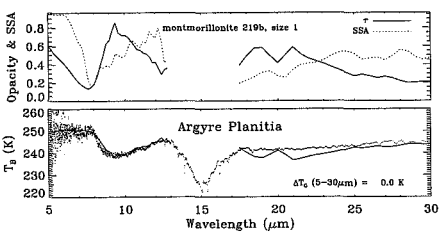


Figure 2a. In the bottom panel, model brightness temperature spectra (T_B (K), solid lines) for various compositions and dust particle size distributions are compared to a Mariner 9 IRIS spectrum (data points) of Mars atmospheric dust during the 1971-1972 Mars global dust storm, over the Argyre Planitia region (see Table 1). The top panel presents the wavelength dependence of the modeled opacity (solid line) and single-scattering albedo (dashed line) of the modeled dust. The observed spectrum indicates the strong atmospheric 15-μm CO_2 absorption between 13 and 17 μm, which is not modeled. Modeled dust features centered at 9 μm and broadly present beyond 18 μm, reflect absorption by cold atmospheric dust above warmer blackbody radiation by the Mars surface. The atmospheric dust model presented is the Toon et al. model of montmorillonite 219b with particle size distribution 1.

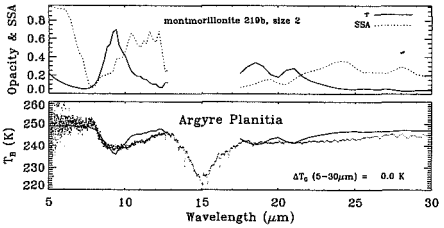


Figure 2b. Same as Figure 2a, but for montmorillonite 219b with the smaller particle size distribution 2.

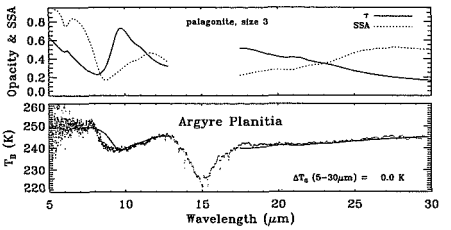


Figure 2d. Same as Figure 2a, but for a Hawaiian palagonite composition with the very broad particle size distribution 3.

the observed IRIS brightness temperature spectrum (data points) of the Argyre Planitia region (see Table 1). The surface emissivity was assumed independent of wavelength for this modeled spectrum. The top half of Figure 2a indicates the wavelength-dependent single-scattering albedo (SSA, dotted line) and extinction opacity (solid line) for the modeled dust properties. In this case, the peak 9-μm dust opacity is ~ 0.8. The fit of the Toon et al. dust model to the IRIS spectrum appears very good in the 9-μm region but much less satisfactory in the 20-μm region, consistent with the original analysis of *Toon et al.* [1977]. In Figure 2b, we present the fit of the montmorillonite 219b optical constants for a smaller particle size distribution (size 2, $r_{cw\mu}$ = 1.2 μm), indicating the sensitivity of the detailed shape of 9-μm absorption band and the 18- to 30-μm dust absorption to the dust particle sizes. Notice that the larger dust particle sizes (size 1, $r_{cw\mu}$ = 2.7 μm) better fit the IRIS spectra beyond 15 μm with a wavelength-independent surface emissivity, whereas the smaller dust particle sizes (size 2, $r_{cw\mu}$ = 1.2 μm) would require a significant (~5 K) decrease in the modeled surface brightness temperature between 5 and 30 μm to fit the IRIS spectra in the same region. This conclusion holds true for the palagonite as well as the montmorillonite composition. The palagonite model, presented in Figure 2c for the smaller dust particle sizes (palagonite, size 2, $r_{cw\mu}$ = 1.2 μm) and wavelength-dependent (dashed line) and wavelength-independent (solid line) models of the surface brightness temperature, provides a reasonable fit the observed 9- to 12-μm dust absorption, but does not present sufficient absorption in the 8- to 9-μm wavelength region. On the other hand, this model leads to an absence of structured absorption in the 20-μm wavelength region, which is much more consistent with the IRIS spectral observations in this region. We present results for the third, broad particle size distribution with a palagonite composition in Figure 2d (palagonite, size 3, $r_{cw\mu}$ = 1.8 μm), which incorporates a wavelength-independent surface brightness temperature. The model fit to the IRIS observations is very good for the 9- to 12-, 20-, and 30-μm regions, but still does not provide sufficient absorption in the 8- to 9-μm region. For an additional compositional comparison, we present a model calculation for the basalt optical constants provided by *Toon et al.* [1977]. Figure 2e indicates a best fit for these basalt optical constants and the Toon et al. size distribution (size 1). This model does not fit either side of the 9-μm absorption very well, but does lead to much subdued absorption structure at 20 μm. All five of the model cases presented in Figures 2a-2e indicate 9-μm extinction opacities of 0.7-0.8.

Hellas Planitia

Both the Argyre and Hellas Planitia spectra were observed at moderate emission angles (53° and 47°, respectively). The Hellas spectrum (Figure 3) was taken over the deeper western portion of the Hellas basin (Table 1), but still exhibits fairly moderate 9-μm dust absorption relative to the two other southern regions presented (i.e., $\tau_{9\mu m}$~ 0.6 versus 0.8 and 1.1). The three basic models of dust are compared to the Hellas observation in Figures 3a-3c. These include the standard Toon et al. model (Figure 3a, montmorillonite 219b, size 1), the palagonite model with a factor-of-two reduction in particle sizes (Figure 3b, size 2), and the palagonite model with a very broad particle size distribution (Figure 3c, size 3). These comparisons are very similar to those presented for the Argyre Planitia region. The standard Toon et al. model provides the best fit to the observed 9-μm absorption, but

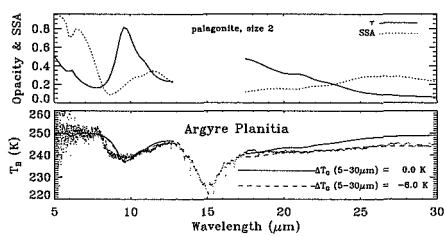


Figure 2c. Same as Figure 2a, but for a Hawaiian palagonite composition with particle size distribution 2, including a case in which the surface brightness temperature decreases by 6 K between 5 and 30 μm (dashed line).

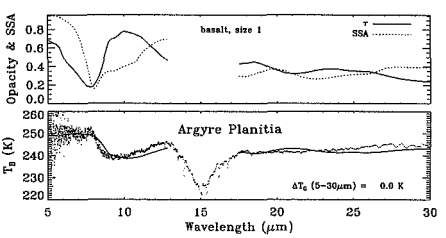


Figure 2e. Same as Figure 2a, but for a basalt composition with the Toon et al. particle size distribution 1.

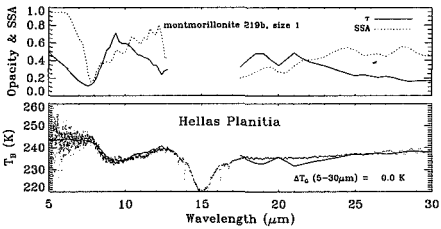


Figure 3a. Same as Figure 2 for an IRIS spectrum over Hellas Planitia (see Table 1). The modeled dust case is the Toon et al. model of montmorillonite 219b with particle size distribution.

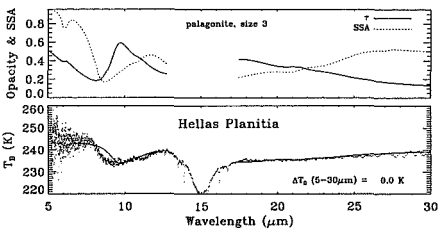


Figure 3c. Same as Figure 3a, but for a Hawaiian palagonite composition with the very broad particle size distribution 3.

overestimates the 20-μm feature with distinctive structure that is not present in the observations. It also leads to an incorrect ratio of the visible/9-μm dust extinction opacity, as described below. The palagonite model with size distribution 2 (Figure 3b) does not match the detailed shape of the observed 9-μm absorption. It also does not provide sufficient absorption in the 20- to 30-μm region, such that it would require a significant decrease (~6 K) in the surface emissivity in this wavelength region to fit the observed 20- to 30-μm brightness temperatures. The palagonite model with the very broad size distribution 3 (Figure 3c) provides the best overall match to the IRIS spectrum, although it still does not exhibit sufficient absorption in the 8- to 9-μm region.

Solis Planum

We present the three basic models of Mars atmospheric dust for comparison to the Solis Planum spectrum in Figures 4 a-4c. These models include the Toon et al. case (Figure 4a, montmorillonite 219b, size 1, $r_{cw\mu}$ = 2.7-μm, wavelength-independent surface emissivity), the palagonite model with a narrow size distribution (Figure 4b, palagonite, size 2, $r_{cw\mu}$ = 1.2 μm, wavelength-dependent and independent surface emissivities) , and the palagonite model with a very broad size distribution (Figure 4c, palagonite, size 3, $r_{cw\mu}$ = 1.8 μm, wavelength-independent surface emissivity). The comparisons of Figures 4a-4c suggest similar conclusions to those indicated for the Argyre Planitia and Hellas Planitia model-data comparisons, although all of the model fits for the Solis Planum spectrum are somewhat poorer in the 7-9 and 18-30 μm regions. The 20-μm absorption of the montmorillonite model is clearly out of line with the observed spectrum. The palagonite models do not produce sufficient absorption in the 8- to 9-μm region, but do provide a better match to the observed 20 μm region than the montmorillonite composition. Palagonite with the small, narrow particle size distribution (Figure 4b, size 2, $r_{cw\mu}$ = 1.2 μm) fits the observed spectrum in the 18- to 30-μm region if the surface brightness temperature is decreased by ~15 K between 5 and 30 μm. This is a substantially larger wavelength gradient in surface emissivity than required for the Argyre Planitia and Hellas Planitia observations. Although it is possible that the higher emission angle of observation for the Solis Planum spectrum (61°) may lead to a more extreme wavelength dependence for surface emissivity, the requirement of a 15 K gradient in surface brightness temperature between 5 and 30 μm appears unlikely in light of the analysis of IRIS emission pair sequences by *McMillan et al.* [1993]. Palagonite with the broad particle size distribution 3 (Figure 4c) provides a reasonable approximation to the observed spectrum in the 18- to 30-μm region without requiring a substantial wavelength dependence for the surface emissivity. We also present a basalt model calculation for the Solis Planum observation, since this spectrum indicates the largest 9-μm dust absorption of the spectra we analyzed. Figure 4d displays the resulting best fit model spectrum with the Toon et al. size distribution (size 1) and basalt optical constants, which does not match the observed absorption structure of the 9- or 20-μm regions well, but does exhibit appropriate levels of absorption at 30 μm due to the larger particle sizes employed.

Another distinction in the Solis Planum IRIS spectrum is a relative maximum at 7- to 8-μm wavelength, which is not clearly apparent in the spectra for the other regions studied. A similar feature is observed in several of the IRIS spectra analyzed by *Toon et al.* [1977)] Model calculations with quartz showed

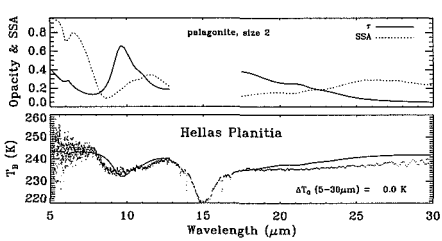


Figure 3b. Same as Figure 3a, but for a Hawaiian palagonite composition with particle size distribution 2.

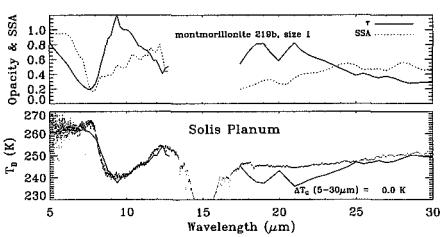


Figure 4a. Same as Figure 2 for an IRIS spectrum over Solis Planum (see Table 1). The model of atmospheric dust presented is the Toon et al. model of montmorillonite 219b with particle size distribution 1.

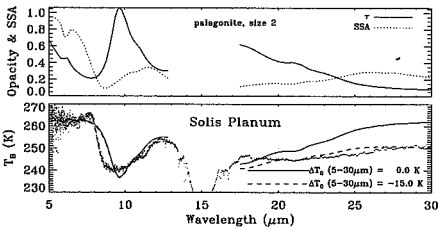


Figure 4b. Same as Figure 4a, but for a Hawaiian palagonite composition with particle size distribution 2, including a case in which the surface brightness temperature decreases by 15 K between 5 and 30 μm (dashed line).

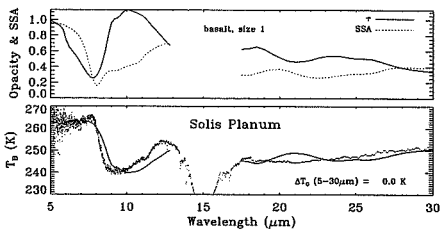


Figure 4d. Same as Figure 4a, but for a basalt composition with the Toon et al. particle size distribution 1.

somewhat similar structure, and there are weak CO_2 bands in the 7.4-μm region which may contribute to this feature [*Toon et al.*, 1977]. Notice also that the palagonite models lead to weak absorptions just beyond 6 μm which are apparently related to the presence of water in the palagonite sample [*Roush et al.*, 1991].

Southern Chryse Planitia

The southern Chryse observation, taken near the crater Galilaei, represents the highest emission angle (74°) and lowest dust opacity ($\tau_{9\mu m} \sim 0.5$) of the four spectra presented. The same three basic models of dust are presented against the observed spectrum in Figures 5a-5c. The results of these comparisons are very similar to those for the Solis Planum region in particular. The standard Toon et al. model (Figure 5a, montmorillonite 219b, size 1, $r_{cw\mu}$ = 2.7 μm) leads to deep, structured 20-μm absorption that is not observed and a very low ratio for the visible/9-μm extinction opacity ratio. The palagonite model for the smaller, narrow size distribution (Figure 5b, size 2, $r_{cw\mu}$ = 1.2 μm) does not fit the observed 20- to 30-μm absorption, unless an unreasonably large decrease in the surface emissivity is assumed for this wavelength region. The palagonite model with the broad particle size distribution (Figure 5c, size 3) provides the best relative fit among the visible, 9-μm, and 30-μm absorption spectrum of the dust, but neither palagonite model provides sufficient absorption in the 8-to 9-μm region.

All of the dust models we consider fail to match the observations in some way. A smaller particle size distribution than the Toon et al. model is required to match the observed visible/9-μm opacity ratio and the Phobos near-IR observations, as discussed below. Simply decreasing the mode radius of the Toon et al. particle size distribution by a factor of 2 (size 2) helps to match these observations, but appears to under predict the 20- to 30-μm extinction opacity of the dust. A very broad particle size distribution with a very small modal radius for the dust particle sizes (size 3) may provide the best fit to the observed variation in the dust extinction opacity between visible and 30-μm wavelengths. A dust composition of montmorillonite provides the best fit to the shape of the observed 9-μm absorption, but is not consistent with the shape of the observed 20-μm absorption of the dust. The modeled palagonite composition can provide a good fit to the observed 20-μm and 9- to 12-μm absorption of the dust, but lacks sufficient absorption in the 8- to 9-μm region. Palagonite also exhibits UV and visible absorption roughly consistent with observations of Mars atmospheric dust (see below), while montmorillonite does not. Several effects may contribute to the poor match of the modeled palagonite to the observed 8- to 9-μm absorption of the dust. The composition of the modeled palagonite is unlikely to be the same as for the Mars dust. Terrestrial palagonites exhibit a wide range of compositions, depending on the composition of the basalt which is weathered to produce the palagonite. The frequency of peak 8- to 9-μm absorption in igneous silicates is known to be sensitive to the SiO_2 content in particular [*Hanel et al.*, 1972]. The SiO_2 content of the modeled palagonite (31%) is much lower than the SiO_2 content of montmorillonite 219b (63%), which is quite a bit higher than the Mars soil SiO_2 content (42-43%) determined from the Viking Lander X ray fluorescence experiment [*Clark et al.*, 1982]. *Singer* [1982] analyzed a terrestrial palagonite sample containing ~48% SiO_2.

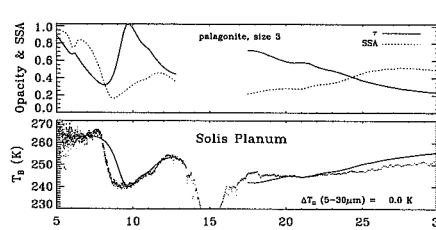


Figure 4c. Same as Figure 4a, but for a Hawaiian palagonite composition with the very broad particle size distribution.

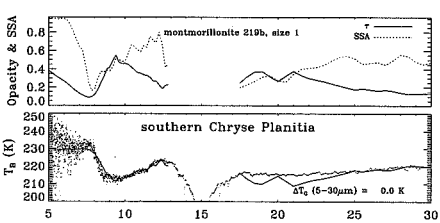


Figure 5a. Same as Figure 2 for an IRIS spectrum over southern Chryse Planitia (see Table 1). Modeled dust case is the Toon et al. model of montmorillonite 219b with particle size distribution 1.

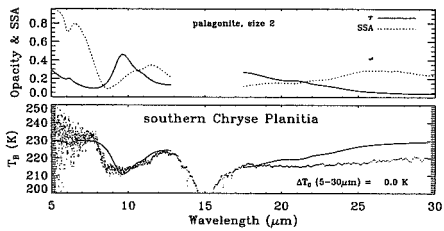


Figure 5b. Same as Figure 5a, but for a Hawaiian palagonite composition with particle size distribution 2.

Furthermore, the method of determining optical constants for a heterogeneous material such as palagonite involves a certain degree of modeling (see *Roush et al.*, 1991], which may introduce uncertainties in the model-observation comparisons. The nonspherical shape of Mars dust may also affect the detailed modelling of the strong 9-μm absorption of the dust. *West et al.* [1989] have shown that the frequencies of IR emission bands of ammonia haze are affected by the aerosol particle shapes. In a study of terrestrial cirrus clouds, *Mitchell and Arnot* [1994] have demonstrated the large effect that different ice crystal habits can have on infrared cloud optical depth, which varies as a function of wavelength. Finally, we do not know the detailed nature of the Mars surface emissivity within the 9-μm dust absorption band. Since the surface exhibits a comparable composition to the dust at some level, it is likely the surface possesses 9-μm structure in its emissivity. This effect may be small, but we do not have sufficiently clear atmosphere observations to determine potential brightness temperature variations of the Mars surface within the 9-μm absorption band. Analyses of Mariner 6 and 7 IRS observations, taken during relatively clear Mars atmospheric conditions, may help to address this issue in the near future [T. Z. Martin, Mariner 6/7 infrared spectrometer: Data set restoration, submitted to *Journal of Geophysical Research*, 1991).

Viking IRTM 9-μm EPF Modeling and the Visible/9-μm Extinction Opacity Ratio

The visible/9-μm ratio of extinction opacities is an important diagnostic of the ratio of submicron-to-micron particle sizes for Mars dust [e.g., *Toon et al.*, 1977]. *Martin* [1986] analyzed Viking IRTM 9-μm radiances to obtain a time sequence of the dust 9-μm opacity over the two Viking Lander sites; and determined a visible/9-μm extinction opacity ratio of 2.5, based on a comparison of his 9-μm opacities with the lander solar extinction observations at visible wavelengths [*Pollack et al.*, 1977, 1979]. As noted by *Zurek* [1982] and *Martin* [1986], the visible/9-μm opacity ratio implied by the *Toon et al.* [1977] model of Mars dust (~1) is substantially smaller than determined by the observational analysis of Martin (≥ 2). We provide a separate measurement of this key opacity ratio based on analysis of the Viking EPF sequences in the IRTM visible (solar band) and 9-μm channels. The solar band dust opacities were derived from a multiple scattering analysis of the EPF multiple emission angle observations of roughly 20 regions on Mars, presented by *Clancy and Lee* [1991]. We employ the IR radiative transfer code, described above, to retrieve the 9-μm dust opacities for 13 of these regions. Because the IR and solar band measurements from the IRTM experiment are obtained simultaneously, the solar band and 9-μm dust opacities retrieved from these EPF sequences are uniquely contemporaneous and co-located. The 9-μm EPF sequences also provide a somewhat better measurement of the 9-μm dust opacity than the single emission angle 9-μm observations analyzed by *Martin* [1986], in that the emission angle dependence and the absolute value of the 9-μm brightness temperature are used to constrain the 9-μm dust opacity. There are, of course, far fewer EPF sequences than the standard IRTM mapping 9-μm observations analyzed by *Martin* [1986].

We employ the same 9-μm dust opacity inversion algorithm described by *Martin* [1986]. The surface brightness temperature over the IRTM 9-μm bandpass is assumed constant (no wavelength or emission angle dependence) and is based on the 7-μm brightness temperature from the IRTM 7-μm channel brightness. The atmospheric temperature profile is constructed from the IRTM 7- and 15-μm channel brightness temperatures, with the assumption of a 10-20 K offset between the surface and lower atmosphere temperatures. This temperature offset is chosen according to the local time of the individual EPF sequences modeled [e.g., see *Martin*, 1986]. The radiative scattering/absorption calculations employ the IRTM bandpass weighting and model dust parameters for both montmorillonite 219b and palagonite with size distributions 1 and 2, respectively. In Figures 6a and 6b we present model fits to 9-μm EPF

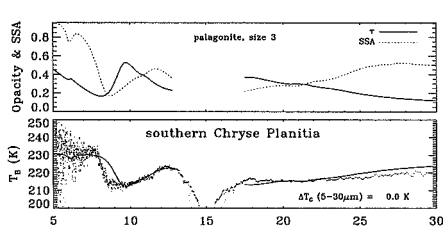


Figure 5c. Same as Figure 5a, but for a Hawaiian palagonite composition with the very broad particle size distribution 3.

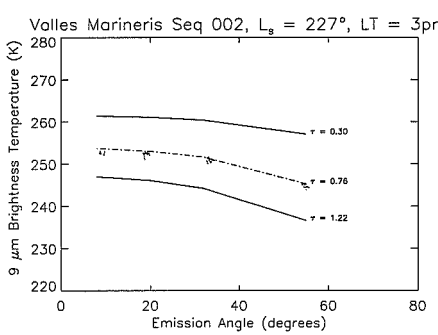


Figure 6a. Emission-phase-function (EPF) sequence observed with the Viking IRTM 9-μm channel over Valles Marineris during the 1977 global dust storm. The smooth model curves represent predicted emission angle dependences for the 9-μm brightness temperature for three atmospheric 9-μm dust opacities, based on the Toon et al. dust model.

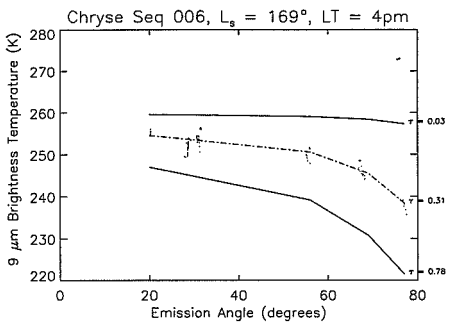


Figure 6b. An EPF 9-μm sequence observed over Chryse Planitia is compared to three 9-μm opacities, based on the Toon et al. dust model.

sequences over Valles Marineris (9-10°S, 71-72°W) and Chryse Planitia (23-24°N, 48-49°W). The observations are indicated by individual points; model results are presented as lines for three values of the zenith 9-μm optical depth, for the montmorillonite 219b, size 1 dust properties. In Figure 6c, we show the fit of the palagonite, size 2 dust properties to the same Chryse Planitia observations of Figure 6b. Both of these models indicate best fit values of ~0.3 for the 9-μm dust optical depth for the Chryse Planitia EPF sequence. Similarly, either dust model leads to a 9-μm dust opacity of ~0.8 for the Valles Marineris EPF sequence.

Figure 7 presents the visible/9-μm opacity ratio versus the visible opacity, based upon the 9-μm dust extinction opacities derived from 13 such EPF sequences and the solar band extinction opacities derived for the same 13 sequences [from *Clancy and Lee,* 1991]. For visible opacities less than 1, the visible/9-μm opacity ratio is consistent with a value of 2. For visible opacities greater than 1, there appears to be an increase in the visible/9-μm opacity ratio to values as high as 3-4. However, the uncertainties in these ratios are very large for the higher dust opacities such that this apparent trend of the visible/9-μm opacity ratio with dust opacity is not reliably defined. The largest source of uncertainty in the opacity ratios of Figure 7 result from the

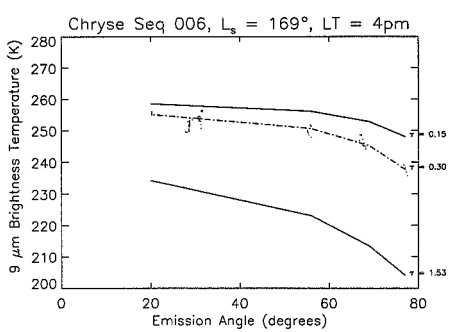


Figure 6c. The same EPF sequence of Figure 6b is compared to three 9-μm opacities, based on a Hawaiian palagonite composition and size distribution 2 model for the dust.

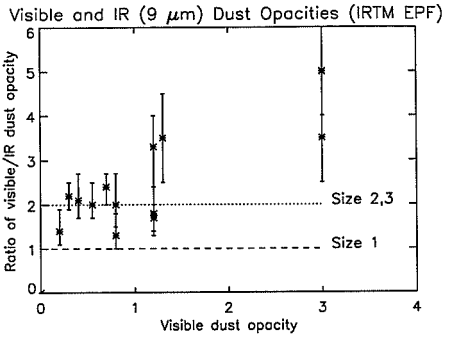


Figure 7. The ratio of visible and 9-μm dust opacities obtained from solar band [*Clancy and Lee,* 1991] and 9-μm Viking IRTM EPF sequences (such as in Figure 6) are plotted versus the visible (solar band) dust opacity for 13 analyzed EPF sequences. The predicted ratio of visible/9-μm dust opacity is presented for the Toon et al. dust model (dashed line, montmorillonite 219b with size distribution 1) and the Hawaiian palagonite model with smaller dust particle sizes (dotted line, size distributions 2, 3)

model dependent uncertainties in the visible dust opacities, which are particularly dependent on the dust single-scattering albedo for the higher dust opacities. Uncertainties in the 9-μm dust opacities are contributed primarily by the uncertain surface brightness and atmospheric temperatures employed in the modelling. *Martin* [1986] estimated uncertainties in 9-μm dust opacities, using this technique, which are less than 15% for the range of 7- and 9-μm IRTM brightness temperatures employed in this current analysis. The difference in the 9-μm opacity between the two models of the dust properties is small, for the reasons described above. Also indicated on Figure 7 are the predicted visible/9-μm extinction opacity ratios for the three basic models of Mars dust considered in this analysis. A ratio of unity results from the Toon et al. model (montmorillonite 219b, size 1, $r_{cw\mu}$ = 2.7 μm). The two palagonite models (size 2, $r_{cw\mu}$ = 1. 2 μm; and size 3, $r_{cw\mu}$ = 1.8 μm) both lead to visible/9-μm opacity ratios of 2. Higher ratios result directly from smaller average particle sizes for the dust. Although dust composition plays a lesser role in determining the visible/9-μm opacity ratio, montmorillonite actually requires substantially smaller particle sizes than the modeled palagonite to obtain a visible/9-μm opacity ratio of ~2. For the same particle size distribution width, montmorillonite 219b would require $r_{cw\mu}$ = 0.7 μm to exhibit a visible/9-μm opacity ratio of 2.

The Phobos Infrared Spectroscopic Modeling Dust Particle Size Determination

Another important measurement regarding the dust particle sizes are the Phobos near-IR (1.9 and 3.7 μm) occultation measurements of dust extinction profiles [e.g., *Krasnopolsky et al.,* 1989]. A more recent analysis of these data in terms of the dust particle sizes is given by *Chassefiere et al.* [1992], who find that the effective radius ($r_{cw\mu}$) of the dust increased from ~0.9 to 1.3 μm between the altitudes of 25 and 15 km. These are

significant measurements because all of the particle sizes considered for Mars dust place the effective radii of the dust at sizes comparable to near IR wavelengths. Hence, an optimum measurement of the wavelength dependence in the dust scattering extinction efficiency would be at wavelengths of 2-4 μm. The results of *Chassefiere et al.* [1992] are generally consistent with the smaller particle sizes we argue for on the basis of the visible/9-μm dust opacity ratios. The width of the particle size distribution from their analysis is quite similar to the size distribution 2 in Figure 1b, which we argue is not consistent with the 20- to 30-μm absorption of the dust indicated in the IRIS spectra. However, Chassefiere et al. indicate very large uncertainties (±50 %) in their determination of the particle size distribution width.

Dust UV/Visible Absorption

Although the detailed variation of the atmospheric dust absorption at visible and ultraviolet wavelengths is not well measured, it is known that the dust exhibits single-scattering albedos of order 0.4-0.6 at 3000 Å [*Pang and Ajello et al.*, 1977] and 0.86 [*Pollack et al.*, 1977] to 0.92 [*Clancy and Lee*, 1991] in the solar band average (~6000-10000 Å). A reanalysis of the multifilter photometry of Mars dust from the Viking landers is provided in this special issue by *Pollack et al.* [this issue]. The ultraviolet/visible absorption of the dust has led to arguments for anatase (TiO_2 [*Pang and Ajello*, 1977]) and magnetite [*Pollack et al.*, 1977, 1979] components in the Mars atmospheric dust. In fact, the elemental composition of the modeled palagonite is consistent with significant, nanocrystalline components of titanium and iron oxides (4.1 % TiO_2, 14.6 % Fe_2O_3, 1 % FeO [*Roush et al.*, 1991]). We compute ultraviolet (3000 Å) and visible (6000 Å) single-scattering albedos for palagonite dust, based on 3000-10000 Å optical constants of another Mauna Kea palagonite taken from *Clark et al.* [1990]. The sharp rise in the imaginary index of this palagonite over the 0.3- to 0.7-μm wavelength region is shown in Figure 1a. Ultraviolet single-scattering albedos of 0.56 and 0.57 are derived for particle size distributions 2 and 3 (Figure 1b), respectively. Visible single-scattering albedos of 0.94, and 0.93 are computed for palagonite with the same particle size distributions.

These values are reasonably consistent with the range of observed single-scattering albedos for Mars atmospheric dust. Of course, no conclusions regarding the detailed composition of Mars dust can be drawn for these very coarse comparisons of single-scattering albedos. They merely point out the plausibility of a "palagonite-like" composition for Mars atmospheric dust in terms of their broad UV-visible absorption properties. It is very unlikely that the composition of the modeled palagonite applies to Mars dust. But it appears possible that a "single component" model for Mars atmospheric dust, based on a heterogeneous, nanocrystalline or amorphous weathering product of basalt, may account for the complete set of ultraviolet-to-30 μm observations of Mars atmospheric dust.

Conclusions

1. A combined analysis of the Viking EPF visible and 9-micron sequences indicates a visible-to-IR dust extinction ratio of ≥ 2, consistent with the earlier comparison by *Martin* [1986] of 9-μm IRTM dust opacity determinations and the Viking lander extinction opacity observations [*Pollack et al.*, 1977]. The observed ratio of ≥ 2 is significantly different from the predicted ratio of 1.0 from the current model of Mars atmospheric dust.

2. The ratios of visible and 9-μm dust extinction opacities for a palagonite or montmorillonite composition are much closer to the observed ratio for smaller particle sizes of the dust (r_{mode}=0.17 versus 0.40 μm). Palagonite dust for two size distribution models presented (size 2 and 3, Figure 2) yields a visible-to-IR extinction ratio of 2, consistent with the measurements of this ratio, as well as the Phobos near-IR extinction measurements of the dust particle sizes [*Chassefiere et al.*, 1992]. A montmorillonite composition would require smaller particle sizes ($r_{cw\mu}$=0.7 μm, r_{mode}=0.10 μm) than palagonite to obtain the same visible-to-IR opacity ratio of 2, such that it might not be consistent with the Phobos particle size determination for the dust, and would no longer fit the IRIS 9-μm or 18- to 30-μm absorptions well.

3. Palagonite dust can also lead to a much improved fit to the IRIS observations near 20-μm wavelengths, since the sample of palagonite modeled displays no distinct 20-μm absorption. However, the same sample of palagonite does not exhibit sufficiently strong absorption in the 8.5-μm wavelength region relative to the IRIS measurements of this absorption. This difference is significant within the IRIS measurement uncertainties, but may be accommodated by a different Mars palagonite composition and/or the uncertainties in detailed modelling of the 9-μm dust absorption.

4. Palagonite dust with a $r_{cw\mu}$ of 1.2-1.8 μm (r_{mode} < 0.2 microns) leads to single-scattering albedos (w_o) which are roughly consistent with their observed values at ultraviolet (w_o~0.4 - 0.6) and (w_o~0.9-0.95) visible wavelengths. Hence, a single component model for Mars atmospheric dust may explain the existing observations of the dust absorption from 0.3 to 20 μm within the modeling and measurement uncertainties.

5. Comparison of palagonite and montmorillonite dust with the same size distribution gives similar 9- to 30-μm opacities, apart from the deep doublet structure at 20-μm and positive frequency shift of the 9-μm band for montmorillonite versus the modeled palagonite. However, the average 18- to 30-μm opacity is very different for $r_{cw\mu}$=2.7 versus 1.2 μm (r_{mode} = 0.40 versus 0.17 μm) particle sizes of either composition. The key uncertainty in differentiating size distributions 1 and 2 (which possess roughly the same size distribution width, r_{eff} variance~0.4 μm) on the basis of the 18- to 30-μm dust opacity is the degree to which the surface emissivity may vary from 5 to 30 μm. The large decreases (6-12 K) in surface brightness temperatures between wavelengths of 5 and 30 μm required by size distribution 2 seem unlikely, and indicate the possibility of a third model for the Mars atmospheric dust size distribution. In this case, a very broad particle size distribution (r_{eff} variance~0.8 μm) is modeled to provide the observed, relative variation of the dust opacity between visible, 9, and 30-μm wavelengths without the requirement of a substantial wavelength dependence for the surface emissivity. This particle size distribution leads to large numbers of very fine dust particles (r_{mode} < 0.02 microns) to provide the observed visible/9-μm opacity ratio and significant numbers of μm-sized particles ($r_{cw\mu}$=1.8 μm) to fit the 18- to 30-μm region of the IRIS spectral observations.

In summary, we submit that palagonite with a smaller and broader particle size distribution than that of the Toon et al. determination provides a much improved model to Mars atmospheric dust. Since palagonite is a common and fairly primitive weathering product of terrestrial basalts, palagonite could sensibly be considered as a major surface component on Mars [e.g., *Soderblom et al.*, 1978; *Singer*, 1982; *Roush et al.*, 1993]. The weakness or absence of reflectance signatures from clay minerals on the surface of Mars [*Bell and Crisp*, 1993; *Clark*, 1992] could also be consistent with a palagonite

composition for Mars dust if the conditions for basalt weathering on Mars produced sufficiently anhydrous forms of palagonite. On the other hand, such clay bands may be spectrally masked by another component [*Orenberg and Handy*, 1992] or the clay minerals, themselves, may be dehydroxylated [*Burns*, 1993]. Variations in palagonite composition and texture are large for terrestrial samples [*Evans et al.*, 1981; *Golden et al.*, 1993], which can lead to significant variations in their 7- to 20-μm absorption properties [*Roush et al.*, 1993]. It is noteworthy that the palagonite composition modeled in this work is very low in SiO_2 abundance (31%, *Roush et al.* [1991]) as compared to the SiO_2 content of the montmorillonite 219b sample (63%, *Toon et al.* [1977]). As noted by *Hanel et al.* [1972] and *Toon et al.* [1977], the 8- to 9-μm absorption properties of silicates are particularly sensitive to the SiO_2 abundance. In addition, the Mie scattering theory employed in the radiative transfer modelling may not accurately determine the 9-μm band position and shape [e.g., *West et al.*, 1989], and uncertainties in the wavelength-dependent emissivity of the surface may prevent an accurate fit. In any case, the key distinction in our new model versus the old model of Mars dust may follow less from differences in composition as from the low degree of crystallinity of palagonite versus montmorillonite, and the smaller particle sizes derived for the atmospheric dust. Both conditions tend to suppress the 20-μm absorption relative to the 9-μm absorption and increase the visible/IR dust opacity ratio. Such a change in the model of Mars atmospheric dust is clearly relevant to a similar controversy regarding the degree of crystallinity and clay mineralization in the soil covering the Mars bright regions [e.g., *Soderblom*, 1992; *Banin et al.*, 1992].

It is also worth emphasizing that the particle sizes of Mars dust are remarkably small for all three models. The modal radii of the atmospheric dust particles range from <.02 μm for our preferred, broad size distribution model to 0.40 μm for the *Toon et al.* [1977] model. The method of lifting Mars dust into the atmosphere is not entirely certain, but may involve saltation of larger particles [e.g., r ~100 μm *Greeley et al.*, 1980)]into a much finer dust component (e.g., $r < 2$ μm). The longer settling times of the fine dust particles would lead to their availability at the surface/atmosphere boundary, subject to many other processes such as surface bonding and coagulation. Our model of smaller particle sizes for Mars dust (r_{mode}=0.17 μm or < 0.02 μm) would increase the settling lifetimes for Mars atmospheric dust. The very fine particles (r< 0.02 μm) might require condensation processes for their removal from the atmosphere over annual timescales, or they may in fact remain as a constant background aerosol in the Mars atmosphere. Current dynamical models of Mars dust transport point to the importance of the dust size distribution on the transport and temporal evolution of atmospheric dust [*Murphy et al.*, 1993]. Smaller average dust particle sizes also lead to significant changes in the radiative forcing of the Mars atmospheric dust on atmospheric temperatures. The factors-of-2 or more decreases in particle sizes for our models versus the *Toon et al.* [1977] model of the dust are directly related to the factor-of-2 increase in the visible/IR opacity ratio between the same models. Changing this opacity ratio can dramatically influence dust heating/cooling rates for the Mars atmosphere. In fact, the thermal modeling analysis of *Haberle and Jakosky* [1991] requires a visible/IR opacity ratio of 2-2.5 to reproduce the sign of the temperature change within the lower Mars atmosphere during the initiation of 1977 global dust storm. In this case, the difference in the visible/IR opacity ratio between the dust particle size models leads to conditions of atmospheric cooling versus heating.

Acknowledgments. The authors are indebted to S. Murchie, J. Bishop, B. Clark, and R. Burns for helpful discussions regarding our current understanding of the Mars surface soil. We also benefited from very constructive reviews by J. Bell and J. Crisp. Assistance in preparation of the manuscript was provided by A. Alfaro, J. Lucero, and B. Dalton. Funding for this work was provided under a NASA Mars Surface and Atmosphere Through Time (MSATT) grant.

References

Banin, A., B. C. Clark, and H. Wänke, Surface chemistry and mineralogy, in *Mars*, edited by H. Kieffer, B. Jakosky, C. Snyder, and M. Matthews, pp. 595-625, University of Arizona Press, Tucson, 1992.

Bell, J., and D. Crisp, Groundbased imaging spectroscopy of Mars in the near-infrared: Preliminary results, *Icarus*, *104*, 2-19, 1993.

Bell, J. F., III, R. V. Morris, and J. B. Adams, Thermally altered palagonitic tephra: A spectral and process analog to the soils and dust of Mars, *J. Geophys. Res.*, *98*, 3373-3385, 1993.

Bishop, J., S. Murchie, S. Pratt, J. Mustard, and C. Pieters, The importance of environmental conditions in reflectance spectroscopy of laboratory analogs for Mars surface materials, *LPI Tech. Rep.*, *93-06*, 4-6, 1993.

Burns, R., Rates and mechanism of chemical weathering of ferromagnesian silicate minerals on Mars, *Geochim. Cosmochim. Acta*, *57*, 4555-4579, 1993.

Chassefiere, E., J. E. Blamont, V. A. Krasnopolsky, O. I. Korablev, S. K. Atreya, and R. A. West, Vertical structure and size distributions of Martian aerosols from solar occultation measurements, *Icarus*, *97*, 46-69, 1992.

Christensen, P. R., Martian dust mantling and surface composition: Interpretation of thermophysical properties, *J. Geophys. Res.*, *87*, 9985-9998, 1982.

Christensen, P. R., The spatial distribution of rocks on Mars, *Icarus*, *68*, 217-238, 1986.

Clancy, R. T. ., and S. W. Lee, A new look at dust and clouds in the Mars atmosphere: Analysis of emission-phase-function sequences from global Viking IRTM observations, *Icarus*, *93*, 135-158, 1991.

Clark, B. C., A. K. Baird, R. J. Weldon, D. M. Tsusaki, L. Schnabel, and M. P. Candelaria, Chemical composition of Martian fines, *J. Geophys. Res.*, *87*, 10,059-10,067, 1982.

Clark, R. N., Mars: Limits on mineralogy from reflectance spectroscopy, *Bull. Am. Astron. Soc.*, *24*, 986, 1992.

Clark, R. N., and T. L. Roush, Reflectance spectroscopy: Quantitative analysis techniques for remote sensing, *J. Geophys. Res.*, *89*, 6329-6340, 1984.

Clark, R. N., G. A. Swayze, R. B. Singer, and J. B. Pollack, High resolution reflectance spectra of Mars in the 2.3-μm region: Evidence for the mineral scapolite, *J. Geophys. Res.*, *95*, 14,463-14,480, 1990.

Evans, D. L., T. G. Farr, and J. B. Adams, Spectral reflectance of weathered terrestrial and martian surfaces, *Proc. Lunar. Planet. Sci. Conf. 12th*, 1473-1479, 1981.

Gladstone, G. R., J. W. Kaminski, R. Link, and J. C. McConnell, Cloud radiance modeling: Phase II, contract KM147-4-1041, Can. Dept. of the Environ., Ottawa, Ontario, 1984.

Golden, D. C., R. V. Morris, D. W. Ming, H. V. Lauer Jr., and S. R. Yang, Mineralogy of three slightly palagonitized basaltic tephra samples from the summit of Mauna Kea Hawaii, *J. Geophys. Res.*, *98*, 3401-3411, 1993.

Greeley, R., R. N. Leach, B. R. White, J. D. Iversen, and J. B. Pollack, Threshold wind speeds for sand on Mars: Wind tunnel simulations, *Geophys. Res. Lett.*, *7*, 121-124, 1980.

Haberle, R. M., and B. J. Jakosky, Atmospheric effects on the remote determination of thermal inertia on Mars, *Icarus*, *90*, 187-204, 1991.

Hanel, R., et al., Investigation of the Martian environment by infrared spectroscopy on Mariner 9, *Icarus*, *17*, 423-442, 1972.

Hansen, J. E., and L. D. Travis, Light scattering in planetary atmospheres, *Space Sci. Rev.*, *16*, 527-610, 1974.

Hapke, B., *Theory of reflectance and emittance spectroscopy*, 455 pp., Cambridge University. Press, New York., 1993

Hunt, G. E., Thermal infrared properties of the Martian atmosphere, 4,

Predictions of the presence of dust and ice clouds from Viking IRTM spectral measurements, *J. Geophys. Res.*, *84*, 2865-2874, 1979.

Hunt, G. E., L. M. Logan, and J. W. Salisbury, Mars: Components of infrared spectra and the composition of the dust cloud, *Icarus*, *18*, 459-469, 1973.

Krasnopolsky, V. A., V. I. Moroz, A. A. Krysko, O. I. Korablev, V. S. Zhegulev, A. V. Grigoriev, A. Yu. Tkachuk, V. A. Parshev, J. E. Blamont, and J. P. Goutail, Solar occultation sounding of the Martian atmosphere by the Phobos spacecraft, *Nature*, *341*, 603-604, 1989.

Martin, T. Z., Thermal infrared opacity of the Mars atmosphere, *Icarus*, *66*, 2-21, 1986.

McMillan, W. W., J. C. Pearl, and B. J. Conrath, Mariner 9 revisited: Mars Observer or bust, *Bull. Am. Astron. Soc.*, *25*, 1061, 1993.

Mitchell, D. L., and W. P. Arnott, A model predicting the evolution of ice particle size spectra and radiative properties of cirrus clouds. Part I. Microphysics, *J. Atmos. Sci.*, *51*, 817-832, 1994

Morris, R. V., J. L. Gooding, H. V. Lauer Jr., and R. B. Singer, Origins of the Mars-like spectral and magnetic properties of a Hawaiian palagonitic soil, *J. Geophys. Res.*, *95*, 14,427-14,435, 1990.

Murchie, S., J. Mustard, J. Bishop, J. Head, C. Pieters, and S. Erard, Spatial variations in the spectral properties of bright regions on Mars, *Icarus*, *105*, 454-468, 1993.

Murphy, J.R., et al., Martian global dust storms: Zonally symmetric numerical simulations including size dependent particle transport, *J. Geophys. Res.*, *98*, 3197-3220, 1993.

Orenberg, J., and J. Handy, Reflectance spectroscopy of palagonite and iron-rich montmorillonite clay mixtures: Implications for the surface composition of Mars, *Icarus*, *96*, 219-225, 1992.

Pang, K., and J. M. Ajello, Complex refractive index of Martian dust: Wavelength dependence and composition, *Icarus*, *30*, 63-74, 1977.

Pang, K., and C. W. Hord, Mariner 9 ultraviolet spectrometer experiment: 1971 Mars' dust storm, *Icarus*, *18*, 481-488, 1973.

Pollack, J., D. Colburn, R. Kahn, J. Hunter, W. Van Camp, C. Carlston, and M. Wolfe, Properties of aerosols in the Mars atmosphere, as inferred from Viking Lander imaging data, *J. Geophys. Res.*, *82*, 4479-4496, 1977.

Pollack, J., D. Colburn, F. M. Flasar, R. Kahn, C. Carlston, and D. Pidek, Properties and effects of dust particles suspended in the Martian atmosphere, *J. Geophys. Res.*, *84*, 2929-2945, 1979.

Pollack, J., M. E. Ockert-Bell, and M. K. Shepard, Viking Lander image analysis of Martian atmospheric dust, *J. Geophys. Res.*, this issue.

Roush, T. L., Characterization of the spectral reflectance of mafic silicates, hydrated silicates, and hydrated silicate-water ice mixtures in the 0.6 to 4.5 μm wavelength region and applications to planetary science, Ph.D. dissertation, 130 pp., Univ. of Hawaii, Honolulu, 1987.

Roush, T. L., Infrared optical properties of Mars soil analog materials: Palagonites, *LPI Tech. Rep. 92-04, Part I*, 32-33, 1992.

Roush, T. L., and D. Blake, Characterization of Mauna Kea palagonite using transmission electron microscopy, *Lunar Planet. Sci. Conf.* XXII, (abstract), Houston, 1139-1140, 1991.

Roush, T., J. Pollack, and J. Orenberg, Derivation of midinfrared (5-25 μm) optical constants of some silicates and palagonite, *Icarus*, *94*, 191-208, 1991.

Roush, T. L., D. L. Blaney, and R. B. Singer, The surface composition of Mars as inferred from spectroscopic observations, in *Remote Geochemical Analysis: Elemental and Mineralogical Composition*, edited by C. Pieters and P. Englert, pp. 367-394, Cambridge University Press, New York, 1993.

Singer, R., Spectral evidence for the mineralogy of high albedo soils and dust on Mars, *J. Geophys. Res.*, *87*, 10159-10168, 1982.

Soderblom, L. A., The composition and mineralogy of the Martian surface from spectroscopic observations: 0.3 μm to 50 μm, in *Mars*, edited by H. Kieffer, B. Jakosky, C. Snyder, and M. Matthews, pp. 557-593, University of Arizona Press, Tucson, 1992.

Soderblom, L. A., K. Edwards, E. M. Eliason, E. M. Sanchez, and M. P. Charette, Global color variations on the Martian surface, *Icarus*, *34*, 446-464, 1978.

Toon, O. B., J. B. Pollack, and C. Sagan, Physical properties of the particles composing the Martian dust storm of 1971-1972, *Icarus*, *30*, 663-696, 1977.

West, R. A., G. S. Orton, B. T. Draine, and E. A. Hubbell, Infrared absorption features for tetrahedral ammonia ice crystals, *Icarus*, *80*, 220-223, 1989.

Zurek, R. W., Martian great dust storms: An update, *Icarus*, *50*, 288-310, 1982.

R. T. Clancy, Space Science Institute, Suite 294, 1234 Innovation Drive, Boulder, CO 80303-0588.

S. W. Lee, Laboratory for Atmospheric and Space Physics, Campus Box 392, University of Colorado at Boulder, Boulder, CO 80309-0392.

G. R. Gladstone, Southwest Research Institute, P. O. Drawer 28510, 6220 Culebra, San Antonio, TX 78284.

W. McMillan, Department of Physics, University of Maryland Baltimore County, Baltimore, MD 21228.

T. Roush, NASA Ames Research Center, MS 245-3, Moffett Field, CA 94035-1000.

(Received March 2, 1994; revised June 14, 1994 ;
accepted July 15, 1994.)

Page intentionally left blank

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5265–5275, MARCH 25, 1995

Temperature measurements of a Martian local dust storm

Ralph Kahn
Jet Propulsion Laboratory, California Institute of Technology, Pasadena

Abstract. A technique for estimating the ground and near-ground atmospheric temperatures within a Martian local dust storm is presented. It is applied to soundings taken by the Viking orbiter infrared thermal mapper (IRTM) instrument at four times-of-day for one storm. Essentially, a comparison is made between infrared radiances emerging from the storm interior and those from the region surrounding the storm. Particle extinction properties are assumed to be independent of position in the storm region, and scattering properties must be selected arbitrarily. For the storm studied here, the ground temperature in the interior is at least 6 K cooler, whereas the near-ground atmospheric temperature may be less than or comparable to, those of the surroundings. The thermal structure of the storm interior did not change measurably between 11.5 and 16.6 hours local time. These observations favor theories of dust storm development in which regional winds rather than local, dust-driven convection inititate the mobilization of dust from the surface. It is also concluded that the optical properties of dust particles in this local storm differ from those observed by Mariner 9 during the 1971-1972 global dust storm.

1. Introduction

Although the occurrence of local and regional dust storms on Mars is well documented [*Briggs et al.* 1979; *Peterfreund and Kieffer* 1979; *Zurek* 1982], there is still some question as to the means by which dust is initially mobilized at the Martian surface [e.g., *Greeley et al.*, 1992]. The idea that regional winds superimposed on the general circulation provide the shear stress necessary to lift dust at the surface is supported both by theory and observation [*Leovy et al.*, 1973; *Zurek et al.*, 1992]. Similarly, there is evidence that direct heating of the surface creates free convective vortices that mobilize dust locally as dust devils [*Gierasch and Goody*, 1973; *Ryan and Lucich*, 1983; *Thomas and Gierasch*, 1985]. Constraints on the evolution of temperature within Martian local dust storms would provide important clues as to the mechanisms that operate under different environmental conditions on Mars.

This paper demonstrates a technique that places some very crude but relevant constraints on the temperatures of Martian local dust storms. The method relies upon broadband infrared soundings from the infrared thermal mapper (IRTM) experiment flown on the Viking mission to Mars. The IRTM made observations over an extensive part of the planet, covering a wide range of seasons and local conditions. Although four of the five detector channels were designed to examine surface emissions [*Kieffer et al.*, 1972], the substantial infrared opacity of the dusty Martian atmosphere, even during relatively clear conditions, produced a measurable atmospheric contribution to the observed infrared radiances [*Kieffer et al.*, 1977]. These data represent the most detailed and extensive spatial and temporal record of the Martian atmosphere yet available. In particular, IRTM data include observations made during the development of a few local dust storms [*Peterfreund and Kieffer*, 1979]. Thermal signatures at these sites show distinct spectral characteristics that contain information about the thermal behavior of the storms.

Current knowledge of infrared absorption and scattering properties of particles in the Martian atmosphere, the surface emission properties, and our understanding of the vertical distribution of optical depth in the Martian sky, are limited. Interpretation of broadband radiances, which sample the surface together with the lower atmosphere, is therefore only partly constrained by existing data. This paper explores a number of the most reasonable hypotheses about the atmospheric optical and thermal structure for consistency with the IRTM observations over a dust storm. The technique essentially makes a comparison between infrared radiances emerging from the storm interior and those coming from the area surrounding the storm. Data taken over the Viking Lander 1 (VL1) site are also examined.

The paper begins by presenting the IRTM data used in this study, and the model used to simulate IRTM spectra. Then the thermal signatures of an area adjacent to a local dust storm are examined. In the dust storm surroundings, the vertical temperature structure is easier to constrain than in the interior. IRTM soundings in the surroundings are used to obtain ratios of particle extinction optical properties from constraints on the temperature structure. Particle scattering properties must be selected arbitrarily for this calculation, but we can place sufficient constraints on the atmospheric optical depth and vertical temperature structure to obtain meaningful extinction coefficients. Next, an attempt is made to calculate the mean atmospheric infrared spectral extinction over the VL1 site, as was done in the dust storm surroundings. However, the low atmospheric optical depth at the VL1 site makes the inversion particularly sensitive to the unknown ground temperature and emissivity. Finally, particle properties are assumed to be independent of position in the storm region. With the help of the constraints on particle properties from the storm surroundings, IRTM soundings of the storm interior are used to obtain an

Paper number 94JE02766.
0148-0227/95/94JE-02766$05.00

effective dust cloud temperature and the ground temperature at the storm center.

The results of these studies represent a step toward constructing thermodynamic models of dust storm generation and growth. The data examined in this paper are of low surface spatial resolution, so it is not possible to perform critical tests on length scales small enough to rule out either of the proposed dust storm mechanisms. However, this work establishes the viability of a technique that can be applied to data with higher spatial resolution.

2. IRTM Dust Storm Data

The data used in this study were taken by the Viking Orbiter IRTM experiment in four broadband channels centered near 7, 9, 11, and 20 μm. The IRTM channel passbands are given in Figure 1. The 7-, 9-, and 11-μm bands are each about 2 μm wide. The 20-μm band is nearly 6 μm wide. Following conventions established in earlier presentations of IRTM data [e.g., *Kieffer et al.*, 1977], the observed radiances are given in terms of brightness temperatures T_7, T_9, T_{11}, and T_{20}, measured in the bands centered near 7, 9, 11, and 20 μm, respectively.

The IRTM instrument made 28 simultaneous soundings: four measurements directed at each of seven spots on the surface. The instrument contained four telescopes. Each telescope sampled light from all seven spots, and focused this radiation onto seven detectors. The telescopes were aimed so that the same seven surface locations, labeled spots one to seven, were observed by each telescope. One telescope served seven 20-μm detectors, another provided light for seven 11-μm detectors, so there is an 11-μm and a 20-μm measurement for every spot in the IRTM data. A third telescope made seven albedo measurements, which are not used in this study. The fourth telescope had three 7-μm detectors, three 9-μm detectors, and one 15-μm channel to measure atmospheric emissions. Therefore, simultaneous 7- and 9-μm measurements cannot occur at any spot. For this reason all the spectral measurements used here are compared with the 11-μm data, which was obtained for every location. Further details of the IRTM are given by *Kieffer et al.* [1977].

A local dust storm that occurred in the Solis Planum region of Mars during summer in the southern hemisphere (L_S = 227°) was selected for this study. The storm appeared in four separate sequences of IRTM observations on the same day, between 11.5 hours and 16.6 hours local time [*Peterfreund and Kieffer*, 1979]. The orbiter camera recorded a dust cloud in this region on the subsequent day. Figure 2 illustrates the dust storm data. Each graph contains the spectral brightness temperature differences along an east to west traverse, transecting the region of the storm with the most pronounced temperature differences relative to the surroundings. Spectral temperature differences are calculated from pairs of observations with the same telescope spot number to minimize sampling errors in regions of large thermal contrast.

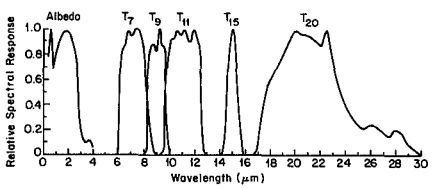


Figure 1: IRTM passbands [from *Kieffer et al.*, 1977].

Although brightness temperature values exhibit a 3 K to 5 K scatter for adjacent measurements, temperature differences taken in this way are consistent to better than 1 K in most instances. Some sampling error does occur, since for fixed temperature, the slope of the Planck function depends on wavelength. This causes the spectral detectors to have varying sensitivity to thermal inhomogeneity in the field of view [*Martin et al.*, 1979; *Jakosky*, 1979]. The spatial resolution on the Martian surface of the four sequences shown lies between 220 km and 230 km.

For this study, data were selected within a 3°-4° strip centered about the latitude shown, so the centers of the sampling footprints are within 180 to 240 km of the latitude of the traverse. The physical size of the dust storm, as judged from the perturbation of the temperature in IRTM contour maps [*Peterfreund and Kieffer*, 1979] is about 450 km by 850 km. This size is comparable to that of the cloud photographed in the same region on the following day by the orbiter cameras. In summary, the available data average out any thermal features with length scales less than 200 km, and there are three to four resolution elements across the storm itself.

The accuracy of the brightness temperatures varies among the channels. The absolute errors are less than 0.5 K in the T_7, T_9, T_{11}, and T_{20} bands for the temperature ranges considered here, according to *Kieffer et al.* [1977].

The storm center is taken to be at the site of minimum T_{11}. (Maps of T_{11}, showing the locations of the minima for several storms, are given by *Peterfreund and Kieffer* [1979].) In Figure 2, note that $T_7 - T_{11}$ reaches a maximum near the storm center. In three of the traverses, $T_9 - T_{11}$ has a maximum near the storm center, whereas $T_{20} - T_{11}$ shows a relative minimum. The temperature differences are asymmetric about the extremes, with a considerably steeper gradient in $T_\lambda - T_{11}$ (where $\lambda = 7, 9$, or 20) to the east of the storm center, and a progressively shallower slope to the west for later measurements, except for $T_{20} - T_{11}$ in the fourth sequence. From these data, representative models of the dust storm center and surroundings were constructed, to be used in the numerical simulations. These are presented in Table 1. For this storm, the brightness temperatures decrease about 35 K, the $T_7 - T_{11}$ and $T_9 - T_{11}$ contrasts increase, whereas $T_{20} - T_{11}$ decreases in the dust storm interior relative to the surroundings.

3. Numerical Simulation of Spectra

To explore the dependence of infrared signatures on physical conditions, synthetic spectra were produced to compare with the IRTM data. A two-stream radiative transfer model similar to that of *Toon et al.* [1977] was used. Toon et al. analyzed data from the infrared interferometer spectrometer (IRIS) that flew on the Mariner 9 mission to Mars. The instrument obtained high-resolution spectra for wavelengths between 5 and 50 μm. Vertical temperature profiles were calculated from an analysis of CO_2 bands in the IRIS data. The dust model required input values of the ground temperature, the optical depth at one wavelength, and the particle extinction and scattering properties. Toon et al. assumed spherical particles, and let the real and imaginary parts of the index of refraction at each wavelength, as well as the particle size distribution, be free parameters giving the particle scattering properties in their model. They obtained particle indices of refraction, size distribution, and atmospheric optical depth at each wavelength by comparing the model results with the IRIS data. The data analyzed by Toon et al. were taken during the Martian global dust storm of 1971-1972, when the atmospheric dust loading was near its maximum.

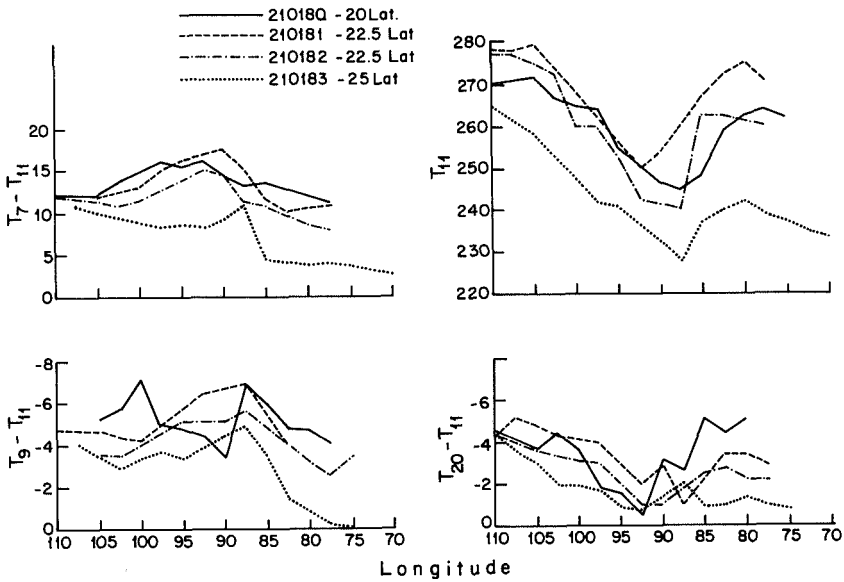


Figure 2: IRTM dust storm data. Each longitude traverse contains data from a latitude strip 3° to 4° wide. Temperature differences (K) and the 11-μm brightness temperature are plotted as functions of longitude for the four IRTM sequences. Sequence 210180 was taken at 11.5 hours local time, with a surface resolution of 220 km, 210181 is at 12.8 hours and 230-km resolution, 210182 at 14.6 hours and 230-km resolution, and 210183 at 16.6 hours and 225-km resolution.

For the present study, source functions were obtained at all levels using upward and downward intensities calculated from the two-stream equations. This method was selected because it is the simplest one that contains two parameters describing Martian dust properties; this is as much information about dust properties as is available in the data. The method also allows direct comparison with the results of Toon et al. The technique approximates the angular dependence of the radiation field due to scattering anisotropy by taking hemispherical average values in two directions (appendix). To obtain the emergent intensity, the source function in the upward direction is integrated along the path selected by the viewing angle, thereby preserving the geometrical angular dependence of the result. The equation for the upward two-stream intensity is written as a second-order differential equation with two-point boundary values, solved using a standard numerical method [*Lindzen and Kuo*, 1969]. The downward two-stream intensity is subsequently obtained as an initial value problem using a Runge-Kutta method of order 3.

Table 1. Representative Dust Storm Thermal Signature

	Storm Center	Storm Surrounding
$T_7 - T_{11}$	16.	11.
$T_9 - T_{11}$	-7.	-5.
$T_{20} - T_{11}$	0. to -2.	-4.
T_{11}	240.	275.

Details of the calculation are given in the appendix. The procedure was tested by calculating the examples given by Toon et al., and obtaining the same results to two significant figures or better.

Unlike the high-resolution IRIS data, this study relies upon broadband observations, with a direct measurement of the atmospheric temperature at only one level, near 24 km, obtained with the IRTM 15-μm channel. The particle scattering properties at each wavelength are given by two parameters, the single scattering albedo ($\tilde{\omega}_o$) and the integral of the normalized single scattering phase function weighted by the cosine of the scattering angle, which is called the asymmetry factor (β). These parameters cannot be determined from the IRTM data available in the region of interest; initially, they are assumed to be equal to those derived by Toon et al. The sensitivity of the results to small changes in these values is examined later. Vertical temperature structure, spectral optical depth, and ground temperature are free parameters, as discussed below.

To produce a model spectrum, the ground temperature, the atmospheric thermal profile, the vertical distribution of optical depth, the spectral extinction coefficients $Q_e(\lambda)$, the scattering parameters $\tilde{\omega}_o(\lambda)$ and $\beta(\lambda)$, and the surface emissivity must be specified. The model allows arbitrary selection of vertical temperature structure. It is parameterized here by the ground temperature T_g, the atmospheric temperature at the ground T_o, the boundary layer height Z_{BL} and boundary layer lapse rate Γ_{BL}, the atmospheric lapse rate Γ_a, and the height of the tropopause Z_{TROP}, above which the temperature is taken to be independent of

height. These quantities are illustrated in Figure 3a. The thermal structure used for dust storm interior models, which are discussed in a later section, are shown in Figure 3b. The emission angle for most of the dust storm observations used in this study lies between 30° and 45°. In most of the models, this parameter is set to 35°, which happens to be the angular region where the two-stream method most closely approximates detailed radiative transfer calculations for expected particle size distributions [*Toon et al.*, 1977]. The calculated spectral brightness temperature differences vary less than 1 K when the input emission angle is changed by 5° around the nominal value selected.

4. Models of the Dust Storm Surroundings

We seek information about atmospheric and ground temperatures in the interior of a dust storm. The general radiative transfer problem requires a knowledge of both the particle properties and the vertical temperature structure. Reasonable estimates of some of the particle properties can be obtained using an assumed vertical temperature in the dust storm surroundings. Some parameters in the temperature profile of the storm interior can then be retrieved using the deduced particle properties.

In this section, we calculate the ratios of the spectral extinction coefficients required to reproduce the observed spectra for the dust storm surroundings. These results will depend upon the unknown vertical temperature structure in the dust storm surroundings, so a range of possible values is obtained for the spectral extinction coefficients, based upon different choices of T_g, T_o, and Γ_{BL}. However, with the available data, reasonable limits are placed on the values of these parameters in the storm surroundings. Extinction coefficients are represented as ratios $R_\lambda = Q_\lambda / Q_{11}$, where Q_λ is the extinction coefficient in the wavelength band λ, which may be 7, 9, or 20 μm, and Q_{11} is the extinction coefficient in the 11-μm band. For each assumed temperature profile, three ratios of extinction coefficients are calculated, together with a value of the optical depth at 11 μm, to match the four observed brightness temperatures.

Since the ratios of extinction coefficients are less dependent on the assumed temperature profile than are the optical depths at each wavelength, these ratios are retrieved for later use in the dust storm interior models. For the interior models, the 11-μm dust cloud optical depth (τ_{11c}) is treated as a free parameter. With these inputs, and the observed spectrum at the storm center, values of T_g and the effective dust cloud temperature T_c are constrained using a least squares scheme.

Figure 4 shows the results of numerical modeling of the dust storm surroundings. We now discuss the choice of parameter space and sensitivity of the results to each of the parameters.

The uncertainty in R_λ is primarily due to the indeterminacy of T_g. T_g was varied from 288 K to 292 K. Near midday, and where the atmospheric visible optical depth is less than 1, T_g is expected to be greater than the atmospheric temperature, based on theoretical modeling [*Gierasch and Goody*, 1972; *Pollack et al.*, 1979]. The largest observed brightness temperature in the dust storm surrounding is 286 K, which is therefore a lower bound on the choice of T_g. The upper bound on T_g in the parameter space is arbitrary. A reasonable guess at an upper bound on T_g is made by noting that T_g is a sensitive function of the 11-μm atmospheric optical depth. There is no direct measurement of the 11-μm optical depth, but the visible optical depth in the dust storm surroundings is typically less than 1, as evidenced by the appearance of ground features in the orbiter images of local dust storms [*Briggs et al.*, 1979], including one taken in the same region as the storm used in this study, on the subsequent day.

Next, the 11-μm optical depth needs to be related to the estimated visible optical depth. The relative particle extinction cross sections for Martian dust in the visible and infrared regions is not yet well established, but probably has a value around 2 [*Zurek*, 1982; *Martin*, 1986]. *Pollack et al.* [1979] use visible properties obtained from the Viking Lander camera experiments and infrared properties derived from the Mariner 9 IRIS measurements, thereby assuming that the atmospheric particles are the same in these two situations. Using these data, the 11-μm optical depth would be less than the visible optical depth.

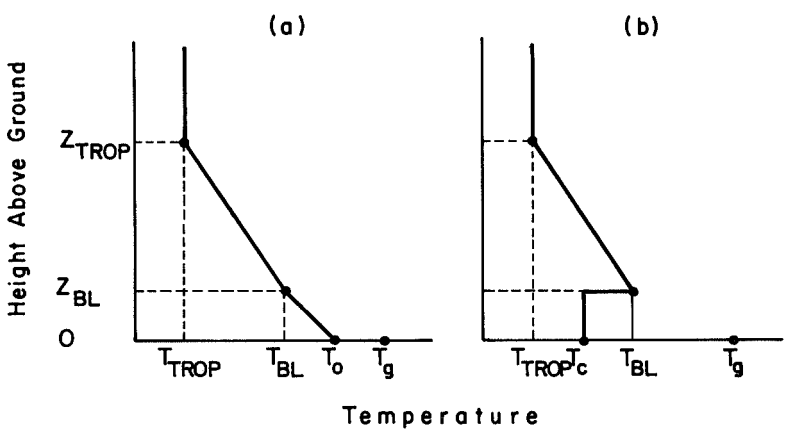


Figure 3: Parameters for model thermal profiles. (a) Model parameters for atmospheric temperature structure. Here lapse rates $\Gamma_{BL} = (T_0 - T_{BL})/Z_{BL}$ and $\Gamma_a = (T_{BL} - T_{TROP}) / (Z_{TROP} - Z_{BL})$. (b) Model parameters for dust storm interior temperature structure.

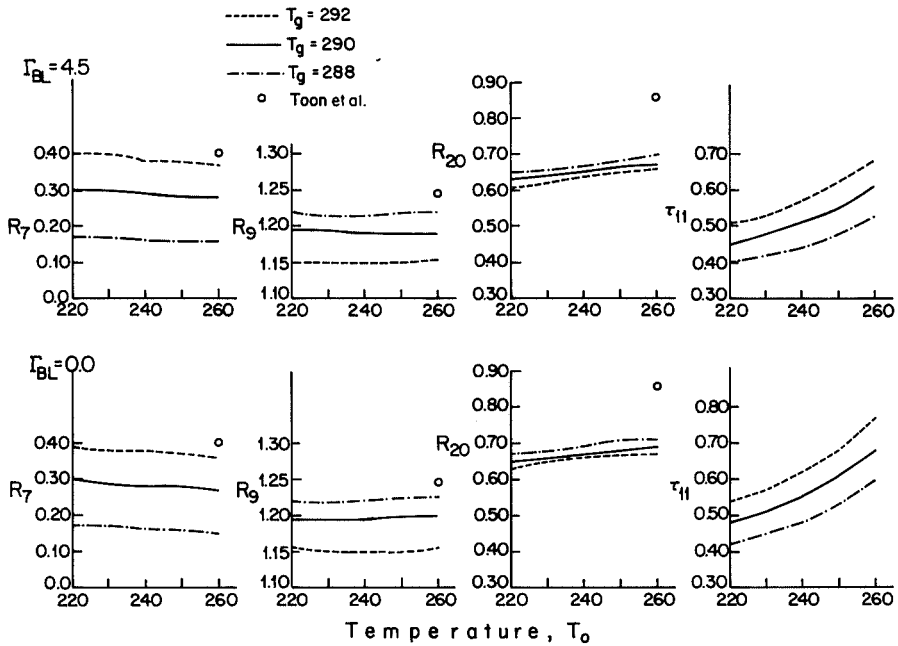


Figure 4: IRTM dust storm surroundings models. Derived ratios of extinction coefficients (R_9/R_{11}) are plotted against the atmospheric temperature T_o for three choices of T_g and two choices of Γ_{BL}. Here, Z_{ZROP} was fixed at 50 km, and $T_{24\,km}$ at 200 K for all cases. Circles show the corresponding ratios for Mariner 9 dust particles.

Unfortunately, we later conclude that the particle properties observed over the dust storm must differ from those of the Mariner 9, so we have no way to make a direct comparison between visual and 11-μm opacities, which would set a firm upper bound on the value of T_g. However, the infrared extinction cross sections for micrometer-sized dust particles are typically smaller than the visible extinction cross sections. Therefore, values of T_g were selected for which the 11-μm optical depth is less than 1. The chosen range of T_g values results in a 30% uncertainty in R_7, and smaller uncertainties in R_9 and R_{20}.

The surface emissivities were all chosen to be equal to 1. For the most probable surface materials, the 7-μm emissivity will be near 1 [*Martin et al.*, 1979]. Emissivities in the 9-, 11-, and 20-μm bands will be more dependent upon surface composition and structure, and may be as low as 0.85 for some likely situations [*Kieffer et al.*, 1972; *Conel*, 1969], but probably fall between 0.92 and 1 [*Christensen*, 1982]. The ground temperature affects the radiance in the 7 μm band more than the other bands. The 10% to 15% uncertainty in surface emissivity can contribute as much as a 10% error to the derived extinction coefficient ratios, particularly R_7. It is more likely that the 9- and 11-μm emissivities, if less than 1, are within a few percent of each other, so that R_9 will exhibit smaller errors.

Values of T_o (see Figure 3a) are taken to fall between 220 K and 260 K, which covers the range of possible values consistent with radiative-convective models of the Martian atmosphere in the afternoon at midlatitudes [*Gierasch and Goody*, 1972; *Pollack et al.*, 1979]. These models were confirmed to a limited degree by values of T_o measured at the Viking Lander sites by the Viking meteorology experiment [*Hess et al.*, 1977]. The upper limit to this range is also constrained by limiting the value of τ_{11} to unity. The boundary layer height was varied between 3 and 8 km, and adiabatic (4.5 K/km) and isothermal lapse rates in the boundary layer were chosen as limiting cases. The range of boundary layer height includes the theoretical boundary layer thicknesses in the afternoon [*Gierasch and Goody*, 1968; *Pollack et al.*, 1979], and covers the dust cloud height estimates made from Viking imaging and infrared data [*Peterfreund and Kieffer*, 1979]. The atmospheric lapse rate above the boundary layer was determined by fixing the atmospheric temperature at 24 km to be 200 K, as measured by the 15-μm channel of the IRTM instrument in the region of the storm observation [*Martin and Kieffer*, 1979]. An exponential vertical distribution of optical depth, with a scale height of 10 km, was used for all models of the dust storm surroundings, to match the results obtained from Viking lander imaging experiments [*Kahn et al.*, 1981]. This implies that the dust is well mixed in the lower atmosphere. For exponential vertical distributions of optical depth, the results are not noticeably affected by the choice of tropopause height, since only a few percent of the total radiation comes from the atmosphere above 20 km in these models. For example, the deduced ratios of extinction coefficients vary less than 2.5% when the tropopause is

moved from 50 to 64 km, for a model with other inputs set in the middle of the parameter space.

Figure 4 presents a series of model results covering the parameter space discussed above. For all cases, mean values of $\tilde{\omega}_o$ and β weighted by the Planck function are used, as obtained for the particles observed by Mariner 9. This assumption is discussed subsequently. The weighted average of property A at temperature T is defined as:

$$\bar{A}_i = \int A(\lambda) B(\lambda,T) f_i(\lambda)\, d\lambda \Big/ \int B(\lambda,T) f_i(\lambda)\, d\lambda \qquad (1)$$

where $B(\lambda,T)$ is the Planck function at temperature T and $f_i(\lambda)$ is the relative IRTM response for band i at wavelength λ, given by *Kieffer et al.* [1977]. Values of the dust particle parameters calculated in this manner are shown in Table 2. From Figure 4, τ_{11} increases with T_o, but the ratios R_λ are virtually independent of T_o and Γ_{BL}. R_7 depends heavily on the choice of T_g, whereas R_9 and R_{20} are much less affected. This confirms that the 7-μm channel samples the ground much more than the other channels, as was suggested by the higher observed brightness temperatures at 7 μm by the IRTM instrument. Table 3 presents the range of values of the extinction coefficient ratios deduced from Figure 4. For comparison, the ratios of extinction coefficients for the particles observed by Mariner 9 [*Toon et al.*, 1977] are also given in Table 3, and are plotted as circles in Figure 4. The derived extinction coefficient ratios are not consistent with those of the Mariner 9 particles. Note that by increasing T_g above 292 K, R_7 can be made to fit, but then R_9 would exhibit an even poorer fit to the Mariner 9 data. Other evidence that the properties of Martian atmospheric dust particles vary in space and time has appeared in several sources [*Clancy and Lee*, 1991; *Kahn et al.*, 1992].

The values of $\tilde{\omega}_o$ and β were assumed to be equal to those of the Mariner 9 particles. There is insufficient information to make a direct determination of these values from the Viking data. The particle scattering properties are treated as free parameters in this study, avoiding the questions of particle composition and size distribution. The sensitivity of the conclusions to these choices is tested below. A number of perturbation studies were also performed on the values of $\tilde{\omega}_o$ and β, although there is no formal basis on which to limit the choices. For 10% changes in $\tilde{\omega}_o$ and β, we obtain less than a 5% change in the derived value of Q_7, a 10% change in Q_9 and Q_{11} and less than a 3% change in Q_{20}. One alternative to this approach is to vary the particle composition and possibly the size distribution to obtain particle optical parameters which yield agreement between the model and the observed spectra. This approach has been employed by *Hunt* [1979], who concludes that the thermal signature of a dust cloud is reproduced using dust composed of two common terrestrial minerals, with the same size distribution as the Mariner 9 particles. However, this result relies on specific assumptions about the temperature structure, and especially upon the particle size distribution and allowed composition of the particles. (For example, water ice was not included as a possible contributor to the column optical depth.) These assumptions are of necessity arbitrary, since there are only four directly measured quantities in the problem. The introduction of an additional free parameter increases the difficulty in extracting valid information from the model.

Table 2. IRTM Channel-Mean-Weighted Properties at T = 220 K for Dust Observed by Mariner 9 [*Toon et al.*, 1977]

	7 μm	9 μm	11 μm	20 μm
Q_e, cm^2	4.88×10^{-8}	1.49×10^{-7}	1.21×10^{-7}	1.04×10^{-7}
$\tilde{\omega}_o$	0.453	0.432	0.636	0.351
β	0.780	0.506	0.559	0.318

Table 3. Particle Extinction Coefficient Ratios

	Dust Storm Surrounding	Mariner 9
R_7	0.29 (+0.11, -0.25)	0.403
R_9	1.18 (±0.06)	1.24
R_{20}	0.67 (±0.04)	0.860

The derived ratios of extinction coefficients are model dependent, particularly with regard to the assumed scattering properties of the particles. The most likely values of R_λ, based on a wide range of possible thermal conditions, are given in Table 3. In summary, for all reasonable thermal profiles, the observed particle properties differ from those of the Mariner 9 dust. Given the arbitrary assumption that the particle scattering properties are similar to the Mariner 9 dust, this approach sets useful limits on the ratios of extinction coefficients from the IRTM data.

5. Viking Lander 1 Sol 0 Models

The best available constraints on Martian atmospheric structure were acquired over the VL1 site at the time of the VL1 entry (sol 0). The vertical temperature profile was measured directly by probes on the entry vehicle [*Sieff and Kirk*, 1977]. We also have constraints on the vertical distribution of optical depth locally from lander camera experiments [*Kahn et al.*, 1981]. This appears to be an ideal place to test the procedure for retrieving particle extinction properties from IRTM data that was presented in the previous section. With the additional information, the particle properties would seem to be better determined than for the dust storm surroundings case. However, the unknown surface emissivities, ground temperature, and particle scattering properties are still required for the model. The atmospheric visible optical depth is more than a factor of 2 lower at the VL1 site than in the dust storm surroundings, based on imaging data. The differences between brightness temperature values in the IRTM channels are much smaller at VL1 than in the dust storm surroundings. This is a manifestation of the lower optical depth at the VL1 site, where most of the radiation observed at the IRTM detector comes directly from the ground.

In Table 4 the IRTM observations are listed which most closely correspond to the VL1 site at sol 0. The landing occurred at 16.13 hours local time. Table 4 shows the orbiter revolution and IRTM sequence identification number, mean viewing angle of the observation, approximate local time of day, and surface resolution of the observations. It also contains the mean and standard deviation of the IRTM 11-μm brightness temperature, and of the IRTM channel differences used in this study. For each IRTM sequence, all the data in a 1° box around the landing site have been accumulated, and the brightness temperature differences were calculated using pairs of measurements made simultaneously for the same telescope spot number. From these data, a representative thermal signature was abstracted (Table 5), to be compared with synthetic spectra. These results are modeled using a temperature profile with T_o = 241.5 K, the top of the boundary layer at 7.5 km with temperature 215.95 K, and a 140 K

Table 4. VL1 Site Data Summary Giving Mean and Standard Deviation for the 11-μm Brightness Temperature and Spectral Brightness Temperature Differences

Revolution/ Sequence	Emission Angle	Local Hour	Resolution, km	T_{11}	$T_7 - T_{11}$	$T_9 - T_{11}$	$T_{20} - T_{11}$
20/105	27.7	15.39	10.7	248.6 (0.78)	2.55 (0.35)	1.07 (0.06)	.57 (0.31)
20/106	17.1	15.49	9.0	245.6 (0.43)	2.40 (0.26)	1.07 (0.15)	.73 (0.28)
20/107	17.5	15.57	8.8	245.0 (0.49)	2.10 (0.10)	0.87 (0.23)	.52 (0.45)
27/102	19.1	15.90	9.1	243.1 (0.35)	2.07 (0.12)	1.20 (0.20)	.90 (0.49)
27/103	5.6	15.94	8.1	242.9 (0.97)	1.97 (0.25)	1.03 (0.23)	.21 (0.43)
27/104	16.3	15.98	8.4	241.7 (0.77)	1.73 (0.12)	1.17 (0.15)	.63 (0.50)
27/105	30.3	16.00	10.0	240.8 (0.67)	1.87 (0.21)	0.67 (0.50)	.44 (0.59)
35/128	40.5	15.62	13.8	247.5 (0.46)	3.23 (0.45)	2.30 (0.42)	3.1 (0.23)
35/129	31.2	15.65	11.0	247.4 (0.79)	2.20 (0.17)	1.50 (0.10)	.34 (0.22)
37/113	23.0	15.61	9.5	248.2 (0.37)	2.13 (0.15)	1.67 (0.15)	.67 (0.25)

tropopause at 57.5 km [*Sieff and Kirk*, 1977]. The scale height for optical depth is set at 10 km. There is no direct measurement of T_g, since the ground temperature sensor is partly shaded during the afternoon [*Moore et al.*, 1977]. Values of T_g between 248 K and 258 K are expected, based on theoretical modeling [*Pollack et al.*, 1979; *Moore et al.*, 1977].

The results are shown in Figure 5. For comparison, particle extinction ratios for Mariner 9 observations are given as symbols in the diagram. The average particle properties required in the vertical column over the VL1 site as deduced from this calculation differ from both the Mariner 9 particles and the properties deduced for the dust storm area (Figure 4). There are many possible reasons for the difference in calculated particle properties. These include differences in particle composition and size distribution, differences in surface emissivity, and the possible presence of additional constituents in the vertical column, as for example, water ice haze. The VL1 site results are also much more susceptible to errors in the data, since the observed spectral temperature differences here (less than 2 K) are much smaller than in the dust storm surroundings cases (4 K to 11 K).

The calculations in Figure 5 assume $\tilde{\omega}_0$ and β to be the same as their Mariner 9 values, and the spectral emissivity of the surface was set to 1. The total atmospheric optical depth at the VL1 site is smaller than in the dust storm area, so the ground emissivity for this case has a much greater effect on the net radiance at the detector than in the previous calculations. Examination of the surface rocks at the VL1 site suggest basaltic composition [*Binder et al.*, 1977]. If the surface in the vicinity of the Lander is composed of basalt, the average surface emissivity in each of the IRTM bands is expected to range between 0.9 and 1.0 [*Hovis and Callahan*, 1966]. However, the emissivity for pure solid basalt reaches a minimum of about 0.83 near 10 μm, which will contribute to the 9- and 11-μm bands to varying degrees, depending on the size of the surface particles. The 7-μm emissivity is 0.97 or greater for particles larger than 0.105 mm. These data suggest that R_7 may be higher than is shown in Figure 5 if the surface is composed of pure basalt. No conclusion can be drawn regarding R_9. Reasonable assumptions about surface materials could produce temperature differences on the order of those observed in the IRTM spectrum. In summary, the particle extinctions derived from the information currently available at the VL1 site are likely to be seriously affected by the arbitrary choice of T_g and surface emissivity.

Table 5. Viking Lander 1 Site Midafternoon Spectral Brightness Temperature Differences, Near Sol 0

Temperature Difference	Degrees Kelvin
$T_7 - T_{11}$	2.0
$T_9 - T_{11}$	1.0
$T_{20} - T_{11}$	0.5
T_{11}	245.8

6. The Dust Storm Interior

Using the range of particle properties deduced in the dust storm surroundings, we now construct a model of the atmosphere at the dust storm interior. The dust cloud is treated as a boundary layer phenomenon, with a characteristic temperature T_c and an optical depth at 11 μm τ_{11c}. Above the cloud, the vertical structure of the atmosphere is assumed to be similar to the surroundings, and unaffected by the presence of the cloud. This structure is illustrated in Figure 3b. The 15-μm brightness temperature measurements, which sample at about 24 km above the surface, are unaltered by the presence of the dust cloud [*Peterfreund and Kieffer*, 1979]. Visible images of local dust clouds also exhibit sharp upper boundaries [*Briggs et al.*, 1979]. In addition, observed radiances are heavily weighted to the more optically dense lower atmosphere even in the dust storm surroundings, as discussed in section 4. Since the dust cloud has a larger optical depth than the boundary layer of the dust storm surroundings, the contribution of the atmosphere above the cloud top to the total radiance will not have an important effect upon the results.

The temperature structure for the dust storm interior model contains two parameters, T_g and T_c, that are fit in a least squares sense by the four spectral radiances obtained over the dust storm interior (see Figure 3b). For this procedure to be meaningful, the four atmospheric weighting functions must overlap in the region of the cloud, since there is only one free parameter in the vertical temperature profile above the ground. Knowledge of the vertical distribution of optical depth within the cloud is insufficient to justify a more elaborate temperature structure model. The

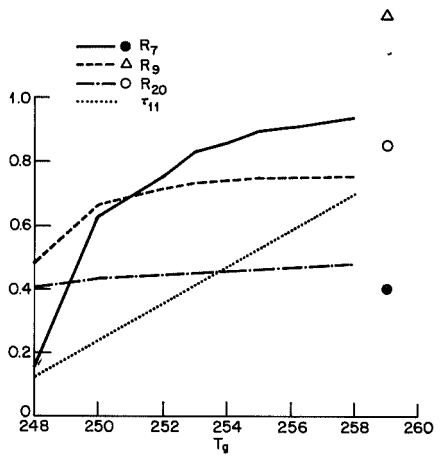


Figure 5: Viking Lander 1 site models. Ratios of extinction coefficients and the 11-μm optical depth are plotted as functions of ground temperature for a model in which the atmospheric temperature is derived from in situ measurements taken during spacecraft landing. Symbols show the corresponding ratios for Mariner 9 dust particles.

atmospheric weighting functions must also be insensitive to the rest of the atmospheric structure, and must sample the ground relative to the cloud to varying degrees. The last two properties are strongly suggested by the results shown in Figure 4, and can be checked posthoc by the convergence and stability of the least squares procedure. From Figure 4, T_g is the most sensitive to the value of T_7, whereas of all the channels, the response of T_o is greatest to changes in T_{20}. The relative independence of the atmospheric contribution to the measurements can be evaluated by calculating the overlap integral of the weighting functions, which is a measure of the coincidence of the regions sampled by any pair of channels:

$$W_{i,j} = \left| 1 - \int_o^\infty K_i K_j \, dz \Big/ \int_o^\infty K_i^2 \, dz \right| \qquad (2)$$

where K_i is the derivative of the transmission function for spectral band i [e.g., *Rogers*, 1976]. This quantity is equal to zero if the two channels sample exactly the same vertical region of the atmosphere, and approaches 1 if the channel sampling functions do not overlap. For $W_{i,j}$ between about 0.0 and 0.3, $1\text{-}W_{i,j}$ is approximately equal to the fractional overlap of the two channel sampling functions in these models. In the dust storm surroundings, the distribution of optical depth leads to values of $W_{i,j}$ which range between 0.16 and 0.3, given the models in section 4. Thus the overlap of sampling functions is very high. Increasing the optical depth in the boundary layer, to simulate a dust cloud, produces values of $W_{9,11}$, $W_{11,20}$, and $W_{9,20}$ between 0.20 and 0.25. These three channels sample the atmosphere at nearly the same level. Overlap with the 7-μm channel may be as small as 50% for some models. This channel samples closer to the ground than the others, and also receives a larger proportion of net radiation from the ground itself. By calculating one representative atmospheric temperature in the inversion scheme, in addition to the ground temperature, the value of T_g will be aliased if the cloud is not isothermal. However, this effect is small, as is shown in the discussion of perturbation studies below.

Figure 6 illustrates the least squares solution to T_g and T_c for the dust storm interior as a function of τ_{11c}. The particle properties obtained for the dust storm surroundings, as given in Table 3, have been used. The vertical dust distribution is discontinuous, with the cloud top set at 4 km, and the scale height of dust set to 10 km in both the cloud and the atmosphere above. The 11-μm optical depth of the atmosphere above the cloud is fixed at 0.4, which is equal to the 11-μm atmospheric optical depth above 4 km in the middle of the parameter space for the dust storm surrounding models. The atmospheric temperature at the cloud top is taken as 220 K, and the atmospheric lapse rate is determined so that at 24 km the temperature is 200 K, to agree with the IRTM measurements at 15 μm [*Martin et al.*, 1979], as was done in the cloud-free area. Convergence is achieved to better than 1 K per channel, which is comparable to the experimental accuracy of the IRTM, for τ_{11c} less than 3. Solutions are less accurate for large cloud optical depth, and for $\tau_{11c} = 5$ the least squares solution for T_9 deviates by 5 K from the observed value.

In Figure 6, T_g below the dust cloud is less than 280 K for all choices of τ_{11c}. For the dust storm surroundings model, the absolute lower bound on T_g came to 286 K. The value of T_c is less certain, and depends strongly on the value of τ_{11c}. In particular, T_c varies most rapidly when $\tau_{11c} < 3$. This is explained as follows: from Figure 4, T_g is most sensitive to the value of T_7, and in Table 3, the ratio of extinction coefficients for the 7- and 11-μm bands is about one third. So when τ_{11c} is

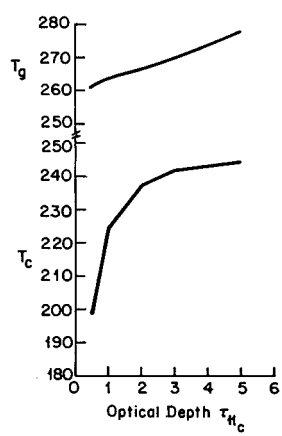


Figure 6: Dust storm interior model. Ground temperature (upper curve) and cloud temperature (lower curve) are plotted as functions of the 11-μm dust cloud optical depth. The dust storm top is fixed at 4 km, and the 11-μm optical depth above 4 km is set to 0.4. $T_{4\,km}$ is set to 220 K and $T_{24\,km} = 200$ K.

increased to 3, the 7-μm optical depth reaches unity, and the ratio of ground to atmospheric radiation becomes less dependent upon τ_{11c}. No specific conclusions can be drawn about the effective temperature of the dust cloud as related to the boundary layer temperature in the surroundings because of the large uncertainty in both these quantities. Theoretical modeling of the boundary layer temperature in the midafternoon at a similar latitude in the northern hemisphere, for early summer, yields temperatures near 230 K [*Pollack et al.*, 1979]. The increased visual optical depth (which is still less than 1) and the greater solar input in the dust storm surroundings over those of the VL1 site models can raise this temperature of the order of 10 K (Figures 7 and 8 in *Pollack et al.* [1979]. This places the current estimate of the near-ground temperature at the upper end of the range of possible values of T_c. From these results, note that $T_g - T_c$ is less than 40 K in the dust storm interior. These observations are discussed in the next section.

A number of perturbation studies have been performed on the results in Figure 6. In general, the least squares character of the method makes the solution less sensitive to changes in a single-input quantity than the single-point inversions of the previous sections. Varying $\tilde{\omega}_0$ and β by 10% causes T_g to change by less than 3 K, whereas T_c will change less than 6 K. Altering the model brightness temperatures T_9, T_{11}, and T_{20}, independently by up to 3 K affects T_g less than 1 K, and T_c less than 2 K. If T_7 is perturbed while τ_{11c} is less than 3, the change in T_g will be comparable to the size of the perturbation. For $\tau_{11c} > 3$, altering T_7 or the 7-μm extinction coefficient will have an increasingly large effect on T_g. For $\tau_{11c} > 5$, the input parameters do not produce a satisfactory fit to the observed radiances.

In summary, an attempt has been made to set meaningful limits on T_g and T_c based on the assumption that particle properties do not vary between the dust storm surrounding and the interior. T_c is not well constrained by the available data, but it is probably in the lower range of the likely near-ground temperatures in the surroundings. Below the storm, T_g must be lower than the ground temperature in the surroundings.

7. Discussion

One of the key questions that may be addressed by studying the thermal structure of a Martian local dust storm is whether local thermodynamic effects play an important role in generating the surface shear stress necessary to lift dust. Small, intense convective cells occur in the form of dust devils on Earth, in places where the planetary boundary layer is statically unstable over regions tens to hundreds of meters wide [*Ryan and Carroll*, 1970]. Similar features were found on Mars in the Viking orbiter camera images [*Thomas and Gierasch*, 1985]. A surface resolution about 3 orders of magnitude greater than that available in the data of this study would be required to observe these features with infrared imaging. However, the results of the present work suggest that at least over surfaces on the scale of 200 km, vigorous convection at ground level is unlikely in the interior of the dust storm. For the storm studied here, the ground temperature in the storm interior is at least 6 K cooler than the ground temperature of the surroundings. Also, thermal contrast between the ground and the atmospheric boundary layer ($T_g - T_o$) in the storm surroundings at midafternoon is 30 to 60 K over most of the parameter space. From Figure 4, the dust storm interior shows this difference to be less than 40 K for τ_{11c} between 1 and 3, suggesting that the boundary layer in the cloud region may be more stable against convection than in the surroundings. Both these observations are contrary to expectation for a situation in which direct heating at ground level predominates in the storm interior.

The evolution of the storm as shown in Figure 2 makes the possibility that the convective process operates even at the storm edges appear unlikely. The thermal structure of the storm over regions 200 km in size does not change measurably between 11.5 and 16.6 hours local time. If the storm were generated by direct thermal forcing at the ground during part of this period, the static stability would be low at first, and vigorous convection would transport heat and carry dust from the ground [*Gierasch and Goody*, 1973]. In this scenario, the static stability would first decrease in the region of cloud formation as the atmosphere was heated by direct absorption of solar radiation by the dust near the ground. Later, as the dust filled the boundary layer, increased atmospheric heating would cause the static stability to increase. The calculations of this study do not rule out the presence of a region of vigorous convection at the cloud tops, nor the possibility of dust plumes or dust devils on a smaller spatial scale. The observations presented here are more favorable to arguments that the storm is maintained either by the flow of cold air in the form of density currents, or the combined effects of winds forced on scales larger than the size of the storm [*Leovy et al.*, 1973; *Pollack et al.*, 1979]. These two hypotheses may be tested when the direction of storm motion can be compared with regional slopes and prevailing winds. Both mechanisms would allow the ground temperature to be relatively cool in the region where dust is being lifted, and neither mechanism is excluded by the apparently unchanging vertical thermal gradient in the boundary layer within the dust storm during the late morning and afternoon.

This study of the Viking IRTM data in the area of one local dust storm led to the conclusion that the dust in the cloud region differs from that observed during the 1971-1972 global dust storm observed by Mariner 9. This may be due to differences in the type of dust distributed in regions of the planet, or to selectivity in the mechanisms which lift dust in the two instances. Given the assumptions that the particle scattering (but not extinction) properties are similar to those observed by Mariner 9, and that dust properties do not vary horizontally in the storm region, it is possible to obtain an estimate of the ground temperature and to achieve a crude determination of the atmospheric temperature near the ground within a local dust storm. This technique may be useful in further exploring the thermodynamics of Martian dust storms when a more extensive data set is available.

Appendix: Two-Stream Radiative Transfer Scheme

Following *Toon et al.* [1977], the equation of radiative transfer in a plane parallel atmosphere is written for monochromatic radiation, given the positive direction as upward and optical depth τ as zero at the top:

$$\mu\, dI(\tau,\mu,\theta)/d\tau = I(\tau,\mu,\theta) - J(\tau,\mu,\theta) \qquad \text{(A1)}$$

where I is the intensity, μ is the cosine of the emission angle, θ is the scattering angle, and J is the source function. In the two-stream approximation,

$$\frac{1}{\sqrt{3}}\, df/d\tau = f - J^{+} \qquad \text{(A2)}$$

$$-\frac{1}{\sqrt{3}}\, dg/d\tau = g - J^{-} \qquad \text{(A3)}$$

Here f and g are the upward and downward two-stream intensities, respectively.

$$J^{+} = \varpi_o/2\,(1+\beta)\,f + \varpi_o/2\,(1-\beta)\,g + (1-\varpi_o)B_\nu(\tau) \qquad \text{(A4)}$$

With $B_\nu(\tau)$ as the Planck function at frequency ν for the temperature at atmospheric level τ, and

$$J^{-} = \varpi_o/2\,(1+\beta)\,g + \varpi_o/2\,(1-\beta)\,f + (1-\varpi_o)B_\nu(\tau) \qquad \text{(A5)}$$

Rewriting equations (A2) and (A3) using (A4) and (A5),

$$f' = af - bg - c \qquad \text{(A6)}$$

$$g' = -ag + bf + c \qquad \text{(A7)}$$

where:

$$\begin{aligned} a &= \sqrt{3}\left(1-\varpi_o/2\,(1+\beta)\right) \\ b &= \sqrt{3}/2\,\varpi_o(1-\beta) \\ c &= \sqrt{3}(1-\varpi_o)B_\nu(\tau) \end{aligned} \qquad \text{(A8)}$$

ϖ_0 and β are the single-scattering properties discussed in the text. By cross-differentiating equations (A6) and (A7), a second-order differential equation for f is obtained:

$$f'' + (b'/b)\,f' + \left(\left(b^2 - a^2\right) + (a/b)\,b' - a'\right) f = (c/b)\,b' - (a+b)\,c - c'$$

where primes are derivatives with respect to τ. When the particle properties are constant over the whole atmosphere, this equation reduces to

$$f' + \left(b^2 - a^2\right) f = -(a+b)\,c - c' \qquad \text{(A9)}$$

The lower boundary condition is

$$f(\tau^*) = \varepsilon B_\nu(T_g) \qquad \text{(A10)}$$

where τ^* is the total atmospheric optical depth and ε is the ground emissivity. The boundary condition at the top of the atmosphere is chosen as

$$g(0) = 0 \qquad \text{(A11)}$$

Using (A2), the upper boundary condition becomes

$$f'(0) - af + c = 0 \qquad \text{(A12)}$$

The solution to (A9), with (A10) and (A12), is obtained using the method of *Lindzen and Kuo* [1969]. Then g is obtained as an initial value problem using a Runge-Kutta method with (A11) as the initial value. The solution to the equation of radiative transfer for the upward intensity is

$$f(0) = \varepsilon B_\nu(T_g) e^{-\frac{\tau^*}{\mu}} + \int_0^{\tau^*} e^{-\frac{t}{\mu}}\, J(t)\, dt/\mu \qquad \text{(A13)}$$

where $J(t)$ is the source function. $J(t)$ is replaced with the source function in the upward direction J^+ from equation (A4), and the integration is performed numerically to obtain the emergent intensity. Contributions from atmospheric gases are neglected. The CO_2 opacity averaged over the 7-, 9-, 11-, or 20-μm IRTM passbands is at least 3 orders of magnitude less than the observed optical depth in these bands for a 7-mbar atmosphere.

Acknowledgments. I thank R. Goody and R. Zurek for their encouragement of this work, F. Palluconi for providing initial access to the IRTM data used in this study, and T.Z. Martin and F. Palluconi for careful reviews of the manuscript. The early part of this work is supported in part by NASA grant NSG 7398 to Harvard University. Later portions of this project were performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration, through the NASA Planetary Atmospheres Program.

References

Binder, A.B., R.E. Arvidson, E.A. Guinness, K.L. Jones, E.C. Morris, T.A. Mutch, D.C. Pieri, and C. Sagan, The geology of the Viking Lander 1 site, *J. Geophys. Res., 82*, 4439-4451, 1977.

Briggs, G.A., W.A. Baum, and J. Barnes, Viking orbiter imaging observations of dust in the Martian atmosphere, *J. Geophys. Res., 84*, 2795-2820, 1979.

Christensen, P.R., Martian dust mantling and surface composition: Interpretation of thermophysical properties, *J. Geophys. Res., 87*, 9985-9998, 1982.

Clancy, R.T., and S.W. Lee, A new look at dust and clouds in the Mars atmosphere: Analysis of emission-phase-function sequences from global Viking IRTM observations, *Icarus , 93*, 135-158, 1991.

Conel, J.E., Infrared emissivities of silicates: Experimental results and a cloudy atmosphere model of spectral emission from condensed particulate mediums, *J. Geophys. Res., 74*, 1614-1634, 1969.

Gierasch, P.J., and R.M. Goody, A study of the thermal and dynamical structure of the Martian lower atmosphere, *Planet. Space Sci.*, 16, 615-646, 1968.

Gierasch, P.J., and R.M. Goody, The effect of dust on the temperature of the Martian atmosphere, *J. Atmos. Sci., 29*, 400-402, 1972.

Gierasch, P.J., and R.M. Goody, A model of a Martian great dust storm, *J. Atmos. Sci., 30*, 169-179, 1973.

Greeley, R., N. Lancaster, S. Lee, and P. Thomas, Martian aeolian processes, sediments, and features, in *MARS*, edited by H.H. Kieffer et al., pp. 730-766, University of Arizona Press, Tucson, 1992.

Hess, S.L., R.M. Henry, C.B. Leovy, J.A. Ryan, and J.E. Tillman, Meteorological results from the surface of Mars: Viking 1 and 2, *J. Geophys. Res., 82*, 4559-4574, 1977.

Hovis, W.A., and W.R. Callahan, Infrared reflectance spectra of igneous rocks, tuffs, and red sandstone from 0.5 to 22 microns, *J. Opt. Soc. Am., 56*, 639-643, 1966.

Hunt, G.E., Thermal infrared properties of the Martian atmosphere, 4, Predictions of the presence of dust and ice clouds from Viking IRTM spectral measurements, *J. Geophys. Res., 84*, 2865-2874, 1979.

Jakosky, B.M., The effects of nonideal surfaces on the derived thermal properties of Mars, *J. Geophys. Res., 84*, 8252-8262, 1979.

Kahn, R., R.M. Goody, and J.B. Pollack, The Martian twilight, *J. Geophys. Res., 86*, 5833-5838, 1981.

Kahn, R.A., T.Z. Martin, R.W. Zurek, and S.W. Lee, The Martian dust cycle, in *MARS*, edited by H.H. Keiffer et al., pp. 1017-1053, University of Arizona Press, Tucson, 1992.

Kieffer, H.H., G. Neugebauer, G. Munch, S.C. Chase Jr., and E. Miner, Infrared thermal mapping experiment: The Viking Mars orbiter, *Icarus, 16*, 47-56, 1972.

Kieffer, H.H., T.Z. Martin, A.R. Peterfreund, B.M. Jakosky, E.D. Miner, and F.D. Palluconi, Thermal and albedo mapping of Mars during the Viking primary mission, *J. Geophys. Res., 82*, 4249-4291, 1977.

Leovy, C.B., R.W. Zurek, and J.B. Pollack, Mechanisms for Mars dust storms, *J. Atmos. Sci., 30*, 749-762, 1973.

Lindzen, R.S., and H.L. Kuo, A reliable method for the numerical integration of a large class of ordinary and partial differential equations. *Mon. Weather Rev., 97*, 732-734, 1969.

Martin, T.Z., Thermal infrared opacity of the Martian atmosphere, *Icarus, 66*, 2-21, 1986.

Martin, T.Z., and H.H. Kieffer, Thermal infrared properties of the Martian atmosphere, 2, The 15-micron band measurements, *J. Geophys. Res., 84*, 2843-2852, 1979.

Martin, T.Z., A.R. Peterfreund, E.D. Miner, H.H. Kieffer, and G.E. Hunt, Thermal infrared properties of the Martian atmosphere, 1, Global behavior at 7, 9, 11, and 20 microns, *J. Geophys. Res., 84*, 2830-2842, 1979.

Moore, H.J., R.E. Hutton, R.F. Scott, C.R. Spitzer, and R.W. Shorthill, Surface materials of the Viking Lander sites, *J. Geophys. Res., 82*, 4497-4523, 1977.

Peterfreund, A.R., and H.H. Kieffer, Thermal infrared properties of the Martian atmosphere, 3, Local dust storms, *J. Geophys. Res., 84*, 2853-2863, 1979.

Pollack, J.B., D.S. Colburn, F.M. Flasar, R. Kahn, C.E. Carlston, and D. Pidek, Properties and effects of dust particles suspended in the Martian atmosphere, *J. Geophys. Res., 84*, 2929-2945, 1979.

Rogers, C.D., Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation, *Rev. Geophys., 14*, 609-624, 1976.

Ryan, J.A., and J.J. Carroll, Dust devil wind velocities: Mature state, *J. Geophys. Res., 75*, 531-541, 1970.

Ryan, J.A., and R.D. Lucich, Possible dust devils, vortices on Mars, *J. Geophys. Res., 88*, 11005-11011, 1983.

Sieff, A., and D.B. Kirk, Structure of the atmosphere of Mars on summer at midlatitudes, *J. Geophys. Res., 82*, 4364-4378, 1977.

Thomas, P.C., and P.J. Gierasch, Dust devils on Mars, *Science, 230*, 175-177, 1985.

Toon, O.B., J.B. Pollack, and C. Sagan, Physical properties of the particles composing the Martian dust storm of 1971-72, *Icarus, 30*, 663-696, 1977.

Zurek, R.W., Martian great dust storms: An update, *Icarus, 50*, 288-310, 1982.

Zurek, R.W., J.R. Barnes, R.M. Haberle, J.B. Pollack, J.E. Tillman, and C.B. Leovy, Dynamics of the Martian atmosphere, in *MARS*, edited by H.H. Keiffer et al., pp. 835-933, University of Arizona Press, Tucson, 1992.

R. Kahn, Jet Propulsion Laboratory, MS 169-237, 4800 Oak Grove Drive, Pasadena, CA 91109.

(Received March 31, 1994; revised October 23, 1994; accepted October 25, 1994.)

Page intentionally left blank

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5277–5284, MARCH 25, 1995

Atmospheric effects on the mapping of Martian thermal inertia and thermally derived albedo

Joan N. Hayashi
Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder

Bruce M. Jakosky
Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder
and Department of Geological Sciences, University of Colorado, Boulder

Robert M. Haberle
Space Sciences Division, NASA Ames Research Center, Moffett Field, California

Abstract. We examine the effects of a dusty CO_2 atmosphere on the thermal inertia and thermally derived albedo of Mars and we present a new map of thermal inertias. This new map was produced using a coupled surface atmosphere (CSA) model, dust opacities from Viking infrared thermal mapper (IRTM) data, and CO_2 columns based on topography. The CSA model thermal inertias are smaller than the 2% model thermal inertias, with the difference largest at large thermal inertia. Although the difference between the thermal inertias obtained with the two models is moderate for much of the region studied, it is largest in regions of either high dust opacity or of topographic lows, including the Viking Lander 1 site and some geologically interesting regions. The CSA model thermally derived albedos do not accurately predict the IRTM measured albedos and are very similar to the thermally derived albedos obtained with models making the 2% assumption.

Introduction

The thermal inertia of a material is a measure of its temperature response to energy input or loss; the temperature of a low thermal inertia material responds quickly, while the temperature of a high thermal inertia material responds slowly. Understanding thermal inertias has implications for understanding surface properties and surface temperatures. For silicate materials, thermal inertia is principally determined by the thermal conductivity and can be used to infer particle sizes at the surface. Also, a knowledge of surface temperatures is necessary to understand the formation of water and carbon dioxide frost and ice which, in turn, influences both the geology and the climate.

The most widely used thermal inertia data for Mars assumes an atmosphere whose only effect is a contribution to the downgoing radiation which is constant and equal to 2% of the maximum solar insolation (the "2% assumption") [*Kieffer* et al., 1977; *Palluconi and Kieffer*, 1981]. *Haberle and Jakosky* [1991] investigated the differences in the thermal inertia derived by making the 2% assumption versus that derived by including the effect of a dusty CO_2 atmosphere and sensible heat exchange with the surface. Specifically, they explored this effect for the location of the Viking Lander 1 (VL1). *Bridges* [1994] did first-order corrections for the effects of elevation on determining the thermal inertia and the effect of variations in pressure on translating thermal inertias into particle sizes. Bridges used both the 2% model and elements of *Haberle and Jakosky*'s [1991] analysis to obtain two particle size maps. To correct the 2% model thermal inertias, Bridges assumed that at 0 km the back radiation was 2% of the maximum solar insolation, but that this decayed with elevation with a scale height of 10 km. To incorporate the results of Haberle and Jakosky, Bridges first assumed that the back radiation was 4% of the maximum solar insolation at an elevation of -2 km, based on Haberle and Jakosky's result for a CO_2 atmosphere with no dust at the VL1 site. He then assumed that both this 4% back radiation and the dust opacity (fixed at 0.4 at -2 km) decayed with elevation with a scale height of 10 km. He then interpreted the resulting thermal inertias in terms of particles sizes, accounting for the variation of conductivity, and thus thermal inertia, with pressure.

In this paper we utilize some aspects of both the *Haberle and Jakosky* [1991] and the *Bridges* [1994] analyses. We use the coupled surface-atmosphere (CSA) model presented by *Haberle and Jakosky* [1991] to investigate the effects of a dusty CO_2 atmosphere on the thermally derived albedo. The thermally derived albedo is the surface albedo which, together with the thermal inertia, provides model surface temperatures which best match the observed temperatures; it may differ from a measured albedo. We also present a new map of thermal inertia obtained utilizing the CSA model. This map differs from *Bridges*' [1994] thermal inertias, as

Paper number 94JE02449.
0148-0227/94/94JE-02449$05.00

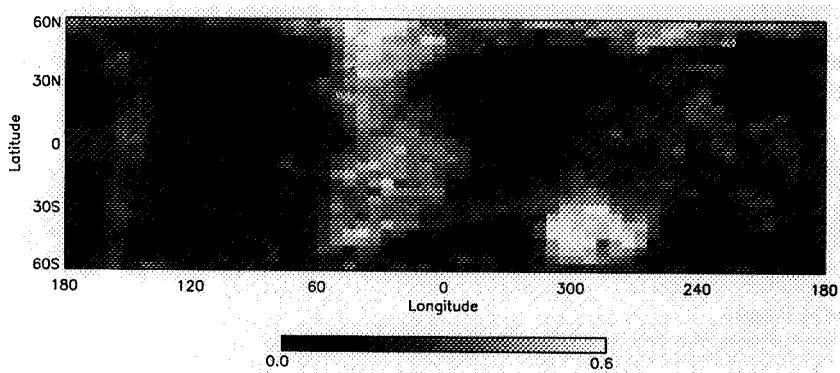


Figure 1a. Map of the time averaged visible dust opacity for the period L_s = 345° - 125°.

the effect of the atmosphere is calculated with a one-dimensional radiative-convective model rather than scaled from the maximum solar insolation. However, in interpreting the thermal inertias for certain regions of Mars in terms of particle sizes, we utilize results from *Bridges* [1994] to account for elevation dependent pressure variations

Method

The CSA model is a one-dimensional diurnal model; it is described in detail by *Haberle and Jakosky* [1991]. The atmospheric portion of the model is a radiative-convective model which allows for absorption and emittance of radiation by CO_2 and dust, convective adjustment of the lower atmosphere, and sensible heat exchange between the surface and atmosphere. The surface and subsurface portion of the model allows for thermal diffusion and exchange of energy with the atmosphere. The CSA model parameters which were varied in this study were thermal inertia, surface albedo, latitude, atmospheric pressure at the surface, and atmospheric dust opacity. In contrast, the parameters which were varied in a diurnal model making the 2% assumption (the 2% model) were thermal inertia, surface albedo, and latitude; the 2% model temperatures do not depend on atmospheric surface pressure or dust opacity.

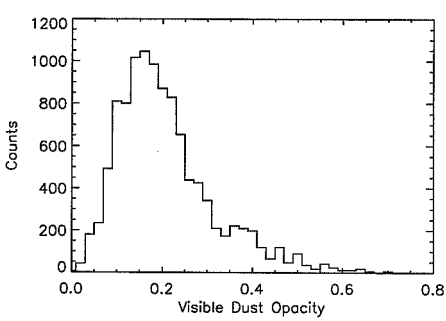


Figure 1b. Histogram of the time averaged visible dust opacity.

Two avenues of investigation were used to examine the differences between the 2% model and the CSA model. First, the differences in both the thermal inertias and thermally derived albedos due to varying dust opacities were explored at a single latitude and surface pressure. Second, the differences in the models' mapped thermal inertias and thermally derived albedos due to spatial variations in latitude, dust opacities, and surface pressure were explored.

First, the differences between the CSA model and the 2% model at a single latitude and surface pressure were investigated for the latitude and pressure of the VL1 site. *Haberle and Jakosky* [1991] examined the difference in thermal inertia between the two models; we explored the difference in thermally derived albedo and also included the differences in thermal inertia for completeness. The 2% model was run at a series of albedos and thermal inertias to obtain minimum and maximum surface temperatures, which were then fit using the CSA model to derive inertia and albedo. To examine the effects on thermally derived albedo, the CSA model was run at a given thermal inertia and various albedos and dust opacities. Minimum and maximum surface temperatures were obtained for each CSA model combination of albedo and dust opacity. The 2% model minimum and maximum temperatures were then interpolated to fit these CSA model temperatures and determine the 2% model albedo which best corresponded to the CSA model albedo and dust opacity. To examine the effects on thermal inertia a similar procedure was used, except the CSA model temperatures were obtained for a given albedo and various thermal inertias and dust opacities.

Second, to examine the spatial differences between the 2% model and the CSA model thermal inertias and thermally derived albedos, it was necessary to first produce maps of the CSA model thermal inertias and thermal inertias. As the *Palluconi and Kieffer* [1981] 2%

Table 1. Thermal Inertias and Thermally Derived Albedos Calculated Using the Temperature at the Surface (I_S^*, A_S^*) and at the Top of the Atmosphere (I^*, A^*) for Two Different Pairs of I and A

	I=83.72, A_S=0.25	I=272, A_S=0.25
I^*	182	431
I_S^*	170	412
A^*	0.19	0.18
A_S^*	0.18	0.18

From *Paige et al.* [1994].

thermal inertias and thermally derived albedos were mapped with each map point or bin having a resolution of 2° latitude by 2° longitude for latitudes between ±60°, we chose to map the CSA model thermal inertias and thermally derived albedos at the same resolution.

The latitude, surface pressure, dust opacity, and minimum and maximum temperatures were needed for each map point to determine the CSA model thermal inertia and thermally derived albedo. Surface temperatures were obtained by running the 2% model for each map point using the *Palluconi and Kieffer* [1981] thermal inertias and thermally derived albedos at L_S=0°. As the *Palluconi and Kieffer* [1981] thermal inertias and thermally derived albedos were chosen to best reproduce the observed 20-μm brightness temperatures, using them in the 2% model should reasonably reproduce the actual temperatures. Note that Palluconi and Kieffer did not publish their map of the thermally derived albedo; we obtained this map directly from them.

The pressure for each bin was determined from the digital topographic maps, assuming 6.1-mbar pressure at 0-km elevation and a constant scale height of 10 km. The pressure of 6.1 mbar is the nominal average pressure at 0 km, and is the constant pressure surface to which the elevations are referenced in the topographic map [*U.S. Geologial Survey* (USGS), 1989]. Seasonal pressure variations are ignored, as they are of the same order as the pressure variations introduced by the uncertainties in the topography (±1.0 to ±1.5 km, depending on the latitude [*USGS*, 1989]).

The dust opacity was approximated by averaging *Martin and Richardson's* [1993] 9-μm dust opacities for the period between L_S=345°-125° (Figure 1). The period L_S=345°-125° was the same period which Palluconi and Kieffer used to obtain their thermal inertias and thermally derived albedos. The visible dust opacities were obtained by assuming the generally accepted 9-μm to visible dust opacity ratio of 0.5. Discussions regarding the infrared to visible opacity ratios may be found in the works by *Zurek* [1982], *Martin* [1986], *Haberle and Jakosky* [1991], and *Clancy et al.* [this issue].

The CSA model was run at L_S=0° for a variety of combinations of latitudes, albedos, thermal inertias, pressures, and dust opacities to produce a multidimensional array of minimum and maximum temperatures. At each map point the CSA model temperatures for the grid of albedos and thermal inertias were interpolated to the latitude, pressure, and dust opacity at that point. Next were found the three combinations of CSA model albedos and thermal inertias which had minimum and maximum temperatures lying closest to the 2% model minimum and maximum temperatures for each bin. The 2% model temperatures for each bin were then used to interpolate between those three points and obtain the CSA model thermal inertia and thermally derived albedo.

Error Analysis

There are several sources of error, and we can estimate the magnitude of error in the CSA model by several different means.

First, in our analysis we exclude those points with a CSA model thermally derived albedo less than 0.08 or greater than 0.36, as those albedos lie outside of the interpolation region. The number of bins which are thus excluded is less than 2% of the total bins. As such a small percentage of bins is excluded, we consider that this has a negligible contribution to the error.

Errors arising from uncertainties in the surface pressure due to seasonal pressure variations and uncertainties in the topography are small in both the thermally derived albedo and the thermal inertia. Model comparisons using different pressures suggest the errors are less than approximately 0.01 in albedo and less than approximately 10 W $s^{1/2}$ m^{-2} K^{-1} in thermal inertia. We use SI units for thermal inertia; however, for brevity, we hereafter omit the units. To convert to units of 10^{-3} cal cm^{-2} $s^{-1/2}$ K^{-1}, divide by 41.86. Allowing the infrared to visible dust opacity ratio to vary from 0.4 to 0.55 gives rise to small uncertainties in albedo (< 0.01) and thermal inertia (< 5).

Errors are also introduced into the CSA thermal inertia and thermally derived albedo due to fitting temperatures as would be observed at the top of the atmosphere (20-μm infrared thermal mapper (IRTM) brightness temperatures) with CSA model surface temperatures. The CSA model included the absorption and emission of radiation by dust and CO_2 in the atmosphere and their effects on the surface temperature; however, the CSA model did not calculate the effects of this absorption and emission on the observed brightness temperature. *Paige et al.* [1994] examined this problem. At latitudes of 60° - 85° and an L_S = 110°, they calculated both surface temperatures and temperatures at the top of the atmosphere for several different combinations of atmo-

Table 2. Relevant Portion of the Wentworth Grain Size Classification

Classification	Grain Size, μm
Coarse silt	31.25-62.5
Very fine sand	62.5-125
Fine sand	125-250
Medium sand	250-500
Coarse sand	500-1000
Very coarse sand	1000-2000
Granules	2000-4000

After *Wentworth* [1922].

spheric conditions, surface thermal inertia (I), and surface albedo(A_S). They then fit the temperatures at the top of the atmosphere (20-μm brightness temperatures) with their basic thermal model, which did not include any atmospheric contribution, to obtain I^* and A^*. They also fit the surface temperatures with their basic thermal model to obtain I_S^* and A_S^*. Thus the difference between I and A_S and I^* and A^* is the difference between having both included atmospheric effects and calculating the brightness temperature correctly or having ignored atmospheric effects. The difference between I and A_S and I_S^* and A_S^* is the difference between having included atmospheric effects to calculate the surface temperature, but not calculating the brightness temperature correctly, or having ignored atmospheric effects. To estimate the error introduced between including atmospheric effects and calculating the brightness temperature or including atmospheric effects but only calculating the surface temperature, we compare the differences between *Paige et al.*'s I^* and A^* and I_S^* and A_S^*. Table 1 summarizes their results for a latitude of 60° and their atmosphere 5. Atmosphere 5 was chosen, as it most closely resembled the CSA model atmosphere; atmosphere 5 had a visible dust opacity of 0.2 as well as an infrared to visible dust opacity ratio of ~0.5. The difference between I^* and I_S^* is 10 to 20 and the difference between A^* and A_S^* is 0.0 to 0.01. Therefore, the error introduced by utilizing surface temperature rather than the brightness temperature at the top of the atmosphere to derive thermal inertia and albedo is small when the visible dust opacity is 0.2 or less. Based on the distribution of dust opacities shown in Figure 1b, we conclude that the error in the thermal inertia introduced by ignoring the effects of the dust on the upgoing radiation is usually less than 20 and is almost always less than 40.

Another source of error is that *Palluconi and Kieffer* [1981] fit an annual thermal model to the IRTM measurements from L_S=344°-125°, while we used their thermal inertia and thermally derived albedo in a diurnal 2% model to calculate temperatures which were matched by the diurnal CSA model temperatures. To estimate the errors due to running the models at only one L_S (L_S=0°), we also ran the 2% and CSA models at another L_S (L_S=55°). At southern high latitudes, CO_2 frost formed on the surface, and we therefore ignored the bins between -39° and -59° latitude (approximately 20% of our total bins) in this analysis. The difference between the CSA model thermally derived albedo at L_S=0° and L_S=55° had a mean of 0.01 and a standard deviation of 0.03. The difference between the CSA model thermal inertia at L_S=0° and L_S=55° had a mean of -0.7 and a standard deviation of 8.7. The error introduced by using a single L_S is small for the following reason. *Palluconi and Kieffer* fit an annual model (which included seasonal effects) to real temperatures (which obviously include seasonal effects); however, we then compared results from two diurnal models, and thus the lack of seasonal effects very nearly canceled out.

Finally, an error is introduced by assuming the surface emissivity is always unity. As discussed by *Christensen* [1982], the thermal emissivity is correlated with albedo. At albedos greater than 0.28 the emissivity is unity, but at lower albedos the emissivities can approach 0.9. In the case where the actual emissivity is 0.9, assuming a model emissivity of 1.0 results in albedo being underestimated by about 0.1 and thermal inertia being overestimated by up to 40.

Based on the above considerations, we estimate the errors in the CSA model thermal inertia are of the order of 10-20 for regions of high albedo (≥ 0.28) and may be as much as 50-60 for regions of low albedo (≤ 0.28).

There is also an additional uncertainty which is introduced in translating thermal inertias into particle

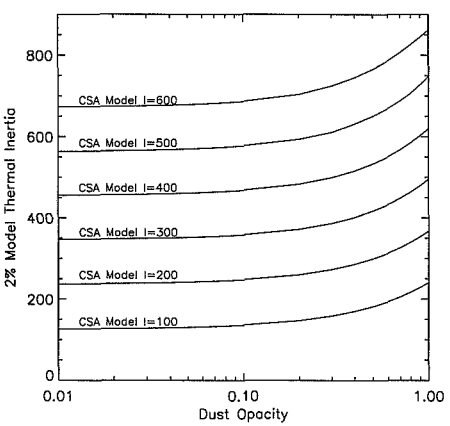


Figure 2. Contours of constant CSA model thermal inertias plotted on dust opacity versus 2% model thermal inertia.

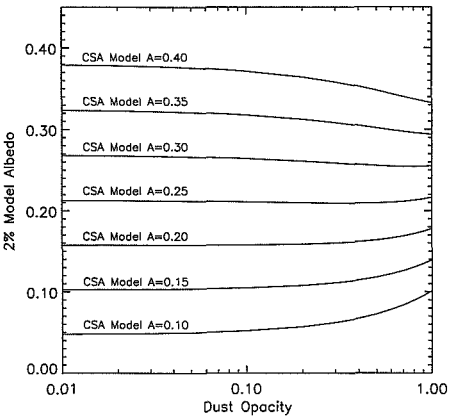


Figure 3. Contours of constant CSA model thermally derived albedos plotted on dust opacity versus 2% model thermally derived albedo.

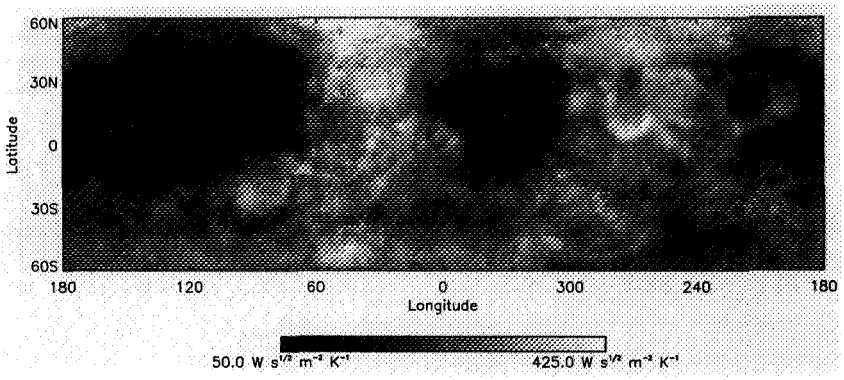


Figure 4. Map of the CSA model thermal inertia.

size. For geologic materials, the thermal inertia is principally determined by the material's conductivity; the conductivity is in turn a function of particle size and pressure. Obviously, a single particle size is not an actual representation of the surface. The particle size inferred from the thermal inertia is an effective particle size, which assumes that the surface is composed of uniform particles. Also, although some data exist on the variation of conductivity with pressure [*Masamune and Smith*, 1963; *Wechsler and Glaser*, 1965], the scatter within a particular data set as well as scatter between data sets adds additional uncertainties in converting thermal inertias into particle sizes. Due to these uncertainties, we have chosen to interpret particle sizes corresponding to the thermal inertias in terms of *Wentworth's* [1922] grain size classification (see Table 2) rather than stating specific particle sizes.

Results

In this section we first present the effects of varying dust opacities on the thermal inertia and albedo at a single latitude and surface pressure. Next, we explore the differences in the CSA and 2% model thermal inertias and thermally derived albedos due to spatial variations in latitude, dust opacities, and surface pressure.

The effect of varying dust opacity on the CSA model thermal inertia versus the 2% model thermal inertia is shown in Figure 2, which is similar to *Haberle and Jakosky's* [1991] Figure 13. This figure shows the difference between the 2% model-derived thermal inertias and the CSA model-derived thermal inertias for a variety of dust opacities and a CSA model albedo of 0.30. The 2% model thermal inertia is always larger than the CSA model thermal inertia, with the difference being greatest at largest dust opacities. Therefore, particle sizes inferred from the 2% model thermal inertias are always too large.

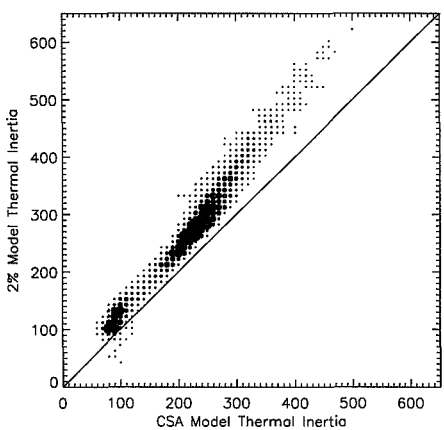


Figure 5. Two-dimensional histogram of CSA model thermal inertia versus 2% model thermal inertia. The larger the circle, the more map bins which have the corresponding 2% model thermally derived albedo and measured albedo.

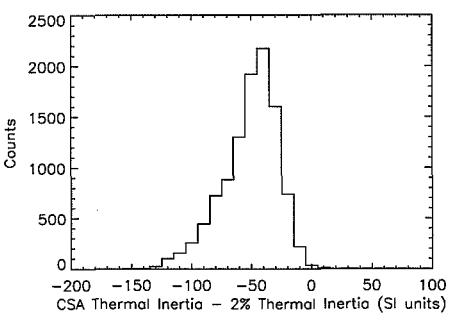


Figure 6. Histogram of CSA model thermal inertia - 2% model thermal inertia.

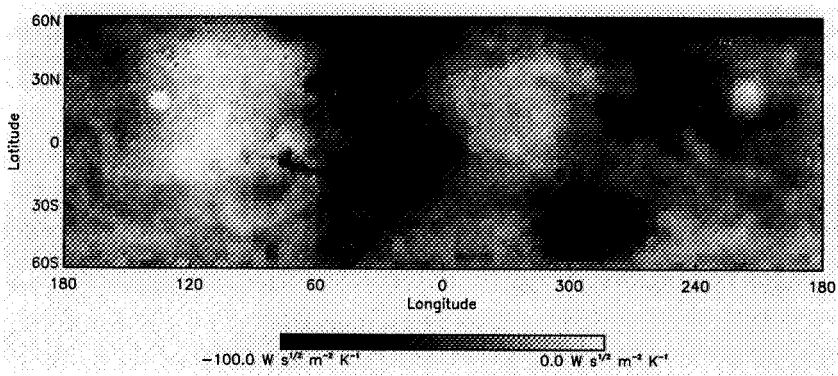


Figure 7. Map of CSA model thermal inertia - 2% model thermal inertia.

Figure 3 is a similar figure for the thermally derived albedo. This figure shows the difference between the 2% model thermally derived albedos and the CSA model thermally derived albedos for a variety of dust opacities and a CSA model thermal inertia of 200. At low dust opacities the 2% model thermally derived albedos are always smaller than the CSA model thermally derived albedos. As there is little dust present at such low opacities, this is attributed solely to the greenhouse effect of the CO_2 in the atmosphere. At high dust opacities, the behavior is different at high and low albedos. At low albedo and increasing dust opacity, a constant CSA thermally derived albedo is matched by a larger 2% model thermally derived albedo. This is attributed to scattering by dust in the atmosphere. However, at high albedo and increasing dust opacity, a constant CSA thermally derived albedo is matched by a smaller 2% model thermally derived albedo. This is attributed to warming of the atmosphere due to the presence of dust.

Next, we look at the spatial differences between the CSA and the 2% model thermal inertias and the significance of the CSA model-derived thermal inertias. Figure 4 is a map of the CSA model-derived thermal inertias. Figure 5 shows the relationship between the CSA model and the 2% model thermal inertias as a two-dimensional histogram. The 2% model thermal inertias are larger, and the difference between the two is greatest at large thermal inertia. Figure 6 is a histogram of the difference between the two model thermal inertias. The difference has a mean of -46 and a standard deviation of 23.

A map of the differences between the CSA model and the 2% model thermal inertias is presented in Figure 7.

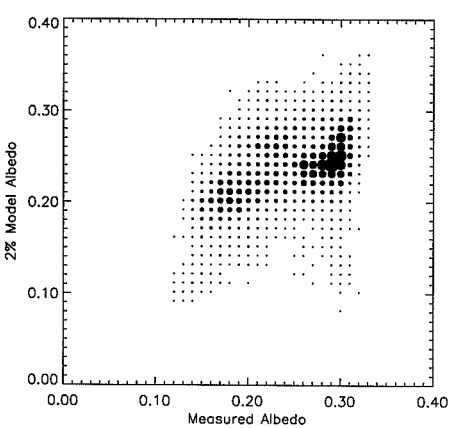


Figure 8. Two-dimensional histogram of IRTM measured albedo versus 2% model thermally derived albedo.

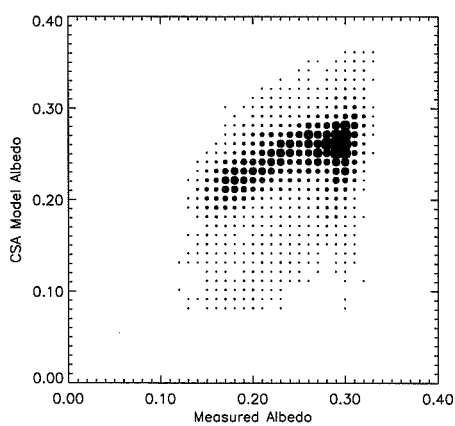


Figure 9. Two-dimensional histogram of IRTM measured albedo versus CSA model thermally derived albedo.

Areas with a high dust opacity (e.g., Acidalia, Chryse, and Hellas Planitias) have some of the largest differences; the CSA model thermal inertias are about 100 smaller than the 2% model thermal inertias for these areas. Some regions which are topographic lows (e.g., Valles Marineris, Vastitas Borealis, and Isidis Planitia) also have large differences (~-80 to -100) between the two model thermal inertias. Therefore, particle sizes for these regions based on the 2% model thermal inertias would be too large. For example, the 2% model thermal inertia in Hellas is about 340, while the CSA model thermal inertia is about 230. As Hellas has a high albedo (~0.33), the emissivity should be unity and the uncertainties are only about 20. From *Bridges'* [1994] Figure 2, which graphically shows the relationship between pressure, thermal inertia, and particle size, and using *Wentworth's* [1922] grain size classification (Table 2), the 2% model thermal inertia would correspond to medium sand, while the CSA model thermal inertia would correspond to fine sand; the CSA grain sizes are about 40% smaller than the 2% grain sizes. However, Acidalia Planitia, which has a 2% model thermal inertia of about 440 and a CSA model thermal inertia of about 350, has an albedo of about 0.16. Due to this low albedo and presumably low emissivity, the thermal inertias may be overestimated by 60. Thermal inertias of 440 and 350 would correspond to coarse sand and medium sand, respectively, but if the thermal inertias were less by 60, the 2% thermal inertia would correspond to grain sizes near the boundary between coarse sand and medium sand, while the CSA thermal inertia would correspond to medium sand. In both cases the CSA grain sizes are about 25% smaller than the 2% grain sizes.

Some areas of topographic highs (e.g., Olympus Mons, the Tharsis Volcanoes, Alba Patera, and Elsium Mons) have CSA model thermal inertias which are very similar to or larger than the 2% model thermal inertias. These regions have differences of -5 to +60 and appear as bright regions on Figure 7. For example, the 2% model thermal inertias at Ascraeus Mons are about 95, while the CSA model thermal inertias are about 110. The albedo is about 0.29 at Ascraeus Mons, so the uncertainties in thermal inertia should be less than 20. These thermal inertias correspond to fine sand; the CSA grain sizes are 25% smaller than the 2% grain sizes.

The Viking Lander sites are perhaps the most studied on Mars due to additional data provided by the landers which is not available for other regions. Both lander sites have high albedos, presumably near-unit emissivity, and uncertainties in thermal inertia of about 20. The VL1 site has a 2% model thermal inertia of 347 and a CSA model thermal inertia of 266, with a difference of -81. Both these thermal inertias correspond to medium sand, with the CSA grain sizes again about 25% smaller. The Viking Lander 2 (VL2) site has a difference between the model thermal inertias of -57, which is close to the mean difference. The 2% model thermal inertia is 356, and the CSA model thermal inertia for the site is 299. The 2% thermal inertia corresponds to grain sizes close to the boundary between medium and coarse sand, while the CSA thermal inertia corresponds to medium sand; the CSA grain sizes are about 25% smaller. Although the model thermal inertias for VL2 have a difference close to the mean, it lies within a few degrees of some of the largest differences between the model thermal inertias.

Thus, the importance of utilizing a CSA model thermal inertia is greater than indicated by the mean and standard deviation of the difference between the two models, as some of the most geologically interesting regions as well as the two lander sites are located within or near the regions which show the largest difference between the two models.

Finally, the CSA model thermally derived albedos are similar to the 2% model thermally derived albedos; neither of the thermally derived albedos accurately model the measured albedos. As seen in Figure 8, a two-dimensional histogram produced from maps of the 2% model thermally derived albedo and the IRTM measured albedos, the 2% model overestimates the albedo at low albedo and underestimates the albedo at high albedo. Figure 9 is a two-dimensional histogram of the CSA model thermally derived albedos and the measured albedos and shows a pattern similar to Figure 8. The difference between the 2% model and CSA model thermally derived albedos has a mean of 0 and a standard deviation of 0.02. Therefore, including the effects of a dusty CO_2 atmosphere does not account for the difference between the thermally derived albedos and the measured albedos. The cause of this difference is not currently understood.

Conclusions

Including a dusty CO_2 atmosphere when modeling the thermal inertia of Mars is an improvement over making the 2% assumption. The mean difference between the CSA model thermal inertias and the 2% model thermal inertias is of the order of 50 units. The CSA model thermal inertias are generally smaller, indicating that particle sizes have probably been overestimated for much of the planet. However, the importance of utilizing the CSA model thermal inertias may be greater than indicated by the mean difference, as the largest differences occur at or near some of the more geologically interesting areas as well as near the lander sites. The map of CSA model thermal inertias presented in Figure 4 is currently the best estimate of Martian thermal inertias between -60° and 60° latitude.

Acknowledgments. We thank Mike Mellon for many useful discussions and for assistance with preparing the figures. We also thank Dave Paige and Phil Christensen for their thoughtful, constructive reviews. This research was supported in part by JPL through contract 957571 from the Mars Observer project and by NASA through grant NAGW-771.

References

Bridges, N. T., Elevation-corrected thermal inertia and derived particle size on Mars and implications for the Tharsis Montes, *Geophys. Res. Lett., 21*, 785-788, 1994.

Christensen, P. R., Martian dust mantling and surface composition: Interpretation of thermophysical properties, *J. Geophys. Res., 87*, 9985-9998, 1982.

Clancy, R. T., S. W. Lee, G. R. Gladstone, W. McMillan, and

T. Roush, A new model for Mars atmospheric dust based upon analysis of ultraviolet through infrared observations from Mariner 9, Viking, and Phobos, *J. Geophys. Res.*, this issue.

Haberle, R. M., and B. M. Jakosky, Atmospheric effects on the remote determination of thermal inertia on Mars, *Icarus*, *90*, 187-204, 1991.

Kieffer, H. H., T. Z. Martin, A. R. Peterfreund, B. M. Jakosky, E. D. Miner, and F. D. Palluconi, Thermal and albedo mapping of Mars during the Viking Primary Mission, *J. Geophys. Res.*, *82*, 4249-4291, 1977.

Martin, T. Z., Thermal infrared opacity of the Mars atmosphere, *Icarus*, *66*, 2-21, 1986.

Martin, T. Z. and M. I. Richardson, New dust opacity mapping from Viking Infrared Thermal Mapper data, *J. Geophys. Res.*, *98*, 10,941-10,949, 1993.

Masamune, S., and J. M. Smith, Thermal conductivity of beds of spherical particles, *Ind. Eng. Chem. Fundam.*, *2*, 136-143, 1963.

Paige, D. A., J. E. Bachman, and K. D. Keegan, Thermal and albedo mapping of the polar regions of Mars using Viking Thermal Mapper observations, 1, North polar region, *J. Geophys. Res.*, in press, 1994.

Palluconi, F. D., and H. H. Kieffer, Thermal inertia mapping of Mars from 60°S to 60°N, *Icarus*, *45*, 415-426, 1981.

U.S. Geological Survey, Topographic maps of the western equatorial, eastern equatorial, and polar regions of Mars, *U.S. Geol. Surv. Misc. Invest. Ser. Map*, I-2030, scale 1:15,000,000, 1989.

Wechsler, A. E., and P. E. Glaser, Pressure effects on postulated lunar materials, *Icarus*, *4*, 335-352, 1965.

Wentworth, C. K., A scale of grade and class terms for clastic sediments, *J. Geol.*, *30*, 377-392, 1922.

Zurek, R. W., Martian great dust storms: An update, *Icarus*, *50*, 288-310, 1982.

R. M. Haberle, Space Sciences Division, MS 245-3, NASA Ames Research Center, Moffett Field, CA 94035.

J. N. Hayashi and B. M. Jakosky, Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, CO 80309.

(Received February 24, 1994; revised August 22, 1994; accepted September 19, 1994.)

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5285–5296, MARCH 25, 1995

55-91
322810

Thermal inertias in the upper millimeters of the Martian surface derived using Phobos' shadow

Bruce H. Betts,[1] Bruce C. Murray, and Tomáš Svítek[2]
Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena

Abstract. The first thermal images of Phobos' shadow on the surface of Mars, in addition to simultaneous visible images, were obtained by the Phobos '88 Termoskan instrument. The best observed shadow occurrence was on the flanks of Arsia Mons. For this occurrence, we combined the observed decrease in visible illumination of the surface with the observed decrease in brightness temperature to calculate thermal inertias of the Martian surface. The most realistic of our three models of eclipse cooling improves upon our preliminary model by including nonisothermal initial conditions and downward atmospheric flux. Most of our derived inertias fall within the range 38 to 59 J m^{-2} $s^{-1/2}$ K^{-1} (0.9 to 1.4 10^{-3} cal cm^{-2} $s^{-1/2}$ K^{-1}), corresponding to dust-sized particles (for a homogeneous surface), consistent with previous theories of Tharsis as a current area of dust deposition. Viking infrared thermal mapper (IRTM) inertias are diurnally derived and are sensitive to centimeter depths, whereas the shadow-derived inertias sample the upper tenths of a millimeter of the surface. The shadow-derived inertias are lower than those derived from Viking IRTM measurements (84 to 147), however, uncertainties in both sets of derived inertias make conclusions about layering tenuous. Thus, near-surface millimeter versus centimeter layering may exist in this region, but if it does, it is likely not very significant. Both eclipse and diurnal inertias appear to increase near the eastern end of the shadow occurrence. We also analyzed a shadow occurrence near the crater Herschel that showed no observed cooling. This analysis was limited by cool morning temperatures and instrument sensitivity, but yielded a lower bound of 80 on eclipse inertias in that region. Based upon our results, we strongly recommend future spacecraft thermal observations of Phobos' shadow, and suggest that they will be most useful if they improve upon Termoskan's geographic and temporal coverage and its accuracy.

1. Introduction

Mars' moon Phobos orbits Mars in a roughly 8-hour, circular, equatorial orbit at an altitude of approximately 6000 km. During the time periods surrounding the Martian equinoxes, Phobos casts a completely penumbral shadow on the surface of Mars' equatorial regions during portions of each orbit. A passing of the shadow would be viewed by an observer on the surface as a partial eclipse lasting roughly 20 s. The Termoskan instrument on board the Soviet Phobos '88 spacecraft obtained the first thermal images of Phobos' shadow on the surface of Mars. Simultaneous visible images were also obtained.

Termoskan was an optical-mechanical scanning radiometer with one visible channel (0.5-1.0 μm) and one thermal infrared channel (8.5-12.0 μm) (see *Selivanov et al.* [1989], *Murray et al.* [1991], and *Betts* [1993] for more information). The instrument was fixed to the spacecraft, pointing in the antisolar direction. Termoskan used a scanning mirror to build up the north-south component of image panoramas and the spacecraft's motion to build the east-west component.

Termoskan observed the shadow of Phobos on the surface of Mars during two of its four panoramas. Using these observations, we have been able to combine the observed decrease in visible illumination of the surface with the observed decrease in brightness temperature to calculate thermal inertias of the uppermost tenths of a millimeter of the Martian surface. Thermal inertia, a bulk measure of the resistance of a unit surface area to changes in temperature, is commonly used to characterize the insulating properties of planetary surfaces. It is defined as $I = (k\rho c_p)^{1/2}$, where k is the thermal conductivity, ρ is the bulk density, and c_p is the specific heat of the material. Low-inertia materials have smaller thermal skin depths and heat and cool more quickly than high inertia materials. In this paper, thermal inertia values are given in SI units (J m^{-2} K^{-1} $S^{-1/2}$) with inertia values in the units (10^{-3} cal cm^{-2} K^{-1} $S^{-1/2}$) often used for the Martian surface [e.g., *Kieffer et al.*, 1977] given in parentheses following the SI values.

Murray et al. [1991] presented a summary of the preliminary results of an analysis of the best observed and least complicated occurrence of the shadow (near Arsia Mars; see Figure 1). We now present for the first time (1) a refined analysis of that shadow occurrence that incorporates two additional model factors: nonisothermal start conditions and atmospheric re-radiation, and that includes inertias as a function of longitude in the shadowed region; (2) an analysis of another shadow occurrence which was observed near the crater Herschel; (3) a detailed description of the models used; (4) a comparison with the diurnal inertias of *Hayashi et al.* [this issue], which are based upon the combined surface atmosphere (CSA) model of *Haberle*

[1]Now at the San Juan Capistrano Research Institute, San Juan Capistrano, California.
[2]Now at Orbital Sciences Corporation, Dulles, Virginia.

Paper number 95JE00226.
0148-0227/95/95JE-00226$05.00

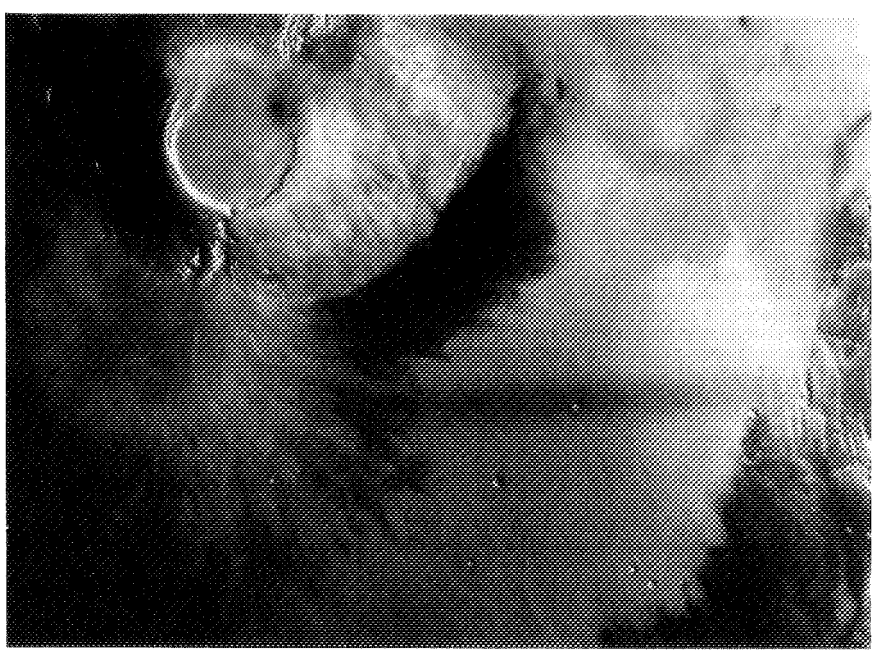


Table 1. Phobos Shadow Occurrences Within the Termoskan Data

Occurrence	Scan	Location	West Longitudes Covered, deg	Times of Day Covered, hours	Notes
1	3	Flanks of Arsia Mons	110 to 120	9.3 to 10.1	Analyzed here.
2	3	S. of western half of Valles Marineris	75 to 105	10.5 to 12.7	
3	4	Eastern end is in the crater Herschel	228 to 255	7.4 to 9.5	No cooling apparent in thermal data. Analyzed here.
4	4	North of Ma'adim Vallis	172 to 210	10.8 to 13.8	No visible channel data for eastern end of shadow. Dropped lines in visible data that do exist.

All shadow occurrences are centered approximately upon 14°S latitude.

and Jakosky [1991]; and (5) implications for the planning and analysis of future similar observations.

Section 2 of this paper gives an overview of the observations and the thermal models we used. Section 3 compares the Termoskan thermal data with the model results to derive thermal inertias. Section 4 discusses the results, and Section 5 discusses and the potential for future Phobos shadow research. The appendix describes in detail the three thermal models, including the inputs that were used from the Termoskan data.

2. Overview of Observations and Modeling

The elongated shape of the Phobos shadow (Figure 1) in the Termoskan panoramas was due to a fortuitous combination of the scanning nature of the instrument, unusual orbital geometry (the spacecraft and Phobos were nearly co-orbiting), the antisolar orientation of the instrument, and a slight rocking of the spacecraft. The similarity of the spacecraft's orbit to Phobos' orbit combined with Termoskan's antisolar orientation (zero solar phase angle) conspired to put Termoskan's instantaneous field of view near the location of the Phobos shadow as it traveled across Mars' surface. However, because the spacecraft and moon were not actually in the same place, this orientation alone would have missed observing the shadow. "Fortunately," the spacecraft, and hence Termoskan, rocked slightly back and forth. Thus, Termoskan's instantaneous field of view rocked into and out of observing the shadow (see *Murray et al.* [1991] or *Betts* [1993] for more detailed descriptions of the observations).

Termoskan's field of view rocked into observing Phobos' shadow four times, twice in each of two observing sessions, both on March 26, 1989. The four shadow occurrences, their locations, and their basic differences are summarized in Table 1. All are centered roughly on 14°S latitude.

Unfortunately, in terms of ease of analysis, each of the shadow occurrences has different characteristics. The Arsia Mons occurrence (number 1 in Table 1) was the best observed and is able to yield the most information. Hence, it is the focus of our modeling and analysis presented here.

We also present an analysis of occurrence 3 (the "Herschel occurrence"); however, that analysis is necessarily simplified and much less accurate because no cooling from the passage of the shadow was detected. Thus, only a lower bound on thermal inertia could be determined by making assumptions about what level of cooling could be detected in the Termoskan data.

Occurrences 2 and 4 are both more complex than the Arsia occurrence. The shadows themselves appear longer (cover a wider range of longitudes), and the E-W shadow profiles do not appear as smooth and symmetric (smoothly darkening, then brightening) as the Arsia shadow profile. These two factors imply that the spacecraft rocking motion was slower and that the assumption of uniform rocking motion (used in our analysis and discussed in the appendix) is less likely to be valid. In addition, the greater apparent lengths of the shadows imply that a wider range of local times of day, surrounding temperatures, and albedos were observed. Occurrence 4 additionally lacks complete visible data coverage.

For the Arsia and Herschel occurrences, we used the observed drop in the visible flux within the shadowed area to model the solar insolation as a function of both actual time since the beginning of eclipse and position in the scan. We then used this in three different one-dimensional, finite difference thermal models for homogeneous surfaces (adapted from *Clifford et al.* [1987]). Model 1 was our preliminary model [*Betts et al.*, 1990; *Murray et al.*, 1991]. Model 2 does not assume that the pre-eclipse temperatures are constant with depth as Model 1 did. Instead it uses a temperature with depth profile derived from a diurnal thermal model (as discussed in the appendix). *Haberle and Jakosky* [1991] concluded that atmospheric effects are less important for eclipse-derived thermal inertias than for diurnally derived thermal inertias, based upon theoretical considerations

Figure 1. Phobos shadow images. (Top) Termoskan visible and (Bottom) thermal images showing the Phobos shadow occurrence on the flanks of Arsia Mons. North is towards the top. Contrast has been enhanced to emphasize the shadow. The elongated shape of the shadow is due to a fortuitous combination of the scanning nature of the instrument, unusual orbital geometry (the spacecraft and Phobos were nearly co-orbiting), the antisolar orientation of the instrument, and a slight rocking motion of the spacecraft. Note that the shadow is observed first (i.e., further west because the instrument built up images from west to east) in the visible, then later (to the east) in the thermal. This is due to the delay in cooling after the onset of the shadow.

involving effects of the short duration of the eclipse relative to a diurnal period, and the relatively high daytime temperatures (which cause surface emission to dominate atmospheric emission). In order to test this and to ascertain the magnitude of atmospheric effects, we created model 3 by adding a downward atmospheric flux term to model 2. Model 3 is the most realistic of the three models, but due to the uncertainty in the downward atmospheric flux, it is useful to consider model 2 separately. For the Arsia occurrence, we present the results of all three models here for comparison and completeness (see the appendix for detailed descriptions of all three models.)

3. Comparison of Data With Models

3.1. Arsia Occurrence

To best compare Termoskan eclipse temperatures with model results, we used temperature drops within the eclipse rather than absolute temperatures. Variations in absolute temperature are observed outside the eclipse due to variations in albedo, inertia, and time of day. These factors also will affect the observed temperatures within the eclipse. To minimize these effects, we did two things to the data used for comparison. First, for points within and outside the shadow, we averaged 10-pixel (in the east-west direction) by 1-pixel strips. Second, the data we compare with models are estimated temperature drops within the eclipse. We find these temperature drops by averaging the temperatures of points to either side (north and south) of the shadow and then subtracting the temperature observed in the middle of the shadow. Using these temperature drops rather than absolute temperatures is particularly effective at removing time of day effects. Also, because the area happens to be reasonably bland thermally, gradual albedo and diurnal thermal inertia variations are also removed. In addition, by using temperature drops, we reduce the effects of atmospheric scattering, which should act approximately equally inside and outside the shadow.

To plot the data versus the model curves, these temperature drops were subtracted from 255 K, the pre-eclipse temperature used in the models. Figure 2 shows the data plotted versus model 1 curves. Figure 3 shows an analogous plot for model 2, Figure 4 shows curves for model 3 with a downward atmospheric flux (A_f) of 20% of the absorbed pre-eclipse solar flux, and Figure 5 shows curves for model 3 with A_f = 10%. Data error bars are discussed in section 4.2.1., and although they are of interest in interpreting physical results, they do not affect comparisons of different models. Most of the data fall between inertias of 38 (0.9) and 50 (1.2) for model 1, between 41 (1.0) and 67 (1.6) for model 2, between 35 (0.8) and 50 (1.2) for model 3 with A_f = 20%, and between 38 (0.9) and 59 (1.4) for model 3 with A_f = 10%. In all cases, inertias are slightly higher later (towards the east) in the shadowed region.

3.2. Herschel Occurrence

For the Herschel occurrence the shadow appears clearly in Termoskan's visible channel, but cooling from the shadow does

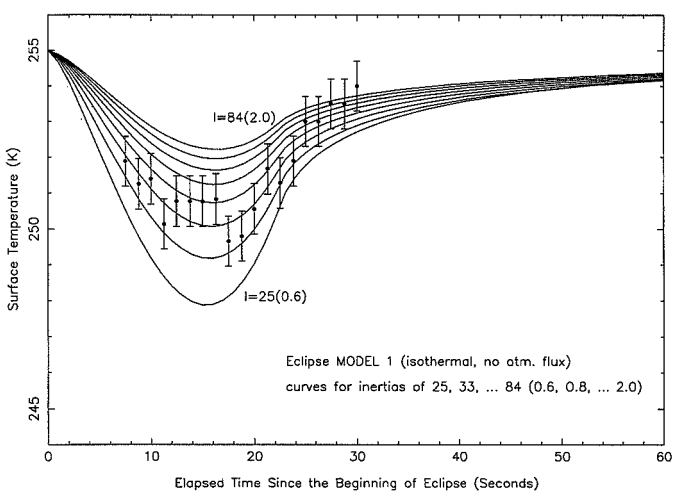


Figure 2. Eclipse model 1 results. The lines in this plot are model surface temperatures for various thermal inertias from 25 to 84 J m^{-2} $s^{-1/2}$ K^{-1} (0.6 to 2.0 in units of 10^{-3} cal cm^{-2} K^{-1} $s^{-1/2}$). The dots are observational data retrieved from the Termoskan thermal infrared channel for the shadow occurrence near Arsia Mons. They are derived from 10-line averages of temperature drops at the center of the shadow relative to comparable points outside the shadow. Size of the error bars is discussed in the text. The temperature drops have been subtracted from 255 K to facilitate comparison with our model results. We chose 255 K for the model because it is a typical temperature from the region surrounding the shadow. Most of the observational data values fall between model curves corresponding to thermal inertias of 38 (0.9) to 50 (1.2). Results in this figure are the same as those presented by *Murray et al.* [1991].

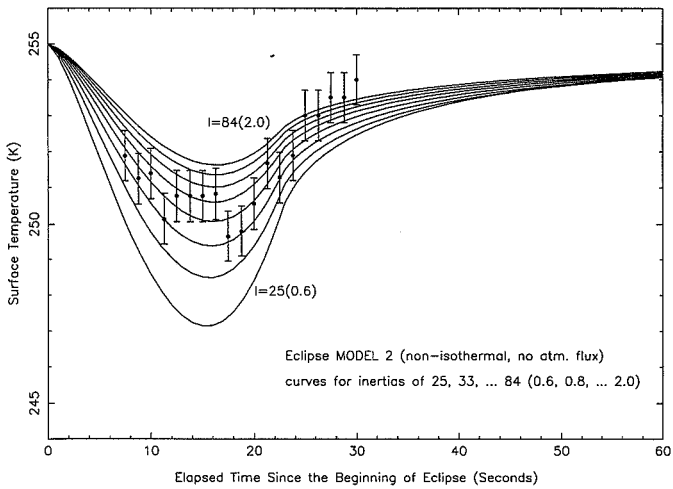


Figure 3. Eclipse model 2 results. Same as Figure 2 but for model 2 results. Model 2 added nonisothermal with depth pre-eclipse conditions to model 1. Most of the observational data values fall between model curves corresponding to thermal inertias of 41 (1.0) to 67 (1.6).

not show up in the thermal channel. Thus, for this occurrence, we cannot measure temperature drops. However, we can estimate the level of cooling that could have been detected by Termoskan, and compare that value to model results to derive a lower bound on eclipse inertias for the region near the center of the shadow. Modeling was carried out as before but with the appropriate changes in time of day, insolation, albedo, and starting temperatures. We only ran one of our models: Model 3

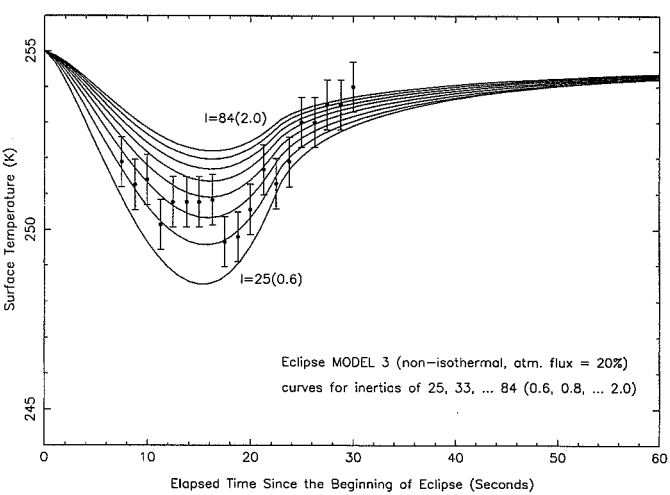


Figure 4. Eclipse model 3, A_f = 20%. Same as Figure 3 but for model 3 results. Model 3 added atmospheric downward flux to model 2. The curves plotted here are for a downward atmospheric flux of 20% of the absorbed pre-eclipse solar flux. Most of the observational data values fall between model curves corresponding to thermal inertias of 35 (0.8) to 50 (1.2).

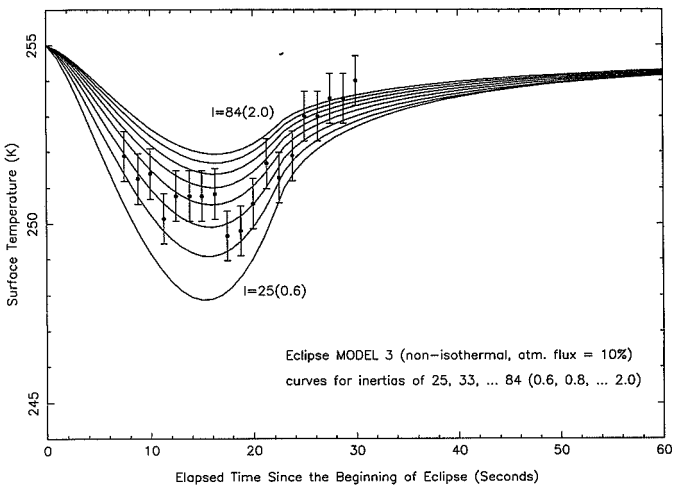


Figure 5. Eclipse model 3, A_f = 10%. Same as Figure 4 but with 10% downward atmospheric flux. Most of the observational data values fall between model curves corresponding to thermal inertias of 38 (0.9) to 59 (1.4). We consider this the most realistic model of those presented for the Arsia occurrence.

with A_f = 10%, which we determined to be our most realistic model (discussed below). We compared the maximum amount of modeled cooling for several thermal inertias with an estimate of the amount of cooling that could have been detected in the Termoskan data either by image processing or digital data analysis.

We estimate for this occurrence that we could have detected shadow cooling corresponding to a drop in the infrared channel of approximately 3 DN (data numbers) relative to the surrounding unshadowed area. This is based upon tested image processing of the data, comparisons with other occurrences of the shadow, and attempts to detect simulated shadow cooling effects. Infrared channel variations in this region due to inherent terrain variations are also on the order of 2 or 3 DN, but we are aided in looking for shadow cooling by the image nature of the data and the fact that we know where the shadow should be based upon the visible channel data.

For the temperatures in the Herschel region, 3 DN corresponds to approximately 2.5 K. Comparing with our model results, we find that a 2.5 K temperature drop corresponds to an eclipse inertia of approximately 80 (2). Thus, the data provide an approximate lower bound of 80 (2) for the thermal inertia in the upper millimeter of the region near the center of the shadow (approximately 14°S, 241°W). At first, it may seem counterintuitive that no observed cooling would only constrain the inertia to be above a value as low as 80 (2). After all, the very obvious cooling observed in the Arsia occurrence corresponds to inertias near 40 (1). However, the following factors make more accurate constraint impossible: (1) The Herschel occurrence was much cooler than the Arsia occurrence due primarily to the earlier morning observations as well as the inherently higher diurnal inertias in the Herschel region. Termoskan had less sensitivity at cooler temperatures (see *Murray et al.* [1991] or *Betts* [1993] for the infrared response characteristics), e.g., 3 DN at 230 K correspond to about 2.5 K, but 3 DN at 255 K (characteristic of the Arsia occurrence) correspond to approximately 1.5 K. (2) Due to the strong dependence of the radiative cooling rate upon temperature, less shadow-induced cooling occurs when the initial temperatures are cooler. (3) The surface temperature drop during eclipse is very nonlinear with inertia, which causes less precision distinguishing between high inertia values than between low inertia values. For example, models of the Herschel occurrence yield temperature drops of 2.5 K, 1.3 K, 0.9 K, 0.7 K, and 0.6 K for inertias of 84 (2), 168 (4), 252 (6), 336 (8), and 420 (10), respectively.

4. Discussion

The results from the Arsia occurrence show that model 2 raises the derived inertias compared to the less realistic (isothermal) model 1. Adding atmospheric downward flux in model 3 reduces the derived inertias. First, we consider why the differences between the models cause these effects. Then, we consider the models' implications for the surface of Mars.

4.1. Model Differences and Realistic Inertias

Here we compare and evaluate our three models using the Arsia occurrence. Including the nonisothermal, initial temperature-versus-depth profile in model 2 allows the model surface to cool more quickly. For these low inertias, this dominates the simultaneous opposite (heating) effect from the additional pre-eclipse insolation caused by the inclusion of a nonzero conduction term in the surface boundary condition (see equations (3) and (5) in the appendix). Overall, the faster cooling causes higher inertias to be derived in model 2 than in model 1.

The atmospheric downward flux added in model 3 stays constant throughout the eclipse (the validity of this assumption is discussed in the appendix), and thus keeps the total insolation higher throughout the eclipse. This causes the surface to cool more slowly. Thus, including more downward atmospheric flux causes lower inertias to be derived.

Model 3 should be the most realistic model because it includes both nonisothermal starting conditions and downward atmospheric flux. The next question is, what value of atmospheric flux in model 3 is the most realistic? This is hard to answer precisely, but fortunately the most likely possibilities make little difference in our results and do not affect our eventual scientific conclusions. The atmospheric flux depends upon the atmospheric optical depth, which is not well known for the time of the observations. Analyses of data from other Phobos '88 instruments have given optical depths that range from about 0.2 to 0.6, with the most favored values somewhere near the middle of that range. According to modeling by *Haberle and Jakosky* [1991, Figure 4], for this range of optical depths and the local time of day of the observations, the atmospheric flux will range from 6% to 11% of the solar flux. Thus, the 20% model run shown in Figure 4 is probably an extreme and should give lower bounds on inertias. Even the 10% run shown in Figure 5 probably gives inertias that are somewhat low. This is particularly true because the shadow was observed on a surface roughly 9 km (almost 1 scale height) above the 6.1 mbar reference altitude used in the Haberle and Jakosky figure that the flux percentages came from. Atmospheric flux contributions will be reduced at higher altitudes.

The model 2 runs, which do not have any atmospheric flux contribution, should give inertia upper bounds. The real values should lie between the model 2 and model 3-20% values and are probably a little higher but not too far from the model 3-10% values. Thus, inertias of 38 (0.9) and 59 (1.4) probably bracket the majority of the eclipse-derived inertias in this shadow region. Ironically, because the nonisothermal effects nearly balance the atmospheric effects for these inertias, these results are nearly the same as the model 1 results (which were originally reported in *Murray et al.* [1991] and *Betts et al.* [1990]).

4.2. Implications for the Martian Surface

4.2.1. Arsia occurrence. The inertias derived from all the models are consistent with dust-sized particles. This is consistent with previous studies that proposed that the Tharsis region has a dust covering and is currently an area of dust deposition [*Kieffer et al.*, 1977; *Zimbelman and Kieffer*, 1979; *Palluconi and Kieffer*, 1981; *Christensen*, 1986b]. It is also reasonably consistent with the low rock abundances (about 5%) in this region [*Christensen*, 1982, 1983, 1986a]. Assuming a homogeneous surface, inertias of 38 to 59 (0.9 to 1.4) likely imply particle sizes of approximately 5 to 10 microns (μm) [*Haberle and Jakosky*, 1991; *Kieffer et al.*, 1973; *Jakosky*, 1986]. Note, however, that these values are theoretical extrapolations from laboratory data that only go down to particle sizes of approximately 30 μm [*Jakosky*, 1986]. Also, all of these particle sizes are for homogeneous surfaces, thus the actual surface could be composed of all 2-μm particles and an occasional rock or any number of other combinations that yield a thermally averaged surface of 5- to 10-μm particles.

For the very low inertias involved in this study, diurnally derived inertias are representative of thermal skin depths of 1 or 2 cm. Due to the short duration of the eclipse relative to the length of a Martian day, the eclipse-based inertia determinations are sensitive to thermal skin depths of only a few tenths of a millimeter. Comparison of inertias derived by the two methods can indicate the degree of layering within the upper centimeters of the surface.

We compare our eclipse inertias for this region with Viking infrared thermal mapper (IRTM)-based diurnal inertias (2° x 2° bins) calculated by both *Hayashi et al.* [this issue] and *Palluconi and Kieffer* [1981]. *Hayashi et al.* [this issue] used the combined surface-atmosphere model of *Haberle and Jakosky* [1991] along with the results of *Palluconi and Kieffer* [1981] and average dust opacities from *Martin and Richardson* [1993] in order to more accurately model the atmospheric contribution to the surface heat balance. Our eclipse inertias of 38-59 (0.9-1.4) compare to diurnal inertias for the same region of 86-133 (2.1-3.2) [*Hayashi et al.*, this issue] and 92-147 (2.2-3.5) [*Palluconi and Kieffer*, 1981]. Note that the atmospheric corrections of *Hayashi et al.* [this issue] have relatively little effect on inertias when compared with *Palluconi and Kieffer* [1981] for this region due to relatively high altitude and low average dust opacities. Corrections are significantly larger elsewhere on Mars [*Hayashi et al.*, this issue].

Thus, our derived eclipse inertias are significantly lower than diurnally derived inertias. However, the uncertainties in both sets of inertias make conclusions about near-surface layering tenuous. Most of the uncertainty for the Arsia occurrence eclipse inertias comes from a combination of approximating temperature drops within the eclipse (about ±0.4 K) and the precision of the instrument (about 0.5 K). Because these errors are independent and presumably random, they combine to provide an uncertainty of approximately ±0.7 K (which is the data uncertainty plotted in Figures 2 through 5). This translates to an inertia uncertainty of less than ±12 (0.3) for the majority of the inertias derived. It is worse for higher inertias where there is less temperature distinction between inertias, and for locations far from the shadow center. In addition, there are uncertainties due to possible inaccuracies in model inputs such as atmospheric downward flux (±4 (0.1), based upon trying several values), the temperature versus depth profile (±4 (0.1), would be worse for higher inertias), and other parameters such as visible flux, flux drop with time, and exact duration of eclipse (±8 (0.2), based upon tests varying these parameters). Thus, the total uncertainty for the Arsia occurrence is approximately ±16 (0.4). So far, we have assumed that the uncertainty was the same, plus or minus. In reality the minus uncertainty is less than 16 (0.4) because inertia values are more easily distinguished at lower inertias. Note also that these uncertainties are for the Arsia occurrence and will vary even for the Termoskan instrument based on time of day, surface temperature, and inertia. Therefore, future observations can be better optimized than the Termoskan observations as discussed in section 5.

Hayashi et al. [this issue] state their uncertainty as 10-20 (0.2-0.5) for regions where albedo ≥0.28 and less than 50-60 (1.2-1.4) for albedos ≤0.28. The average albedo for the Arsia occurrence was 0.27 (it was 0.17 for the Herschel occurrence) [based upon *Pleskot and Miner*, 1981]. For the Arsia occurrence, we assume that the emissivity component of error in the *Hayashi et al.* [this issue] inertias is small since the average albedo of the region is nearly that of the cutoff albedo (0.28), above which emissivity errors are negligible (because the emissivity does effectively equal 1 as assumed [*Christensen*, 1982]). However, since it is slightly below this cutoff, we include some emissivity error for comparison's sake. In this case, we have assumed that

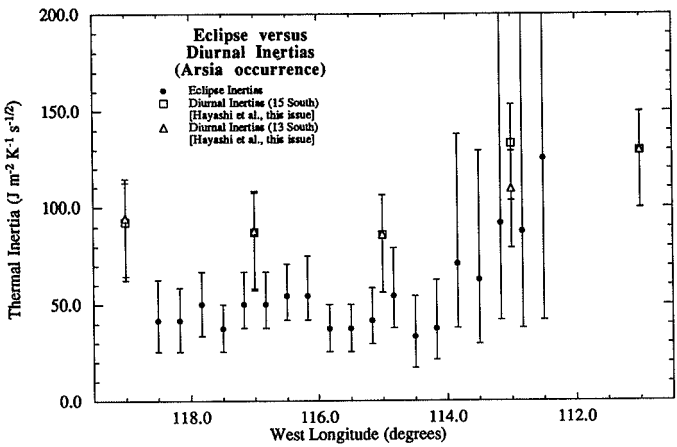


Figure 6. Eclipse versus diurnal inertias for the Arsia occurrence. Eclipse inertias from our preferred model (model 3, A_f = 10%, see Figure 5) plotted along with the diurnal inertias (2° x 2° binned) of *Hayashi et al.* [this issue] for the same region. Note that the error bars (discussed in text) often overlap, but that the eclipse inertias are consistently lower than the diurnal inertias. Also note that both sets of inertias trend upwards towards the eastern end of the shadow. Eclipse errors also increase towards the east because the method is less sensitive far from the center of the eclipse and at higher inertias.

the emissivity error is less than 10 (0.2) in the *Hayashi et al.* [this issue] diurnal inertias for the Arsia occurrence region.

Figure 6 compares eclipse inertias (with uncertainties found independently for each point) with the *Hayashi et al.* [this issue] inertias, which are plotted with uncertainties of +20 and -30 (emissivity error will only act in the negative direction since it causes the inertias to be overestimated). Note that the eclipse inertias are, as discussed, consistently lower than the diurnal inertias, but that error bars often overlap. Also, notice that the eclipse inertias trend upwards towards the east as do the diurnal inertias. Unfortunately, the eclipse errors become large in that region because the eclipse is ending and the inertias are higher. However, the trend appears in the last few eclipse inertia points, making the simultaneous upward trend of both eclipse and diurnal inertias more believable than the error bars from any one point would indicate.

Thus, due to the uncertainties in both our eclipse inertias and in diurnally derived inertias, we are unable to say for certain that there is layering in the Arsia occurrence region. If there is, we can say that it is mostly fine dust over mostly less fine dust. The easiest and most interesting locations to look for layering with future missions will be in regions that have higher diurnal inertias (indicative of average particle sizes significantly larger than dust) than those of the Arsia occurrence.

The few tenths of a millimeter sampled by eclipse cooling measurements should present a fair representation of the surface that is sensed by optical and near-infrared instruments. Thus, for the region studied, it is unlikely that optical and near-infrared instruments would sample any significant amount of bare (not covered by dust) rock surfaces.

4.2.2. Herschel occurrence. The inertia lower bound of 80 (2) that we derive for the Herschel occurrence is consistent with diurnal inertias for the shadowed region: 198-252 (4.7-6.0) [Hayashi et al., this issue] and 239-301 (5.7-7.2) [Palluconi and Kieffer, 1981]. These diurnal inertias are consistent with a homogeneous surface of sand-sized particles. Unfortunately, because we are only able to derive a lower bound for this occurrence, and because that lower bound is so low, we are unable to derive any strong physical conclusions about the surface. If eclipse inertias are actually as low as 80 (2), then one could say there is a dust covering, but since they also could be significantly larger than 80 (2), the surface may be totally devoid of layering. Note, that if more advantageous observations (midday, greater sensitivity, etc.) are obtained by future spacecraft, the question of layering can be addressed effectively by the shadow analysis method.

4.3. Summary

Our analyses have shown that thermal and visible observations of the shadow of Phobos can be used to derive physical information about the upper millimeters of the Martian surface. Combining our preferred eclipse model with the Termoskan thermal data implies inertias for the Arsia Mons shadow occurrence that mostly fall within the range 38 to 59 (0.9 to 1.4). These inertias correspond to dust-sized particles (5 to 10 μm for a homogeneous surface based upon theoretical extrapolations of laboratory data). The presence of dust at the surface is consistent with previous theories of Tharsis as a current area of dust deposition. Thus, most of the upper couple centimeters is likely composed of unbonded, few micron dust particles. There may be a slight degree of layering, and the eclipse inertias and diurnal inertias both appear to increase towards the eastern end of the shadow. However, uncertainties

make these last two conclusions tenuous. Analysis of a shadow occurrence near the crater Herschel places a lower bound on eclipse inertia of 80 (2).

These findings obviously do not preclude layering or higher inertias elsewhere. Questions of geographic variability of the upper millimeter of the surface can be addressed to some extent in the future using the two other Termoskan observed occurrences of the Phobos shadow, but can best be addressed by future missions with optimized observations.

5. Future Phobos Shadow Research

Our results indicate that analyses of Martian surface cooling, due to the passage of the shadow of Phobos, can yield unique and interesting results about the current physical state of the upper millimeters of the Martian surface. However, the fortuitous Termoskan observations of the shadow are very limited in geographic coverage and, with the exception of the Arsia occurrence, are very limited in accuracy. Because of these limitations, some of the key questions that could be addressed by this type of observation remain unanswered, e.g., is there a dust layer of at least millimeters in thickness covering much of Mars?, or are dust layers regional in nature? Thus, we recommend that observations of the shadow be attempted with future missions. Future observations should improve upon the Termoskan data in geographic coverage, accuracy of the observations, and temporal coverage.

Ideal characteristics of future observations would include targeting midday shadow passages to maximize the pre-eclipse temperature, which will maximize the shadow temperature drops, the distinction between different inertias during eclipse, and the signal to noise ratio; utilizing framing rather than (Termoskan-like) scanning imaging systems so the entire image is taken simultaneously, and taking several images as the shadow crosses the surface so that pre-eclipse, eclipse, and posteclipse temperatures are more accurately known for any given location; and having at least one broad band thermal channel and one broad band visible channel.

Observations will need to be taken during periods surrounding the equinoxes when the shadow will cross the Mars surface. The least shadow distortion and atmospheric interaction will occur in the equatorial region at equinox and during periods near that time. Because Phobos has an orbital period of approximately 8 hours, there should be ample opportunity to observe several shadow crossings at several latitudes. The most advantageous orbits for the observations will be equatorial (as is Phobos' orbit) so that the shadow can be followed across the surface and so pre-eclipse and posteclipse observations can easily be obtained. One ideal observing situation would be to have the instruments on a lander on the Mars facing side of Phobos. Then, observations could be taken merely by pointing in an antisolar direction. They will be particularly useful if the field of view is wide enough and if observations are taken frequently enough so a number of locations are observed, with temporal coverage for each location providing pre-eclipse, eclipse, and posteclipse temperatures.

Appendix: Thermal Models of the Eclipse Cooling

In order to derive thermal inertias from the eclipse observations, we modeled the cooling of the surface for several thermal inertias and compared the results to the observed cooling. We present here descriptions and results from three different, but related, models: model 1, the isothermal model, which assumes that initially all depths are at the same temperature; model 2, the nonisothermal model, which utilizes an initial temperature with depth profile derived from the *Clifford et al.* [1987] diurnal thermal model; and model 3, the nonisothermal atmospheric model, which adds a downward atmospheric flux term to model 2. All three models are presented here for comparison and completeness. Model 1 was used to produce the results in *Murray et al.* [1991]. Model 2 improves on model 1. Model 3 is the most complete model because it includes atmospheric re-radiation; however, the amount of atmospheric flux is highly dependent upon poorly known atmospheric conditions at the time of the observations. Thus, it is useful to consider model 2 separately from model 3.

All of the models solve the heat equation:

$$\frac{\partial T}{\partial t} = \frac{1}{\rho c}\frac{\partial}{\partial z}\left(k\frac{\partial T}{\partial z}\right) \tag{1}$$

where T is temperature, t is time, ρ is density, c is specific heat, k is thermal conductivity, and z is depth. This equation is solved using the same numerical method used by *Clifford et al.* [1987]. This is essentially the same diurnal model described by *Kieffer et al.* [1977]. Modeling the eclipse rather than diurnal and seasonal temperature variations requires several important differences between the eclipse models and the overall Clifford et al. model. Parameters such as the timescales used and inputs such as the decrease in visible flux with time in the eclipse must be changed. Where possible, we use the actual Termoskan data to make these changes, rather than relying on theoretical calculations. For example, we use the Termoskan visible data to derive the decrease in absorbed solar flux with time. Using the data where possible decreases the potential error introduced by poorly known factors such as the complex geometry including spacecraft rocking, atmospheric effects, and uncertainties in albedos.

All of the models produce model temperatures as a function of time in eclipse for a given value of thermal inertia. Actual thermal inertias are estimated by first running a model for several values of thermal inertia; then, model temperatures are compared with Termoskan data temperatures to find inertias for given values of time eclipsed. This comparison was discussed in section 3.

In the rest of this section we describe the three models by pointing out their differences from the *Clifford et al.* [1987] and *Kieffer et al.* [1977] models. In the first five points below, we present the elements that apply to all three models. Then, in the last two points, we address the differences in the three models. Hence, the following are the eclipse model features that are different from the *Clifford et al.* [1987] and, by analogy, the *Kieffer et al.* [1977] model.

1. Timescales (depth scales). We have adjusted all time-related variables to account for the 23-s eclipse timescale, rather than diurnal timescales. The iteration time, along with the modeled inertia, determines the thickness and depths of the model compartments based upon finite difference stability criteria [e.g., *Clifford et al.*, 1987]. We changed the interval between iterations from several minutes to 0.001 s. Models with smaller iteration times were tested but showed no appreciable difference from the models run with 0.001-s iterations.

2. Use of orbital information and albedos. Neither Mars orbital information nor albedo information is used in the eclipse models. How these are avoided will become more clear in the specific discussions of elements below. We generalize this point here to emphasize the major differences with the *Clifford et al.*

[1987] model. Note that although the eclipse models do not use orbital information or albedo information directly, models 2 and 3 do use them indirectly because we ran the *Clifford et al.* [1987] model to generate the initial temperature with depth profiles (discussed in point 6 below). Errors in albedo introduce only very minor errors when used in this indirect fashion.

3. Downward atmospheric flux term. The downward atmospheric flux model term is set equal to 0 rather than 2% of the noontime solar insolation. Models 1 and 2 have no downward atmospheric flux and model 3 incorporates it elsewhere within the model.

4. Initial surface temperature. Rather than calculating this from orbital and albedo considerations, the initial (pre-eclipse) surface temperature is assigned a value that is a representative of observed Termoskan temperatures for regions just outside the shadow (e.g., 255 K for the Arsia Mons occurrence).

5. Relative insolation as a function of time. In order to model the cooling as a function of time eclipsed, the model needs a description of the total absorbed insolation as a function of time within the eclipse. For models 1 and 2 this is equivalent to the absorbed solar insolation as a function of time. For model 3 this includes atmospheric as well as solar insolation as discussed more fully in point 7 below. We derive the shape of the insolation-versus-time function from the visible data and the magnitude from the thermal data. In point 7 below we consider the magnitude of the pre-eclipse insolation used, which differs for each of the three models.

Here we discuss the shape of the insolation versus time function. We use this same relative insolation function in all three models. To find the relative insolation, we did the following: A - calculated the duration of the eclipse for a point on the surface; B - derived a function that connects location within the shadow in the data to the amount of time that location has been in eclipse; C - used the visible channel eclipse profile to fit the decrease in absorbed solar insolation as a function of location within the shadow; and D - combined B and C to derive relative insolation as a function of time in eclipse. These steps are accomplished in the following manner.

A. Combining the maximum width of a north-south profile, the orbital speed of Phobos, and the rotational speed of Mars gives an eclipse duration of just under 23 s.

B. We assumed the E-W rocking motion to be uniform over the brief Arsia Mons shadow observation. This assumption is justified both by the short duration of the observation and by the smoothness and shape of the E-W visible shadow profile. Thus, we were able to develop a linear relation between the number of lines into eclipse and the time a location had been in eclipse using the observed E-W size of the shadow (approximately 180 lines for the Arsia occurrence) and the eclipse duration calculated above (approximately 22.6 s). So, for example, the 90th line after the beginning of eclipse had been in some form of eclipse for approximately 11.3 s.

C. The next step was to determine the relative decrease in solar insolation as a function of location within the shadow. We assumed that the visible channel signal was linearly proportional to visible flux, consistent with the visible detector's characteristics and with all approximate flux calibrations of the visible channel [*Betts*, 1993]. Thus, the E-W visible channel signal profile should approximate the visible flux decrease within the shadow. We used this profile to quadradically fit the relative decrease in solar insolation as a function of location within the shadow. Note that in the darkest part of the shadow the flux decreased by approximately 30%. The advantage of using the visible channel profile instead of theoretical calculations of the flux decrease is that it already incorporates atmospheric effects as well as geometric distortion due to the spacecraft rocking. Position within the visible channel can be directly correlated with position within the thermal channel to within approximately 1 pixel [*Betts*, 1993; 1992].

D. Combining B and C gives the relative absorbed solar insolation as a function of time eclipsed, $f(t)$. This function can vary from 0 to 1, where 1 represents the pre-eclipse insolation.

The only remaining step for modeling the insolation is to get an absolute pre-eclipse absorbed solar insolation that this relative function can be mated with. This is discussed separately below in point 7. First, we discuss the initial temperature with depth profile used because that will be important for point 7.

6. Initial temperature with depth profile. Model 1 is isothermal before the onset of eclipse with all depths set equal to the pre-eclipse representative surface temperature (255 K for the Arsia occurrence). Models 2 and 3 utilize more realistic nonisothermal subsurface pre-eclipse temperature profiles. As mentioned, establishing a reasonable estimate for this temperature with depth profile is the only place in the models where either albedo or Mars orbit information is used. Even here, they are used indirectly. Before each run of either models 2 or 3, we ran our adaptation of the *Clifford et al.* [1987] diurnal model utilizing the appropriate season (L_s = 18°), approximate latitude (14°S), and approximate albedo (0.27 for the Arsia occurrence and 0.17 for the Herschel occurrence as taken from the 1° x 1° binned bolometric albedos of *Pleskot and Miner* [1981]). Using the diurnal model output corresponding to the local time of day of the eclipse, about 10 H for the Arsia occurrence (where 24 H = 1 Martian day), we estimated the pre-eclipse temperature as a function of depth with a linear approximation:

$$T_j = T_i - \frac{D_j(T_{sd} - T_{fd})}{D} \tag{2}$$

where T_j is the temperature of the jth compartment, T_i is the initial surface temperature, i.e., 255 K for the Arsia occurrence, D_j is the depth of the top of the jth eclipse depth compartment, T_{sd} is the surface temperature in the diurnal model, T_{fd} is the temperature of the first depth compartment in the diurnal model, and D is the depth of the first depth compartment in the thermal model. Note that this pre-eclipse temperature with depth profile depends upon the thermal inertia used in the diurnal model. Thus, for each value of inertia run in the eclipse model, we first modeled the depth profile using that inertia in the *Clifford et al.* [1987] diurnal model. This gave new values of T_{sd}, T_{fd}, and D to use in equation (2) and the eclipse model. Then, for a given inertia, all values of T_j could be found as a linear function of D_j.

7. Pre-eclipse absorbed insolation. Finally, we need the value of the pre-eclipse absorbed insolation to use with the relative function found in part 5 above. To do this, we use the surface equilibrium boundary condition discussed by *Kieffer et al.* [1977] and elsewhere:

$$\frac{S(1-A)\cos i}{R^2} + F_a = \varepsilon\sigma T^4 - k\frac{dT}{dz} - L\frac{dM}{dt} \tag{3}$$

where S is the solar constant, A is the bolometric Bond albedo, i is the solar incidence angle, R is the Mars heliocentric distance, F_a is the downward atmospheric flux, ε is the surface emissivity, σ is the Stefan-Boltzmann constant, T is the surface temperature, k is the thermal conductivity, z is the depth, dT/dz is the change

in temperature with depth evaluated at the surface, L is the CO_2 latent heat of condensation, and M is the mass per unit area of CO_2 frost.

The terms on the left in equation (3) represent the total absorbed insolation, i.e., what we need to know for pre-eclipse conditions. The first term on the left is the absorbed solar insolation. The second term on the left represents the downward infrared atmospheric flux that reaches the surface. This thermal infrared flux is assumed to be absorbed completely by the surface. The first term on the right is the surface emission. Here we assume the emissivity, ε, to be 1, as was done by *Kieffer et al.* [1977]. The second term on the right is the surface conduction term. The third term on the right is the term representing latent heat from carbon dioxide sublimation or condensation. The temperatures involved in this shadow analysis are far above the carbon dioxide condensation temperature, so this term is set to zero.

The three models differ in which terms are ignored in calculating the total absorbed insolation. Model 3 only ignores the CO_2 latent heat term. Model 2 also ignores the atmospheric term, and model 1 additionally ignores the conduction terms.

In principle, the total absorbed insolation could be calculated directly in any of these models from the terms on the left side of equation (3). Instead, we choose to solve for the total insolation by finding the terms on the right side. There are three reasons to do this. One, by more explicitly using the Termoskan data, we avoid uncertainties in albedo, and the orbital geometry, and we significantly reduce the effects of atmospheric scattering. Two, utilizing the same initial surface temperature, 255 K, that is already used in the pre-eclipse temperature profile will result in greater initial consistency within the model. Three, in model 3, to calculate the terms on the left side, we also would need to directly calculate the atmospheric downward flux. This is very difficult to do accurately due to the poor knowledge of the state of the Martian atmosphere at the time of the observations. By using the surface emission and conduction terms to determine the total insolation, we avoid these problems.

We summarize the differences between our three eclipse models in Table 2. Here we describe in detail the differences in the insolations used. In model 1, which ignores the atmospheric, conduction, and latent heat terms, the pre-eclipse absorbed insolation, I_0, is approximated by σT^4, where σ is the Stefan-Boltzmann constant and T is a representative value of the surface temperature derived from the data itself (255 K for the Arsia occurrence).

Noneclipsed surface temperatures, and the corresponding noneclipsed absorbed insolations, will change over the region covered by the eclipse and differ from the representative surface temperature (255 K for Arsia); however, the short duration and small area covered by the Arsia eclipse keep these variations small. Also, as discussed in section 3, we use temperature drops within the eclipse rather than absolute temperatures for comparison with the model to minimize errors caused by insolation or initial temperature errors in the model.

Ignoring the conduction term, $k(dT/dz)$, in model 1 is consistent with the initial isothermal assumption in this model which implies $dT/dz = 0$. The insolation as a function of time, $I(t)$, combines I_0 with the relative flux decrease as a function of eclipsed time, $f(t)$, that was derived from the visible channel data in step 5. Thus, for model 1,

$$I(t) = I_0 f(t) \tag{4}$$

Model 2 differs from model 1 by approximating the pre-eclipse absorbed insolation, I_0, by

$$I_0 = \sigma T^4 - k\frac{dT}{dz} \tag{5}$$

Note that including the conduction term here is consistent with the nonisothermal initial conditions of this model. After the initial temperature profile with depth is derived for this model as described in step 6 above, dT/dz is approximated by Taylor expansions of the first three depth steps and their temperatures [*Clifford et al.*, 1987]. Note that for the midmorning Arsia eclipse observation, dT/dz is negative. Thus, I_0 will be larger in model 2 than in model 1. The total insolation as a function of time, $I(t)$, is found in model 2 in the same way as in model 1, using equation (4).

Model 3 determines the value of I_0 in the same manner as model 2, using equation (5). Model 3 differs by assuming I_0 to be composed of an atmospheric component (downward infrared flux) in addition to the solar component. Thus, model 3 differs from model 2 in the calculation of $I(t)$. The relative flux drop as a function of time derived in step 5 is based upon the visible channel, which sensed only the solar flux drop. We assume that the absorbed downward atmospheric flux, I_a, remains constant throughout the eclipse. This assumption is very reasonable because (1) the shadowed portion of the atmosphere represents less than approximately 0.5% of the volume of atmosphere that contributes infrared atmospheric flux to the shadowed surface; and (2) the flux from the shadowed atmosphere will only drop on the order of 10-20% even assuming very large amounts of atmospheric cooling, so the total atmospheric flux variation will be less than 0.1%. Thus, in our model, I_a remains constant throughout the eclipse, whereas the absorbed solar insolation, I_s, will vary with eclipse time as described by the function $f(t)$ that was derived in step 5. Thus, rather than using equation (4) as models 1 and 2 did, model 3 uses

$$I(t) = I_a + I_s f(t) \tag{6}$$

Here, I_0, which is what we derived from the surface emission and conduction terms in equation (5), represents the combination of I_a and I_s, i.e.,

$$I_0 = I_s + I_a \tag{7}$$

I_a will depend strongly upon the amount of dust in the atmosphere as well as to some extent the elevation of the surface. In order to conveniently model different values of I_a, we define I_a

Table 2. Summary of Eclipse Model Differences

Model	Pre-eclipse Temperature With Depth	Includes Atmospheric Flux	I_0	$I(t)$
1	Isothermal	No	σT^4	$I_s f(t)$
2	Nonisothermal	No	$\sigma T^4 - k(dT/dz)$	$I_s f(t)$
3	Nonisothermal	Yes	$\sigma T^4 - k(dT/dz)$	$I_a + I_s f(t)$

I_0, pre-eclipse total absorbed insolation = I_s when there is no atmospheric downward flux, $I(t)$, total insolation as a function of time in eclipse; I_s, pre-eclipse absorbed solar insolation; I_a, absorbed atmospheric insolation; $f(t)$, relative absorbed solar insolation (on a scale of 0 to 1) as a function of time in eclipse; t, time.

as a fraction, A_f, of I_s. Thus,

$$I_a = A_f I_s \quad (8)$$

Combining equations (6), (7), and (8) gives an equation dependent upon only I_0 and A_f:

$$I(t) = \frac{I_0(f(t) + A_f)}{1 + A_f} \quad (9)$$

which reduces to equation (4) when $A_f = 0$. So, using equation (9), the model can be run for several values of A_f to determine the effects of different amounts of downward atmospheric flux.

Acknowledgments. We thank Ken Herkenhoff and Ken Edgett for their helpful and detailed reviews, Doug Nash, Dewey Muhleman, and Andy Ingersoll for helpful comments on the manuscript; Joan Hayashi for providing digital versions of the Hayashi et al. [this issue] inertias; and Bruce Jakosky, David Paige, and Hugh Kieffer for insightful discussions. We also thank A. Selivanov, M. Naraeva, V. Kharlamov, and Y. Gektin for assistance with the data, and R. Eby, R. McLendon, K. Hudelson, and S. Welch for assistance with tasks at SJI. Funding for this research was provided by NASA grants NAGW-1426 and NAGW-2491. Division of Geological and Planetary Sciences, California Institute of Technology contribution 5384.

References

Betts, B. H., Edited Termoskan and raw data files and descriptions, on the *PDS (Planetary Data System) Phobos '88 CD-ROM*, vol. PHB-1001, May 1992.

Betts, B. H., Thermal and visible studies of Mars using the Termoskan data set, Ph.D. thesis, Calif. Inst. of Technol., Pasadena, 1993.

Betts, B. H., T. Svitek, M. L. Santee, B. C. Murray, D. Crisp, D. A. Paige, M. Naraeva, and A. Selivanov, Preliminary quantitative assessment and analysis of Phobos 88 Termoskan observations of Mars (abstract), *Lunar Planet. Sci., XXI*, 77-78, 1990.

Christensen, P. R., Martian dust mantling and surface composition: Interpretation of thermophysical properties, *J. Geophys. Res.*, *87*, 9985-9998, 1982.

Christensen, P. R., The distribution of rocks on Mars (abstract), *Lunar Planet. Sci., XIV*, 109-110, 1983.

Christensen, P. R., The spatial distribution of rocks on Mars, *Icarus*, *68*, 217-238, 1986a.

Christensen, P. R., Regional dust deposits on Mars: Physical properties and global distribution, *J. Geophys. Res.*, *91*, 3534-3546, 1986b.

Clifford, S. M., C. J. Bartels, and E. P. Rubenstein, The Mars thermal model (MARSTHERM): A FORTRAN 77 finite-difference program designed for general distribution, Lunar and Planetary Institute, Houston, Tex., 1987.

Haberle, R. M., and B. M. Jakosky, Atmospheric effects on the remote determination of thermal inertia on Mars, *Icarus*, *90*, 187-204, 1991.

Hayashi, J. N., B. M. Jakosky, and R. M. Haberle, Atmospheric effects on the mapping of Martian thermal inertia and thermally derived albedo, *J. Geophys. Res.*, this issue.

Jakosky, B. M., On the thermal properties of the Martian fines, *Icarus*, *99*, 117-124, 1986.

Kieffer, H. H., S. C. Chase Jr., E. Miner, G. Münch, and G. Neugebauer, Preliminary report on infrared radiometric measurements from the Mariner 9 spacecraft, *J. Geophys. Res.*, *78*, 4291-4312, 1973.

Kieffer, H. H., T. Z. Martin, A. R. Peterfreund, B. M. Jakosky, E. D. Miner, and F. D. Palluconi, Thermal and albedo mapping of Mars during the Viking primary mission, *J. Geophys. Res.*, *82*, 4249-4291, 1977.

Martin, T. Z. and M. I. Richardson, New dust opacity mapping from Viking Infrared Thermal Mapper Data, *J. Geophys. Res.*, *98*, 10941-10949, 1993.

Murray, B. C., et. al., Preliminary assessment of Termoskan observations of Mars, *Planet. Space Sci.*, *39*(1/2), 237-265, 1991.

Palluconi, F. D., and H. H. Kieffer, Thermal inertia mapping of Mars from 60°S to 60°N, *Icarus*, *45*, 415-426, 1981.

Pleskot, L. K., and E. D. Miner, Time variability of Martian bolometric albedo, *Icarus*, *45*, 179-201, 1981.

Selivanov, A. S., M. K. Naraeva, A. S. Panfilov, Yu. M. Gektin, V. D. Kharlamov, A. V. Romanov, D. A. Fomin, and Ya. Ya. Miroshnichenko, Thermal imaging of the surface of Mars, *Nature*, *341*, 593-595, 1989.

Zimbelman, J. R., and H. H. Kieffer, Thermal mapping of the northern equatorial and temperate latitudes of Mars, *J. Geophys. Res.*, *84*, 8239-8251, 1979.

B. H. Betts., San Juan Capistrano Research Institute, 31872 Camino Capistrano, San Juan Capistrano, CA 92675. (e-mail: betts@sji.org)

B. C. Murray, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125. (e-mail: bcm@earth1.gps.caltech.edu)

T. Svítek, Orbital Sciences Corporation, 21700 Atlantic Blvd., Dulles, VA 20166.

(Received March 1, 1994; revised January 13, 1995; accepted January 13, 1995.)

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5297–5307, MARCH 25, 1995

Mid-infrared transmission spectra of crystalline and nanophase iron oxides/oxyhydroxides and implications for remote sensing of Mars

James F. Bell III[1]
National Research Council/NASA Ames Research Center, Space Science Division, Moffett Field, California

Ted L. Roush[2]
Department of Geosciences, San Francisco State University, San Francisco, California

Richard V. Morris
Planetary Science Branch, NASA Johnson Space Center, Houston, Texas

Abstract. Ferric-iron-bearing materials play an important role in the interpretation of visible to near-IR Mars spectra, and they may play a similarly important role in the analysis of new mid-IR spacecraft spectral observations to be obtained over the next decade. We review existing data on mid-IR transmission spectra of ferric oxides/oxyhydroxides and present new transmission spectra for ferric-bearing materials spanning a wide range of mineralogy and crystallinity. These materials include 11 samples of well-crystallized ferric oxides (hematite, maghemite, and magnetite) and ferric oxyhydroxides (goethite, lepidocrocite). We also report the first transmission spectra for purely nanophase ferric oxide samples that have been shown to exhibit spectral similarities to Mars in the visible to near-IR and we compare these data to previous and new transmission spectra of terrestrial palagonites. Most of these samples show numerous, diagnostic absorption features in the mid-IR due to Fe^{3+}–O^{2-} vibrational transitions, structural and/or bound OH, and/or silicates. These data indicate that high spatial resolution, moderate spectral resolution mid-IR ground-based and spacecraft observations of Mars may be able to detect and uniquely discriminate among different ferric-iron-bearing phases on the Martian surface or in the airborne dust.

Introduction and Background

Ferric (Fe^{3+}) iron-bearing minerals are present on the surface of Mars. Information regarding the crystalline or amorphous nature of the iron-bearing (and other) surface materials on Mars can provide insight into the availability of liquid water at the surface and the duration, mode, and extent of chemical weathering. The presence of ferric iron has been established by a variety of remote sensing and in situ analysis measurements performed over the past four decades by Earth-based telescopes and the Mariner, Viking, and Phobos 2 spacecraft missions [e.g., *Dollfus*, 1955; *Sharanov*, 1961; *Adams and McCord*, 1969; *McCord and Westphal*, 1971; *Hargraves et al.*, 1977; *Toulmin et al.*, 1977; *Clark et al.*, 1982; *Pollack et al.*, 1977; *Soderblom et al.*, 1978; *Singer*, 1982; *Bell et al.*, 1990; *Erard et al.*, 1991; *Bell*, 1992; *Soderblom*, 1992; *Murchie et al.*, 1993; *Roush et al.*, 1993]. However, since the Viking Landers did not carry any instruments capable of determining mineralogy, the identities of the specific iron-bearing phases remain uncertain.

The existing remote sensing data indicate that Mars ferric minerals exist in two distinct but co-existing phases: a poorly crystalline or perhaps even amorphous phase similar to nanophase ferric oxides [e.g., *Morris et al.*, 1989; *Morris and Lauer*, 1990; *Banin et al.*, 1993] and much more crystallized ferric oxide phases such as hematite and magnetite [e.g., *Pollack et al.*, 1977; *Morris et al.*, 1989; *Bell et al.*, 1990]. Analysis of visible to near-IR spectra and comparisons to laboratory mineral samples and spectral analogs have shown that these two phases are intimately associated: the bulk ferric oxides seem to occur within a matrix of the more fine-grained, possibly nanophase, ferric oxides, with the former accounting for specific charge-transfer absorption features and the latter being the primary pigmenting agent in the Martian soils [e.g., *Morris et al.*, 1990, 1993; *Bell et al.*, 1993; *Golden et al.*, 1993].

In the near- to mid-infrared (400-4000 cm^{-1}; 2.5-25 μm), spectral features arise from vibrational motions of atoms and molecules which comprise the materials. Fundamental modes of these vibrations yield IR absorption features that are much more intense than any associated combination and/or overtones of these modes, and hence IR remote sensing observations are extremely sensitive to minor concentrations

[1]Also at University of Washington, Geological Remote Sensing Laboratory, Seattle.
[2]Also at NASA Ames Research Center, Space Science Division, Moffett Field, California.

Paper number 94JE01389.
0148-0227/95/94JE-01389$05.00

of these absorbing species [e.g., *Farmer*, 1974; *Ryskin*, 1974; *Gadsden*, 1975; *Salisbury et al.*, 1991]. Mid-IR ground-based and spacecraft observations of Mars have been presented and discussed by *Hunt et al.* [1973], *Toon et al.* [1977], *Pollack et al.* [1990], *Roush et al.* [1991, 1992], and *Moersch et al.* [1993] among others.

This study represents one aspect of a broader project intended to provide data in the mid- and far-infrared spectral regions for both well-characterized iron oxides/oxyhydroxides and poorly crystalline or amorphous materials like palagonites and nanophase ferric oxides [*Roush*, 1992; *Bell and Roush*, 1993; *Roush and Bell*, this issue]. In this paper we review the currently available transmission data for many ferric oxide/oxyhydroxide minerals that could plausibly be considered to exist on Mars [e.g., *Bell*, 1992], present new transmission measurements from a suite of well-characterized ferric-bearing materials, and discuss the implications of these data on the interpretations of existing and potential future remote sensing observations of Mars.

Existing Data

Mid-IR transmission spectra of uncharacterized samples of hematite (α-Fe_2O_3), goethite (α-FeOOH), lepidocrocite (γ-FeOOH), akaganeite (β-FeOOH), and magnetite (Fe_3O_4) were discussed and reviewed by *Nyquist and Kagel* [1971], *Hunt et al.* [1973], *Farmer* [1974], *Ryskin* [1974], and *Gadsden* [1975]. The results in the mid-IR can be summarized as follows: hematite spectra showed strong absorption features near 633, 544, and 465 cm^{-1} (15.8, 18.4, and 21.5 μm); goethite spectra showed features near 3125, 893, 800, 667, 592, and 461 cm^{-1} (3.2, 11.2, 12.5, 15.0, 16.9, and 21.7 μm); lepidocrocite spectra showed bands near 3333, 1163, 752 cm^{-1} (3.0, 8.6, and 13.3 μm); akaganeite has absorption bands near 3333, 847, 676, and 450 cm^{-1} (3.0, 11.8, 14.8, and 22.2 μm); and magnetite has absorption bands near 578 and 380 cm^{-1} (17.3 and 26.3 μm). Little effort was made in these reviews to assign a specific origin to these features, except for the OH-bearing species. For those samples, bands near 3333 cm^{-1} (3.0 μm) were assigned to the OH stretching fundamental; bands near 893 to 1163 cm^{-1} (8.6 to 11.2 μm) were assigned to OH in-plane bending modes; and bands near 752 to 800 cm^{-1} (12.5 to 13.3 μm) were assigned to OH out-of-plane bending modes [*Ryskin*, 1974]. The anhydrous phases exhibited many fewer absorption features in the mid-IR than the OH-bearing phases.

More recently, *Salisbury et al.* [1991] have presented mid-IR transmission and reflectance spectra for four petrographically and chemically analyzed samples of hematite, goethite, and magnetite. The major bands identified from our analysis of their digital data are similar to those discussed above and are included in Table 1 below.

New Samples

Because there can be considerable physical, chemical, and structural variability among samples, we obtained transmission measurements of a different set of well-characterized crystallized ferric materials for comparison with previously obtained lab data and currently available mid-IR Mars remote sensing data. We also report the first mid-IR transmission measurements of purely nanophase ferric oxides shown to be good spectral analogs to Mars at visible wavelengths. Finally, we measured the transmission spectrum of a naturally occurring palagonite sample for comparison with the nanophase and well-crystallized ferric materials and with previously reported transmission measurements of palagonites [*Roush*, 1992; *Crisp and Bartholomew*, 1992; *Roush et al.*, 1993].

The well-crystallized ferric oxide samples include four hematites (HMS3, HMS13, HMS14, HMS15), three goethites (GTS2, GTS3, GTS5), two maghemites (γ-Fe_2O_3; MHS3, MHS4), one lepidocrocite (LPS2), and one magnetite (MTS4). All of these samples are submicron powders that have been chemically and physically characterized by *Morris et al.* [1985] using a variety of laboratory techniques including Mössbauer and visible to near-IR reflectance spectroscopy, X-ray fluorescence, X-ray diffraction (XRD), thermogravimetric analysis (TGA), vibrating sample magnetometry (VSM), and transmission electron microscopy. All of the samples were found to be reasonably pure, with the major impurities being SiO_2 and SO_3 which varied in abundance between samples at the 0.5 to 2.5% level [*Morris et al.*, 1985].

The nanophase iron oxide samples include two pure synthetic samples that have been heated to different temperatures (TNA13 at 210° and 260°C, and TNA30 at 215° and 290°C), one synthetic ferrihydrite sample (S11-4-DGS), and one sample of nanophase ferric oxide that was dispersed in a silica gel matrix (S6FN28). The pure TNA13 and TNA30 samples were prepared by oxidation in air of trinuclear acetato hydroxy iron(III) nitrate as described by *Morris et al.* [1991]. The low-temperature samples (210° and 215°C) have been characterized as nanocrystalline by Mössbauer and reflectivity measurements that do not show any indication of crystallized ferric oxide phases [*Morris et al.*, 1991]. This same study showed that the higher-temperature samples (260° and 290°C) begin to show evidence of crystalline hematite. *Golden et al.* [1994] have determined a mean particle size of 6±1 nm for TNA13/210 and 8.5±1 nm for TNA13/260 using high-resolution transmission electron microscopy techniques. These authors also used electron diffraction to determine that the composition of the TNA13 nanophase particles is hematite and/or ferrihydrite. The S11-4-DGS synthetic ferrihydrite sample is a 2-line ferrihydrite (verified by Mössbauer spectroscopy), prepared and provided to us by D.G. Schulze. The S6FN28 sample dispersed in silica gel was prepared by impregnating gel particles having 6-nm-diameter pores with a ferric nitrate solution, followed by calcination at 550°C [*Morris et al.*, 1989]. The sample was characterized chemically and physically by *Morris et al.* [1989] using Mössbauer and visible to near-IR reflectance spectroscopy, instrumental neutron activation analysis, XRD, TGA, and VSM. The same techniques were used by *Bell et al.* [1993] to characterize the naturally occurring palagonitic tephra sample PH-1. This sample underwent post emplacement thermal alteration, which oxidized some of the preexisting basaltic tephra particles. The ferric iron mineralogy of the sample measured here, from the finest size fraction of PH-1 (< 20 μm), was found to be dominated by nanophase ferric oxide, with minor amounts of well-crystallized hematite and magnetite. Other phases identified in the sample include olivine, pyroxene, plagioclase, and very minor smectite clay [*Bell et al.*, 1993].

Measurement Technique

A small amount (≈1 mg) of each sample was mixed with KBr powder while both were immersed in alcohol. After the al-

cohol evaporated, 200 mg of the resultant mixture was pressed into a cohesive pellet using a standard 13-mm KBr pellet die. Using the facilities at the NASA Ames Astrochemical Laboratory, with the cooperation of Robert Bohn, Scott Sandford, and Louis Allamandolla, the transmission spectrum of each pellet was obtained at a resolution of 4 cm^{-1} using a Nicolet 7199 Fourier transform spectrometer in the 4000 to 400 cm^{-1} spectral region, and was subsequently ratioed to the transmission of a pure KBr pellet.

Results

Well-Crystallized Iron Oxides/Oxyhydroxides

Our transmission measurements of pure crystalline ferric oxide and oxyhydroxide powders are shown in Figures 1 and 2. A compilation of the major spectral features in these data is presented in Table 1. The transmission spectra of the four hematite samples (Figure 1) show strong absorptions in the 450 to 600 cm^{-1} (16 to 22 μm) region separated by a narrow maximum near 500 cm^{-1} (20 μm). These spectra are consistent with previously reported data for hematite [e.g., *Farmer*, 1974; *Gadsden*, 1975; *Salisbury et al.*, 1991]. Because of the known purity of the samples and the relative weakness of transmission features due to adsorbed or structurally contaminating water, it is likely that the 450 to 600 cm^{-1} bands are the result of Fe^{3+}–O^{2-} vibrational transitions in the hematite hexagonal close packed (hcp) structure [*Klein and Hurlbut*, 1985]. The strong drop-off in transmission towards the highest frequencies for samples HMS13 and HMS15 most likely reflects the decrease in signal due to scattering of light out of the incident beam as the wavelength of the incident radiation begins to approach the mean particle size of the samples. *Morris et al.* [1985] showed that both of these samples have a much larger mean particle size than the other two hematites studied here; this observation is consistent with our scattering interpretation.

The transmission spectra of our two maghemite samples (Figure 1) are substantially more complex. There is a broad similarity to the hematite spectra in that there are strong absorption features in the 440 to 700 cm^{-1} region with a maximum near 500 cm^{-1}, but these strong bands exhibit much more detailed fine structure than the similarly located features in hematite. Since again there is little evidence for a major spectral influence of water or OH on the spectra, we would conclude that these detailed and apparently characteristic spectral features are caused by Fe^{3+}–O^{2-} vibrational transitions specific to this cubic close packed (ccp) polymorph of hematite.

Spectra of our three goethite samples (Figure 2) show more and narrower features below 1000 cm^{-1} than any of the anhydrous iron oxides in Figure 1. In addition, a strong OH vibrational feature near 3120 cm^{-1} is apparent. These goethite spectra are consistent with previously reported spectra [e.g., *Ryskin*, 1974; *Gadsden*, 1975; *Salisbury et al.*, 1991]. The bands near 900 cm^{-1} and 800 cm^{-1} have been interpreted as in-plane and out-of-plane OH bending transitions by *Ryskin* [1974]. Based on these assignments and the measurements of the anhydrous phases discussed above, we interpret the remaining strong absorption features near 430 and 600 cm^{-1} as Fe^{3+}–O^{2-} vibrational transitions diagnostic of the goethite hcp crystal structure [*Sherman et al.*, 1982; *Klein and Hurlbut*, 1985].

The spectrum of our lepidocrocite sample (Figure 2) is similar to the goethite spectra in that there are many narrow features at the lowest frequencies and a very strong OH vibrational feature near 3100 cm^{-1}, but there are important differences. First, the OH vibrational band is much broader in lepidocrocite than it is in goethite. Second, bands interpreted as the in-plane and out-of-plane OH bending transitions [*Ryskin*, 1974] near 1160 and 750 cm^{-1} are shifted substantially between the γ and α oxyhydroxide phases. And third, a complex of bands occurs near 420 to 600 cm^{-1} that may be due to Fe^{3+}–O^{2-} vibrational transitions and which are shifted relative to the Fe^{3+}–O^{2-} band positions of other ferric oxides and oxyhydroxides. The difference between the lepidocrocite and goethite spectra must be due in large part to structural differences between these hcp and ccp polymorphs [e.g., *Schwertmann and Taylor*, 1989].

While we do not reproduce the data here, *Nyquist and Kagel* [1971] show a spectrum of akaganeite (the β phase polymorph of goethite) that also shows a complex spectrum exhibiting many features similar to those seen for goethite and lepidocrocite but occurring at slightly different frequencies (Table 1). It appears that there are enough systematic structural differences between goethite and its structural polymorphs to allow each to have very distinct and diagnostic mid-IR spectral signatures.

The transmission spectrum of our final crystalline sample, magnetite, is shown in Figure 2. This spectrum is similar to a previously measured spectrum by *Salisbury et al.* [1991] in that it shows very few spectral features. Only one strong band is seen near 580 cm^{-1} (17.2 μm) which, because of the lack of evidence for OH or water contamination, is most likely due to Fe^{3+}–O^{2-} vibrational transitions in the cubic magnetite structure. The steep drop-off in transmission at the highest frequencies is most likely caused by the general visible to near-IR opacity of magnetite (resulting from very strong Fe^{2+} and Fe^{3+} charge transfer transitions; e.g., *Morris et al.*, 1985), although an additional contribution from the increasing effects of scattering cannot be ruled out.

None of the samples show strong absorption features associated with any of the low-level impurities (SiO_2, SO_3, etc.) measured by *Morris et al.* [1985]. Weak features near 1000 to 1100 cm^{-1} are most likely caused by the small amount of SiO_2 in these samples, which is the most abundant impurity besides water. Other weak features near 1050 cm^{-1} and 2900 to 3000 cm^{-1} are probably caused by C–H vibrations from residual alcohol used in the sample preparation procedure.

Nanophase Ferric Oxides and Palagonites

Spectra of our nanophase ferric oxide and palagonite samples are shown in Figures 3 and 4. Also shown in Figure 4 are transmission spectra of other palagonite samples obtained previously by *Roush* [1992] and *Crisp and Bartholomew* [1992]. The spectra of these materials are substantially different from those of the well-crystallized ferric oxides/oxyhydroxides. The primary differences are that these materials do not exhibit as many features as the crystalline materials and the spectral contrast of the features that exist in the nanophase and some of the palagonite samples is much lower than for the well-crystallized samples. Most of the samples show distinct absorptions in the 450 to 1200 cm^{-1} (8.3 to 22.2 μm) region. By comparison with the crystalline sample spectra, it is probable that bands occurring in the 450

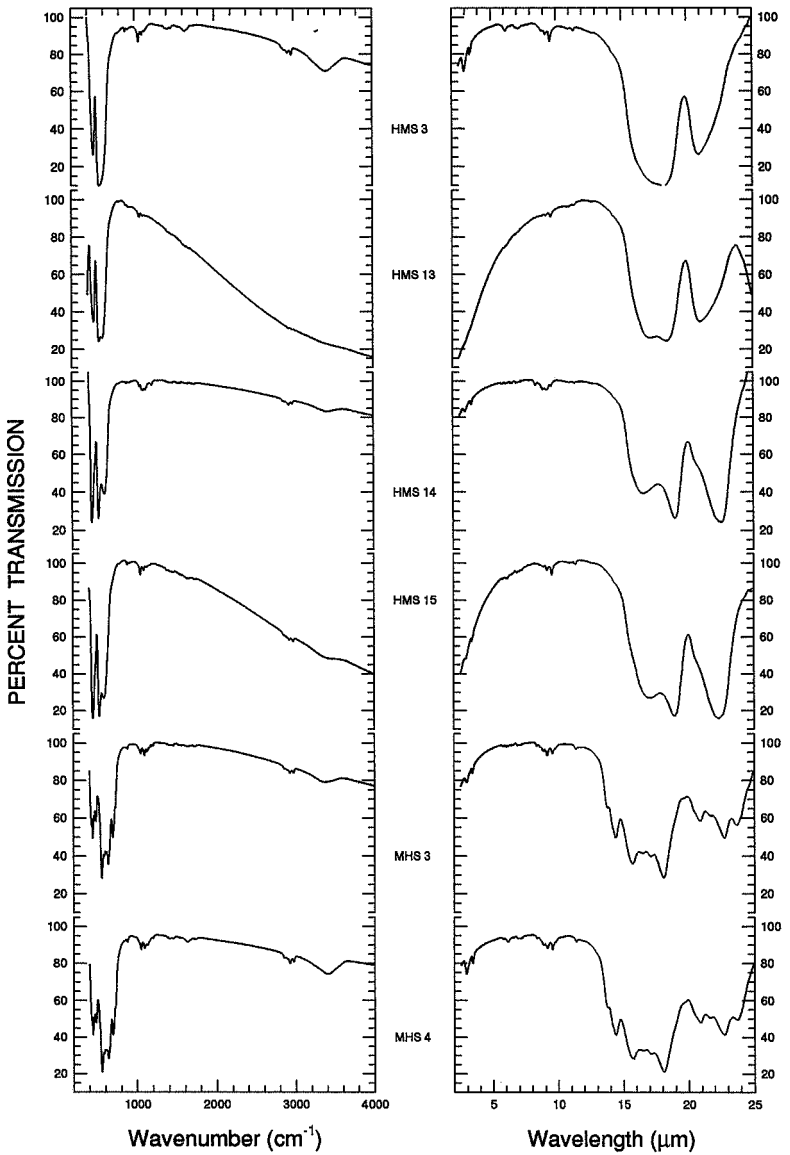


Figure 1. Transmission spectra of well-characterized and well-crystallized submicron powders of the iron oxides hematite (HMS series) and maghemite (MHS series). Data are plotted in frequency on the left and wavelength on the right.

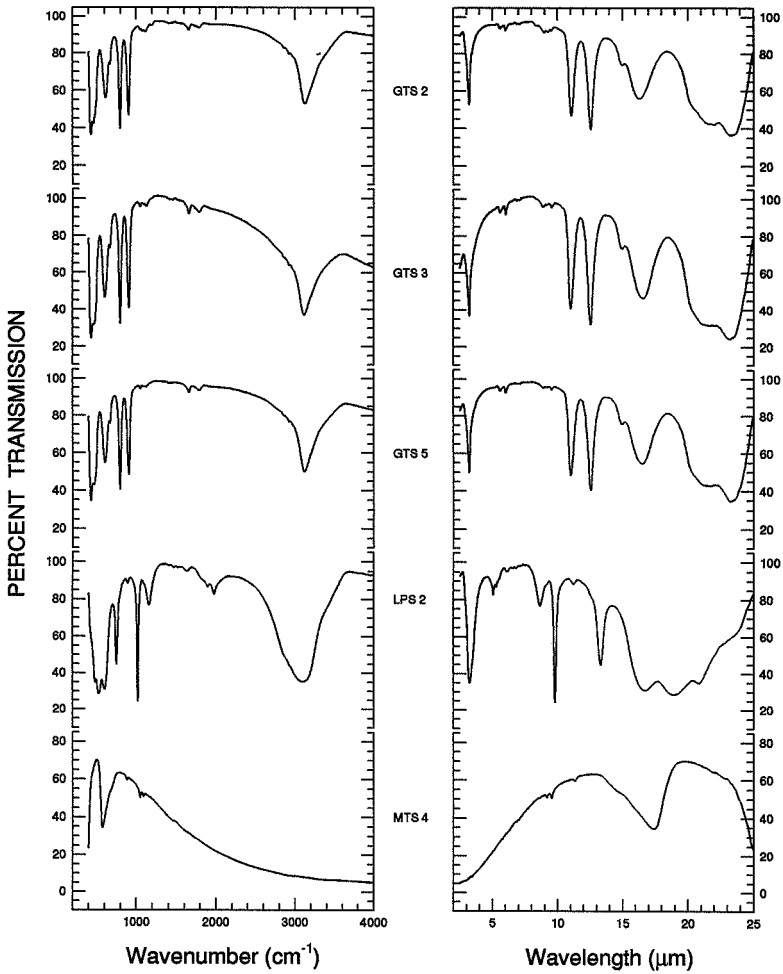


Figure 2. Transmission spectra of well-characterized and well-crystallized submicron powders of the iron oxyhydroxides goethite (GTS series) and lepidocrocite (LPS2). A magnetite (MTS4) spectrum is shown at the bottom. Data are in frequency on the left and wavelength on the right.

to 600 cm^{-1} region result from Fe^{3+}–O^{2-} vibrational transitions in the very poorly crystalline or even amorphous structures of these materials. This interpretation is supported by examining the systematic differences in the spectra of the TNA series samples (Figure 3) that were heated to different temperatures. For example, TNA13/210, shown by *Morris et al.* [1991] to contain only nanophase hematite, shows a moderately strong band near 598 cm^{-1} and a shoulder near 450 cm^{-1}. However, TNA13/260, which shows evidence for a small amount of crystalline hematite in Mössbauer spectra [*Morris et al.*, 1991], has stronger, more well-defined features near 455 and 544 cm^{-1}. A similar pattern is seen for the TNA30 sample. It appears, then, that not only is there evidence in these nanophase materials for Fe^{3+}–O^{2-} vibrational transitions at these mid-IR frequencies, but that there are systematic variations in band position that may be characteristic of the nanophase crystal structure (Table 1).

Most of the nanophase samples show a general decrease in transmission towards the higher frequencies. As in the well-crystallized samples above, this effect is most likely a

Table 1. Major Features in Mid-IR Transmission Spectra of Ferric Materials

Sample	Frequency of Strongest Transmission Features, cm^{-1}	Notes
	Well-Crystallized Hematite (α-Fe_2O_3)	
HMS3	**478,548**,1408,1630,**3402**	[1]
HMS13	**476,542**d	[1]
HMS14	**444,525,602**,3408	[1]
HMS15	**447,528**,1458,1653,3467	[1]
NMNHB7373	**467,546**,1037,1635,**3492**	[2]
	Well-Crystallized Maghemite (γ-Fe_2O_3)	
MHS3	**440**t,**553**d,**636**d,694d,879,1049d,1093t,2927t,3350	[1]
MHS4	**442**t,**553**d,**636**d,694d,879,1051d,1093t,1637,2925t,**3408**	[1]
	Well-Crystallized Magnetite ($Fe^{2+}Fe_2^{3+}O_4$)	
MTS4	**577**,700s,881,1051,1092	[1]
NMNH114111	**579**	[2]
	Well-Crystallized Goethite (α-FeOOH)	
GTS2	**430**d,**613**d,**796,904**,1655,1792,**3118**	[1]
GTS3	**432**d,**602**d,**798,908**,1662,1792,**3116**	[1]
GTS5	**430**d,**606**d,**796,908**,1660,1790,**3116**	[1]
NMNH152500	**461**d,**590**d,667,**799,901**,1065,1659,1788,**3098**,3430	[2]
	Well-Crystallized Lepidocrocite (γ-FeOOH)	
LPS2	420s,**480,528,598,752**,893,**1020,1159**,1637,**1979**d,**3099**	[1]
	Well-Crystallized Akaganeite (β-FeOOH)	
Sample 386	**450,675,850**,1150,**1340**t,1520,1625,2300,**3350**d	[3]
	Poorly-Crystallized or Nanophase Ferric Oxides	
PH-1 < 20 μm	**467**,577,791,**1099,1643**,3427	[4]
TNA13/210	450s,**602**d,1049d,1458,2974t	[5,6]
TNA13/260	**444,536**d,3410	[5,6]
TNA30/215	460s,**571**d,1049d,**1421,1541**,3367	[5]
TNA30/290	**453,534**d,1049,3367	[5]
S6FN28	**465**,590s,**802,1092**d,1635,2976t,**3427**	[7]
S6 Silica Gel	**465**,590s,**798,1090**d,1635,2974t,**3423**	[7]
S11-4-DGS	**459,598,1385**	[8]

Strong, well-characterized bands are shown in boldface. An "s" after the frequency means that the band occurs as a shoulder with an approximate center at the given value. A "d" or a "t" after the frequency indicates that the feature is a doublet or a triplet. Only the frequency of the strongest absorption is listed for close doublets or triplets.

Notes: [1], Samples characterized by *Morris et al.* [1985]; [2], Samples measured and characterized by *Salisbury et al.* [1991], band minima determined by us; [3], Sample measured and characterized by *Nyquist and Kagel* [1971]; [4], Sample characterized by *Bell et al.* [1993]; [5], Samples characterized by *Morris et al.* [1991]; [6], Additional sample characterization by *Golden et al.* [1994]; [7], Samples characterized by *Morris et al.* [1989]; [8], Sample prepared and characterized by D.G. Schulze.

manifestation of the increased effect of scattering at shorter wavelengths. All of the palagonites (Figure 4) and sample S6FN28 (Figure 3) show a relatively strong band centered near 1000-1100 cm^{-1} (9-10 μm). We interpret this feature as the Si–O vibrational transition common in silicate minerals. For S6FN28 this band arises because the nanophase ferric particles were impregnated in a silica gel matrix [*Morris et al.*, 1989]. After examining the S6FN28 spectrum and comparing it to the spectrum of the silica gel matrix alone (Figure 3), we have been unable to identify any features that are unique to the nanophase ferric oxide component in S6FN28. The 1000-1100 cm^{-1} band in PH-1, 91-1, and the Pahala Ash samples comes from the presence of other silicate minerals in these naturally occurring materials [*Roush*, 1992; *Crisp and Bartholomew*, 1992; *Bell et al.*, 1993]. The band at 1643 cm^{-1} (6.1 μm) in PH-1 may also result from silicate minerals like olivine or pyroxene [e.g., *Nash et al.*, 1993], or from an H–O–H bending fundamental of structural water [*Ryskin*, 1974], all of which are known to exist in minor amounts in PH-1. A weak feature at a similar position can also be seen in the S6FN28 and silica gel spectra. Bands centered near 3400 cm^{-1} in the 91-1 and Pahala Ash spectra result from the OH stretching fundamental of structural water in these samples. This feature is absent in PH-1 presumably because that palagonite sample had much of its water driven off by a thermal alteration event [*Bell et al.*, 1993].

Discussion and Implications

Transmission spectra such as these cannot be used to directly deduce the expected remote-sensing signatures from ferric-iron-bearing materials on the Martian surface or in the airborne dust because the samples were measured in an environment (pressed KBr pellet) that is nothing like that which is likely to occur in nature. Specifically, these spectra do not reveal the same important and diagnostic effects of scattering

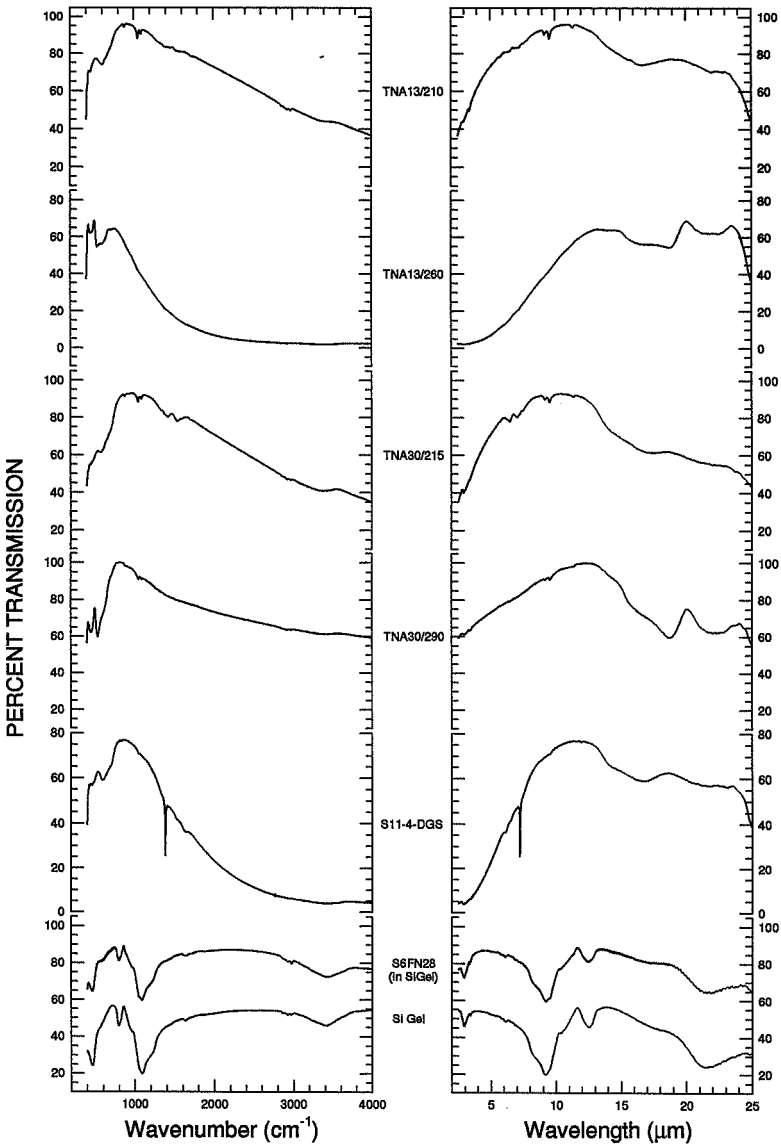


Figure 3. Transmission spectra of well-characterized nanophase ferric materials. The TNA series nanophase iron oxide samples were produced in the lab by oxidation in air of trinuclear acetato hydroxy iron(III) nitrate [*Morris et al.*, 1991], the S11-4-DGS synthetic ferrihydrite sample was prepared by D.G. Schulze, and the S6FN28 nanophase iron oxide sample was produced in the lab by calcining silica gel impregnated with ferric nitrate solutions in air at 550°C [*Morris et al.*, 1989]. The transmission spectrum of S6FN28 is shown along with a spectrum of the silica gel material alone that has been offset by 30% relative to S6FN28. Data are in frequency on the left and wavelength on the right.

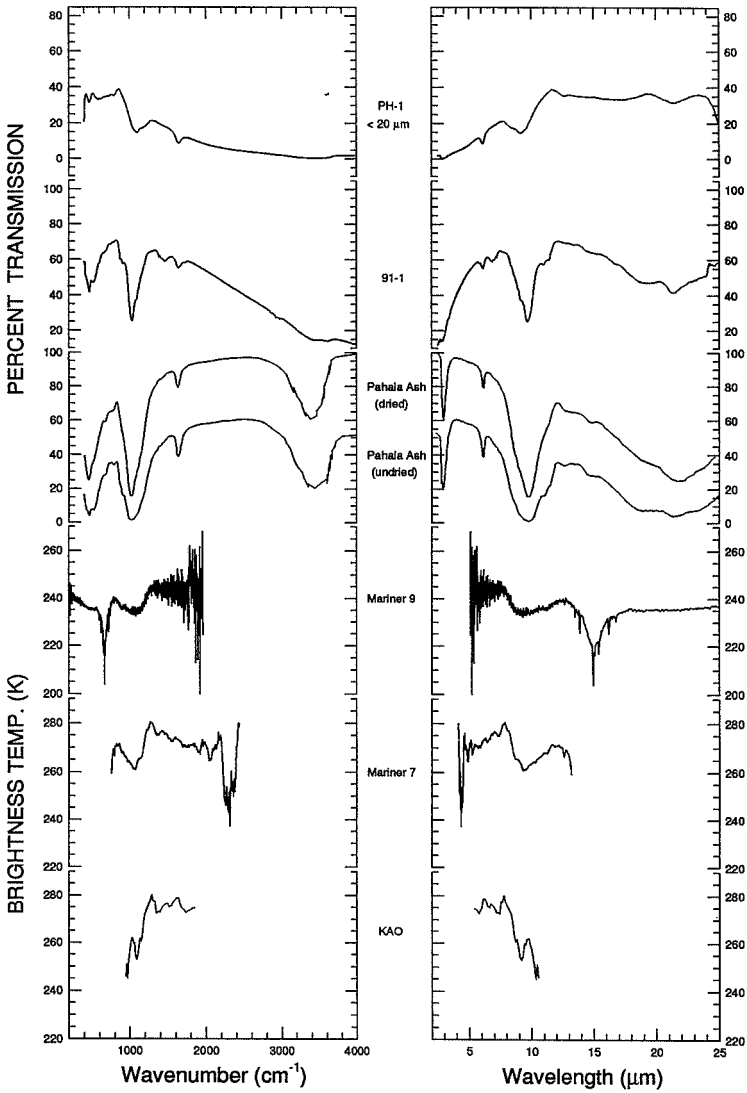


Figure 4. Transmission spectra of Mars-analog palagonites from the island of Hawaii compared to mid-IR Mars spectra from Mariner 7, Mariner 9, and the KAO. PH-1 < 20 µm is a spectrum of the finest seive fraction of a naturally occurring thermally altered palagonitic tephra from Mauna Kea [*Bell et al.*, 1993]. Spectrum 91-1 is from a fine fraction (< 38 µm) sample of Pahala Ash collected near South Point and measured by *Roush et al.* [1992]. The two other Pahala Ash spectra are from a different sample of the Pahala Ash unit than 91-1 and were measured by *Crisp and Bartholomew* [1992]. One of the spectra is from a subsample of Pahala Ash dried in the laboratory. The Mars spectra are plotted as brightness temperature, and representative spectra are shown from Mariner 9 IRIS [*Hanel et al.*, 1972]. Mariner 7 IRS (T.Z.Martin, Mariner 6/7 infrared spectrometer: Data set restoration, submitted to *Journal of Geophysical Research*, 1991), and 1988 Kuiper Airborne Observatory observations (*Pollack et al.*, 1990]. Data are in frequency on the left and wavelength on the right.

that will certainly be evident in real Martian data [see *Salisbury et al.*, 1991; *Salisbury and Wald*, 1992; *Nash et al.*, 1993]. However, this study has demonstrated that both crystallized and nanophase ferric oxides and ferric oxyhydroxides exhibit complex spectral features at mid-IR wavelengths. Well-crystallized ferric oxides/oxyhydroxides have strong, well-resolved absorption features primarily in the 400 to 1200 cm^{-1} (8.3 to 25.0 μm) region. Analysis by *Roush and Bell* [this issue] of the *Salisbury et al.* [1991] mid-IR reflectance spectra of crystalline hematite, goethite, and magnetite and the assumption of Kirchoff's law (emissivity = 1 – reflectivity) indicates that the absorption features seen in our transmission spectra of well-crystallized samples should be detectable as high-contrast features in remotely sensed Mars emissivity spectra. The nanophase materials and palagonites, however, typically show much weaker mid-IR absorption features. By inference then, we can expect that emissivity features in the spectra of these poorly-crystallized materials both in the lab and potentially in the Martian environment will be subdued relative to those seen in the crystalline materials. In fact, the subtle nature of emissivity features in palagonite samples PH-1 and 91-1 has been demonstrated by *Roush and Bell* [this issue]. Thus, our transmission measurements provide a starting point for understanding how to interpret one aspect of existing and future Mars mid-IR observations in that they provide information on the variations in absorption band positions of ferric iron bearing materials as a function of composition and structural environment.

Figures 1-4 clearly indicate that ferric-bearing materials spanning a wide range of mineralogy and crystallinity exhibit many strong and diagnostic absorption features in the mid-IR. To date, no conclusive evidence exists for the presence of any of the Fe^{3+}–O^{2-} features in ground-based or spacecraft Mars data sets (Figure 4). *Hunt et al.* [1973] conducted a preliminary examination of the Mariner 7 and Mariner 9 spectra and included hematite and goethite in their spectral searches. They favored montmorillonite as explaining the 9.5- and 21-μm features seen in the Mars data, although the presence of kaolinite and, to a lesser extent, hematite and goethite, could not be completely excluded. *Toon et al.* [1977] searched many of the Mariner 9 dust storm spectra for evidence of ferric oxides or oxyhydroxides. They found no conclusive evidence for hematite absorption features below 600 cm^{-1}, but were not able to rule out the possible occurrence of substantial amounts of goethite either mixed with or coating other materials.

It is important to consider how the Martian atmosphere could potentially interfere with the interpretation of ferric mineral bands in mid-IR remote sensing observations. Radiative transfer models show that CO_2 accounts for all of the major mid-IR absorption features in the Martian atmosphere [e.g., *Crisp*, 1990]. Strong CO_2 features occur near 550 to 750 cm^{-1}, 2000 to 2400 cm^{-1}, and 3400 to 3800 cm^{-1}. The two higher-frequency bands are not likely to cause problems in the detection of ferric materials, but the lowest frequency band (15 μm) will be a concern because of overlap with Fe^{3+}–O^{2-} vibrational transitions (Table 1; Figure 4). Because these absorption features occur primarily in the unsaturated wing at frequencies lower than the center of the CO_2 band, however, it will most likely be possible to model the predicted atmospheric spectral response and remove or at least understand its signature in the spectra of surface materials and atmospheric dust [e.g., *Toon et al.*, 1977; *Pollack et al.*, 1990]. But because of CO_2 band saturation, it is unlikely that any useful surface spectral signatures could be detected from 630 to 700 cm^{-1}, except possibly from surface landers.

Experience with high spatial and spectral resolution ground-based and spacecraft visible to near-IR spectroscopic observations indicates that increasing the spatial and spectral resolution allows for more detailed characterization of spectral heterogeneity on Mars [e.g., *Bell et al.*, 1990; *Murchie et al.*, 1993; *Mustard and Bell*, 1994]. Ferric minerals are an obvious and important component of the optical surface of Mars, and high spatial and spectral resolution ground-based or spacecraft mid-IR surveys like those of *Moersch et al.* [1993] or of the Mars Observer/Surveyor TES instrument [*Christensen et al.*, 1992] will have a much greater ability than previous investigations to detect and spectrally characterize the Martian surface. An understanding of the spectral contribution of ferric-bearing materials in the mid-IR will be an important aspect of the analysis and interpretation of these new data sets.

Conclusions

Our results suggest that high spatial resolution, moderate spectral resolution thermal infrared observations of Mars may provide significant insight into the ferric mineralogy of that planet. The subdued nature of the spectral features in the poorly crystallized samples indicates that any future mid-IR remote sensing investigations must be able to achieve a high signal-to-noise ratio in order to detect and characterize such materials. Continued laboratory study of the reflectance and emissivity behavior of crystalline and nanophase ferric oxides/oxyhydroxides and other iron-bearing analog materials like palagonites may prove quite important for the interpretation of mid-IR remote sensing spectral observations of Mars.

Acknowledgments. This research was supported by grants from the National Research Council/National Academy of Sciences and the NASA Planetary Geology Program. We thank Robert Bohn, Scott Sandford and Louis Allamandolla for assistance with and access to the NASA Ames Astrochemical Laboratory FTIR spectrometer facility. We also thank Joy Crisp for graciously providing digital versions of the Pahala Ash palagonite transmission spectra. We thank Phil Christensen, Paul Lucey, and Jack Salisbury for constructive reviews of an earlier draft of this paper. We note that the new transmission data presented here are available digitally to interested researchers. Please send an Internet electronic mail request for more information to the first author.

References

Adams, J.B., and T.B. McCord, Mars: Interpretation of spectral reflectivity of light and dark regions, *J. Geophys. Res., 74*, 4851-4856, 1969.

Banin, A., T. Ben Shlomo, L. Margulies, D.F. Blake, R.L. Mancinelli, and A.U. Gehring, The nanophase iron mineral(s) in Mars soil, *J. Geophys. Res., 98*, 20,831-20,853, 1993.

Bell, J.F., III, Charge-coupled device imaging spectroscopy of Mars, 2, Results and implications for Martian ferric mineralogy, *Icarus, 100*, 575-597, 1992.

Bell, J.F., III, and T.L. Roush, Thermal emission measurements (5-25 μm) of Hawaiian palagonitic soils with implications for Mars, Mars: Past, Present, and Future–Results from the MSATT Program, *LPI Tech. Rep., 93-06*, 2-3, 1993.

Bell, J.F., III, T.B. McCord, and P.D. Owensby, Observational evidence of crystalline iron oxides on Mars, *J. Geophys. Res., 95*, 14,447-14,461, 1990.

Bell, J.F., III, R.V. Morris, and J.B. Adams, Thermally altered palagonitic tephra: A spectral and process analog to the soils and dust of Mars, *J. Geophys. Res., 98,* 3373-3385, 1993.

Christensen, P.R., D.L. Anderson, S.C. Chase, R.N. Clark, H.H. Kieffer, M.C. Malin, J.C. Pearl, J. Carpenter, N. Bandiera, F.G. Brown, and S. Silverman, Thermal emission spectrometer experiment: Mars Observer mission, *J. Geophys. Res., 97,* 7719-7734, 1992.

Clark, B.C., A.K. Baird, R.J. Weldon, D.M. Tsusaki, L. Schnabel, and M.P. Candelaria, Chemical composition of Martian fines, *J. Geophys. Res., 87,* 10,059–10,067, 1982.

Crisp, D., Infrared radiative transfer in the dust-free Martian atmosphere, *J. Geophys. Res., 95,* 14,577-14,588, 1990.

Crisp, J., and M.J. Bartholomew, Mid-infrared spectroscopy of Pahala Ash palagonite and implications for remote sensing studies of Mars, *J. Geophys. Res., 97,* 14,691-14,699, 1992.

Dollfus, A., Étude des planètes par la polarisation de leur lumière, Ph.D. thesis, Univ. of Paris, 1955. (Also available as *NASA Tech. Trans., TTF-188,* 1964.)

Erard, S., J.-P. Bibring, J. Mustard, O. Forni, J.W. Head, S. Hurtrez, Y. Langevin, C.M. Pieters, J. Rosenqvist, and C. Sotin, Spatial variations in composition of the Valles Marineris and Isidis Planitia regions of Mars derived from ISM data, *Proc. Lunar Planet. Sci. Conf., 21st,* 437-456, 1991.

Farmer, V.C., The anhydrous oxide minerals, in *The Infrared Spectra of Minerals, Mineral. Soc. Monogr. 4,* (edited by V.C. Farmer) pp. 183-204, Mineralogy Society, London, 1974.

Gadsden, J.A., *Infrared Spectra of Minerals and Related Inorganic Compounds,* 277 pp., Butterworths, London, 1975.

Golden, D.C., R.V. Morris, D.W. Ming, H.V. Lauer Jr., and S.R. Yang, Mineralogy of three palagonitic soils from the summit of Mauna Kea, Hawaii, *J. Geophys. Res., 98,* 3401-3411, 1993.

Golden, D.C., R.V. Morris, D.W. Ming, and H.V. Lauer Jr., High resolution transmission electron microscopy (HRTEM) of nanophase ferric oxides (abstract), *Lunar Planet. Sci. Conf., XXV,* 437-438, 1994.

Hanel, R.B., B. Conrath, W. Hovis, V. Kunde, P. Lowman, W. Maguire, J. Pearl, J. Pirraglia, C. Prabhakara, and B. Schlachman, Investigation of the Martian environment by infrared spectroscopy on Mariner 9, *Icarus, 17,* 423-442, 1972.

Hargraves, R.B., D.W. Collinson, R.E. Arvidson, and C.R. Spitzer, The Viking magnetic properties experiment: Primary mission results, *J. Geophys. Res., 82,* 4547-4558, 1977.

Hunt, G.R., L.M. Logan, and J.W. Salisbury, Mars: Components of infrared spectra and the composition of the dust cloud, *Icarus, 18,* 459-469, 1973.

Klein, C., and C.S. Hurlbut Jr., *Manual of Mineralogy (after James D. Dana),* pp. 295-319, John Wiley, New York, 1985.

McCord, T.B., and J.A. Westphal, Mars: Narrowband photometry, from 0.3 to 2.5 microns, of surface regions during the 1969 apparition, *Astrophys. J., 168,* 141-153, 1971.

Moersch, J., P. Nicholson, S. Squyres, J. Van Cleve, P. Lee, T. Hayward, J. Houck, and J. Miles, Thermal infrared observations of Mars during the 1993 opposition, *Bull. Am. Astron. Soc., 25,* 1033, 1993.

Morris, R.V., and H.V. Lauer Jr., Matrix effects for reflectivity spectra of dispersed nanophase (superparamagnetic) hematite with application to martian spectral data, *J. Geophys. Res., 95,* 5101-5109, 1990.

Morris R.V., H.V. Lauer Jr., C.A. Lawson, E.K. Gibson Jr., G.A. Nace, and C. Stewart, Spectral and other physicochemical properties of submicron powders of hematite (α-Fe_2O_3), maghemite (γ-Fe_2O_3), magnetite (Fe_3O_4), goethite (α-FeOOH), and lepidocrocite (γ-FeOOH), *J. Geophys. Res., 90,* 3126-3144, 1985.

Morris, R.V., D.G. Agresti, H.V. Lauer Jr., J.A. Newcomb, T.D. Shelfer, and A.V. Murali, Evidence for pigmentary hematite on Mars based on optical, magnetic, and Mössbauer studies of superparamagnetic (nanocrystalline) hematite, *J. Geophys. Res., 94,* 2760-2778, 1989.

Morris, R.V., J.L. Gooding, H.V. Lauer, Jr., and R.B. Singer, Origins of Marslike spectral and magnetic properties of a Hawaiian palagonitic soil, *J. Geophys. Res., 95,* 14,427-14,434, 1990.

Morris, R.V., H.V. Lauer, Jr., D.G. Schulze, and R.G. Burns, Preparation and characterization of a nanophase hematite powder (abstract), *Lunar Planet. Sci. Conf., XXII,* 927-928, 1991.

Morris, R.V., D.C. Golden, J.F. Bell III, H.V. Lauer Jr., and J.B. Adams, Pigmenting agents in Martian soils: Inferences from spectral, Mössbauer, and magnetic properties of nanophase and other iron oxides in Hawaiian palagonitic soil PN-9, *Geochim. Cosmochim. Acta, 57,* 4597-4609, 1993.

Murchie, S., J. Mustard, J. Bishop, J. Head, C. Pieters, and S. Erard, Spatial variations in the spectral properties of bright regions on Mars, *Icarus, 105,* 454-468, 1993.

Mustard, J.F., and J.F. Bell III, New composite reflectance spectra of Mars from 0.4 to 3.14 μm, *Geophys. Res. Lett., 21,* 353-356, 1994.

Nash, D.B., J.W. Salisbury, J.E. Conel, P.G. Lucey, and P.R. Christensen, Evaluation of infrared emission spectroscopy for mapping the Moon's surface composition from lunar orbit, *J. Geophys. Res., 98,* 23,535-23,553, 1993.

Nyquist, R.A., and R.O. Kagel, *Infrared Spectra of Inorganic Compounds,* 495 pp., Academic, San Diego, CA, 1971.

Pollack, J.B., D. Colburn, R. Kahn, J. Hunter, W. Van Camp, and B. Baldwin, Properties of aerosols in the Martian atmosphere, as inferred from Viking Lander data, *J. Geophys. Res., 82,* 4479-4496, 1977.

Pollack, J.B., T.L. Roush, F. Witteborn, J. Bregman, D. Wooden, C. Stoker, O.B. Toon, D. Rank, B. Dalton, and R. Freedman, Thermal emission spectra of Mars (5.4-10.5 μm): Evidence for sulfates, carbonates, and hydrates, *J. Geophys. Res., 95,* 14,595-14,627, 1990.

Roush, T.L., Infrared optical properties of Mars soil analog materials: Palagonites, in Workshop on Chemical Weathering on Mars, *LPI Tech. Rep., 92-04,* Part 1, 32-33, 1992.

Roush, T.L., and J.F. Bell III, Thermal emission measurements 2000-400 cm^{-1} (5-25 μm) of Hawaiian palagonitic soils and their implications for Mars, *J. Geophys. Res.,* this issue.

Roush, T., F. Witteborn, P. Lucey, A. Graps, and J. Pollack, Thermal infrared observations of Mars (7.5-12.8 μm) during the 1990 opposition (abstract), *Lunar Planet. Sci. Conf., XXII,* 1137-1138, 1991.

Roush, T., F. Witteborn, J. Bregman, D. Rank, A. Graps, and J. Pollack, Thermal infrared spectra (5.5-9.2 μm) of Mars obtained from the Kuiper Airborne Observatory (abstract), *Lunar Planet. Sci. Conf., XXIII,* 1181-1182, 1992.

Roush T.L., D.L. Blaney, and R.B. Singer, The surface composition of Mars as inferred from spectroscopic observations, in *Remote Geochemical Analysis: Elemental and Mineralogical Composition,* edited by C. Pieters and P. Englert, pp. 367-393, Cambridge University Press, New York, 1993.

Ryskin, Y.I., The vibrations of protons in minerals: Hydroxyl, water, and ammonium, in *The Infrared Spectra of Minerals, Mineral. Soc. Monogr. 4,* (edited by V.C. Farmer) pp. 137-181, Mineralogy Society, London, 1974.

Salisbury, J.W., and A. Wald, The role of volume scattering in reducing spectral contrast of Reststrahlen bands in spectra of powdered minerals, *Icarus, 96,* 121-128, 1992.

Salisbury, J.W., L.S. Walter, N. Vergo, and D.M. D'Aria, *Infrared (2.1-25 μm) Spectra of Minerals,* 267 pp., Johns Hopkins University Press, Baltimore, Md., 1991.

Schwertmann, U., and R.M. Taylor, Iron oxides, in *Minerals in Soil Environments, Book Ser., no. 1,* (edited by J.B. Dixon and S.B. Weed) pp. 379-438, Soil Science Society of America, Madison, Wisc., 1989.

Sharanov, V.V., A lithological interpretation of the photometric and colorimetric studies of Mars. *Sov. Astron., 5,* Engl. Transl., 199-202, 1961.

Sherman D.M., R.G. Burns, and V.M. Burns, Spectral characteristics of the iron oxides with application to the Martian bright region mineralogy, *J. Geophys. Res., 87,* 10,169-10,180, 1982.

Singer, R.B., Spectral evidence for the mineralogy of high-albedo soils and dust on Mars, *J. Geophys. Res., 87*, 10,159-10,168, 1982.

Soderblom, L.A., The com sition and mineralogy of the Martian surface from spectroscopic observations: 0.3-50 μm, in *Mars*, edited by H. Kieffer et al., pp. 557-593, University of Arizona Press, Tucson, 1992.

Soderblom, L.A., K. Edwards, E.M. Eliason, E.M. Sanchez, and M.P. Charette, Global color variations on the Martian surface, *Icarus, 34*, 446-464, 1978.

Toon, O.B., J.B. Pollack, and C. Sagan, Physical properties of the particles composing the Martian dust storm of 1971-1972, *Icarus, 30*, 663-696, 1977.

Toulmin, P., III, A.K. Baird, B.C. Clark, K. Keil, H.J. Rose Jr., R.P. Christian, P.H. Evans, and W.C. Kelliher, Geochemical and mineralogical interpretation of the Viking inorganic chemical results, *J. Geophys. Res., 82*, 4625-4634, 1977.

J.F. Bell III and T.L. Roush, NASA Ames Research Center, Space Science Division, M/S 245-3, Moffett Field, CA 94035-1000. (fax: 415-604-6779; email: jimbo@anarchy.arc.nasa.gov)

R.V. Morris, Code SN4, Planetary Science Branch, NASA Johnson Space Center, Houston, TX 77058. (email: morris@snmail.jsc.nasa.gov)

(Received March 1, 1994;
accepted May 22, 1994.)

Page intentionally left blank

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5309–5317, MARCH 25, 1995

Thermal emission measurements 2000-400 cm^{-1} (5-25 μm) of Hawaiian palagonitic soils and their implications for Mars

Ted L. Roush[1]

Department of Geosciences, San Francisco State University, San Francisco, California

James F. Bell III

National Research Council, NASA Ames Research Center, Moffett Field, California

Abstract. The thermal emission of two palagonitic soils, common visible and near infrared spectral analogs for bright soils on Mars, was measured over the wavelength range of 5 to 25 μm (2000 to 400 cm^{-1}) for several particle size separates. All spectra exhibit emissivity features due to vibrations associated with H_2O and SiO. The maximum variability of emissivity is $\approx$20% in the short wavelength region (5 to 6.5 μm, 2000 to 1500 cm^{-1}), and is more subdued, <4%, at longer wavelengths. The strengths of features present in infrared spectra of Mars cannot be solely provided by emissivity variations of palagonite; some other material or mechanism must provide additional absorption(s).

Introduction and Summary

Insight into the availability of liquid water, and the duration, mode, and extent of weathering occurring throughout Mars' history can be provided by understanding the mineralogy of the surface materials on Mars. Si and Fe are the two most abundant elements on Mars [*Toulmin et al.*, 1977; *Clark et al.*, 1977, 1982], but how these are combined into surface minerals has not been directly determined. The mineralogy is inferred from in situ analyses [*Toulmin et al.*, 1977; *Clark et al.*, 1977, 1982; *Hargraves et al.*, 1977, 1979; *Oyama et al.*, 1978; *Banin and Rishpon*, 1978, 1979; *Huguenin*, 1982; *Quinn and Orenberg*, 1993], and spectroscopic observations (reviewed in *Soderblom* [1992] and *Roush et al.* [1993], also see *Pollack et al.* [1977], *Adams et al.* [1986], *Erard et al.* [1991], *Bell* [1992], *Mustard et al.* [1993], *Murchie et al.* [1993], and *Mustard and Bell* [1994]).

Interpretations of spectroscopic observations of Mars reveal that the ferric mineralogy occurs in two distinct forms: (1) nanophase or truly amorphous Fe^{3+}-bearing materials [*Morris and Lauer*, 1990; *Morris et al.*, 1989] that spectrally resemble certain terrestrial palagonites [e.g., *Singer et al.*, 1979; *Evans and Adams*, 1980; *Singer*, 1982; *Morris et al.*, 1990; *Bell et al.*, 1993; *Roush et al.*, 1993]; and (2) well-crystalline ferric oxides [*Pollack et al.*, 1977; *Morris et al.*, 1985, 1989; *Bell et al.*, 1990; *Bell*, 1992]. The available data indicate that the poorly-crystalline "palagonite-like" phases are spectrally dominant [e.g., *Hunt et al.*, 1973; *Toon et al.*, 1977; *Soderblom*, 1992; *Roush et al.*, 1993) and that the well-crystalline ferric oxides cannot constitute an abundance of more than about 4-8 wt % [*Bell*, 1992]. Thus palagonites are commonly used as visible and near-infrared spectral analogs for bright soils on Mars. Hawaiian palagonitic soils contain abundant Si and Fe, and laboratory studies have revealed nanophase iron oxides as the coloring agent in certain samples [*Morris et al.*, 1990; *Bell et al.*, 1993; *Golden et al.*, 1993].

Only transmission [*Roush*, 1992; *Roush et al.*, 1993] and reflectance [*Bruckenthal*, 1987; *Roush et al.*, 1988; *Crisp and Bartholomew*, 1992; *Bishop et al.*, 1993] measurements of palagonites have been reported in the mid- to far-infrared. Remotely sensed thermal infrared data measure the energy emitted from planetary surfaces. For ideal conditions [see *Christensen and Harrison*, 1993], when all emitted and reflected energy is measured, emittance is related to reflectance via Kirchoff's law (emittance = 1 - reflectance). Such conditions are rarely met in nature and only occasionally in the laboratory.

On Mars the spectrally dominant phase resembles terrestrial palagonites hence, it is important to understand the mid- and far-infrared emissivity of such materials. We report the measured emissivity of a Mauna Kea palagonitic soil whose transmission spectra have been previously reported [*Roush*, 1992] and of a thermally altered volcanic tephra [*Bell et al.*, 1993]. Both samples represent analogs for formation mechanisms capable of producing highly altered secondary weathering products on Mars. Such information can be used for interpretation of past (e.g., Mariner 6, 7, and 9) and future spacecraft (e.g., Mars Global Surveyor, MGS), as well as continuing Earth-based [e.g., *Pollack et al.*, 1990; *Roush et al.*, 1991, 1992; *Moersch et al.*, 1993], thermal infrared observations of Mars.

The measured thermal emission of the two palagonites exhibits complex spectral behavior within the range of 2000 to 400 cm^{-1} (5 to 25 μm). Emissivity contrasts are subdued

[1]Also at NASA Ames Research Center, Moffett Field, California.

Paper number 94JE02448.
0148-0227/95/94JE-02448$05.00

($\leq 4\%$) in the 1250 to 400 cm^{-1} (8 to 25 μm) range for large particle sizes (>90 μm) and are slightly enhanced (approximately 6%) for finer particle sizes (<90 μm). The most prominent spectral contrast is observed for all samples and particle sizes in the 1250 to 2000 cm^{-1} (5 to 8 μm) region with contrasts of up to 20% for the finer particle sizes.

All emissivity spectra exhibit features due to vibrations associated with molecular H_2O. Features associated with the Si-O Christiansen frequency (CF) suggest that the samples studied here are chemically and mineralogically different. The observed behavior of the CF suggests some compositional segregation is associated with particle size separation. All spectra exhibit subtle emissivity features that may be associated with ferric oxides and/or ferric hydroxides. However, more definitive assignments must await emissivity measurements of fine-grained samples of these types of materials. Whether or not nanophase ferric oxides and/or ferric hydroxides are consistent with the spectral data also must await measurements of these materials.

Comparisons of the palagonite emissivities to infrared measurements of Mars indicate that H_2O features are less obvious than those of terrestrial materials, supporting previous suggestions of dehydrated Martian surface materials [see *Soderblom*, 1992; *Roush et al.*, 1993]. This suggests that mid-IR spectra can be used for readily identifying the presence of H_2O-bearing regions on Mars. The palagonites are capable of producing spectral features, associated with Si-O and Fe-O, at frequencies (ν) consistent with the spacecraft data but cannot produce the magnitude of the features seen in the observational data. This suggests that (1) if materials similar to these palagonites are a major surface component on Mars, then high-quality (signal-to-noise) thermal observations are required to assess the relatively weak emissivity contrasts expected; and (2) the observed spectral variations of Mars are likely due to contributions of other materials or mechanisms, such as atmospheric aerosol and gas absorption. The implication for Mars is that compositional variability between similar materials is detectable and could be mapped on a global scale.

Experiment

Sample Description and Preparation

Palagonitic soil. Palagonitic soil 91-1 is a sample of Pahala Ash [*Macdonald et al.*, 1983] and was collected near South Point on the island of Hawaii. Samples from this site have been previously studied using reflectance spectroscopy [*Singer*, 1982; *Bruckenthal*, 1987; *Crisp and Bartholomew*, 1992]. After the sample was dried at room temperature, the bulk sample passing through a 2-mm sieve was retained. Elemental abundances of the bulk sample, determined by X ray fluorescence analysis, are provided in Table 1 along with the hygroscopic moisture percentage and the specific surface area. Subsequentially, the sample was dry sieved into eight particle size separates using 0.02-, 0.045-, 0.09-, 0.15-, 0.25-, 0.5-, 1-, and 2-mm sieves. A histogram of the relative weight percentage of each particle size fraction is shown in Figure 1.

Volcanic tephra. Volcanic tephra PH-1 from the Puu Huluhulu (Mauna Kea) cinder cone on the island of Hawaii has been previously studied in the visible and near-infrared and is a sample thermally altered by local magmatic activity [*Bell et al.*, 1993]. The sample was wet sieved using freon into seven size fractions using 0.02-, 0.045-, 0.09-, 0.15-, 0.25-, 0.5-, and 1-mm sieves. The relative abundances of each particle size fraction are shown in Figure 1. Fe and water concentrations were determined by instrumental neutron activation analysis [INAA] and thermogravimetric analysis [TGA] and are listed in Table 1. Additional analytical techniques have been used to characterize the mineralogy of this sample [*Bell et al.*, 1993].

Table 1. Elemental Abundances and Other Physical Parameters of Samples

Element	Weight Percentage			
	91-1*	VOL07†	Pahala Ash‡	PH-1§
SiO_2	29.9	38.10	38.40	n.a.
Al_2O_3	9.63	27.30	12.00	n.a.
CaO	4.38	3.16	2.08	5.49
MgO	5.19	2.38	3.54	n.a.
Na_2O	0.63	0.80	4.19	n.a.
Fe_2O_3	10.4	12.34	11.2	n.a.
FeO	n.a.	1.60	0.3	10.88
MnO	0.15	0.22	0.12	n.a.
TiO_2	1.89	3.07	2.33	n.a.
P_2O_5	0.18	1.60	0.12	n.a.
Lost on ignition	28.0	22.46	24.4	n.a.
Total	96.67	101.09	99.8	n.a.
Hygroscopic moisture,‖ (% by wt. at 105°C)	13.38	n.a.	n.a.	5.2
Specific surface area,‖ m^2/g	189.0	n.a.	n.a.	n.a.

Here, n.a. not available.
*determined from the <2 mm sample.
†from *Singer* [1982], determined from the ≤ 38 μm sample.
‡from *Bruckenthal* [1987], determined from the ≤ 38 μm sample.
§determined by INAA and TGA.
‖(A. Banin, personal communication, 1993).

Thermal Emission Measurements and Calculations

Emissivity spectra were acquired at the Arizona State University Thermal Emission Laboratory [*Anderson et al.*, 1991]. Each particle size separate was poured into a copper sample cup and the powder was leveled with a teflon-coated laboratory spatula. The cup was placed in an oven and

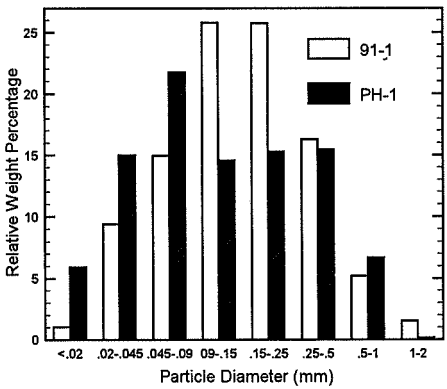


Figure 1. Histogram of relative mass fractions of the particle sizes for samples 91-1 and PH-1.

heated at 80°C overnight, a procedure that likely removes loosely bound water but at these low temperatures tightly bound and/or chemically adsorbed water should be retained [e.g., *Bruckenthal*, 1987]. The samples were transferred to the emission chamber, and particle size separates of sample 91-1 and PH-1 were heated at 70.0±0.2°C and 80.0±0.2-°C, respectively, during data acquisition. Spectra were also acquired of a blackbody source at liquid nitrogen temperature and approximately 80°C. From these three measurements the sample emissivity can be derived [*Christensen and Harrison*, 1993]. Each spectrum consists of 200 scans, obtained at a constant resolution of 4 cm^{-1}, over the entire spectral domain using a Fourier transform spectrometer. The rms peak-to-peak variability in the derived emissivity is ≈4-6% in the 2000 -1400 cm^{-1} region and improves to ≈0.4-0.6% at <1400 cm^{-1}, as illustrated in Figure 2.

To more readily identify emissivity features, we smoothed the original 4 cm^{-1} data using a Gaussian filter deconvolution program with a distance weighting scheme and a 3-σ despiking algorithm. Figure 2 shows the results for various Gaussian half-widths. We have used a five-channel half-width at half maximum convolution on all subsequent data presented here because this compromise retains the overall broad features while eliminating high-frequency peak-to-peak variations. Figure 3 illustrates the emissivity spectra of all particle size separates of sample 91-1 and all spectra exhibit maxima at 1980, 1870, 1650, 1520, ≈1200, and 675 cm^{-1}. Figure 4 illustrates the emissivity spectra of all particle size separates of sample PH-1 and in this case maxima occur near 1950, 1850, 1640, 1550, 1520, between 1250 and 1300, and 675 cm^{-1}.

As seen in Figures 3 and 4, the maximum variability of emissivity, ≈10-20%, is observed at the highest ν, 1500 to 2000 cm^{-1}, and is most pronounced for the finest grain sizes. At lower ν the emissivity variation is subdued, typically <4%, but again, more pronounced in the finest grain sizes. The subdued emissivities observed at low ν is consistent with model calculations of the palagonite emissivity [*Pollack et al.*, 1990, Figure 12a] and the emissivity of Pahala Ash calculated from reflectance [*Crisp and Bartholomew*, 1992, Figures 4 and 5].

Discussion

Spectral Features

Emissivity maxima near 675 cm^{-1} of both samples (see tick marks in Figures 5 and 6) are most likely associated with the Si-O bending fundamentals of silicate minerals and rocks [*Farmer*, 1974; *Salisbury et al.*, 1991; *Nash et al.*, 1993], although some ferric oxides and oxyhydroxides exhibit features near this ν [*Bell et al.*, this issue]. We have compared our emissivity spectra to the transmission spectra of crystalline iron oxides presented by *Bell et al.* [this issue] and identify only poorly defined corresponding features in the emissivity spectra near 460 and 550 cm^{-1} (see tick marks in Figures 5 and 6). As indicated above, nanophase iron oxides appear to be the coloring agents in certain palagonites [*Morris et al.*, 1990; *Bell et al.*, 1993; *Golden et al.*, 1993]. *Bell et al.* [this issue] show the transmission spectra of several nanophase iron oxides. The spectra of these materials exhibit absorption minima at ν similar to those of the crystalline iron oxides and oxyhydroxides. However, the

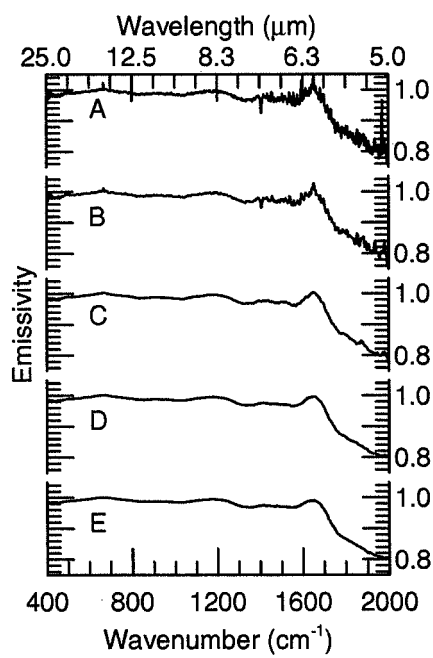


Figure 2. Smoothing of original data (curve A) using 1 (curve B), 5 (curve C), 10 (curve D), and 15 (curve E) channel half-width at full maxima Gaussians. Curve C retains the overall broad spectral features while eliminating the high frequency peak-to-peak variations.

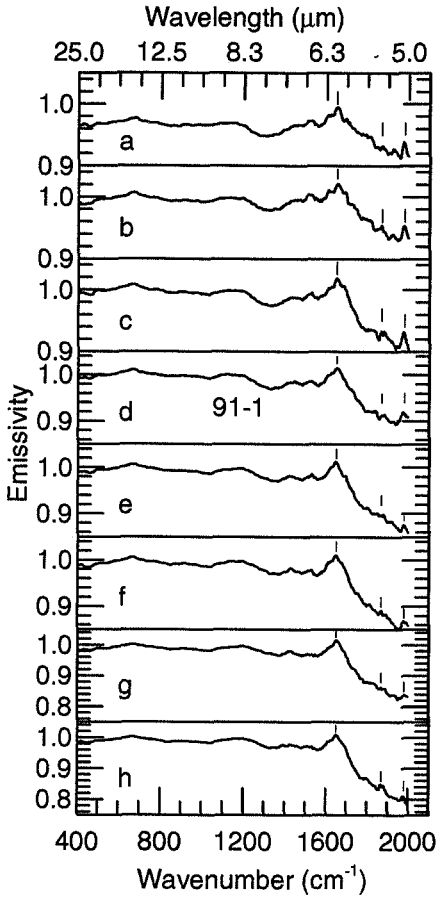


Figure 3. Measured emissivity, in the 400 to 2000 cm^{-1} spectral domain, of sample 91-1 for grain sizes of (a) 1-2 mm, (b) 0.5-1 mm, (c) 0.25-0.5 mm, (d) 0.15-0.25 mm, (e) 90-100 μm, (f) 45-90 μm, (g) 20-45 μm, and (h) <20 μm. Each large tick on the vertical axis is 0.1 and each small tick is 0.02 units in emissivity. Vertical tick marks indicate spectral features discussed in the text.

features are more subdued for the nanophase iron oxides when compared to the crystalline materials. Thus, by analogy to the crystalline materials, we would suspect that the nanophase iron oxides would produce more subdued features when measured in emission. This conclusion would be consistent with nanophase oxides being the coloring agent in the palagonites studied here.

Emissivity maxima near 1200 cm^{-1} (91-1) and between 1250 and 1300 cm^{-1} (PH-1) (see pointers on Figures 5 and 6) most likely correspond to the CF commonly associated with the silicate minerals and rocks. The CF position has been related to the polymerization and geometry of Si-tetrahedra and is observed at $>\nu$ for highly polymerized silicates than for poorly polymerized silicates [*Conel*, 1969; *Logan et al.* 1973; *Walter and Salisbury*, 1989; *Lucey*, 1991, *Nash et al.*, 1993]. Based on these previous studies, the apparent difference of CF seen in the two samples here suggests that the Si-bearing component in 91-1 is less polymerized than in PH-1. The CF of both samples appear to shift to $>\nu$ with decreasing particle size (see Figures 5 and 6). *Logan et al.* [1973] reported a CF shift to $<\nu$ with decreasing particle size for individual minerals, yet no CF shift is seen in the spectra of *Salisbury and Wald* [1992].

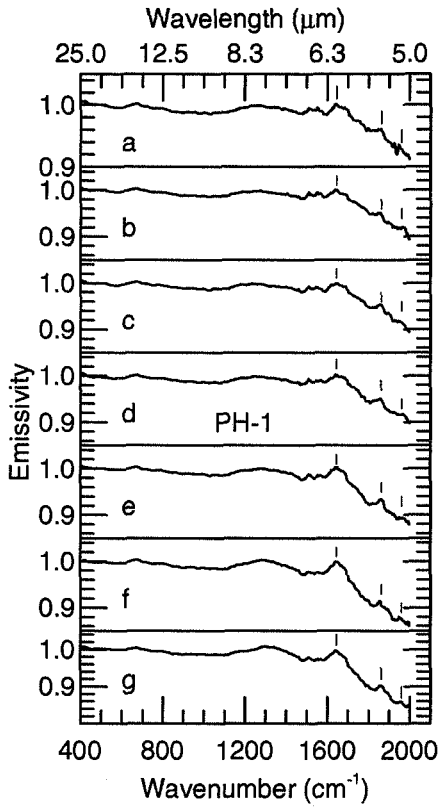


Figure 4. Measured emissivity, in the 400 to 2000 cm^{-1} spectral domain, of sample PH-1 for grain sizes of (a) 0.5-1 mm, (b) 0.25-0.5 mm, (c) 0.15-0.25 mm, (d) 90-100 μm, (e) 45-90 μm, (f) 20-45 μm, and (g) <20 μm. Each large tick on the vertical axis is 0.1, and each small tick is 0.02 units in emissivity. Vertical tick marks indicate spectral features discussed in the text.

The fine fraction of PH-1 is depleted in poorly polymerized silicates [*Bell et al.*, 1993]. Thus the CF shift to $>\nu$ with decreasing particle size is consistent with compositional rather than particle size changes. Although more complete analysis of 91-1 is required, we speculate that compositional rather than grain size effects are responsible for the observed CF shift.

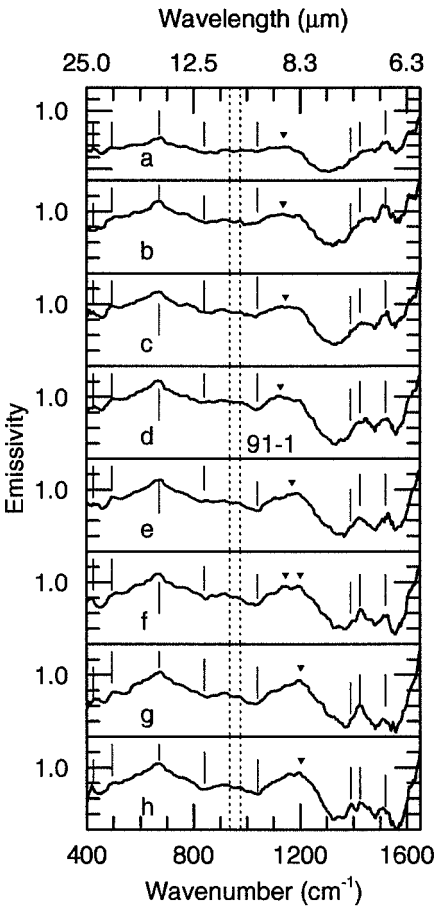


Figure 5. Enlargement of the emissivity, in the 400 to 1650 cm^{-1} spectral domain, of sample 91-1 for grain sizes of (a) 1-2 mm, (b) 0.5-1 mm, (c) 0.25-0.5 mm, (d) 0.15-0.25 mm, (e) 90-100 μm, (f) 45-90 μm, (g) 20-45 μm, and (h) <20 μm. Each large tick on the vertical axis is 0.05, and each small tick is 0.01 units in emissivity. Vertical tick marks indicate spectral features discussed in the text and inverted triangles indicate the position of the local maxima associated with the CF as discussed in the text.

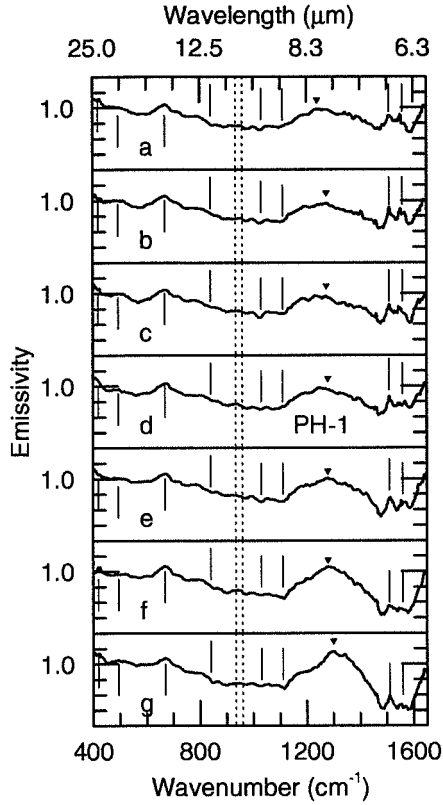


Figure 6. Enlargement of the emissivity in the 400 to 1650 cm^{-1} spectral domain, of sample PH-1 for grain sizes of (a) 0.5-1 mm, (b) 0.25-0.5 mm, (c) 0.15-0.25 mm, (d) 90-100 μm, (e) 45-90 μm, (f) 20-45 μm, and (g) <20 μm. Each tick on the vertical axis is 0.01 units in emissivity. Vertical tick marks indicate spectral features discussed in the text, and inverted triangles indicate the position of the local maxima associated with the CF as discussed in the text.

The spectral features in the 1380 to 1560 cm^{-1} region (see tick marks in Figures 5 and 6) occur where carbonates, nitrates, and hydrocarbons all have fundamental modes. More detailed chemical and mineralogical analyses of these samples are required before a specific assignment can be suggested.

The 1650 (91-1) and 1640 (PH-1) cm^{-1} emissivity maxima (see tick marks in Figures 3 and 4) occur at ν typical of the H-O-H bending fundamental, and Table 1 indicates water is associated with these samples. The presence of emissivity maxima in all spectra suggests that water is present in all grain sizes and may be enhanced in the finest grain sizes, as

indicated by an emissivity decrease $\nu \geq 1650$ cm^{-1}. This is consistent with the reported inverse correlation of water content and grain size for PH-1 [*Bell et al.*, 1993]. Although Mars is a desiccated environment compared to the Earth, hydrous materials have been suggested as surface and atmospheric aerosol constituents based on observed spectral features (see *Soderblom* [1992] and *Roush et al.* [1993] for more details, and also *Erard et al.* 1991; and *Murchie et al.*, 1993).

We provide no specific assignment for the emissivity maxima located in the 1950-1980 and 1850-1870 cm^{-1} regions (see tick marks in Figures 3 and 4). Reflectance spectra of many silicates exhibit features at similar ν to those observed here [*Salisbury et al.*, 1991; *Nash et al.*, 1993]. *Salisbury and Walter* [1989] suggested such features represent overtone and/or combination modes of the lower ν fundamental modes. Figures 5 and 6 are an expanded view of the emissivity in the 400-1650 cm^{-1} region for 91-1 and PH-1, respectively, and on each figure are dotted lines at the predicted fundamental ν. There appear to be very weak emissivity maxima present in the 91-1 spectra at the predicted ν but the correspondence for the PH-1 spectra is absent.

Comparison with Martian Data

There have been a variety of spectral observations of Mars in the infrared from spacecraft, Earth-based telescopes, and airborne telescopes (see *Soderblom* [1992] and *Roush et al.* [1993] for recent reviews). All of these observations have measured the total flux from Mars, including (1) flux emitted from the surface; and (2) absorption and scattering due to Martian atmospheric aerosols and gases. Quantitative comparisons of those observations to the measured emissivities are beyond the scope of this paper. Here we provide a qualitative comparison to the observational data. As a result, we address whether the measured emissivity of the palagonites studied here produce spectral features at ν consistent with similar features seen in the observational data, and whether or not surface emissivity alone can produce the intensity of features seen in the observational data.

Spectra of the Martian surface (5260-700 cm^{-1}, 1.9-14.1 μm) were acquired during the 1969 fly-by of Mars by Mariner Mars 7 (MM7) [*Herr et al.*, 1972] (also T. Z. Martin, Mariner 6/7 infrared spectrometer: Data set restoration, submitted to *Journal of Geophysical Research*, 1991) for a small region of the planet with the best spatial resolution sampling $\approx$825 km^2 areas. Thousands of Martian spectra (2000-2000 cm^{-1}, 5-50 μm) were acquired during the 1971-1972 Mariner Mars 9 (MM9) orbital mission [*Hanel et al.*, 1972], providing global coverage, but the best spatial resolution came from areas sampling $\approx 10^4$ km^2. In 1988 infrared observations of Mars were obtained from the Kuiper Airborne Observatory (KAO) (1850-950 cm^{-1}, 5.4-10.5 μm) [*Roush et al.*, 1989; *Pollack et al.*, 1990] for limited regions of the surface and sampled areas of $\approx 4 \times 10^6$ km^2. Figures 7, 8, and 9 compare the scaled emissivity of the coarsest and finest palagonite samples to a spectrum of a Martian bright region obtained by MM7, MM9, and KAO, respectively.

The MM7, MM9, and KAO spectra are all flat in the 1600 to 2000 cm^{-1} region, unlike either palagonite sample. This suggests compositional differences between the palagonite samples and the materials comprising the surface and/or atmospheric dust of Mars. Because we believe that the $\approx$1640 cm^{-1} band indicates water associated with the palagonites, this difference may be related to the relative hydration state of the Martian surface materials. If this is the case, then the weaker features in the observational data suggest that the Martian surface materials are less hydrous than the palagonites studied here, and is consistent with previous interpretations, based on near- and mid-infrared spectra [see *Soderblom* 1992; *Roush et al.*, 1993].

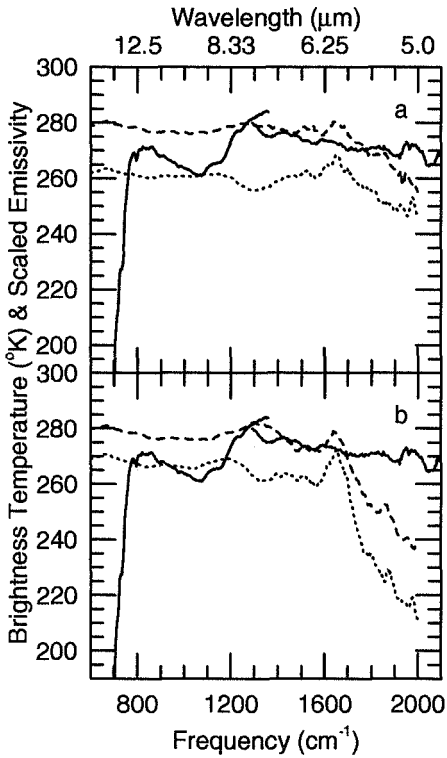


Figure 7. Comparison of Mariner 7 spectrum of Mars, 12°.1 latitude, 3.6° longitude (solid line), to 91-1 (dotted line), scaled to 270°K, and PH-1 (dashed line) scaled to 280°K, for both (a) coarse and (b) fine grain size samples.

The observational spectra all exhibit a minimum centered near 1050-1075 cm^{-1} with relative depths ranging from $\approx$1% to 10%. Additionally, the MM9 spectra have a minimum near 470 cm^{-1}, with relative depths varying from $\approx$1% to 6%. Based on analysis of the MM9 data by *Toon et al.* [1977] both of these band depths are variable and a function of emission angle. The three spectra shown by *Toon et al.* [1977, Figure 9] illustrate that the band depths are correlated with the atmospheric dust loading. As shown in Figure 10, Martian atmospheric gases alone exhibits

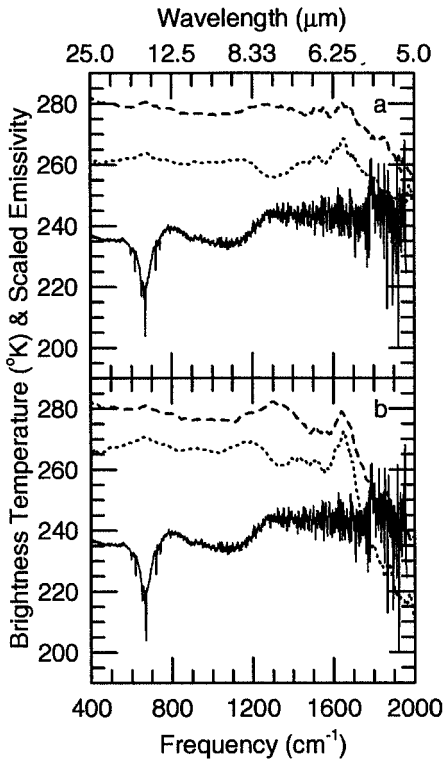


Figure 8. Comparison of Mariner 9 infrared interferometer spectrometer (IRIS) spectrum of a bright region on Mars, -39.8° latitude, 306.4° longitude (solid line), obtained during heavy atmospheric dust loading, to 91-1 (dotted line), scaled to 270°K, and PH-1 (dashed line), scaled to 280°K, for both (a) coarse and (b) fine grain size samples.

≈10% to 15% absorptions in the 1040-1075 cm^{-1} region and ≈5% absorptions near 450 cm^{-1}. However, these features will be weakened by aerosol scattering and are too narrow to account for the full breadth of the absorptions seen in the MM7, MM9, and KAO data. Emissivity spectra of PH-1 (all grain sizes) and 91-1 (grain sizes <500 μm) both exhibit minima near 1050 and 460 cm^{-1}. As mentioned above, the 1050 and 460 cm^{-1} features occur in spectral regions commonly attributed to Si-O stretching and bending modes and perhaps Fe-O stretching modes for the latter feature. This suggests that, if one ascribes these features in the observations solely to surface emissivity, then there may be some compositional similarities between the Si- and/or Fe-bearing materials contained in the palagonites studied here and those materials on the surface of Mars. However, the maximum relative depth seen in the palagonite emissivity spectra do not appear capable of engendering the strength of the feature seen in the observational spectra. Some other material(s) (aerosols or gases) or mechanism(s) contribute to these features. For example, *Pollack et al.* [1990] reproduced the strength of the 1075 cm^{-1} feature in the KAO data using models that included contributions due to atmospheric dust and CO_2 opacity.

Significance to Future Remote Sensing Observations of Mars

The MGS is intended to carry a duplicate of the ill-fated Mars observer thermal emission spectrometer, which was designed to return data in the 1600-200 cm^{-1} spectral domain and provide global coverage of Mars at a spatial resolution of 3x3 km [*Christensen et al.*, 1992]; representing an increase of ≈2 orders of magnitude in spatial resolution

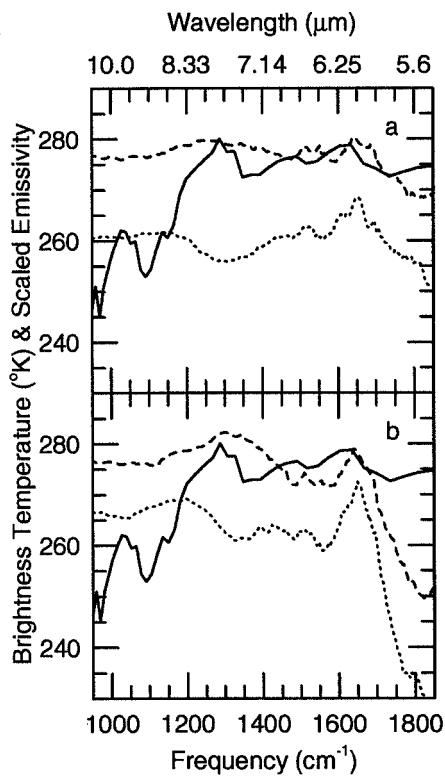


Figure 9. Comparison of Kuiper Airborne Observatory spectrum of Mars (solid line), -24° latitude, 132° longitude, to 91-1 (dotted line), scaled to 270°K, and PH-1 (dashed line), scaled to 280°K, for both (a) coarse and (b) fine grain size samples.

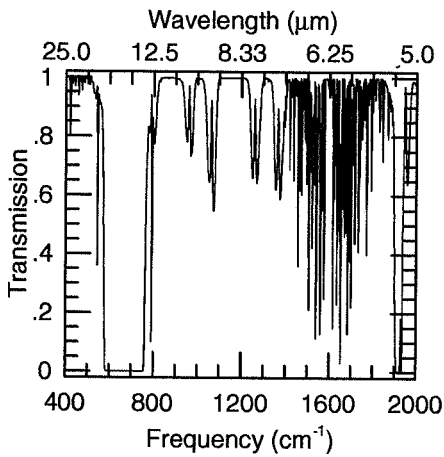


Figure 10. One-way transmission of the Martian atmosphere for nominal surface pressure (6.1 mbar) that includes absorptions due to CO_2, CO, H_2O, O_2, and O_3 gases but does not include any atmospheric dust absorption [after *Crisp*, 1990].

over existing data. Analyses of recent spectral imaging of Mars [*Bell et al.*, 1990; *Bibring et al.*, 1990; *Erard et al.*, 1991; *Bell*, 1992; *Mustard et al.*, 1993; *Murchie et al.*, 1993; *Mustard and Bell*, 1994] illustrate how increasing spatial resolution can reveal compositional differences in areas once believed homogeneous. If MGS can achieve the same nominal spatial resolution, then our knowledge regarding the surface composition of Mars will surely improve.

Acknowledgments. This research was supported by the Mars Observer Participating Scientist Program and NASA's Planetary Geology and Geophysics Program via RTOP 151-01-60-01. We extend our gratitude to Phil Christensen, Paul Barbera, Steve Ruff, and the rest of the staff and students at the Arizona State University Thermal Emission Spectrometer Facility for their assistance during acquisition of the emissivity data. We thank Mary Jane Bartholomew for assistance with the spectral smoothing program and Richard V. Morris and D.C. Golden for the compositional information of PH-1. We thank Wendy Calvin for her review that helped improve the manuscript.

References

Adams, J.B., M.O. Smith, and P.E. Johnson, Spectral mixture modeling - A new analysis of rock soil types at the Viking Lander site, *J. Geophys. Res., 91*, 8098-8112, 1986.

Anderson, D.L., J. Carpenter, P.R. Christensen, and P.W. Barbera, Development of a sample chamber for thermal emission spectroscopy (abstract), *Lunar Planet. Sci., XXII*, 21-22, 1991.

Banin, A., and J. Rishpon, Experimental simulation of the Viking Labeled Release (LR) results with iron-adsorbed smectite clay minerals, *Life Sci. Space Res., 17*, 59-64, 1978.

Banin, A., and J. Rishpon, Smectite clays in Mars soil: Evidence for their presence and role in Viking biology experimental results, *J. Mol. Evol., 14*, 133-152, 1979.

Bell, J.F., III, Charge-coupled device imaging spectroscopy of Mars, *Icarus, 100*, 575-597, 1992.

Bell, J.F., III, T.B. McCord, and P.D. Owensby, Observational evidence of crystalline iron oxides on Mars, *J. Geophys. Res., 95*, 14,447-14,461, 1990.

Bell, J.F., III, R.V. Morris, and J.B. Adams, Thermally altered palagonitic tephra: A spectral and process analog to the soils and dust of Mars, *J. Geophys. Res., 98*, 3373-3385, 1993.

Bell, J.F., III, T.L. Roush, and R.V. Morris. Mid-infrared transmission spectra of crystalline and nanophase iron oxides/oxyhydroxides: Implications for Mars, *J. Geophys. Res.*, this issue.

Bibring, J.-P., *et al.*, ISM observations of Mars and Phobos: First results, *Proc. Lunar Planet. Sci. Conf., 20th*, 461-471, 1990.

Bishop, J., S. Murchie, S. Pratt, J. Mustard, and C.Pieters, The importance of environmental conditions in reflectance spectroscopy of laboratory analogs for Mars surface materials, *LPI Tech. Rep., 93-06, Part 1*, 4-6, 1993.

Bruckenthal, E.A. The dehydration of phyllosilicates and palagonites: Reflectance spectroscopy and differential scanning, calorimetry, M.S. thesis, 182 pp., Univ. of Hawaii, 1987.

Christensen, P.R., and S.T. Harrison, Thermal emission spectroscopy of natural surfaces: Application to desert varnish coatings on rocks, *J. Geophys. Res., 98*, 19,819-19,834, 1993.

Christensen, P.R., D.L. Anderson, S.C. Chase, R.N. Clark, H.H. Kieffer, M.C. Malin, J.C. Pearl, J. Carpenter, N. Bandiera, F.G. Brown, and S. Silverman, Thermal emission spectrometer experiment: Mars Observer mission, *J. Geophys. Res., 97*, 7719-7734, 1992.

Clark, B.C., A.K. Baird, H.J. Rose Jr., P. Toulmin III, P.R. Christensen, W.C. Kelliher, A.J. Castro, C.D. Rowe, K. Kiel, and G. Huss, The Viking X ray fluorescence experiment: Analytical methods and early results, *J. Geophys. Res., 82*, 4577-4594, 1977.

Clark, B.C., A.K. Baird, R.J. Weldon, D.M. Tsusaki, L. Schnabel, and M.P. Candelaria, Chemical composition of Martian fines, *J. Geophys. Res., 87*, 10,059-10,067, 1982.

Conel, J.E., Infrared emissivities of silicates: Experimental results and a cloudy atmosphere model of spectral emission from condensed particulate mediums, *J. Geophys. Res., 74*, 1614-1634, 1969.

Crisp, D., Infrared radiative transfer in the dust-free Martian atmosphere, *J. Geophys. Res., 95*, 14,577-14,588, 1990.

Crisp, J. and M.J. Bartholomew, Mid-infrared spectroscopy of Pahala Ash palagonite and implications for remote sensing studies of Mars, *J. Geophys. Res., 97*, 14,691-14,699, 1992.

Erard, S., J.-P. Bibring, J. Mustard, O. Forni, J.W. Head, S. Hurtrez, Y. Langevin, C.M. Pieters, J. Rosenqvist, and C. Sotin, Spatial variations in composition of Valles Marineris and Isidis Planatia Regions of Mars derived from ISM data, *Proc. Lunar Planet. Sci. Conf., 21st*, 437-455, 1991.

Evans, D.L., and J.B. Adams, Amorphous gels as possible analogs to Martian weathering products, *Proc. Lunar Planet. Sci. Conf., 11th*, 757-763, 1980.

Farmer, V.C., *The Infrared Spectra of Minerals, Monogr. 4*, 539 pp., Mineralogical Society, London, 1974.

Golden, D.C., R.V. Morris, D.W. Ming, H.V. Lauer Jr., and S.R. Yang, Mineralogy of three slightly palagonitized basaltic tephra samples from the summit of Mauna Kea Hawaii, *J. Geophys. Res., 98*, 3401-3411, 1993.

Hanel, R., *et al.*, Investigation of the Martian environment by infrared spectroscopy on Mariner 9, *Icarus, 17*, 423-442, 1972.

Hargraves, R.B., D.W. Collinson, R.E. Arvidson, and C.R. Spitzer, The Viking magnetic properties experiment: Primary mission results, *J. Geophys. Res., 82*, 4547-4558, 1977.

Hargraves, R.B., D.W. Collinson, R.E. Arvidson, and P.M. Cates, Viking magnetic properties experiment: Extended mission results, *J. Geophys. Res., 84*, 8379-8384, 1979.

Herr, K.C., P.B. Forney, and G.C. Pimental, Mariner Mars 1969 infrared spectrometer, *Appl. Opt., 3*, 493-501, 1972.

Huguenin, R.L., Chemical weathering and the Viking biology experiments on Mars, *J. Geophys. Res., 87,* 10,069-10,082, 1982.

Hunt, G.R., L.M. Logan, and J.W. Salisbury, Mars: Components of infrared spectra and the composition of the dust cloud, *Icarus, 18,* 459-469, 1973.

Logan, L.M., G. Hunt, J. Salisbury, and S. Balsamo, Compositional implications of Christiansen frequency maximums for remote sensing applications, *J. Geophys. Res., 78,* 4983-5003, 1973.

Lucey, P.G., Comparison of thermal emission spectroscopic measurements of the lunar surface: 1968-1990, *Proc. Lunar Planet. Sci. Conf., 21st,* 417-423, 1991.

Macdonald, G.A., A.T. Abbot, and F.L. Peterson, *Volcanoes in the Sea,* 517 pp., University of Hawaii Press, Honolulu, 1983.

Moersch, J., P. Nicholson, S. Squyres, J. Van Cleve, P. Lee, T. Hayward, J. Houck, and J. Miles, Thermal infrared observations of Mars during the 1993 opposition (abstract), *Bull. Am. Astron. Soc., 25,* 1033, 1993.

Morris, R.V., and H.V. Lauer Jr., Matrix effects for reflectivity spectra of dispersed nanophase (superparamagnetic) hematite with application to Martian spectral data, *J. Geophys. Res., 95,* 5101-5109, 1990.

Morris, R.V., H.V. Lauer Jr., C.A. Lawson, E.K. Gibson Jr., G.A. Nace, and C. Stewart, Spectral and other physicochemical properties of submicron powders of hematite (α-Fe_2O_3), maghemite (γ-Fe_2O_3), magnetite (Fe_3O_4), goethite (α-FeOOH), and lepidocrocite (γ-FeOOH), *J. Geophys. Res., 90,* 3126-3144, 1985.

Morris, R.V., D.G. Agresti, H.V. Lauer Jr. J.A. Newcomb, T.D. Shelfer, and A.V. Murali, Evidence for pigmentary hematite on Mars based on optical, magnetic and Mössbauer studies of superparamagnetic (nanocrystalline) hematite, *J. Geophys. Res., 94,* 2760-2778, 1989.

Morris, R.V., J.L. Gooding, H.V. Lauer Jr., and R.B. Singer, Origin of Mars-like spectral and magnetic properties of a Hawaiian palagonitic soil, *J. Geophys. Res., 95,* 14,427-14,434, 1990.

Murchie, S., J. Mustard, J. Bishop, J. Head, C. Pieters, and S. Erard, Spatial variations in the spectral properties of bright regions on Mars, *Icarus, 105,* 454-468, 1993.

Mustard, J. F., and J. F. Bell III, New composite reflectance spectra of Mars from 0.4 to 3.14 μm, *Geophys. Res. Lett., 21,* 353-356, 1994.

Mustard, J.F., S. Erard, J.-P. Bibring, J. Head, S. Hurtez, Y. Langevine, C. Pieters, and C. Sotin, The surface of Syrtis Major: Composition of the volcanic substrate and mixing with altered dust and soil, *J. Geophys. Res., 98,* 3387-3400, 1993.

Nash, D.B., J.W. Salisbury, J.E. Conel, P.G. Lucey, and P.R. Christensen, Evaluation of infrared emission spectroscopy for mapping the Moon's surface composition from lunar orbit, *J. Geophys. Res., 98,* 23,535-23,552, 1993.

Oyama, V.I., B.J. Berdahl, F. Woeller, and M. Lehwalt, The chemical activities of the Viking biology experiments and the arguments for the presence of superoxides, peroxides, γ-Fe_2O_3 and carbon suboxide polymer in the Martian soil, *Life Sci. Space Res., 16,* 3-8, 1978.

Pollack, J.B., D. Colburn, R. Kahn, J. Hunter, W. Van Camp, C.E. Carlston, and M.R. Wolf, Properties of aerosols in the Martian atmosphere, as inferred from Viking lander data, *J. Geophys. Res., 82,* 4479-4496, 1977.

Pollack, J.B., T.L. Roush, F. Witteborn, J. Bregman, D. Wooden, C. Stoker, O.B. Toon, D. Rank, B. Dalton, and R. Freedman, Thermal emission spectra of Mars (5.4-10.5 μm): Evidence for sulfates, carbonates, and hydrates, *J. Geophys. Res., 95,* 14,595-14,627, 1990.

Quinn, R., and J. Orenberg, Simulations of the Viking gas exchange experiment using palagonite and Fe-rich montmorillonite as terrestrial analogs: Implications for the surface composition of Mars, *Geochim. Cosmochim. Acta, 57,* 4611-4618, 1993.

Roush, T.L., Infrared optical properties of Mars soil analog materials: Palagonites, *LPI Tech. Rep., 92-04, Part 1,* 32-33, 1992.

Roush, T.L. and J.F. Bell III, Transmission measurements (4000-400 cm^{-1}, 2.5-25μm) of ferric oxides and ferric oxyhydroxides: Implications for Mars (abstract), *Lunar Planet. Sci. Conf. XXV,* 1161-1162, 1994.

Roush, T.L., E.A. Roush, R.B. Singer, and P.G. Lucey, Preliminary analysis of recent 2.2-4.2 μm telescopic observations of Elysium, Mars: Implications for crystallinity and hydration state of surface materials, *LPI Tech Rep., 88-05,* 111-113, 1988.

Roush, T., J. Pollack, C. Stoker, F. Witteborn, J. Bregman, D. Wooden, and D. Rank, CO_3^{2-} and SO_4^{2-}-bearing anionic complexes detected in martian atmospheric dust (abstract), *Lunar Planet. Sci., XX,* 928-929, 1989.

Roush, T., F. Witteborn, P. Lucey, A. Graps, and J. Pollack, Thermal infrared observations of Mars (7.5-12.8 μm) during the 1990 opposition (abstract), *Lunar Planet. Sci., XXII,* 1137-1138, 1991.

Roush, T.L., F. Witteborn, J. Bregman, D. Rank, A. Graps, and J. Pollack, Thermal infrared spectra (5.5-9.2 μm) of Mars obtained from the Kuiper Airborne Observatory (abstract), *Lunar Planet. Sci. XXIII,* 1181-1182, 1992.

Roush, T.L., D.L. Blaney, and R.B. Singer, The surface composition of Mars as inferred from spectroscopic observations, in *Remote Geochemical Analysis: Elemental and Mineralogical Composition,* edited by C. Pieters and P. Englert, 367-394 pp., Cambridge University Press, New York, 1993.

Salisbury, J.W., and A. Wald, The role of volume scattering in reducing spectral contrast of Restrahlen bands in spectra of powdered minerals, *Icarus, 96,* 121-128, 1992.

Salisbury, J.W., and L.S. Walter, Thermal infrared (2.5-13.5 μm) spectroscopic remote sensing of igneous rock types on particulate planetary surfaces, *J. Geophys. Res., 94,* 9192-9202, 1989.

Salisbury, J.W., L.S. Walter, N. Vergo, and D. M. D'Aria, *Infrared (2.1-25 μm) Spectra of Minerals,* 267 pp., Johns Hopkins University Press, Baltimore, Md., 1991.

Singer, R.B., Spectral evidence for the mineralogy of high-albedo soils and dust on Mars, *J. Geophys. Res., 87,* 10,159-10,168, 1982.

Singer, R.B., T.B. McCord, R.N. Clark, J.B. Adams, and R.L. Huguenin, Mars surface composition from reflectance spectroscopy: A summary, *J. Geophys. Res., 84,* 8415-8426, 1979.

Soderblom, L.A., The composition and mineralogy of the Martian surface from spectroscopic observations: 0.3 μm to 50 μm, in *Mars,* edited by H.H. Kieffer, B.M. Jakosky, C.W. Snyder, and M.S. Matthews, 557 pp., University of Arizona Press, Tucson, 1992.

Toon, O.B., J.B. Pollack, and C. Sagan, Physical properties of the particles composing the Martian dust storm of 1971-1972, *Icarus, 30,* 663-696, 1977.

Toulmin, P., III, A.K. Baird, B.C. Clark, K. Keil, H.J. Rose Jr., P.R. Christensen, P.H. Evans, and W.C. Kelliher, Geochemical and mineralogical interpretation of the Viking inorganic chemical results, *J. Geophys. Res., 82,* 4625-4634, 1977.

Walter, L.S., and J.W. Salisbury, Spectral characterization of igneous rocks in the 8- to 12-μm region, *J. Geophys. Res., 94,* 9203-9213, 1989.

J.F. Bell and T.L. Roush, NASA Ames Research Center, MS 245-3, Moffett Field, CA 94035-1000. (e-mail: roush@barsoom.arc.nasa.gov)

(Received March 1, 1994; revised August 23, 1994; accepted September 19, 1994.)

Page intentionally left blank

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5319–5328, MARCH 25, 1995

Hematite, pyroxene, and phyllosilicates on Mars: Implications from oxidized impact melt rocks from Manicouagan Crater, Quebec, Canada

Richard V. Morris,[1] D. C. Golden,[2] James F. Bell III,[3] and H. V. Lauer Jr.[4]

Abstract. Visible and near-IR reflectivity, Mössbauer, and X ray diffraction data were obtained on powders of impact melt rock from the Manicouagan Impact Crater located in Quebec, Canada. The iron mineralogy is dominated by pyroxene for the least oxidized samples and by hematite for the most oxidized samples. Phyllosilicate (smectite) contents up to ~15 wt % were found in some heavily oxidized samples. Nanophase hematite and/or paramagnetic ferric iron is observed in all samples. No hydrous ferric oxides (e.g., goethite, lepidocrocite, and ferrihydrite) were detected, which implies the alteration occurred above 250°C. Oxidative alteration is thought to have occurred predominantly during late-stage crystallization and subsolidus cooling of the impact melt by invasion of oxidizing vapors and/or solutions while the impact melt rocks were still hot. The near-IR band minimum correlated with the extent of aleration (Fe^{3+}/Fe_{tot}) and ranged from ~1000 nm (high-Ca pyroxene) to ~850 nm (bulk, well-crystalline hematite) for least and most oxidized samples, respectively. Intermediate band positions (900-920 nm) are attributed to low-Ca pyroxene and/or a composite band from hematite-pyroxene assemblages. Manicouagan data are consistent with previous assignments of hematite and pyroxene to the ~850 and ~1000 nm bands observed in Martian reflectivity spectra. Manicouagan data also show that possible assignments for intermediate band positions (900-920 nm) in Martian spectra are pyroxene and/or hematite-pyroxene assemblages. By analogy with impact melt sheets and in agreement with observables for Mars, oxidative alteration of Martian impact melt sheets above 250°C and subsequent erosion could produce rocks and soils with variable proportions of hematite (both bulk and nanophase), pyroxene, and phyllosilicates as iron-bearing mineralogies. If this process is dominant, these phases on Mars were formed rapidly at relatively high temperatures on a sporadic basis throughout the history of the planet. The Manicouagan samples also show that this mineralogical diversity can be accomplished at constant chemical composition, which is also indicated for Mars from analyses of soil at the two Viking landing sites.

Introduction

Oxidative alteration of impact melt sheets is reported at many terrestrial impact structures, including Manicouagan [*Floran et al.*, 1978] and West Clearwater Lake [*Phinney et al.*, 1978], both in Quebec, Canada, and the Ries basin [*Pohl et al.*, 1977; *Newsom et al.*, 1986] in Germany. Oxidation apparently occurs shortly after the impact by an influx of oxidizing vapors and/or fluids while the rocks are still hot but below solidus temperatures [*Phinney et al.*, 1978]. Hematite is reported as an oxidation product at all three impact structures [*Floran et al.*, 1978; *Phinney et al.*, 1978; *Pohl et al.*, 1977]. Because meteoritic impact also occurs under oxidizing surface conditions on Mars, a number of studies [e.g., *Newsom*, 1980, 1986; *Kieffer and Simonds*, 1980; *Allen et al.*, 1982; *Clifford*, 1993] have advocated that a significant fraction of Martian soil may consist of erosional products of hydrothermally altered (oxidized) impact melt sheets. Because of the identification of hematite in visible and near-IR spectral data of Martian bright regions [e.g., *Morris et al.*, 1989; *Bell et al.*, 1990; *Morris and Lauer*, 1990; *Morris et al.*, 1993], this view is at least consistent with the above studies of terrestrial impact melts. The mineral assemblages in terrestrial impact melt rocks that have undergone oxidative alteration are thus putative mineral assemblages for the Martian surface, and their optical properties relevant to the mineralogic interpretation of Martian spectral data. We report here compositional, Mössbauer, and spectral data (visible and near-IR) for powders of impact melt rocks from Manicouagan Crater.

[1]NASA Johnson Space Center, Houston, Texas.
[2]Dual, Inc., Houston, Texas.
[3]National Research Council/NASA Ames Research Center, Space Sciences Division, Moffett Field, California.
[4]Lockheed ESC, Houston, Texas

Paper number 94JE01500.
0148-0227/95/94JE-01500$05.00

Sample and Methods

Samples. We studied 23 samples of powdered coherent rock from the Manicouagan impact melt sheet. All rock samples were powdered in an alumina mortar and passed through a 90 μm sieve. *Floran et al.* [1976, 1978] report chemical and petro-graphic data for 16 of these samples.

Analytical methods. A Cary-14 spectrophotometer configured with a 14-cm-diameter integrating sphere was used to obtain visible and near-IR reflectance spectra. Elscint and Ranger Mössbauer spectrometers using $^{57}Co(Rh)$ sources were used to obtain iron Mössbauer spectra (FeMS). Mössbauer spectra were fit to theoretical line shapes using an in-house

computer program (JSCFIT). Saturation magnetizations (J_s) were calculated from magnetization data obtained on a Princeton Applied Research model 151 vibrating sample magnetometer. Concentrations of major (Fe, Ca, Na, and K) and minor and trace (Sc, Cr, Rb, Sr, Ba, and rare earth elements) elements were obtained by instrumental neutron activation analysis (INAA). A Scintag XDS 2000 X ray diffractometer using CuKα radiation was employed to obtain powder X ray diffraction (XRD) patterns. Mg-saturated, oriented samples of clay-size fractions (<2 μm obtained by gravity sedimentation of an aqueous dispersion) with and without glycol solvation were used to determine if phyllosilicates (and smectite in particular) are present. More detailed information on analytical procedures is given by *Morris et al.* [1989, 1990], *Golden et al.* [1993], and *Bell et al.* [1993].

Results and Discussion

Mössbauer Mineralogy

Iron Mössbauer spectroscopy (FeMS), through the position of spectral lines, is sensitive to both the oxidation state and mineralogy of iron. The relative intensities of spectral lines provide quantitative information about the relative distribution of iron among its oxidation states and iron-bearing mineralogies. Because the technique is sensitive only to iron (specifically the isotope ^{57}Fe which has 2.2% natural abundance) and thus blind to iron-free phases (e.g., feldspar and quartz), the absolute modal abundance of iron-bearing phases cannot be determined. However, if the concentration of iron in the iron-bearing phases can be determined or estimated (e.g., from the mineralogical identification), the relative modal proportions of iron-bearing phases can be estimated.

FeMS spectra (293 K) for samples having the highest and lowest proportions of ferric iron are shown in Figure 1. A total of nine components, not all present in all samples, were needed to fit these spectra. The Mössbauer parameters for the nine components are compiled in Table 1. These parameters are values averaged over all spectra where they could be determined without constraining peak positions. The number of spectra used to calculate averages and the maximum relative area observed for each component are also given. The mineralogy associated with six of the nine components is obvious from inspection of the spectra and values of their Mössbauer parameters. The hematite sextet (well-crystalline α-Fe_2O_3, bulk-Hm) is a major component of all the heavily oxidized samples, and it dominates the spectrum of MAN-74-177. Pyroxene ($(Ca,Mg,Fe)SiO_3$; unresolved ferrous doublets from M1 and M2 sites) is the major component of the least oxidized samples. The two-sextet pattern characteristic of magnetite (Fe_3O_4; tetrahedral A and octahedral B sites) is apparent in the spectra for MAN-74-608A and 512A. A weak doublet from ferrous iron in ilmemite ($FeTiO_3$) is also evident between the pyroxene doublets in MAN-74-512A. These six mineralogical assignments are consistent with petrographic studies of these samples [*Floran et al.*, 1978].

The remaining three components are two ferric doublets ($Fe3D_1$ and $Fe3D_2$) and one ferrous doublet ($Fe2D_1$). $Fe3D_2$ and $Fe2D_1$ both have their highest relative contributions in the MAN-74-342A spectrum (Figure 1). The high-velocity line of the ferrous doublet is evident near 2.1 mm/s. MAN-74-342A is the only sample where this line is more intense than the high-velocity lines of the pyroxene doublets. $Fe3D_1$ is most

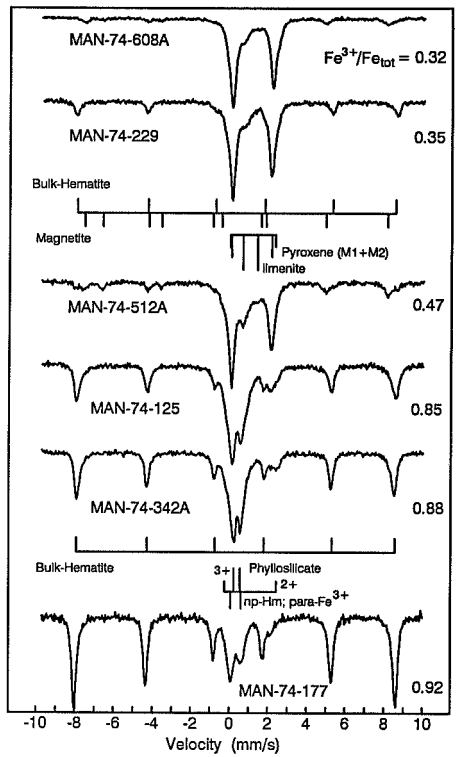


Figure 1. Mössbauer spectra for Manicouagan impact melt rocks having lowest and highest values of Fe^{3+}/Fe_{tot}.

apparent in the spectrum of MAN-74-177, and its quadrupole splitting is larger than that for $Fe3D_2$. The mineralogy associated with these three components is not readily assignable just from positions of spectral lines.

In order to constrain mineralogical assignments for $Fe3D_1$, $Fe3D_2$, and $Fe2D_1$, we calcined MAN-74-342A and 608A in air at 400, 500, 800, and 1000°C. For MAN-74-342A, $Fe2D_1$ and $Fe3D_2$ are not present by 400 and 800°C, respectively. For MAN-74-608A, magnetite and pyroxene are not present by 500 and 1000°C, respectively. The 1000°C spectra for both samples are shown in Figure 2. Both are characterized by a hematite sextet and an asymmetric ferric doublet ($Fe3D_1$). These spectra are very similar to that for MAN-74-177. The observations are that the $Fe3D_1$ component is stable at high temperatures and that the $Fe2D_1$ component is easily oxidized at a lower temperature than ferrous iron in pyroxene. The temperatures where $Fe2D_1$ and $Fe3D_2$ disappear from the Mössbauer spectra (by 800°C) are similar to the decomposition temperatures for smectites [e.g., *Borchardt*, 1977]. Because XRD data (see below) show that smectite is

Table 1. Average Mössbauer Parameters (293 K) for the Nine Components, Number of Samples Used for Averages, and Maximum Area Observed for Component

Component	IS[a], mm/s	QS, mm/s	B_{hf}, T	Number of Samples	Maximum Area, %[b]	Mineralogy
$Fe3S_1$	0.38(1)	-0.20(2)	51.4(2)	14	67	Hematite
$Fe3S_2$	0.29	0.01	49.0	2	8	Magnetite, A-site
$Fe23S_1$	0.70	0.00	45.7	2	9	Magnetite, B-site
$Fe3D_1$	0.38(1)	0.70(1)	-	2	40	np-Hm; para-Fe^{3+}
$Fe3D_2$	0.40	0.37	-	1	25	Phyllosilicate, Fe^{3+}
$Fe2D_1$	1.13	2.60	-	1	5	Phyllosilicate, Fe^{2+}
$Fe2D_2$	1.12(1)	2.06(2)	-	19	54	Pyroxene, M2-site
$Fe2D_3$	1.13	2.57	-	2	12	Pyroxene, M1-site
$Fe2D_4$	1.07	0.67	-	1	4	Ilmenite

[a]IS = isomer shift; QS = quadrupole splitting; B_{hf} = hyperfine field. Isomer shifts are reported relative to Fe metal at room temperature.

[b]Calculated using a value of 1.21 as the ferric to ferrous recoilless fraction ratio [*DeGrave and Van Alboom*, 1991].

present in our highly altered samples, we assign $Fe2D_1$ and $Fe3D_2$ to phyllosilicates (smectite).

Possible assignments for $Fe3D_1$ include pseudobrookite solid solutions from oxidation of ilmenite, nanophase hematite (np-Hm) from oxidation/decomposition of silicate phases including mafic minerals [*Straub et al.*, 1991] and phyllosilicates [*Weldon et al.*,1982; *Moskowitz and Hargraves*, 1982], and/or paramagnetic ferric iron (para-Fe^{3+}) in structural sites of silicate/aluminosilicate phases. The assymetry of the doublet suggests contributions from more than one source. Np-Hm is probably a major contributor to $Fe3D_1$ because a sharp band near 450 nm, which would be indicative of Fe^{3+} at well-defined structural sites, is not present. Np-Hm does not have well-defined bands [*Morris et al.*, 1989; *Morris and Lauer*, 1990]

Although the mineralogical assignments made above for $Fe3D_1$, $Fe3D_2$, and $Fe2D_1$ are not definitive, we can reasonably rule out other potential mineralogies. Lepidocrocite (γ-FeOOH) and superparamagnetic goethite (α-FeOOH) are not reasonable assignments for the ferric doublets because the phases have decomposition temperatures below 400°C [e.g., *Morris and Lauer*, 1981]. The sextet from crystalline goethite is also not present in any sample. The Mössbauer parameters for siderite ($FeCO_3$) are also not consistent with the data. The low sulfur content of these samples (<20 ug/g [*Floran et al.*, 1978]) implies that jarosites, sulfates, low-spin ferrous sulfides (e.g., FeS_2), and other minerals that contain iron and sulfur are not present in significant amounts.

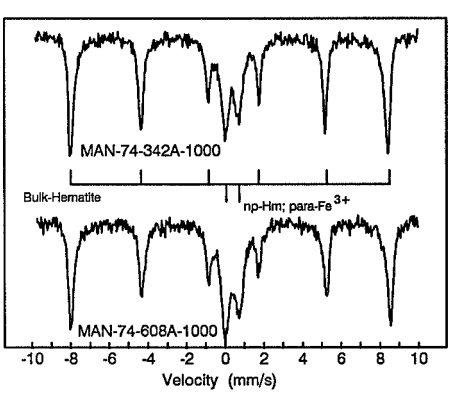


Figure 2. Mössbauer spectra of MAN-74-342A and MAN-74-608A after calcining at 1000°C.

Relative component areas for each sample are given in Table 2. These areas include a correction for different recoilless fractions for ferrous and ferric ions. We used a value of 1.21, which is the average value we calculated from the data of *De Grave and Van Alboom* [1991], as the ferric to ferrous recoilless fraction ratio. The fraction of Mössbauer spectral area resulting from ferric iron is highly variable and ranges from 0.32 to 0.92. To a first approximation, this range represents (for iron bearing mineralogies) a variation from an assemblage dominated by pyroxene to one dominated by bulk-Hm. Together, the two mineralogies account for an average of 62% of the total iron. In Figure 3, Fe^{3+}/Fe_{tot} determined from FeMS data is compared with the same ratio calculated from chemical data reported by *Floran et al.* [1978]. Also plotted is the prealteration value of Fe^{3+}/Fe_{tot} (0.21) estimated by *Floran et al.* [1978]. The agreement between chemical and Mössbauer data is very good, and all samples we studied have undergone oxidative alteration. The oxidative alteration probably took place at temperatures above 250°C, where hematite is stable relative to hydrous phases like goethite and ferrihydrite [e.g., *Morris and Lauer*, 1981]. Smectites generally decompose at 500-700°C [e.g., *Borchardt*, 1977], so that samples with a significant amount of smectite (e.g., MAN-74-342A) probably did not see temperatures higher than this during oxidative alteration.

Examination of the data in Table 2 shows that sample saturation magnetization (J_s) correlates with the relative area of magnetite (or titanomagnetite), so that its oxidation product is weakly magnetic hematite (or titanohematite) rather than strongly magnetic maghemites (or titanomaghemites). It may be that the relatively high temperatures for oxidative alteration at Manicouagan favored alteration of (titano)magnetites to hematites rather than to maghemites. As reported by *Allan et*

Table 2. Percentage of Total Fe Associated With Iron-Bearing Phases (Mössbauer Mineralogy), Ratio of Fe^{3+}/Fe_{tot}, Sample Saturation Magnetization (J_s), and Position of Near-IR Band Minimum for Manicouagan Impact Melt Rocks

Sample	Hematite (bulk Hm)	np-Hm; para-Fe^{3+}	Phyllo-silicate	Magnetite	Ilmenite	Pyroxene	J_s of Sample, Near-IR			
							FeMS	Chemistry	A m^2/kg	Band, nm
MAN-74-15	42	30	2	9	0	18	0.81	0.74	0.34	935
MAN-74-74C	47	27	16	0	0	11	0.89	-	0.32	845
MAN-74-83	29	28	0	7	0	36	0.63	-	0.14	930
MAN-74-95A	37	25	14	0	0	24	0.73	0.73	0.04	850
MAN-74-97	57	24	5	0	0	14	0.86	0.78	0.08	840
MAN-74-125	36	40	8	0	0	15	0.85	-	0.25	875
MAN-74-131	9	35	0	15	1	39	0.56	0.48	0.73	920
MAN-74-133	31	40	6	7	0	16	0.76	0.65	0.30	900
MAN-74-176	40	30	1	0	0	29	0.71	0.70	0.08	910
MAN-74-177	67	19	6	0	0	8	0.92	0.84	0.06	830
MAN-74-217	37	25	4	0	0	34	0.66	0.62	0.08	910
MAN-74-229	13	21	1	0	0	65	0.35	0.34	0.03	1000
MAN-74-243	15	32	0	6	0	48	0.51	0.46	0.33	980
MAN-74-342A	45	18	30	0	0	7	0.88	-	0.05	855
MAN-74-390	21	23	0	11	0	45	0.52	0.54	0.69	940
MAN-74-396	28	29	0	0	0	44	0.56	0.55	0.05	920
MAN-74-407	17	22	0	3	0	58	0.41	0.41	0.12	940
MAN-74-506	17	33	0	10	0	40	0.58	0.64	0.64	910
MAN-74-512A	5	28	0	17	4	45	0.47	-	1.30	935
MAN-74-608A	0	21	1	12	1	64	0.32	-	0.73	980
MAN-74-685E	17	37	0	11	0	35	0.62	-	0.64	920
MAN-75-3	28	33	0	10	0	28	0.69	0.73	0.51	930
MAN-75-20	34	28	0	0	0	38	0.62	0.59	0.16	910

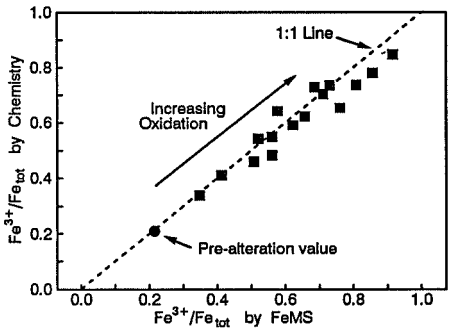


Figure 3. Comparison of Fe^{3+}/Fe_{tot} determined by chemical analysis and by Mössbauer spectroscopy.

al. [1988], low-temperature oxidation to (titano)maghemites has been observed in highly magnetic soils from Brazil.

X-Ray Diffraction and Chemical Composition

Four samples (MAN-74-177, 217, 342A, and 608A) were analyzed by X ray diffraction techniques for the presence of phyllosilicates. The Mg-saturated, oriented samples all had lines at ~1.45 nm (very weak for MAN-74-608A), which is indicative of the presence of phyllosilicates. After glycerol solvation of MAN-74-177, 217, and 342A, only a single phyllosilicate line at 1.8 nm, which is diagnostic for smectite, was present. Glycerol solvation of MAN-74-608A showed that the small amount of phyllosilicate is smectite, chlorite, and mica in approximately equal proportions. A qualitative estimate of the total phyllosilicate content is 1-2, 7, 12, and 15 wt % for MAN-74-608A, 177, 217, and 342A, respectively. The amount of phyllosilicate in these samples is significantly larger than we observed previously in Hawaiian palagonitic soils [*Golden et al.*, 1993; *Morris et al.*, 1993]. *Floran et al.* [1978] and *Allen et al.* [1982] also report smectite in Manicouagan samples based on petrographic and X ray diffraction data, respectively.

From their petrographic observations, *Floran et al.* [1978] concluded that smectite formed as a part of the oxidative alteration of silicate glass and mafic minerals (especially olivine). Our limited XRD data are consistent with this view. MAN-74-608A, which is the least oxidatively altered (lowest Fe^{3+}/Fe_{tot}) sample we studied, has the least smectite. Because Mössbauer data are sensitive only to iron-bearing smectite and because the iron content of the smectites is not known and may be variable, smectite contents determined by Mössbauer mineralogy (Table 2) would not be expected to correlate with the degree of alteration as measured by Fe^{3+}/Fe_{tot}. In addition, smectite can be destroyed if oxidative alteration took place at

Table 3. Compositional Data for Impact Melt Rocks from Manicouagan Crater

	All 24 Samples		12 Samples: Highest Fe^{3+}/Fe_{tot}		12 Samples: Lowest Fe^{3+}/Fe_{tot}		Nine Samples (This Study)	
	Average	SD	Average	SD	Average	SD	Average	SD
			Concentrations, wt %					
SiO_2	57.75	1.18	57.90	1.09	57.60	1.25		
TiO_2	0.77	0.06	0.79	0.03	0.75	0.08		
Al_2O_3	16.51	0.62	16.63	0.72	16.39	0.47		
Fe_2O_3	3.98	0.89	4.66	0.54	3.29	0.58		
FeO	2.29	0.74	1.69	0.35	2.88	0.50		
MnO	0.11	0.02	0.10	0.04	0.11	0.01		
MgO	3.50	0.53	3.25	0.41	3.75	0.53		
CaO	5.84	0.95	5.54	0.35	6.29	1.18	5.40	0.70
Na_2O	3.85	0.27	3.96	0.27	3.67	0.19	3.88	0.22
K_2O	3.04	0.54	3.12	0.19	2.94	0.47	3.00	0.80
P_2O_5	0.22	0.03	0.24	0.02	0.21	0.03		
H_2O	1.73	0.36	1.84	0.31	1.63	0.37		
Fe	4.58	0.28	4.57	0.23	4.54	0.15	4.55	0.41
			Concentrations, μg/g					
Sc	18.08	1.97	17.26	1.05	18.40	1.96	18.19	0.75
Cr							75	14
Rb	70.9	28.8	77.4	25.5	76.2	34.6	57.0	6.9
Ba	1276	306	1459	201	1098	213	1440	257
La	37.76	5.37	39.5	5.2	35.6	4.7	37.4	5.0
Ce	78	10	81	9	74	9	76	10
Sm	6.64	0.68	6.91	0.63	6.41	0.61	6.35	0.76
Eu	1.85	0.24	1.91	0.24	1.85	0.21	1.72	0.17
Tb	0.810	0.166	0.820	0.160	0.830	0.190	0.740	0.070
Yb	2.38	0.17	2.43	0.17	2.37	0.16	2.26	0.17
Lu	0.363	0.033	0.375	0.032	0.356	0.036	0.341	0.026

The first set of columns is the average composition and its standard deviation for all 24 samples analyzed by *Floran et al.* [1978]. The second and third sets of columns are equivalent data for samples having the 12 highest and 12 lowest values of Fe^{3+}/Fe_{tot}, respectively. The last set is for nine samples analyzed in this study.

sufficiently high temperatures (500-800°C). This may explain, for example, the correspondence between the Mössbauer spectrum for MAN-74-177 and those for MAN-74-608A and 342A heated to 1000°C and the relatively lower proportion of smectite in MAN-74-177 compared to 342A (Table 2).

Despite large differences in mineralogy from hematite-dominated to pyroxene-dominated assemblages for Manicouagan impact melt rocks, their chemical composition is very homogeneous. As discussed by *Floran et al.* [1978], there are no statistically significant positional (vertical, lateral, and radial) chemical variations within the impact melt sheet. In Table 3, we used the same chemical data to show that the process of oxidative alteration also did not impart statistically significant variations in chemical composition. The average chemical composition for the 12 samples having the highest value of the Fe^{3+}/Fe_{tot} (most oxidized) is the same as that for the 12 least oxidized samples. The average concentration of iron (expressed as Fe) for the two groups is essentially identical, for example. In Table 3 we also report the average composition determined from INAA data for nine additional Manicouagan samples. There is no statistical difference between these data and those reported by *Floran et al.* [1978].

Reflectivity Spectra

Reflectivity spectra for samples with the lowest and highest proportions of ferric iron are shown in Figure 4 (Mössbauer spectra in Figure 1). In agreement with Mössbauer mineralogy, the reflectivity spectra of the most oxidized and least oxidized samples (three each) are dominated by the spectral characteristics of hematite and pyroxene, respectively. The changes in spectral slope near 520 and 620 nm, the relative reflectivity maximum and the band minimum (or plateau) centered near 860 nm for the most oxidized samples are all characteristic of bulk-Hm (i.e., crystalline hematite) [e.g., *Morris et al.*, 1985]; these samples are light pink to salmon in color. The two-band signature of othopyroxene or type-B clinopyroxene [e.g., *Cloutis and Gaffey*, 1991] is clearly evident in the spectra of the least oxidized samples, which are light gray in color. The band minima range from 950 and ~1800 nm (MAN-74-512A) to 1040 and >2100 nm (MAN-74-407). Our spectral data do not extend to long enough wavelengths to define well the minimum for the second band. The spectral features of bulk-Hm and pyroxene are superimposed on a relatively featureless ferric absorption edge through the visible and near-IR. This ferric edge is attributed to nanophase hematite, which has appropriate spectral properties [*Morris et al.*, 1989; *Morris and Lauer*, 1990] and/or paramagnetic ferric iron observed in the Mössbauer data.

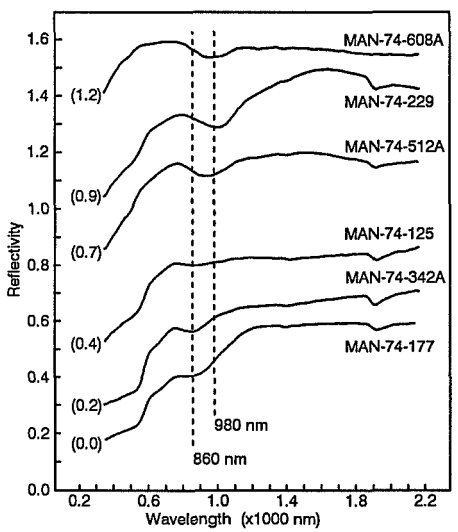


Figure 4. Diffuse reflectivity spectra of powders of Manicouagan impact melt rocks having lowest and highest values of Fe^{3+}/Fe_{tot}. Spectra are offset for clarity.

No bands attributable to electronic bands of ferrous and/or ferric iron in smectite are apparent in the spectra. Based on Mössbauer mineralogy, sample MAN-74-342A has the highest proportion of iron-bearing smectite. As discussed above, the spectral features resulting from iron are dominated by those for bulk-Hm. The bands at ~1400 and 1900 nm, which are combination and overtone bands of H_2O fundamental vibrations, probably result in part if not wholly from smectites. Note that the reflectivity spectrum of MAN-74-608A, which is the least oxidized sample and contains only a trace of phyllosilicates, has no discernible bands at 1400 and 1900 nm.

Figure 5 is a correlation of the position of the near-IR band and the ratio Hm/(Hm+Px); the six samples shown in Figures 1 and 4 are indicated in the figure. The hematite used in this calculation is bulk-Hm; np-Hm does not have a well-defined band minimum [*Morris et al.*, 1989; *Morris and Lauer*, 1990]. A linear correlation is indicated by the data. We considered two extreme models to explain this correlation. (1) The linear correlation results from simple (optical) mixing between pyroxene-dominated and hematite-dominated assemblages (Figure 5, model 1). This could result from progressive oxidation of pyroxene to hematite. Figure 6 does show that the position of the near-IR band minimum does correlate with total-sample Fe^{3+}/Fe_{tot}. (2) The linear correlation results from a fortuitous relationship among optically dominant phases (two pyroxenes and hematite) and Hm/(Hm+Px) (Figure 5, model 2). Based on microprobe data, *Floran et al.* [1978] showed that two main compositional groups of pyroxene were present before oxidative alteration. The average composition of these pyroxenes is $Wo_{40}En_{35}Fs_{15}$ (high-Ca pyroxene) and $Wo_4En_{58}Fs_{38}$ (low-Ca pyroxene). According to the study of pyroxenes by *Cloutis and Gaffey* [1991], these compositions generally yield (assuming two-band pyroxenes) bands in the range 920±10 and 1870±70 nm for low-Ca pyroxene and 1000±30 and 2200±150 nm for high-Ca pyroxene. As shown in Figure 5 model 2, the data can be grouped into low-Ca and high-Ca groups of pyroxenes on this basis. The average values of Hm/(Hm+Px) associated with these two groups are ~0.14 and ~0.41, respectively.

Can high-Ca pyroxenes be reasonably associated with relatively lower whole-rock Hm/(Hm+Px) ratios? As shown in Figure 6, this is equivalent to high-Ca pyroxenes being associated with the least oxidized rocks. The stability order in the terrestrial weathering environment is olivine <

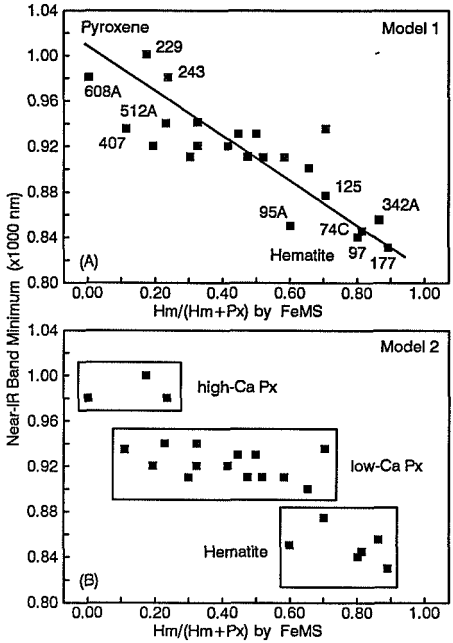


Figure 5. Position of near-IR band minimum as a function of Hm/(Hm + Px) as determined by Mössbauer spectroscopy. The ratio Hm/(Hm + Px) is ratio of the molar amount of Fe associated with hematite (bulk Hm) to the molar amount of Fe associated with hematite plus pyroxene.

hypersthene (low-Ca pyroxene) < augite (high-Ca pyroxene) [e.g., *Huang*, 1977]. On this basis, we would expect low-Ca pyroxene to be preferentially in the least oxidized rocks, which is contrary to observation. However, differential alteration could still be consistent with the data if the rocks in the high-Ca pyroxene group had, prior to alteration, significantly more high-Ca pyroxene than low-Ca pyroxene. The normative calculations of *Floran et al.* [1978] indicate that this is in fact the case. For samples that where both reflectivity and normative data are available, prealteration values of augite/(augite+hypersthene) are 0.72 and 0.39 for high-Ca and low-Ca groups, respectively.

In summary, it is clear from reflectivity, chemical, and petrographic data that both high-Ca and low-Ca pyroxenes are present. Although these pyroxenes are characterized by different band minima, oxidation of each to hematite could produce overlapping mixing relationships between their band minima and that for hematite (model 1). However, it is also possible that only three band minima occur (980, 920, and 850 nm for high-Ca and low-Ca pyroxene and hematite, respectively) and that these phases, from a special set of circumstances, are optically dominant in samples with increasing whole-rock values of Hm/(Hm+Px). This is reasonable for hematite, but not necessarily for the pyroxenes. As discussed next, examination of several samples with 900- to 920-nm band minima shows that likely candidates for both models are present among the samples.

In Figure 7 are Mössbauer and reflectivity spectra for five samples having band minima between 900 and 920 nm. The Mössbauer data show that both pyroxene and (bulk) hematite are present. If the 800- to 1000-nm region is not considered, optical manifestations of bulk-Hm include a relative reflectivity maximum near 750 nm, and inflections near 520 and 620 nm. A broad band centered in the 1600- to 2300-nm region is a manifestation of pyroxene. If any of these hematite or pyroxene features are present, a corresponding features in the 800- to 1000-nm region might be expected. The hematite features at 520, 620, and 750 nm are most evident in the spectrum of MAN-74-217, much less intense in 176 and 506, and virtually absent in 131 and 133. A ~2000 nm band is clearly present for MAN-74-506, and present but much weaker for the other four samples. For MAN-74-506, 131, and 133, the spectral contrast for the ~2000-nm band may be reduced relative to the 800- to 1000-nm band because magnetite (which is optically opaque [e.g., *Morris et al.*, 1985]) is present and because the ~2000-nm band is normally less intense in pyroxenes [*Cloutis and Gaffey*, 1991]. Thus, optical contributions in the 800- to 1000-nm region from both hematite and pyroxene seem likely for MAN-74-217 and 176. Optical contributions from pyroxene alone seem to dominate the spectra of MAN-74-506, 131, and 133.

Application to Mars

Reflectivity Spectra

Spectral data from the Imaging Spectrometer for Mars (ISM) instrument on the Phobos 2 spacecraft are currently the highest spatial resolution data for Martian surface materials. In the 800- to 1300-nm region, these data show band minima extending from ~850 to ~980-nm [*Murchie et al.*, 1993; *Mustard et al.*, 1993]. In agreement with previous interpretations of the same band from lower spatial resolution data [e.g., *Morris et al.*, 1989; *Morris and Lauer*, 1990; *Bell et al.*, 1990], bands near ~850 nm were assigned by *Murchie et al.* [1993] to bulk (well-crystalline) hematite. Bands near 980 nm

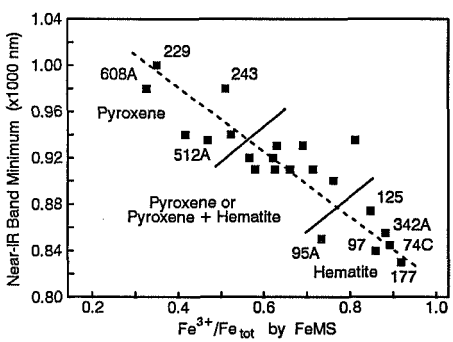


Figure 6. Position of near-IR band minimum as a function of Fe^{3+}/Fe_{tot}.

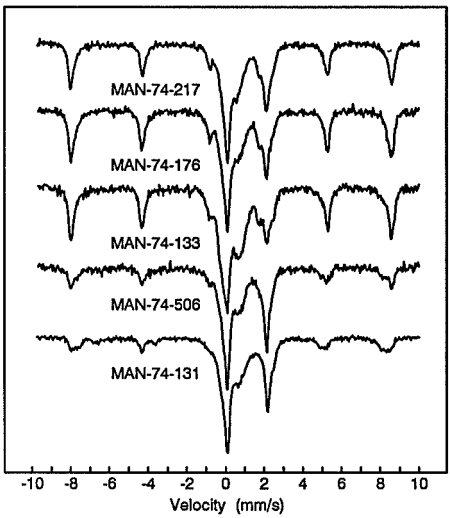


Figure 7a. Mössbauer spectra for five Manicouagan impact melt rocks having near-IR band minima in the region 900-920 nm.

were assigned to calcic pyroxene or possibly an intimate mixture of pigeonite and augite [*Mustard et al.*, 1993]. There seems to be general agreement on the mineralogical assignment of hematite and pyroxene to the bands near 860 and 980 nm, respectively. Can unequivocal mineralogical assignments be made for band minima at intermediate wavelengths?

In Figure 8 are Phobos 2 ISM spectra [*Murchie et al.*, 1993] whose band minima range from ~850 nm (hematite) to intermediate band positions near 920 nm. *Mustard et al.* [1993] consider that the origin of intermediate positions is areal mixing between low-albedo (optically pyroxene) material of volcanic origin and highly altered (optically hematite) bright dust or soils. *Murchie et al.* [1993] consider several models but favor the interpretation that these spectra indicate the presence of an additional ferric-bearing phase such as goethite (α-FeOOH), ferrihydrite ($5Fe_2O_3.9H_2O$), or jarosite ($(K,Na,H_3O)Fe_3(SO_4)_2(OH)_6$). As discussed next, we advocate here an alternate hypothesis: that the intermediate band centers represent intimate hematite-pyroxene assemblages and/or pyroxene analogous to the Manicouagan impact melt rocks.

In Figure 9 we plot reflectivity spectra between 800 and 1300 nm for selected impact melt rocks from Manicouagan Crater. As in Figure 8, the spectra are arranged in order of increasing wavelength of the band minimum. If we did not have Mössbauer mineralogy and magnetic and chemical data for these samples, we would be in a similar position for their interpretation as we are for the Phobos 2 data. Namely, we would with a high degree of confidence assign hematite and pyroxene to the ~850 and 980- to 1000-nm band positions. We would not be able to make unequivocal assignments for intermediate band positions. Possible assignments would include intimate hematite-pyroxene assemblages, the ferric-bearing phases suggested by *Murchie et al.* [1993] (goethite, ferrihydrite, and jarosite), other ferric bearing phases (lepidocrocite and maghemite [*Morris et al.*, 1985]; nontronite [*Singer*, 1982]), and some pyroxenes [*Cloutis and Gaffey*, 1991]. However, Mössbauer mineralogy and other data constrain the Manicouagan spectra to only two of the possible interpretations: hematite-pyroxene assemblages and/or pyroxene. Reasonable, but not unequivocal, assignments for Martian spectra having band minima near 910 nm are also hematite-pyroxene assemblages and/or pyroxene.

There is conflicting and controversial evidence for the existence of phyllosilicate minerals on Mars. *Toulmin et al.* [1977] inferred the existence of nontronite using mass balance constraints from elemental abundances derived from Viking X ray fluorescence data. *Toon et al.* [1977] interpreted Mariner 9 mid-IR spectra as indicating the presence of montmorillonite in the airborne dust. Features near 1500 and 2300 nm in near-IR telescopic observations of *McCord et al.* [1982] were interpreted as evidence for (Al,Fe,Mg)-OH absorption in phyllosilicates. The features near 2300 nm were examined in more detail by *Clark et al.* [1990b], who interpreted the absorptions as evidence of the bicarbonate-bearing mineral scapolite rather than of phyllosilicates. More recently, *Bell and Crisp* [1993], *Murchie et al.* [1993], and *Bell et al.* [1994] have presented evidence from ISM and ground-based spectra for subtle features in the 2200- to 2400-nm region that could be indicative of phyllosilicates.

We found no unequivocal evidence in our reflectivity data of Manicouagan samples for electronic transitions of ferric and ferrous iron in structural sites of phyllosilicates, in spite of the fact that up to ~15 wt % smectite is observed in our samples. This circumstance implies that the smectite does not contain much iron. This is consistent with our FeMS data which show that only 3 of 23 samples have more that 10% of their total

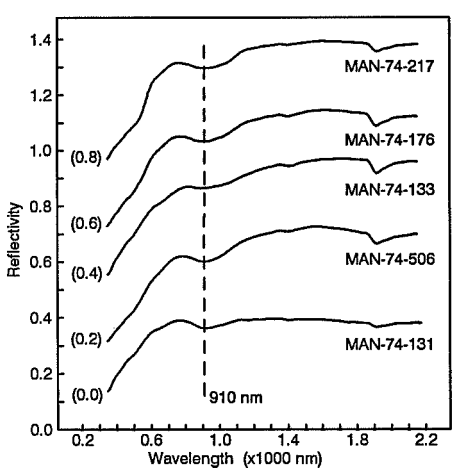


Figure 7b. Same as Figure 7a for reflectivity spectra.

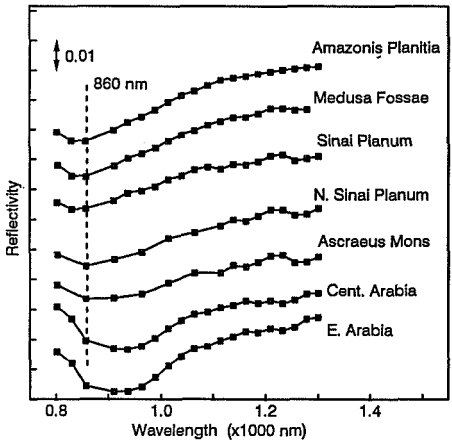


Figure 8. Phobos 2 ISM data from 800 to 1300 nm [from *Murchie et al.*, 1993].

iron associated with phyllosilicates. More definitive spectral evidence for phyllosilicates would be (Al,Mg,Fe)-OH transitions near 2300 nm [e.g., *Clark et al.*, 1990a,b]. Unfortunately, our data do not extend beyond 2100 nm.

Impact Processes on Mars

Oxidative alteration of the Manicouagan impact melt rocks is thought to occur shortly after the impact by influx of oxidizing vapors and/or fluids while the rocks are still hot but below solidus temperatures [*Floran et al.*, 1978]. For Manicouagan, the solidus is estimated at about 915°C [*Simonds et al.*, 1978]. Based on the petrographic observations of *Floran et al.* [1978] and *Phinney et al.* [1978] for Manicouagan and West Clearwater Lake impact melt rocks, respectively, Ti-bearing (and relative large) hematite particles are formed by oxidative alteration of primary titanomagnetite. Essentially Ti-free hematite particles extending in size from micron-sized particles (bulk-Hm) to particles below the limit of optical microscopes (nanophase hematite) are formed from glass and mafic minerals. Because of their small size and high number density, the essentially Ti-free particles are likely the dominant ferric iron pigmenting phase. In addition to hematite, *Floran et al.* [1978] and *Phinney et al.* [1978] include hydrous ferric oxides as potential oxidative alteration products; however, our data (primarily the Mössbauer mineralogy) show that only hematite is present, at least for Manicouagan samples. An explanation for the absence of ferrihydrite and ferric oxyhydroxide phases (e.g., goethite and lepidocrocite) and presence of hematite is that the oxidative alteration took place at temperatures above 250-300°C where those phases are thermally unstable relative to hematite [*Morris and Lauer*, 1981]. The phyllosilicates (primarily smectite) must also have formed (from silicate minerals and glass, according to *Floran et al.* [1978]) above these temperatures.

Impacts have obviously been an important surface modification and alteration process throughout Martian history. A number of studies [*Newsom*, 1980; *Allen et al.*, 1982; *Clifford*, 1993] have advocated that a significant portion of Martian soil represents erosional products of hydrothermally altered impact melt sheets. For example, *Clifford* [1993] has calculated an equivalent global impact melt layer of 521 m. By analogy with Manicouagan impact melt rocks and in agreement with observables for Mars, oxidative alteration of Martian impact melt sheets above 250°C and subsequent erosion could produce rocks and soils with variable proportions of hematite (both bulk and nanophase), pyroxene, and phyllosilicates as iron-bearing mineralogies. The Manicouagan samples show that this mineralogical diversity can be accomplished at constant chemical composition, which is also indicated for Mars from analyses of soil at the two Viking landing sites [*Clark et al.*, 1982]. If this process is volumetrically important, these phases on the Martian surfaces were formed rapidly at relatively high temperatures on a sporadic basis throughout the history of the planet. This process is in contrast to slow, relatively continuous, low-temperature processes involving dissolution of ferrous-bearing phases and precipitation of low-temperature and ferric-bearing phases such as goethite, lepidocrocite, and ferrihydrite [e.g., *Banin et al.*, 1993; *Burns*, 1993]. In this case, the heat associated with repetitive meteoritic impact and volcanic processes would convert hydrous ferric oxides to their anhydrous forms (e.g., hematite) as observed for Hawaiian palagonitic soil PH-1 [*Bell et al.*, 1993].

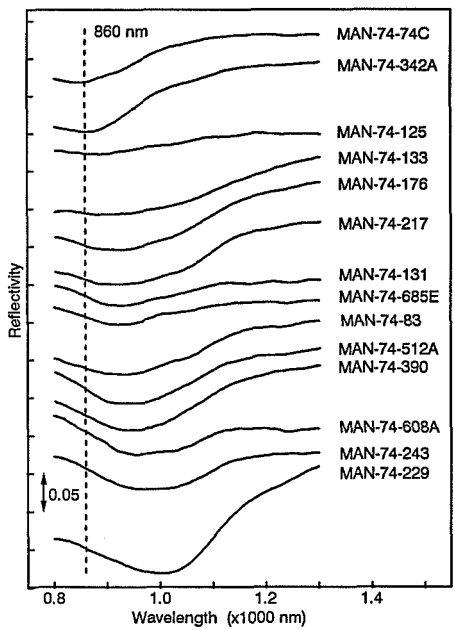


Figure 9. Manicouagan reflectivity data from 800 to 1300 nm. As in Figure 8, the spectra are arranged according to band minimum from lowest to highest wavelength.

References

Allan, J. E. M., J. M. D. Coey, M. Resende, and J. D. Fabris, Magnetic properties of iron-rich oxisols, *Phys. Chem. Miner., 15*, 470-475, 1988.

Allen, C. C., J. L. Gooding, and K. Keil, Hydrothermally altered impact melt rock and breccia: Contributions to the soil of Mars, *J. Geophys. Res., 87*, 10,083-10,101, 1982.

Banin, A., T. Ben-Shlomo, L. Margulies, D. F. Blake, R. L. Mancinelli, and A. U. Gehring, The nanophase iron mineral(s) in Mars soil, *J. Geophys. Res., 98*, 20,831-20,853, 1993.

Bell, J. F., III, J.F. and D. Crisp, Groundbased imaging spectroscopy of Mars in the near-infrared: Preliminary results, *Icarus, 104*, 2-19, 1993.

Bell, J. F., III, T. B. McCord, and P. D. Owensby, Observational evidence of crystalline iron oxides on Mars, *J. Geophys. Res., 95*, 14,447-14,461, 1990.

Bell, J. F., III, R. V. Morris, and J. B. Adams, Thermally altered palagonitic tephra: A spectral and process analog to the soil and dust of Mars, *J. Geophys. Res., 98*, 3373-3385, 1993.

Bell, J. F., III, J. B. Pollack, T. R. Geballe, D. P. Cruikshank, and R. Freedman, Spectroscopy of Mars from 2.0 to 2.5 microns during the 1993 opposition: Atmospheric vs. mineralogic origin of absorption features, *Icarus*, in press, 1994.

Borchardt, G. A., Montmorillonite and other smectite minerals, in *Minerals in Soil Environments*, pp. 293-330, Soil Science Society of America, Madison, Wisc., 1977.

Burns, R. G., Rates and mechanisms of chemical weathering of ferromagnesian silicate minerals on Mars, *Geochim. Cosmochim. Acta, 57*, 4555-4574, 1993.

Clark, B. C., A. K. Baird, R. J. Weldon, D. M. Tsusaki, L. Schnabel, and M. P. Candelaria, Chemical composition of Martian fines, *J. Geophys. Res., 87*, 10,059-10,067, 1982.

Clark, R. N., T. V. V. King, M. Klejwa, and G. A. Swayze, High spectral resolution reflectance spectroscopy of minerals, *J. Geophys. Res., 95*, 12,653-12,680, 1990a.

Clark, R. N., G. A. Swayze, R. B. Singer, and J. B. Pollack, High-resolution reflectance spectra of Mars in the 2.3 μm region: Evidence for the mineral scapolite, *J. Geophys. Res., 95*, 14,463-14,480, 1990b.

Clifford, S. M., A model for the hydrologic and climatic behavior of water on Mars, *J. Geophys. Res., 98*, 10,973-11,016, 1993.

Cloutis, E. A., and M. J. Gaffey, Pyroxene spectroscopy revisited: Spectral-compositional correlations and relationship to geothermometry., *J. Geophys. Res., 96*, 22,809-22,826, 1991.

De Grave, E., and A. Van Alboom, Evaluation of ferrous and ferric Mössbauer fractions, *Phys. Chem. Minerals, 18*, 337-342, 1991.

Floran, R. J., C. H. Simonds, R. A. Grieve, W. C. Phinney, J. L. Warner, M. J. Rhodes, B. M. Jahn, and M. R. Dence, Petrology, structure and origin of the Manicouagan melt sheet, Quebec, Canada: A preliminary report, *Geophys. Res. Lett., 3*, 49-52, 1976.

Floran, R. J., R. A. Grieve, W. C. Phinney, J. L. Warner, C. H. Simonds, D. P. Blanchard, and M. R. Dence, Manicougan impact melt, Quebec, 1, Stratigraphy, petrology, and chemistry, *J. Geophys. Res., 83*, 2737-2759, 1978.

Golden, D. C., R. V. Morris, H. V. Lauer Jr., and S. R. Yang, Mineralogy of three slightly palagonitized tephra samples from the summit of Mauna Kea, Hawaii, *J. Geophys. Res., 98*, 3401-3411, 1993.

Huang, P. M., Feldspars, olivines, pyroxenes, and amphiboles, in Minerals in Soil Environments, pp. 553-602, Soil Science Society of America, Madison, Wisc., 1977.

Kieffer, S. W., and C. H. Simonds, The role of volatiles and lithology in the impact cratering process, *Rev. Geophys., 18*, 143-181, 1980.

McCord, T. B., R. N. Clark,. and R. B. Singer, Mars: Near-infrared spectral reflectance of surface regions and compositional implications, *J. Geophys. Res., 87*, 3021-3032, 1982.

Morris, R. V., and H. V. Lauer Jr., Stability of goethite (α-FeOOH) and lepidocrocite (γ-FeOOH) to dehydration by UV radiation: Implications for their occurrence on the Martian surface, *J. Geophys. Res., 86*, 10,893-10,899, 1981.

Morris, R. V., and H. V. Lauer Jr., Matrix effects for reflectivity spectra of dispersed nanophase (superparamagnetic) hematite with application to Martian spectral data, *J. Geophys. Res., 95*, 5101-5109, 1990.

Morris, R. V., H. V. Lauer Jr., C. A. Lawson, E. K. Gibson Jr., G. A. Nace, and C. Stewart, Spectral and other physicochemical properties of submicron powders of hematite (α-Fe_2O_3), maghemite (γ-Fe_2O_3), magnetite (Fe_3O_4), goethite (α-FeOOH), and lepidocrocite (γ-FeOOH), *J. Geophys. Res., 90*, 3126-3144, 1985.

Morris, R. V., D. G. Agresti, H. V. Lauer Jr., J. A. Newcomb, T. D. Shelfer, and A. V. Murali, Evidence for pigmentary hematite on Mars based on optical, magnetic, and Mössbauer studies of superparamagnetic (nanocrystalline) hematite, *J. Geophys. Res., 94*, 2760-2778, 1989.

Morris, R. V., J. J. Gooding, H. V. Lauer Jr., and R. B. Singer, Origins of Marslike spectral and magnetic properties of a Hawaiian palagonitic soil, *J. Geophys. Res., 95*, 14,427-14,434, 1990.

Morris, R. V., D. C. Golden, J. F. Bell III, H. V. Lauer Jr., and J. B. Adams, Pigmenting agents in Martian soils: Inferences from spectral, Mössbauer, and magnetic properties of nanophase and other iron oxides in Hawaiian palagonitic soil PN-9, *Geochim. Cosmochim. Acta, 57*, 4597-4609, 1993.

Moskowitz, B. M., and R. B. Hargraves, Magnetic changes accompanying the thermal decomposition of nontronite (in air) and its relevance to Martian mineralogy, *J. Geophys. Res., 87*, 10,115-10,128, 1982.

Murchie, S., J. Mustard, J. Bishop, J. Head, and C. Pieters, Spatial variations in the spectral properties of bright regions on Mars, *Icarus, 105*, 454-468, 1993.

Mustard, J. F., S. Erard, J. P. Bibring, J. W. Head, S. Hurtrez, Y. Langevin, C. M. Pieters, and C. J. Sotin, The surface of Syrtis Major: Composition of the volcanic substrate and mixing with altered dust and soil, *J. Geophys. Res.*, 98, 3387-3400, 1993.

Newsom, H. E., Hydrothermal alteration of impact melt sheets with implications for Mars, *Icarus, 44*, 207-216, 1980.

Newsom, H.. E., G. Graup, T. Sewards, and K. Keil, Fluidization and hydrothermal alteration of the suevite deposit at the Ries Crater, West Germany, and implications for Mars, J. Geophys. Res., 91, E239-E251, 1986.

Phinney, W. C., C. H. Simonds, A. Cochran, and P. E. McGee, West Clearwater, Quebec impact structure, Part II: Petrology, *Proc. Lunar Planet. Sci. Conf. 9th*, 2659-2693, 1978.

Pohl, J., D. Stoffler, H. Gall, and K. Emstson, The Ries impact crater, in *Impact and Explosion Cratering*, pp. 343-404, Pergamon, New York, 1977.

Simonds, C. H., R. J. Floran, P. E. McGee, W. C. Phinney, and J. L. Warner, Petrogenesis of melt rocks, Manicouagan impact structure, Quebec, *J. Geophys. Res., 83*, 2773-2788, 1978.

Singer, R. B., Spectral evidence for the mineralogy of high-albedo soils and dust on Mars, *J. Geophys. Res., 87*, 10,159-10,168, 1982.

Straub, D. W., R. G. Burns, and S. F. Pratt, Spectral signature of oxidized pyroxenes: Implications to remote sensing of terrestrial planets, *J. Geophys. Res., 96*, 18819-18830, 1991.

Toon, O. B., J. B. Pollack, and C. Sagan, Physical properties of the particles composing the Martian dust storm of 1971-1972, *Icarus, 30*, 663-696, 1977.

Toulmin, P., III, A. K. Baird, B. C. Clark, K. Keil, H. J. Rose Jr., R. P. Christian, P. H. Evans, and W. C. Kelliher, Geochemical and mineralogical interpretation of the Viking inorganic chemical results, *J. Geophys. Res., 82*, 4625-4634, 1977.

Weldon, R. J., W. M. Thomas, M. B. Boslough, and T. J. Ahrens, Shock-induced color changes in nontronite: Implications for the Martian fines, *J. Geophys. Res., 87*, 10,102-10,114, 1982.

J. F. Bell III, National Research Council/NASA Ames Research Center, Space Sciences Division, Moffett Field, CA 94035.

D. C. Golden, Dual, Inc., Houston, TX 77058.

H. V. Lauer Jr., Lockheed ESC, Houston, TX-77058

R. V. Morris, Code SN4, NASA Johnson Space Center, Houston, TX 77058. (e-mail: morris@snmail.jsc.nasa.gov)

(Received March 3, 1994; accepted June 4, 1994.)

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5329–5340, MARCH 25, 1995

S ≠ NC: Multiple source areas for Martian meteorites

Allan H. Treiman
Lunar and Planetary Institute, Houston, Texas

Abstract. The Shergottite, Nakhlite, and Chassigny (SNC) meteorites are igneous rocks from Mars, inferred previously to have been propelled into space (and to Earth) by a single impact event. However, the sum of petrologic, chemical, and chronological data suggests two distinct sites of origin: S from one, NC from another. Differences between the S and NC include distributions of terrestrial residence ages, cosmic ray exposure ages, impact shock effects, pre-terrestrial aqueous alteration products, bulk magma compositions, trace element magma chemistries, and mantle source compositions. If the S and NC did not come from the same site on Mars, selection criteria for determining source craters are relaxed, and a number of craters become candidate sources. Formation of two meteorite source craters on young Martian surfaces (≤1300 Ma) within the last 20 m.y. (spanning the cosmic ray exposure ages) seems to require either smaller source craters or greater cratering rates than allowed in current theories. Both of these options suggest that meteorites from nonvolcanic and ancient Martian surfaces ought to be more common than the known types of Martian meteorites. The absence of such meteorites may reflect the physical properties of nonvolcanic and ancient surfaces, or a failure to recognize these rock types as meteorites. The ALH84001 meteorite, long classified as a diogenite, may be such a sample of the ancient Martian crust.

Introduction: Rocks From Mars

Three small classes of basaltic meteorites, collectively called the SNCs (for Shergottites, Nakhlites, and Chassigny), are samples of the Martian crust, and so are of surpassing interest to the planetary sciences. Lacking Martian samples returned by spacecraft missions, the SNC meteorites have stood as proxy returned samples for petrologic and geological studies [*McSween*, 1985]. Unlike a sample return mission, however, the source sites of the SNCs are not known, so their utility is greatly diminished. There have been attempts to determine the Martian source of the SNCs, all of which have assumed or concluded that the meteorites came from a single source crater [*Wood and Ashwal*, 1981; *Nyquist*, 1983; *Vickery and Melosh*, 1987; *Mouginis-Mark et al.*, 1992]. *Mouginis-Mark et al.* [1992] located nine potential source craters in Tharsis, but none are consistent with constraints imposed by all of the SNCs. Multiple source craters on Mars have been considered but rejected [*Bogard et al.*, 1984; *Vickery and Melosh*, 1987], partly on cratering and orbital dynamic bases and partly because all SNCs were inferred to have the same crystallization age.

However, differences between the S and NC meteorites are significant, and suggest that these two groups of meteorites experienced different histories from mantle differentiation to the time they landed on Earth. These differences are most consistent with the S originating at a different impact crater than the NC [*Treiman*, 1993a, 1994a]. The inference of multiple source craters has broad implications for the geology and petrology of the SNC meteorites, and furthermore for impact cratering rates on Mars, impact mechanics, and meteoritics.

Paper number 94JE02184.
0148-0227/94/94JE-02184$05.00

The Martian Meteorites

The most convincing evidence that the SNC meteorites are from Mars is that they contain within them a trapped gas component identical in every measurable respect to Mars' current surface atmosphere [*Bogard et al.*, 1984; *Wiens and Pepin*, 1988]. The Mars atmosphere is unique among gas reservoirs in the solar system, notably for its abundance ratios of $^{14}N/^{15}N$, $^{40}Ar/^{36}Ar$, $^{129}Xe/^{132}Xe$, etc. [*Pepin*, 1989], leaving essentially no doubt that the SNCs formed on Mars. Even lacking this comparison (based on Viking lander data), there is a strong case for the SNCs' Martian origin, based on their volatile element contents and oxidation states, indicating a large source planet, their young crystallization ages (1300 and ~180 Ma) and trace element fractionations, indicating recent and recurrent igneous activity, presence of hydrous magmatic minerals, indicating a hydrous mantle or crust, and their preterrestrial aqueous alteration materials, indicating an active hydrosphere within the last 1300 to ~180 m.y. [e.g., *Wood and Ashwal*, 1981; *McSween*, 1985; *Gooding*, 1992]. Of all bodies in the solar system, only Mars and the Earth could yield such rocks. A terrestrial origin is precluded by the oxygen isotope composition and geochemistry of the SNCs [*Clayton and Mayeda*, 1983].

The shergottites, abbreviated here as S, include basaltic and cumulate igneous rocks with abundant low-calcium pyroxenes. The Shergotty, Zagami and EETA79001 meteorites are basalts, with or without additional (cumulate) pyroxene [*Stolper and McSween*, 1979; *McSween and Jarosewich*, 1983]. EETA79001 contains two basalt lithologies, A and B, in igneous contact. The ALHA77005 and LEW88516 meteorites are websterites or lherzolites, cumulate rocks of olivine and low-calcium pyroxene with lesser augite and plagioclase [*McSween et al.*, 1979; *Jagoutz*, 1989; *Harvey et al.*, 1993; *Treiman et al.*, 1994]. The EETA79001A lithology contains xenoliths and xenocrysts from a websterite or lherzolite, here called EETA79001X. For convenience, the Martian lherzolites are grouped here with the

shergottites as S, although they are not shergottites in the strict definition [*McSween et al.*, 1979; *Lindstrom et al.*, 1994]. Amphibole of igneous origin is present inside pyroxene grains in most S [*Treiman*, 1985, and unpublished data, 1994].

The nakhlite meteorites, abbreviated here as N, are Nakhla itself, Lafayette, and Governador Valadares [*Berkley et al.*, 1980]. The N are augite-rich igneous rocks with significant proportions of olivine and very little low-calcium pyroxene. They are interpreted to be cumulate rocks, formed from a Ca-rich basaltic magma [*Harvey and McSween*, 1992a; *Treiman*, 1993b]. Amphibole of igneous origin is reported in inclusions in nakhlite olivines [*Harvey and McSween*, 1992a].

The Chassigny meteorite, C here, is a dunite, interpreted as an igneous cumulate [*Floran et al.*, 1978]. Its most abundant pyroxene is augite, some of which may also be cumulate, as may some of its less abundant low-Ca pyroxene [*Boynton et al.*, 1976; *Wadhwa and Crozaz*, 1994]. Plagioclase are not abundant. Olivines in Chassigny contain magmatic inclusions with common grains of water-bearing amphibole, and rare biotite [*Floran et al.*, 1978; *Johnson et al.*, 1991; *Watson et al.*, 1993].

ALH84001 is a newly recognized Martian meteorite, previously classified as a diogenite [*Mittlefehldt*, 1994]. Data on ALH84001 are limited, although it is clearly not an S, N, or C. It will be discussed separately later.

S≠NC: Evidence for Multiple Source Craters

The SNC meteorites have consistently been inferred or concluded to originate from a single impact crater on the Martian surface [*Wood and Ashwal*, 1981; *Nyquist*, 1982; *McSween*, 1985; *Vickery and Melosh*, 1987; *Mouginis-Mark et al.*, 1992], although some studies have considered multiple crater origins [*Bogard et al.*, 1984; *Vickery and Melosh*, 1987; *Jones*, 1989]. A single crater origin was considered most likely in that the SNC meteorites were inferred to have all crystallized at 1300 m.y. [*Wood and Ashwal*, 1981; *Shih et al.*, 1982], and because impact ejection was considered to be an unusual or unique event [*Wood and Ashwal*, 1981; *Nyquist*, 1982]. The idea that normal impacts could have ejected the SNCs seemed to be directly refuted by the glaring absence of meteorites from the Moon. Since these early works, many lunar meteorites have been found in Antarctica and elsewhere, the crystallization ages of the S are now known to be different from those of NC [*Jones*, 1986], and it has been shown that normal meteorite impacts can eject rocks from Mars [*Vickery and Melosh*, 1987]. But the possibility that the SNC meteorites came from multiple source craters is still rejected [e.g., *Mouginis-Mark et al.*, 1992].

My thesis here is that petrologic, chemical, and chronological evidence suggests that the SNCs originated at two distinct sites on Mars. The evidence is summarized in Table 1 in reverse chronological order, from which it appears that the S and NC have been distinct through their whole histories, from source mantle, through magma composition, through low-temperature aqueous alteration, through impact ejection, and through encounter with Earth.

The following discussions document the data and inferences summarized in Table 1, starting with the present and working backward through time (not the order of strongest argument). I have relied on synthesis and summary references where possible, so as not to further burden the reference list; apologies are offered for not citing every relevant contribution.

Terrestrial Residence Age

The terrestrial residence ages of meteorites, the time intervals since they fell to Earth, can be determined either by observation of a fall or by analysis of a meteorite for certain cosmogenic radioactive nuclides [e.g., *Nishiizumi et al.*, 1986]. As summarized in Table 1, the terrestrial ages of the S and NC overlap, but appear to have sampled different age populations. Thus, the ages are consistent with S and NC originating in separate meteoroid populations before coming to Earth.

Table 1. Summary of Selected Properties of Martian Meteorites, Reverse Chronologic Order

Property	S	NC
Terrestrial exposure, years	32 - 200,000	<200
Cosmic ray exposure age, Ma	2.8±0.3, 0.5	11±1
Shock effects	maskelynite after plagioclase; abundant shock glasses; implanted atmosphere gas; disturbed isotopic systems;	original plagioclase preserved; shock glass rare; no implanted atmosphere; no isotopic disturbances
Pre-terrestrial aqueous alteration: silicates (where present)	aluminosilicate clay? poorly crystalline?	smectite-iron oxide; well-crystalline
Crystallization age, Ma	~180	~1300
Magma composition: First pyroxene to crystallize	low Ca (orthopyroxene, pigeonite)	high Ca (augite)
Magma geochemistry: incompatible	depleted versus CI	enriched versus CI
Magma source: mantle	depleted before 1300 Ma; depleted again at 180 Ma	depleted before 1300 Ma; enriched at ~1300 Ma
Magma source: crust	no "crustal" components	two "crustal" components

See text for full explanation.

Shergottite meteorites (in the broad definition above) have been observed to fall (Shergotty, Zagami), and have been collected at three separate sites in Antarctica. Residence ages for these are: ALHA77005, 200,000 years [*Nishiizumi et al.*, 1986, 1992; *Freundel et al.*, 1986]; EETA79001, 12,000±2,000 years [*Nishiizumi et al.*, 1986; *Jull and Donahue*, 1988]; and LEW88516, < 50,000 years [*Nishiizumi et al.*, 1992]. On the other hand, all of the NC meteorites were seen to fall or found shortly after fall and contain no products of terrestrial weathering [*Graham et al.*, 1985; *Gooding et al.*, 1991; *Treiman et al.*, 1993].

The terrestrial residence ages of the S and NC overlap, as meteorites from both groups have been observed to fall. But the absence of NC meteorites from Antarctica may be significant in suggesting that no (or few) NC meteorites fell over the past million years represented by stony meteorites in the Antarctic collections [*Nishiizumi et al.*, 1989]. If the historical proportion of S versus NC falls, two versus four, was the proportion that fell in Antarctica, the probability of finding three S and no NC meteorites from Antarctica is 1/27, or 0.037. This value may be of statistical significance; it is not likely to reflect collection bias, as many pyroxenite and dunite meteorites (like N and C, respectively) have been collected in Antarctica.

Cosmic Ray Exposure Age

The cosmic ray exposure age of a meteorite is the time interval that it spent being irradiated by solar and galactic cosmic rays; the age can be measured by abundances of some isotopes produced during irradiation, the cosmogenic nuclides. Exposure ages for the NC are ~11 Ma, much older than exposure ages for S, which are 2.8 or 0.5 Ma [e.g., *Bogard et al.*, 1984; *Nishiizumi et al.*, 1986; *Treiman et al.*, 1994]. The differences in exposure ages is significant, and has been interpreted in terms of ejection from the parent planet [*Jones*, 1989], disruption of parent meteoroids in space [e.g., *Bogard et al.*, 1984], or a combination of both mechanisms [*Treiman*, 1993c]. Separate source sites (and ejection events) for S and NC would require different exposure ages, but the ages do not necessarily preclude a single source crater [*Bogard et al.*, 1984; *Vickery and Melosh*, 1987].

Shock Metamorphism

Effects of shock metamorphism are apparent in all of the SNC meteorites, presumably reflecting the trauma they suffered on ejection from Mars by meteorite impact. The S have uniformly been shocked to much higher levels than the NC, suggesting significantly different shock histories. Different shock histories are consistent with and probably required by the two-crater scenario, but are also consistent with ejection as different objects from different fields around a single impact.

The shock experienced by the S was "strong to very strong," stages S5-S6 of *Stöffler et al.* [1991]. Petrographic manifestations of this shock include transformation of original plagioclase to maskelynite and to liquid, intense cracking, planar deformation elements, polygonization and even melting of pyroxene and olivine, formation of the high-pressure phases ringwoodite and majorite, and common melting of the bulk rocks [*McSween*, 1985; *Treiman et al.*, 1994]. These data are consistent with peak shock pressures overall of 30-54 GPa, and localized peak pressures in excess of 75 GPa [*Stöffler et al.*, 1991]. This intense shock event has been held responsible for disturbance of radio-isotope chronometers [*Shih et al.*, 1982], and for implantation of Martian atmospheric gases [*Bogard et al.*, 1984; *Wiens and Pepin*, 1988; *Ott and Begemann*, 1985]. The age of the shock event is not known, but may be 15±15 m.y. for ALHA77005 [*Jagoutz*, 1989].

On the other hand, the NC meteorites are effectively "unshocked to weakly shocked," stages S1 to S3 of *Stöffler et al.* [1991]. Petrographic manifestations of these low shock levels include the presence of original plagioclase, olivine with sharp to slightly undulatory extinction, absence of high-pressure phases, and rarity of shock melts [*Floran et al.*, 1978; *Berkley et al.*, 1980]. Greater shock pressures are apparent locally in the NCs, where pyroxenes and olivines may be granulated or twinned, and in which shock melts occur rarely [*Melosh et al.*, 1983]. These data are consistent with overall peak shock pressures of approximately 15 GPa, with localized peak pressures up to 56 GPa. Compared to S, the NC show no indication that shock has affected their radiochronometer systems, nor that shock has had any effect on their trace gas contents [*Ott*, 1988; *Drake et al.*, 1993].

Preterrestrial Aqueous Alterations

Preterrestrial aqueous alteration materials (hydrous silicates, oxides, and ionic salts) in the SNC meteorites are important here as telltales to the compositions of waters at their emplacement site(s) near the surface of Mars. The mineralogy and compositions of the alteration products are different in NC and S, suggesting that their parent magmas were emplaced into terrains of different geochemistry or with different mechanisms of alteration. The geochemistry and mechanisms of alteration are not yet known, and will be a fruitful field of research. However, the differences in alteration materials between the S and NC are great enough to suggest distinctly different altering fluids, consistent with different emplacement sites and post-igneous histories.

The presence of preterrestrial hydrous alteration products in achondrite meteorites has been difficult to establish, even though early students of the N suggested that their "iddingsite" was probably preterrestrial [*Bunch and Reid*, 1975; *Boctor et al.*, 1976]. Stratigraphic proof that the SNC meteorites contained products of preterrestrial aqueous alterations was provided first by *Gooding and Muenow* [1986] and *Gooding et al.* [1988] for the EETA79001 shergottite and by *Gooding et al.* [1991] and *Treiman et al.* [1993] for the N. The mineralogy of the alteration materials are summarized in Table II of *Gooding* [1992]; *Treiman et al.* [1993] provide additional data. *Wentworth and Gooding* [1993] found evidence for preterrestrial Fe-Mg phosphate in ALHA77005, but the remaining secondary minerals in it and LEW88516 could have formed during Antarctic weathering [*Smith and Steele*, 1984]. There have not been extensive searches for preterrestrial alteration products in Shergotty or Zagami.

All of the analyzed SNC meteorites contain ionic salts of preterrestrial or possible preterrestrial origin (see references above). The most common are alkaline earth sulfates (both Ca and Mg) and carbonates (Ca, Mg, Fe, Mn). Silicate alterations are known only in the N and in the shergottite EETA79001, so the comparison of Table 1 is between those groups. The silicate alteration materials in EETA79001 are hydrous aluminous silicates of variable composition and unknown structure [*Gooding and Muenow*, 1986]; some grains are consistent with illite clay, while others are described as "S,Cl-bearing micabole" in that chemical analyses could not be rationalized in terms of

known mineral stoichiometries. The compositional variations and inability to match known stoichiometries strongly suggests that these materials are poorly crystalline, although their crystallinity has not been measured. There are no iron oxide phases in the alteration assemblage.

The alteration materials in the N are coarser and more abundant than in EETA79001, and so are better known. In the N, olivine and pyroxenes have been altered to "iddingsite," an intimate mixture of smectite clay (low Al, high Fe) and iron oxides (hematite and ferrihydrite) [*Gooding et al.*, 1991; *Treiman et al.*, 1993]. These clays are "coarsely" crystalline, with individual crystallites to 5 μm across.

There are some data on the isotopic compositions of alteration materials in the SNC meteorites, which show a few systematic differences between the S and NC. Oxygen in water in the NC has $\Delta^{17}O$ significantly greater than oxygen in the anhydrous silicates [*Karlsson et al.*, 1992], suggesting that NC were altered by a water reservoir that was not equilibrated with the source region of the NC magmas. On the other hand, only one sample of an S meteorite (one aliquot of Zagami) showed a similar enrichment; oxygen in water in Shergotty, EETA79001, and the other Zagami aliquot had $\Delta^{17}O$ near that of the bulk silicates [*Karlsson et al.*, 1992]. This difference in oxygen composition is echoed in heavy noble gas isotopes [*Drake et al.*, 1993], but not in deuterium/hydrogen ratios [*Watson et al.*, 1994].

Crystallization Ages

Ages of igneous crystallization are among the most important data consistent with separate sources for the NC and S meteorites (Table 1). Crystallization age measurements by K-Ar, Rb-Sr, Sm-Nd, and U-Th-Pb radiochronometer systems are summarized by *McSween* [1985] and *Jones* [1986]. For the S, *McSween* [1985] found no basis at that time to choose among 1300, 350, and 180 Ma as being the true ages of crystallization. However, the current understanding of SNC chronology makes it very unlikely that the S crystallized at the same time as the NC, consistent with origins at separate craters.

For the NC, these radiochronometers are all concordant within error at 1300 Ma [*McSween*, 1985]. Given the absence of strong shock effects (see above) or extensive metamorphism in the nakhlites, this age is almost certainly that of igneous crystallization [*Berkley et al.*, 1980; *Harvey and McSween*, 1992b].

The radiochronology of the S is complex; it is accepted here that they crystallized from magma at ~180 Ma [*Jones*, 1986, 1989]. All of the S have at least one radio-isotope chronometric age of ~180 Ma, which originally had been ascribed to shock metamorphism [*Shih et al.*, 1982; *McSween*, 1985]. However, *Jones* [1986] argued that this view was inconsistent with the igneous chemical zoning retained in the minerals of most S (especially Zagami and EETA79001). In particular, the Zagami shergottite retains igneous zoning in all its major minerals, and yields ages of ~180 Ma from Rb-Sr, Sm-Nd, and U-Th-Pb radiochronometers. This age for Zagami is almost certainly that of igneous crystallization, and thus *Jones* [1986] argued that the ~180 Ma ages for other S also represent igneous crystallization. It remains clear, however, that some isotopic systems have been perturbed through shock or assimilation of foreign material [*Jones*, 1986, 1989]. Confirmation of Jones' interpretation came from *Jagoutz* [1989], who found a comparable crystallization age for ALHA77005 and an age of shock metamorphism of 15± 15 Ma, and from *Chen and Wasserburg* [1993], who report a single Pb-loss event at ~170 Ma for LEW88516. In the absence of serious challenges to *Jones'* [1986] arguments, it seems most likely that the S all crystallized from magmas at about 180 Ma.

Magma Composition

Crystallization sequence (silicates). The sequence in which minerals crystallized in an igneous rock is a direct consequence of the composition of the rock's parent magma. The crystallization sequence is readily determined in petrographic examination, and is particularly useful for cumulate igneous rocks, which do not have the compositions of magmas. This discussion focuses entirely on silicate minerals and ignores chromite and other oxide minerals.

The different crystallization sequences in the S and NC meteorites, summarized in Table 2, indicate that they formed from magma groups that were fundamentally different and must have originated in source areas of different mineral proportions, as first noted by *Stolper et al.* [1979]. Subsequent studies, new meteorites, and additional samples of known meteorites have validated and extended *Stolper et al.'s* [1979] inferences (Table 2), and it remains clear that N and C formed from similar magmas, all S formed from similar magmas, and that the NC and S magma groups are fundamentally different (Table 2).

Among the S, the earliest crystallizing (and most abundant) pyroxenes are low-calcium varieties, either orthopyroxene or pigeonite. Low-calcium pyroxene was the earliest mineral to crystallize in the shergottites proper: Shergotty, Zagami, and the A and B lithologies of EETA79001 [*Stolper and McSween*, 1979; *Steele and Smith*, 1982; *McSween and Jarosewich*, 1983; *McCoy et al.*, 1992; *Treiman and Sutton*, 1992]. In the other shergottite lithologies (ALHA77005, xenoliths in EETA79001; LEW88516), olivine was the first silicate mineral to crystallize, followed by low-calcium pyroxenes [*McSween et al.*, 1979; *Steele and Smith*, 1982; *McSween and Jarosewich*, 1983; *Lundberg et al.*, 1990; *Harvey et al.*, 1993; *Treiman et al.*, 1994].

Among the NC meteorites, the first pyroxene to crystallize was augite [*Bunch and Reid*, 1975; *Boctor et al.*, 1976; *Floran et al.*, 1978; *Berkley et al.*, 1980]. The position of olivine in nakhlite crystallization has been uncertain, but there is now agreement that olivine and augite crystallized together [*Harvey and McSween*, 1992a; *Treiman*, 1993b]. Low-calcium pyroxene (pigeonite) appears later in crystallization, in part as a replacement of olivine [*Berkley et al.*, 1980; *Harvey and McSween*, 1992b]. In Chassigny, olivine was the earliest silicate phase to crystallize, followed most probably by augite and pigeonite [*Floran et al.*, 1978; *Longhi and Pan*, 1989; *Johnson et al.*, 1991]. *Wadhwa and Crozaz* [1994] suggest that pigeonite crystallized before augite, but definitive textural relationships have not yet been found.

Trace elements. Although great effort has been expended in the analysis and interpretation of trace element compositions of the SNC meteorites, only the grossest comparison is adequate here: parent magmas of the S were depleted in incompatible elements (e.g., the light rare earth elements, K, U, Th etc.), and the parent magmas of the NC were enriched in those elements, as shown in Figure 1. The differences between incompatible element abundance patterns in S and NC are so great that no simple petrologic process could produce both groups at once. The differences in trace element abundances between S and NC are not required or predicted by their bulk compositional differences, and so are independent constraints on their origins.

Table 2. Crystallization Sequences For Silicate Minerals In The Martian Meteorites

Meteorite Group	Sequences of Crystallizing Minerals							
S	OL	→	OL+OPX/PIG	→	OPX/PIG	→	PIG +AUG	⇒ PIG+AUG+PL
	ALHA77005 LEW88516 EETA79001X				EETA79001A,B Shergotty Zagami			
NC	OL	→	OL+AUG	→	OL+AUG+PIG	→	AUG+PIG+PL	
	Chassigny		Nakhla Lafayette Governador Valadares					

For each meteorite group (S or NC), a generalized crystallization sequence is given left to right. Each meteorite is listed at the point where it enters the crystallization sequence. Mineral abbreviations are: OL, olivine; OPX, orthopyroxene (low Ca); PIG, pigeonite pyroxene (low Ca); AUG, augite pyroxene (high Ca); PL, plagioclase. Meteorite names are standard, except for EETA79001X, which refers to xenoliths of pyroxenite/websterite in EETA79001A.

Magma Source Regions

The source regions for basaltic magmas can be studied only indirectly, because melting and subsequent fractionation can obscure many clues about them. However, the available evidence suggests that there was no direct petrogenetic relationship between the sources for the S and NC magmas, although some indirect relationship is likely [*Jones*, 1989].

Among the least ambiguous clues to magma source regions are radio-isotope ages and initial ratios: abundance ratios of radiogenic to nonradiogenic isotopes at the time of igneous crystallization. These initial ratios are preserved through igneous fractionation, and so reflect the source(s) of the magmas. The most detailed analysis of initial isotope values in the SNCs is by *Jones* [1989], who tried to find a coherent petrogenetic scheme to relate the SNCs' initial radio-isotope ratios for ^{87}Sr (from ^{87}Rb), ^{143}Nd (from ^{147}Sm), ^{206}Pb (from ^{238}U), ^{207}Pb (from ^{235}U), and ^{208}Pb (from ^{232}Th).

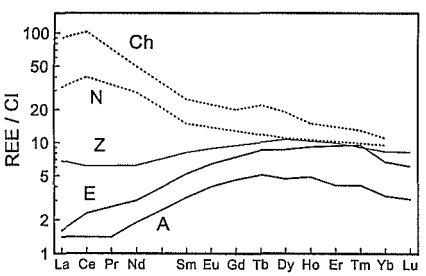


Figure 1. Rare earth element abundances in parent magmas of SNC meteorites, normalized to CI abundances. Nakhlite (N) and Chassigny (C) parent magmas from *Wadhwa and Crozaz* [1994]. "Shergottite" parent magmas are: Z, bulk Zagami [*Shih et al.*, 1982] (similar to Shergotty), which contains little cumulus pyroxene [*Treiman and Sutton*, 1992]; A, bulk ALHA77005 [*Burghele et al.*, 1983], identical to the parent magma [*Lundberg et al.*, 1990] and similar to LEW88516 [*Harvey et al.*, 1993; *Treiman et al.*, 1994].

The NC have essentially identical radiometric ages (Table 1) and initial isotope ratios [*Jones*, 1989]. Although the NC are enriched in incompatible elements (Rb, Nd, U, Th), they were derived from a source region that was depleted in these elements; at least part of the depletion must have occurred within a few tens of millions of years after planet formation [*Nyquist et al.*, 1991]. The enrichment in incompatible elements must have happened immediately before the NC parent magmas were formed (1300 m.y.), and is the only geochemical event evident in the NC source region after planet formation and initial differentiation (see references in *Jones* [1989]). Specifically, there is no evidence that the NC parent magmas assimilated any crustal components that are recognizable isotopically [*Jones*, 1989].

The S do not all have the same initial isotopic ratios; this heterogeneity has been interpreted as a mixing of the isotopic signatures of three reservoirs, one mantle and two "crustal" [*Jones*, 1989]. The mantle reservoir could reasonably have been the NC source mantle after normal isotopic evolution to ~180 Ma. This mantle must have become further depleted in Nd immediately before generation of the S parent magmas, but without depletion of other incompatible elements [*Jones*, 1989]. However, this scenario may require additional reservoirs to accommodate abundances of ^{142}Nd (from extinct ^{146}Sm) and ^{206}Pb [*Jones*, 1989; *Nyquist et al.*, 1991].

Thus, it is reasonable that both the NC and S were derived from mantle materials that were similar before 1300 m.y., but thereafter experienced divergent geochemical processing. Derivation from similar mantle materials does not necessarily imply that the NC and S formed in the same geographical spot. On Earth, mantle provinces of common differentiation age can extent for thousands of kilometers [*Bell and Blenkinsop*, 1989], and there is no a priori reason to expect that mantle provinces on Mars might be significantly smaller.

Summary

Thus, the sum of petrologic, chemical, and chronological data in Table 1 suggests that the S and NC followed entirely separate histories, from at least the inferred mantle differentiation event at 1300 Ma to their arrival on Earth. Put bluntly, the S and NC have had very little in common since 1300 Ma, except for their

shared origin on Mars. No single property listed in Table 1 would conclusively show that the S and NC formed at separate sites on Mars; a reasonable explanation or rationalization could be found for each difference. But the intersection of these many reasonable explanations is so small, so unlikely, that formation at two separate sites on Mars seems the simplest explanation of the pervasive differences between the S and the NC.

Single Crater Scenarios

In most previous studies, the SNC meteorites have been inferred to come from a single impact event on the Martian surface. I have shown above that petrologic, geochemical, and chronological evidence suggests separate source craters for S and NC. However, single crater scenarios are themselves inconsistent with available chemical and physical inferences.

Ejection of all SNC meteorites from Mars in a single impact event at ~180 Ma is the scenario most favored in the literature [*Wood and Ashwal*, 1981; *Nyquist*, 1983; *Bogard et al.*, 1984; *Vickery and Melosh*, 1987; *Mouginis-Mark et al.*, 1992]. This scenario is no longer tenable because (as discussed above) the ~180 Ma event recorded in S is not shock but igneous crystallization [*Jones*, 1986, 1989; *Jagoutz*, 1989].

Ejection of the SNCs from a single crater at 11 Ma (the cosmic ray exposure age for NC) might be possible if the crater ejection zone overlapped distinct S and NC terrains, or hit a veneer of S over NC. This scenario is not tenable either. *Vickery and Melosh* [1987] showed that it requires a crater of ≥85 km diameter on a young (~180 Ma) volcanic surface or completely obliterating a young volcanic center on an older (1300 Ma) volcanic surface. In the former case, there are no such craters on young surfaces [*Vickery and Melosh*, 1987; *Mouginis-Mark et al.*, 1992]; the latter case is completely ad hoc. In addition, ejection and subsequent orbit evolution must somehow ensure that (1) of materials exposed to cosmic rays for 11 m.y., only NC lithologies arrive at Earth, and (2) of materials exposed to cosmic rays for 2.5 m.y. or less, only S lithologies arrive at Earth. Neither scenario seems likely.

S ≠ NC: Implications of Multiple Source Craters

If the SNC meteorites left Mars in more than one impact event, at least some current understandings of Mars and its surface processes must be revised. First and foremost, accepted cratering scenarios must be revised to permit smaller impacts capable of ejecting solid material from Mars, to allow greater current (or near past) rates of crater production on Mars, or both. Second, many craters on Mars become acceptable as S or NC source sites, by eliminating the restriction that all S and NC come from the same site. And third, the lack of ancient or nonvolcanic lithologies among Martian meteorites suggests some additional problems in the recognition of meteorites, in the physical properties of Martian surface materials, or in impact mechanics.

Ejection Mechanics and Cratering Rate

The scenario favored here, where S and NC meteorites hailed from different source craters on Mars, was evaluated by *Vickery and Melosh* [1987] and rejected as improbable. Indeed, under current models of impact ejection processes and cratering rates, it is quite unlikely that Mars would have experienced an impact each into ~180 Ma (for S) and ~1300 Ma (for NC) volcanic surfaces. This deficit in SNC-ejecting craters of proper ages and sizes can be ameliorated by postulating greater cratering rates (short- or long-term), smaller permissible crater sizes, or both. From the SNCs alone, it is not possible to separate the effects of cratering rate and crater size.

Ejection of meteorites from planetary surfaces is most likely by a surface spallation mechanism [*Vickery and Melosh*, 1987]. The size of the largest ejecta fragments produced by an impact is directly dependent on the size of the impact, and thus on the size of the impact crater produced. If ejecta need be no larger than 0.5-m diameter, current theory predicts that impacts producing craters of ≥12 km diameter are adequate, provided the impactor velocity is greater than about 10 km/s [*Vickery and Melosh*, 1987]. If the ejecta need be larger (e.g., to satisfy constraints of cosmic ray exposure) larger craters are required: ejection of 3-m fragments requires a crater of ≥50 km diameter, and ejection of 15-m fragments requires a crater of ≥175 km diameter [*Vickery and Melosh*, 1987].

Current rates of crater production on Mars are poorly known, and may be significantly greater than suggested in commonly used models (Table 3). *Strom et al.* [1992] and *Mouginis-Mark et al.* [1992] accept *Neukum and Hiller's* [1981] second model (NH-2 in Table 3) of relatively low crater production rates, which are comparable to the nominal rate for the Moon over the last 3 Gyr. An updated version of that model is given by *Neukum* [1983], cited below as N-83. *Tanaka et al.* [1988, 1992] accept *Hartmann et al.'s* [1981] preferred cratering rate model (HT in Table 3), with approximately twice Moon's nominal crater production rate. *Hartmann et al.* [1981] recognized the possibility of greater cratering rates, up to twice their preferred rate. Recent analyses of crater populations on Earth and of asteroid and comet populations near the Earth suggest that cratering rates over the last 360 m.y. (at least) are 2-3 times the nominal average lunar rate [*Grieve*, 1984; *Shoemaker et al.*, 1990]. Further, the population of Earth-approaching asteroids and the abundance and sources of meteorites from the Moon suggest even higher rates of crater production [*Rabinowitz*, 1993; *Warren*, 1994a,b]. The 3xHT model of Tables 3-6 accommodates these inferred higher rates by assuming a crater production rate on Mars of 3 times *Hartmann et al.'s* [1981] preferred model.

To evaluate the consistency of these models of meteorite ejection and cratering rates, Table 3 shows predicted cratering rates and numbers of craters produced on Martian surfaces of various ages over the last 20 m.y., approximately twice the cosmic ray exposure ages of the NC. There is no a priori reason to begin counting meteorite-forming events at 11 m.y., to just cover and include the NC ejection event (Table 1); ejecta-forming events on Mars at 20 m.y. could yield meteorites on Earth with only slightly lower probability than events at 11 m.y. (Figure 1 of *Wetherill* [1984]). Tables 4 and 5 assume that the given crater production models were followed over the past 250 m.y. and 1800 m.y., because the choice of crater production model influences the absolute ages of Martian surfaces inferred from crater counts. These time intervals were chosen to include the crystallization ages of the S and NC, respectively (180 Ma and 1300 Ma, Table 1); as before, there are no a priori reasons for limiting these time intervals to be the crystallization ages.

The assumed cratering rate has a large, compounded effect on the calculated number of potential S or NC source craters because increased crater production decreases the estimated absolute ages of surfaces based on crater count statistics. Thus, increased cratering rates yield more craters per unit time, and more surface area younger than a given age. The effects on this compounding effect is can be seen by comparing the HT and 3xHT models in Table 4; a tripling of the cratering rate yields

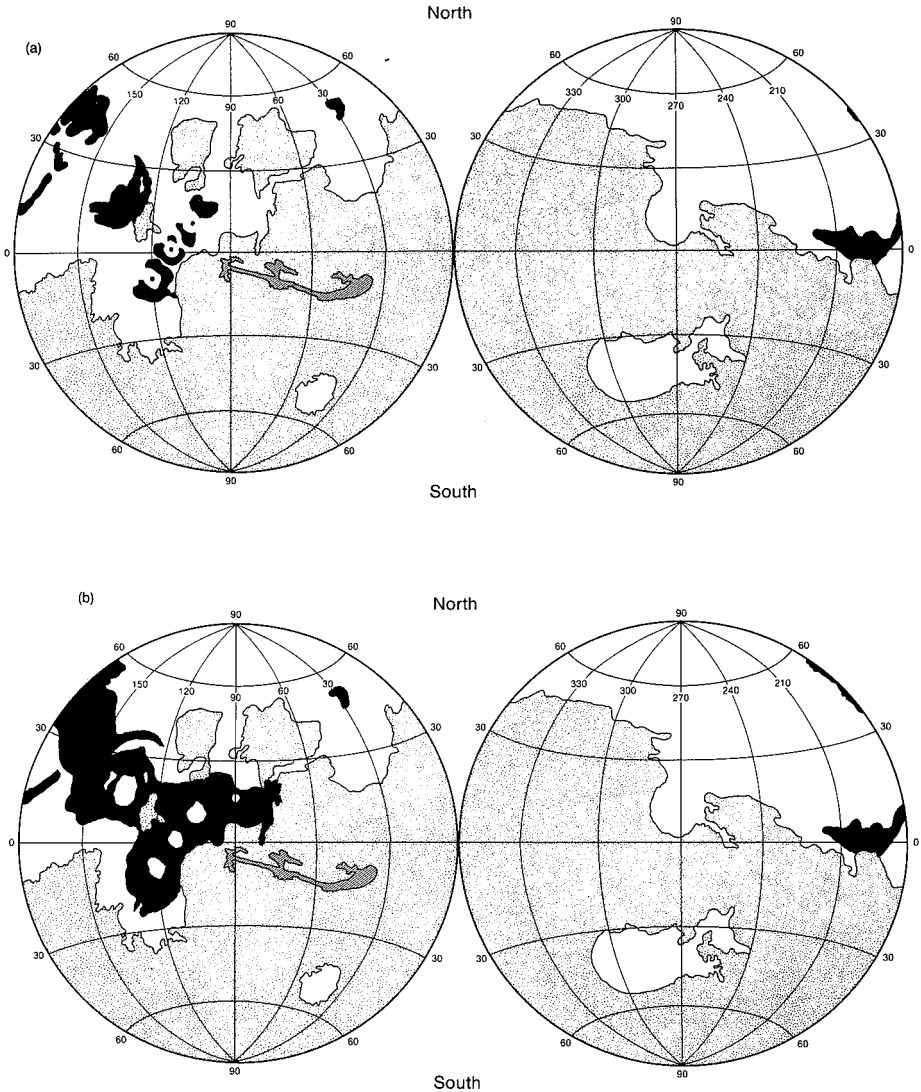


Figure 2. Potential Martian source areas for shergottite (S) meteorites, shown in black on map of Mars. Stippled areas are heavily cratered highlands, diagonal stripes show Valles Marineris. a) The HT crater production model (Table 3), potential sources restricted to basaltic units of upper Amazonian age. Includes units Aop, Aos, At6, Aa5, Aa4, and Achu of *Scott et al.* [1987] and *Tanaka et al.* [1988]. b) The 3xHT crater production model, potential sources restricted to basaltic units of upper and middle Amazonian ages. Includes units in Figure 2a and Aa3, Aa2, Aoa4, Aoa3, Aoa2, and At5 of *Scott et al.* [1987] and *Tanaka et al.* [1988].

approximately 10 times the number of likely source craters. As shown in Figure 2, the HT model implies that only surfaces of late Amazonian age (by crater densities) are <250 m.y. old [*Tanaka et al.*, 1992]; volcanic surfaces of this age are present only on a portion of the Tharsis plateau (Olympus Mons itself, and some flow fields surrounding the Olympus Mons and the Tharsis Montes) and as scattered patches in Acidalia and Arcadia Planitia [*Scott et al.*, 1987]. On the other hand, the 3xHT model implies that surfaces of middle and late Amazonian are <250 m.y. old; in addition to the above late Amazonian volcanic surfaces, one must include most of the remainder of Tharsis, and most of Amazonis and Arcadia Planitia (Figure 2) [*Scott et al.*, 1987].

It is also assumed here that there is a high probability that at least some fragments from each ejection impact are transported to Earth, and thus may appear in the meteorite record [*Wetherill*, 1984; *Melosh and Tonks*, 1993].

Craters: 1 S, 1 NC. A scenario in which S and NC meteorites hailed from different source craters on Mars was considered explicitly by *Vickery and Melosh* [1987] and rejected because it would require a crater ≥12 km diameter for NC (ejecta of <0.5 m diameter) and a crater ≥50 km diameter for S (ejectum of ≥3 m diameter). The latter is most unlikely (Table 4) and in fact no craters of this size are present on likely surfaces on Mars [*Mouginis-Mark et al.*, 1992]. Based on Tables 4 and 5, the three crater production models predict in the last 20 m.y. the following numbers of potential S and NC source craters: model NH-2, 0.001 and 0.13; model N-83, 0.02 and 0.4; and HT 0.03 and 3.2. None of these models predict a high likelihood of separate craters for S and NC meteorites.

But production rates for SNC-ejecting craters are strongly dependent on the details of the physical models involved. For instance, if one accepted a high crater production rate of 3xHT, one would expect production of 0.3 such ≥50 km S source craters and 2.2 such ≥12 km NC source craters in 20 m.y. (Tables 4, 5; Figure 2), within the range of plausibility. Similarly, if one accepts HT crater production and a minimum crater diameter to eject the S of ≥16 km, one would expect production of 0.31 S source craters in 20 m.y. (Table 4), again within the range of plausibility. At this point, it seems more reasonable to accept higher cratering rates, like 3xHT, given evidence for comparably greater cratering rates on the Earth and Moon [*Grieve*, 1984; *Shoemaker et al.*, 1990; *Warren*, 1994a,b]. If this 3xHT crater production rate were coupled with ejection of S from craters of ≥25 km diameter, half the nominal size limit of *Vickery and Melosh* [1987], one would expect to find a likely source crater for S (Table 4).

Vickery and Melosh [1987] faulted this scenario for failing to account for the absence of meteorites from ancient and non-volcanic terrains of Mars. This point is significant and is considered below, but does not constitute a death-blow to the scenario.

Craters: 2 S, 1 NC. A related scenario for SNC origins holds that each exposure age group (Table 1) represents a separate impact event on Mars, Case 4 of *Vickery and Melosh* [1987], and recommended by *Jones* [1989]. This scenario is attractive in that the EETA79001 shergottite, which has a cosmic ray exposure age of 0.5 Ma, is in some ways distinct from the other S. EETA79001 is the only S meteorite that has an igneous contact, contains xenoliths, contains unadulterated Martian atmosphere, contains aluminosilicates among its pre-terrestrial alteration materials, contains the high-pressure shock minerals ringwoodite and majorite, and contains strong evidence for a second assimilated crustal component [*Bogard et al.*, 1984; *McSween*, 1985; *Gooding and Muenow*, 1986; *Jones*, 1989]. On the other hand, EETA79001 has essentially the same crystallization age, mineral chemistry and magma chemistry (bulk and trace element) and petrography as the other S, and its xenoliths are quite similar to the ALHA77005 and LEW88516 shergottites [*McSween*, 1985; *Treiman*, 1993b; *Wadhwa et al.*, 1994].

This scenario is attractive in that it requires no modifications

Table 3. Expected Production of Impact Craters on Mars, Last 20 m.y.: Craters Larger than Given Diameter per 10^6 km^2

Crater	Crater Production Model			
Diameter km	NH-2	N-83	HT	3xHT
4	0.15	0.22	1.1	3.4
8	0.04	0.09	0.3	0.9
12	0.016	0.05	0.1	0.4
16	0.009	0.03	0.08	0.2
25	(0.003)	0.015	0.03	0.09
32	(0.0016)	0.009	0.02	0.06
50	(0.0005)	0.003	0.008	0.02
64	(0.0003)	0.0014	0.005	0.014

NH-2: Model 2 of *Neukum and Hiller* [1981], calibrated for craters smaller than 20 km diameter, values in parentheses may not be accurate. N-83: From *Neukum* [1983]. HT: Hartmann-Tanaka model: *Hartmann et al.* [1981] and *Tanaka et al.* [1988]. 3xHT: Cratering rate 3 times HT.

Table 4. Expected Production of Impact Craters on Mars, Last 20 m.y.: Craters Larger than Given Diameter on Volcanic Surfaces Aged <250 Ma

Crater	Crater Production Model			
Diameter km	NH-2*	N-83	HT†	3xHT‡
4	0.30	0.44	4.6	40
8	0.08	0.18	1.2	10.
12	0.03	0.11	0.5	4.7
16	0.018	0.07	0.3	2.7
25	(0.006)	0.03	0.13	1.1
32	(0.003)	0.018	0.08	0.7
50	(0.001)	0.006	0.03	0.3
64	(0.0005)	0.003	0.02	0.2

See Table 3 for crater production models; values in parentheses may not be accurate. Areas of volcanically resurfaced terrains drawn and calculated from *Tanaka et al.* [1988], including also unit Achu of lava flows covered with a thin veneer of channel sediments (K. Tanaka, personal communication, 1994).

* Half of upper Amazonian: 2.0 x 10^6 km^2.

† Upper Amazonian: 4.1 x 10^6 km^2.

‡ Middle and upper Amazonian: 11.8 x 10^6 km^2.

to impact ejection theory. *Vickery and Melosh* [1987] calculated that the ejecta from each impact event need be no bigger than 0.5-m diameter and the ejecting craters need only be larger than 12-km diameter [*Vickery and Melosh*, 1987]. Production of so many suitable craters in the last 20 m.y. is likely given a 3xHT crater production model, and barely tenable under the HT model, two craters from an expected 0.5 (Tables 4, 5). However, the scenario does require deriving all the petrologically and chronologically similar S from multiple craters, and may be inconsistent with the total mass of Mars material transferred to Earth [*Vickery and Melosh*, 1987]. It also fails to accord with the absence of meteorites from ancient and nonvolcanic terrains of Mars, but this point is considered below. I do not feel this scenario is probable, but find no cause to reject it outright.

Source Craters

It is not within the scope of this paper to exhaustively review potential source craters for the S and NC meteorites, especially in light of *Mouginis-Mark et al.'s* [1992] compilation of potential source craters. *Mouginis-Mark et al.* [1992] accepted the scenario that all of the SNCs came from a single source crater, and thus tried to find craters on surfaces of ~180 Ma (the S) that could potentially have had access to 1300 Ma basaltic rocks (the NC). Their search included craters of ≥10 km diameter, recognizing that there were no craters of ≥50 km in the youngest (late Amazonian) volcanic terrains, such as are required in the models of *Vickery and Melosh* [1987]. Of the 25 candidate craters on Tharsis, none satisfied all constraints, particularly that it have access to S and NC lithologies, so *Mouginis-Mark et al.* [1992] listed their nine most likely candidates, with recommending and detracting features for each.

However, if the S and NC meteorites did not come from a single crater, all of the most likely craters in *Mouginis-Mark et al.* [1992] become potential sources for one or the other meteorite group. Based on ages of surfaces into which they were emplaced, one might suggest that their craters 1, 3, or 7-9 might be the source for S (7 is most likely because it is largest), and craters 2 and 4-9 might be the source for NC [*Mouginis-Mark et al.*, 1992].

Table 5. Expected Production of Impact Craters on Mars, Last 20 m.y.: Craters Larger than Given Diameter on Volcanic Surfaces Aged <1800 Ma

Crater	Crater Production Model			
Diameter km	NH-2§	N-83§	HT¶	3 x HT**
4	1.1	1.7	27	186
8	0.3	0.7	7.0	48
12	0.13	0.4	3.2	22
16	0.07	0.3	1.81	12.4
25	(0.02)	0.12	0.76	5.2
32	(0.01)	0.07	0.47	3.2
50	(0.004)	0.02	0.20	1.3
64	(0.002)	0.01	0.12	0.8

See Table 3 for crater production models, values in parentheses may not be accurate. Areas of volcanically resurfaced terrains drawn and calculated from *Tanaka et al.* [1988], including also unit Achu of lava flows covered with a thin veneer of channel sediments (K. Tanaka, personal communication, 1994).

§ Upper and half of middle Amazonian: 7.7×10^6 km^2.

¶ Amazonian: 24.1×10^6 km^2.

** Hesperian and Amazonian: 54.9×10^6 km^2.

Other Lithologies and Ages?

If the S and NC originated from different craters, ejection from Mars cannot have been caused by a unique impact, and similar impacts should have occurred on other surfaces on Mars. Where are the meteorites from these Martian surfaces? This question led *Vickery and Melosh* [1987], among others, to reject scenarios with multiple SNC source craters. However, given the evidence of Table 1, this question remains as a challenge to planetary geologists, impact dynamists, and meteoriticists. It should be noted that lunar meteorites, 10 total now (excluding paired samples), appear to be a representative sampling (by exposed area) of the major lunar surface units: highlands, mare basalt, KREEP-rich rock [*Lindstrom et al.*, 1994; *Warren*, 1994a,b].

To quantify the number of potential meteorite-ejecting impacts on Mars, one can compare the values of Tables 5 and 6, which show the numbers of craters of given sizes on volcanic surfaces ≤1800 Ma and on the whole planet. The HT crater production model implies ~6 times as many meteorite-producing craters over the whole planet as on volcanic surfaces ≤1800 Ma. Naively, one might expect approximately five times as many ancient and nonvolcanic Martian meteorites as young volcanic Martian meteorites. The 3xHT crater production model implies ~3 times as many meteorite-producing craters over the whole planet as in the ≤1800 Ma volcanic surfaces. Again naively, one might expect twice as many ancient and nonvolcanic Martian meteorites as young volcanic Martian meteorites.

The lack of Martian meteorites from ancient and nonvolcanic terrains admits many possible explanations. First, older and nonvolcanic surfaces might not contain coherent enough material to be ejected as 0.5-m or larger fragments [*Jones*, 1989]. Sedimentary, especially aeolian, surfaces are likely to be weak. Older volcanic surfaces are, however, interpreted to be coherent and strong [*Tanaka and Golombek*, 1989; *Mouginis-Mark et al.*, 1992]. Second, the physical properties of older surface materials may preclude ejection of meteorite-sized fragments. Calculations of impact ejection by spallation [*Vickery and Melosh*, 1987] considered ejection only from coherent rock surfaces. Spallation as a meteorite ejection mechanism is likely to be less effective from a weak (low tensile strength) surface (equation 1 of *Vickery and Melosh* [1987]; *Jones* [1989]), e.g., loosely packed,

Table 6. Expected Production of Impact Craters on Mars in Last 20 m.y.: Craters Larger than Given Diameter on Whole Planet

Crater	Crater Production Model			
Diameter km	NH-2*	N-83	HT	3 x HT
4	21	32	160	490
8	5.5	13	40	130
12	2.4	7.6	20	60
16	1.3	5.	11	33
25	(0.44)	2.2	4.6	14
32	(0.23)	1.3	2.8	8.4
50	(0.07)	0.4	1.2	3.5
64	(0.04)	0.2	0.7	2.2

See Table 3 for crater production models.

heavily fractured, brecciated, or weathered surface material, as some of the shock energy will be expended compacting and deforming the material. Perhaps rocks of the Hesperian-age ridged plains, the most extensive volcanic unit on Mars, are too broken or weathered for impact ejection (K. Tanaka, personal communication, 1994). And third, perhaps materials from the older surfaces do fall to Earth and are not recognized as Martian or meteoritic [*Lindstrom et al.*, 1994]. The ALH84001 meteorite, of Martian origin, is an orthopyroxenite, and was misclassified and studied for eight years as a common diogenite (*Mittlefehldt* [1994], see below). Given the varied geology of Mars, one might reasonably expect Martian meteorites to include altered basaltic rocks, siltstone, sandstone, and conglomerate, and possibly to include rhyolite, granite, or anorthosite. But no such meteorites have been reported, notably from the blue ice areas of Antarctica, where terrestrial rocks are rare to absent (R.P. Harvey, personal communication, 1994). However, it is not clear that the scientific community would seriously consider a reported "sandstone from the sky" as possibly being a legitimate meteorite.

And ALH84001?

The ALH84001 meteorite, long thought to be a diogenite, has recently been reclassified as a unique Martian meteorite, not belonging to the S, N, or C groups [*Mittlefehldt*, 1994]. ALH84001 is an orthopyroxenite, ruling out any affinity to NC by bulk magma composition and crystallization sequence arguments (Table 1 and above). Also unlike the NC, ALH84001 is strongly shocked (plagioclase converted to maskelynite). Affinities to the S, though stronger, are also questionable. By mineralogy, ALH84001 could be a cumulate member of S, similar to the xenoliths in EETA79001 (Table 1) [*McSween and Jarosewich*, 1983]. However, the parent magma of ALH84001 was enriched in incompatible elements [*Mittlefehldt*, 1994], while S are depleted. In addition, ALH84001 contains abundant products of aqueous alteration, dominantly Fe-Mg carbonates, while the S contain only rare products of alteration materials, dominantly Ca-bearing carbonates and sulfates [*Gooding*, 1992, and references therein]. And unlike any SNC, ALH84001 appears to show evidence for multiple shock events: a late event that yielded maskelynite after plagioclase and caused cracking and faulting; and an early event responsible for pervasive crush zones (fault gouge), subsequently annealed and recrystallized.

Although data on ALH84001 are limited and preliminary, it seems likely that it is much older than the other Martian meteorites. A great age is suggested by the presence of recrystallized crush zones [*Mittlefehldt*, 1994; *Treiman*, 1994b], which appear to record an early impact brecciation followed by annealing at high temperatures. This event was probably not the impact that lofted ALH84001 from Mars; for the other SNCs, the ejecting impact events did not cause recrystallization, even of maskelynite [*Duke*, 1968], suggesting a very low time-temperature history. ALH84001 can best be compared with lunar "recrystallized cataclastic rocks" of highland derivation [*Stöffler et al.*, 1980]. Following this analogy, ALH84001 could possibly be from the Martian highlands, a product of Martian magmatism more than 3 Gyr ago.

Acknowledgments. I am greatly indebted to J. Jones for his papers on the SNCs and for many fruitful discussions. Additional discussions with W. Kiefer, A. Vickery, M.M. Lindstrom, D. Mittlefehldt, P. Warren, and R. Herrick (and many others) have helped clarify my thoughts. G. McGill assisted with access to some cratering literature. Abstracts of this work were originally presented at the MSATT Final Conference in November, 1993 and the XXV Lunar and Planetary Science Conference. Reviews by N. Barlow, J. Jones, G. Ryder, K. Tanaka, and H. McSween Jr. were quite helpful. Lunar and Planetary Institute contribution 838.

References

Bell, K., and J. Blenkinsop, Neodymium and strontium isotope chemistry of carbonatites, in *Carbonatites: Genesis and Evolution*, edited by K. Bell, pp. 278-300, Hyman Unwin, Winchester, Mass., 1989.

Berkley, J.L., K. Keil, and M. Prinz, Comparative petrology and origin of Governador Valadares and other nakhlites, *Proc. Lunar Planet. Sci. Conf., 11th*, 1089-1102, 1980.

Boctor, N.Z., H.O.A. Meyer, and G. Kullerud, Lafayette meteorite: Petrology and opaque mineralogy, *Earth Planet. Sci. Lett., 32*, 69-76, 1976.

Bogard, D.D., L.E. Nyquist, and P. Johnson, Noble gas contents of shergottites and implications for the Martian origin of SNC meteorites, *Geochim. Cosmochim. Acta, 48*, 1723-1739, 1984.

Boynton, W.V., P.M. Starzyk, and R.A. Schmitt, Chemical evidence for the genesis of the ureilites, the achondrite Chassigny, and the nakhlites, *Geochim. Cosmochim. Acta, 40*, 1439-1447, 1976.

Bunch, T.E., and A.M. Reid, The nakhlites, part I, Petrography and mineral chemistry, *Meteoritics, 10*, 303-315, 1975.

Burghele, A., G. Dreibus, H. Palme, W. Rammensee, B. Spettel, G. Weckwerth, and H. Wänke, Chemistry of shergottites and the Shergotty parent body (SPB): Further evidence for the two-component model of planet formation (abstract), *Lunar Planet. Sci., XIV*, 80-81, 1983.

Chen, J.H., and G.J. Wasserburg, LEW88516 and SNC meteorites (abstract), *Lunar Planet. Sci., XXIV*, 113-114, 1993.

Clayton, R.N., and T.K. Mayeda, Oxygen isotopes in eucrites, shergottites, nakhlites, and Chassigny, *Earth Planet. Sci. Lett., 62*, 1-6, 1983.

Drake, M.J., T. Owen, T. Swindle, and D. Musselwhite, Noble gas evidence of an aqueous reservoir near the surface of Mars more recently than 1.3 Ga (abstract), *Lunar Planet. Sci., XXIV*, 431-434, 1993.

Duke, M.B., The Shergotty meteorite: Magmatic and shock metamorphic features, in *Shock Metamorphism of Natural Materials*, edited by B.M. French and N.M. Short, pp. 613-621, Mono Books, San Francisco, 1968.

Floran, R.J., M. Prinz, P.F. Hlava, K. Keil, C.E. Nehru, and J.R. Hinthorne, The Chassigny meteorite: A cumulate dunite with hydrous amphibole-bearing melt inclusions, *Geochim. Cosmochim. Acta, 42*, 1213-1229, 1978.

Freundel, M., L. Schulz, and R.C. Reedy, Terrestrial ^{81}Kr-^{83}Kr ages of Antarctic meteorites, *Geochim. Cosmochim. Acta, 50*, 2663-2673, 1986.

Gooding, J.L., Soil mineralogy and chemistry on Mars: Possible clues from salts and clays in SNC meteorites, *Icarus, 99*, 28-41, 1992.

Gooding, J.L., and D.W. Muenow, Martian volatiles in shergottite EETA79001: New evidence from oxidized sulfur and sulfur-rich aluminosilicates, *Geochim. Cosmochim. Acta, 50*, 1049-1059, 1986.

Gooding, J.L., S.J. Wentworth, and M.E. Zolensky, Calcium carbonate and sulfate of possible extraterrestrial origin in the EETA79001 meteorite, *Geochim. Cosmochim. Acta, 52*, 909-915, 1988.

Gooding, J.L., S.J. Wentworth, and M.E. Zolensky, Aqueous alteration of the Nakhla meteorite, *Meteoritics, 26*, 135-143, 1991.

Graham, A.L., A.W.R. Bevan, and R. Hutchison, *Catalog of Meteorites, 4th Edition*, University of Arizona Press, Tucson, 1985.

Grieve, R.A.F., The impact cratering record in recent time, *Proc. Lunar Planet. Sci. Conf., 14th, Part 2, J. Geophys. Res., 89*, suppl., B403-B408, 1984.

Hartmann, W.K., et al., Chronology of planetary volcanism by

comparative studies of planetary cratering, chapter 8 in *Basaltic Volcanism on the Terrestrial Planets*, edited by Basaltic Volcanism Study Project, pp. 1049-1128, Pergamon, New York, 1981.

Harvey, R.P., and H.Y. McSween Jr., The parent magma of the nakhlite meteorites: Clues from melt inclusions, *Earth Planet. Sci. Lett., 111*, 467-482, 1992a.

Harvey, R.P., and H.Y. McSween Jr., Petrogenesis of the nakhlite meteorites: Evidence from cumulate mineral zoning, *Geochim. Cosmochim. Acta, 56*, 1655-1663, 1992b.

Harvey, R.P., M. Wadhwa, H.Y. McSween Jr., and G. Crozaz, Petrography, mineral chemistry, and petrogenesis of Antarctic shergottite LEW88516, *Geochim. Cosmochim. Acta, 57*, 4769-4784, 1993.

Jagoutz, E., Sr and Nd isotope systematics in ALHA77005: Age of shock metamorphism in shergottites and magmatic differentiation of Mars, *Geochim. Cosmochim. Acta, 53*, 2429-2441, 1989.

Johnson, M.C., M.J. Rutherford, and P.C. Hess, Chassigny petrogenesis: Melt compositions, intensive parameters, and water contents of Martian (?) magmas, *Geochim. Cosmochim. Acta, 55*, 349-366, 1991.

Jones, J.H., A discussion of isotopic systematics and mineral zoning in the shergottites: Evidence for a 180 m.y. igneous crystallization age, *Geochim. Cosmochim. Acta, 50*, 969-977, 1986.

Jones, J.H., Isotopic relationships among the shergottites, nakhlites and Chassigny, *Proc. Lunar Planet. Sci. Conf., 19th*, 465-474, 1989.

Jull, A.J.T., and D.J. Donahue, Terrestrial ^{14}C age of the Antarctic shergottite EETA79001, *Geochim. Cosmochim. Acta, 52*, 1309-1311, 1988.

Karlsson, H.R., R.N. Clayton, E.K. Gibson Jr., and T.K. Mayeda, Water in SNC meteorites: Evidence for a Martian hydrosphere, *Science, 255*, 1409-1411, 1992.

Lindstrom, M.M., A.H. Treiman, and D.W. Mittlefehldt, Pigeonholing planetary meteorites: The lessons of misclassification of EET87521 and ALH84001 (abstract), *Lunar Planet. Sci., XXV*, 797-798, 1994.

Longhi, J., and V. Pan, The parent magmas of the SNC meteorites, *Proc. Lunar Planet. Sci. Conf., 19th*, 451-464, 1989.

Lundberg, L.L., G. Crozaz, and H.Y. McSween Jr., Rare earth elements in minerals of the ALHA77005 shergottite and implications for its parent magma and crystallization history, *Geochim. Cosmochim. Acta, 54*, 2535-2547, 1990.

McCoy, T.J., G.J. Taylor, and K. Keil, Zagami: Product of a two-stage magmatic history,. *Geochim. Cosmochim. Acta, 56*, 3571-3582, 1992.

McSween, H.Y. Jr., SNC meteorites: Clues to Martian petrologic evolution?, *Rev. Geophys., 23*, 391-416, 1985.

McSween, H.Y. Jr., and E. Jarosewich, Petrogenesis of the Elephant Moraine A79001 meteorite: Multiple magma pulses on the shergottite parent planet, *Geochim. Cosmochim. Acta, 47*, 1501-1513, 1983.

McSween, H.Y. Jr., L.A. Taylor, and E.M. Stolper, Allan Hills 77005: A new meteorite type found in Antarctica, *Science, 204*, 1201-1203, 1979.

Melosh, H.J., and W.B. Tonks, Swapping rocks: Ejection and exchange of surface materials among the terrestrial planets (abstract), *Meteoritics, 28*, 398, 1993.

Melosh, H.J., A.H. Treiman, and R.A.F. Grieve, Olivine composition glass in the Chassigny meteorite: Implications for shock history (abstract), *EOS Trans. A.G.U., 64*, 254, 1983

Mittlefehldt, D.W., ALH84001, a cumulate orthopyroxenite member of the SNC meteorite group, *Meteoritics, 29*, 214-221, 1994.

Mouginis-Mark, P.J., T.J. McCoy, G.J. Taylor, and K. Keil, Martian parent craters for the SNC meteorites, *J. Geophys. Res., 97*, 10,213-10,336, 1992.

Neukum, G., *Meteoritenbombardment und Datierung Planetarer Oberflachen, Habilitationsschrift*, 186 pp., Ludwig-Maximilian-Universität, Munich, 1983.

Neukum, G., and K. Hiller, Martian ages, *J. Geophys. Res., 86*, 3097-3121, 1981.

Nishiizumi, K., J. Klein, R. Middleton, D. Elmore, P.W. Kubik, and J.R. Arnold, Exposure history of shergottites, *Geochim. Cosmochim. Acta, 50*, 1017-1021, 1986.

Nishiizumi, K, D. Elmore, and P.W. Kubik, Update on terrestrial ages of Antarctic meteorites, *Earth Planet. Sci. Lett., 93*, 299-313, 1989.

Nishiizumi, K., J.R. Arnold, M.W. Caffee, R.C. Finkel, and J. Southon, Exposure histories of Calcalong Creek and LEW 88516 meteorites (abstract), *Meteoritics, 27*, 270, 1992.

Nyquist, L.E., Do oblique impacts produce Martian meteorites?, *Proc. Lunar Planet. Sci. Conf., 13th, Part 2, J. Geophys. Res., 88*, suppl., A785-A798, 1983.

Nyquist, L.E., C.L. Harper, H. Wiesmann, B. Bansal, and C.-Y. Shih, $^{142}Nd/^{144}Nd$ in SNCs and lunar samples: Implications for early differentiation of a heterogeneous Martian (?) mantle (abstract), *Meteoritics, 26*, 381, 1991.

Ott, U., Noble gases in SNC meteorites: Shergotty, Nakhla, Chassigny, *Geochim. Cosmochim. Acta, 52*, 1937-1948, 1988.

Ott, U., and F. Begemann, Are all the 'Martian' meteorites from Mars?, *Nature, 317*, 509-512, 1985.

Pepin, R.O., Atmospheric compositions: Key similarities and differences, in *Origin and Evolution of Planetary and Satellite Atmospheres*, edited by S.K. Atreya, J.B. Pollack, and M.S. Matthews. pp. 291-305, University of Arizona Press, Tucson, 1989.

Rabinowitz, D.L., The size distribution of the Earth-approaching asteroids, *Astrophys. J., 407*, 412-427, 1993.

Scott, D.H., K.L. Tanaka, R. Greeley, and J.E. Guest, Geologic maps of the western equatorial, eastern equatorial and polar regions on Mars, *U.S. Geol. Surv. Maps I-1802-A,B,C*, 1987.

Shih, C.-Y., L.E. Nyquist, D.D. Bogard, G.A. McKay, J.L. Wooden, B.M. Bansal, and H. Wiesmann, Chronology and petrogenesis of young achondrites, Shergotty, Zagami, and ALHA77005: Late magmatism on a geologically active planet, *Geochim. Cosmochim. Acta, 46*, 2323-2344, 1982.

Shoemaker, E.M., R.F. Wolfe, and C.S. Shoemaker, Asteroid and comet flux in the neighborhood of the Earth., in Global Catastrophes in Earth History: An Interdisciplinary Conference on Impacts, Volcanism, and Mass Mortality, edited by V.L. Sharpton and P.D. Ward, *Geol. Soc. Amer., Spec. Pub. 247*, 155-170, 1990.

Smith, J.V., and I.M. Steele, Achondrite ALHA77005: Alteration of chromite and olivine, *Meteoritics, 19*, 121-133, 1984.

Steele, I.M., and J.V. Smith, Petrography and mineralogy of two basalts and olivine-pyroxene-spinel fragments in achondrite EETA79001, *Proc. Lunar Planet. Sci. Conf. 13th, Part 1, J. Geophys. Res., 87*, suppl., A375-A384, 1982.

Stöffler, D., H.-D. Knöll, U.B. Marvin, C.H. Simonds, and P.H. Warren, Recommended classification and nomenclature of lunar highland rocks – A committee report. in *Proceedings of the Conference on the Lunar Highlands Crust*, edited by J.J. Papike and R.B. Merrill, pp. 51-70, Pergamon, New York, 1980.

Stöffler, D., K. Keil, and E.R.D. Scott, Shock metamorphism of ordinary chondrites, *Geochim. Cosmochim. Acta, 55*, 3845-3867, 1991.

Stolper, E., and H.Y. McSween Jr., Petrology and origin of the shergottite meteorites, *Geochim. Cosmochim. Acta, 43*, 1475-1498, 1979.

Stolper, E.M., H.Y. McSween Jr., and J.F. Hays, A petrogenetic model of the relationships among achondritic meteorites, *Geochim. Cosmochim. Acta, 43*, 589-602, 1979.

Strom, R.G., S.K. Croft, and N.G. Barlow, The Martian impact cratering record, in *Mars*, edited by H.H. Kieffer, B.M. Jakosky, C.W. Snyder, and M.S. Matthews. pp. 383-423, University of Arizona Press, Tucson, 1992.

Tanaka, K.L. and M. Golombek, Martian tension fractures and the formation of grabens and collapse features at Valles Marineris, *Proc. Lunar Planet. Sci. Conf., 19th*, 383-396, 1989.

Tanaka, K.L., N.K. Isbell, D.H. Scott, R. Greeley, and J.E. Guest, The resurfacing history of Mars: A synthesis of digitized, Viking-based geology, *Proc. Lunar Planet. Sci. Conf., 18th*, 665-678, 1988.

Tanaka, K.L., D.H. Scott, and R. Greeley, Global stratigraphy, in

Mars, edited by H.H. Kieffer, B.M. Jakosky, C.W. Snyder, and M.S. Matthews, pp. 345-382, University of Arizona Press, Tucson, 1992.

Treiman, A.H., Amphibole and hercynite spinel in Shergotty and Zagami: Magmatic water, depth of crystallization, and metasomatism, *Meteoritics, 20*, 229-243, 1985.

Treiman, A.H., The Martian sources of the SNC meteorites (two, not one), and what can and can't be learned from the SNC meteorites (abstract). *Mars: Past, Present, and Future -- Results from the MSATT Program*, pp. 49-51, Lunar and Planetary Institute, Houston, 1993a.

Treiman, A.H., The parent magma of the Nakhla (SNC) meteorite, inferred from magmatic inclusions, *Geochim. Cosmochim. Acta, 57*, 4753-4768, 1993b.

Treiman, A.H., Xenoliths in the EETA79001 shergottite: Geological and astronomical implications of similarities to the ALHA77005 and LEW88516 shergottites (abstract), *Meteoritics, 28*, 451, 1993c.

Treiman, A.H., Two source areas for the SNC meteorites: Petrologic, chemical, and chronologic evidence (abstract), *Lunar Planet. Sci., XXV*, 1413-1414, 1994a.

Treiman, A.H., An ancient age for ALH84001: Petrographic evidence for multiple shock events (abstract), *Meteoritics, 29*, 542, 1994b.

Treiman, A.H., and S.R. Sutton, Petrogenesis of the Zagami meteorites: Inferences from synchrotron X-ray (SXRF) microprobe analyses of pyroxenes, *Geochim. Cosmochim. Acta, 56*, 4059-4074, 1992.

Treiman, A.H., J.L. Gooding, and R.A. Barrett, Preterrestrial aqueous alteration of the Lafayette (SNC) meteorite, *Meteoritics, 28*, 86-97, 1993.

Treiman, A.H., G.A. McKay, D.D. Bogard, M.-S. Wang, M.E. Lipschutz, D.W. Mittlefehldt, L. Keller, M.M. Lindstrom and D. Garrison, Comparison of the LEW88516 and ALHA77005 Martian meteorites: Similar but distinct, *Meteoritics, 29*, 581-592, 1994.

Vickery, A.M., and H.J. Melosh, The large crater origin of SNC meteorites, *Science, 237*, 738-743, 1987.

Wadhwa, M., and G. Crozaz, Trace and minor elements in minerals of nakhlites and Chassigny: Clues to their petrogenesis, *Geochim. Cosmochim. Acta, 58*, in press, 1994.

Wadhwa, M., H.Y. McSween Jr., and G. Crozaz, Petrogenesis of the shergottite meteorites inferred from minor and trace element microdistributions, *Geochim. Cosmochim. Acta, 58*, in press, 1994.

Warren, P.H., Petrologic insights regarding lunar and planetary meteorite launch processes (abstract), *Lunar Planet. Sci., XXV*, 1463-1464, 1994a.

Warren, P.H., Lunar and Martian meteorite delivery services, *Icarus, 110*, in press, 1994b.

Watson, L.L., I.D. Hutcheon, S. Epstein, and E.M. Stolper, D/H ratios and water contents of amphiboles in Chassigny and Shergotty (abstract), *Meteoritics*, 28, 456-457, 1993.

Watson, L.L., S. Epstein, and E.M. Stolper, The abundance and stable isotopic composition of volatiles released from weathering products during stepped heating of Nakhla and Lafayette (abstract), *Lunar Planet. Sci., XXV*, 1471-1472, 1994.

Wentworth, S.J., and J.L. Gooding, Weathering features and secondary minerals in Antarctic shergottites ALHA77005 and LEW88516 (abstract), *Lunar Planet. Sci., XXIV*, 1507-1508, 1993.

Wetherill, G.W., Orbital evolution of impact ejecta from Mars, *Meteoritics, 19*, 1-13, 1984.

Wiens, R.C., and R.O. Pepin, Laboratory shock emplacement of noble gases, nitrogen, and carbon dioxide into basalt, and implications for trapped gases in shergottite EETA79001, *Geochim. Cosmochim. Acta, 52*, 295-308, 1988.

Wood, C.A, and L.D. Ashwal, SNC meteorites: Igneous rocks from Mars?, *Proc. Lunar Planet. Sci. Conf., 12th*, 1359-1375, 1981.

Allan H. Treiman, Lunar and Planetary Institute, 3600 Bay Area Boulevard, Houston TX 77058-1113. (e-mail: treiman @lpi.jsc.nasa.gov)

(Received February 28, 1994; revised July 14, 1994; accepted August 24, 1994.)

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5341–5349, MARCH 25, 1995

Simultaneous adsorption of CO_2 and H_2O under Mars-like conditions and application to the evolution of the Martian climate

Aaron P. Zent and Richard C. Quinn
SETI Institute, Space Sciences Division, NASA Ames Research Center, Moffett Field, California

Abstract. The Martian regolith is the most substantial volatile reservoir on the planet; estimates of its adsorbed inventory have been based on simple measurements of the adsorption of either water or CO_2 in isolation. Under some conditions, H_2O can poison adsorbate surfaces, such that CO_2 uptake is greatly reduced. We have made the first measurements of the simultaneous adsorption of CO_2 and H_2O under conditions appropriate to the Martian regolith and have found that at H_2O monolayer coverage above about 0.5, CO_2 begins to be displaced into the gas phase. We have developed an empirical expression that describes our co-adsorption data and have applied it to standard models of the Martian regolith. We find that currently, H_2O does not substantially displace CO_2, implying that the adsorbate inventories previously derived may be accurate, not more than 3–4 kPa (30–40 mbar). No substantial increase in atmospheric pressure is predicted at higher obliquities because high-latitude ground ice buffers the partial pressure of H_2O in the pores, preventing high monolayer coverages of H_2O from displacing CO_2. The peak atmospheric pressure at high obliquity does increase as the total inventory of exchangeable CO_2 increases.

1. Introduction

The regolith, cap, and atmosphere constitute an integral part of the Martian climate system. Water and carbon dioxide exchange between these three primary reservoirs on a variety of timescales. Over astronomical timescales of 10^5–10^6 years, geologically significant amounts of volatiles can exchange, profoundly affecting the Martian climate.

The exchange of H_2O and CO_2 have always been modeled separately, and the exchange patterns predicted by numerical climate models have been described separately [e.g., *Farmer and Doms*, 1979; *Toon et al.*, 1980; *Fanale et al.*, 1982, 1986; *Zent et al.*, 1987]. This practice omits any interactions between the two species from consideration. In particular, the competition for a finite number of adsorption sites in the Martian regolith, a phenomenon known as co-adsorption, may significantly alter the distribution of CO_2 from that predicted in single volatile models. In previously reported laboratory measurements, the presence of H_2O vapor severely poisons the CO_2 adsorptive capacity of commercial zeolites [*Scott*, 1968; *Carter and Husain*, 1974]. However, these measurements were made under conditions very different from those encountered in the Martian regolith.

The objective of this paper is to measure the simultaneous adsorption of CO_2 and H_2O under conditions that are appropriate to the Martian regolith, and to apply those measurements to a numerical model of the Martian CO_2 inventory, and quasi-periodic climate change on Mars.

We undertook this study because CO_2 adsorption measurements, made under conditions of zero relative humidity, have been used to estimate the inventory of exchangeable CO_2 in the Martian atmosphere + cap + regolith [e.g., *Fanale and Cannon*, 1971], and also to predict the evolution of the Martian climate caused by oscillation of orbital elements [e.g., *Fanale and Cannon*, 1974; *Toon et al.*, 1980; *Zent et al.*, 1987; *François et al.*, 1990] (see *Kieffer and Zent* [1992] for review). Failure to account for a potentially first-order process such as adsorptive competition jeopardizes the conclusions reached in previous studies.

Paper number 94JE01899.

2. Background

Previous Climate Models

Previously published climate models rely on the assumption of adsorptive equilibrium of CO_2 between the atmosphere and regolith. Numerical representations of adsorptive equilibrium are based on empirical fits to data acquired under conditions of zero relative humidity (R_H) [*Fanale and Cannon*, 1971, 1978; *Zent et al.*, 1987]. One assumes that the isotherms so derived correctly describe the behavior of the entire Martian regolith. One can then solve the heat diffusion equation for annually averaged insolation and determine the temperature profile to any depth of interest. One must assume values for the thickness of the regolith, for its thermal conductivity, and for the planetary heat flux.

The temperature profile of the regolith as a function of latitude, combined with the atmospheric pressure, permit one to calculate how much CO_2 is adsorbed on the regolith as a function of latitude and depth. The current exchangeable CO_2 inventory is estimated as that which simultaneously satisfies the adsorptive equilibrium, the observed atmospheric pressure, and the absence of large polar CO_2 caps. By varying assumptions of the depth of the regolith and its specific surface area, a variety of CO_2 inventories consistent with observations can be generated.

The atmospheric pressure history over an obliquity cycle

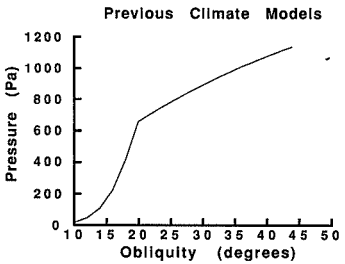


Figure 1. Previous models of the evolution of the Martian climate with obliquity were calculated without regard to the effects of H_2O. This example is after *Zent et al.* [1987].

can then be found by recalculating the thermal gradient at each obliquity, and rebalancing the CO_2. If the atmospheric pressure exceeds the vapor pressure at polar temperatures, permanent polar caps are assumed to exist, and they are then allowed to fix the atmospheric pressure. The adsorbate inventory is recalculated in accord with the new atmospheric pressure, and the remainder of the CO_2 inventory is assumed to be in the quasi-permanent polar caps. This climate model approach was developed by *Fanale et al.* [1982] and the same scheme will be employed in this study.

Previous climate models predict that at obliquities slightly lower than at present (roughly 25°), permanent CO_2 polar caps will form, trapping a significant fraction of the CO_2 inventory. At higher obliquities, the high-latitude regolith releases CO_2, but this is somewhat compensated by cooling of the low-latitude regolith, which takes up some of the CO_2 [*Toon et al.*, 1980; *Fanale et al.*, 1982]. The result is a slight increase in pressure as the obliquity increases (Figure 1).

The Role of H_2O

In laboratory measurements of the simultaneous adsorption of water and CO_2, the adsorptive abundance of CO_2 was substantially depressed from its $R_H = 0$ levels. *Carter and Husain* [1974] passed a gas stream carrying CO_2 and H_2O through a packed zeolite bed at 273 K, and examined the breakthrough curves of the two gases. The principle is identical to that used in gas chromatography, where the greater the adsorptive coverage on the bed material, the longer the gases take to appear at the end of the bed. They found that CO_2 breakthrough was significantly faster when water was present in the gas stream, indicating that the CO_2 adsorptive capacity of the zeolite material was being poisoned. Importantly, the breakthrough curve for water was unaffected by the presence of CO_2, indicating that water adsorption was unaffected by CO_2. The explanation of this asymmetry is that the lifetime of an adsorbed molecule on a surface is dependent on the partial molar enthalpy of adsorption as

$$\tau = \tau_0 \exp\left(\frac{Q}{RT}\right) \quad (1)$$

where τ_0 is a characteristic frequency for the vibration of the adsorbent, characteristically of the order of 10^{-14} s. In Table 1 the characteristic lifetime is shown as a function of Q for 298 K. Typical values for the molar enthalpy of adsorption for water and CO_2 are of the order of 4.48×10^4 J mol^{-1} [*Mooney et al.*, 1952] and 1.96×10^4 J mol^{-1} [*Zent et al.*, 1987], respectively. At 210 K, this would imply adsorbed lifetimes of 7×10^{-3} s for H_2O and 4×10^{-9} s for CO_2, a difference of 6 orders of magnitude. CO_2 is present at roughly 800 times the number density of H_2O in pore gases, and so collides and adsorbs more frequently, but the difference in adsorbed lifetime suggests that adsorbed H_2O is essentially unaffected by the presence of CO_2 at Mars-like conditions. We shall see that our data support this conclusion.

Scott [1968] also reported that CO_2 was displaced from adsorption sites by the presence of H_2O. In his study, co-adsorption for purification of He streams in gas-cooled nuclear reactors was studied experimentally at adsorbent temperatures of 298 K, by passing a He stream at 10–30 atm over zeolite beds, and metering in both water and CO_2. He found that sorbed CO_2 was irreversibly replaced by sorbed water and that CO_2 loading was dependent on water concentration.

If CO_2 is also substantially displaced from Martian materials, then the current estimates of the adsorptive capacity of the regolith may be too high. The Martian regolith is at $R_H = 1$ throughout the latitude-depth domain in which ground ice is stable, roughly poleward of 40° latitude, with notable outliers in high-albedo areas [*Paige*, 1992; *Mellon and Jakosky*, 1993]. Our objective in this paper is to determine how H_2O and CO_2 compete for adsorption sites as a function of temperature and their respective partial pressures at Mars-like conditions, and then apply that relationship to a numerical model of quasi-periodic climate variations throughout the Martian obliquity cycle.

3. Laboratory Effort

Experiment

In order to test the adsorptive capacity of Martian regolith materials, we measured co-adsorption of H_2O and CO_2 under conditions appropriate to the Martian regolith, as well as the isotherms of the individual species in isolation.

The manifold used for the adsorption measurements (Figure 2) allows the temperature of the soil, and the partial pressures of CO_2 (99.99%) and H_2O to be controlled. The partial pressure of CO_2 is controlled with a gas-regulating valve and measured with a capacitance manometer. The partial pressure of the H_2O is controlled by submerging the H_2O reservoir in a cryogenic bath. The temperature of the H_2O reservoir is measured with a thermocouple temperature probe in direct contact with the ice. Doubly distilled H_2O

Table 1. Lifetime of Adsorbed Molecules as a Function of Heat of Adsorption

Q, J/mol	t at 298 K, s
1×10^2	5.21×10^{-14}
1×10^3	7.49×10^{-14}
5×10^3	2.51×10^{-13}
1×10^4	2.83×10^{-12}
2×10^4	1.60×10^{-10}
5×10^4	2.90×10^{-5}
1×10^5	1.68×10^4

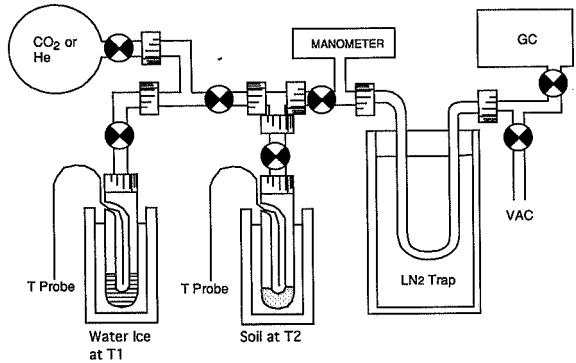


Figure 2. Schematic rendering of the manifold used in the co-adsorption experiment.

was degassed by placing it in the reservoir, freezing with liquid nitrogen, degassing, melting, and then refreezing. Four cycles of freezing, followed by degassing and melting, were carried out.

We used palagonite, collected from the summit of Mauna Kea, and sieved to <38 μm, as our Mars analog material. The palagonite is conditioned before each run by heating it overnight to 120°C under vacuum. The temperature of the soil is controlled with a second cryogenic bath and measured with a thermocouple probe in direct contact with the soil. The soil is kept at a higher temperature than the H_2O reservoir in order to keep the relative humidity in the soil chamber below unity.

The dead space in the soil sample is calculated by expanding He into the soil reservoir and noting the pressure drop. It is assumed that there is negligible He adsorption.

It was determined that adsorptive equilibration takes at least four hours. A series of experiments was performed in which the only variable was the equilibration time. Analyses on all samples exposed 4 hours or longer yield the same adsorbate loading, within experimental error. For one of our 210 K measurements, we carried out the exposure for 12 hours to investigate kinetic barriers to equilibration. Since our soil samples were small (roughly 0.1 g), four hours was adequate time to charge them with their equilibrium load, even at 210 K.

After equilibration, the adsorbed gases are desorbed from the sample, trapped, and analyzed by gas chromatography. To trap the CO_2 and H_2O, the soil chamber is heated to 120°C for 2 hours, and the desorbed gases are captured in a U-trap cooled in a liquid nitrogen bath. Spectroscopic and differential scanning calorimetry studies of palagonite have shown that virtually all H_2O is desorbed at this temperature [*Bruckenthal*, 1987].

After the trapping step is complete, the trap is removed from the manifold, and desorbed volatiles are analyzed using a Varian 3400 gas chromatograph (GC) with thermal conductivity detector. To analyze a sample, the trap is heated to 120°C to minimize gas adsorption on the walls of the trap, and a sample is introduced into the GC column using a gas sampling valve also heated to 120°C. A 6 ft × 1/8 inch Hayesep N 60/80 mesh column, heated to 140°C was used to separate the gases. Peak elution areas are calculated using an onboard integrator.

Results

Measurements. We first measured zero relative humidity CO_2 isotherms in order to establish a baseline against which to compare the effects of H_2O. We measured the $R_H = 0$ isotherms by simple pressure drop; the GC was not used. We used these isotherms to calculate a specific surface area of 94.8 $m^2\ g^{-1}$ for the palagonite. Brunauer, Emmett, and Teller (BET) specific area determinations made using N_2 at 77 K were 116 $m^2\ g^{-1}$. The data were comparable to previously measured data for CO_2 on palagonite [*Zent et al.*, 1987], when corrected for differences in the specific surface area of the soil. Table 2 shows the data for the zero relative humidity CO_2 isotherms, which we will use as controls to examine the amount of desorption caused by H_2O.

Table 2. The Zero Relative Humidity Isotherms of CO_2 Under Appropriate Conditions of Partial Pressure and Temperature

Adsorbed CO_2, g/g	Pressure, Pa	Temperature, K
0.0179	15.3	155.8
0.0223	125.3	155.8
0.0265	148.0	155.8
0.0284	266.6	155.8
0.0302	393.3	155.8
0.0319	615.1	155.8
0.0057	72.0	210.0
0.0089	197.8	210.0
0.0110	384.5	210.0
0.0128	571.8	210.0
0.0144	750.1	210.0
0.0154	962.2	210.0
0.0169	1259.8	210.0
0.0037	267.0	242.0
0.0052	510.4	242.0
0.0064	754.9	242.0
0.0074	1000.6	242.0
0.0086	1281.5	242.0

Table 3. Adsorbed H_2O, Measured Without Any CO_2 Present

Adsorbed H_2O, g/g	Pressure, Pa	Temperature, K
0.00446	0.47	210.0
0.04076	611.60	294.8
0.04042	611.60	294.9
0.03833	611.60	294.6

We also made a few measurements of H_2O in isolation, with the gas chromatograph. The data gathered on H_2O in isolation are shown in Table 3. The soil temperature was 294.5 K (room temperature), and the temperature of the H_2O ice was 273.15 K, maintained by an ice-water bath. The partial pressure of H_2O was therefore 611.6 Pa, and the relative humidity in the soil chamber was 0.193. High adsorptive coverages, 40 mg g^{-1}, were measured.

We next measured the co-adsorption isotherms at a variety of temperatures, partial pressures, and relative humidities (Table 4). At T_{soil} = 210 K, we found that H_2O has limited effectiveness in displacing CO_2 from its R_H = 0 coverage. In the first series of measurements, we buffered the ice at 184 K (R_H = 0.013), in the second, at 206 K (R_H = 0.47), and in a third at 209 K (R_H = 0.87). In the final case the measured H_2O coverage is only about 7% of a complete monolayer, and the CO_2 adsorptive coverage is reduced by about 25% from its zero relative humidity coverage. The data are shown in Figure 3, along with the zero relative humidity data for comparison. This result is significant because these conditions are characteristic of much of the regolith. Further, the CO_2 adsorption data measured via gas chromatography at the lowest humidity nearly overlie the isotherms measured via pressure drop, giving us confidence in both techniques.

At higher temperatures (T_{soil} = 242 K) the presence of water begins to be felt. We measured co-adsorption isotherms at $R_H = 9.3 \times 10^{-3}$ and 0.5. At R_H = 0.5 the

Table 4. Co-Adsorption Data

P CO_2, Pa	Adsorbed CO_2, mg/g	P H_2O, Pa	Adsorbed H_2O, mg/g	T Soil, K
190.0	9.23	0.00933	0.249	210.
192.0	8.55	0.00933	0.193	210.
217.3	9.69	0.00933	0.225	210.
380.0	10.41	0.00933	0.177	210.
496.1	12.73	0.00933	0.289	210.
701.9	13.64	0.00933	0.273	210.
800.2	16.48	0.47	0.812	210.
809.4	14.73	0.47	2.12	210.
684.1	15.55	0.47	0.83	210.
738.6	15.38	0.47	1.17	210.
860.7	16.20	0.47	1.12	210.
926.3	14.96	0.47	0.916	210.
889.5	11.35	0.615	2.06	210.
871.5	11.92	0.615	2.00	210.
867.0	7.12	0.47	2.91	210.
469.5	3.55	14.21	8.35	243.
705.9	2.96	14.21	13.48	243.
822.4	5.22	14.21	6.57	243.
706.3	4.84	0.362	0.29	243.
770.6	5.61	0.362	0.49	243.
1006.8	6.24	0.362	1.00	243.
1084.	6.30	18.1	12.4	243.

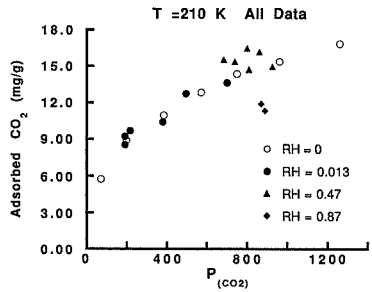


Figure 3. The T = 210 K adsorption data for CO_2, at R_H = 0 and higher relative humidities.

measured H_2O coverage is about 0.25 monolayers, and the CO_2 coverage is depressed by roughly 30% from its dry coverage. The data are shown in Figure 4.

The adsorption of H_2O is virtually independent of the partial pressure of CO_2. This is what is expected from previously published work. Surface area analyses suggest that at R_H = 0.47, there is only about 4% of a monolayer coverage at this partial pressure. The more water in the system, the more error in the data.

Empirical fits. The R_H = 0 CO_2 data can be fit by least squares to a logarithmic form as follows:

$$\rho_a = A_s \delta P^{\gamma} T^{\beta} \tag{2}$$

where A_s is the specific surface area of the soil, 9.48×10^4 m^2 kg^{-1}, and δ, γ, and β are unknowns to be determined. We find a best fit to the data for $\delta = 5.749 \times 10^1$, $\gamma = 0.2788$, and $\beta = -4.0711$. Figure 5 shows the adsorption data and the isotherms fit to them.

We assume that the adsorptive coverage of H_2O is independent of the CO_2 partial pressure or coverage. This assumption is supported by our data, by the data from *Carter and Husain* [1974], and by consideration of the relationship between ΔH_a and adsorbed lifetimes. By taking the co-adsorption data described below, we are able to fit an isotherm of the same form as (2) to the H_2O data. The corresponding values for the H_2O isotherm constants are

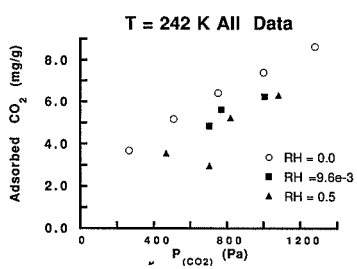


Figure 4. The T = 242 K adsorption data for CO_2, at R_H = 0 and higher relative humidities.

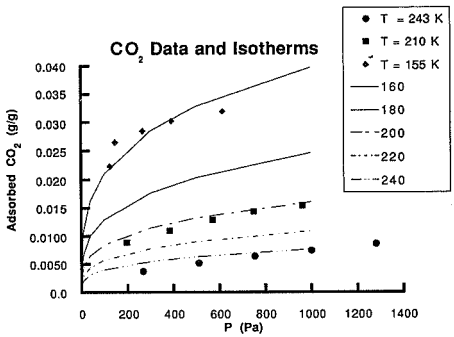


Figure 5. CO_2 data and isotherms fit to (2).

$\delta = 0.7842$, $\gamma = 0.57504$, and $\beta = -3.209$. The H_2O isotherms are shown in Figure 6, along with the data.

There are several difficulties with completing orderly isotherms for H_2O. One such problem derives from our use of cryogenic baths to control the partial pressure of H_2O in the experiment. There are only certain partial pressures available, limited by the supply of cryogenic chemicals, which have specified eutectics. Second, the technique we have employed assumes that all H_2O is desorbed from the soil and is vaporized in the LN_2 trap prior to injection into the GC column. If either of these conditions fails to apply, the technique will underestimate the adsorbed H_2O. The H_2O measurements in the co-adsorption data show a considerable standard deviation, which may arise from violation of any of these assumptions, or which may simply arise from hysteresis in the adsorption branch of the isotherm.

In deriving an empirical expression that will describe co-adsorption, we assume a modified form of the Langmuir isotherm. Studies of mixed gas adsorption are fragmentary because of difficulties with experimental and theoretical investigations [*Jaroniec and Madey*, 1988]. Discussion of the theoretical foundations of co-adsorption is beyond the scope of this paper. The problem is particularly bedeviling for heterogeneous surfaces, where many different functional groups are exposed to the adsorbates, and a variety of adsorption energies characterize individual adsorption and desorption events. Our approach is to derive our own empirical equation, which we use to fit to the data, and for subsequent interpolation. Although the derivation depends upon a number of unrealistic assumptions, if thought of as an empirical representation rather than a physical explanation, it serves well [*Carter and Husain*, 1974; *Sircar and Kumar*, 1983; *Golden and Sircar*, 1994]. We assume that only one molecule can adsorb at a given site; if f represents the fraction of collisions that result in CO_2 adsorption, then the adsorption rate is

$$f\tfrac{1}{4}n_{CO_2}\bar{c}(1 - \theta_c - \theta_w) \tag{3}$$

where θ_c is the fraction of adsorption sites already occupied by CO_2 and θ_w is the fraction of sites occupied by H_2O. The desorption rate depends on the lifetime of adsorbed molecules and the coverage, so at equilibrium,

$$f\tfrac{1}{4}n_{CO_2}\bar{c}(1 - \theta_c - \theta_w) = \nu \exp(\beta/T)\theta_c \tag{4}$$

or

$$\frac{fn_{CO_2}\bar{c}}{4\nu \exp(\beta/T)} = \frac{\theta_c}{(1 - \theta_c - \theta_w)} \tag{5}$$

if we let $n_{CO_2} = PV/RT$, per the ideal gas equation, and take the root mean square velocity of the molecules as

$$\bar{c} = \left(\frac{8kT}{\pi m}\right)^{1/2} \tag{6}$$

then we can define a constant that is independent of the temperature and partial pressure,

$$\alpha = \frac{fV(8k/\pi m)^{1/2}}{4R\nu} \tag{7}$$

so (5) reduces to

$$\frac{\alpha P}{T^{1/2} \exp(\beta/T)} = \frac{\theta_c}{(1 - \theta_c - \theta_w)} \tag{8}$$

which can be solved for θ_c, the fractional monolayer coverage of CO_2, as

$$\theta_c = \frac{(1 - \theta_w)\alpha P}{\alpha P + T^{1/2} \exp(\beta/T)} \tag{9}$$

This equation can be linearized by rearranging, and taking the log of all terms,

$$\ln(1 - \theta_c - \theta_w) - \ln(\theta_c) + \ln(P) - \ln(T^{1/2}) = \frac{\beta}{T} - \ln(\alpha) \tag{10}$$

and solved for α and β. We set the left-hand side of (10) to Y and set $X = 1/T$. In calculating the value of Y, we use the measured co-adsorption data as well as the zero relative

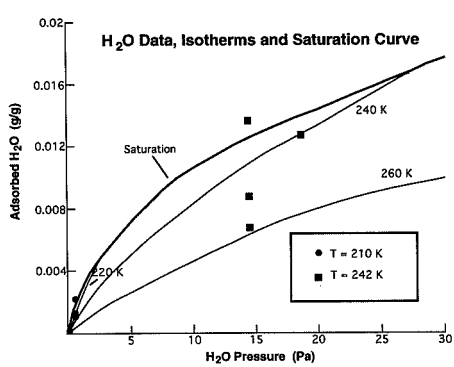


Figure 6. The H_2O data used in fitting the coefficients for (2).

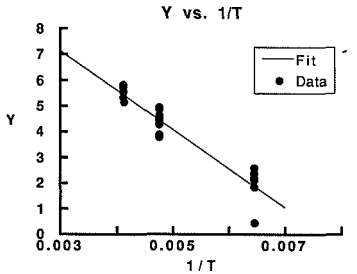


Figure 7. The co-adsorption data converted to X and Y per (10), and plotted along with our fit.

humidity data (i.e., $\theta_w = 0$ in some cases). The data, and our best fit, are shown in this format in Figure 7. We find the least squares values for the empirical constants to be $\alpha = 7.512 \times 10^{-6}$ and $\beta = -1541.52$. These values are used in (9) to calculate the adsorptive CO_2 coverage of the regolith.

4. Climate Model

Description

The numerical model employed to extrapolate the laboratory results to the Martian climate is similar in many respects to previously published models of quasi-periodic climate change on Mars.

The regolith is divided into chunks, bounded by latitude and depth, from the pole to the equator and from the bottom of the regolith to the surface. Some thickness must be assumed for the powdered component of the regolith, which may be expected to follow the adsorption isotherms measured in the laboratory; this is in distinction to the megaregolith, which is composed largely of blocks and fractured basement and which is insignificant in terms of specific surface area exposed to the atmospheric gases.

We assume that the heat diffusion equation describes the flow of heat through the Martian regolith:

$$\frac{dT}{dt} = \frac{K}{\rho c}\frac{d^2T}{dz^2} \tag{11}$$

where T is temperature, t is time, K is the thermal conductivity of the regolith (which must be assumed), ρ is the regolith density (we assume 2000 kg m^{-3}), and c is the heat capacity of the regolith. We use a temperature-dependent heat capacity applicable to dry silicates over the appropriate temperature range [*Winter and Sarri*, 1969]. The upper boundary condition is the surface temperature appropriate to the annual average insolation at each latitude, and the lower boundary condition is

$$\frac{dT}{dz} = -\frac{Q}{K} \tag{12}$$

where Q is the planetary heat flow (0.03 J m^{-2} s^{-1}) and K is the thermal conductivity of the regolith (0.8 W m^{-2}). Given these assumptions, a characteristic temperature is calculated for each chunk of the regolith. We take that temperature as the average temperature of the bounding grid points for each chunk.

We next assume, per the laboratory data and literature, that the adsorptive coverage of H_2O is independent of the partial pressure of CO_2. The equilibrium H_2O coverage of each chunk is calculated according to (2). The annual average vapor pressure of H_2O in the atmosphere is calculated by averaging the vapor pressure over the poles at 20 time steps throughout the Martian year. Because of the nonlinearity of the vapor pressure curve, one cannot calculate the annual average vapor pressure from the annual average temperature. The vapor pressure expression is

$$P = a \exp(b/T) \tag{13}$$

where $a = 3.56 \times 10^{12}$ (Pa) and $b = -6141.7$ K. If the regolith is saturated at the temperature of a given regolith chunk, then the partial pressure of that chunk is assumed to be controlled by ice at the local temperature.

The adsorptive coverage of CO_2 can then be found in the following way. The solution to (10) will provide the fractional monolayer coverage on a per kilogram basis, for a given chunk of the regolith, if θ_w, T, and P are provided. The first two terms are already calculated for each chunk. The pressure must be found iteratively as described below.

The calculation proceeds by assuming a value for the total amount of exchangeable CO_2 in the atmosphere + regolith system, and finding the atmospheric pressure P, such that

$$\sum CO_2 = PA_m g + 2\rho_c M_{\theta_c} \sum_{z=1}^{9}\sum_{l=1}^{10} \frac{V_{z,l}(1-\theta_w)\alpha P}{\alpha P + T_{z,l}^{1/2}\exp(\beta/T_{z,l})} \tag{14}$$

where the final term is summed over each chunk of regolith through each depth z and latitude l. The fractional monolayer coverage must be multiplied by the mass of a monolayer, M_{θ_c}, and the density of the regolith ρ_c. The term $V_{z,l}$ is the volume of each chunk of regolith, and A_m is the total surface area of Mars. The last term is doubled because we sum over only one hemisphere. If the atmospheric pressure P is greater than the equilibrium vapor pressure over CO_2 for the temperature predicted for the cap, then quasi-permanent CO_2 caps are assumed to exist and to set the atmospheric pressure. The mass of the caps is then found from

$$M_c = \sum CO_2 - PA_m g - 2rM_{\theta_c} \sum_{z=1}^{9}\sum_{l=1}^{10} \frac{V_{z,l}(1-\theta_w)\alpha P}{\alpha P + T_{z,l}^{1/2}\exp(\beta/T_{z,l})} \tag{15}$$

Results

CO_2 inventory. Straightforward inspection of the data in Figures 3 and 4, which represent conditions throughout much of the regolith, suggests that H_2O is inefficient in displacing CO_2 under current conditions. More elaborate numerical modeling bears out this impression.

Although the total inventory of exchangeable CO_2 (ΣCO_2) is supplied to the model, it is possible to constrain its value. Given assumptions of the specific surface area of the regolith materials (A_s), the geothermal heat flux Q, and the cap

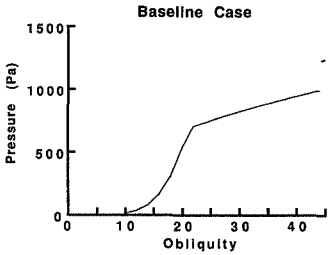


Figure 8. The predicted variation of atmospheric pressure with obliquity for the baseline case where Z = 75 m and ΣCO_2 = 722.8 kg m^{-2} (cf. Figure 1).

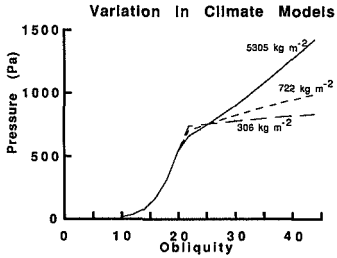


Figure 9. The full range of variation predicted by our model for the largest, baseline, and smallest exchangeable CO_2 inventory.

albedo, there are fixed combinations of ΣCO_2 and total regolith thickness Z for which the model predicts the observed conditions at the present orbital configuration. As baselines, we assume that the specific surface area of the Martian regolith materials is 17 m^2 g^{-1}, consistent with the conclusions of *Ballou et al.* [1978], and planetary heat flow of 0.03 J m^{-2} s^{-1} [*Fanale et al.*, 1982]. We use a CO_2 cap albedo of 0.6, invariant with the obliquity of the planet.

The criteria for accepting a combination of ΣCO_2 and Z are that the model must predict an atmospheric pressure of 750 Pa and produce no significant CO_2 caps. We find a range of plausible inventories, shown in Table 5. Unless the finely powdered regolith is deeper than 100 m on a global average, then the most likely exchangeable CO_2 inventory is in the range of 1–4 kPa, which is consistent with estimates of the exchangeable inventory made without consideration of co-adsorption [e.g., *Zent et al.*, 1987]. The full range of possible exchangeable inventories is bounded by 750 Pa, which is what is observed in the atmosphere at present, and which implicitly implies no significant powdered regolith; the upper limit is 19 kPa, equivalent to 5053 kg m^{-2}, which is the appropriate value for 1 km of powdered regolith.

Climate evolution. At obliquities higher than at present, the atmospheric pressure depends on the assumed values of Z and ΣCO_2. We can select any combination of Z and ΣCO_2 that appear in Table 5 and explore the evolution of the system as the obliquity and eccentricity of Mars vary. Our baseline is Z = 75 m and ΣCO_2 = 722.8 kg m^{-2}, and we vary the obliquity of the planet through its full range of 10°–45° [*Bills*, 1990; *Ward and Rudy*, 1991]. Figure 8 shows the predicted atmospheric pressure history for our baseline case, as a function of obliquity. In comparison with Figure 1, we can see that including the physics of co-adsorption does not significantly alter the predicted climate history of Mars. At the highest obliquities, the atmospheric pressure increases to about 1000 Pa (10 mbar). As the obliquity of the planet decreases, so does the atmospheric pressure, as the high-latitude regolith cools and adsorbs more CO_2. The H_2O adsorbed load of the high-latitude regolith actually decreases, since more H_2O molecules find ice to be the lowest energy state.

Table 5. The Combinations of Powdered Regolith Thickness and Exchangeable CO_2 Inventory That Yield P = 750 Pa and No Substantial Permanent CO_2 Caps

Regolith Thickness, m	Σ CO_2, kg m^{-2}
0	199.4
15	306.54
30.	412.40
50.	551.69
75.	722.82
125.	1055.18
200.	1529.26
500.	3152.42
1000.	5053.03

We explore the effects of our baseline assumptions in Figure 9, where the most extreme combinations of Z and ΣCO_2 are assumed. Where 1 km of powdered regolith is assumed, and 19 KPa of exchangeable CO_2, we find that atmospheric pressures increase to 1440 Pa at θ = 45°. Quasi-permanent polar caps appear as the obliquity goes below 20.6°.

If we assume only 15 m of powdered regolith over the planet, and an exchangeable inventory of 1.1 kPa, the pressure increases to no more than 830 Pa at θ = 45°. There is simply insufficient CO_2 to affect a significant increase in the atmospheric pressure. Quasi-permanent polar caps form as the obliquity drops below 21.3°.

At lower obliquities the atmospheric pressure is buffered by the presence of polar caps, and the temperature of those caps controls the atmospheric pressure. We have assumed a CO_2 cap albedo of 0.6. The first appearance of a quasi-permanent polar cap as obliquity decreases and the poles cool also depends on the albedo of polar CO_2 ice. If we assume that the albedo of the permanent cap is 0.65, instead of 0.6, a permanent cap is established at 24° for the baseline case. The atmospheric pressure is lower at any given obliquity as the caps grow brighter, and the caps are more massive.

5. Discussion

H_2O Data

The validity of the results in this paper depends upon the assumption that (2), with the proper coefficients, accurately describes the adsorption of H_2O on Martian regolith materials. Unfortunately, the partial molar enthalpy of adsorption calculated from (2) is only 1.04 × 10^4 J mol^{-1}, which is somewhat lower than is typically reported.

Fortunately, because we have been compelled to make repeated measurements of adsorption at the same soil temperature and H_2O partial pressure, we have enough data to associate standard deviations with our H_2O measurements. Typically, adsorption measurements made at the same conditions have standard deviations less than 30%. While this is certainly less accuracy than we could wish for, it gives us confidence that our adsorption isotherms are not in error by more than a few tens of a percent. As we shall see below, our qualitative conclusion, that H_2O cannot substantially displace CO_2 from its adsorption sites, is robust in the face of that level of uncertainty.

The fundamental reason that co-adsorption does not alter the climatic evolution as dramatically as might have been expected is that ground ice acts to buffer the partial pressure of H_2O throughout much of the regolith. In numerical experiments in which ground ice was prevented from buffering the pore vapor pressure, increase in the atmospheric H_2O vapor pressure with obliquity caused higher H_2O adsorption, and significant CO_2 desorption. Without the buffering effects of the ground ice, the atmospheric pressure reaches 3–4 kPa at obliquities of 40°, as virtually all adsorbed CO_2 is eventually displaced.

As demonstrated by *Anderson et al.* [1967], at any relative humidity, the adsorptive coverage of H_2O on the Martian regolith decreases with decreasing temperature. Our adsorption data support this conclusion, as do the data of *Anderson et al.* [1978]. For example, we found no more than about 3 mg H_2O g^{-1} soil at soil temperatures of 210 K, even when the relative humidity approached unity. By contrast, we measured 40 mg H_2O g^{-1} soil at $R_H = 0.2$ when the soil was at 294 K. In the middle- to high-latitude Martian regolith, where temperatures are usually below 210 K and where most adsorbed CO_2 is held, the upper limit on H_2O coverage is extremely low. Assuming that the surface area occupied by an adsorbed H_2O molecule is of the order of 10.5 $Å^2$, and using our sample specific area of 94.8 m^2 g^{-1}, we estimate that a monolayer of adsorbed H_2O weighs 27 mg g^{-1} soil. Since the upper limit on adsorbed H_2O is around 3 mg g^{-1} at 210 K, we find that no more than about 10% of a monolayer is likely at these low temperatures. Since displacement is linear with the fractional monolayer coverage of a competing adsorbent, the effect of the water on the total CO_2 inventory is no more than 10%. It is for this reason that an uncertainty in H_2O coverage of 30%, or 3% of a monolayer, does not affect our conclusion that H_2O is inefficient at present in displacing CO_2 into the vapor phase.

Extent of Regolith

A second major issue that arises from this study is the uncertainty in the total mineral surface area exposed to the Martian atmosphere. The total area is usually represented as the product of the specific surface area of the regolith and its uniform depth across the planet; however, many deposits with different volumes and characteristic specific surface areas probably cover the surface. It is important to determine the correct structure for modeling the Martian regolith, since most models of volatile inventory and climate history were all built on the assumption of adsorptive equilibrium between the atmosphere and 10^2–10^3 m of fine-grained, powdered regolith. It is interesting to compare our understanding of the regolith structure on the Moon to the thickness of the regolith postulated for Mars.

The structure of the lunar regolith, at least at the Apollo sites, is reasonably well understood. Active seismic experiments were undertaken to sound the shallow stratigraphy of the lunar crust in both the highlands and the mare. The salient conclusion from our point of view is that the fine-grained component of the lunar regolith is only 5–15 m thick.

At Apollo 14, in what is thought to be Imbrium ejecta, the uppermost layer detected in seismic experiments was 8.5 m thick, and is interpreted as the highly gardened material that was found across the lunar surface [*Cooper et al.*, 1974]. It is this powdered material that is analogous to the material put in the adsorption apparatus.

Apollo 16 was in the lunar highlands, and breccia fragments recovered from the site have been dated from 4.2 to 3.9 Ga [*Maurer et al.*, 1978]. The thickness of the fine-grained surface layer is 12.2 m.

The question that arises is, To what extent does the existence of an atmosphere and chemical weathering on Mars imply that Martian fines are much more abundant than lunar fines? While a complete investigation of this issue is plainly beyond the scope of this paper, it is important.

Comparison With Viking Gas Exchange Experiment

We can make one comparison with data in order to assess the utility of our palagonite sample as a Mars-analog material. The Viking spacecraft performed soil analyses in the course of the Viking Gas Exchange experiment. When soil from the VL1 site was humidified, after equilibrating with the atmosphere at 291 K, 9800 nmol of CO_2 cm^{-3} were released [*Oyama and Berdahl*, 1977]. The assumption made by *Ballou et al.* [1978] in analyzing this experiment is that the H_2O displaced all of the adsorbed CO_2. Based upon the equations determined in this experiment, we predict the release of 12,850 nmol of CO_2 cm^{-3}, assuming full desorption of the CO_2 equilibrated at conditions of 700 Pa and 291 K. The discrepancy between our prediction and the observed CO_2 displacement is about 30%, again suggesting the degree of uncertainty in our empirical representations.

6. Summary

We have measured the simultaneous adsorption of water and carbon dioxide on fine-grained palagonite at conditions appropriate to the Martian regolith. We find that water competes poorly for adsorption sites at low temperatures and pressures characteristic of the Martian regolith. Therefore, previous upper limits estimated for the adsorbed CO_2 inventory on Mars, 3–4 kPa, remain valid. The most likely values depend on the total exposed silicate surface area. We use 720 kg m^{-2} and 75 m of powdered regolith as a baseline.

At higher obliquities, the atmospheric pressure history is dependent on the total inventory of CO_2. When the total inventory is large (19 kPa) and the powdered regolith is deep (1 km), the atmospheric pressure goes up to around 1440 Pa and $\theta = 45°$. If the total inventory is assumed to be very small (1.1 kPa) and the powdered regolith shallow (15 m), atmospheric pressures do not increase significantly with obliquity.

Buffering of pore gases by hard-frozen permafrost is primarily responsible for limiting the effectiveness of H_2O in displacing adsorbed CO_2. The maximum H_2O coverage decreases with temperature, and hence CO_2 displacement becomes less effective at lower temperatures.

Major uncertainties in the system include the utility of palagonite to represent the Martian near-surface materials, and the total surface area of regolith material exposed to the Martian atmosphere, and hence the exchangeable CO_2 inventory. Future investigations will focus on these issues.

Acknowledgments. This research was supported by the Mars Surface and Atmosphere Through Time program under RTOP 155-01-60-02 and the Planetary Geology Program under RTOP 155-01-60-11. The authors would like to thank Christopher McKay for helpful discussions. We would also like to thank Fraser Fanale and an anonymous reviewer for constructive reviews.

References

Anderson, D. M., E. S. Gaffney, and P. F. Low, Frost phenomena on Mars, *Science*, *155*, 319–322, 1967.

Anderson, D. M., M. J. Schwarz, and A. R. Tice, Water vapor adsorption by sodium montmorillonite at −5°C, *Icarus*, *34*, 638–644, 1978.

Ballou, E. V., P. C. Wood, T. Wydeven, M. E. Lehwalt, and R. E. Mack, Chemical interpretation of Viking Lander 1 life detection experiment, *Nature*, *271*, 644–645, 1978.

Bills, B. G., The rigid body obliquity history of Mars, *J. Geophys. Res.*, *95*, 14,137–14,153, 1990.

Bruckenthal, E. A., The dehydration of phyllosilicates and palagonites: Reflectance spectroscopy and differential scanning calorimetry, Masters thesis, Univ. of Hawaii, Honolulu, 1987.

Carter, J. W., and H. Husain, The simultaneous adsorption of carbon dioxide and water vapour by fixed beds of molecular sieves, *Chem. Eng. Sci.*, *29*, 267–273, 1974.

Cooper, M. R., R. L. Kovach, and J. S. Watkins, Lunar near-surface structure, *Rev. Geophys.*, *12*, 291–308, 1974.

Fanale, F. P., and W. A. Cannon, Adsorption on the Martian regolith, *Nature*, *230*, 502–504, 1971.

Fanale, F. P., and W. A. Cannon, Exchange of adsorbed H_2O and CO_2 between the regolith and atmosphere of Mars caused by changes in surface insolation, *J. Geophys. Res.*, *79*, 3397–3402, 1974.

Fanale, F. P., and W. A. Cannon, Mars: The role of the regolith in determining atmospheric pressure and the atmosphere's response to insolation changes, *J. Geophys. Res.*, *83*, 2321–2325, 1978.

Fanale, F. P., J. R. Salvail, W. B. Banerdt, and R. J. Saunders, Mars: The regolith-atmosphere-cap system and climate change, *Icarus*, *50*, 381–407, 1982.

Fanale, F. P., J. R. Salvail, A. P. Zent, and A. Z. Postzwko, Global distribution and migration of subsurface ice on Mars, *Icarus*, *67*, 1–18, 1986.

Farmer, C. B., and P. E. Doms, Global and seasonal variations of water vapor on Mars and the implications for permafrost, *J. Geophys. Res.*, *84*, 2881–2888, 1979.

François, L. M., J. C. G. Walker, and W. R. Kuhn, A numerical simulation of climate changes during the obliquity cycle on Mars, *J. Geophys. Res.*, *95*, 14,761–14,778, 1990.

Golden, T. C., and S. Sircar, Gas adsorption on silicalite, *J. Colloid Interface Sci.*, *162*, 182–188, 1994.

Jaroniec, M., and R. Madey, *Physical Adsorption on Heterogeneous Solids*, Elsevier, New York, 1988.

Kieffer, H. H., and A. P. Zent, Quasi-periodic climate change on Mars, in *Mars*, edited by H. Kieffer et al., pp. 1180–1218, University of Arizona Press, Tucson, 1992.

Maurer, P., P. Eberhardt, J. Geiss, N. Grogler, A. Stettler, G. M. Brown, A. Peckett, and U. Krahenbuhl, Pre-Imbrian craters and basins: Ages, compositions, and excavation depths of Apollo 16 breccias, *Geochim. Cosmochim. Acta*, *42*, 1687–1720, 1978.

Mellon, M. T., and B. M. Jakosky, Geographic variations in the thermal and diffusive stability of ground ice on Mars, *J. Geophys. Res.*, *98*, 3345–3364, 1993.

Mooney, R. W., A. G. Keenan, and L. A. Wood, Adsorption of water vapor by montmorillonite, II, Effect of exchangeable ions and lattice swelling as measured by X ray diffraction, *J. Am. Chem. Soc.*, *74*, 1371–1374, 1952.

Oyama, V. I., and B. J. Berdahl, The Viking gas exchange experiment: Results from Chryse and Utopia surface samples, *J. Geophys. Res.*, *82*, 4669–4676, 1977.

Paige, D. A., The thermal stability of ground ice on Mars, *Nature*, *356*, 43–45, 1992.

Scott, C. D., Removal of water and carbon dioxide from helium coolants by molecular sieves, *Nucl. Sci. Eng.*, *34*, 214–223, 1968.

Sircar, S., and R. Kumar, Adiabatic adsorption of bulk binary gas mixtures: Analysis by constant pattern model, *Ind. Eng. Chem. Process Des. Dev.*, *22*, 271–280, 1983.

Toon, O. B., J. B. Pollack, W. Ward, J. A. Burns, and K. Bilski, The astronomical theory of climate change on Mars, *Icarus*, *44*, 552–607, 1980.

Ward, W. R., and D. J. Rudy, Resonant obliquity of Mars?, *Icarus*, *94*, 160–164, 1991.

Winter, D. F., and J. M. Sarri, A particulate thermophysical model of the lunar soil, *Astrophys. J.*, *156*, 1135–1151, 1969.

Zent, A. P., F. P. Fanale, and S. E. Postawko, Carbon dioxide: Adsorption on palagonite and partitioning in the Martian regolith, *Icarus*, *71*, 241–249, 1987.

R. C. Quinn and A. P. Zent, NASA Ames Research Center, MS 245-7, Moffett Field, CA 94035-1000. (e-mail: zent@barsoom.arc.nasa.gov)

(Received March 17, 1994; revised June 13, 1994; accepted July 21, 1994.)

Page intentionally left blank

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5351–5368, MARCH 25, 1995

Evidence of ancient continental glaciation in the Martian northern plains

Jeffrey S. Kargel,[1] Victor R. Baker,[2] James E. Begét,[3] Jeffrey F. Lockwood,[4] Troy L. Péwé,[5] John S. Shaw,[6] and Robert G. Strom[7]

Abstract. Whorled ridges, spaced about 2-6 km and forming lobate patterns with lobe widths of about 150 km, occur at many locations in the northern plains of Mars, commonly in close association with sinuous troughs that contain medial ridges. These landforms resemble moraines, tunnel channels, and eskers found in terrestrial glacial terrains, such as the midcontinent of North America. Some Martian landscapes may have formed by disintegration of continental glaciers that covered much of the northern plains into the early Amazonian (i.e., late in Martian geologic history). Meltwater processes apparently were important in the collapse of these hypothesized ice sheets; hence, the glaciers apparently were wet based in part. Whereas striking similarities exist among areas of the northern plains and some glaciated Pleistocene terrains on Earth, there are also important differences; notably, drumlin fields, such as those in many glacial landscapes on Earth, are rare, absent, or not yet resolved in images of the Martian northern plains. Another major difference is that postglacial fluvial and other water-related modifications (especially erosion) of Pleistocene terrains are substantial, but similar modifications are not observed in the northern plains; a virtually complete and sudden decline in the activity of liquid surface water following glaciation in the northern plains seems to be implied. The climatic implications of the hypothesized Martian glaciers and their decline are unclear. We investigate two possibilities, alternatively involving a relatively warm paleoclimate and the modern Martian climate. The hypothesized ice sheets in the basins within the northern plains (generally at elevations lower than -1 km) suggest a relationship of these frozen bodies of water with former regional lakes or seas, which may have formed in response to huge discharges of water from Martian outflow channels. This possible relationship has been modeled. Glaciers may have evolved from seas by their progressive freezing and then grounding and sublimational redistribution of sea ice. The transition to glaciation may have taken several million years if the climate was very cold, comparable to today's, or tens of thousands of years if the climate was as warm as modern Antarctica. A glacierized sea may have involved an extended period of glaciolacustrine and ice shelf processes.

1. Introduction

The hypothesis that Mars was glaciated carries important implications for its geologic, hydrologic, climatic, and possible biological evolution. Glaciation may have contributed to development of outflow channels [*Lucchitta*, 1982; *Tricart* 1988]; formation of volcanic features [*H. Faul*, 1973, unpublished; *Hodges and Moore*, 1979; *Lucchitta*, 1981; *Battistini*, 1987; *Chapman*, 1994]; and erosion and deposition in the northern plains [*Carr*, 1984; *Lucchitta et al.*, 1986; *Baker et al.*, 1991; *Scott and Underwood*, 1991; *Kargel et al.*, 1992; *Lockwood et al.*, 1992; *Shaw et al.*, 1992], in the Argyre and Hellas impact basins, and in certain mountainous regions [*Kargel and Strom*, 1990, 1992; *Kargel et al.*, 1991; *Kargel*, 1993). The modern Martian polar caps, described in detail by *Thomas et al.* [1992], may even be active glaciers, albeit slowly moving ones [*Budd et al.*, 1986; *Fisher*, 1993]. Others have hypothesized large amounts of ground ice or ice sheets in the northern plains, but not necessarily mobile glaciers or ice sheets of continental extent [*Carr and Schaber*, 1977; *Allen*, 1979; *Rossbacher and Judson*, 1981; *Mouginis-Mark*, 1985; *Squyres and Carr*, 1986; *Squyres et al.*, 1987; *Christensen*, 1989; *Costard*, 1989; *Jöns*, 1991; *Lucchitta*, 1993; *Costard and Kargel*, 1994).

In this report we examine aspects of the Martian northern plains where landscapes which we interpret to be glacial are widespread. It is difficult with available data to rule out proposed alternate explanations for individual landforms (witness the recent debate over photogeologic interpretations of sinuous ridges in the Martian southern hemisphere [*Kargel and Strom*, 1992; *Metzger*, 1992; *Ruff*, 1994]). Notwithstanding this unavoidable uncertainty, physical analogies of Martian landforms and entire landscapes with terrestrial glacial terrains offers a strong case for the hypothesis of ancient Martian continental glaciation.

Our conclusions are ideally consistent with those of *Chapman* [1994], who mapped the area of eastern Utopia Planitia near Elysium, where features occur that she interprets to have formed by volcano-ice (glacier?) interactions.

[1] U.S. Geological Survey, Flagstaff, Arizona.
[2] Department of Geosciences and Department of Planetary Sciences, University of Arizona, Tucson.
[3] Department of Geology and Geophysics, University of Alaska, Fairbanks.
[4] Sahuaro High School, Tucson, Arizona.
[5] Department of Geology, Arizona State University, Tempe.
[6] Department of Geography, University of Alberta, Edmonton, Alberta, Canada.
[7] Lunar and Planetary Laboratory and Department of Planetary Sciences, University of Arizona, Tucson.

Paper number 94JE02447.
0148-0227/95/94JE-02447$05.00

Specifically, we and *Chapman* [1994] have concluded that (1) suites of landforms are morphologically consistent with ancient glacier activity, (2) large, thick ice sheets existed in the northern plains, (3) the ice sheets principally occupied low ground and enveloped high terrain up to a level of about -1 km elevation, and (4) the ice sheets were active during the Early Amazonian (i.e., relatively late in Martian history).

2. Observations

2.1. Thumbprint Terrain

"Thumbprint terrain" (TT) is among the most widespread terrains in the northern plains. Named for its resemblance in Viking Orbiter images to fingerprints, TT consists of parallel, en echelon, or nested sets of regularly spaced curvilinear ridges or aligned hills. The ridges are 0.5-2.5 km wide and 1-40 km long. The characteristic spacing of thumbprint ridges is 2-6 km. Whorled lobes of TT are 75-150 km wide. The heights of thumbprint ridges are not known exactly, but they probably occur in the entire range from about 10 m to 200 m high.

Thumbprint terrain is distinctive, although it occurs in a variety of gradational morphologic types; two typical examples are shown in Figures 1a and 1b. Some unusual types, which occur mainly in Isidis Planitia, may represent a variant of the same or a completely different phenomenon (Figures 1c and 1d).

Twenty-two areas of TT, between 3,000 and 420,000 km^2 each, have been identified in the northern plains, generally a few hundred kilometers from the cratered highlands boundary and at elevations between 0 and -2 km (Figure 2). In addition, two small patches of TT occur in the Hellas Basin. TT has not been found in the cratered highlands, in wrinkle-ridge plateaus (Hesperian ridged plains), or on the Tharsis and Elysium plateaus. Thus, all known occurrences of TT are restricted to topographic basins [*Lockwood et al.*, 1992].

Thermal inertia, derived from Viking orbiter infrared thermal mapper measurements, relates to the effective grain sizes of surface rocks [*Christensen and Moore*, 1992]. All known areas of TT in the northern plains have moderate or high thermal inertias (0.006 to 0.012 cal^{-1} cm^{-2} s$^{-1/2}$ K^{-1}), values that are inconsistent with air fall dust deposits [*Edgett and Christensen*, 1991]. Sixteen of the 22 areas of TT have high thermal inertias (similar to or higher than that of the boulder-strewn Viking 2 landing site), which probably indicates materials that are significantly coarser than Martian dunes [*Edgett and Christensen*, 1991]. The other 6 areas have moderate thermal inertias, consistent with coverage by sand; however, these values also could be produced by partly block-covered surfaces that have fewer boulders than the Viking 2 site.

2.2. Associated Landforms

Scott and Underwood [1991] recognized that TT is associated with branching troughs and medial ridges in Utopia Planitia (Figure 3a). *Scott* [1983] had interpreted similar troughs in Ares Vallis as meander relics. Significantly, TT is closely associated with troughs in at least nine other areas of the northern plains; where image resolution is sufficient to resolve small details, some troughs at each location have medial ridges. Figures 3b and 4 show a similar landscape in Arcadia Planitia. TT ridges in Arcadia Planitia are 5-40 km long, they are spaced ~4 km apart, and they form whorled lobes (concave to the NE) ~100 km wide.

TT in Arcadia Planitia merges southward into an anastomosing system of troughs (Figures 3b, 4). The troughs, 2-6 km wide, are fairly well integrated, although some are isolated. TT-associated troughs in Arcadia Planitia, as in several other areas of the northern plains, contain medial ridges (Figures 1b, 3a, 4b, 4c), which range from 400 to 1200 m in width and are probably a few tens of meters high. Medial ridges are 1 to 40 km long; they exhibit abrupt variations in width and, like the troughs, many terminate abruptly. Some medial ridges exhibit low-order branching (Figure 4c). Smooth plains (smooth at resolutions of tens to hundreds of m/pixel) appear to embay the southern ends of the troughs in Arcadia Planitia. Embayment of troughs by smooth plains is the rule in other occurrences of this type of landscape.

Other features in Arcadia Planitia include mounds 1-5 km across; some are flat topped, others are rounded, and many possess basal scarps or terraces (Figure 4*b*). Similar mounds occur near other areas of TT and, as in Arcadia Planitia, many occur within sinuous troughs (Figures 1*b*, 4*b*). Some mounds connect segments of medial ridges.

This association of TT, sinuous troughs, medial ridges, mounds, and smooth plains occurs widely in the northern plains but apparently nowhere else on Mars. As discussed below, we favor a glacial interpretation for its origin. Accordingly, we refer to this type of landscape as the Northern Plains Glacial Association (NPGA).

3. Origin of the Northern Plains Glacial Association

3.1. Interpretations of Landforms

Suggested origins of TT are accumulations of eolian sediment that were deposited at successive stands of retreating mantles of dust or ice [*Carr and Schaber*, 1977], solifluction lobes or ice-cored periglacial ridges [*Rossbacher and Judson* 1981], ice-pushed or wave-built beach ridges or other near-shore lacustrine deposits [*Parker et al.*, 1989, 1993; *Scott et al.*, 1992], ice-shoved glaciotectonic ridges [*Carr*, 1984], or moraines [*Lucchitta*, 1981; *Battistini*, 1987; *Scott and Underwood*, 1991; *Kargel et al.*, 1992; *Lockwood et al.*, 1992; *Shaw et al.*, 1992]. It is uncertain if all types of TT are variations of the same basic landform or if different phenomena might be represented by the more unusual variants. The similarity of geographic occurrence, scale, spacing, and whorled patterns of different types of TT suggests that there was a shared genetic process. As examined below, the characteristics of TT and regional landform associations support a glacial origin.

Scott and Underwood [1991] interpreted TT as ice-thrust ridges, associated troughs as meltwater channels, and the medial ridges as fluvial bedforms. *Chapman* [1994] observed similar troughs with medial ridges in eastern Utopia Planitia near Elysium in association with features that she interprets as having been produced by subglacial volcanic activity (table mountains, moberg ridges, and jökulhlaup deposits). Although the specific mechanism of formation of the troughs and medial ridges and the origin of meltwater may have varied from place to place, we agree with *Scott and Underwood* [1991] and *Chapman* [1994] that these trough-ridge features probably were formed by meltwater and wet-based glaciers (i.e., glaciers having basal temperatures at the pressure-melting point of ice). We interpret TT as moraines or other ice-marginal deposits, associated troughs as tunnel channels (subglacially eroded fluvial channels), the medial ridges as eskers (subglacial stream-deposited sediment ridges), the mounds as possible kames (ice-contact, stratified glaciolacustrine deposits) or table mountains (subglacially erupted volcaanoes), and the smooth plains as proglacial or post-glacial lacustrine sediment.

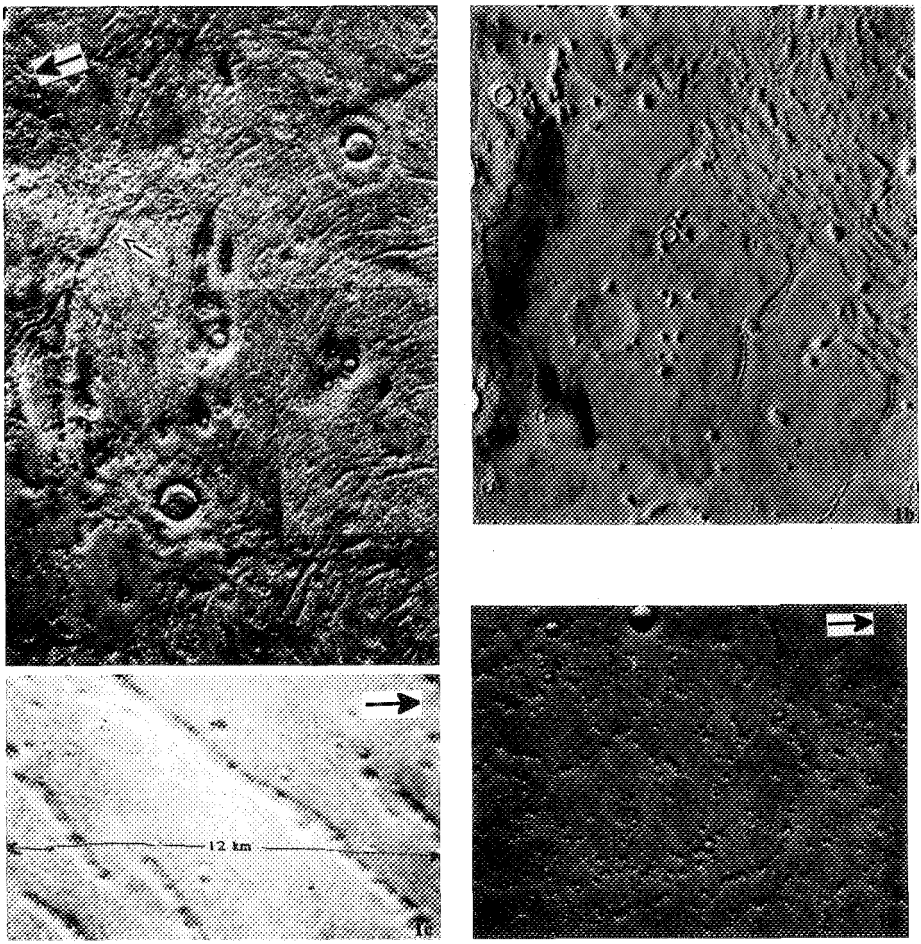


Figure 1. Different forms of thumbprint terrain (TT), which is interpreted as types of moraines or glaciotectonic ridges. Solar illumination is approximately from the bottom of each image. (*a*) This example of TT consists of relatively continuous ridges with inverted V-shaped or rounded profiles. In places, the topographic expression of TT is very slight and is mainly discernible as bright arcs. The prominent ridge at top center (small arrow) occurs where many smaller ridges converge from two sides, which suggests that it may be an interlobate moraine. Mosaic of Viking prbiter (VO) frames 057B52 to 057B57, western Utopia Planitia, latitude 50.4°N, longitude 289.0°. North given by large arrow (upper left). Scene is 125 x 189 km. (*b*) Another form of TT (upper right) is characterized by curvilinear alignments of disconnected hills and short ridges with rounded tops. This scene shows another type of ridge (interpeted as possible eskers), relatively long and narrow compared to TT ridges, that trend approximately orthogonally to the TT; some of these narrow ridges occur in shallow troughs (interpreted as possible tunnel channels), and others connect mounds (possible kames). This type of narrow ridge has otherwise been interpreted as ice-shelf features [*Lucchitta et al.*, 1986], linear pingos [*Scott*, 1983], pressure ridges at the edges of a mud ocean or giant mud flows [*Jöns*, 1991], or barrier beaches [*Parker et al.* 1993]. VO frame 608A06, southwestern Utopia Planitia, latitude 36.5°N, longitude 277.8°. North up. Scene width 102 km. (*c*) Another possible form of TT consists of long linear or arcuate ridges whose crests bear summit pits or grooves; the ridges locally have flat tops. VO frame 146S23, Isidis Planitia, latitude 16.6°N, longitude 277.8°. (*d*) Yet another possible variation of TT is distinguished by lineaments or arcs of disconnected hills and short ridges, many with summit pits. Also present are disorganized groups of hills and short ridges, recognized as possible TT only by a close association with and gradation into a more obvious type of TT. VO frame 146S13, Isidis Planitia, latitude 16.1°N, longitude 277.2°. Scene width ~32 km.

Apparently like their Martian counterparts, terrestrial tunnel channels characteristically contain or merge into eskers. The scales, sinuosities, and integration of terrestrial and Martian tunnel channels and channel-esker complexes are similar in many cases, including (1) single, isolated channels and eskers and (2) complex networks of anastomosing, rectilinear, and low-order dendritic channels and eskers [*Armstrong and Tipper*, 1948; *Farrand*, 1969; *Tipper*, 1971; *Wright*, 1973; *Shaw and Healy*, 1977; *Ehlers*, 1981; *Krüger*, 1983; *Christiansen*, 1987; *Boyd et al.*, 1988; *Attig et al.*, 1989; *Hebrand and Åmark*, 1989; *Mooers*, 1990; *Shaw and Gorrell*, 1991; *Shaw et al.*, 1992; *Brennand and Sharpe*, 1993; *Kargel*, 1993; *Brennand*, 1994; *Brennand and Shaw*, 1994]. Figure 5 shows an example of anastomosing tunnel channels on Earth, some of which contain medial eskers. Tunnel channel-esker complexes are commonly attributed to progressive channelization of subglacial floods during the collapse of glaciers; alternatively, eskers may represent a rejuvenation of subglacial flow under a more normal hydraulic regime than that which caused the formation of channels [*Brennand and Sharpe*, 1993].

Sinuous eskerlike ridges have been described from the southern hemisphere of Mars, including the south polar region, the Hellas Basin, the Argyre Basin, and highlands adjoining Argyre. Widely variable interpretations have been applied to those ridges [*Kargel and Strom*, 1992; *Kargel*, 1993; *Metzger*, 1992; *Ruff*, 1994]. Although some of those ridges are associated with channels [*Kargel*, 1993], most are not, and none are closely associated with TT, which generally seems to be lacking in the southern hemisphere. Thus, it appears that the sinuous ridges of the southern hemisphere of Mars, though probably glaciogenic, may have formed under different circumstances than TT-associated sinuous ridges in the northern plains.

TT resembles the form and scale of common types of recessional moraines deposited by Pleistocene ice sheets across central North America [*Geological Society of America*, 1959; *Wright and Frey*, 1965] and northern Europe [*Fyfe*, 1990]. Recessional moraines in the Great Lakes region form whorled lobes ~75-150 km wide (Figure 6), the same as Martian TT (Figure 3). Moraines in this region are spaced 3-25 km apart, averaging a little greater than (but overlapping) the spacing of TT ridges. Individual moraines in the Great Lakes region are 1-25 km wide, also averaging a little greater than (but overlapping) the widths of TT ridges. The cross profiles of Great Lakes moraines include sharp-crested, rounded, and flat-topped varieties, and some also have summit depressions (kettles) produced by melting of buried ice, analogous to the variety of cross profiles observed for Martian TT ridges. The height of these moraines is commonly 10-60 m. Some moraines illustrated in Figure 6 have nearly constant widths and heights for tens of kilometers, whereas others have discontinuous, "beaded" longitudinal profiles, both analogous to some types of Martian TT. These, like most moraines, are composed mostly of poorly sorted sediment and contain a large proportion of boulders and cobbles.

Several other types of ice-marginal deposits on Earth, including sublacustrine moraines, structures formed by subglacial deformation of water-saturated till, and ice-thrust ridges, also form spatially periodic whorl patterns similar to Martian TT [*Hoppe*, 1957, 1960; *Barnett and Holdsworth*, 1974; *Lundqvist*, 1981; *Sorensen*, 1983; *Aber et al.*, 1989; *Zilliacus*, 1989; *Beaudry and Prichonnet*, 1991]. This general similarity of forms produced by distinct glacial processes broadly supports the hypothesis of glaciation on Mars, but it leaves the exact mechanism of formation of TT unclear. Although we seek a unified interpretation of TT on Mars, it is entirely possible that different potentially glacier-related

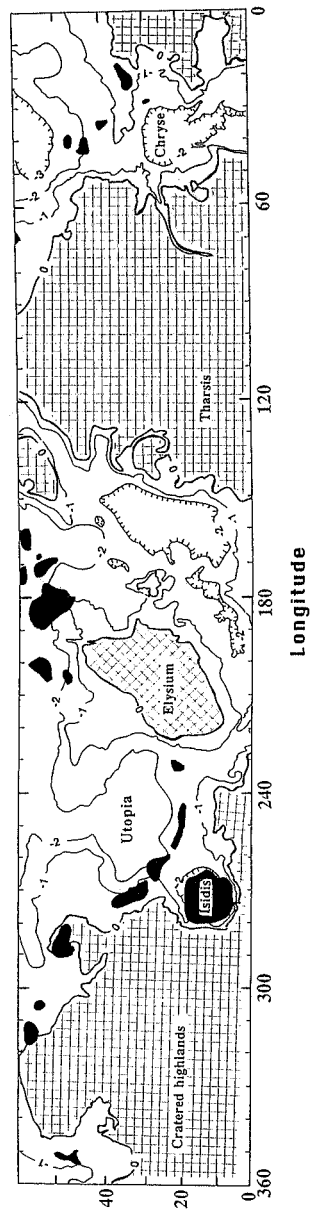


Figure 2. Distribution of TT (solid pattern) in the northern plains of Mars (modified after *Lockwood et al.* [1992]). Some areas were not imaged with sufficient clarity for detection of TT. In addition, some areas of TT exist at latitudes as high as 67°N, and there are two occurrences in the southern hemisphere Hellas Basin. Contour interval 1 km.

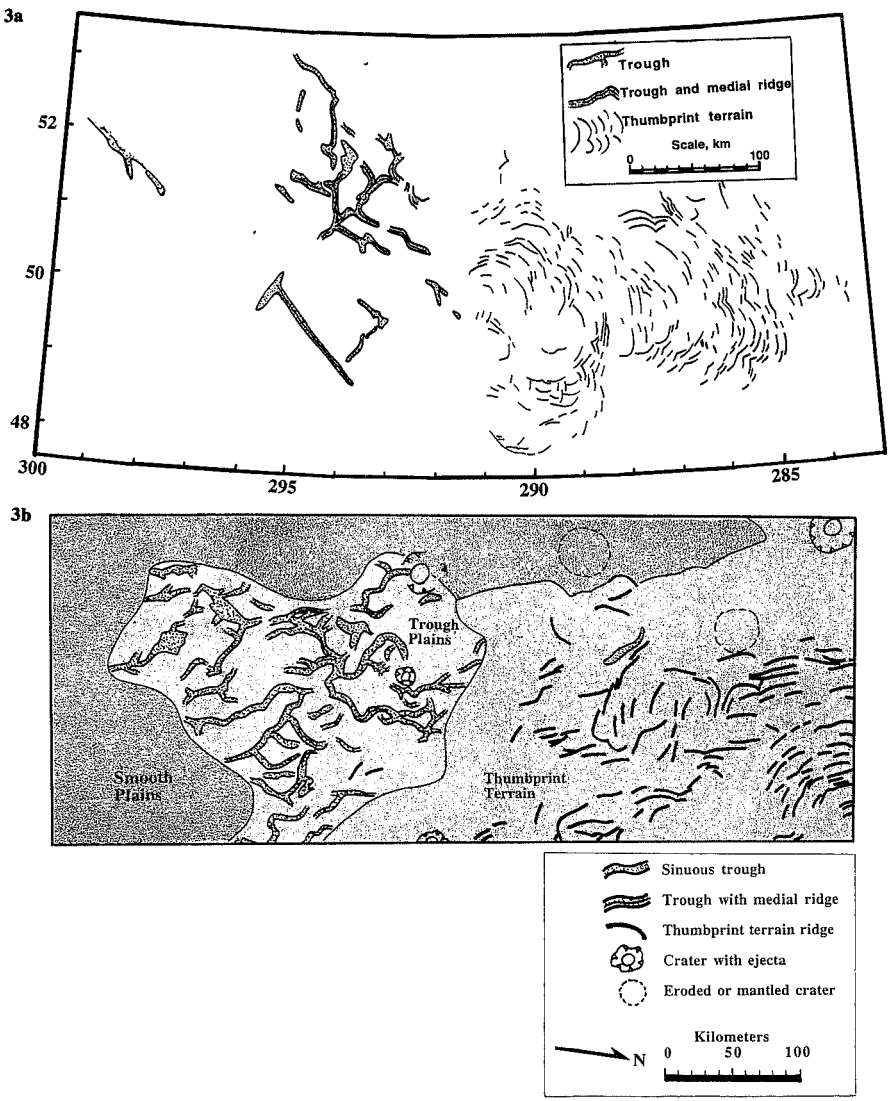


Figure 3. Generalized geologic/geomorphologic maps of two terrains of possible glacial origin (see text for detailed interpretations). (*a*) Northwesternmost Utopia Planitia showing the association between whorled ridges of TT and a system of sinuous to rectilinear troughs, many of which contain medial ridges. (*b*) Part of Arcadia Planitia showing three major terrains. Compare with corresponding image mosaic (Figure 4a). Smooth plains embay troughs at their southern and western extremities (left) and mantle old impact craters (top center). The dominant landforms illustrated here modify old cratered plains, which postdate the ancient period of heavy bombardment. Mapped from Viking Orbiter images 776A32, 776A34, and 810A29.

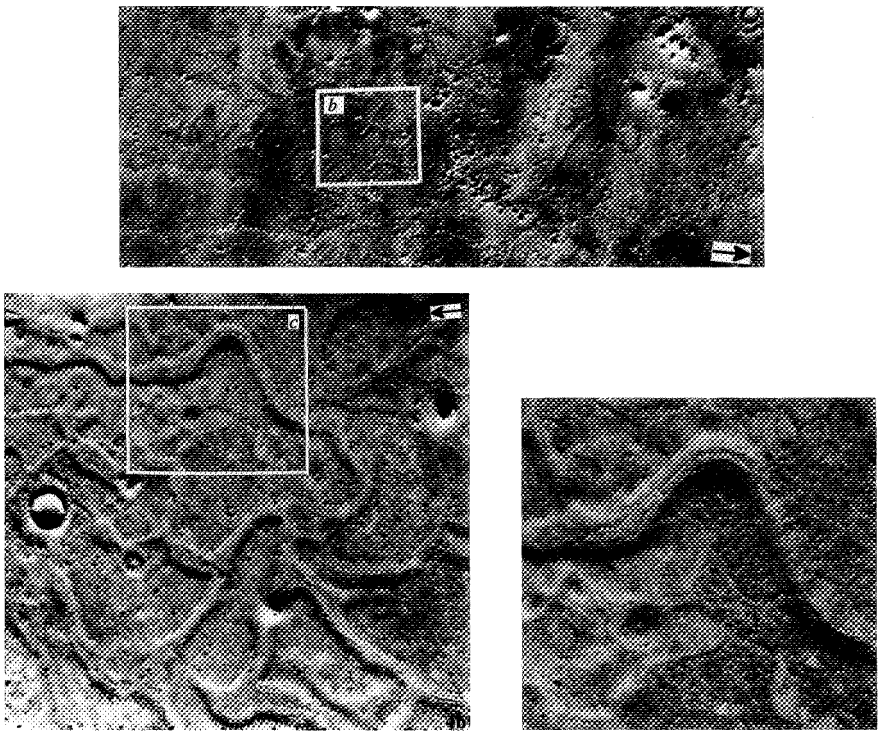


Figure 4. Landscape in Arcadia Planitia. Illumination in each scene is from below. (*a*) Entire area mapped in Figure 3b. Viking orbiter images 776A32, 776A34, and 810A29. Scene 250 x 615 km. Outlined box shows area of Figure 4b. North orientation given by arrow. (*b*) Troughs with medial ridges (interpreted as tunnel channels and eskers) and mounds (kames?). Note the scattered mounds, some of which are on the floors of channels; some mounds are rounded (below center) and others have flat tops (upper right); many have basal scarps or terraces. Also note rotation of this scene compared to the same area in Figure 3a, so that solar illumination is from the bottom of each image. VO image 076B86. Scene width 90 km. Outline shows area of Figure 4c. North orientation given by arrow. (*c*) Enlargement of part of Figure 4b, showing a branching medial ridge in a trough. Below the trough is a steep-sided, pitted mound (possibly an eroded impact crater, kame, or table mountain). Scene width ~34 km.

processes may have operated at different times and locations to produce the different forms of TT.

Terrestrial recessional moraine complexes commonly form by several processes, resulting in composite ice-marginal deposits that locally include lodgement tills, push moraines, glaciotectonic thrusts, flow tills, crevasse fill, sediment diapirs, kames, eskers, and deltas. Whatever the exact mechanism(s) of formation of TT, the best terrestrial analogs involve liquid water in the proglacial environment and/or in subglacial till. Terrestrial glaciers that are completely frozen to their beds commonly cause little erosion and do not generally produce notable recessional moraine complexes. One may infer that thumbprint terrain on Mars was formed with the involvement of liquid water as well as ice, if indeed the moraine interpretation generally is correct.

The smooth plains in the NPGA are interpreted as proglacial lacustrine sediments. On Earth, glaciolacustrine or glaciomarine plains commonly embay or mantle tunnel channels and eskers [*Armstrong and Tipper*, 1948; *Ehlers and Linke*, 1989]. One cannot rule out the possibility that other types of deposits may be components of the Martian smooth plains, but the interpretation of these plains as lacustrine sediments is consistent with the suggested importance of water in the formation of TT, troughs, and medial ridges.

The mounds in Arcadia Planitia (Figure 4b), Utopia Planitia (Figure 1b), and several other areas of NPGA may be kames. Terrestrial kames, commonly similar in shape to the Martian mounds, are characteristically associated with eskers, tunnel channels, moraines, and other features that were formed by ice stagnation [*Zoltai*, 1965; *Krüger*, 1983; *Ehlers*

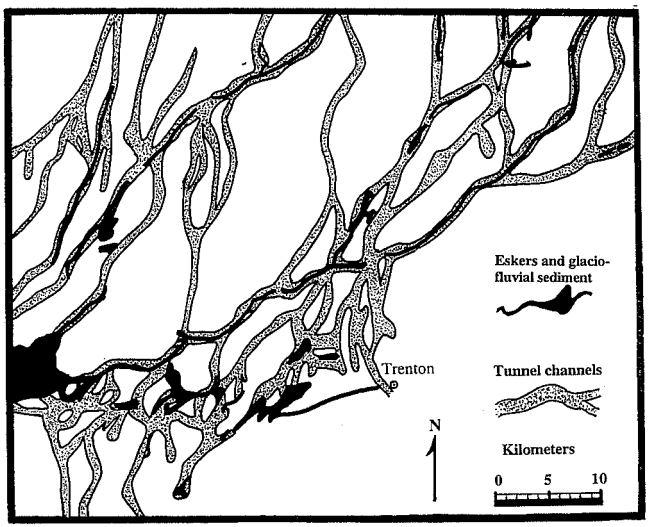


Figure 5. Map of tunnel channels and medial eskers in eastern Ontario, near town of Trenton, just north of Lake Ontario [after *Shaw and Gorrell*, 1991].

and Linke, 1989; *Mooers*, 1990; *Péwé and Reger*, 1993]. In some places on Earth, eskers originate directly from kames, and the eskers and kames occur together within tunnel channels [*Mooers*, 1990]; similar associations may be shown in Figures 1b and 4b. A basal scarp or terrace, such as those of some Martian kame-like mounds, can be eroded into kames by waves in a proglacial lake [*Zoltai*, 1965], or a basal scarp can reflect the shape of the ice-walled lake basin in which a kame was deposited [*Péwé and Reger*, 1993].

Regional geomorphologic mapping of glacial landscapes, recently with the aid of orbital imagery, has provided important insights for the sources and motion of glaciers and for glacial processes [e.g., *Aylsworth and Shilts*, 1989; *Boulton and Clark*, 1990; *Clark*, 1994]. Viking imagery generally lacks the high spatial resolutions and lateral continuity necessary for comparable regional mapping, although certain features, especially thumbprint terrain, are bold enough that fairly complete maps are possible for some regions. The general trends of curvature of TT ridges in Utopia and Arcadia Planitiae can be interpreted as markers of the northerly sources of ice and of an advance generally toward the highlands. If the smooth plains of the NPGA are lacustrine deposits, then lakes may have been impounded between the ice sheet on one side and the highland/lowland boundary on the other; this would be analogous to the late Pleistocene glacial Lake Agassiz, which was impounded between the Laurentide ice sheet to the north and basement rocks and glacial sediment deposits to the south [*Teller and Clayton*, 1983]. Isidis Planitia and other basins within the northern plains (notably near Deuteronilus Mensae) contain enclosed whorls of TT, possibly indicating progressive decay of isolated bodies of ice in depressions within the northern lowlands.

The close association of channel-ridge features with TT and the other features described here is a strong point favoring the glaciation hypothesis. However, these associations do not cover the full spectrum of glacial landscapes on Earth, nor are the observed associations always identical with the most common occurences on Earth. Two notable differences are (1) the apparent absence of large drumlin fields in the Martian northern plains, and (2) the usual convexity of TT with respect to channel-ridge features on Mars, contrasted by the usual concavity of terrestrial moraine sets with respect to tunnel channel-esker complexes. We don't know of any theory of glacial landscape genesis that would prevent moraine sets from being convex toward tunnel channels and eskers. The common association of TT with channel-ridge complexes on Mars indicates that this relationship is important. Perhaps it is indicating that subglacial outflows occurred first, the glaciers then receded or partly floated from the area of tunnel channels and eskers, and then the moraine sets formed during a phase of steadier retreat and more normal hydrologic regime; this could represent a possible reversal of the usual sequence on Earth. The apparent absence of drumlins in the Martian northern plains is discussed below.

3.2. Absence of Candidate Drumlin Fields in the Northern Plains

Drumlins are elongate, streamlined hills produced by glacial action; they are commonly 500 to 1000 m long, 100 to 500 m wide, and tens of meters high. Large fields of drumlins have not been observed in the Martian northern plains, though they may exist in Hellas [*Kargel et al.*, 1991; *Kargel and Strom*, 1992]. It is difficult to rule out the presence of drumlin fields in the northern plains, considering that there are relatively few Viking images with the high resolutions and low sun inclinations that would be necessary for their resolution. There are some good Landsat images of swarms of drumlins on Earth, but these are exceptional and generally are acquired under conditions of very low solar inclinations. Terrestrial drumlins commonly are more difficult to resolve in satellite images than are dunes of

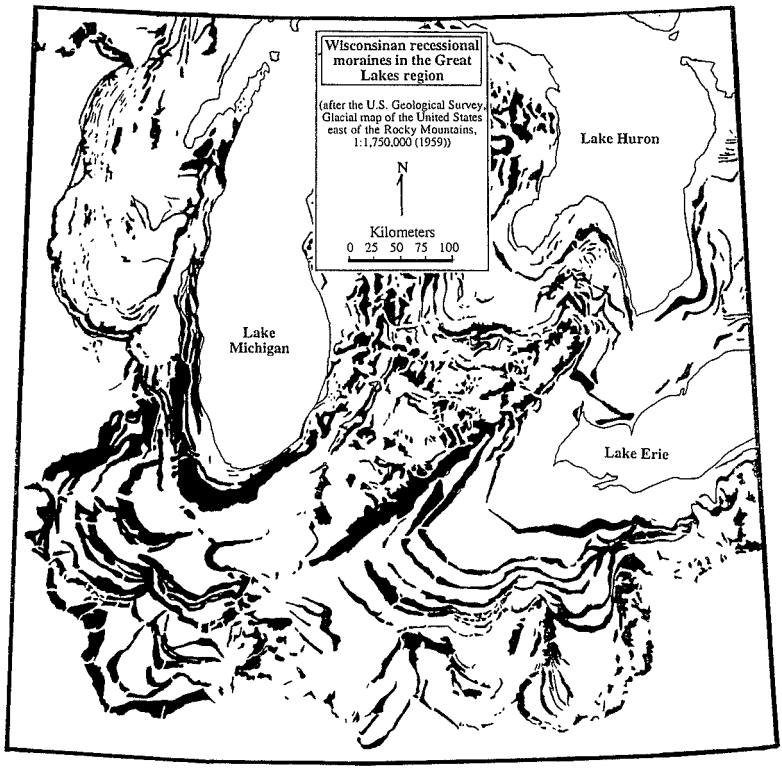


Figure 6. Map showing the pattern of ice-marginal recessional moraines associated with the retreat of the Laurentide ice sheet from the Great Lakes region.

comparable width, because drumlins commonly have lower height:width ratios than many dunes, and drumlins usually do not have steep edges near the angle of repose. Nevertheless, the apparent absence of large fields of drumlins in the northern plains may be real and significant.

Swarms of drumlins in many glaciated terrains on Earth are variously interpreted to have formed either by (1) catastrophic subglacial meltwater floods that overtopped tunnel channel rims and extended in vast sheets across the landscape [e.g., *Shaw and Kvill*, 1984; *Shaw et al.*, 1989; *Shoemaker*, 1990, 1992] or (2) plastic deformation of water-saturated subglacial till [*Aylsworth and Shilts*, 1989; *Boyce and Eyles*, 1991].

If terrestrial tunnel channels and moraines were invariably associated with drumlins, a similar association might be expected on Mars. But this is not the case: for example, tunnel channels in northern Germany are not associated with drumlins [*Ehlers*, 1981]. The north German channels are exceptionally deep and are cut into easily erodible sedimentary rocks. The lack of drumlins in this region of Europe can be explained utilizing the subglacial flood theory of drumlin formation [e.g., *Shaw and Kvill*, 1984] if deep channels contained the subglacial floods. Thus, without sheet floods, drumlins were not formed. It may be that large Martian channels also carried flood discharges, accounting for the absence of drumlins between the channels. Alternatively, under the deforming-till theory of drumlin formation [e.g., *Boulton*, 1987], the interchannel areas may have lacked the large amounts of till or other unlithified sediment that would be necessary for the formation of drumlins; the formation of Martian drumlins by this mechanism on a scale that would have been visible in Viking orbiter images would have required the prior existence of a widespread blanket of unlithified till-like sediment at least several tens of meters thick. As yet another variation of this interpretation of drumlins, the theory of moulding of terrestrial drumlins by ice streams or surges [e.g., *Clark*, 1994] would imply that the Martian glaciers in the area of the TT may not have exhibited the surge behavior thought to be necessary for drumlin formation; a lack of surge behavior could be caused by a general condition of a frozen glacier bed in a wide zone up to the glacier terminus and a restriction of basal melting to isolated regions well behind the terminus in areas of thick ice or elevated geothermal gradient.

3.3. Models of Martian Glaciation

Implications of basinal associations of glacial terrains on Mars. As described above, putative glacial terrains in the northern hemisphere are concentrated near an elevation of -1 km, seeming to imply a connection with former bodies of ponded liquid water [*Lucchitta et al.*, 1986; *Parker et al.*, 1989, 1993; *Scott et al.*, 1992; *Chapman*, 1994]. Although some interpreted glacial terrains occur at high elevations in the southern hemisphere, glaciations in all of these regions also may be related to pondings of water in major depressions (Argyre, Hellas, and the South Polar Basin). Atmospheric transfer of moisture from basins and precipitation over uplands can explain evidence of glaciation in the Charitum Montes of Argyre [*Kargel and Strom*, 1992], on Malea Planum near Hellas [*Kargel et al.*, 1991], and in mountains near the South Pole [*Kargel*, 1993]. There are no hard constraints available on the rate of such transfers, hence, no constraints on prevalent atmospheric temperature and humidity. In the case of the putative glacial landscapes of Dorsa Argentea [*Kargel*, 1993], the cause may have been gravity flows of water and ice from the South Polar Basin and ultimate drainage into Argyre. Thus, ponded water may have been instrumental in the origin of all known glacial terrains on Mars. It is unclear whether these bodies of water may have induced major changes in the thermal regime of the Martian climate [*Baker et al.*, 1991; *Kargel and Strom*, 1992] or merely supplied the water necessary for glaciation.

The glacial interpretation of these vast and widely distributed landscapes on Mars cannot be divorced from the theoretical problems of the formation of large Martian glaciers. We have considered two very different groups of models for the origin of the NPGA and the hypothesized glaciers. One model is roughly based on an Earth-like hydrologic cycle that included the growth and collapse of the Laurentide ice sheet. Another group of models is based on a unique Martian history that is directly linked with the fluvial origin of outflow channels and the formation of large lakes or seas in the northern plains.

Laurentide ice sheet analog. The growth and collapse of Pleistocene ice sheets may have had important aspects in common with the growth and demise of Martian ice sheets. The Laurentide ice sheet in the area of glacial Lake Agassiz and the Great Lakes [*Geological Society of America*, 1959; *Krüger*, 1983; *Teller and Clayton*, 1983; *Ehlers and Linke*, 1989] is a possible partial analog for the ice sheets in the northern plains. A possible scenario of glacial geologic events in Arcadia Planitia follows.

1. An ice sheet was deposited or accumulated over the area of Arcadia Planitia that now is occupied by troughs and TT. The ice sheet was mobile for long enough to entrain a substantial quantity of rocky debris (which later was deposited as eskers, moraines, and glaciolacustrine plains).

2. Subglacial geothermal meltwater (produced by slow basal melting caused by the normal geothermal gradient and heat flow) or supraglacial meltwater (produced by seasonal atmospheric or solar heating of the glacier's surface) was released at the glacier bed and caused erosion of tunnel channels during an early period of increasing discharge, probably during stagnation of the ice sheet.

3. Eskers were deposited in tunnel channels some time after maximum water flow during ice stagnation, and kames were deposited in thaw pits and collapse depressions as water discharge virtually terminated.

4. Smooth plains were deposited in a proglacial lake.

5. TT recessional moraines or glaciotectonic ridges were formed. TT may have formed where the glacier terminated directly at the lake, on dry ground following dissipation of the lake, or in subglacial structures near the glacier's terminus.

6. The lake and the remnants of the stagnated glacier evaporated, leaving "freeze-dried" glacial landscapes intact with little opportunity for degradation since then.

As an alternative to stage 2, the sources of liquid water might have been extraglacial. Catastrophic discharges of ground water from the highlands may have formed the outflow channels and then pooled between the ice sheet and the highland/lowland escarpment. Hydraulic pressure at the front of the ice sheet may have lifted the ice, resulting in "reverse outflows" from the lake into the glacier-covered area along the base of the ice sheet [after the model of *Shoemaker* 1992], thereby producing tunnel channels and eskers.

Freezing of Martian lakes and their transition to glaciers. Besides the apparent absence of drumlins in the northern plains, another important difference with respect to terrestrial glacial landscape analogs is that postglacial fluvial modification of Martian glacial terrains is nil, despite the fact that hundreds of millions or billions of years have elapsed since their formation. If the Martian glaciers, like the Laurentide ice sheet, ablated largely by surface melting, thus implying warm conditions, then why are the glacial landforms virtually pristine at the resolutions of available images? This important aspect will be discussed below, but for now it is sufficient to say that the differences between glacial terrains on Earth and Mars warrant consideration of other glacier models for Mars.

The concentration of putative glacial landforms near an elevation of -1 km in the northern plains and especially their apparent absence at high elevations may indicate a connection with paleolakes or seas, as suggested by *Kargel and Costard* [1993] and *Chapman* [1994]. Of course an ice-surfaced lake or sea is not a glacier. But is it possible that a frozen lake/sea may evolve into a glacier and thereby preserve a similar geographic relationship as previously held by the lake?

The thermal evolution of a Martian lake is considered here in two stages. First, we model the time required for the lake to start to freeze, which turns out to be a fairly short period. Second, we model the progressive solidification of the lake, including the redistribution of ice caused by sublimation and reprecipitation. These models result in the grounding of lake ice and the eventual formation of substantial surface topography on the ice mass, so that glacial flow at a finite rate would be inevitable. Our purpose is not to model any specific Martian glacier, but simply to illustrate our point that a Martian lake can evolve into a true glacial ice sheet.

If the water discharged by outflow channels had already started to freeze by the time it entered the lake basin, the surface freezing of the lake would begin as soon as it formed. However, if the lake water originated from a deep aquifer, it may initially have been warm, so that some period of time would be required before freezing would begin. This period can be calculated for reasonable assumptions of the pertinent conditions. We consider two widely differing cases. In case "a," the annually averaged surface temperature on Mars is 200 K, initial water temperature is 274 K, and water depth is 100 m; in case "b," Martian surface temperature is 250 K, initial water temperature is 293 K, and water depth is 1000 m. In both cases we assume for the sake of simplicity that there is complete mixing (by gravity flows and wind-driven waves and currents) and we neglect evaporative enthalpic loss; hence, cooling of the lake can be approximated as the integrated loss of thermal energy from the lake at a rate very roughly estimated by the difference of the upward and downward blackbody radiative fluxes. Present climatic conditions in the middle latitudes of the northern plains correspond

approximately to case a, where freezing would commence in about 19 days. Freezing would begin in case b after about 18 years. These are upper limits, because (1) any evaporative heat loss would increase the rate of cooling, (2) development of density stratification could allow commencement of freezing of the upper thermal boundary layer before the entire lake mass had cooled to the freezing point, and (3) an atmosphere with optically thin spectral windows could radiate more effectively than assumed. Hence, a Martian lake would not have had a liquid surface for long, unless the ambient surface temperature was very close to or above the freezing point.

Once freezing had commenced, the lake ice would have gradually thickened. We consider the detailed thermal evolution of an idealized Martian lake that was big enough to explain the scale and distribution of putative glacial landforms described above. The scenario modeled below includes the progressive thickening of lake ice owing to the combined effects of surface cooling, normal low levels of geothermal heating, sublimation of ice from the warmer half of the lake, and cold trapping of vapor on the colder half. The lake basin is assumed to be 3000 km wide and sinusoidal, reaching a depth of 2000 m in the middle. So that the modeled topography roughly reflects an idealized representation of the northern plains, we set the initial lake shore at an elevation of -1000 m, so that the center of the lake bed is at -3000 m. This situation roughly corresponds to a Martian paleosea that extends from about latitude 30°N to latitude 80°N. To put the amount of water in perspective, consider the idealized case that this seaway extends around the globe in this latitude belt; the volume of water is ~5 x 10^{16} m^3, corresponding to a global layer of water 340 m deep, consistent with other estimates of Mars' surface and subsurface water inventory [*Carr*, 1986]. The volume of water could be substantially less if the northern plains had several large lakes instead of a single deep seaway.

The freezing of the lake is modeled under two sets of climatic conditions, described below. Geothermal heat flow is modeled as the Earth's global average heat flow today [*Fowler*, 1990], which is roughly equal to Mars' heat flow 2 b.y. ago [*Stevenson et al.*, 1983], a possible age of Martian glaciation. (To the extent that volcanic activity may have contibuted to geothermal heating, local inputs of geothermal heat may have exceeded the levels considered here.) The effects of surface cooling are modeled by the solution to the Stefan problem [*Turcotte and Schubert*, 1982]; however, unlike this solution, which involves the thickness of the ice layer increasing throughout time at a rate inversely proportional to time, in the present model the effects of constant sublimation from the ice surface eventually results in a steady state ice thickness (e.g., the southern part of the lake in model 2, see Figure 8). The thickness of ice also depends on the input of geothermal heat or the addition of precipitated ice. These effects are all included. It is also assumed that the ice is in hydrostatic equilibrium wherever ice overlies water. However, the effects of large-scale glacial creep are not considered.

Model 1 corresponds approximately to the modification (cooling) of current conditions [*Zurek et al.*, 1992] caused by coverage of the northern plains by a dirty ice mantle in the southern ablation areas and a bright ice mantle in the northern areas of deposition. The annual average surface temperature is assumed to vary linearly across the lake from 190 K at its southern (temperate latitude) edge to 160 K at its northern (polar) edge. The lake is assumed to be a hydrologically closed system. Initial rates of net sublimation vary linearly across the lake from 0.5 mm/yr at the southern edge (about the same as the sublimation rate at the edge of the modern northern polar cap [*Jakosky and Haberle* 1992] to no net sublimation at the lake's center; net deposition varies linearly from nothing at the center to 0.5 mm/yr at the northern edge. As ice eventually recedes from the warmer, southern reach of the frozen lake, where sublimation rates are highest, part of the moisture source dries up, so that deposition rates in the northern part of the lake decline in response. The linear variation of the rates of sublimation and accumulation is an assumption rather than a result of a rigorous climate model, which has not yet been developed for the pertinent conditions and which would be expected to produce a more complex pattern of ice mass exchange.

Model 2 represents relatively mild conditions comparable to much of Greenland and Antarctica today. Annual average surface temperatures are assumed to vary linearly across the lake from 250 K at its southern edge to 200 K at its northern edge. Initial rates of net sublimation vary from 10 cm/yr at the lake's southern edge to nothing at its center; initial rates of net deposition vary from nothing at the lake's center to 10 cm/yr at its northern edge. As in model 1, as ice recedes in the warm ablation zone of the lake, deposition in the northern depositional zone slows.

The results of models 1 and 2 are shown in Figures 7 and 8, respectively. The complete solidification of the model Martian lake requires about 10^5 years under cold climatic conditions (Figure 7). Extensive redistribution of ice by sublimation and recondensation takes millions of years. At the end of 5 m.y., the ice has redistributed itself in such a way that glacial creep likely would be significant. However, this glacier is cold-based almost throughout its bed, so that ice-rock interactions might be fairly minor. Normal low levels of geothermal heating may induce basal melting if any part of the glacier exceeds several kilometers in thickness, so that formation of interesting landforms are not excluded by a cryogenic climate such as today's. High local levels of geothermal heating (e.g., caused by subglacial volcanism) could cause basal melting beneath very thin ice sheets.

In the case of a relatively warm, Antarctic-type environment (Figure 8), sublimation and recondensation occur more rapidly, so that substantial topography develops on the grounded part of the ice even within a few tens of thousands of years. A warm climate results in a glacier that discharges into an ice-covered lake. The lake does not completely solidify, and large parts of the glacier are wet-based and other parts consist of an ice shelf that is grounded at both ends. Glacial flow and meltwater processes probably would produce interesting landforms, especially near the grounding lines.

3.4. Subglacial Flooding Mechanisms and Paleoclimatic Implications

The glacial hypothesis implies that an ice sheet once existed where none is now. The wide distribution of TT (Figure 2) suggests either one very large or several smaller ice sheets. The formation of one or more ice sheets may have been a response to the inundation by water of the northern plains [*Baker et al.*, 1991]. Glacier formation may have been either a direct response due to the progressive solidification of large lakes and grounding of the ice (Figures 7-8) or an indirect response due to the development of a warmer, wetter Martian climate involving widespread snowfall and ice accumulation. Hence, the hypothesized existence of ancient ice sheets in itself does not sharply constrain the paleoclimatic conditions. Requirements are (1) a large mass of surface water or ice, (2) severely or moderately cold conditions, and (3) lateral gradients in the climate such that ice accumulation or lake freezing rates were nonuniform, thus causing the growth of a large body of ice with a surface topographic gradient. In any case, the former existence of an ice sheet at middle and

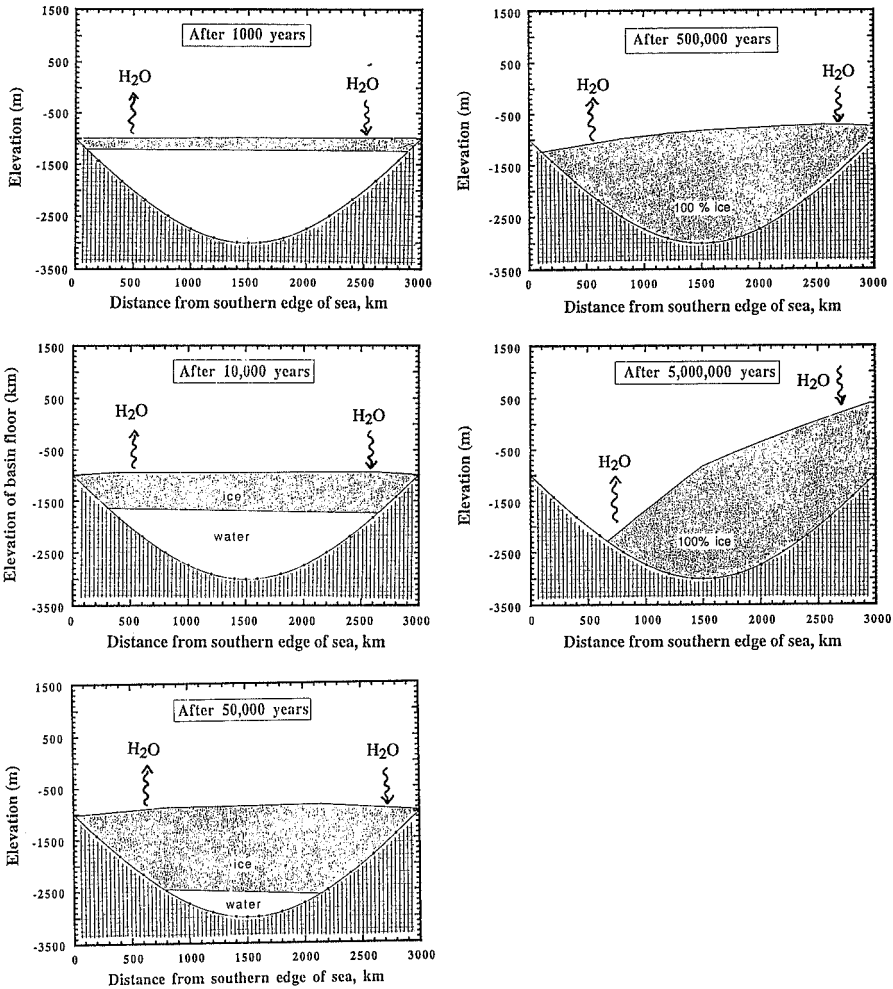


Figure 7. Model of the solidification and transition to a glacier of an idealized Martian sea under climatic conditions that are comparable to today's conditions in the northern plains, modified (cooled) by the high albedo of the ice. The high profile of the ice in the northern part of the frozen lake develops because of sublimational redistribution of ice. Details of the models are provided in the text.

low latitudes would indicate that the Martian environment (but not necessarily the climate) was once very different from today's.

Considerable liquid water apparently was involved in the formation of the NPGA. The climatic implications, if any, depend on the mechanism whereby meltwater was generated. The Laurentide and Fennoscandian ice sheets are possible analogs for ancient Martian ice sheets; modern Martian polar caps and polar layered deposits are another model. Liquid water could affect the beds of both types of ice sheets, yet the water would be generated by different mechanisms; hence, very different paleoclimates would be implied. Any model must be reconciled with the nearly pristine condition of the Martian glacial terrains (when viewed at resolutions on the order of 100 m per pixel).

The rapid disintegration of Earth's late Pleistocene ice sheets was prompted by a warming climate. If Pleistocene glacial terrains are acceptable analogs of the Martian terrains,

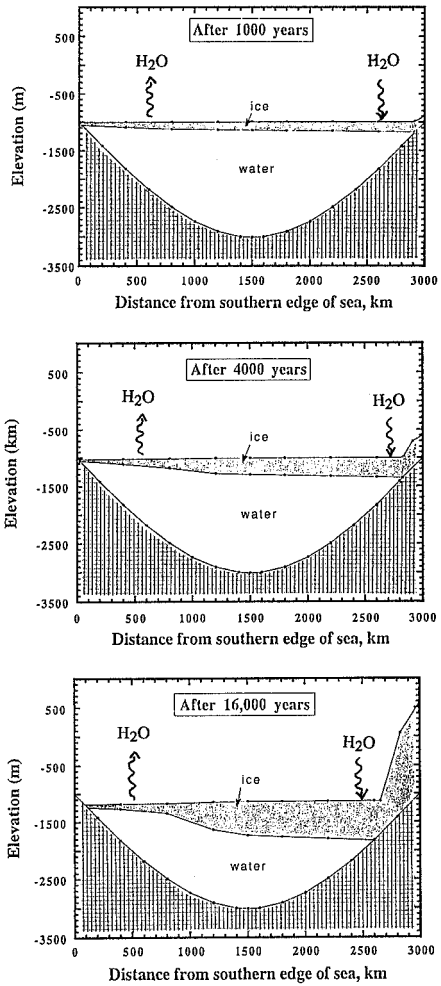


Figure 8. Model of the solidification and transition to a glacier of an idealized Martian sea under climatic conditions that are relatively warm, comparable to the Antarctic. The high profile of the ice in the northern part of the frozen lake develops because of sublimational redistribution of ice. Details of the models are provided in the text.

then a similar, relatively warm climate may be implicated in deglaciation on Mars. However, it is not clear whether a warm climate is necessary for deglaciation or whether warming simply was the particular condition that triggered the demise of Pleistocene ice sheets. It is possible that, given a continuing supply of moisture regardless of surface temperatures, ice sheets tend to thicken until subglacial geothermal melting occurs. This idea is supported by the occurrence of basal meltwater accumulations ("lakes") beneath some of the coldest and thickest parts of the East Antarctic ice sheet [*Oswald*, 1975]. If basal meltwater accumulates, megafloods or ice streaming (or surging) may result. In the most extreme cases (as with the Laurentide and Fennoscandian ice sheets), this water-modified dynamic behavior may spur deglaciation.

If deglaciation on Mars was caused by a warm climate, then there must have been very rapid climatic cooling and drying following deglaciation. Otherwise, even a few tens of thousands of years of landform degradation in a warm, humid climate would have left a mark (such as glacial features that are partly blanketed by postglacial sediment, overprinted by normal rainfall runoff valleys or gullies, or otherwise degraded); no such degradation is observed. Perhaps deglaciation occurred rapidly in a modest greenhouse climate [*Fanale et al.*, 1992] that was further modified by a brief period of high obliquity [*Kieffer and Zent*, 1992; *Mellon and Jakosky*, 1993], which resulted in warm summers at high and middle latitudes; no sooner had the ice melted than cold, dry conditions set in again and effectively freeze-dried the glacial terrains.

Alternatively, Mars always may have been too cold for significant surface melting. If the modern polar layered deposits are a better analog for the ancient Martian ice sheets, and geothermal basal melting provided the water, or if the catastrophic eruption of highland acquifers and outflow channels provided the water, then no particular paleoclimatic conditions are implied. Ablation of the ice sheets and exposure of subglacial water-formed landscapes may have been accomplished by the slow sublimation of glacier ice under exceedingly cold, dry conditions or relatively warm, dry periglacial conditions; once the terrain beneath such an ice sheet was exposed, there would have been no liquid water to degrade the landscapes.

Regardless of the Martian paleoclimate, the floods that formed putative Martian tunnel channels, eskers, kames, and other features may have analogs in Earth's Pleistocene history. Subglacial floods, some of enormous magnitude, and consequent formation of vast networks of tunnel channels, eskers, moraines, kames, and other glacial features, were common during the late evolution of the Laurentide and Fennoscandian ice sheets [*Wright*, 1973; *Shaw and Kvill*, 1984; *Attig et al.*, 1989; *Shaw*, 1989; *Shaw et al.*, 1989; *Sharpe and Cowan*, 1990; *Shoemaker* 1990, 1992; *Rains et al.*, 1993; *Brennand*, 1994; *Brennand and Shaw*, 1994]. The glaciated and isostatically depressed Hudson Bay area may have filled with subglacial water and then quickly emptied, perhaps repeatedly [*Shoemaker*, 1992], causing rapid formation of extensive glacial landscapes. Alternatively, such landscapes may have formed by rapid ice advances promoted by subglacial till streams (mobile subglacial layers of water-saturated clay- and silt-rich sediment) or deformable till beds [*Begét*, 1986; *Alley et al.*, 1989; *Clark*, 1994] (instead of or in addition to basal meltwater floods). Either way, on Mars as well as on Earth, meltwater probably is the crucial agent that modified ice-sheet behavior, thus resulting in the formation of the observed landscapes.

3.5. Alternative Glacier-Ice Compositions

The ambiguities of possible paleoclimatic implications of putative Martian glaciers as presented above are substantial. An added uncertainty pertains to the composition of the glacier ice. Hitherto, we have assumed that Martian glaciers were made of ordinary water ice. Glaciers on Earth are composed of water ice because H_2O is the only abundant atmospheric constituent that can also exist on the surface of Earth as a relatively soft and abundant solid near a state of equilibrium with the atmosphere. The fact that H_2O can also

exist as a surface or subglacial liquid is not necessary for the existence of glaciers, but liquid water has crucial implications for glacier dynamics and glacial landform genesis.

The major roles of water and ice in Martian geologic history, as presently understood, make H_2O a reasonable candidate for the major constituent of Martian glaciers. However, Mars also has abundant CO_2 in its atmosphere; like H_2O, surface condensates of CO_2 are stable in certain regions of Mars (either as a pure phase or as a mixed clathrate hydrate [*Milton*, 1974; *Clark and Mullin*, 1976]). CO_2-clathrate-hydrate glaciers are possible on Mars. Furthermore, given the pressures that are likely at the bases of hypothetical clathrate-hydrate glaciers (several tens of bars), basal clathrate decomposition to a liquid phase plus vapor (or to two immisible liquids) is also possible. Hence, CO_2-clathrate-hydrate glaciers may behave like ordinary ice glaciers in many respects, though significant differences in some of the physical properties of these two icy substances may account for some aspects of glacial terrains which differ on Earth and Mars.

3.6. Chronology of Glaciation

The NPGA modifies plains that have mapped ages of late early Hesperian, late Hesperian, and early Amazonian, thus providing an upper limit on the age of the hypothesized glacial period(s). We have confirmed Hesperian ages of two areas of NPGA on the basis of crater counts, which show typical Hesperian values of N4 (the number of craters larger than 4 km in diameter per million km^2). However, there are strong indications that these terrains were greatly modified well into the Amazonian (Figure 9). Four regions where glaciation is thought to have been an important process have N1-N1.41 (number of craters between 1.0 and 1.41 km in diameter per million km^2) that are far less than would have formed in the time since the origins of these terrains. By contrast, areas that have not been glaciated, according to our interpretation, have approximately the densities of small craters that would be expected according to their geologic ages (determined from the densities of large craters). Hence, unglaciated terrains yield almost concordant stratigraphic

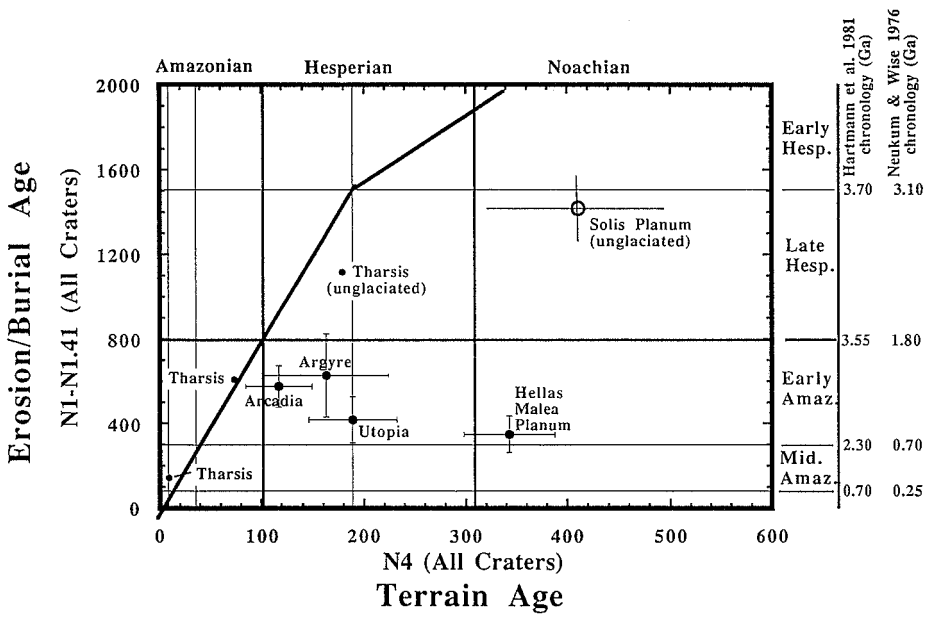


Figure 9. Crater densities (numbers of craters per unit area) and possible absolute ages of four terrains that are thought to have been glaciated (Arcadia, Utopia, and Argyre Planitiae, and Malea Planum near Hellas) and four terrains that provide no morphologic indications of ever having been glaciated (three areas of the Tharsis plateau and an area of Solis Planum). N4 is the number of craters larger than or equal to 4 km in diameter per million square kilometers and is equated with the geologic terrain age. N1-N1.41 is the number of craters between 1.0 and 1.41 ($\sqrt{2}$) km in diameter per million square kilometers; this parameter is modeled as the age of erosional or depositional modification (glaciation, if it occurred), as described in the text. "All craters" means that the graphed data do not include any attempt to morphologically classify the craters as to their state of preservation. If a terrain has not experienced a substantial erosional or depositional event, then the stratigraphic ages given by N4 and N1-N1.41 should be concordant (near the bold sloping lines, which represent the production size-frequency cratering curve). The crater production function, which changed at the end of "late heavy bombardment" (approximated here as the end of the early Hesperian), and the definition of these crater densities in terms of stratigraphic ages is from *Tanaka* [1986]. Proposed translations of crater densities into absolute ages are given on the right-hand axis.

ages for N4 and N1-N1.41, whereas putative glacial terrains yield sharply discordant ages.

A simple interpretation of Figure 9 is that some process obliterated a large fraction of small craters during the early Amazonian while preserving large ones in the four putative glacial terrains. We assume, for the sake of simplicity, that (1) the terrains of interest formed during a single discrete event, (2) a single later event erased craters smaller than 1.41 km completely but did not erase craters larger than 4 km, and (3) no other events have altered the crater size distribution. It follows that the relative model geologic age of terrain formation is given by N4 and that the model age of modification is given by N1-N1.41. The choice of these diameters is partly from the requirements that the numbers of craters must be statistically meaningful and the resolution limits of the imagery must support the diameters of craters counted (counts of craters smaller than 1 km in diameter generally would be unreliable).

The choice of diameters for craters that would and would not be susceptible to obliteration by continental glaciation can be supported on the basis of crater dimensions. Figure 10 shows the dimensions of several component features of impact craters. Obviously, large craters are more difficult to erode or bury completely than small ones. If the depth of glacial erosion and burial was typically 100-200 m, comparable to the depths of Martian tunnel channels and the heights of thumbprint ridges (and typical of the average amount of rock removed by the Pleistocene glaciations on Earth [*White*, 1972, 1988; *Bell and Laine*, 1985]), then kilometer-sized craters would generally be obliterated while craters larger than 4 km would be preserved, although in degraded form.

If a terrain has not experienced a substantial erosional or depositional event, then the stratigraphic ages given by N4 and N1-N1.41 should be concordant; otherwise N1-N1.41 should be younger than N1. The validity of these ages as lower limits depends on an assumption that other geologic processes subsequent to the putative glaciation did not erode or bury many craters. According to the cratering time scale of *Neukum and Wise* [1976], the early Amazonian spanned the period from 1800 m.y. to 700 m.y. ago; according to Figure 9, glaciation may have occurred any time during this period. Thus, glaciation seems to have occurred relatively late in Martian history, long after heavy bombardment. The early Amazonian age suggested for these widespread glacial events is consistent with the stratigraphic age of thick deposits interpreted to have a glacial origin in eastern Utopia Planitia around the western flank of Elysium [*Chapman*, 1994].

The duration of glaciation is very poorly constrained. Analogies with the Laurentide ice sheet suggest a glacially active period longer than 10^5 years. Alternatively, if the ice sheet was analogous to polar layered deposits or if they formed as the evolved remnants of frozen lakes, the ice sheet's construction and sublimation may have required tens of millions of years; however, the primary glacial landscapes formed by wet-based glaciers may have formed during a few weeks to a few years in a given area (the typical durations implied by episodes of catastrophic flooding). A very long lifetime of the ice sheets, at least millions of years, is suggested by the fact that numerous volcanic landforms thought to have been produced by volcanic interactions with an ancient ice sheet are preserved in eastern Utopia Planitia [*Chapman*, 1994]; it is improbable that all this volcanic activity would have occurred when an ice sheet was present unless it was there for a long time.

4. Conclusions and Possible Observational Tests

Many Martian terrains in the northern plains appear to have been formed by active, wet-based, continental glaciers.

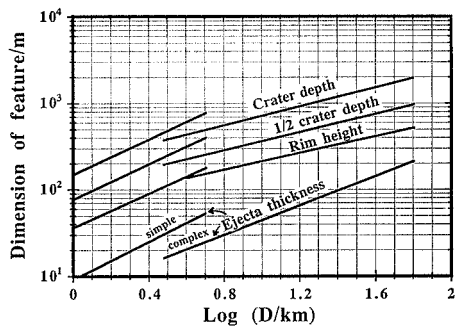


Figure 10. Depth, rim height, and average ejecta thickness of Martian impact craters versus crater diameter D (in km, log scale). Crater depth and rim height are from *Pike and Davis* [1984]. Ejecta thickness was calculated from the excavated volumes determined from the data of *Pike and Davis* [1984] and *Baker et al.* [1993].

These terrains are similar to some terrestrial glacial landscapes that formed under Pleistocene glacial climatic conditions. Such conditions are very warm by current Martian standards and may have allowed the summertime stability of liquid water on the Martian surface. Alternatively, the putative glacial landscapess could have originated at the warm bases of ultracold-surfaced glaciers or glacierized remnants of frozen seas; in this case, the observations are consistent with very cold near-surface conditions that may not have allowed the stable existence of liquid water on Mars' surface. The putative glacial landscapes in the northern hemisphere are concentrated near an elevation of -1 km and appear to onlap higher terrain. Possible glacial terrains in the southern hemisphere likewise occur in major basins (Argyre, Hellas, and the South Polar Basin) but also on nearby highlands. Southern highland glaciations may be related to regional condensation and accumulation of snow or frost from water that evaporated from ancient lakes in the major basins. In the northern hemisphere, recondensation may have tended to occur at cold northerly latitudes, where highlands are absent.

These new observations and interpretations do not allow us to confirm or reject the possibility that polar wandering may have affected the locations of ancient polar deposits on Mars [*Schultz and Lutz*, 1988]. If polar wandering was an important factor, wandering must have been rapid enough to have produced apparent paleopole positions during the early Amazonian that appear to be very widely distributed on Mars. More detailed and precise geochronologic data, preferably absolute radiometric ages, would be required to provide a more definitive constraint regarding possible polar wandering.

An early Amazonian age (recent by Martian standards) has now been established for five regions of putative glacial landscapes on Mars (Argyre, Hellas, Arcadia Planitia, western Utopia Planitia, and eastern Utopia Planitia; the latter determined by *Chapman*, 1994]). A single, widespread episode of glaciation affecting wide expanses of the northern and southern hemispheres of Mars is the simplest explanation, although the age uncertainties allow for substantial diachroneity, and there are weak indications that Mars may have experienced earlier glaciations as well [*Kargel and Strom*, 1992].

A puzzling characteristic of Mars' putative glacial landscapes is that they are so well preserved, although they must be at least hundreds of millions of years old. This and other differences of the Martian glacial terrains with respect to

terrestrial glacial terrains suggests that an unelucidated aspect of Martian glaciation was involved. The postglacial climate of Mars may have been extremely cold (perhaps similar to today's climate), thus preventing erosion; or else warm, humid conditions during the postglacial period would have to have been transient ($<10^3$-10^4 years?).

In sum, the hypothesized Martian glaciation does not, at the current level of understanding, require a unique set of paleoclimatic conditions or glacier composition. However, the capabilities of Mars missions planned for the next 10 years could allow (1) the study of the detailed morphology, sedimentology, and petrography of ancient glacial landforms and deposits, and (2) the determination of the composition, thickness, and climatic conditions of origin of the existing Martian polar caps. In so doing, we may hope to obtain the desired tight constraints on the Martian paleoclimate during the hypothesized ancient era of glaciation and deglaciation.

If terrestrial glacial landforms are any guide at all to Mars, the glacial interpretations of Mars imply certain specific characteristics of the landforms and sediment that constitute these landscapes. Unpiloted Mars space missions of the near future may contribute substantially to our understanding of Martian glaciation (Table 1). These missions will include the orbital Mars Global Surveyor spacecraft (launches in 1996 and 1998) and the Mars Surveyor landers (which will include small rovers with limited mobility, starting with Pathfinder in 1996-1997 and new launches about every 2 years). Global imaging at moderate resolution should improve coverage of the northern plains, thus allowing a more complete mapping of thumbprint terrain and putative tunnel channels and eskers. High-resolution imaging of representative areas may show some of the fine details of putative glacial features. Global laser altimetry may contribute additional knowledge of the morphology and large-scale topography of putative glacial landforms and terrains. Many putative glacial terranes and landforms are expected to be very bouldery, and this should show up in high-resolution images and as high values in maps of thermal inertia generated from thermal emission spectrometer data. The proposed Mars Polar Pathfinder lander may tell us the composition of ices and the abundance of dust in the upper meter of the existing northern Martian polar cap and the thickness of the ice cap at the landing site; hence, we

Table 1. Possible Observational Tests by Spacecraft

Putative Feature	Characteristics Not Readily Discernible in Viking Data	Types of Additional Analysis Needed; Expected Result if the Glaciation Hypothesis is Correct
Eskers	Sorted fluvial sediments; sand, gravel, boulders; cross-bedding.	MOC on MGS*: rare boulders; beds observable; esker forms. Laser altimeter on MGS: esker morphology; cross low divides. TES on MGS†: High thermal inertia, unless dust-covered. Rover: Sand, gravel, and cobbles; cross bedding coomon.
Tunnel channels	Fluted bedrock or bouldery bed. Some channels trend downhill, others uphill.	MOC on MGS: Boulders; anastomosed channels, some eskers. Laser altimeter: Downhill and uphill trends. TES on MGS: High thermal inertia, unless dust-covered. Rover: Gravelly, bouldery bed; bedrock fluted, water-polished.
End moraines and till plains	Poorly sorted and some sorted sediments; clay, silt, sand, gravel, and matrix-supported boulders, some faceted and striated; diverse lithologies. Outwash plains and channels.	MOC on MGS: many 1- to 5-m boulders; moraine morphologies; meltwater features such as small channels and outwash plains. Laser altimeter: moraine morphologies. TES on MGS: High thermal inertia, unless dust-covered. Rover imaging: Fine sediment, gravel, and matrix-supported boulders of diverse lithologies; a few percent of boulders striated and faceted; some melt-water-deposited beds and lenses possible; small diapiric structures possible. Rover micropetrography and APX‡: diverse mineralogic and chemical compositions of lithic clasts.
Glaciolacustrine plains	Mainly clayey-silt, some outsized dropstones; sediments draped over topography up to level of possible wave-cut terraces; varved deposits.	MOC on MGS: Very smooth; if wind- or water- etched, bedding may be visible and laterally extensive; iceberg- ploughed furrows and dropstone boulders possible. Laser altimeter: Very smooth, originally horizontal surface (now may be gently warped by rebound). TES on MGS: Low to moderate thermal inertia. Rover imaging: Very smooth plains; fine-grained mudstones; dropstone boulders and glaciolacustrine gravel lenses possible. Short drill core: fine laminae and varves. Returned sample: diverse mineralogic and chemical compositions of lithic clasts. Stable isotopes: Authigenic precipitates indicate equilibration near 273 K (lower if exotic ice composition).

* MOC, Mars observer camera planned for Mars Global Surveyor (MGS).
† TES, Thermal emission spectrometer on Mars Global Surveyor.
‡ APX, Alpha-proton X ray backscatter analyzer on Pathfinder and Mars Surveyor landers.

may soon know whether the present northern polar cap is likely to be flowing and, therefore, whether it constitutes a true glacial ice cap [*Fisher*, 1993].

Proof or disproof of the Mars glaciation hypothesis beyond all reasonable doubts may require a traverse to a putative esker or moraine by a rover (preferably one more mobile than those planned for Mars Surveyor) or perhaps analysis of a short drill core from a proposed glacial lake deposit. For instance, a rover could show (or not show) that a putative esker, true to this interpretation, consists of fairly well sorted fluvial sediment, rather than sand dunes, basalt flows, or some other lithology that would be predicted on the basis of another hypothesis of origin. A rover whose mission included a drive to a putative moraine or over a till plain may reveal striated facets of several percent of the boulders and cobbles therein (Table 1). This would probably be taken as reasonable proof of the theory of glaciation on Mars, at least in its bare outlines; but it would also reasonably demonstrate that the glacier was wet-based, because only wet-based glaciers are believed to commonly grind boulders at the glacier bed. However, this still would not prove that the Martian climate was warm because a normal flux of geothermal energy can cause basal melting of cold-surfaced glaciers.

Elucidation of the paleoclimatic conditions during the putative glaciation probably will require detailed micropetrographic and stable isotopic analyses of sediments, especially those deposited in glacial lakes. Very powerful evidence might include analyses of stable, radiogenic, and radioactive isotopes in evaporite or cryogenic salts and authigenic precipitates that should record their ages and temperatures of formation. Model diagenetic temperatures would have a bearing on the composition of the glaciers (ordinary water ice versus CO_2-clathrate hydrate or CO_2 ice) and the paeloclimate. Petrographic analyses of short drill cores taken from glaciolacustrine deposits (but sampled from beneath eolian dust or sand that may have accumulated) may reveal varves, which would constrain the lake processes and their duration.

Acknowledgments. We acknowledge helpful reviews from Tracy Brennand, Derrick Hirsch, Baerbel Lucchitta, Jeff Moore, David Scott, Doris Weir, and an anonymous reviewer. The development of ideas was assisted by discussions of alternative glacial and nonglacial hypotheses by many individuals, notably James Aber, Richard Alley, Mary Chapman, Chris Clark, Ken Edgett, David Fisher, David Hopkins, Steve Metzger, Jeff Moore, Tim Parker, Richard Reger, and Ken Tanaka.

References

Aber, J.S., D.G. Croot, and M.M. Fenton, *Glaciotectonic Landforms and Structures*. pp. 29-42, Kluwer Academic Publishers, Norwell, Mass., 1989.

Allen, C.C., Volcano-ice interactions on Mars, *J. Geophys. Res., 84*, 8048-8059, 1979.

Alley, R.B., D.D. Blankenship, S.T. Rooney, and C.R. Bentley, Sedimentation beneath ice shelves — The view from ice stream B, *Mar. Geol., 85*, 101-120, 1989.

Armstrong, J.E., and H.W. Tipper, Glaciation in north-central British Columbia, *Am. J. Sci., 246*, 283-310, 1948.

Attig, J.W., D.M. Mickelson, and L. Clayton, Late Wisconsin landform distribution and glacier-bed conditions in Wisconsin, *Sediment. Geol., 62*, 399-405, 1989.

Aylsworth, J.M., and W.W. Shilts, Bedforms of the Keewatin ice sheet, Canada, *Sediment Geol., 62*, 407-428, 1989.

Baker, V.R., R.G. Strom, V.C. Gulick, J.S. Kargel, G. Komatsu, and V.S. Kale, Ancient oceans, ice sheets, and the hydrological cycle on Mars, *Nature, 352*, 589-594, 1991.

Baker, V.R., V.C. Gulick, and J.S. Kargel, Water resources and hydrogeology of Mars, in *Resources of Near-Earth Space*, edited by J. Lewis, M.S. Matthews, and M.L. Guerriri, pp. 765-798, University of Arizona Press, Tucson, 1993.

Barnett, D.M., and G. Holdsworth, Origin, morphology, and chronology of sublacustrine moraines, Generator Lake, Baffin Island, Northwest Territories, Canada, *Can. J. Earth Sci., 11*, 380-408, 1974.

Battistini, R., Indices morphologiques d'anciens glaciers dans les des basses latitudes de Mars: leur relation avec l'hydrolithosphère, *Internord, 18*, 399-411, 1987.

Beaudry, L.M., and G. Prichonnet, Late glacial De Geer moraines with glaciofluvial sediment in the Chapais area, Québec (Canada), *Boreas, 20*, 377-394, 1991.

Begét, J.E., Modelling the influence of till rheology on Pleistocene glacier flow and glacier profiles, southern Laurentide ice sheet, *J. Glaciol., 32*, 235-241, 1986.

Bell, M., and E.P. Laine, Erosion of the Laurentide region of North America by glacial and glaciofluvial processes, *Quat. Res., 23*, 154-174, 1985.

Boulton, G.S., A theory of drumlin formation by subglacial sediment deformation, in *Drumlin Symposium*, edited by J. Menzies and J. Rose), pp. 25-80, A.A. Balkema, Rotterdam, 1987.

Boulton, G.S., and C.D. Clark, A highly mobile Laurentide ice sheet revealed by satellite images of glacial lineations, *Nature, 346*, 813-817, 1990.

Boyce, J.I., and N. Eyles, Drumlins carved by deforming till streams below the Laurentide ice sheet, *Geology, 19*, 787-790, 1991.

Boyd, R., D.B. Scott, and M. Douma, Glacial tunnel valleys and Quaternary history of the outer Scotian shelf, *Nature, 333*, 61-64, 1988.

Brennand, T.A., Macroforms, large bedforms and rythmic sedimentary sequences in subglacial eskers, south-central Ontario: implications for esker genesis and meltwater regime, *Sediment. Geol., 91*, 9-56, 1994.

Brennand, T.A., and D.R. Sharpe, Ice-sheet dynamics and subglacial meltwater regime inferred from form and sedimentology of glaciofluvial systems: Victoria Island, District of Franklin, Northwest Territories, *Can. J. Earth Sci., 30*, 928-944, 1993.

Brennand, T.A., and J. Shaw, Tunnel channels and associated landforms, south-central Ontario: Their implications for ice-sheet hydrology, *Can. J. Earth Sci., 31*, 505-522, 1994.

Budd, W.F., D. Jensen, J.H.I. Leach, I.N. Smith, and U. Radok, The north polar ice cap of Mars as a steady-state system, *Polarforschung, 56*, 43-63, 1986.

Carr, M.H., Mars: A water-rich planet? *Icarus, 56*, 187-216, 1986.

Carr, M.H., and J.G. Schaber, Martian permafrost features, *J. Geophys. Res., 82*, 4039-4054, 1977.

Carr, M.H. (ed.), The geology of the terrestrial planets, NASA Spec. Publ. SP-469, 230, 1984.

Chapman, M.G., Evidence, age, and thickness of a frozen paleolake in Utopia Planitia, Mars, *Icarus, 109*, 393-406, 1994.

Christensen, P.R., Lahars in the Elysium region of Mars, *Geology, 17*, 203-206, 1989.

Christensen, P.R., and H.J. Moore, The Martian surface layer, in: *Mars*, edited by Kieffer, H.H., B.M. Jakosky, C.W. Snyder, and M.S. Matthews, pp. 686-729, University of Arizona Press, Tucson, 1992.

Christiansen, E.A., Verendrye valley and the Glidden esker, Saskatchewan: Subglacial and ice-walled features in southeastern Saskatchewan, Canada, *Can. J. Earth Sci., 24*, 170-176, 1987.

Clark, B.R., and R.P. Mullin, Martian glaciation and the flow of solid CO_2, *Icarus, 27*, 215-228, 1976.

Clark, C.D., Large-scale ice-moulding: a discussion of genesis and glaciological significance, *Sediment. Geol., 91*, 253-254, 1994.

Costard, F.M., The spatial distribution of volatiles in the Martian hydrolithosphere, *Earth, Moon, and Planets, 45*, 265-290, 1989.

Costard, F.M., and J.S. Kargel, Outwash plains and thermokarst on Mars, *Icarus*, 1994 (in press).

Edgett, K.S., and P.R. Christensen, The particle size of Martian aeolian dunes, *J. Geophys. Res., 96*, 22,765-22,776, 1991.

Ehlers, J., Some aspects of glacial erosion and deposition in north Germany, *Ann. Glaciol., 2*, 143-146, 1981.

Ehlers, J., and G. Linke, The origin of deep buried channels of Elsterian age in northwest Germany, *J. Quat. Sci., 4*, 255-265, 1989.

Fanale, F.P., S.E. Postawko, J.B. Pollack, M.H. Carr, and R.O. Pepin, Mars: Epochal climate change and volatile history, in *Mars*, edited by H.H. Kieffer, B.M. Jakosky, C.W. Snyder, and M.S. Matthews, pp. 1135-1179, The University of Arizona Press, Tucson, 1992.

Farrand, W.R., The Quaternary history of Lake Superior, *Proc. Conf. Great Lakes Res. 12th*, 181-197, 1969.

Fisher, D.A., If Martian ice caps flow: Ablation mechanisms and appearance, *Icarus, 105*, 501-511, 1993.

Fowler, C.M.R., *The Solid Earth*, p. 234, Cambridge University Press, New York, 1990.

Fyfe, G.F., The effect of water depth on ice-proximal glaciolacustrine sedimentation: Salpausselkä I, southern Finland, *Boreas, 19*, 147-164, 1990.

Geological Sopciety of America, *Glacial Map of the United States east of the Rocky Mountains*, 1:750,000, 2 sheets, 1959.

Hartmann, W.K., et al., Chronology of planetary volcanism by comparative studies of planetary cratering, in *Basaltic Volcanism on the Terrestrial Planets*, pp. 1049-1127, Pergamon, New York, 1981.

Hebrand, M., and M. Åmark, Esker formation and glacier dynamics in eastern Skåne and adjacent areas, southern Sweden, *Boreas, 18*, 67-81, 1989.

Hodges, C.A., and H.J. Moore, The subglacial birth of Olympus Mons and its aureoles, *J. Geophys. Res., 84*, 8061-8074, 1979.

Hoppe, G., Problems of glacial morphology and the ice age, *Geogr. Ann., XXXIX*, 1-18, 1957.

Hoppe, G., Glacial morphology and inland ice recession in northern Sweden, *Int. Geogr. Congr., XIX*, 193-212, 1960.

Jakosky, B.M., and R.M. Haberle, The seasonal behavior of water on Mars, in *Mars*, edited by H.H. Kieffer, B.M. Jakosky, C.W. Snyder, and M.S. Mathhews, pp. 969-1016, University of Arizona Press, Tucson, 1992.

Jöns, H.-P., Der Mars, *Geogr. Rundsch., 43*, 98-109, 1991.

Kargel, J.S., Geomorphic processes in the Argyre-Dorsa Argentea region of Mars (abstract), *Lunar Planet. Sci., XXIV*, 753-754, 1993.

Kargel, J.S., and R.G. Strom, Ancient glaciation on Mars (abstract), *Lunar Planet. Sci., XXI*, 597-598, 1990.

Kargel, J.S., and R.G. Strom, Ancient glaciation on Mars, *Geology, 20*, 3-7, 1992.

Kargel, J.S., R.G. Strom, and N. Johnson, Glacial geology of the Hellas region of Mars (abstract), *Lunar Planet Sci. , XXII*, 687-688, 1991.

Kargel, J., R. Strom, J. Lockwood, and J. Shaw, Subglacial and glaciomarine processes in the Martian northern plains (abstract), *Lunar Planet Sci., XXIII*, 657-658, 1992.

Kargel, J.S., and F.M. Costard, 1993, Possible occurrence and origin of massive ice in Utopia Planitia (abstract), In *Workshop on the Martian Northern Plains: Sedimentological, Periglacial, and Paleoclimatic Evolution, LPI Tech. Rep. 93-04*, pp. 7-8, edited by J.S. Kargel, J. Moore, and T. Parker, Workshop held Aug. 12-14, 1993, in Fairbanks, Alaska. Lunar and Planetary Institute, Houston, 1993.

Kieffer, H.H., and A.P. Zent, Quasi-periodic climate change on Mars, in *Mars*, edited by H.H. Kieffer, B.M. Jakosky, C.W. Snyder, and M.S. Matthews, pp. 1180-1218, The University of Arizona Press, Tucson, 1992.

Krüger, J., Glacial morphology and deposits in Denmark, in *Glacial Deposits in Northwest Europe*, edited by J. Ehlers, pp. 181-191, A.A. Balkema, Rotterdam, 1983.

Lockwood, J.F., J.S. Kargel, and R.G. Strom, Thumbprint terrain on the northern plains: A glacial hypothesis (abstract), *Lunar Planet. Sci., XXIII*, 795-796, 1992.

Lucchitta, B.K., Mars and Earth: Comparison of cold-climate features, *Icarus, 45*, 264-303, 1981.

Lucchitta, B.K., Ice sculpture in the Martian outflow channels, *J. Geophys. Res.,87*, 9951-9973, 1982.

Lucchitta, B.K., H.M. Ferguson, and C. Summers, Sedimentary depsosits in the northern lowlands plains, Mars, *Proc. Lunar Planet. Sci. Conf. 17th, Part 1, J. Geophys. Res., 91, suppl.*, E166-E174, 1986.

Lucchitta, B.K., Ice in the northern plains: relic of a frozen ocean? In *Workshop on the Martian Northern Plains: Sedimentological, Periglacial, and Paleoclimatic Evolution, LPI Tech. Rep. 93-04.*, edited by J.S. Kargel, J. Moore, and T. Parker, pp. 9-10, Workshop held Aug. 12-14, 1993, in Fairbanks, Alaska, Lunar and Planetary Institute, Houston, 1993.

Lundqvist, J., Moraine morphology: Terminological remarks and regional aspects, *Geogr. Ann., 63A*, 127-138, 1981.

Mellon, M.T., and B.M. Jakosky, The distribution of ground ice on Mars, In *Workshop on the Martian Northern Plains: Sedimentological, Periglacial, and Paleoclimatic Evolution, LPI Tech. Rep. 93-04*, edited by J.S. Kargel, J. Moore, and T. Parker, p. 12, workshop held Aug. 12-14, 1993, in Fairbanks, Alaska, Lunar and Planetary Institute, Houston, 1993.

Metzger, S.M., The eskers of New York state: Formation process implications and esker-like features on the planet Mars (abstract), *Lunar Planet. Sci, XXIII*, 901-902, 1992.

Milton, D.J., Carbon dioxide hydrate and floods on Mars, *Science, 183*, 654-656, 1974.

Mooers, H.D., A glacial-process model: The role of spatial and temporal variations in glacier thermal regime, *Geol. Soc. Am. Bull., 102*, 243-251, 1990.

Mouginis-Mark, P.J., Volcano-ground ice interactions in Elysium Planitia, *Icarus, 64*, 265-284, 1985.

Neukum, G., and D.U. Wise, Mars: A standard crater curve and possible new time scale, *Science, 194*, 1381-1387, 1976.

Oswald, G.K.A., Investigation of sub-ice bedrock characteristics by radio-echo sounding, *J. Glaciology, 15*, 75-87, 1975.

Parker, T.J., R.S. Saunders, and D.M. Schneeberger, Transitional morphology in the west Deuteronilus Mensae region of Mars: Implications for modification of the lowland/upland boundary, *Icarus, 82*, 111-145, 1989.

Parker, T.J., D.S. Gorsline, R.S. Saunders, D.C. Pieri, D.C., and D.M. Schneeberger, Coastal geomorphology of the Martian northern plains, *J. Geophys. Res., 98*, 11,061-11,078, 1993.

Péwé, T.L., and R.D. Reger, *Richardson and Glenn Highways, Alaska: Guidebook to Permafrost and Quaternary Geology, reprinted ed.*, 263 pp., State of Alaska, Department of Natural Resources, Division of Geological and Geophysical Surveys, Fairbanks, 1993.

Pike, R.J., and P.A. Davis, Toward a topographic model of Martian craters from photoclinometry (abstract), *Lunar Planet. Sci., XV*, 645-646, 1984.

Rains, B., J. Shaw, R. Skoye, D. Sjogren, and D. Kvill, Late Wisconsin subglacial megaflood paths in Alberta, *Geology, 21*, 323-326, 1993.

Rossbacher, L.A., and S. Judson, Ground ice on Mars: Inventory, distribution, and resulting landforms, *Icarus, 45*, 39-59, 1981.

Ruff, S.W., Comparison of Mars sinuous ridges with terrestrial linear dunes: Observations from the field, *Lunar Planet. Sci., XXV*, 1171-1172, 1994.

Schultz, P.H., and A.B. Lutz, Polar wandering of Mars, *Icarus, 73*, 91-141, 1988.

Scott, D.H., Meander relics: Direct evidence of extensive flooding on Mars, in *Conference on Planetary Volatiles, LPI Tech. Rep. 83-01*, edited by R. Pepin and R. O'Connell, pp.157-159, 1983.

Scott, D.H. and J.R. Underwood Jr., Mottled terrain: A continuing Martian enigma, *Proc. Lunar Planet.Sci. Conf. 21st*, 627-634, 1991.

Scott, D.H., M.G. Chapman, J.W. Rice, Jr., and J.M. Dohm, New evidence of lacustrine basins on Mars: Amazonia and Utopia Planitiae, *Proc. 22nd Lunar and Planet. Sci. Conf.*, Lunar and Planetary Institute, 53-62, 1992.

Sharpe, D.R., and W.R. Cowan, Moraine formation in northwestern Ontario: Product of subglacial fluvial and glaciolacustrine sedimentation, *Can. J. Earth Sci., 27*, 1478-1486, 1990.

Shaw, J., Drumlins, subglacial meltwater floods, and ocean responses, *Geology, 17*, 853-856, 1989.

Shaw, J., and G. Gorrell, Subglacially formed dunes with bimodal and graded gravel in the Trenton Drumlin Field, Ontario, *Geogr. phys. Quat., 45*, 21-34, 1991.

Shaw, J., and T.R. Healy, The formation of the Labyrinth, Wright Valley, Antarctica, *N.Z. J. Geol. Geophys., 20*, 933-947, 1977.

Shaw, J., and D. Kvill, A glaciofluvial origin for drumlins of the Livingstone Lake area, Saskatchewan, *Can. J. Earth Sci., 21*, 1442-1459, 1984.

Shaw, J., D. Kvill, and B. Rains, Drumlins and catastrophic subglacial floods, *Sediment. Geol., 2*, 177-202, 1989.

Shaw, J., J.S. Kargel, and R.G. Strom, Terrestrial subglacial landforms as analogs for Martian landscapes (abstract), *Lunar Planet. Sci., XXIII*, 1273-1274, 1992.

Shoemaker, E.M., Subglacial floods and the formation of low-relief ice-sheet lobes, *J. Glaciol., 38*, 105-112, 1990.

Shoemaker, E.M., Water sheet outburst floods from the Laurentide ice sheet, *Can. J. Earth Sci., 29*, 1250-1264, 1992.

Sorensen, R., Glacial deposits in the Oslofjord area, in *Glacial deposits in North-west Europe*, edited by J. Ehlers, pp. 19-40, A.A. Balkema, Rotterdam, 1983.

Squyres, S.W., and M.H. Carr, Geomorphic evidence for the distribution of ground ice on Mars, *Science, 231*, 249-252, 1986.

Squyres, S.W., D.E. Wilhelms, and A.C. Moosman, Large-scale volcano-ground ice interactions on Mars, *Icarus, 70*, 385-408, 1987.

Stevenson, D.J., T. Spohn, and G. Schubert, Magnetism and thermal evolution of the terrestrial planets, *Icarus, 54*, 466-489, 1983.

Tanaka, K.L., The stratigraphy of Mars, *Proc. Lunar Planet. Sci. Conf. 17th, Part 1, J. Geophys. Res., 91, suppl.*, E139-E158, 1986.

Teller, J.T., and L. Clayton, L. (eds.), *Glacial Lake Agassiz, Geol. Assoc. Can. Spec. Pap., 26*, 1-15, 1983.

Thomas, P., S. Squyres, K. Herkenhoff, A. Howard, and B. Murray, Polar deposits of Mars, *in Mars*, edited by H.H. Kieffer, B.M. Jakosky, C.W. Snyder, and M.S. Matthews, pp. 767-798, University of Arizona Press, Tucson, 1992.

Tipper, H.W., Glacial geomorphology and Pleistocene history of central British Columbia, *Geol. Surv. Can.Bull., 196*, 1-89, 1971.

Tricart, J.L.F., Environmental change of planet Mars demsonstrated by landforms, *Z. Geomorph., 32*, 385-407, 1988.

Turcotte, D.L., and G. Schubert, *Geodynamics: Applications of Continuum Physics to Geological Problems*, 450 pp., John Wiley, New York, 1982.

White, W.A., Deep erosion by continental ice sheets, *Geol. Soc. Am. Bull., 83*, 1037-1056, 1972.

White, W.A., More on deep glacial erosion by continental ice sheets and their tongues of distributary ice, *Quat. Res., 30*, 137-150, 1988.

Wright, H.E., Jr., Tunnel valleys, glacier surges, and subglacial hydrology of the Superior lobe, Minnesota, *Geol. Soc. Am. Mem., 136*, 251-276, 1973.

Wright, H.E., and D.G. Frey (eds.), *The Quaternary of the United States*, pp. 15-128, Princeton University Press, Princeton, 1965.

Zilliacus, H., Genesis of De Geer moraines in Finland, *Sediment. Geol., 62*, 309-317, 1989.

Zoltai, S.C., Glacial features of the Quetico-Nipigon area, Ontario, *Can. J. Earth Sci.*, 2, 247-269, 1965.

Zurek, R.W., J.R. Barnes, R.M. Haberle, J.B. Pollack, J.E. Tillman, and C.B. Leovy, Dynamics of the atmosphere of Mars, in *Mars*, edited by H.H. Kieffer, B.M. Jakosky, C.W. Snyder, and M.S. Matthews, pp. 835-933, University of Arizona Press, Tucson, 1992.

Victor R. Baker, Department of Geosciences, University of Arizona, Tucson, AZ 85721.

James E. Begét, Department of Geology and Geophysics, University of Alaska, Fairbanks, AK 99775.

Jeffrey S. Kargel, U.S. Geological Survey, 2255 N. Gemini Dr., Flagstaff, AZ 86001. (email: JKargel@Iflag2.wr.usgs. gov)

Jeffrey F. Lockwood, Sahuaro High School, Tucson, AZ.

Troy L. Péwé, Department of Geology, Arizona State University, Tempe, AZ 85287.

John S. Shaw, Department of Geography, University of Alberta, Edmonton, Alberta, T6G 2H4, Canada.

Robert G. Strom, Lunar and Planetary Laboratory, University of Arizona, Tucson, AZ 85721.

(Received May 6, 1994; revised September 1, 1994; accepted September 19, 1994.)

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5369–5379, MARCH 25, 1995

Low-temperature and low atmospheric pressure infrared reflectance spectroscopy of Mars soil analog materials

Janice L. Bishop[1]

Departments of Chemistry and Geological Sciences, Brown University, Providence, Rhode Island

Carlé M. Pieters

Department of Geological Sciences, Brown University, Providence, Rhode Island

Abstract. Infrared reflectance spectra of carefully selected Mars soil analog materials have been measured under low atmospheric pressures and temperatures. Chemically altered montmorillonites containing ferrihydrite and hydrated ferric sulfate complexes are examined, as well as synthetic ferrihydrite and a palagonitic soil from Haleakala, Maui. Reflectance spectra of these analog materials exhibit subtle visible to near-infrared features, which are indicative of nanophase ferric oxides or oxyhydroxides and are similar to features observed in the spectra of the bright regions of Mars. Infrared reflectance spectra of these analogs include hydration features due to structural OH, bound H_2O, and adsorbed H_2O. The spectral character of these hydration features is highly dependent on the sample environment and on the nature of the H_2O/OH in the analogs. The behavior of the hydration features near 1.9 μm, 2.2 μm, 2.7 μm, 3 μm, and 6 μm are reported here in spectra measured under a Marslike atmospheric environment. In spectra of these analogs measured under dry Earth atmospheric conditions the 1.9-μm band depth is 8-17%; this band is much stronger under moist conditions. Under Marslike atmospheric conditions the 1.9-μm feature is broad and barely discernible (1-3% band depth) in spectra of the ferrihydrite and palagonitic soil samples. In comparable spectra of the ferric sulfate-bearing montmorillonite the 1.9-μm feature is also broad, but stronger (6% band depth). In the low atmospheric pressure and temperature spectra of the ferrihydrite-bearing montmorillonite this feature is sharper than the other analogs and relatively stronger (6% band depth). Although the intensity of the 3-μm band is weaker in spectra of each of the analogs when measured under Marslike conditions, the 3-μm band remains a dominant feature and is especially broad in spectra of the ferrihydrite and palagonitic soil. The structural OH features observed in these materials at 2.2-2.3 μm and 2.75 μm remain largely unaffected by the environmental conditions. A shift in the Christiansen feature towards shorter wavelengths has also been observed with decreasing atmospheric pressure and temperature in the midinfrared spectra of these samples.

1. Introduction

Spectral experiments on particulate, ferric-containing analogs to Martian soil have been performed to facilitate interpretation of remotely acquired spectra. These analog materials include chemically altered montmorillonites that contain nanophase ferric oxyhydroxides or ferric sulfate complexes, a palagonitic soil and ferrihydrite. Since reflectance spectra of such minerals and soils are environmentally sensitive, accurate spectral measurements of these samples as Mars analogs requires measurement under Marslike atmospheric conditions. The results presented here encompass spectral experiments of laboratory analog materials to Martian bright region soils measured in ar environmental chamber under controlled atmospheric and temperature conditions. As the atmospheric pressure and temperature are reduced in these experiments the partial pressure of H_2O in the sample chamber is decreased and the samples are dehydrated.

The presence of smectites in bright region surface material on Mars has been suggested based on analyses of chemical measurements from Viking [*Toulmin et al.*, 1977; *Baird et al.*, 1977] and is further supported by identification of smectites in some of the SNC meteorites [*Gooding et al.*, 1991; *Treiman et al.*, 1993]. Spectroscopic analyses in the visible to near infrared have shown that chemically altered, ferric-enriched smectites prepared in the laboratory (e.g., ferrihydrite-bearing montmorillonites) exhibit important similarities to the soils on Mars [*Bishop et al.*, 1993a; *Bishop*, 1994]. Visible absorption bands centered near 0.91 μm are found in ferrihydrite and ferrihydrite-bearing montmorillonites (J. L. Bishop et al., Reflectance spectroscopy of ferric sulfate-bearing montmorillonites as Mars soil analog materials, submitted to *Icarus*, 1994; hereinafter referred to as submitted manuscript) and also in the bright soils of Arabia on Mars [*McCord et al.*, 1982; *Murchie et al.*, 1993a]. Montmorillonites containing hydrated ferric sulfate

[1]Now at Deutsche Forschungsanstalt für Luft- und Raumfahrt, Berlin, Germany.

Paper number 94JE03331.
0148-0227/95/94JE-03331 $05.00

complexes exhibit a ferric absorption feature near 0.88 μm (J. L. Bishop et al., submitted manuscript, 1994), which is also observed in the surface material at Lunae Planum [*McCord et al.*, 1977; *Murchie et al.*, 1993b]. Geochemical and spectral analyses of palagonitic soils indicate that they may have formed on Mars as well [*Allen et al.*, 1981; *Singer*, 1982]. Spectral measurements of palagonitic soils indicate that these materials have a ferric band centered near 0.86 μm and a spectral brightness of about 30-40% reflectance, near that of soils on Mars in the bright regions [*Singer*, 1982; *Morris et al.*, 1990].

Smectites, ferrihydrite, and palagonitic soils all contain structural OH bonded to metal cations, adsorbed H_2O, and bound H_2O. The infrared absorption features due to OH and H_2O have been measured under variable environmental conditions in smectites [*Prost*, 1975; *Bruckenthal*, 1987; *Coyne et al.*, 1990a,b; *Johnston et al.*, 1992; *Bishop et al.*, 1994] and in palagonitic soils [*Bruckenthal*, 1987]. Preliminary experiments involving a variety of Mars analogs under differing environmental conditions examined the influence of exposure history on water content and absorption features due to H_2O [*Bishop et al.*, 1993b]. These experiments showed that reflectance spectra measured under dry conditions of ferrihydrite, ferric sulfate-bearing montmorillonite, and palagonitic soils contain sufficient bound H_2O to retain a strong 3-μm band similar to that observed in spectra of Mars. Spectra measured under similar conditions of naturally occurring smectites (on Earth) and also many ferric-bearing smectites exhibit a weaker 3-μm band.

Low-temperature reflectance spectra of hydrated minerals have exhibited changes in the band shape of the hydration features [*Clark*, 1981a]. Reflectance spectra of Mars soil analog materials dehydrated under reduced atmospheric pressures have also shown variations in the spectral character of the hydration features [*Bishop et al.*, 1993b]. In the low atmospheric pressure experiments, as well as in some of the low-temperature experiments, these changes are due to dehydration of the sample. Additional experiments involving low-pressure and low-temperature reflectance spectra found changes in the hydration features and in the Christiansen feature of Mars soil analog materials [*Bishop and Pieters*, 1994]. The effects of the environment on water in clays, oxides, and salts are manifested in the spectral hydration bands measured via reflectance spectroscopy. Environmental influences on the spectral hydration features are especially important in interpreting observations of Mars where subtle spatial variations in the strengths of cation-OH and H_2O absorptions have been observed in telescopic spectra [*Bell and Crisp*, 1993] and in spectra measured by the French near-infrared imaging spectrometer ISM on the Phobos 2 spacecraft [*Bibring et al.*, 1990; *Arnold*, 1992; *Murchie et al.*, 1992, 1993a]. For these reasons, reflectance spectra are measured in this study of Mars analog materials under the pressure and temperature conditions of temperate regions on Mars.

2. Background

2.1. Spectral Properties of Mars

Soderblom [1992] and *Roush et al.* [1993] have recently produced reviews of the spectroscopic properties of the surface of Mars. The spectral properties of the Martian atmosphere [e.g., *Rosenqvist et al.*, 1992], H_2O ice in the north polar cap [e.g., *Clark and McCord*, 1982], and CO_2 ice in the south polar cap [e.g., *Calvin*, 1990] have also been studied. Reflectance spectra of the bright areas on Mars are characterized by ferric crystal field transitions in the extended visible region [*Adams and McCord*, 1969; *Singer et al.*, 1979; *Sherman et al.*, 1982; *Bell et al.*, 1990]. These spectra include absorptions near 0.5 μm and 0.9 μm, an inflection near 0.6 μm, and a reflectance maximum at 0.7-0.8 μm. The absorption features in the spectra of bright regions on Mars have been assigned to octahedrally coordinated Fe^{3+} transitions [*Sherman et al.*, 1982; *Burns*, 1993] and exhibit band centers ranging from 0.85 to 0.92 μm [*McCord et al.*, 1977; *Singer*, 1982; *Bell et al.*, 1990; *Murchie et al.*, 1993a,b]. This implies heterogeneous ferric-containing phases or possibly heterogeneous ferric mineralogy in the Martian bright regions.

The shape of the 3-μm absorption band indicates a prevalence of bound molecular H_2O in Martian surface materials [*Houck et al.*, 1973; *Singer et al.*, 1990; *Blaney and McCord*, 1989; *Erard et al.*, 1991]. Since CO_2 in the Martian atmosphere obscures the 2.7-2.8 μm region, features due to structural OH species cannot be detected in currently available spectra of the surface materials. A weak 2.2-μm feature, characteristic of structural OH in aluminous clays, has been observed in telescopic spectra of Mars [*Bell and Crisp*, 1993] and in ISM spectra of some areas on Mars [*Arnold*, 1992; *Murchie et al.*, 1993a,b]. This suggests that aluminous clay silicates commonly found on Earth occur in some of the bright regions of Mars as well. Variation in the strengths of the 2.2-μm and 3-μm features in spectra of bright and dark regions on Mars [*Murchie et al.*, 1993a] indicates heterogeneity in the type, crystallinity or quantity of silicates present in the surface material, as well as variations in the water content.

2.2. Atmospheric Properties of Mars

Surface temperatures on Mars span the range 130 to 290 K [*Kieffer et al.*, 1977]. Spring temperatures vary from 170 to 230 K in north temperate latitudes, and from 200 to 250 K in south temperate latitudes [*Davies*, 1979]. Atmospheric pressures on the surface of Mars were measured by the Viking landers and vary from 680 Pa (6.7 mbar) to 1000 Pa (9.9 mbar) [*Hess et al.*, 1979]. From mass spectrometry on the Viking landers, *Owen et al.* [1977] report a $CO_2(g)$ abundance in the Martian atmosphere near 95%. Climatic models indicate that atmospheric CO_2 pressures vary from 0.1 to 15 mbar [*Fanale et al.*, 1982]. Atmospheric water vapor on Mars varies diurnally and regionally across the planet from about 1 to 100 perceptible microns [*Farmer et al.*, 1977]. This water vapor abundance indicates the depth of water precipitation that would result if all of the water in the atmosphere precipitated on the surface [*Carr*, 1981].

2.3. Spectral Experiments Under Low Temperatures

Clark [1981a] presented reflectance spectra of montmorillonites mixed with ice at low-temperatures. A decrease in the intensity of the 1.4-μm and 1.9-μm features and a decrease in the albedo were observed as the temperature was lowered [*Clark*, 1981a]. This is explained by low-temperature adsorption isotherms for montmorillonite, which showed decreases in the amount of adsorbed water under reduced temperatures [*Anderson et al.*, 1967]. Reflectance experiments

involving H_2O or CO_2 frosts have observed temperature dependent and grain size dependent spectral behavior [*Kieffer*, 1970; *Clark*, 1981b; *Lucey and Clark*, 1985]. The bandwidths and intensities of the spectral features varied with grain size and temperature, but the band positions remained constant in these studies.

2.4. Spectral Properties of Mars Analogs

Reflectance spectra are shown in Figure 1 for a group of Mars soil analogs from 0.3 to 3.5 μm. These spectra of ferrihydrite, ferrihydrite-bearing montmorillonite, ferric sulfate-bearing montmorillonite and a palagonitic soil from Haleakala, Maui, were measured using a biconical configuration in the Nicolet Fourier transform interferometer (FTIR) under H_2O and CO_2-purged conditions from 1.2 to 3.5 μm, then scaled to bidirectional spectra measured from 0.3 to 1.2 μm using the reflectance experiment laboratory (RELAB) spectrometer at Brown University. Composite observational spectra of Olympus Mons, Mars [*Mustard and Bell*, 1994] are shown in Figure 1 for comparison. The telescopic spectrum was scaled to the ISM spectrum at 0.77 μm, and a model was used to remove the strong atmospheric bands. Absorption features due to OH and H_2O are observed in these spectra near 1.4, 1.9, 2.2 or 2.3, 2.75, and 2.8-3.1 μm. The visible and near-infrared spectral features of the altered montmorillonites and ferrihydrite are described in detail by *Bishop et al.* [1993a, submitted manuscript, 1994] and of palagonitic soils by *Singer* [1982], *Bruckenthal* [1987], *Morris et al.* [1990] and *Golden et al.* [1993].

2.4.1. Ferrihydrite. Ferrihydrite is a nanophase, reddish brown ferric oxyhydroxide. The structure contains Fe^{3+} in tetrahedral and octahedral sites bound to O, OH, or H_2O, and the formula $Fe_{4.3}(O,OH,H_2O)_{12}$ has been assigned based on X ray powder diffraction [*Eggleton and Fitzpatrick*, 1988]. Reflectance spectra of ferrihydrite exhibit subtle ferric absorption features in the extended visible region including a reflectance maximum at 0.80 μm and a reflectance minimum at 0.92-0.93 μm [*Bishop et al.*, 1993a]. Infrared spectra of ferrihydrite include weak features due to structural OH at 1.41 μm, 2.3-2.4 μm, and 2.75 μm, while stronger and broader features due to water are present near 1.95 μm and 3 μm [*Bishop et al.*, 1993a].

2.4.2. Ferric-bearing montmorillonites. Montmorillonite belongs to the smectite mineral group, which is characterized spectrally by features due to adsorbed H_2O, bound H_2O, and structural OH. Chemisorbed water in smectites is bound to the interlayer surfaces or cations and is a necessary part of the smectite structure, while physisorbed water is adsorbed on the interlayer or grain surfaces with longer (looser) bonds at lower energies (longer wavelengths). This chemisorbed and physisorbed water will be informally referred to as "bound" and "adsorbed" water. The term "structural OH" indicates the OH molecules octahedrally coordinated to Al and other cations in the smectite layers.

Spectral features in smectites due to bound water are observed at 1.41 μm, 1.91 μm, ~2.9 μm, ~3.1 μm, and near 6 μm [*Prost*, 1975; *Cariati et al.*, 1983; *Bishop et al.*, 1994], due to adsorbed water at 1.46 μm, 1.97 μm, ~3.0 μm, and near 6 μm [*Prost*, 1975; *Cariati et al.*, 1983; *Bishop et al.*, 1994], and due to structural OH at 1.41 μm, 2.2-2.35 μm, 2.75 μm, and 10-12 μm [*Farmer*, 1974; *Bishop et al.*, 1994]. Montmorillonites doped with ferric salts (e.g., ferrihydrite-bearing and ferric sulfate-bearing montmorillonite) have shown ferric absorptions near 0.5 μm and near 0.9 μm, that are dependent on the chemical conditions of the doping procedure [*Banin et al.*, 1993; *Bishop et al.*, 1993; J. L. Bishop et al., submitted manuscript, 1994]. The spectrum of ferrihydrite-bearing montmorillonite (Figure 1) exhibits a ferric

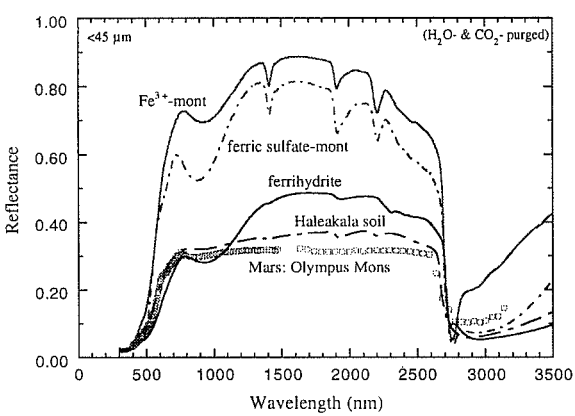


Figure 1. Reflectance spectra are shown from 0.3 to 3.5 μm for Mars analogs and Mars. The spectrum of Olympus Mons, Mars, is a composite spectrum created from ISM and telescopic data [*Mustard and Bell*, 1994]. The reflectance spectra of the ferrihydrite, ferrihydrite-bearing montmorillonite, ferric sulfate-bearing montmorillonite and a palagonitic soil from Haleakala, Maui, were all measured under H_2O- and CO_2- purged conditions at 23°C.

absorption band centered near 0.92 μm, a reflectance maximum near 0.75 μm and a subtle shoulder near 0.6 μm. The ferric sulfate-bearing montmorillonite spectrum (Figure 1) exhibits a ferric absorption band at 0.88 μm, a reflectance maximum at 0.70 μm, and a more pronounced shoulder near 0.6 μm.

2.4.3. Palagonitic soil. Palagonite is defined as weathered basaltic volcanic glass [*Bouška*, 1993] and is formed via heating, hydration, and devitrification of this glass [*Bates and Jackson*, 1984]. During palagonitization the ferric to ferrous iron ratio increases and the formation of smectites and zeolites is common [*Bouška*, 1993]. Based on the thermodynamics of weathering processes, *Gooding and Keil* [1978] suggested that basaltic glass may have weathered to form clay-rich products on Mars. Palagonitic soils have been proposed as analogs to bright soils on Mars based on basaltic elemental abundances and spectral properties [e.g., *Allen et al.*, 1981; *Singer*, 1982]. Mineralogical studies of variably palagonitized basaltic tephra from Mauna Kea, Hawaii, have found smectites and iron oxides in the most altered fine-grained particles of several samples [*Golden et al.*, 1993].

Reflectance spectra of palagonitic soil exhibit subtle ferric absorption features in the extended visible region that are highly dependent on the nature of the sample. Features such as a broad shallow absorption band near 0.86 μm, a reflectance maximum near 0.75 μm, and an inflection near 0.6 μm are observed in the fine particle size fractions of many samples [*Singer*, 1982; *Morris et al.*, 1990; *Golden et al.*, 1993]. Infrared spectra of palagonitic soils vary depending on the composition and crystallinity of the sample, but generally contain broad features near 1.9 μm and 3 μm, as well as one or more weak features in the range 2.2-2.35 μm [*Singer*, 1982; *Morris et al.*, 1990; *Bishop et al.*, 1993b]. Infrared absorption spectra of <0.2 μm size fractions of palagonitic soils exhibit features near 3620 cm^{-1} (2.76 μm) and 880 cm^{-1} (11.4 μm), indicative of smectite clays [*Golden et al.*, 1993]. The spectrum of the palagonitic soil from Haleakala (Figure 1) exhibits a broad, weak band extending from 0.8 to 0.95 μm, a reflectance maximum near 0.78 μm, and a rounded shoulder near 0.6-0.65 μm.

2.5. Christiansen Feature

The Christiansen frequency is defined as the frequency at which the real part of the refractive index of individual particles equals that of the surrounding medium [*Henry*, 1948; *Conel*, 1969]. For powdered silicates the Christiansen frequency occurs at the frequency of maximum emissivity and maximum transmission [*Salisbury*, 1993]. This feature is characterized by a sharp decrease in albedo in reflectance spectra and occurs where the real index of refraction, n, is near unity and the extinction coefficient, k, is decreasing strongly.

Spectral experiments [*Logan et al.*, 1973] have shown shifts in the Christiansen feature of quartz and other minerals as a function of particle size, sample texture, background temperature and atmospheric pressure. In this Logan et al. study, emission spectra exhibited shifts in the Christiansen feature toward longer wavelengths with decreasing particle size and toward shorter wavelengths with decreasing atmospheric temperature and pressure. Variations in particle size (0-5 μm versus 0-74 μm) produced the largest effect on the Christiansen feature in this suite of emission spectra [*Logan et al.*, 1973]. In a more recent study involving reflectance spectra of fine-grained olivine, the Christiansen feature was observed to broaden with decreasing particle size and perhaps shift toward shorter wavelengths as well [*Hayes and Mustard*, 1994].

3. Experimental Procedures

3.1. Sample Preparation

3.1.1. Ferrihydrite. The ferrihydrite was synthesized by gradually adding potassium hydroxide to a ferric nitrate solution until a pH of 7-8 was reached, according to the procedure for 2-line ferrihydrite described by Schwertmann and *Cornell* [1991]. The resulting moist sample was rinsed with H_2O over a porous frit, lyophilized, and dry sieved to <45 μm. Additional information about this ferric oxyhydroxide ($Fe_5HO_8{\cdot}4H_2O$) sample can be found in the work by *Bishop et al.* [1993a].

3.1.2. Ferric-bearing montmorillonites. SWy-1 montmorillonite (The Clay Minerals Society, Source Clays Repository) was chemically altered in the laboratory to produce a large collection of montmorillonites containing a variety of interlayer ferric oxyhydroxide and ferric sulfate complexes [*Bishop*, 1994; J. L. Bishop et al., submitted manuscript, 1994]. A suspension was formed from the SWy montmorillonite and H_2O, and repeated washing with dilute HCl produced an H-exchanged form of the montmorillonite. The ferrihydrite-bearing and ferric sulfate-bearing montmorillonites included in this study were prepared by adding 0.5 N ferric chloride or ferric sulfate solutions at pH ~2 to the H-montmorillonite. Dilute NaOH was added dropwise until pH 5 and pH 3.5, respectively. The suspensions were then centrifuged and lyophilized to form a fine powder, which was dry sieved to <45 μm. Additional information about the ferrihydrite-bearing montmorillonite (# 35) and ferric sulfate-bearing montmorillonite (# 117) used in these experiments can be found in the work by *Bishop* [1994] and J. L. Bishop et al. (submitted manuscript, 1994).

3.1.3. Palagonitic soil. This sample (#72) was collected from the Haleakala crater, Maui, as clumps several millimeters in diameter coated by a fine-grained, reddish brown soil. This sample was selected because of the smectitelike absorption features observed near 2.2 μm in reflectance spectra. Only that part of the sample able to pass through a <45 μm dry sieve was used in this study.

3.2. Reflectance Spectra

3.2.1. Nicolet. Reflectance spectra were measured relative to a rough gold surface (Infragold) using a Nicolet 740 FTIR in a H_2O- and CO_2-purged environment. A PbSe detector was used from 0.9 μm to 3.2 μm and a deuterated triglycine sulfate (DTGS) detector from 1.8 μm to 25 μm. The sample chamber, including a remotely controlled sample train with multiple sample positions, was purchased from SpectraTech. Bulk powdered samples are measured horizontally in this system. Spectra are averaged from two locations on each of two replicates for each kind of sample.

3.2.2. RELAB. Additional spectra were measured under ambient conditions relative to Halon in the visible and near-infrared using a photomultiplier tube with the bidirectional RELAB spectrometer at Brown University (see J. L. Bishop et al., submitted manuscript, 1994 for analyses of these spectra). A detailed description of this instrument is provided by *Mustard and Pieters* [1989]. The bidirectional spectrometer allows for variable illumination (i) and emergence (e) angles.

Spectra included here were measured at $i = 30°$ and $e = 0°$. The sample dish is rotated during the measurement to eliminate orientation effects.

3.2.3. Environmental chamber. Reflectance spectra were also measured under controlled atmospheric pressure and temperature conditions using an environmental chamber attachment (for the Nicolet 740 FTIR) built by Connecticut Instruments. Reflectance spectra were measured using a liquid nitrogen-cooled mercury-cadmium-telluride (MCT) detector in the range 1.5-20 μm for each sample under ambient pressure and then again at several increments under reduced atmospheric pressures. Care was taken to deflect the IR beam during pressure adjustment to avoid additional heating of the sample. Dry conditions (i.e., the unbound water in natural smectites was lost) were achieved under atmospheric pressures in the range 0.1 to 10 mbar. The humidity in the lab was controlled to approximately 40% relative humidity giving a P_{H_2O} of ~10 torrs (~13 mbar) at 1-bar atmospheric pressure. The partial pressure of H_2O is then lower under reduced atmospheric pressures and temperatures.

The measurement procedure entailed first measuring the gold standard, then the sample at ambient atmospheric pressure and temperature, in each case after purging the air in the chamber for at least 30 min. A mechanical pump was used to evacuate the air from the environmental chamber, and the atmospheric pressure was measured using three pressure gauges, each having different active ranges and units. For low atmospheric pressure experiments spectra were measured of the sample at several pressure settings: P (atmospheric pressure) = P_{amb} (ambient atmospheric pressure) -10 inches (in) Hg, $P = P_{amb}$ -20 in Hg, $P = P_{amb}$ -30 in Hg; P = 10 mm Hg, P = 5 mm Hg, P = 1 mm Hg, P = 0.1 mm Hg; P = 100 μm Hg, 50 μm Hg, 10 μm Hg (~10^{-6} bar).

The spectra shown in Figure 2 were measured under reduced atmospheric pressures and temperatures according to the following procedure. The temperature of the sample chamber was cooled by adding $N_2(l)$ to a dewar attached to the environmental chamber. Atmospheric pressures below about 1 mbar are required for the low-temperature experiments to prevent H_2O condensation on the internal surfaces. Spectra were measured in these experiments at a temperature, $T = 23°C$ and P = 10 mm Hg, 1 mm Hg, 0.1 mm Hg; then at $T = 20°C$, 15°C, and continuing at 5°C intervals as low as possible in an air environment (-10 to -20°C). At this point the atmospheric pressure was pumped down to ~10^{-6} bar, then $CO_2(g)$ was added up to pressures near 1 mbar. The CO_2 molecules assist in heat transfer but do not cause the condensation problems associated with H_2O molecules. Therefore cooler temperatures (near -40°C) were achieved in a CO_2 environment. The temperature range in this system is limited by a thermally conductive disk between the environmental chamber and the $N_2(l)$ dewar, which can be exchanged to achieve the desired temperature range.

Water vapor fine structures are seen near 2.7 μm and 6 μm in the low-pressure and temperature spectra due to differences in the atmospheric water vapor concentration between ambient and reduced atmospheric pressures. Spectra of this atmospheric water vapor were derived by measuring spectra of the gold standard under variable atmospheric pressure and temperature conditions, and dividing these spectra by the initial gold spectrum. In order to remove the water vapor fine structure the spectra acquired under low atmospheric pressure and temperature were divided by an atmospheric water vapor spectrum (ratio of gold spectra) and smoothed using a variable width boxcar smoothing program. Absolute reflectance values were obtained for the calibrated environmental chamber spectra by scaling these to RELAB values.

4. Results

Reflectance spectra are shown of ferrihydrite, ferrihydrite-bearing montmorillonite, ferric sulfate-bearing montmorillonite and a palagonitic soil under variable atmospheric and environmental conditions in Figure 2. Spectra of each sample were measured under ambient atmospheric air pressure at 23°C, under ~1 mbar air pressure at 23°C, and under 1 mbar CO_2 pressure at near -40°C. The spectra shown in Figure 2, panels 2a, 3a, 4a, and 5a, extend from 5500 to 2500 cm^{-1} (1.8 to 4.0 μm); the spectra in Figure 2, panels 2b, 3b, 4b, and 5b, extend from 1800 to 1250 cm^{-1} (5.5 to 8.0 μm); and the spectra in Figure 2, panels 2c, 3c, 4c, and 5c, extend from 1400 to 800 cm^{-1} (7 to 12 μm).

In Figure 2, panel 2a, the broad feature at 1.95 μm and the shoulder at 2.5 μm observed in the ambient ferrihydrite spectrum are reduced in intensity in the spectra measured under low-pressures and temperatures. The spectra measured under low-pressure and temperature conditions also show a weak 2.75-μm feature that is not resolvable in ambient spectra. In Figure 2, panel 3a, the spectra of ferrihydrite-bearing montmorillonite measured under low-pressure and temperature conditions show little change in the features at 2.2 and 2.75 μm due to structural OH, but suppression of the features at 1.9, 2.5 and ~3 μm due to adsorbed and bound H_2O is observed. Similarly, the spectra of ferric sulfate-bearing montmorillonite in Figure 2, panel 4a, show little change at 2.2 and 2.75 μm; however, significant loss in intensity at 1.9, 2.5 and 3 μm is observed. In spectra of the palagonitic soil shown in Figure 2, panel 5a, the 1.9-μm feature is essentially removed, the 2.2-μm band is enhanced, and the broad 2.5- and 3-μm features are weakened in spectra measured under low-pressure and temperature conditions.

The water-bending vibration near 6.1 μm is seen in spectra of each of these samples in Figure 2, panels 2b, 3b, 4b, and 5b. In the ferrihydrite spectra (Figure 2, panel 2b) this feature nearly disappears under reduced atmospheric pressures and temperatures. In Figure 2, panels 3b, 4b, and 5b, this water feature in the spectra of the ferrihydrite-bearing montmorillonite, the ferric sulfate-bearing montmorillonite, and the palagonitic soil is substantially reduced upon decreasing the atmospheric pressure, and reduced again slightly upon lowering the temperature.

For spectra of the montmorillonites and the palagonitic soil shown in Figure 2, panels 3c, 4c, and 5c, a shift in the Christiansen feature was observed towards shorter wavelengths as the atmospheric pressure and temperature were lowered. In the ferrihydrite spectra (Figure 2, panel 2c) there is an absorption feature near 1350 cm^{-1} (7.5 μm) that disappears under dry conditions, thus complicating detection of shifts in the Christiansen feature for ferrihydrite. The $CO_2(g)$ was required to achieve the lower temperatures in these experiments, but did not appear to influence the position of the Christiansen feature (e.g., no change in the Christiansen feature was observed as a function of the CO_2 pressure).

Band depths have been calculated using the method of Clark and *Roush* [1984] for the hydration features in the spectra shown in Figure 2. For the ferrihydrite spectra band minima of 1.95 μm, 2.32 μm, and 2.95 μm were used. For the spectra of

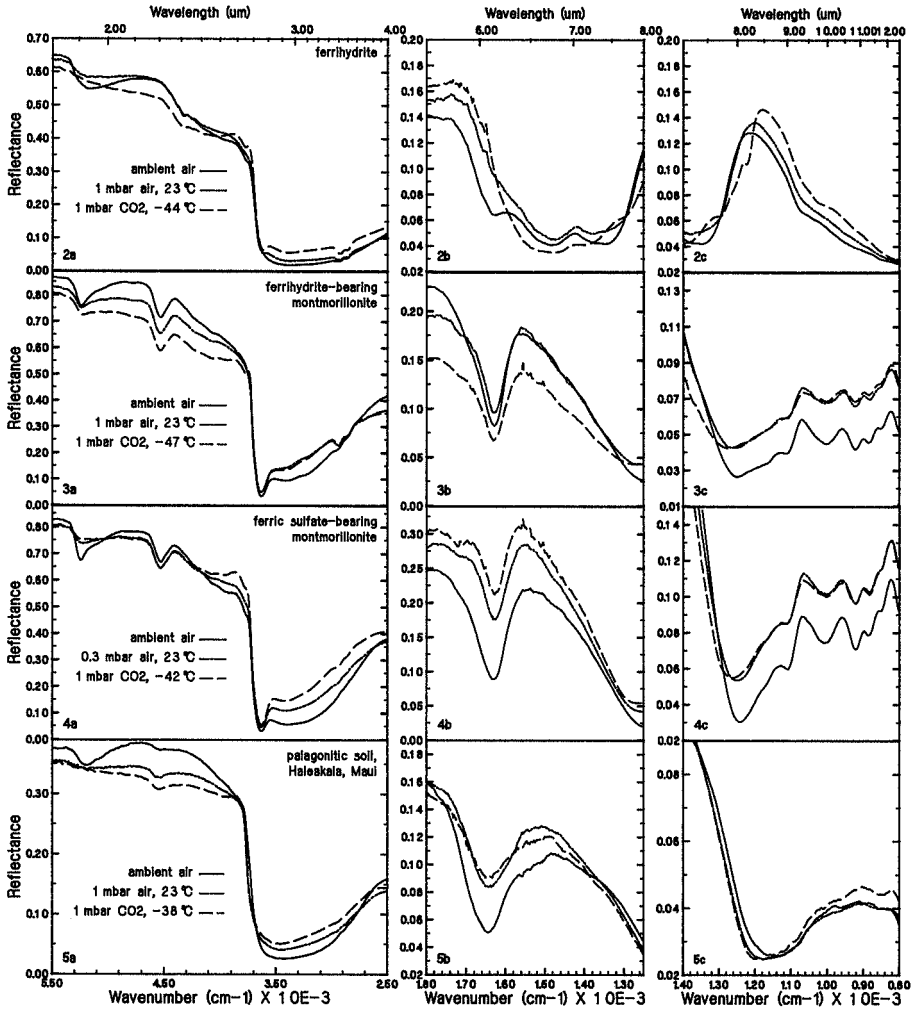


Figure 2. Reflectance spectra of ferrihydrite (panel 2), ferrihydrite-bearing montmorillonite (panel 3), ferric sulfate-bearing montmorillonite (panel 4), and a palagonitic soil (panel 5) measured under variable atmospheric conditions are shown from: (a) 5500 cm^{-1} to 2500 cm^{-1} (1.8 μm to 4.0 μm), (b) 1800 cm^{-1} to 1250 cm^{-1} (5.5 μm to 8.0 μm), and (c) 1400 cm^{-1} to 800 cm^{-1} (7 μm to 12 μm). The spectra shown here were measured under ambient pressure and temperature (solid curve), under reduced atmospheric pressure (patterned curve), and under reduced atmospheric pressures and temperatures (dashed curve).

ferrihydrite-bearing montmorillonite and ferric sulfate-bearing montmorillonite, band minima of 1.91 μm, 2.21 μm, 2.95 μm and ~1630 cm^{-1} (6.14 μm) were used. For the palagonitic soil spectra band minima of 1.93 μm, 2.21 μm, 2.95 μm, and 1645 cm^{-1} (6.08 μm) were used. The band depths for these spectral features are summarized in Table 1.

5. Discussion

5.1. Infrared Features due to Bound and Adsorbed H_2O

Reducing the atmospheric pressure and temperature in these experiments decreases the water vapor in the sample

Table 1. Band Depths for Hydration Features

	Feature, μm								
	1.91	1.93	1.95	2.21	2.32	2.95	6.08	6.14	6.15
Ferrihydrite									
Ambient			14		8	94			
23 °C			6		9	90			
-44 °C			3		11	80			
Ferrihydrite-bearing montmorillonite									
Ambient	12			12		84		55	
23 °C	8			14		72		48	
-47 °C	6			13		73		51	
Ferric sulfate-bearing montmorillonite									
Ambient	17			11		89			63
23 °C	8			12		78			40
-42 °C	6			112		74			32
Palagonitic Soil									
Ambient		8		1		90	64		
23 °C		3		6		83	42		
-38 °C		1		6		79	34		

All values are in percent.

environment. This induces the H_2O molecules not chemically bound to cations or grain surfaces to desorb from the sample. Lowering the atmospheric pressure was found to dehydrate montmorillonites exchanged with several different interlayer cations [*Bishop et al.*, 1994] and montmorillonites containing ferric oxyhydroxides and hydrated ferric sulfate complexes (J. L. Bishop et al., submitted manuscript, 1994). Changes in the spectral hydration features in these experiments depended on the nature of the bonds binding H_2O molecules to the mineral surfaces or interlayer cations. For example, Mg-exchanged montmorillonites exhibit greater decreases in the intensities of the 1.9-μm and 3-μm features than Fe-exchanged montmorillonites, because the interlayer Mg cations are bound to more H_2O molecules [*Bishop et al.*, 1994]. Additional spectral experiments using a Wyoming bentonite showed decreases in the intensities of the features near 1.4 μm and 1.9 μm as the temperature was lowered [*Clark*, 1981a].

Variation in the behavior of the hydration features in spectra of the ferrihydrite, ferrihydrite-bearing and ferric sulfate-bearing montmorillonites and palagonitic soil measured under low atmospheric pressure and temperature conditions (Figure 2) indicates differences in how water is present in these samples. Detailed descriptions of the spectral character of these Mars analogs as a function of atmospheric conditions follow. First, however, there is one trend shared by spectra of all of these samples: reduction in the 2.5 μm shoulder under reduced pressures and temperatures. This feature at 2.5-μm in the ambient spectra of ferrihydrite, montmorillonite, and palagonite is thus apparently due to H_2O adsorbed in a similar manner in each of these samples.

The spectral character of the hydration features in chemically altered montmorillonites is dependent on the ferric interlayer components present. The spectra of the ferrihydrite-bearing montmorillonite in Figure 2, panels 3a and 3b, show a decrease in the intensity of the 3-μm and 6-μm features when the atmospheric pressure was reduced to 1 mbar, but no additional intensity loss was observed upon reducing the pressure further or lowering the temperature. In contrast the spectra of the ferric sulfate-bearing montmorillonite in Figure 2, panels 4a and 4b, show continual loss of intensity of the 3 μm and 6-μm features as the pressure and temperature are lowered. This indicates that the adsorbed H_2O molecules are removed from the ferrihydrite-bearing montmorillonite under relatively mild dehydrating conditions and that additional H_2O in this sample is tightly bound (chemisorbed). For the ferric sulfate-bearing montmorillonites there must be additional H_2O molecules in the system that are more tightly bound than the adsorbed H_2O and less tightly bound than the chemisorbed H_2O. The spectra of both the ferrihydrite-bearing montmorillonite and the ferric sulfate-bearing montmorillonite in Figure 2, panels 3a and 4a exhibit subtle features near 3400 cm^{-1} (2.95 μm) and 3200 cm^{-1} (3.12 μm) when measured under dehydrated conditions. However, spectra of these samples measured in a moist environment show a broad 3-μm band lacking distinguishable features (J. L. Bishop et al., submitted manuscript, 1994).

The H_2O stretching and bending combination band near 1.9 μm lost intensity primarily as a result of dehydration under reduced atmospheric pressure, with only secondary loss in intensity from dehydration under reduced temperature, in spectra of the ferrihydrite-bearing and ferric sulfate-bearing montmorillonites. The band depths decreased from 12% to 6% for the ferrihydrite-bearing montmorillonite spectra and from 17% to 6% for the ferric sulfate-bearing montmorillonite spectra. Although the band minimum is near 1.91 μm in

spectra of both samples, this feature is much broader toward longer wavelengths in spectra of the ferric sulfate-bearing montmorillonite, especially noticeable under dehydrated conditions. This long-wavelength shoulder must be linked to the long-wavelength broadening of the 3-μm band due to the interlayer sulfate complexes in this sample.

These experiments indicate that bound H_2O in the ferrihydrite- bearing and ferric sulfate- bearing montmorillonites exhibit spectral features near 1.91, 2.9, 3.1, 5.9, and 6.1 μm. Adsorbed H_2O and H_2O coordinated to the sulfate groups exhibit broad spectral features near 1.97, 3, and 6 μm that hinder identification of the features due to bound H_2O. Water coordinated to the sulfate groups in the ferric sulfate-bearing montmorillonites is more firmly bound than the adsorbed H_2O, and therefore, broad features near 1.97 and 3 μm can be observed in spectra measured under lower pressures and temperatures. Analyses of these ferric sulfate-bearing montmorillonites, including Mössbauer spectroscopy, differential thermal analysis and X ray diffraction, suggest that small particles of the mineral Schwertmannite and hydrated ferric sulfate complexes surround the montmorillonite grains (J. L. Bishop et al., submitted manuscript, 1994).

The broad combination band centered near 1.95 μm in spectra of ferrihydrite (Figure 2, panel 2a) loses intensity (14% to 3% band depth) and shifts towards 1.9 μm as the pressure and temperature are lowered. The broad 3-μm band due to H_2O stretching vibrations in these spectra decreases in intensity as well (94% to 80% band depth), and a band minimum near 2.95 μm can be detected under dehydrated conditions. The 6.0-μm H_2O bending feature (Figure 2, panel 2b) disappears as the pressure and temperature are reduced, and another feature near 1340 cm^{-1} (7.5 μm) becomes weaker under lower pressures and finally appears as two features at 1400 cm^{-1} (7.1 μm) and 1300 cm^{-1} (7.7 μm) in the low-temperature spectra. It is unclear whether these features at 7.1 and 7.6 μm were present, but masked, in the ambient spectra, or whether a change in the ferrihydrite structure is responsible for these features. Another feature near 1490 cm^{-1} (6.7 μm) broadens and shifts slightly toward shorter wavelengths as the pressure and temperature are decreased. Reflectance spectra of hematite [*Salisbury et al.*, 1991] exhibit absorption features near 6 and 6.5 μm that may be related to those observed for ferrihydrite. Absorption features in reflectance spectra of goethite are observed near 5.5 and 6 μm [*Salisbury et al.*, 1991].

The 1.9-μm H_2O combination band weakens such that it is barely discernible in spectra of the palagonitic soil measured under reduced pressures and temperatures. The broad 3-μm band exhibits continued H_2O loss as the pressure and temperature are decreased (90% to 79% band depth), and a band center is observed near 2.95 μm under low-pressures and temperatures. The H_2O bending feature near 6 μm decreased in intensity, as well, with decreasing pressure and temperature.

In summary, the ferrihydrite and palagonitic soil contain a large amount of adsorbed H_2O and bound H_2O that is not constrained to claylike interlayer sites. Both forms of water are manifested in the spectra of these materials as a broad, strong (90-94% band depth) feature near 3 μm under ambient atmospheric conditions. Upon dehydration under Marslike conditions ferrihydrite and palagonitic soil exhibit a very weak or absent feature near 1.9 μm and a weakened, less broad 3-μm band (although intensity is lost this feature remains very strong). In contrast, spectra of the ferrihydrite-bearing and ferric sulfate-bearing montmorillonites exhibit primarily a sharpening of the 1.9-μm and 3-μm bands when measured under Marslike conditions. These materials contain H_2O bound in specific interlayer sites resulting in features near 1.9 μm, 2.95 μm, and 3.1 μm under dehydrated conditions. Additional H_2O molecules coordinated to sulfate complexes, or adsorbed in the interlayer region or along grain surfaces give broadening near 1.97 μm and 3 μm.

5.2. Infrared Features due to Structural OH

The fundamental OH stretching feature at 2.75 μm and the combination band at 2.2 μm in spectra of the montmorillonites in Figure 2, panels 3a and 4a, remain unchanged as a result of the decreases in atmospheric pressure and temperature. The band depth of the 2.21-μm structural OH combination band is ~12% in spectra of the ferrihydrite-bearing montmorillonite and ferric sulfate-bearing montmorillonite. The fundamental OH bending features at 920 cm^{-1} (10.9 μm), 880 cm^{-1} (11.4 μm), and 850 cm^{-1} (11.8 μm) in spectra of the montmorillonites in Figure 2, panels 3c and 4c, were also unaffected by the environmental changes. Octahedrally coordinated OH groups in smectites are responsible for these features [*Farmer*, 1974]. In montmorillonites, primarily Al cations occupy these octahedral sites, giving rise to the features observed at 2.2, 2.75, and 10.9 μm. Small amounts of Fe and Mg substituting for Al in these sites are responsible for the features at 11.4 and 11.8 μm [*Farmer*, 1974] and the long-wavelength shoulder of the 2.2-μm band [*Bishop et al.*, 1993a].

Spectra of ferrihydrite measured under low-pressures and temperatures (Figure 2, panel 2a) exhibit a fundamental OH stretching feature at 2.75 μm that is masked by the strong 3-μm water band in spectra measured under ambient conditions. The strength of the combination band at 2.3 μm is slightly enhanced in spectra of ferrihydrite under dehydrated conditions. The OH groups are octahedrally coordinated to Fe instead of Al and vibrate at a lower energy (longer wavelength) because the O is more tightly bonded to the Fe cation. This feature observed at 2.3 μm is similar to that observed in nontronite, a smectite that has primarily Fe in the octahedral sites [*Farmer*, 1974]. However, nontronite has strong crystal field bands in visible spectra [*Singer*, 1982; *Burns*, 1993], which are not observed in the spectra of ferrihydrite. Additional mid-IR OH bending features are seen in nontronite near 820 cm^{-1} (12.2 μm) [*Farmer*, 1974; *Stubican and Roy*, 1961], but not seen in the spectra of ferrihydrite in Figure 2, panel 2c, or in *Bishop et al.* [1993a]. Ferrihydrite has only nanophase crystalline order [*Eggleton and Fitzpatrick*, 1988], which may explain why the features observed in spectra of nontronite near 0.6, 0.9, 2.3, 2.75, and 12 μm [*Singer*, 1982; *Burns*, 1993; *Farmer*, 1974; *Stubican and Roy*, 1961], due to octahedrally coordinated Fe and OH species, appear only weakly if at all in the spectra of ferrihydrite.

The spectra of the palagonitic soil in Figure 2, panel 5a, exhibit an OH combination band at 2.2 μm despite the fact that the fundamental OH stretching and bending features are not distinguishable in Figure 2, panel 5c. The 2.2-μm band strength increases from 1% to 6% in spectra of the palagonitic soil as the atmospheric pressure and temperature were reduced. Asymmetry in the broad 3-μm band is probably due to a fundamental OH stretching feature at 2.75 μm that is unresolvable. The presence of a feature near 2.2 μm and possible feature near 2.75 μm indicate a phyllosilicate

component in this soil having Al in octahedral sites coordinated to OH groups. Since smectites have been found in palagonitic soils from Hawaii [*Golden et al.*, 1993], it is likely that the palagonitic soil studied here has partially weathered to montmorillonite.

5.3. Christiansen Feature

Under low atmospheric pressure and low-temperature conditions the Christiansen feature is observed to broaden and shift towards shorter wavelengths in these experiments. The Christiansen feature occurs at 1245 cm^{-1} (8.0 μm) in both of the montmorillonite samples under 1-bar air and 23°C (Figure 2, panels 3c and 4c). This feature shifts gradually to 1270 cm^{-1} (7.9 μm) under lower pressures and temperatures. A similar shift in the Christiansen feature is observed in the palagonitic soil (Figure 2, panel 5c) from 1150 cm^{-1} (8.7 μm) to 1200 cm^{-1} (8.3 μm). The shifts in the Christiansen feature observed here parallel previous observations by *Logan et al.* [1973] of shifts in the Christiansen feature in emission spectra of silicates toward shorter wavelengths with decreasing atmospheric pressure and temperature. The shifts towards shorter wavelength in the Christiansen feature observed here, however, were not accompanied by increased spectral contrast, as was observed by *Logan et al.* [1973].

Possible explanations for the shifts in wavelength of the Christiansen feature in the low-temperature reflectance spectra shown here include apparent changes in the optical properties of the bulk materials as a function of the temperature or a temperature gradient, as a function of dehydration, or as a function of effective sample particle size or texture. There is no precedent for dehydration of a material effecting optical properties. Decreasing the particle size in fine-grained olivine, however, has been observed to broaden the Christiansen feature in reflectance spectra with an apparent shift towards shorter wavelengths [*Hayes and Mustard*, 1994]. In the case of fine-grained soils, the sample texture is a complex mixture of tiny grains clustered together as aggregates. If water molecules adsorbed on the grain surfaces play a role in the ability of these small particles to cluster together, then dehydrating the soils may influence the sample texture and result in an effectively smaller particle size. These and other possible environmental effects on fine-grained soils need to be further evaluated.

6. Conclusions

The Mars soil analogs examined in these experiments are compositionally consistent with the chemical analyses of Viking and exhibit spectral features in the visible part of the spectrum that are similar to features observed in the bright regions on Mars. Reflectance spectra of these analog materials were measured under reduced atmospheric pressure and temperature environments comparable to the temperate zones on Mars. These experiments showed that the structural OH features near 2.2-2.3 μm and 2.75 μm in spectra of the palagonitic soil and ferrihydrite became enhanced when measured under Marslike atmospheric conditions. The spectral features due to structural OH in ferrihydrite-bearing and ferric sulfate-bearing montmorillonite remained unchanged as a result of decreasing the atmospheric pressure and temperature.

The spectral features due to adsorbed H_2O in these samples were lost or weakened under low-pressure and temperature conditions. Therefore broad features centered at 2.9 μm in the ferrihydrite and palagonitic soil are due to bound H_2O, and features at 1.9 μm, 2.9 μm, and 3.1 μm in the low-pressure and temperature spectra of the ferrihydrite-bearing and ferric sulfate-bearing montmorillonites are due to bound H_2O. The ferrihydrite-bearing montmorillonite contains adsorbed H_2O that is lost under relatively mild dehydrating conditions and bound H_2O that is a necessary part of the structure and is resistant to dehydration under low atmospheric pressures and temperatures. In addition to adsorbed and bound H_2O, the ferric sulfate-bearing montmorillonites contain H_2O molecules coordinated to sulfate groups, which are loosely bound to the montmorillonite system.

Spectra of the palagonitic soil when examined under atmospheric pressures and temperatures of temperate regions on Mars, exhibit enhanced near-infrared features, suggesting the presence of montmorillonite. This is further evidence that palagonitization is linked to smectite formation and that palagonitization of basaltic volcanic glass on Mars may be a mechanism for smectite formation in the surface material there.

7. Applications to Mars

Based on chemical composition, visible reflectance spectra, and infrared reflectance spectra acquired under low atmospheric pressures and temperatures, these samples are promising analog materials for the bright regions on Mars. The ferrihydrite-bearing montmorillonite or a mixture containing ferrihydrite and a silicate phase may be present in bright regions such as Arabia where the ferric absorption occurs at 0.91-0.92 μm. The ferric sulfate-bearing montmorillonite may be present in bright regions such as Lunae Planum where the ferric absorption occurs near 0.88 μm. The nature of the ferric and silicate components vary greatly in palagonitic soils; however, many exhibit hematiticlike absorptions and may be present in bright regions having a ferric absorption near 0.85 μm.

The experiments presented here show the sensitivity of the spectral hydration features in soils to the atmospheric environment of the sample, and the importance of measuring analog soils under the appropriate environmental conditions. Under Marslike atmospheric conditions, the palagonitic soil exhibited a structural OH feature with a 6% band strength near 2.2 μm. A feature near 2.2 μm has also been observed by multiple researchers in some of the Martian bright regions [*Arnold*, 1992; *Murchie et al.*, 1992, 1993a; *Bell and Crisp*, 1993]. Spectra of ferrihydrite-bearing montmorillonite and ferric sulfate-bearing montmorillonite samples are characterized by a stronger 2.2-μm feature (11-12% band strength). Spectra of ferrihydrite include a broad feature near 2.3 μm, which may be related to a feature in this region observed in some bright regions on Mars. Under Marslike atmospheric conditions (1) the 1.9-μm feature is substantially weaker in spectra of each of the samples, especially for the palagonitic soil and ferrihydrite, where the 1.9-μm feature is barely discernible, (2) the feature near 2.2-2.3 μm is stronger than the 1.9-μm feature in spectra of each of these analog materials, and (3) a strong 3-μm band is present in the spectra of ferrihydrite, palagonitic soil, and ferric sulfate-bearing montmorillonite. The results presented here are an important step in constraining interpretations of the hydration features in Mars bright region spectra.

Acknowledgments. The authors would like to express gratitude to S. Pratt for technical assistance and to J. Mustard and S. Murchie for discussions of the results and helpful editorial suggestions. Support through the ZONTA Foundation, the NASA Graduate Student Researchers Program at NASA-ARC, Moffett Field, and the Alexander von Humboldt Foundation, Germany, is greatly appreciated by J. Bishop. NASA support for this research under grant NAGW-28 is gratefully acknowledged by C. Pieters. RELAB is a multi-user facility supported by NASA under grant NAGW-748. The Nicolet system was acquired under a grant from the Keck Foundation.

References

Adams, J., and T. B. McCord, Mars: Interpretation of spectral reflectivity of light and dark regions, *J. Geophys. Res., 74*, 4851-4856, 1969.

Allen, C. C., J. L. Gooding, M. Jercinovic, and K. Keil, Altered basaltic glass: A terrestrial analog to the soil of mars, *Icarus, 45*, 347-369, 1981.

Anderson, D. W., E. S. Gaffney and P. F. Low, Frost phenomena on Mars, *Science, 155*, 319-322, 1967.

Arnold, G., Infrarotsondierung planetarer Oberflächen, *Dtsch. Luft Raumfahrt-Nachrichten, 69*, 25, 1992.

Baird, A. K., A. J. Castro, B. C. Clark, P. Toulmin III, H. Rose Jr., K. Keil, and J. Gooding, The Viking X ray fluorescence experiment: Sampling strategies and laboratory simulations, *J. Geophys. Res., 82*, 4595-4624, 1977.

Banin, A., T. Ben-Shlomo, L. Margulies, D. F. Blake, R. L. Mancinelli, and A. U. Gehring, The nanophase iron mineral(s) in Mars soil, *J. Geophys. Res., 98*, 20,831-20,853, 1993.

Bates, R. L., and J. A. Jackson (Ed.), *Dictionary of Geological Terms*, Anchor, New York, 1984.

Bell, J. F., III, and D. Crisp, Groundbased imaging spectroscopy of Mars in the near-infrared: Preliminary results, *Icarus, 104*, 2-19, 1993.

Bell, J. F., III,, T. B. McCord, and P. D. Owensby, Observational evidence of crystalline iron oxides on Mars, *J. Geophys. Res., 95*, 14,447-14,461, 1990.

Bibring, J.-P., *et al.*, ISM observations of Mars and Phobos: First results, *Proc. Lunar Planet, Sci. Conf., 20th*, 461-471, 1990.

Bishop, J. L., Spectroscopic analyses of chemically altered montmorillonites and applications to the soils on Mars, doctoral thesis, Brown Univ., Providence, R.I., 1994.

Bishop, J. L., and C. M. Pieters, Reflectance spectra of Mars soil analogs measured under reduced atmospheric pressures and temperatures, (abstract), *Lunar Planet. Sci. Conf. XXV*, 117-118, 1994.

Bishop, J. L., C. M. Pieters, and R. G. Burns, Reflectance and Mössbauer spectroscopy of ferrihydrite-montmorillonite assemblages as Mars soil analog materials, *Geochim. Cosmochim. Acta, 57*, 4583-4595, 1993a.

Bishop, J. L., S. L. Murchie, S. Pratt, J. F. Mustard, and C. M. Pieters, The importance of environmental conditions in reflectance spectroscopy of laboratory analogs for Mars surface materials, *LPI Tech. Rep., 93-06*, 4-5, 1993b.

Bishop, J. L., C. M. Pieters, and J. O. Edwards, Infrared spectroscopic analyses on the nature of water in montmorillonite, *Clays Clay Miner., 42*, 701-715, 1994.

Blaney, D. L., and T. B. McCord, An observational search for carbonates on Mars, *J. Geophys. Res., 94*, 10,159-10,166, 1989.

Bouška, V., *Natural Glasses*, 354 pp., Ellis Horwood, Chichester, England, 1993.

Bruckenthal, E. A., The dehydration of phyllosilicates and palagonites: Reflectance spectroscopy and differential scanning calorimetry, M.S. thesis, Univ. of Hawaii, Honolulu, 1987.

Burns, R. G., *Mineralogical Applications of Crystal Field Theory*, 2nd ed., Cambridge University Press, New York, 1993.

Calvin, W. M., Additions and corrections to the absorption coefficients of CO_2 ice: Applications to the Martian south polar cap, *J. Geophys. Res., 95*, 14,743-14,750, 1990.

Cariati, F., L. Erre, G. Micera, P. Piu, and C. Gessa, Polarization of water molecules in phyllosilicates in relation to exchange cations as studied by near-infrared spectroscopy, *Clays Clay Miner., 31*, 155-157, 1983.

Carr, M. H., *The Surface of Mars*, 232 pp., Yale University Press, New Haven, Conn., 1981.

Clark, R. N., The spectral reflectance of water-mineral mixtures at Low-temperatures, *J. Geophys. Res., 86*, 3074-3087, 1981a.

Clark, R. N., Water frost and ice: The near-infrared spectral reflectance 0.65-2.5 μm, *J. Geophys. Res., 86*, 3087-3096, 1981b.

Clark, R. N., and T. B. McCord, Mars residual north polar cap: Earth-based spectroscopic confirmation of water-ice as a major constituent and evidence for hydrated minerals, *J. Geophys. Res., 87*, 367-370, 1982.

Clark, R. N., and T. L. Roush, Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications, *J. Geophys. Res., 89*, 6329-6340, 1984.

Conel, J. E., Infrared emissivities of silicates: Experimental results and a cloudy atmosphere model of spectral emission from condensed particular mediums, *J. Geophys. Res., 74*, 1614-1634, 1969.

Coyne, L. M., J. L. Bishop, T. Scattergood, A. Banin, G. Carle, and J. Orenberg, Near infrared correlation spectroscopy: Quantifying iron and surface water in a series of variably cation-exchanged montmorillonite clays, in *Spectroscopy of Minerals and Their Surfaces*, edited by L. Coyne, S. McKeever, and D. Blake, pp. 407-429, American Chemical Society, Washington, D.C., 1990a.

Coyne, L. M., J. L. Bishop, L. Howard, and T. Scattergood, Recent spectroscopic findings concerning clay-water interactions at low humidity: Possible applications to models of Martian surface reactivity, in *4th Symposium on Chemical Evolution and the Origin of Life*, NASA-Ames Research Center, Moffett Field, Calif., 1990b.

Davies, D. W., The relative humidity of Mars' atmosphere, *J. Geophys. Res., 84*, 8335-8342, 1979.

Eggleton, R. A., and R. W. Fitzpatrick, New data and a revised structure model for ferrihydrite, *Clays Clay Miner., 36*, 111-124, 1988..

Erard, S., J.-P. Bibring, J. Mustard, O. Forni, J. W. Head, S. Hurtrez, Y. Langevin, C. M. Pieters, J. Rosenqvist, and C. Sotin, Spatial variations in composition of the Valles Marineris and Isidis Planitia regions of Mars derived from ISM data, *Proc. Lunar Planet. Sci. Conf., 21st*, 437-455, 1991.

Fanale, F. P., J. R. Salvail, W. B. Banerdt, and R. S. Saunders, Mars: The regolith-atmosphere-cap system and climate change, *Icarus, 50*, 381-407, 1982.

Farmer, C. B., D. W. Davies, A. L. Holland, D. D. LaPorte and P. E. Doms, Mars: Water vapor observations from the Viking orbiters, *J. Geophys. Res., 82*, 4225-4248, 1977.

Farmer, V. C., The layer silicates, in *The Infrared Spectra of Minerals*, edited by V. C. Farmer, 539 pp., Mineralogical Society, London, 1974.

Golden, D. C., R. V. Morris, D. W. Ming, H. V. Lauer Jr., and S. R. Yang, Mineralogy of three slightly palagonitized basaltic tephra samples from the summit of Mauna Kea, Hawaii, *J. Geophys. Res., 98*, 3401-3411, 1993.

Gooding, J. L., and K. Keil, Alteration of glass as a possible source of clay minerals on Mars, *Geophys. Res. Lett., 5*, 727-730, 1978.

Gooding, J. L., S. J. Wentworth, and M. E. Zolensky, Aqueous alteration of the Nakhla meteorite, *Meteoritics, 26*, 135-143, 1991.

Hayes, J. E. and J. F. Mustard, Effects of fine particles (<25 μm) on reflectance spectra from 0.3 to 25 μm, (abstract), *Lunar Planet. Sci. Conf., XXV*, 519-520, 1994.

Henry, R. L., The transmission of powder films in the infrared, *J. Opt. Soc. Am., 38*, 775-789, 1948.

Hess, S. L., R. M. Henry, and J. E. Tillman, The seasonal variation of atmospheric pressure on Mars as affected by the south polar cap, *J. Geophys. Res., 84*, 2923-2927, 1979.

Houck, J. R., J. B. Pollack, C. Sagan, D. Schaack, and J. A. Decker, Jr., High-altitude infrared spectroscopic evidence for bound water on Mars, *Icarus, 18*, 470-480, 1973.

Johnston, C. T., G. Sposito, and C. Erickson, Vibrational probe studies of

water interactions with montmorillonite, *Clays Clay Miner., 40*, 722-730, 1992.

Kieffer, H. H., Spectral reflectance of CO_2-H_2O frosts, *J. Geophys. Res., 75*, 501-509, 1970.

Kieffer, H. H., T. Z. Martin, R. Peterfreund, and B. M. Jakosky, Thermal and albedo mapping of Mars during the Viking primary mission, *J. Geophys. Res., 82*, 4249-4291, 1977.

Logan, L. M., G. R. Hunt, J. W. Salisbury, and S. R. Balsamo, Compositional implications of Christiansen frequency maximums for infrared remote sensing applications, *J. Geophys. Res., 78*, 4983-5003, 1973.

Lucey, P. G., and R. N. Clark, Spectral properties of water ice and contaminants, in *Ices in the Solar System*, edited by J. Klinger et al., pp. 155-168, D. Reidel, Norwell, Mass., 1985.

McCord, T. B., R. L. Huguenin, D. Mink, and C. Pieters, Spectral reflectance of Martian areas during the 1973 opposition: Photoelectric filter photometry, *Icarus, 31*, 25-39, 1977.

McCord, T. B., R. N. Clark, and R. B. Singer, Mars: Near-infrared spectral reflectance of surface regions and compositional implications, *J. Geophys Res., 87*, 3021-3032, 1982.

Morris, R. V., J. L. Gooding, H. V. Lauer Jr., and R. B. Singer, Origins of Marslike spectral and magnetic properties of a Hawaiian palagonitic soil, *J. Geophys. Res., 95*, 14,427-14,434, 1990.

Murchie, S., S. Erard, J. Bishop, J. Mustard, J.-P. Bibring, Y. Langevin, J. Head, and C. Pieters, The forms and evolution of water in Martian soil: Evidence from ISM imaging spectroscopy, (abstract), *Lunar Planet. Sci. Conf., XXIII*, 941-942, 1992.

Murchie, S., J. Mustard, J. Bishop, J. Head, C. Pieters, and S. Erard, Spatial variations in the spectral properties of bright regions on Mars, *Icarus, 105*, 454-468, 1993a.

Murchie, S., J. Mustard, S. Erard, and R. Singer, Variations in the Fe mineralogy of bright Martian soil,(abstract), *Lunar Planet. Sci. Conf., XXIV*, 1029-1030, 1993b.

Mustard, J. F., and J. F. Bell III, New composite reflectance spectra of Mars from 0.4 to 3.14 μm, *Geophys. Res. Lett., 21*, 353-356, 1994.

Mustard, J. F., and C. M. Pieters, Photometric phase functions of common geologic minerals and applications to quantitative analysis of mineral mixture reflectance spectra, *J. Geophys. Res., 94*, 13,619-13,634, 1989.

Owen, T., K. Bieman, D. R. Rushneck, J. E. Biller, D. W. Howarth, and A. L. Lafleur, The composition of the atmosphere at the surface of Mars, *J. Geophys. Res., 82*, 4635-4639, 1977.

Prost, R., Étude de l'hydration des argiles: Interactions eau-minéral et mécanisme de la retention de l'eau., II., Étude d'une smectite, *Ann. Agron., 26*, 463-535, 1975.

Rosenqvist, J., P. Drossart, M. Combes, T. Encrenaz, E. Lellouch, J.-P. Bibring, S. Erard, Y. Langevin, and E. Chassefière, Minor constituents in the Martian atmosphere from the ISM/Phobos experiment, *Icarus, 98*, 254-270, 1992.

Roush, T. L., D. L. Blaney, and R. B. Singer, The surface composition of Mars as inferred from spectroscopic observations, in *Remote Geochemical Analyses: Elemental and Mineralogical Composition* , edited by C. M. Pieters and P. Englert, pp. 367-393, Cambridge University Press, New York, 1993.

Salisbury, J. W., Mid-infrared spectroscopy: Laboratory data, in *Remote Geochemical Analysis: Elemental and Mineralogical Composition*, edited by C. M. Pieters and P. A. J. Englert, pp. 79-98, Cambridge University Press, New York, 1993.

Salisbury, J. W., L. S. Walter, N. Vergo, and D. M. D'Aria, *Infrared (2.1-25 μm) Spectra of Minerals*, Johns Hopkins Univ. Press, 1991.

Schwertmann, R. M., and R. M. Cornell, *Iron Oxides in the Laboratory*, VCH Publications, Weinheim, Germany, 1991.

Sherman, D. M., R. G. Burns, and V. M. Burns, Spectral characterization of the iron oxides with application to the Martian bright region mineralogy, *J. Geophys Res., 87*, 10,169-10,180, 1982.

Singer, R. B., Spectral evidence for the mineralogy of high-albedo soils and dust on Mars, *J. Geophys. Res., 87*, 10,159-10,168, 1982.

Singer, R. B., T. B. McCord, R. N. Clark, J. B. Adams, and R. L. Huguenin, Mars surface composition from reflectance spectroscopy: A summary, *J. Geophys Res., 84*, 8415-8426, 1979.

Singer, R. B., J. S. Miller, K. W. Wells, and E. S. Bus, Visible and near-IR spectral imaging of Mars during the 1988 opposition, (abstract), *Lunar Planet. Sci. Conf., XXI*, 1164-1165, 1990.

Soderblom, L. A., The composition and mineralogy of the Martian surface from spectroscopic observations: 0.3 μm to 50 μm, in *Mars*, edited by H. Kieffer et al., pp. 557-593, University of Arizona Press, Tucson, 1992.

Stubican, V., and R. Roy, A new approach to the assignment of infrared absorption bands in layer silicates, *Z Kristallogr., 115*, 200-214, 1961.

Toulmin, P., III, A. K. Baird, B. C. Clark, K. Keil, H. J. Rose Jr., R. P. Christian, P. H. Evans, and W. C. Kelliher, Geochemical and mineralogical interpretation of the Viking inorganic chemical results, *J. Geophys Res., 82*, 4625-4634, 1977.

Treiman, A. H., R. A. Barret, and J. L. Gooding, Preterrestrial aqueous alteration of the Lafayette (SNC) meteorite, *Meteoritics, 28*, 86-87, 1993.

J. L. Bishop, DLR, Institute for Planetary Exploration, Rudower Chaussee 5, D-12489 Berlin, Germany. (email: bishop@terra.pe.ba.dlr.de)

C. M. Pieters, Brown University, Box 1846, Department of Geological Sciences, Providence, RI 02912. (email: pieters@pds.geo.brown.edu)

(Received March 3, 1994; revised November 17, 1994; accepted December 16, 1994.)

Page intentionally left blank

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5381–5395, MARCH 25, 1995

Longitudinal dunes on Mars: Relation to current wind regimes

Pascal Lee and Peter C. Thomas

Center for Radiophysics and Space Research, Cornell University, Ithaca, New York

Abstract. Longitudinal dunes are extremely rare on Mars, but constitute a substantial fraction of terrestrial desert dunes. We report finding isolated examples of longitudinal dunes on Mars and relate their occurrence to expected sand transport regimes. Terrestrial longitudinal dunes form in bimodal and multimodal transport regimes. General circulation models and streak data indicate that bimodal and multimodal transport of sand should be very rare on Mars. Thus the dearth of longitudinal dunes on Mars is consistent with their apparent formation conditions on Earth.

1. Introduction and Summary

Although dunes are prominent features on the surface of Mars that are easily related to common terrestrial physical processes, the relative proportions of dune types on Mars are very different from those on the Earth. The scarcity of longitudinal dunes on Mars compared to their relatively common occurrence on Earth suggests some fundamental differences in the Martian eolian system that may have to be taken into account by models that relate eolian features to present or past climates, sediment sources, or to properties of the dune materials themselves.

This paper uses the characteristics of the few reported examples of Martian longitudinal dunes and the expected wind regimes to investigate the significance of their scarcity. We first briefly review the relationship of dune forms to wind regime and sand supply on the Earth, with particular emphasis on bimodal and multimodal wind regimes and dune forms. We review terrestrial longitudinal dunes and emphasize the findings that their formation requires bimodal or multimodal transport. Then the circumstances of occurrence of dunes on Mars, and particularly of longitudinal ones, are summarized, with a description of two newly found occurrences of longitudinal dunes in the north polar region. The longitudinal dunes occur in restricted areas at 73°N, 57°W and 71.4°N, 52.5°W and appear to be seifs spawned from the asymmetric elongation of barchan horns. We next evaluate the likely sand transport regimes on Mars using results from the Ames general circulation model, observations of wind streaks on Mars, and Viking Lander data. We show that the models and data available allow for only very limited multimodal transport on Mars, if any. We then focus on the relationship of the longitudinal dunes in the north polar erg to the current wind regime. The possibility of dune stabilization by salt cementation is also considered, although, on the basis of the good coincidence of many bedforms with current winds, we suggest that the dunes were formed by recent wind regimes.

Paper number 95JE00225.
0148-0227/95/95JE-00225$05.00

We conclude that the scarcity of longitudinal dunes on Mars is entirely consistent with the unimodal transport regimes expected and observed for most of the planet.

2. Wind Regimes and Dune Forms

2.1. Wind Regimes and Sand Transport

Eolian dunes are bedforms in loose materials transported in saltation by the wind. Most dune forms are distinctive enough that they provide excellent evidence of the process of saltation even when seen with marginal resolution on another planet. Although dunes include a wide variety of sizes and shapes and occur in many geological, topographic, and climatic settings, there are useful generalities relating dune forms to the strength and variation of sand transport. These generalities should apply to dune forms on Mars and many of the worst complicating factors intervening on Earth are absent on Mars: vegetation and other effects of moisture.

The wind regimes to which dune forms can be compared are often described by the drift potential (DP), i.e., the total sand flux in mass/unit length/time calculated from wind measurements, and its relation to the resultant drift potential (RDP), which is the calculated net transport. The DP is the scalar sum of the calculated transport; RDP is the magnitude of the vector sum of the transport. Works such as those by *Fryberger* [1979] and *Wasson and Hyde* [1983] describe the techniques for calculating drift potentials and their diverse applications. The calculation of the drift potential depends upon use of a relation of the mass transport rate, $\dot{m}$, of loose sand to the wind shear velocity, V_*. Many expressions have been derived from theory, wind tunnel testing, and field measurements to relate sand transport rates to shear velocity [e.g., *Bagnold*, 1941; *Kawamura*, 1951; *Kadib*, 1965; *Lettau and Lettau*, 1977; *White*, 1979]. Summary comparisons are given by *Greeley and Iversen* [1985] and *Sarre* [1987]. The expression of *White* [1979] contains the elements essential to models of sand transport:

$$\dot{m} = A\rho(V_* - V_{*t})(V_* + V_{*t})^2/g \qquad (1)$$

where A is a constant, ρ is atmospheric density, V_* is the shear velocity, V_{*_t} is the impact threshold shear velocity,

and g is gravitational acceleration. $\dot{m}$ is in $M^1L^{-1}T^{-1}$ units. V_{*t} of course varies with particle size and density. For purposes of comparison, one particle size and density are assumed, and often, the most easily moved size (~150 μm on Mars [*Greeley and Iversen* 1985]). We can simply write the sand transport quantity, Q, as

$$Q = A(V_* - V_{*t})(V_* + V_{*t})^2 t \quad (2)$$

where A is a constant and t is the time this particular wind condition is active. Q is the magnitude of a vector; its orientation is taken from measurements or a model. Over a period of time, a locale is subject to a variety of winds, and we sum the various Q's to get the DP, and sum the vectors to get the magnitude and orientation of the RDP. The transport variability, TV, is then given by

$$TV = RDP/DP \quad (3)$$

This parameter is used by *Fryberger* [1979] and *Wasson and Hyde* [1983] to indicate the variability in sand transport directions and to relate quantitatively the wind regime to dune morphologies.

The value of TV can vary from 0 to 1. We illustrate some combinations of winds that give different values of TV to provide context for discussion below. Two equal transport vectors, 90° apart, give a value of TV of 0.71. Without at least one sand-transporting wind component more than 90° from another, the value of TV cannot fall below 0.71. Figure 1 shows the calculated sand transport from a set of winds, assumed to blow for equal periods of time, that vary from weakest to strongest by a factor of about 1.5. In plot A the winds range from 1.35 times V_{*t} to 2.07 times; in plot F from 0.77 times to 1.18 times V_{*t}. The only change from A to F is the proportional decrease in all the wind magnitudes. As the winds decrease, the effective transport regime goes from multimodal (low values of TV) to unimodal (high values of TV) because some winds are below threshold and the others have decreasing transport capability. Simply by scaling the winds shown in A by 0.6, the transport changes to that in example E. *Fryberger* [1979] presents sand rose plots and drift potential statistics, along with Landsat images and dune morphology maps for a variety of terrestrial sand seas that can be compared to these idealized examples.

2.2. Sand Supply and Sand Transport

The most commonly used measure of the availability of sand is the equivalent sand thickness (EST). The EST is the average depth of sand if the dunes were spread out evenly. This quantity is difficult to measure accurately, but is fairly simple to estimate approximately because of the general geometric rules that apply to piles of sand formed by saltation accumulation. Measurement of dune widths, lengths, and spacing from aerial or orbital images can give reasonable estimates of sand volumes. *Lancaster and Greeley* [1990] discuss applications of EST measurements to Martian dunes. As discussed by *Wasson*

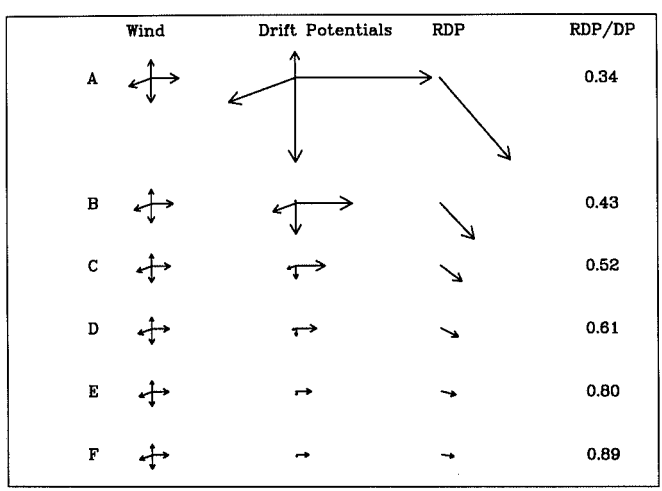


Figure 1. Model drift potentials (DPs) and resultant drift potentials (RDPs) for a four-wind system. The only difference among the examples is the gradual reduction in all winds from A to F; the ratio of wind speeds is the same throughout; as winds decrease, fewer components are above threshold, and transport by those that are above threshold decreases (see equation (1)). Simply by scaling the entire wind system by 0.6 from A to F, the resultant transport has gone from a strong multimodal regime to a weak, unimodal one. Note these are not sand roses; wind and transport arrows are directed downwind.

and Hyde [1983], EST may play a role in determining some dune types.

Drift potential has been used to describe the vigor of sand transport. Its relation to dune form or to EST is complex, though it may in some areas provide some additional discrimination. It is sufficient to emphasize here that strong winds do not mean large accumulations of sand.

2.3. Dune Forms and Wind Regimes

Apart from any influence from vegetation or local obstacles, dune forms are mostly related to the directional variability of effective winds and to the equivalent sand thickness; wind strength and particle size appear to have minimal influence [*Wasson and Hyde*, 1983]. Although morphologic and genetic classifications of dunes abound, it is most useful here to consider four groups: barchan, transverse, longitudinal, and star [see *McKee*, 1979; *Lancaster*, 1982, 1989; *Lettau and Lettau*, 1977]. Subtypes of course are recognized, and supplement the overall classification (e.g., longitudinal dune types of simple, compound, and complex) [*Lancaster*, 1982]. Good summaries of the range of dune morphologies on the Earth are found in the works by *McKee* [1979], *Breed and Grow* [1979], *Fryberger* [1979], and *Fryberger and Goudie* [1981].

Comparisons of terrestrial dune types with wind regimes on regional and continental scales have been made by *Fryberger* [1979], *Wasson and Hyde* [1983], *Tsoar* [1989], and *Lancaster* [1989]. These studies all show that dune types correlate very strongly with wind regime: barchans occur in the most unimodal winds; transverse dunes in slightly less unimodal; longitudinal dunes form along or close to the net transport vector of substantially bimodal or multimodal winds, and star dunes form in the most complex, multimodal winds. To illustrate the generality of these trends, we have replotted results from *Wasson and Hyde* [1983] and from *Fryberger* [1979] in Figure 2. Wasson and Hyde plot RDP/DP versus the equivalent sand thickness. Their data include dunes in Africa, Australia, South and North America, and show good distinction of dune types based on the combination of wind regime and sediment thickness. Note that Wasson and Hyde's data are given with 1 sigma ranges. Fryberger's data come from Africa, Asia, Australia, and North America, and are given as total ranges of RDP/DP for three groups of dunes, with no distinction of sediment thickness.

We emphasize that correlation of dune types with wind regimes involves considerable interpretation and extrapolation. The interpretation rests on classification of forms that are detected in the field and in aerial or orbital images with varying resolutions and lighting. The extrapolation resides primarily in applying wind data recorded at the margins of sand seas to their interiors (maps by *Fryberger* [1979] and *Fryberger and Ahlbrandt* [1979] are instructive). Extrapolation is also involved in assigning dune forms to present winds or to past winds, a problem for the very largest forms and for those that may be stabilized. In general, however, we regard the

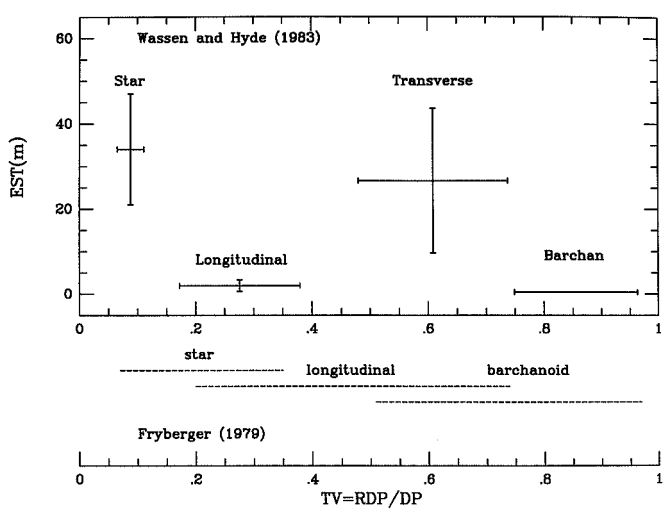


Figure 2. Occurrence of dune types as a function of RDP/DP and equivalent sediment thickness (EST) from *Wasson and Hyde* [1983]; and as a function of RDP/DP from data in *Fryberger* [1979]. Ranges from Wasson and Hyde are 1 sigma's; the ranges from Fryberger are the entire range of occurrences listed in tables in Fryberger [1979]. See text, and compare to transport regimes in Figure 1.

correlation between forms and winds to be good enough for simple, broad comparisons.

2.4. Longitudinal Dunes and Their Significance

The terminology of "linear" and "longitudinal" has caused much confusion and grief. For the most part, linear is taken as a morphological term, longitudinal as defining a form that parallels a resultant transport vector. Linear and longitudinal dunes are dealt with in detail by *Lancaster* [1982] and *Tsoar* [1989].

Longitudinal dunes cover half to two thirds of the Earth's sand seas. They are relatively less common in Asia and in the Americas than in Africa, but are completely absent in only a few sand seas [*Lancaster,* 1982].

The formation of longitudinal dunes has been a subject of controversy. Several kinds of such dunes are often distinguished (e.g., lee dunes, vegetated-linear dunes, and seif dunes [*Tsoar, 1989*]), each kind involving apparently a distinct mode of formation. Lee dunes are easily interpreted, and vegetation is absent from Mars, so our task reduces to comparing seif-like dunes on the two planets. Seif dunes are often sinuous, kinked or convoluted in plan view, and have sharp-edged crests when fresh. Closely spaced seifs sometimes interact and their wiggly crest-lines may merge in Y-shaped junctions. Seif dunes on Earth are commonly found in association with barchans or barchanoid dunes, and can be derived from the asymmetric elongation of barchan horns under a bidirectional wind regime [*Bagnold,* 1941; *Lancaster,* 1980; *Tsoar,* 1989]. Seif dunes may also form under more complex, multimodal wind regimes (see recent review by *Cooke et al.* [1993]). For this work we are chiefly concerned with the findings that longitudinal (seif) dunes appear to require bimodal or multimodal wind regimes to form.

3. Martian Dunes

3.1. Overview

The major masses of dunes on Mars occur in a partial ring of ergs around the north polar deposits at 70-80°N, and in intracrater dune fields 60-80°S. Dunes also are abundant in low-latitude chasma, and are found scattered in other topographic traps, chiefly in craters and near wrinkle ridges. Their global distribution can be related to both wind patterns and to source regions [*Thomas,* 1982; *Greeley et al.,* 1992; *Ward and Doyle,* 1983; *Thomas and Gierasch,* this issue].

The Viking imaging coverage is more than adequate to classify the major dune types and locations on Mars. While some regions are seen only at resolutions of 200 m/pixel, and with less than optimal viewing, the considerable coverage of dunes that exists at well under 100 m/pixel allows confidence in the finding that Martian dunes are overwhelmingly crescentic and transverse [*Breed et al.,* 1979a,b; *Tsoar et al.,* 1979; *Greeley et al.,* 1992].

3.2. Previous Study of Mars Longitudinal Dunes

A few examples of longitudinal dunes on Mars have now been found, one set at 59°S, three others at 70-73°N. All are small local concentrations.

Tsoar et al. [1979] reported dune morphologies that reflect up to three local wind directions at several places in the north polar erg. Two examples of elongate barchans were documented (73°N, 57°W and 77°N, 97°W), where the inferred effective winds differ by about 40°. Other locations show variations in effective winds typically of 60°. All examples are interpreted by Tsoar et al. to show domination of the dune forms by one wind regime and minor modification by others. The dunes reported exhibit dominantly transverse characteristics and true longitudinal dunes were not reported in these areas.

Edgett and Blumberg [1994] have identified perhaps the clearest examples of "linear" dunes so far located on Mars at 59°S, 343°W. The dunes are about 4-5 km in length, and occur between two masses of transverse dunes inside a 70-km-diameter crater. The longitudinal dunes apparently are derived from the upwind transverse dunes and are reformed downwind into another mass of transverse dunes. These are similar to longitudinal dunes emanating from transverse dune masses on the Earth. The identification of only one set of longitudinal dunes in the many dune fields in Noachis suggests to Edgett and Blumberg that local effects such as the topographic funneling of winds through breaches in the crater rim and on the crater floor are responsible for their restricted occurrence. They further note that the forms might also represent the result of long-term changes in the winds that are moving the main dune mass from a position in the southern side of the crater, towards the north. *Edgett and Blumberg* [1994] have also found star dunes at 8.8°S, 270.9°W, in a valley formed by overlapping crater rims.

3.3. Longitudinal Dunes in the North Polar Erg

We have found two sites of longitudinal dunes in the north polar erg. The first is at 73°N, 57°W [*Lee et al.,* 1993], a few tens of kilometers south of one of the occurrences of elongated barchans reported by *Tsoar et al.* [1979]. The dunes are best resolved in Viking Orbiter frames 525B11-14, obtained at L_s = 38° at pixel scales of 29 m (Figure 3). They are clustered within a mass of transverse and barchanoid dunes near the southernmost extent of the north polar erg. The longitudinal dunes occur over an area of the order of 10^2 km^2 and are grouped in sets of subparallel linear ridges. The longest dunes are about 3 km in length, while dune width is approximately 90 m. Spacing between the dunes, measured orthogonally from crest to crest, is typically 150 m, ranging from less than 100 m in the densest parts of the field to about 300 m in the field's more open portions. In several instances, the longitudinal dunes display what appear to be kinks and Y-shaped junctions. The general wiggly or convoluted outline of the dunes suggests that they are of the seif type. Many of the longitudinal dunes originate from barchans or barchanoid ridges by the asymmetric elongation of individual horns (Figure 3b). Many have completely lost any transverse character, and their connection with the transverse dune forms is only inferred by working backwards through the continuum of forms.

The second site is near (71.4°N, 52.5°W) (Figure 4), in the immediate vicinity of a 4-km impact crater, approximately 120 km southeast of the field discussed above. The dunes are imaged at 26 m/pixel in Viking Orbiter frame 487B13, acquired at L_s = 21°. The

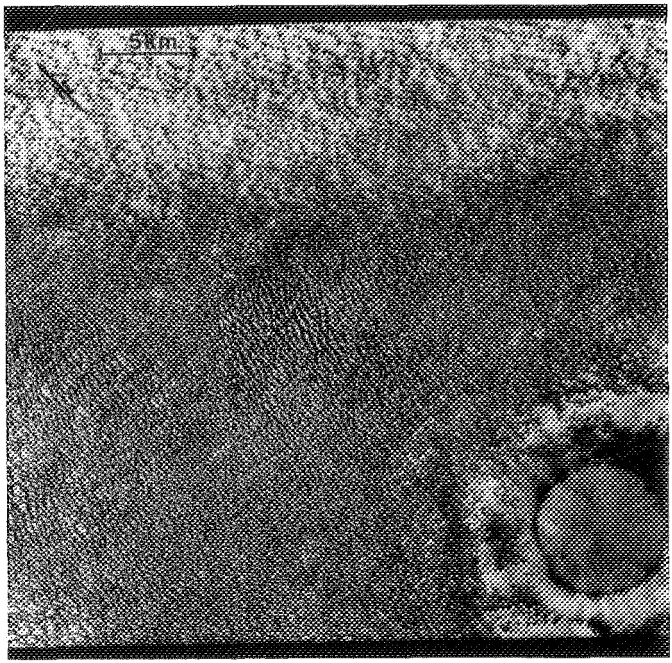


Figure 3a. Barchan-derived longitudinal dunes. The dunes occur within a field comprising many dune forms (individual barchans and barchanoid ridges) at 73°N, 57°W. Viking Orbiter image 525B13, acquired at L_s = 38° (resolution = 28.8 m/pixel).

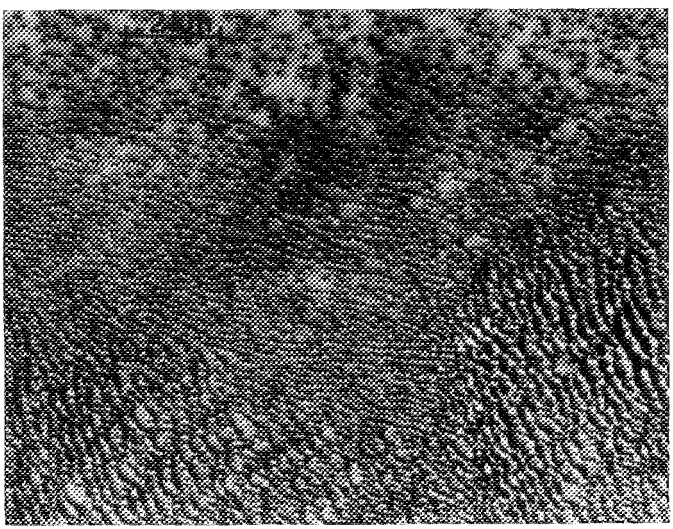


Figure 3b. Enlargement of the central area of Figure 3a, showing the elongation of barchan horns into seifs.

longitudinal dunes occur as sets of linear ridges that run roughly parallel to a NW-SE-trending dune shadow streak. The dunes are best seen downwind from the crater itself, immediately off the southwestern edge of the streak. The longitudinal dunes are associated with transverse bedforms and grade into the barchanoid ridges amassed near the downwind end of the streak. Some longitudinal dunes are at least 1 km long. Dune width is of the order of 90 m. Typical dune spacing is approximately 150 m. These dimensional characteristics are similar to those of the seifs at (73°N, 57°W).

3.4. Summary of Mars Longitudinal Dune Occurrences

Longitudinal dunes occur on Mars only as small isolated groups and largely in areas otherwise dominated by transverse dune forms. There is no occurrence of longitudinal dunes on Mars remotely similar to the regional fields of longitudinal dunes on Earth that can cover well over 10^4 km^2. The Martian longitudinal dunes do occur in areas where multiple effective winds modify barchan and transverse dune forms, produce streamers of dark material at widely divergent angles from dunes, and cause braided dune patterns. True longitudinal dunes, though, are very rare, and are only poorly resolved.

4. Modality of Sand Transport on Mars

4.1. Circulation Models

The major Martian winds are driven by seasonally dependent heating and by mass flux to and from the polar caps. Diurnal slope winds can also be significant. The rapid radiative response of the atmosphere and the low thermal inertia of the surface help produce the most severe seasonal changes in any observed planetary atmosphere [*Magalhães*, 1987]. Seasonal changes on Mars are not only large, they present a strong asymmetry between north and south because southern summer occurs near perihelion, leading to much more vigorous circulation in southern summer than during northern summer. Flux of CO_2 to the polar caps is also asymmetric, and the different mass transferred into and out of each polar region each year drives winds of different strengths in the two hemispheres.

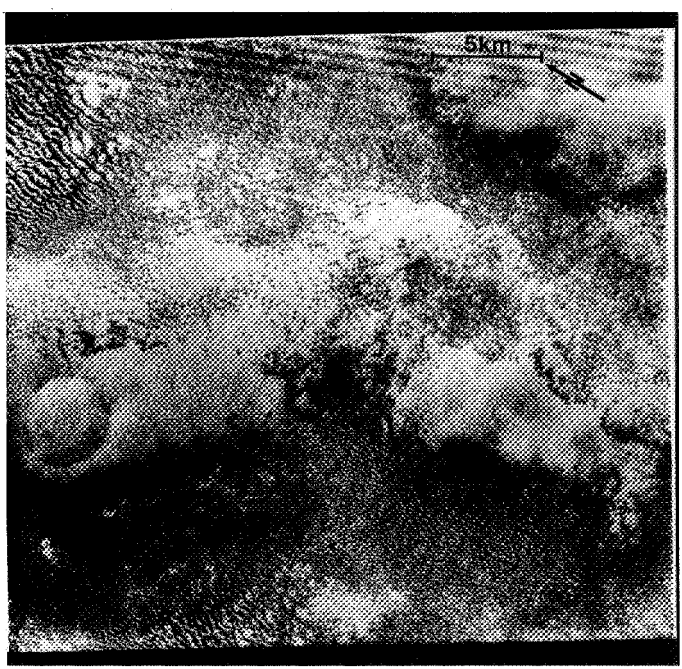


Figure 4. Longitudinal dunes at 71.4°N, 52.5°W. Viking Orbiter image 487B13, acquired at $L_s = 21°$ (resolution = 26 m/pixel). The main region of longitudinal dunes is bounded on three sides by transverse dunes, and on the fourth grades into the dune-free region of the streak extending from the 4-km impact crater. Note that the transverse dune orientations vary by about 30° in this image.

General circulation and other models of the winds on Mars have had considerable (but not total) success in predicting directions and times of winds as shown by wind streaks and dune orientations [*Greeley et al.*, 1993; *French and Gierasch*, 1979; *Magalhães*, 1987], and by cloud motions [*Kahn*, 1983]. However, they rarely predict surface shear stresses that exceed threshold for the most easily moved particles on Mars. As noted by *Greeley et al.* [1992], *Edgett and Blumberg* [1994], and others, this situation likely means either that sand particles on Mars are less dense than thought, or that local enhancement of winds provides enough of an increase to the general circulation to allow the threshold shear to be exceeded. Inasmuch as large dust storms on Mars involve particles that are probably more difficult to set in motion than are sand-sized particles, it is clear that threshold velocities are being exceeded in the present epoch.

General circulation models (GCMs) [*Pollack et al.*, 1990; *Haberle et al.*, 1993] predict very seasonally peaked distributions of strong winds on Mars. Maximum surface calculated stresses occur around L_s = 280° (southern summer, the season of most large dust storms). High surface stresses also are predicted at latitudes 50-70°N and S during seasons of maximum polar sublimation or deposition (spring and fall).

We use the Ames GCM results to evaluate predictions of multimodal transport regimes on Mars. The model is described by *Pollack et al.* [1990], *Haberle et al.* [1993], and *Greeley et al.* [1993]. It calculates atmospheric motion accounting for solar input, optical depth, topography, albedo, thermal inertia, and CO_2 condensation in the atmosphere and on the surface. Assumptions are made about atmospheric and surface properties (such as the ratio of boundary layer thickness to surface roughness). The data used here were calculated for twelve 10-day intervals distributed through the Martian year (L_s = 20, 44, 71, 99, 127, 156, 178, 209, 247, 281, 306, and 349°), recorded in 1.5-hour time steps. The output was supplied by D. Blumberg and includes surface stress components in 9° × 7.5° cells. Nearly all the stresses are below threshold for the most easily moved particles on Mars [*Greeley et al.*, 1993], and thus it might be argued that the models have no applicability to sand transport. However, the models do reproduce the directions of significant fractions of bright and dark streaks [*Greeley et al.*, 1993] and of some cloud motions [*Haberle et al.*, 1993]. In addition, the high stress areas correlate well with areas of higher block abundance [*Christensen*, 1986]. This relation suggests that the models offer reasonable predictions of the locations of the highest wind stresses capable of eroding dust coverings. We adopt the assumption that extreme effects are hidden by cell size smoothing (cells are 440 km by 530 km at the equator) and time step size (1.5 hours), and that local enhancements of the general circulation cause threshold to be exceeded over a significant fraction of the planet at some time during the typical year.

We model the drift potential by scaling the stress results to shear velocity, as the stress is proportional to the shear velocity squared. We assume several different threshold values and thus in effect we scale all the ratio of shear velocities to threshold velocity calculated from the GCM results. This scaling allows for uncertainties in our knowledge of the threshold shear velocity, calculated surface stresses, and amount of any local enhancement of the general circulation. The uniform scaling of course assumes that our uncertainty in the relation of shear velocity to threshold shear velocity is the same for all areas of the planet. The technique is obviously crude but is quite adequate for tabulating the general predictions of the GCM for changes in wind regimes at different points during the Martian year. A choice of a high threshold results in very little calculated transport; a very low value results in more transport, and, as shown in section 2, should result in higher order modality in the transport regime, i.e., lower values of TV.

The output for one typical day's run (16 time steps; the 10-day total for each L_s bin is very similar to the typical day) during each of the 12 periods was summed for DP and RDP at each cell location, using (2). The values of TV for nearly all of these daily transport regimes were over 0.9, indicative of strongly unimodal winds, though a few were less than 0.6. The net transport for each of these 12 representative days was then summed for the whole year with the assumption that each wind regime operated for the same time period. Because of the slightly uneven spacing of the runs available, some minor bias may have been introduced. Orientations of the net transport vector (RDP), its magnitude, and the values of TV were saved for each of the 1000 cell locations. We have varied the threshold velocity in four steps from 0.43 to 0.76 of the inferred nominal threshold shear velocity; a scaling of 1.0 would correspond to a nominal shear stress of 550×10^{-4} N m^{-2} [*Greeley et al.*, 1992].

Predicted net transport directions for the four scaled threshold values are shown in Figure 5. Figure 6a presents the distribution of TV values with latitude for one of the four different assumed threshold values. The average value of TV for each latitude is also plotted. The latitude dependences of average TV for the four assumed threshold values are compared in Figure 6b. Several important characteristics are evident from Figures 5 and 6. (1) Strict application of the nominal 550×10^{-4} N m^{-2} applied to the GCM results yields virtually no effective transport. Reducing the threshold by 44% from the maximum (scaling reduced from 0.76 to 0.43) results in a very great increase in the fraction of the planet predicted to be subject to effective sand transport. The generally widespread distribution of eolian features on Mars, and their frequent alignment with present observed and calculated winds, suggests that the model in Figure 5a is more realistic than that in Figure 5d. Note that the orientation of the net transport changes in some areas as the threshold allows more components to be effective. (2) Even with the lowest assumed threshold, transport regimes are mostly unimodal. For the lowest threshold considered, over 95% of the area is calculated to have a TV of over 0.5, and 75% a TV greater than 0.8. Regardless of how one might change thresholds or slightly modify model inputs, the association of most of the vectors in Figure 5a with winds of a very specific season emphasize that multimodal sand transport on Mars is likely to be very restricted geographically in the absence of local modification of the general circulation. (3) The two polar margin regions and low northern latitudes are predicted to

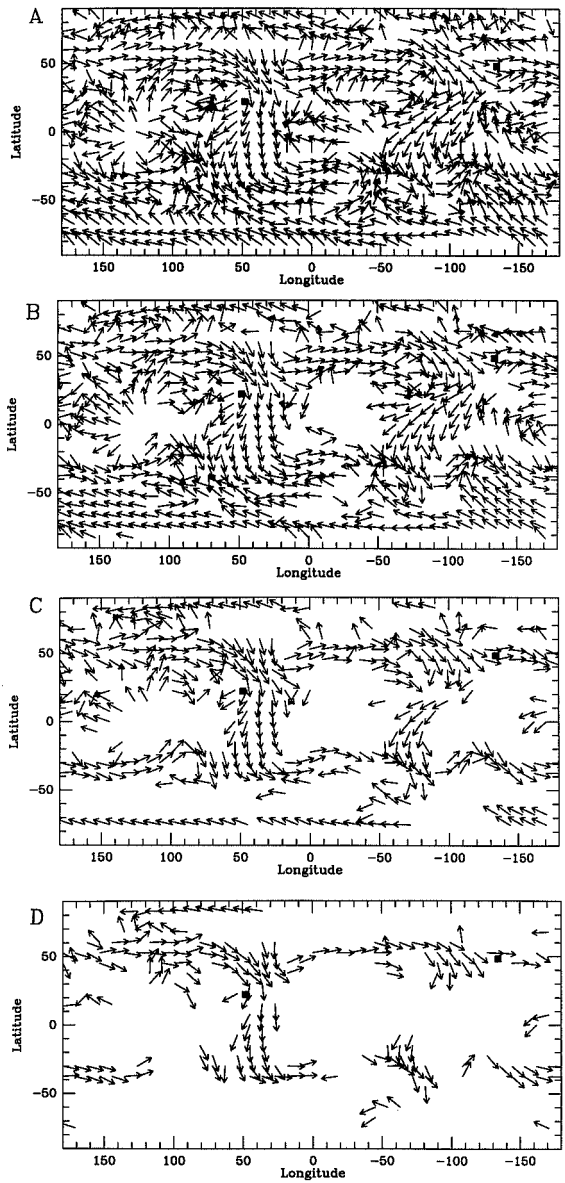


Figure 5. Net annual transport directions calculated from output of the Ames GCM for four scalings of the nominal threshold velocity (see text): threshold scalings are: (a) 0.43, (b) 0.54, (c) 0.65, and (d) 0.76. (A scaling of 1 would correspond to a shear stress of 550×10^{-4} N m^{-2}.) Vectors shown only for cells with transport (winds above threshold). As a ratio of threshold to modeled winds increases, areas with transport of sand decrease.

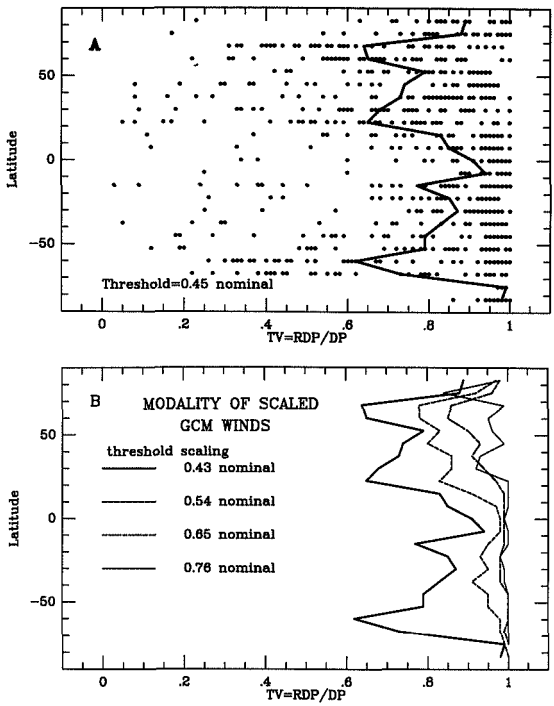


Figure 6. Modality of transport regimes on Mars from Ames GCM output. (a) Results for a threshold shear velocity 0.43 times the nominal shear velocity. Dots mark values of TV for GCM cells (7.5 × 9°); solid line is average value of TV as a function of latitude. (b) Summary of the four different scalings of threshold; additional values are 0.56, 0.65, and 0.76 of the nominal value, corresponding to a threshold of 550 × 10^{-4} N m^{-2}. As threshold is increased, the transport generally becomes more unimodal, as in Figure 1. See text for description of the GCM and the calculations for this figure. Compare to values of TV in Figure 2 for the different terrestrial dune types.

be the most likely areas of multimodal transport. The affinity of the polar margin regions for multimodal regimes is also predicted by the model of *French and Gierasch* [1979] in which strong prograde winds associated with fall CO_2 mass transport to the pole and retrograde spring winds associated with mass loss from the caps imply multimodal transport.

Although the GCM predicts that sand transport under a multimodal wind regime is most likely to occur at high latitudes where indeed the few examples of longitudinal dunes are found, the net transport vectors predicted for the north polar locations above about 60°N are considerably different from what the dune forms, wind streaks, and even clouds suggest. The discrepancy and its significance for the north polar longitudinal dune interpretations are discussed below in section 5. Suffice it to say here that the GCM predicts mostly unimodal winds on Mars, with multimodal regimes likely possible in only restricted locations.

4.2. Wind Streak Data

The vast majority of wind streaks on Mars are interpreted to belong to one of four categories: (1) dust deposits in the lee of craters (bright streaks; Type Ib); (2) erosional scars in dust in the lee of topographic features (dark streaks; Type Id); (3) plumes of dark material blown from dune fields (splotch streaks; Type II); and (4) accumulations of windblown frost (frost streaks). The interpretations are based on a variety of morphologic, time variability, and photometric data, as well as modeling of wind flow over obstacles and of atmospheric stability [*Greeley et al.*, 1992].

The orientations of bright streaks (Type Ib) and dark Type II streaks that emanate from intracrater dunes

overwhelmingly follow the Hadley cell flow of southern summer between latitudes ±40° [*Thomas et al.*, 1984; *Magalhães*, 1987]. The bright streaks cover large areas with consistent wind patterns and with virtually no hint of multiple lobe streaks emanating from any single crater. The most variable of these streaks occur on regional slopes where dunes are essentially absent. Two aspects of the Type II streaks, interpreted to be saltating material blown off intracrater dunes, also suggest very unimodal transport. First, they show a consistent global pattern of flow that is largely coincident with the southern summer Hadley cell flow. Local variations in directions are small except for some demonstrably topographic effects [see *Thomas et al.*, 1984, Figure 6b]. Second, most of the streaks are long compared to their widths. The large length/width ratios of many of these dark streaks suggest that once the sand is blown off the source deposit, it continues to move in a very uniform direction. Effective multimodal winds would be expected to form other streaks from the source deposits or to smear out the streaks. The boundaries of many Type II streaks are smaller streaks parallel to the main streak [*Thomas et al.*, 1984], and indicate only one locally effective wind regime. In some instances the source deposits are resolvable as crescentic dunes with orientations similar to that of the associated dark wind streak (which can extend for well over 100 km).

Type II streaks at high southern latitudes mostly reflect southern spring flow, which is retrograde off the seasonal polar cap, to the northwest [*Thomas et al.*, 1979]. There are a few local streamers from dune fields between 40° and 50°S that show more than one effective wind, but these seem to be a minor component of the streak population and vary in orientation by less than 50°. At high northern latitudes, observational evidence for spring flow is minimal except very near the polar layered deposits [*Thomas,* 1981]. In both polar regions prograde flow outward from the pole during fall and winter is recorded by frost streaks [*Thomas et al.*, 1979]. Winds that occur during times of frost cover might be expected to be ineffective on materials covered by frost. However, the strong variegation of the seasonal frost caps [*James et al.*, 1979] suggests the presence of areas within the caps that are minimally covered or perhaps entirely swept free of frost. This might provide the opportunity for wind to move sand also during the fall and winter seasons.

Crater-related wind streaks (Types Ib and Id) do record multiple winds at several locations on Mars. In the Daedalia region (30°S, 120°W), there are multiple dark (Type Id) streaks from single craters spanning orientations that diverge by as much as 100° [*Veverka et al.*, 1977; *Thomas and Veverka*, 1979]). These winds, however, probably would have TV's of over 0.7 because there are additional winds within the 100° fan. They also are not in areas that have many sand dunes. Some areas of Hesperia (28°S, 245°W) show nearly 180° differences in the winds that form Type Ib bright streaks and those that produce Type Id dark streaks [*Veverka et al.*, 1981]. The wind regime certainly can be classed as multimodal, but because the flow regime responsible for the streaks may occur only in association with the crater topography, a single location might not experience multimodal sand transport. This area, too, has a dearth of visible sand dunes. In many areas where slope winds are expected, there may be a daily reversal of wind directions, which could support multimodal transport. However, the slopes are not areas that accumulate large dune fields.

4.3. Viking Lander Data

Viking Landers operated for several seasons at 22°N, 48°W, and 48°N, 226°W. *Arvidson et al.* [1983] reported that winds above threshold for undisturbed surfaces were rare or absent at the lander locations. Hodographs of lander wind data indicate considerable daily and seasonal variation, but still with dominating winds such that any transport is essentially unimodal. Drifts observed at the Viking Lander 1 site apparently have suffered substantial erosion, but roughly align with crater-related wind streaks seen from orbit. The VL-1 images also show small streaks extending from rocks that roughly coincide in direction with the streaks observed from orbit [*Sagan et al.*, 1977]. Azimuths of these features range from about 180° to 220° and the GCM net transport vector for this area is close to 200° for all assumed thresholds (Figure 4). Wind directions recorded by the VL-1 are strongest toward an azimuth of about 200°-220° [see *Haberle et al.*, 1993]. The Viking Lander 2 site did not have the large range of wind markers seen at the VL-1 site. On the basis of lander meteorology data, GCM predictions for the higher latitude VL-2 site appear to be somewhat less accurate than for the VL-1 site [*Haberle et al.*, 1993].

4.4. Summary of Modality of Expected Sand Transport on Mars

Overall, the GCM results and the wind streak data suggest that multimodal transport of sand should be uncommon on Mars. Viking Lander observations are also consistent with this conclusion. The dearth of longitudinal dunes on Mars can be attributed to this relative lack of bimodal or multimodal transport.

5. Wind Regime of North Polar Longitudinal Dunes

Can the occurrences of longitudinal dunes in the north polar erg be ascribed specifically to the transport regimes modeled in section 4, or is the situation more complex? Because most dunes in the region are transverse, the longitudinal dunes would seem to be local exceptions to the general regime.

5.1. Dunes, Streaks, and GCM Results

Dune forms shown in Figure 3a are plotted in Figure 7 to show the distribution of major dune types: transverse (barchanoid) ridges, barchan dunes, barchans with elongated horns, and longitudinal (seif) dunes. Also plotted are the orientations of some of the most unambiguous wind markers in the area: the aforementioned dunes, the dune shadow streak that extends over 20 km from the large, partially buried impact crater, and dark streaks emanating from dunes and from the crater rim.

In general the region shows the effects of winds toward between 90° and 135° azimuth. Differences in orientations

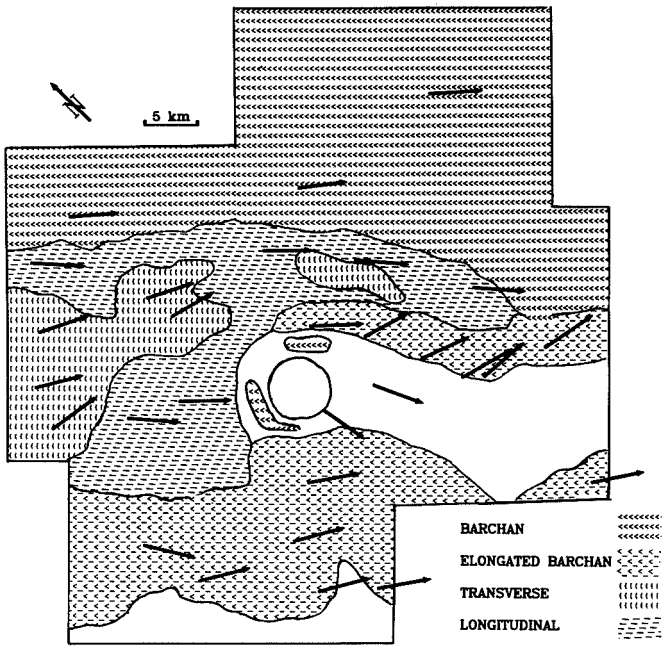


Figure 7. Morphology of dunes shown in Figure 3a. The dunes in this area include a continuum between barchans, barchanoid ridges, elongate barchans, and the strictly longitudinal seifs. Units are delineated as areas dominated by a particular kind of dune. Considerable morphologic variability exists in some units. Refer to Figure 3a for the full variation of forms. Vectors are drawn for the clearest wind-indicating morphologies.

of barchan forms, elongate horns, and nearby longitudinal dunes are nearly all less than 20°. Some elongate barchan horns suggest winds with azimuths of 140-145°. The distal segment of the dune shadow streak may show winds with azimuths of about 145°. Wind directions of about 170° azimuth are indicated by a few dark streaks from the degraded crater rim.

Figure 8 shows the orientations of north polar dunes and streaks to provide context for Figure 7. The overall patterns are discussed in *Thomas and Gierasch* [this issue]. The flow inferred from dunes and streak orientation in the region of the longitudinal dunes is somewhat exceptional for the north polar erg in that winds there do not tend to confine dunes to a latitude belt, but allow material to flow south (Figure 8). Topography (poorly known) may influence the flow in this region as it is a broad trough leading south [*Thomas and Gierasch*, this issue]. Figure 9 compares the north polar dune and streak orientations with the transport vectors predicted from the GCM for different times of the year. The GCM results for the region of 74°N, 54°W where the longitudinal dunes occur are quite different from the winds inferred from the orientation of dunes and streaks: net transport is predicted to be from the east, while dunes and streaks suggest northwesterly winds. Interestingly, however, the pattern shown by the GCM for slightly lower latitudes would match those inferred for the dune areas very well (Figure 9b). *Haberle et al.* [1993] have noted that the GCM predicts a much longer period of flow of easterly winds in the northern spring than the cloud data show [*Kahn*, 1983]. They attribute this difference to the GCM not being able to fully account for changing dust loading during the run period. The GCM results for high southern latitudes are more consistent with the dominant wind markers there, and with the *French and Gierasch* [1979] model. Inversion of the French and Gierasch model for northern latitudes of 60-70° would suggest a bimodal regime with winter winds toward the southeast and spring winds toward the southwest. Winds to the southwest, however, are rarely shown by surface features except above latitude 80°N [*Thomas and Gierasch*, this issue]. We must assume that the asymmetry of dust loading reduces the effectiveness of the spring winds in the far north (as suggested by *Haberle et al.* [1993]), and that the winter, off-pole winds are stronger at latitudes 60-75°N than the Ames GCM output predicts.

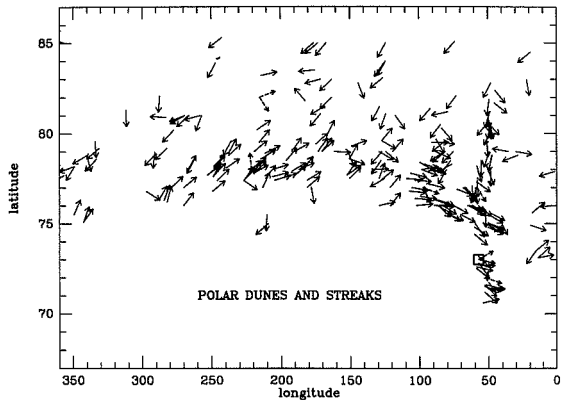


Figure 8. Orientations of north polar dunes and streaks. Plot of downwind directions of dune and streak data discussed by *Thomas and Gierasch* [this issue]. Dashed arrows show winds inferred at several locations by *Tsoar et al.* [1979]. Box encompasses area with longitudinal dunes shown in Figure 3. Most of the polar erg is effectively confined in a latitude belt. The area that includes the longitudinal dunes is the only large area that allows dune material to move far south of the erg.

Do these results invalidate those of section 4.1? That section's primary conclusion was that the GCM predicted little multimodal transport on Mars. If anything, the GCM has given a more multimodal result for the high northern latitudes than is suggested by eolian features. Thus the conclusion of the likely paucity of multimodal transport on Mars remains firm.

The stable transport mode (type and size of dune) varies greatly over scales of only a few kilometers in the areas where the north polar longitudinal dunes occur; this is most dramatically shown by the dune field shadow southeast of the large degraded crater (Figure 7). The variability in winds and sand supply induced by the modest topography of this crater's rim (probably only tens of meters high) were sufficient to eliminate dunes entirely from a zone over 20 km long, even though dune forms at its margins show that sand must have moved through the area that lacks dunes. The change in transport that formed the longitudinal dunes from the barchans or the transverse dunes was probably much more subtle.

All indications are that the north polar longitudinal dunes formed from a slight modification of winds with at least two distinct effective directions, but which differed by less than 50°. Some additional modification may have occurred under other winds, but their action was only poorly preserved in the eolian forms. Inasmuch as the north polar longitudinal dunes are intimately associated with elongate barchans, they almost certainly would fall in the very upper part of TV values applicable for longitudinal dunes in Figure 2.

5.2. Relict Features?

In this study we have compared Martian dune forms to present-day winds. But could the dunes instead have been formed by winds of a different climate? The dunes may effectively be relict in several ways: they could have been physically stabilized; they could still be active but unable to change rapidly enough to reflect current winds; or they could be inactive simply because winds are below threshold.

On the Earth, dune immobilization by cementation (the only form of stabilization that might occur on Mars) usually involves dissolution and reprecipitation of salts [e.g., *Verstappen*, 1968]. Terrestrial dunes made of salt-rich (in particular carbonates) and clay-rich sediments exposed to moisture are those most likely to experience cementation. Once lithified, the dunes may preserve their form over geological timescales, especially if the bedforms are spared significant weathering.

Several lines of evidence point to the availability of abundant salts and clays on Mars, the basic ingredients of typical terrestrial dune cements. Geochemical inventories conducted at the Viking Lander sites suggest that sulfates and chlorides are likely present, possibly along with carbonates and nitrates [*Baird et al.*, 1976; *Toulmin et al.*, 1977]. Salts are also invoked as possible bonding agents responsible for the cementation of the many clods and blocks observed at the Viking Lander sites [*Clark et al.*, 1982], and also for the pelletization of fine-grained silt and clay to allow these materials to be transported by winds as sand-sized particles. From theoretical arguments based on the range of possible thermodynamic conditions prevailing on Mars and on the physical chemistry of salts, *Clark and Van Hart* [1981] suggested that halite (NaCl), magnesium and sodium sulfates [$(Mg,Na_2)SO_4$], and magnesium and calcium carbonates [$(Mg,Ca)CO_3$] are likely abundant on Mars. From a synthesis of remote-sensing observations, *Jakosky and Christensen* [1986] inferred the existence of a global duricrust on Mars formed perhaps by the mobilization of salt ions in a layer

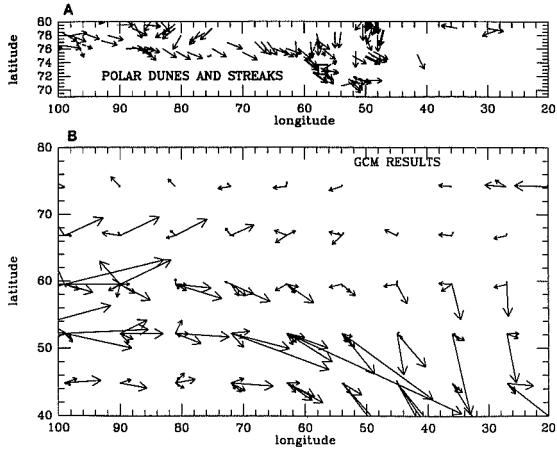


Figure 9. Comparison of part of the north polar dune and streak orientations with GCM results. (a) Dune and streak data as in Figure 8. (b) GCM transport vectors for a threshold of 0.43 times the nominal threshold for the 12 time periods listed in section 4. The pattern of enhanced flow out from the pole occurs at similar longitudes, but at higher latitudes than predicted by the GCM.

of water adsorbed within the regolith. These authors pointed out that the close similarity in the respective concentrations of S and Cl at the two well-separated Viking Lander sites is consistent with the existence of an indurated saline crust of global extent. By analogy with weathering processes operating in salt-rich regions on the Earth, salts were also considered as viable and effective weathering agents on Mars [*Malin*, 1974]. Salt-rich aqueous solutions have also been invoked in the preterrestrial alteration of several SNC meteorites (unusual achondrites thought to have originated from Mars [*McSween*, 1985, 1994]) [*Gooding et al.*, 1991; *Treiman et al.*, 1993; *Wentworth and Gooding*, 1994]. Numerous intracrater and some extracrater albedo features on Mars have also been interpreted in terms of salt or clay-rich playa deposits, some of which were apparently processed by winds into lunettes and parna [e.g. *Lee*, 1993; *Williams and Zimbleman*, 1994; *Greeley and Williams*, 1994]. The features provide additional circumstantial evidence in support of an abundant availability of salts and clays on Mars.

With salt- and clay-rich sediments available, the cementation of Martian dunes could occur if a solvent is available. Liquid water is not currently stable at the Martian surface but water may be adsorbed in a monolayer around near-surface soil particles [*Fanale and Cannon*, 1974]. In Antarctica, a monolayer of unfrozen, adsorbed water is capable of transporting dissolved ions through soils even at temperatures below freezing [*Ugolini and Anderson*, 1973]. From these authors' ion diffusion rates, *Jakosky and Christensen* [1986] indicated that salts could migrate several meters in the Martian upper surface layer over a timescale of the order of 10^5 years, a process which could be facilitated by the diffusion of water into and out of the regolith on a similar timescale. Near-surface crusts could thus form on Mars over a period of 10^5 years, provided, of course, sediments are not disturbed over shorter periods. Wind regimes on Mars, however, experience major astronomically driven changes over timescales of 10^4 years. The season of perihelion controls which season experiences the strongest winds and undergoes a cycle of period 51,000 years. The dunes and streaks in the vicinity of the north polar longitudinal dunes appear to match present fall and winter flow, near southern summer and perihelion, the epoch of highest wind stresses. At the other extreme in the cycle, winds in this region would be expected to be dominated by northern spring and summer winds, toward the southwest. Because the alignment of the largest wind feature in the area, the dune shadow streak, with the present-day wind system is unlikely to be fortuitous, we believe most smaller bedforms near the streak probably also relate to winds that are recent and that have been effective over timescales very short compared to the 51,000-year cycle. Moreover, because dune lithification would appear to require quiescent, diffusive conditions over periods of 10^5 years to be effective, the observed bedforms are likely poorly stabilized.

Thus, although there is the potential for stabilization of dunes on Mars, the indications of current widespread activity and coincidence of many bedforms with current winds argues that it probably has a minor influence at most. The question of what bedforms might lag behind changing winds has been dealt with by *Thomas* [1981, 1982]. While some of the larger dune masses might lag, the overall patterns suggest minimal record of winds from the opposite phase of the 51,000-year cycle of perihelion season.

6. Conclusion

The scarcity of longitudinal dunes on Mars is entirely consistent with the unimodal transport regimes expected and observed for most of the planet. In fact, the scarcity of longitudinal forms on Mars can be taken as reinforcing the generalization of their formation under bimodal or multimodal transport regimes. Spacecraft missions such as Mars Global Surveyor will help amplify our knowledge of present and past eolian action on Mars by providing high-resolution views of dune morphology and global seasonal coverage of evidence of changes that show the orientation and frequency of activity of effective winds.

Acknowledgments. This research was funded in part by NASA grants NAGW-3274 and NAGW-3275. Discussions with A. Bloom, D. Blumberg, K. Edgett, P. Gierasch, R. Greeley, and J. Veverka were helpful. GCM run results were kindly supplied by D. Blumberg on short notice. R. Sullivan and W. Ward provided very helpful and detailed reviews. We also thank Mary Roth for help in preparing the manuscript.

References

Arvidson, R.E., E.A. Guinness, H.J. Moore, J. Tillman, and S.D. Wall, Three Mars years: Viking Lander 1 imaging observations, *Science, 222,* 463-468, 1983.

Bagnold, R.A., *The Physics of Blown Sand and Desert Dunes*, Methuen, New York, 1941.

Baird, A.K., P. Toulmin, III, B.C. Clark, H.J. Rose, Jr., K. Keil, R.P. Christian, and J.L. Gooding, Mineralogic and petrologic implications of Viking geochemical results from Mars: Interim results, *Science*, *194*, 1288-1293, 1976.

Breed, C.S., and T. Grow, Morphology and distribution of dunes in sand seas observed by remote sensing, *U.S. Geol. Surv. Prof. Pap., 1052*, 253-302, 1979.

Breed, C.S., S.G. Fryberger, S. Andrews, C. McCauley, F. Lennartz, D. Gebel, and K. Horstman, *U.S. Geol. Surv. Prof. Pap., 1052*, 305-397, 1979a.

Breed, C.S., M.J. Grolier, and J.F. McCauley, Morphology and distribution of common "sand" dunes on Mars: Comparison with the Earth, *J. Geophys. Res., 84*, 8183-8204, 1979b.

Christensen, P.R., The spatial distribution of rocks on Mars, *Icarus, 68*, 217-238, 1986.

Clark, B.C., and D.C. Van Hart, The salts of Mars, *Icarus, 45*, 370-378, 1981.

Clark, B.C., A.K. Baird, R.J. Weldon, D.M. Tsusaki, L. Schnabel, and M.P. Candelaria, Chemical composition of Martian fines, *J. Geophys. Res.*, *87*, 10059-10067, 1982.

Cooke, R.V., A. Warren, and A.S. Goudis, *Desert Geomorphology*, 526 pp., UCL Press, London, 1993.

Edgett, K.S., and D.G. Blumberg, Star and linear dunes on Mars, *Icarus, 112*, 448-464, 1994.

Fanale, F.P., and W.A. Cannon, Exchange of adsorbed H_2O and CO_2 between the regolith and atmosphere of Mars caused by changes in surface insolation, *J. Geophys. Res.*, *79*, 3397-3402, 1974.

French, R.G., and P.J. Gierasch, The Martian polar vortex: Theory of seasonal variation and observations of eolian features, *J. Geophys. Res., 84*, 4634-4642, 1979.

Fryberger, S.G. Dune forms and wind regime. *U.S. Geol. Surv. Prof. Pap., 1052*, 137-170, 1979.

Fryberger, S.G., and T.S. Ahlbrandt, Mechanisms for the formation of eolian sand seas, *Z. Geomorphol., 23*, 440-460, 1979.

Fryberger, S.G., and A.S. Goudie, Progress report: arid geomorphology, *Prog. Phys. Geogr., 5*, 409-428, 1981.

Gooding, J.L., S.J. Wentworth, and M.E. Zolensky, Aqueous alteration of the Nakhla meteorite, *Meteoritics, 26*, 135-143, 1991.

Greeley, R., and J.D. Iversen, *Wind as a Geological Process on Earth, Mars, Venus, and Titan*, Cambridge University Press, New York, 1985.

Greeley, R., and S.H. Williams, Dust deposits on Mars: The "parna" analog, *Icarus*, *110*, 165-177, 1994.

Greeley, R., N. Lancaster, S. Lee, and P. Thomas, Martian aeolian processes, sediments, and features, in *Mars,* edited by H. Kieffer, B. Jakosky, C. Snyder, and M. Matthews, pp. 730-766, University of Arizona Press, Tucson, 1992.

Greeley, R., A. Skypeck, and J.B. Pollack, Martian aeolian features and deposits: Comparisons with general circulation model results, *J. Geophys. Res., 98*, 3183-3196, 1993.

Haberle, R. M., J.B. Pollack, J.R. Barnes, R.W. Zurek, C.B. Leovy, J.R. Murphey, H. Lee, and J. Schaeffer, Mars atmospheric dynamics as simulated by the NASA Ames general circulation model, 1, The zonal-mean circulation, *J. Geophys. Res., 98*, 3093-3123, 1993.

Jakosky, B.M., and P.R. Christensen, Global duricrust on Mars: Analysis of remote-sensing data, *J. Geophys. Res., 91*, 3547-3559, 1986.

James, P.B., G. Briggs, J. Barnes, and A. Spruck, Seasonal recession of Mars' south polar cap as seen by Viking, *J. Geophys. Res., 84*, 2889-2922, 1979.

Kadib, A. A., A Function of Sand Movement by Wind, *Tech. Rept. HEL-2-12*, Univ. of California, Berkeley, 1965.

Kahn, R., Some observational constraints on the global-scale wind systems of Mars, *J. Geophys. Res., 88*, 10189-10209, 1983.

Kawamura, R., Study of sand movement by wind, Rep. 313, Inst. of Sci. and Tech. 5, Univ. of Tokyo, 1951.

Lancaster, N., The formation of seif dunes from barchans—Supporting evidence for Bagnold's model from the Namib Desert, *Z. Geomorph., NF24*, 160-167, 1980.

Lancaster, N., Linear dunes, *Prog. in Phys. Geogr., 6*, 476-504, 1982.

Lancaster, N., Star dunes, *Prog. in Phys. Geogr., 13*, 67-91, 1989.

Lancaster, N., and R. Greeley, Sediment volume in the north polar sand seas of Mars, *J. Geophys. Res., 95*, 10921-10927, 1990.

Lee, P., Briny lakes on early Mars? Terrestrial intracrater playas and martian candidates (abstract), in Workshop on Early Mars: How Warm and How Wet?, edited by S. Squyres and J. Kasting, *LPI Tech. Rep., 93-03*, Part 1, 17, 1993.

Lee, P., P.C. Thomas, J. Veverka, and S. Calvo, Discovery of longitudinal dunes on Mars (abstract), *Bull. Am. Astron. Soc., 25*, 1038, 1993.

Lettau, K., and H. Lettau, Experimental and micrometeorological field studies of dune migration, in *Exploring the World's Driest Climate,* edited by K. Lettau and H. Lettau, University of Wisconsin Press, Madison, 1977.

Magalhães, J.A., The Martian Hadley circulation: Comparison of "viscous" model predictions to observations, *Icarus, 70,* 442-468, 1987.

Malin, M.C., Salt weathering on Mars, *J. Geophys. Res.*, *79*, 3888-3894, 1974.

McKee, E.D., Introduction to a study of global sand seas, in A Study of Global Sand Seas, edited by E.D. McKee, *U.S. Geol. Surv. Prof. Pap. 1052*, 1979.

McSween, H.Y., Jr., SNC meteorites: Clues to Martian petrologic evolution?, *Rev. Geophys., 23*, 391-416, 1985.

McSween, H.Y., Jr., What we have learned about Mars from SNC meteorites, *Meteoritics, 29*, 757-779, 1994.

Pollack, J.B., R.M. Haberle, J. Schaeffer, and H. Lee, Simulations of the general circulation of the Martian atmosphere, 1, Polar processes, *J. Geophys. Res., 95*, 1447-1473, 1990.

Sagan, C., D. Pieri, P. Fox, R.E. Arvidson, and E.A. Guinness, Particle motion on Mars inferred from the Viking lander cameras, *J. Geophys. Res., 82*, 4430-4438, 1977.

Sarre, R.D., Aeolian sand transport, *Prog. in Phys. Geogr., 11*, 157-182, 1987.

Thomas, P., North-south asymmetry of eolian features in Martian polar regions: Analysis based on crater-related wind markers, *Icarus, 48*, 76-90, 1981.

Thomas, P., Present wind activity on Mars: Relation to large latitudinally zoned sediment deposits, *J. Geophys. Res., 87*, 9999-10,008, 1982.

Thomas, P.C. and P.J. Gierasch, Polar margin dunes and winds on Mars, *J. Geophys. Res.*, this issue.

Thomas, P., and J. Veverka, Seasonal and secular variations of wind streaks on Mars: An analysis of Mariner 9 and Viking data, *J. Geophys. Res., 84*, 8131-8146, 1979.

Thomas, P., J. Veverka, and R. Campos-Marquetti, Frost streaks in the south pola cap of Mars, *J. Geophys. Res., 84*, 4621-4633, 1979.

Thomas, P., J. Veverka, D. Gineris, and L. Wong, "Dust" streaks on Mars, *Icarus, 60*, 161-179, 1984.

Toulmin, P., III, A.K. Baird, B.C. Clark, K. Keil, H.J. Rose Jr., R.P. Christian, P.H. Evans, and W.C. Kelliher, Geochemical and mineralogical interpretation of the Viking inorganic chemical results, *J. Geophys. Res., 82*, 4625-4634, 1977.

Treiman, A.H., R.A. Barrett, and J.L. Gooding, Preterrestrial aqueous alteration of the Lafayete (SNC) meteorite, *Meteoritics, 28*, 86-97, 1993.

Tsoar, H., Linear dunes—forms and formation, *Prog. in Phys. Geogr., 13*, 507-528, 1989.

Tsoar, H., R. Greeley, and A. R. Peterfreund, Mars: The north polar sand sea and related wind patterns, *J. Geophys. Res., 84*, 8167-8180, 1979.

Ugolini, F.C., and D.M. Anderson, Ionic migration and weathering in frozen Antarctic soils, *Soil Sci., 115*, 461-470, 1973.

Verstappen, H.T., On the origin of longitudinal (seif) dunes, *Z. Geomorph., 12*, 200-212, 1968.

Veverka, J., P. Thomas, and R. Greeley, A study of variable features on Mars during the Viking primary mission, *J. Geophys. Res., 82*, 4167-4187, 1977.

Veverka, J., P. Gierasch, and P. Thomas, Wind streaks on Mars: Meteorological control of occurrence and mode of formation, *Icarus, 45*, 154-166, 1981.

Ward, A.W., and K.B. Doyle, Speculation on Martian north polar wind circulation and resultant orientations of polar sand dunes, *Icarus, 55*, 420-431, 1983.

Wasson, R.J. and R. Hyde, Factors determining desert dune type, *Nature, 304*, 337-339, 1983.

Wentworth, S.J., and J.L. Gooding, Carbonates and sulfates in the Chassigny meteorite: Further evidence for aqueous chemistry on the SNC parent planet, *Meteoritics, 29*, 860-863, 1994.

White, B.R., Soil transport by winds on Mars, *J. Geophys. Res., 84*, 4643-4651, 1979.

Williams, S.H., and J.R. Zimbelman, White Rock: An eroded Martian lacustrine deposit, *Geology, 22*, 107-110, 1994.

P. Lee and P.C. Thomas, Center for Radiophysics and Space Research, Cornell University, Space Sciences Building, Ithaca, NY 14853-6801. (e-mail: lee@astrosun.tn.cornell.edu)

(Received March 9, 1994; revised November 28, 1994; accepted January 13, 1995.)

Page intentionally left blank

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5397–5406, MARCH 25, 1995

Polar margin dunes and winds on Mars

Peter C. Thomas and Peter J. Gierasch

Center for Radiophysics and Space Research, Cornell University, Ithaca, New York

Abstract. The approximately concentric arrangement of layered deposits and dune fields at the two Martian poles may reflect a nearly steady state dispersal of material from the polar deposits. Data on effective surface winds from high resolution Viking Images combined with theory of local winds suggest that the northern dunes are in part confined to a latitude band by winds generated by their own low albedo. Dispersal of the dark sand from the southern polar region is not subject to this kind of feedback because the irregular topography prevents areal accumulations sufficiently extensive to produce winds.

Introduction

Both polar regions of Mars have crudely concentric arrangements of frosts and large sediment deposits: perennial frost caps, layered deposits, dune fields, and a variety of "mantling" deposits. However, the frost and sediments show some significant differences between the poles, among them the different perennial ice (water in the north; CO_2 in the south), and the contrast of a widespread (though incomplete) annulus of dunes about the northern layered deposits with scattered dune fields in isolated depressions in the south. Because the layered deposits at both poles appear to have suffered considerable erosional reduction from their maximum volumes, and the dunes appear to have their sources within the layered deposits, the arrangement of the dunes may reveal some of the mechanisms for the climate and geological influences on the evolution of the polar deposits [see *Murray et al.*, 1972; *Squyres*, 1979; *Howard et al.*, 1982; *Thomas*, 1982; *Plaut et al.*, 1988; *Thomas and Weitz*, 1989]. The fundamental geological problem at issue here is the cause of the narrow latitudinal banding of dunes near the north polar deposits, and the much wider zone of intracrater dunes near the south polar cap.

In this study we first briefly review previous work on winds and the polar dunes. Next we present the mapping results for the north and south polar regions and note the geologic settings that may influence winds affecting the dune materials such as albedo contrasts and topography. Then terrestrial sea-breeze theory is reviewed and applied to the dark Martian dune fields that may provide strong surface temperature gradients because of their low albedo. Finally, our conclusion that sea-breeze effects may be important in confining the north polar dunes is discussed as part of the overall evolution of the polar cap sedimentary systems. For this work we have presented only a few schematic drawings of the eolian features in Figure 1; complete sets of images and background on their interpretation for wind directions and eolian processes can be found in the works of *Breed et al.* [1979]; *Tsoar et al.* [1979]; *Greeley et al.* [1992], and many others.

Paper number 94JE02639.
0148-0227/95/94JE-02639$05.00

Previous Studies of Polar Dune Winds

Tsoar et al. [1979] found that most of the north polar dunes had morphologies that show the effects of multiple wind directions. They suggested cyclonic winds due to temperature contrasts of cold, ice-covered areas and ice-free areas would be significant in controlling dune activities. *Thomas* [1982] mapped latitudinal variations of flow on-pole and off-pole that might help confine the dunes in both polar regions. Ward and Doyle [1983] suggested that changing east-west wind components at about 80°N contributed to dune occurrence and orientations. *Howard* [1980] mapped a very consistent pattern of summer flow within the north polar deposits (mostly above 80°N) retrograde out from the pole (from the northeast). *Tsoar et al.* [1979] additionally noted that winds in the polar regions were not simply latitudinally dependent and might fall into regional wind regimes.

An axisymmetric model of winds at the Martian poles by *French and Gierasch* [1979] indicated that the highest winds in the polar regions formed spiral patterns outward from the pole, prograde in the fall during condensation of the annual frost cap (to the southeast in the northern region), and retrograde (to the southwest) in spring. These calculations showed good agreement with south polar wind indicators mapped by *Thomas et al.* [1979]. The general circulation model (GCM) calculations reported by *Greeley et al.* [1993] show strong fall and winter winds from the west at 80°N, and moderate winds from the east in spring. The GCM does not appear to give significant north-south wind components. *Parish and Howard* [1993] used three-dimensional modeling to find that katabatic winds off the polar deposits could be very effective, but their intensity and directions could be substantially altered in summer by the increased insolation on slopes.

The models account for many aspects of the wind flow, but have not quantitatively addressed confinement wind regimes. We address this problem with a sea breeze wind model in conjunction with a digital map of features

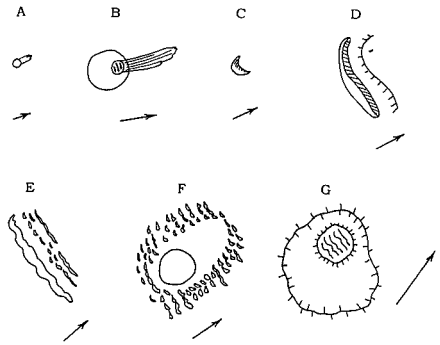


Figure 1. Examples of eolian features mapped with effective wind orientations. (a) Crater related wind streaks. These may be dark or light and are formed by enhanced erosion or deposition of dust downwind of craters or other topographic obstacles. (b) Splotch streaks; also called type II streaks. Dark material deflated from deposits (splotches) in craters. The orientation of the splotches is also a wind indicator. (c) Crescentic dunes. (d) Echo dunes. These form upwind of topographic obstacles. (e) Framing dunes. These large, transverse form are on the upwind margins of dune fields. (f) Dune field shadows. These are areas lacking dunes in the lee of obstacles caused by diversion of sand flow and increased complexity of wind patterns. (g) Intracrater dune fields. These usually form on the downwind side of topographic depressions. Some are discrete dunes, but many are continuous, thick mounds, with tranverse bedforms on top.

together with published data on the dune fields and other eolian features. We study both polar regions since the differences in geological setting may indicate the relative effects of climate and geology on the disposition of surface sediments in the two polar regions.

Methods

For this study we desired a digital data set of wind direction markers that spans different sizes of features (hence possibly different formation timescales), and that had been collected in a uniform manner. Part of the strategy was to make as complete a search as possible, but to deal with only the unambiguous markers. We have examined the higher resolution Viking images of polar regions (latitudes greater than 60°) for eolian features that indicate surface wind directions such as dunes (crescentic, longitudinal, transverse dunes), wind streaks in dust, sand, and frost, and wind shadow patterns in dune fields that indicate effective wind directions. Of particular importance are the sizes of some of the features that may crudely indicate the length of time the effective winds would have had to operate. We have not mapped every dune form that has an interpretable effective wind, but only the representative examples from each image, to give a clear representation of the geographic coverage of types and wind directions rather than a complete catalog of all features. Ambiguous features were avoided. Table 1 lists the features mapped and the typical sizes and some crude estimates of possible relative effective timescales based on parameters from *White* [1979] and limits discussed in *Thomas* [1982]. Because the Martian climate is expected to show hemispherical reversal of some seasonal effects on the timescale of changing season of perihelion (51,000 years full cycle; *Arvidson et al.* [1979]; *Kieffer and Zent* [1992], the larger eolian forms might be the result of a range of climate conditions and effective wind regimes. The data have been entered into a digital file and in conjunction with global wind streak data reported earlier [*Thomas*, 1981; *Thomas et al.*, 1984] allow a much more complete picture of the winds that have done geologic work near the Martian polar regions.

Results: North Polar Region

Figure 2 shows cylindrical projection maps of several of the eolian features in the north polar region: wind streaks, crescentic dunes and associated streaks, and larger dune features such as framing dunes, echo dunes, and dune field shadows (see Figure 1). As with the data of *Howard* [1980], *Tsoar et al.* [1979], the streaks in latitudes 80°-85° show consistent retrograde winds off the cap. The pattern becomes more complex below 80°N, and as shown in many other studies [e.g., *Thomas et al.*, 1984], the pattern in mid-northern latitudes becomes prograde winds merging into the global, southern summer flow.

The crescentic dunes show a complex pattern of flow to the southwest at higher latitudes and flow to the northeast at some longitudes in the regions below 80°N. The only region of apparent through-flowing material is in the longitude range 30°-100°. This material, in fact, feeds into the large low albedo region of Acidalia [see also *Thomas et al.*, 1984; *Christensen*, 1988]. Many of the crescentic dunes in this region, where the dune material is freely moving from the pole, are elongate barchans and even linear dunes (P. Lee et al., Longitudinal dunes on Mars, submitted to *Journal of Geophysical Research,* 1994) and indicate moderately different wind directions for long-term gentle winds, and shorter term, stronger flow. Differences in wind directions inferred from the asymmetric and elongate barchans can be nearly 90°, but are usually 45° or less.

The echo dunes, framing dunes, and dune field shadows represent eolian forms that require much larger timescales to form than do individual crescentic dunes (see Table 1 and Figure 1). Reorientation of a crescentic dune would take 1 to 2 orders of magnitude less mass transport than the rearrangement of a field of dunes such as shown in Figure 1c, or the orientation of the large framing dunes or dunes formed in echo patterns around large topographic obstacles. These forms also show the same patterns of flow that include significant on-pole winds, as do the erosional outliers of polar layered deposits shown in Figure 2d.

The observed flow patterns suggest that the dunes are confined by regional winds, that poleward flow is substantial and has persisted over at least moderate geologic

Table 1. Eolian Features and Possible Timescales

Feature	Applicable Scale	"Measurement" Method	Timescale
Bright and dark "dust" streaks	micrometers to centimeters	periodic images	days
Dark streaks from dunes	millimeters to centimeters	periodic images	seasonal
Dune field bedforms; individual dunes	100-200 m	saltation models	10^3-10^4 years
Echo, framing dunes; dune field shadows	1-30 km	saltation models	10^4-10^5 years
Intracrater dune fields	10-500 m	saltation models	10^4-10^6 years
Latitudinally distributed dunes	100-1000 km	saltation models	10^6-10^7 years
Polar layers	tens of meters	dust and ice deposition models	10^5-10^6 years
Polar layered deposits (PLD)	1-5 km	dust and ice deposition models	10^7 + years
Erosion of troughs in PLD	100 m to 1 km	guess (crater densities)	10^7 years
PLD extent	hundreds of kilometers	guess (crater densities)	10^8 years

Features are arranged approximately in order of increasing volume; see Figure 1 for examples. Applicable scale is the dimension of material that has to be moved to form the feature. For some features, such as dust streaks, individual dunes, and layers this is the thickness of material that must be deposited or removed. For features that require lateral rearrangement to form, the scale is a horizontal motion distance. Measurement method is the means by which the timescale for change has been estimated. Timescale is the time that the features may require under present Martian conditions to reform; see *Thomas* [1982] and *Thomas et al.* [1992] for background. These numbers are meant as only order of magnitude, relative values except for the first two rows.

intervals, and that "leakage" of some of the material south contributes to albedo markings that may extend to the equator.

Results: South Polar Region

Wind markers in the south are very different from those in the north. Dunes are largely confined locally in craters or by wrinkle ridges on plains. Although the surface area covered by dunes is smaller in the south than in the north, the volumes of sand may be just as large as in the north because of deep fill in many of the southern craters [*Thomas*, 1982; see also *Lancaster and Greeley*, 1990]. Material blown from the crater deposits (splotches) forms many streaks (see Figure 1b) in the south, but only a few in the north. Likewise, the concentration of dunes in mounds inside craters in the south, and the lower image resolution in this region result in very few interpretable individual bedforms in the southern data compared to those in the north.

Figure 3 shows the maps of southern wind markers; note that the latitude range is greater than in Figure 2.

The dark, crater-related streaks, likely formed by erosion of thin dust layers and observed to be the most ephemeral of these markers, show a complex pattern with significant on-pole flow (Figure 3a). This on-pole flow is distinct from that in the north in three major ways: it is retrograde, it is shown by the shorter-lived features, and the on-pole wind markers are at lower latitudes than those in the north.

The splotch streaks (Figure 3b) are probably the next most rapidly modified features and show consistent retrograde flow from the polar regions. The splotch orientations within craters, likely representative of a more substantial deposit and longer timescale of modification, also show consistent flow off-pole (Figure 3c).

The orientation of bedforms on dune masses show more variation than do the streaks from splotches (Figure 3d). The bedforms mapped are nearly all transverse ridges and thus have 180° ambiguities. About half these transverse ridges could have been formed by winds similar to those inferred from the streaks, and from the dune field and splotch orientations (Figures 3c and 3e). However, those that could show a southwest or northeast flow require different winds. Such different effective winds could have occurred at different seasons, or different epochs, or could result from local characteristics of flow within the topographic trap (usually craters) of the dune mass.

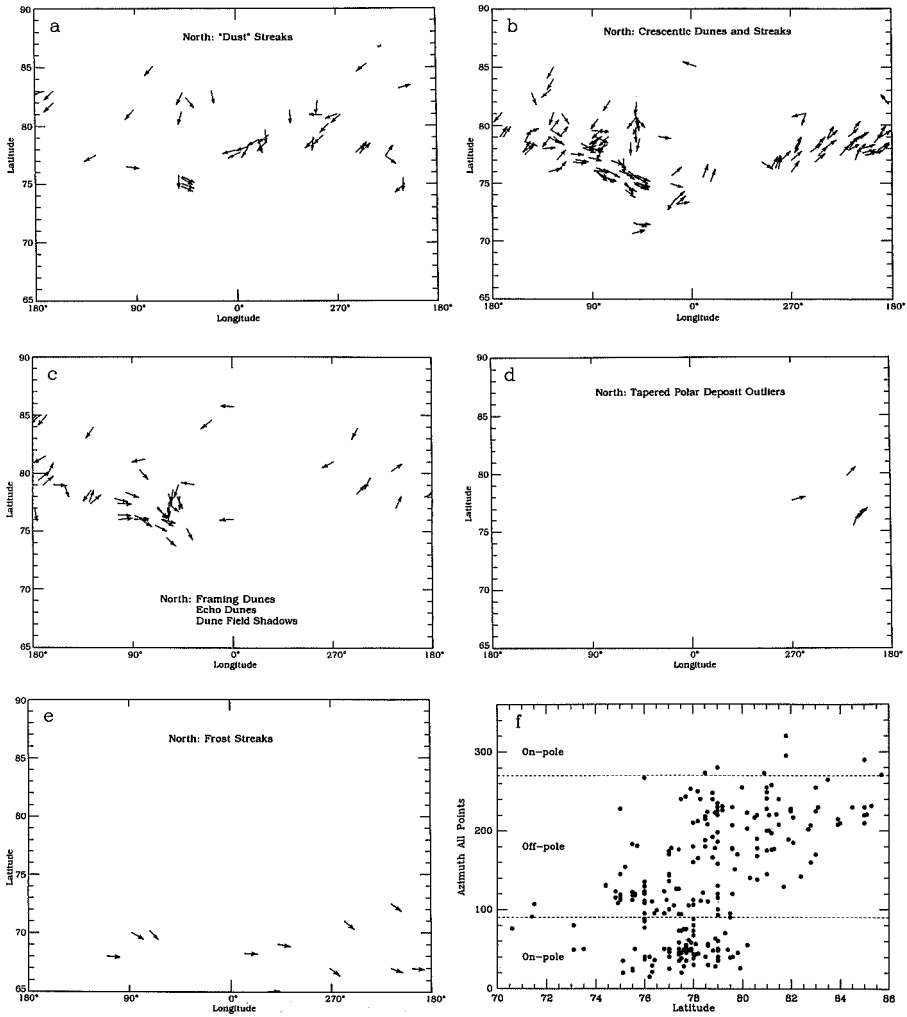


Figure 2. Wind markers in the north polar region of Mars. Data digitized from Viking Orbiter Images. (a) Dust streaks: crater-related streaks. (b) Crescentic dunes and associated streaks. (c) Framing dunes, echo dunes, and dune field shadows. (d) Tapered outliers of polar deposits. (e) Frost streaks in annual CO_2 cap. (f) Wind marker azimuth as a function of latitude. Dashed lines denote boundary of flow from the pole (90°-270°) and flow toward the pole.

The dune field azimuths recorded from the high-resolution images provide strong support for the use of splotch orientations as representing winds over time periods longer than the splotch streaks. Some of the dune fields that are well positioned against crater walls show that the effective winds have had to operate over times adequate to move very large quantities of sand.

For complete comparison of north and south wind markers, frost streaks formed in the annual cap are shown in Figures 2e and 3f. Both show prograde, off-pole flow

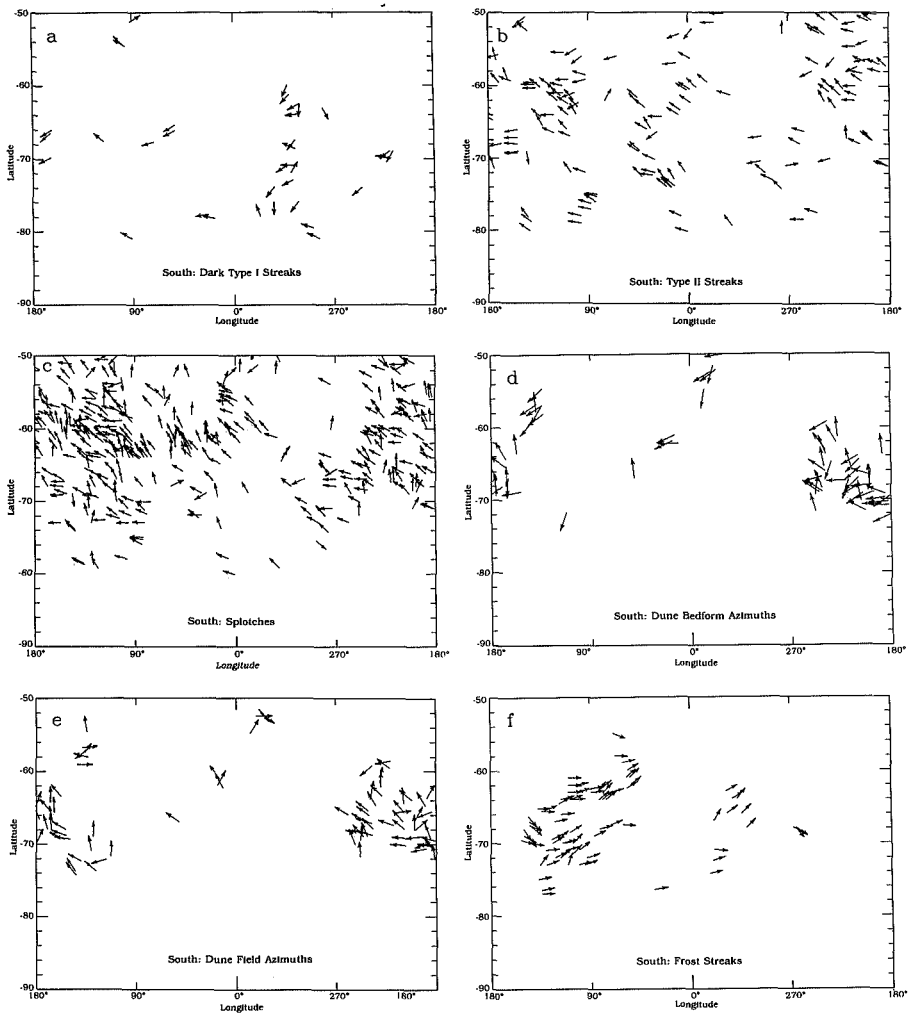


Figure 3. *Wind markers in the south polar region.* (a) Dark, crater-related streaks. (b) Splotch streaks: dark material blown from crater deposits. (c) Crater splotches: these are interpreted as occurring on downwind side of craters. (d) Bedforms on dune fields in craters. These are transverse ridges on a continuous, often deep mantle of material (sand?). Azimuths shown here are interpreted as winds with components to the west; all have 180° ambiguity. (e) Azimuths of dune fields; these are interpreted from morphology of dune field within crater and are similar to splotches in Figures 3b and 3c. (f) Frost streaks in the southern seasonal frost cap.

and were the basis for much of the French and Gierasch model of polar winds.

Geological Settings of Polar Dune Deposits

Before our testing of sea breeze influences on the dune field locations, we discuss some aspects of the geological settings. As noted by several works [*Breed et al.*, 1979; *Howard et al.*, 1982; *Thomas*, 1982) there is strong morphologic evidence for dune sources in the layered deposits, and color measurements support this inference [*Thomas and Weitz*, 1989; *Herkenhoff and Murray*, 1990]. The reexamination of the images for this study has reinforced these data by the absence of indications that any of the dunes in the major chasma and reentrants are moving on-pole. The digital mapping did not include many of the common but subtle streaks emanating from scarps in the polar layered deposits (PLD's) that add to the evidence of recent dune material erosion from the polar layers. The restriction of the on-pole flow to dune areas at and beyond the margins of the northern PLD makes it very difficult to invoke present conditions for surface motion of the dune material to latitudes well above 80°N.

Although the north polar sand seas are often classed as being a ring about the pole, *Lancaster and Greeley* [1990] have emphasized that in terms of volumes of material, the sand seas are fragmented, and are dominated by the material between longitudes 110° and 250°; half the north polar sand is probably in the longitudes 130° to 230°. These dunes are in fact not simply south of the polar deposits; they are substantially contained between the polar deposits and icy outliers.

The complexity of the north polar dunes is well illustrated by maps of *Tsoar et al.* [1979] and *Lancaster and Greeley* [1990]. The plots in Figure 4 include approximate outlines of the major dune volume concentrations from *Lancaster and Greeley* [1990] and show that the regions of significant dune cover occur on latitudinal scales of about 4°-6°, or about 2°-3° (120-180 km) from margins of the dune field to the central regions of maximum dune cover and the region of apparent reversal of wind flow. Dune fields in the south polar region are all less than 70 km in extent; most are less than 20 km across. The south polar dunes are also at lower latitudes and are not confined by ice outliers, but by local depressions such as craters.

Regional topography is poorly known in the Mars polar regions; the digital data provided by the *U.S. Geological Survey* [1993] indicates a possible wide depression leading from about 75°N and longitudes 20°-90°W toward the south; elevations for most of the dune areas are about 0.5 to 3 km below the reference datum. In the south elevations average 2.5-5 km above the reference at 75°S. In

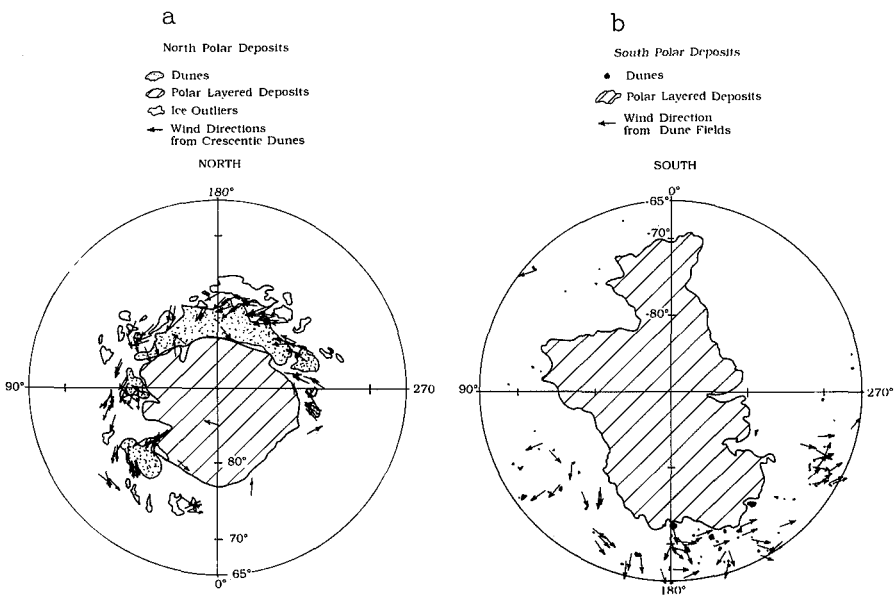


Figure 4. Schematic diagram of major deposits at north and south poles. Outlines of polar deposits after *Tanaka and Scott* [1987]; *Thomas et al.* [1992]. (a) North pole: layered deposits, areas of major dune volumes modified from *Lancaster and Greeley* [1990], wind directions indicated by crescentic dunes, and ice outliers. (b) South pole: layered deposits and associated pitted deposits, dunes, and intracrater dune field orientations.

both poles the layered deposits may reach 1-2 km above the surrounding plains.

Temperature observations from the Viking Orbiter infrared thermal mapper (IRTM) have been reported for the polar areas by *Keegan et al.* [1991], *Paige and Keegan* [1991], and *Paige et al.* [1994]. Data of *Paige et al.* [1994] and data reported by *Tsoar et al.* [1979] for the region 70°-90° and longitude 130°-230° suggest temperature contrasts of as much as 20 K with the ice to the south over distances of a few tens of kilometers. Average maximum summer daytime temperature contrasts are probably about 5-10 K; ice temperatures were approximately 220 K.

The thermal inertia of dunes on Mars is compatible with particle sizes expected to saltate on Mars [*Edgett and Christensen*, 1991, 1994]. The north polar dunes have inertias that may be lower than most other Mars dunes, but are still consistent with saltating particles [*Keegan et al.*, 1991]. The inferred approximate particle sizes means that these materials should be some of the most easily moved materials on Mars [*White*, 1979; *Greeley et al.*, 1992].

Sea Breeze Theory and Extension to Mars

Haberle et al. [1979] carried out a zonally symmetric general circulation computation of Martian polar cap winds and found that strong winds of 20 to 30 m s^{-1} developed in relatively narrow zones near the edge of the seasonal CO_2 cap. Here we shall argue that terrestrial sea breeze theory, which is quite simple and easy to parameterize, can be carried over to the Mars situation and could be used to estimate the local enhancement of winds near cap edges and albedo boundaries where a sharp surface temperature gradient exists. The results are consistent with the cap edge circulations predicted by Haberle et al. and lead to simple analytic expressions that are convenient.

The sea breeze on Earth is well understood. *Rotunno* [1983] presents a review of the subject and discusses a simple linear model that works well for the diurnally periodic sea breeze. *Dalu and Pielke* [1989] extend Rotunno's work to the impulsively driven or the steady case. Modern observations with lidar techniques show that the linear theory provides a good description of the phenomenon [*Banta et al.*, 1993]. A key result is that the circulation develops a horizontal scale determined by the stratification of the atmosphere, independent of the details of the sharp horizontal gradient in heating that produces the flow. The horizontal scale is

$$L = \frac{N}{(\lambda^2 + f^2)^{1/2}} D, \qquad (1)$$

where D is the depth of the heated layer, f is the Coriolis parameter, λ is the frictional or thermal damping rate (these are assumed to be the same), and N is the Brunt frequency. For Mars parameters we shall show that the same physics is valid, and we shall show that the scale L is about 100 km, in good agreement with the width of the zones of dune activity near albedo contrasts in the north polar region.

The clearest way to present the sea breeze model is to reproduce the governing equations from *Dalu and Pielke* [1989]. Let x be the horizontal coordinate parallel to the temperature gradient that forces the circulation, y be along the "coast" or albedo boundary, and z be upward, with x, y, z forming a right-handed system. Let the corresponding velocities be u, v, w. Because the circulation is much wider than its depth (about 100 km compared to about 1 km) the vertical force balance is hydrostatic. The model is

$$\begin{aligned} \partial_t u - fv + \partial_x \Phi &= -\lambda u, \\ \partial_t v + fu &= -\lambda v, \\ \partial_z \Phi &= g\, T/T_0, \\ \partial_x u + \partial_y v &= 0, \\ \partial_t T + T_0(N^2/g)w &= \lambda(T_F - T) \end{aligned} \qquad (2)$$

The notation is standard. The drive for the circulation is a forcing temperature perturbation $T_F(x,z,t)$. At the surface it is equal to the departure of the surface temperature from the mean reference value T_0. We are primarily interested here in the steady case with $\partial_t = 0$, with the flow driven by a mean temperature difference produced by surface albedo variation. Diffusion, radiation, and frictional drag are all formulated in terms of a constant linear damping coefficient λ. Parameter Φ is the pressure perturbation divided by the mean density. Note that a wind parallel to the isotherms, v, is produced only through the effect of the Coriolis parameter f.

The depth of the circulation is determined by the vertical distribution of $T_F(x,z,t)$. This involves boundary layer heating mechanisms and depends on the density of the atmosphere because of its influence on the thermal inertia. On the Earth the effective depth for boundary layer heating effects is typically a few hundred meters, and is verified by direct observations [see *Stull*, 1988]. On Mars there are not yet any observations of the thermal boundary layer thickness, but theoretical treatments indicate that it is a bit larger than 1 km [*Blumsack et al.*, 1973; *Pollack et al.*, 1979]. Thus an important difference between Mars and the Earth is that the boundary layer on Mars is about 5 times deeper.

The appropriate damping rate λ is found to be about 10^{-4} s^{-1} for the Earth, which is of the same order as the Coriolis parameter. A value of this same order is probably also appropriate for Mars. For example, a drag coefficient can be estimated from bulk drag considerations by writing

$$D\partial_t u = -C_D u^2$$

where D is the boundary layer depth and C_D is a dimensionless drag coefficient, which should be the same on Mars or the Earth if the surface roughness characteristics are not very different. This expression leads to an estimate for the boundary layer damping rate of $\lambda = C_D U/D$, where U is a typical wind speed. Wind speeds are higher on Mars than on Earth [*Zurek et al.*, 1992]. With $C_D = 0.003$ [e.g., *Stull*, 1988], $U = 20$ m s^{-1}, and a midlayer depth $D = 0.75$ km, we obtain approximately $\lambda = 10^{-4}$ s^{-1}. In addition, the radiative relaxation rate on Mars for a layer of thickness 1.5 km is about 10^{-4} s^{-1}

(which is part of the reason that this thickness is predicted) [*Goody and Belton*, 1967]. Thus both momentum and heat transfer considerations are consistent with $\lambda = 10^{-4}$ s^{-1}. The uncertainty in this parameter is probably a factor of 3 or so.

Because the Coriolis parameter at high latitudes on Mars is approximately $f = 10^{-4}$ s^{-1}, the same size as λ, there will be appreciable turning of the sea breeze wind by Coriolis effects.

Rotunno [1983] and *Dalu and Pielke* [1989] argue that the sea breeze on Earth is well described by the linear model, and it remains to verify that this will also be true of Mars. Relative to the frictional drag, nonlinear advection terms are of the order of $u/(\lambda L)$, where u is the sea breeze wind itself. With $u = 10$ m s^{-1}, $\lambda = 10^{-4}$ s^{-1}, and $L = 100$ km, we find $u/(\lambda L) = 1$, which indicates that nonlinear terms should have important effects. The same conclusion is reached in the terrestrial case, where both u and L are about a factor of 3 smaller. Nevertheless, the linear theory is a good approximation on Earth, and we shall assume that it also is on Mars.

In summary, the mechanisms that produce the sea breeze on Earth should carry over to Mars and produce local wind systems where there are sharp gradients of the surface temperature (scales smaller than 100 km). The most significant difference between Mars and Earth is that the depth of these flows is greater on Mars and, as a result, the width is also greater. Equation (1), from the results of *Dalu and Pielke* [1989], is used to estimate the width of the circulation. The Brunt frequency is given by

$$N^2 = \frac{g}{T_0}\left(\frac{\partial T}{\partial z} + \frac{g}{c_p}\right) \tag{3}$$

With the acceleration of gravity $g = 380$ cm s^{-2}, specific heat $c_p = 0.7 \times 10^7$ erg g^{-1} K^{-1}, and assuming an isothermal mean temperature, we find $N \approx 10^{-2}$ s^{-1}. Using (1) with $D = 1.5$ km and $\lambda = f = 10^{-4}$ s^{-1} gives $L \approx 100$ km.

The wind speed, according to Dalu and Pielke, is given by

$$u \approx \frac{g}{N}\frac{\Delta T_F}{T_0} \approx (380\, m s^{-1})\frac{\Delta T_F}{T_0} \tag{4}$$

where ΔT_F is the horizontal temperature contrast at the surface. This can be estimated from

$$\sigma T^4{}_S = (1 - A) S, \tag{5}$$

where T_S is the surface temperature and S is the flux of insolation. Letting ΔA be the horizontal albedo contrast and linearizing (5) give

$$\frac{\Delta T_F}{T_0} = \frac{\Delta A}{4(1 - a)} . \tag{6}$$

Maps of north polar region albedos [*Paige et al.*, 1994] show dune area albedos of about 0.15 and adjacent areas with albedos of about 0.3 and higher on ice outliers. Using these albedos in (6) yields fractional temperature differences of about 5%, which implies wind speeds from

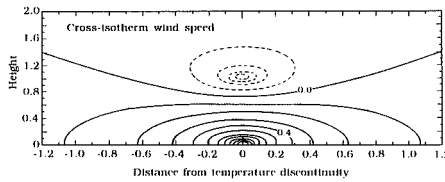


Figure 5. Horizontal wind velocity from sea breeze theory. The units for horizontal and vertical disturbances are about 100 km and 1.5 km, respectively. The wind speed contours are scaled by the maximum value, which would typically be given by (4). The results displayed here are for a step function profile of forcing. Adapted from *Dalu and Pielke* [1989].

(4) of about 20 m s^{-1} (Figure 5). These temperature differences are consistent with the measurements reported by *Keegan et al.* [1991] and *Paige et al.* [1994].

These winds probably do not have sufficient strength to initiate saltation and transport, but because the sea breeze mechanism is linear to a first approximation, the winds are additive and will augment other circulations that might be present. These include the tides, the mean circulation of the polar vortex, baroclinic instabilities, slope winds and local vortices [*Zurek et al.*, 1992]. The local wind systems produced by albedo-induced temperature contrasts probably act in conjunction with other systems to produce the episodes of strongest wind that dominate sediment transport. The topographic contrast of the polar deposits and surrounding plains may enhance the confinement by off-pole flow [*Parish and Howard*, 1993]. The mean direction of the sediment transport will be strongly biased toward that of the albedo-induced flow because maximum winds stresses will occur when all components are aligned. The difficulty of matching predicted wind stress with motion of materials is a familiar problem for Mars [*Greeley et al.*, 1992, 1993; *Edgett and Blumberg*, 1994]: while many of the patterns of wind flow and eolian features match aspects of circulation models, the magnitudes of the calculated surface stresses are often somewhat less than those necessary to initiate motion. There is no doubt that motion occurs, and is occurring, so modeled maximum wind stresses should provide indications of the locations and orientations of the maximum wind transport.

Summary and Discussion

The principal conclusions are as follows:

1. The largest accumulations of dunes in the north polar region are confined by on- and off-pole winds.

2. The albedo contrast of the dunes, especially with icy areas, can generate significant winds into the areas of dunes, thereby providing confining wind components.

3. Although south polar dunes face more topographic obstacles to dispersal from the PLD, the rougher topography prevents areal coverage by low albedo materials sufficient to generate seabreeze type winds that might help confine the dune materials.

Self-regulated dispersal of sand from the poles is only a small part of the evolution of the polar deposits, as the sand is volumetrically small compared to the dust/ice [*Thomas and Weitz*, 1989; *Herkenhoff and Murray*, 1990; *Lancaster and Greeley*, 1990]. However, the observations of the dunes emphasize that the layered deposits are undergoing a long-term change and that current conditions are unlikely to be applicable to their formation, though they could theoretically apply to conditions of some of the unconformities or particular events recorded in the deposits [see *Howard et al.*, 1982].

There are many other differences between north and south that we have not dealt with here, such as the differences in absolute elevations, polar ices, extent of residual caps, and erosional morphology in deposits associated with the layers such as the pitted terrain in the south.

A better understanding of the mechanisms of the polar winds and their relation to overall deposit evolution could come with proposed Mars Global Surveyor orbital instruments. Many of the problems with modeling wind arise from lack of knowledge of topography and details of thermal inertias, both of which could be vastly improved with suggested instruments. Interpretations of geologically effective winds on all size scales and timescales would profit from more higher-resolution images.

Acknowledgments. This work was supported in part by NASA grants NAGW-3274 and NAGW-1997. Technical assistance was provided by B. Carcich, B. Fox, R. Kline, C. Mann, M. Roth, and B. Boettcher. Helpful comments were provided by T. Parish and an anonymous reviewer.

References

Arvidson, R. E., E. A. Guinness, and S. Lee, Differential aeolian redistribution rates on Mars, *Nature, 278*, 533-535, 1979.

Banta, R. M., L. D. Oliver, and D. H. Levinson, Evolution of the Monterey Bay sea-breeze layer as observed by Doppler lidar, *J. Atmos. Sci., 50*, 3959-3982, 1993.

Blumsack, S. L., P. J. Gierasch, and W. R. Wessel, An analytical and numerical study of the Martian planetary boundary layer over slopes, *J. Atmos. Sci., 30*, 66-82, 1973.

Breed, C. S., M. J. Grolier, and J. F. McCauley, Morphology and distribution of common "sand" dunes on Mars: Comparison with the Earth, *J. Geophys. Res., 84*, 8183-8204, 1979.

Christensen, P. R., Global albedo variations on Mars: Implications for active aeolian transport, deposition, and erosion, *J. Geophys. Res., 93*, 7611-7624, 1988.

Dalu, G. A., and R. A. Pielke, An analytical study of the sea breeze, *J. Atmos. Sci., 46*, 1815-1825, 1989.

Edgett, K. S., and P. R. Christensen, The particle size of Martian aeolian dunes, *J. Geophys. Res., 96*, 22,765-22,776, 1991.

Edgett, K. S., and P. R. Christensen, Mars aeolian sand: Regional variations among dark-hued crater floor features, *J. Geophys. Res., 99*, 1997-2018, 1994.

Edgett, K. S., and D. G. Blumberg, Star and linear dunes on Mars, *Icarus*, in press, 1994.

French, R. G., and P. J. Gierasch, The Martian polar vortex: Theory of seasonal variation and observations of eolian features, *J. Geophys. Res., 84*, 4634-4642, 1979.

Goody, R. M., and M. J. S. Belton, Radiative relaxation times for Mars, *Planet. Space Sci., 15*, 247-256, 1967.

Greeley, R., N. Lancaster, S. Lee, and P. Thomas, Martian aeolian processes, sediments, and features, in *Mars*, edited by H. Kieffer et al., pp. 730-766, University of Arizona Press, Tucson, 1992.

Greeley, R., A. Skypeck, and J. B. Pollack, Martian aeolian features and deposits: Comparisons with general circulation model results, *J. Geophys. Res., 98*, 3183-3196, 1993.

Haberle, R. M., C. B. Leovy, and J. B. Pollack, A numerical model of the Martian polar cap winds, *Icarus, 39*, 151-183, 1979.

Herkenhoff, K. E., and B. C. Murray, Color and albedo of the south polar layered deposits on Mars, *J. Geophys. Res., 95*, 1343-1358, 1990.

Howard, A. D., Effect of wind on scarp evolution on the Martian poles (abstract), in Reports of Planetary Geology Program—1980, *NASA Tech. Memo., TM-82385*, 333-335 1980.

Howard, A. D., J. A. Cutts, and K. R. Blasius, Stratigraphic relationships within Martian polar cap deposits, *Icarus, 50*, 161-215, 1982.

Keegan, K. D., J. E. Bachman, and D. A. Paige, Thermal and albedo mapping of the north polar region of Mars (abstract), *Lunar Planet. Sci. Conf., XXII*, 701, 1991.

Kieffer, H. H. and A. P. Zent, Quasi-periodic climate change on Mars, in *Mars*, edited by H. H. Kieffer et al., pp. 1180-1218, University of Arizona Press, Tucson, 1992.

Lancaster, N., and R. Greeley, Sediment volume in the north polar sand seas of Mars, *J. Geophys. Res., 95*, 10,921-10,927, 1990.

Murray, B. C., L. A. Soderblom, J. A. Cutts, R. P. Sharp, D. J. Milton, and R. B. Leighton, Geological framework of the south polar region of Mars, *Icarus, 17*, 328-345, 1972.

Paige, D. A., and K. D. Keegan, Thermal and albedo map of the south polar region of Mars (abstract), *Lunar Planet. Sci. Conf., XXII*, 1013, 1991.

Paige, D. A., J. E. Bachman, and K. D. Keegan, Thermal and albedo mapping of the polar regions of Mars using Viking thermal mapper observations, 1, North polar region, *J. Geophys. Res.*, in press, 1994.

Parish, T. R., and A. D. Howard, Numerical simulation of thermally induced near-surface flows over Martian terrain, LPI Tech. Rep. *93-06*, 35-39, 1993.

Plaut, J. J., R. Kahn, E. A. Guinness, and R. E. Arvidson, Accumulation of sedimentary debris in the south polar region of Mars and implications for climate history, *Icarus, 75*, 357-377, 1988.

Pollack, J. B., D. S. Colburn, F. M. Flasar, R. Kahn, C. E. Carlston, and D. Pidek, Properties and effects of dust particles suspended in the Martian atmosphere, *J. Geophys. Res., 84*, 2929-2945, 1979.

Pollack, J. B., R. M. Haberle, J. Schaeffer, and H. Lee, Simulations of the general circulation of the Martian atmosphere, 1, Polar processes, *J. Geophys. Res., 95*, 1447-1473, 1990.

Rotunno, R., On the linear theory of the land and sea breeze, *J. Atmos. Sci., 40*, 1999-2009, 1983.

Squyres, S. W., The evolution of dust deposits in the Martian north polar region, *Icarus, 40*, 244-261, 1979.

Stull, R. B, *Boundary Layer Meteorology*, 666 pp., Kluwer Academic, Norwell, Mass., 1988.

Tanaka, K. L., and D. H. Scott, Geologic map of the polar

regions of Mars, *U.S. Geol. Surv. Misc. Invest. Ser. Map, kI-1802-C*, 1987.

Thomas, P., North-south asymmetry of eolian features in Martian polar regions: Analysis based on crater-related wind markers, *Icarus, 48*, 76-90, 1981.

Thomas, P., Present wind activity on Mars: Relation to large latitudinally zoned sediment deposits, *J. Geophys. Res., 87*, 9999-10,008, 1982.

Thomas, P., J. Veverka, and R. Campos-Marquetti, Frost streaks in the south polar cap of Mars, *J. Geophys. Res., 84*, 4621-4633, 1979.

Thomas, P., and C. Weitz, Sand dunes and polar layered deposits on Mars, *Icarus, 81*, 185-215, 1989.

Thomas, P., J. Veverka, D. Gineris, and L. Wong, "Dust" streaks on Mars, *Icarus, 60*, 161-179, 1984.

Thomas, P., K. Herkenhoff, A. Howard, B. Murray, and S. Squyres, Polar deposits of Mars, in *Mars*, edited by H. Kieffer et al., pp. 767-798, University of Arizona Press, Tucson, 1992.

Tsoar, H., R. Greeley, and A. R. Peterfreund, Mars: The north polar sand sea and related wind patterns, *J. Geophys. Res., 84*, 8167-8180, 1979.

U.S. Geological Survey, Mars digital topographic map, version 2, vol. 7, Global topography, *CD-ROM VO2007*, Flagstaff, Ariz., 1993.

Ward, A. W., and K. B. Doyle, Speculation on Martian north polar wind circulation and resultant orientations of polar sand dunes, *Icarus, 55*, 420-431, 1983.

White, B. R., Soil transport by winds on Mars, *J. Geophys. Res., 84*, 4643-4651, 1979.

Zurek, R. W., J. R. Barnes, R. M. Haberle, J. B. Pollack, J. E. Tillman, and C. B. Leovy, Dynamics of the atmosphere of Mars, in *Mars*, edited by H. H. Kieffer et al., pp. 835-933, University of Arizona Press, Tucson, 1992.

P. J. Gierasch and P. C. Thomas, Center for Radiophysics and Space Research, 422 Space Sciences Building, Cornell University, Ithaca, NY 14853-6801.

(Received March 1, 1994; revised September 28, 1994; accepted October 6, 1994.)

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5407–5432, MARCH 25, 1995

Geology and landscape evolution of the Hellas region of Mars

Kenneth L. Tanaka and Gregory J. Leonard
U.S. Geological Survey, Flagstaff, Arizona

Abstract. Hellas basin on Mars has been the site of volcanism, tectonism, and modification by fluvial, mass-wasting, and eolian processes over its more than 4-b.y. existence. Our detailed geologic mapping and related studies have resulted in the following new interpretations. The asymmetric distribution of highland massifs and other structures that define the uplifted basin rim suggest a formation of the basin by the impact of a low-angle bolide having a trajectory heading S60°E. During the Late Noachian, the basin was infilled, perhaps by lava flows, that were sufficiently thick (>1 km) to produce wrinkle ridges on the fill material and extensional faulting along the west rim of the basin. At about the same time, deposits buried northern Malea Planum, which are interpreted to be pyroclastic flows from Amphitrites and Peneus Paterae on the basis of their degraded morphology, topography, and the application of a previous model for pyroclastic volcanism on Mars. Peneus forms a distinctive caldera structure that indicates eruption of massive volumes of magma, whereas Amphitrites is a less distinct circular feature surrounded by a broad, low, dissected shield that suggests generally smaller volume eruptions. During the Early Hesperian, a ~1- to 2-km-thick sequence of primarily fined-grained, eolian material was deposited on the floor of Hellas basin. Subsequently, the deposit was deeply eroded, except where armored by crater ejecta, and it retreated as much as 200–300 km along its western margin, leaving behind pedestal craters and knobby outliers of the deposit. Local debris flows within the deposit attest to concentrations of groundwater, perhaps in part brought in by outflow floods along the east rim of the basin. These floods may have deposited ~100–200 m of sediment, subduing wrinkle ridges in the eastern part of the basin floor. During the Late Hesperian and Amazonian, eolian mantles were emplaced on the basin rim and floor and surrounding highlands. Their subsequent erosion resulted in pitted and etched plains and crater fill, irregular mesas, and pedestal craters. Local evidence occurs for the possible former presence of ground ice or ice sheets ~100 km across; however, we disagree with a hypothesis that suggests that the entire south rim and much of the floor of Hellas have been glaciated. Orientations of dune fields and yardangs in lower parts of Hellas basin follow directions of the strongest winds predicted by a recently published general circulation model (GCM). Transient frost and dust splotches in the region are, by contrast, related to the GCM prediction for the season in which the images they appear in were taken.

1. Introduction

The Hellas region of Mars consists of diverse topography, geology, and surface characteristics, yet it is dominated by one particular feature, the Hellas impact basin. Although the basin was formed during early bombardment, it is relatively well preserved and makes up the deepest and broadest (~9 km relief, ~2000 km across) depression on Mars (Figures 1 and 2). Volcanism and channel dissection have modified large expanses of the northeastern and southern parts of the basin rim. Both of these modified areas and the floor of the basin are marked by wrinkle ridges. Also, a large deposit of proposed eolian origin covers the basin floor; this area is presently the site of origin of major dust storms [*Martin and Zurek*, 1993] and of the highest wind stresses on Mars predicted by general circulation models (GCM's) [*Greeley et al.*, 1993].

Paper number 94JEO2804.

In spite of the great storehouse of geologic information contained in the Hellas region, comprehensive mapping of the region has been undertaken only recently at 1:5,000,000 scale (G.J. Leonard and K.L. Tanaka, work in progress, 1995). This study focuses on the interpretation of the geologic and eolian history of the Hellas region. We address several questions and issues of regional to global geologic and climatic significance, including (1) What accounts for the asymmetric spatial distribution of basin massifs? (2) What was the nature of volcanism on the south rim of Hellas? (3) How were large and small channels generated on the basin rim? (4) What is the origin of basin-fill materials, and are they the source for present-day dust storms? (5) What was the overall history of deposition and erosion in the region, and is there a relation to planetwide resurfacing history? (6) How do the observed spatial distribution and orientations of interpreted, present-day eolian streaks and landforms (such as dust and frost splotches, dunes, and yardangs) compare with predicted surface wind stresses and directions based on the GCM of *Greeley et al.* [1993]? We address these questions by way of geomorphologic interpretation, stratigraphic relations, topographic measurements, crater counts, and application of

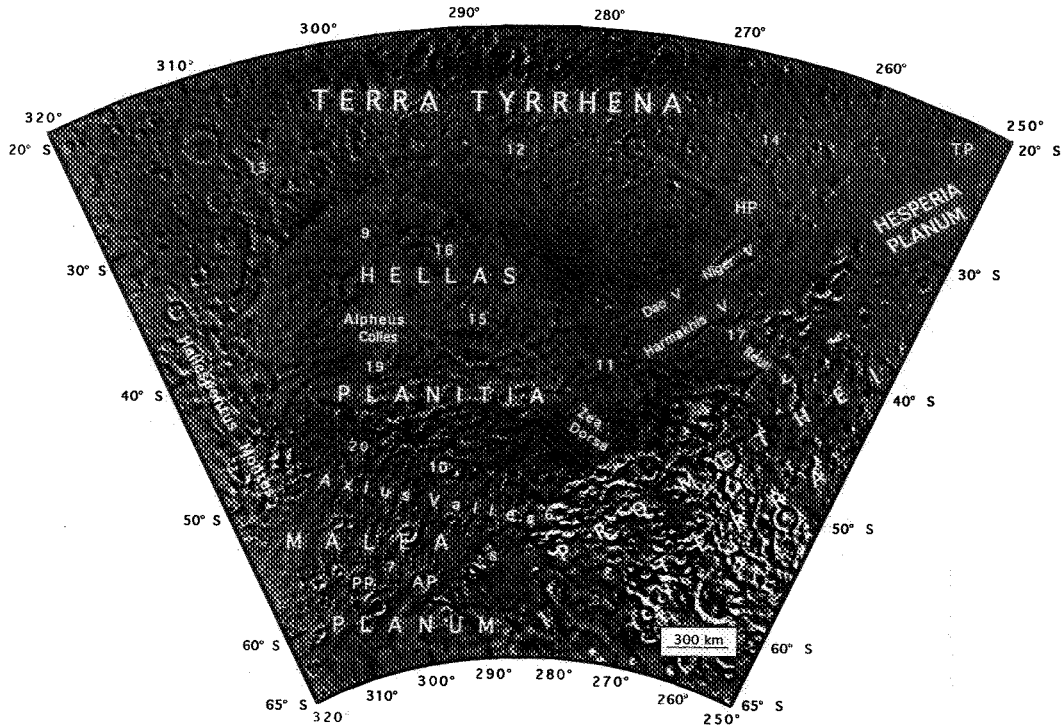


Figure 1. Geographic nomenclature of the Hellas region of Mars. Numbers index image figures in this paper; HP, Hadriaca Patera; TP, Tyrrhena Patera; PP, Peneus Patera; AP, Amphitrites Patera. Photomosaic base derived from Mars Digital Image Mosaic.

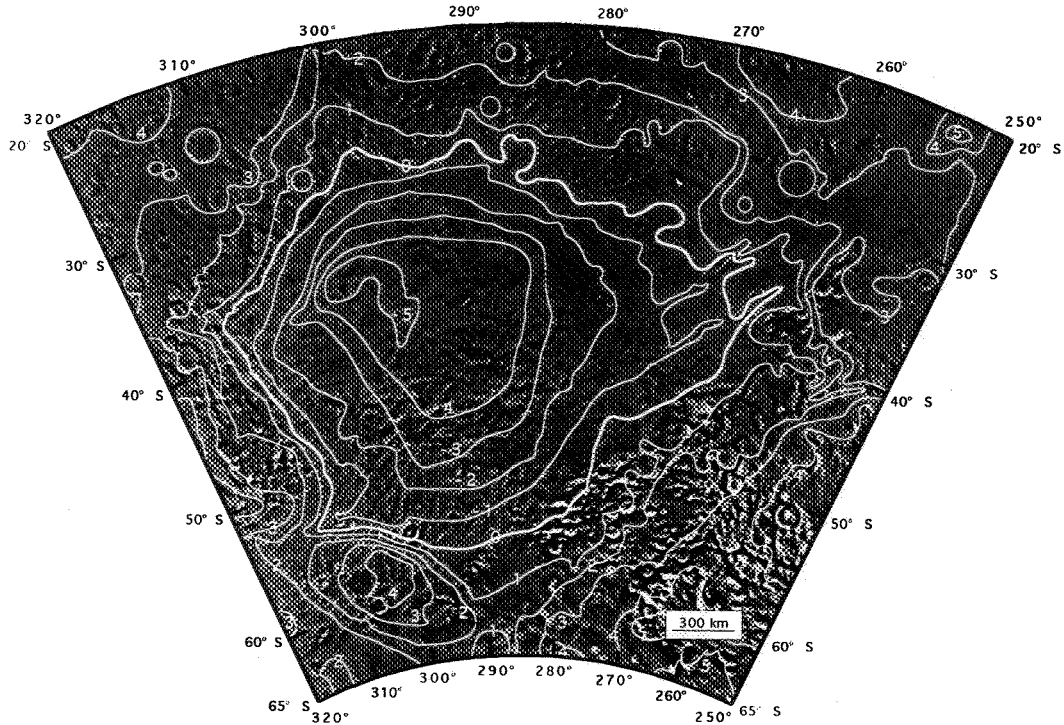


Figure 2. Topographic map of the Hellas region of Mars; contour interval 1000 m [*U.S. Geological Survey*, 1991].

various hypothetical models. Computer enhancement and display of available Viking images have optimized our geologic analyses.

We summarize our findings in the following reconstruction of the geologic and geomorphologic history of the Hellas region. Heavy bombardment of the early crust in the Hellas region resulted in densely cratered terrain during the Early Noachian. A southeast-trending low-angle impact probably formed the Hellas basin cavity and surrounding uplifted rim material. Much of the rim comprises disrupted, large massif blocks of relatively resistant crustal material. Concentric graben (outside the study area) associated with the impact formed soon thereafter. In the Middle Noachian, continued bombardment formed impact craters throughout the study area. Ejecta, fluvial sediment from dissection of basin massifs, and possibly eolian mantles and volcanic deposits began infilling topographic lows, including the Hellas basin, highland areas outside the basin rim, and intermassif and intercrater areas of the basin rim. Such highland modification by local eolian, fluvial, and volcanic activity continued into the Late Noachian. Hellas basin became infilled, probably by lava flows that were apparently sufficiently thick to cause lithospheric sagging leading to wrinkle-ridge formation, normal faulting at Hellespontus Montes, and lava effusion along the edge of the basin interior. Pyroclastic volcanism formed low-relief highland volcanoes on the outer parts of the northeast and south basin rim; large areas of the inner rim slopes were buried by the deposits. A thick sequence of lava flows was emplaced south and west of Amphitrites and Peneus Paterae on Malea Planum. Subsidence under the weight of the volcanic rocks formed wrinkle ridges on the volcanoes and southern lava-flow fields.

Highland-paterae pyroclastic volcanism continued into and culminated during the Early Hesperian, and summit collapse produced circular caldera structures on each volcano. Eruption of massive volumes of magma formed Peneus, a distinctive caldera structure, whereas generally smaller volume eruptions created Amphitrites, a less distinct circular feature surrounded by a broad, low, dissected shield. Runoff and (or) groundwater sapping, accompanied by mass wasting and eolian erosion, produced valleys and troughs on the paterae flanks and neighboring areas of the basin rim; a few large channels also formed (such as Reull Vallis). Lava flows were emplaced in Hesperia Planum and deformed by wrinkle ridges. Most of the Hellas basin floor was covered by a thick sequence of deposits; morphologic and stratigraphic relations and volume estimates indicate the deposits are predominantly eolian in origin, with some contributions of fluvial and volcanic material. The source of the eolian material likely was sand, silt, and dust eroded from highland terrains (particularly along the highland/lowland boundary, which was deeply eroded at this time). Local highland erosion and subsequent mantling probably occurred. During the Late Hesperian, the basin interior deposit was deeply eroded into hummocky and knobby terrain (except where armored by the ejecta of large craters), and its western margin retreated as much as 200–300 km, leaving behind pedestal craters, infilled craters, and knobby outliers. Late-stage effusive volcanism at Tyrrhena Patera produced a large rille structure and a field of lobate lava flows that apparently triggered collapse and outflow erosion, which produced Dao, Niger, and Harmakhis Valles. These channels debouched into Hellas Planitia, and their sediments likely buried much of the interridge plains of the basin ridged unit. Mantling and subsequent partial erosion of the basin rim and adjoining highlands resulted in (1) intracrater and intercrater etched plains material; (2) irregular mesas, pedestal craters, and channel fill on dissected and ridged plains material on the upper part of the basin rim; and (3) pitted plains and pedestal craters superposed on degraded and subdued valleys on northern Malea Planum.

Mantling and mantle erosion continued through the Early and Middle Amazonian on the basin rim and interior. Interior mantles formed along the south edge of the retreated interior deposit and in more deeply eroded, central areas of the interior deposit. Mass-wasting of the interior deposit resulted in debris aprons and flows. On the basin rim, debris aprons formed along the bases of massifs and channel flanks. Late Amazonian activity included local mantling, mass-wasting, and erosion at low intensity throughout the Hellas region. In lower parts of the Hellas basin interior deposit and basin ridged unit, oriented ridges are interpreted as dunes and yardangs formed by strong, seasonal winds (among the strongest presently occurring on Mars). Presently, transient, seasonal frost and dust splotches form throughout the Hellas region and are commonly oriented along co-seasonal trends of GCM predictions.

Overall, our conclusions for inner basin fill and modification are in conflict with *Kargel and Strom* [1992], whose glacial interpretation rests on contemporaneous formation of channels on Malea Planum and infill and modification of deposits in Hellas basin. We find, instead, distinct timing of formation of those features and alternative explanations consistent with our observations and GCM predictions.

2. Stratigraphy

The stratigraphy of the Hellas region is defined by geologic mapping; we draw from our preliminary draft of a geologic map at 1:5,000,000 scale based on Viking images (G.J. Leonard and K.L. Tanaka, work in progress, 1995). Previous regional geologic mapping of note includes the Mars quadrangle series at 1:5,000,000 scale, based on Mariner 9 images [e.g., *Potter,* 1976; *De Hon,* 1977; *Peterson,* 1977], and the global Viking series at 1:15,000,000 scale [*Greeley and Guest,* 1987; *Tanaka and Scott,* 1987]. Also, detailed geologic maps at 1:500,000 scale from high-resolution images have been produced for the Hadriaca and Tyrrhena Paterae areas [*Greeley and Crown,* 1990; *Crown et al.,* 1992]. Our mapping differs from the earlier work in that we used the Viking versus the Mariner data set, and from the later work in that we examine Hellas from a regional perspective versus global and local context.

Stratigraphic assignments are derived from observed overlap and embayment contact relations recorded by geologic mapping (Figure 3) and from crater counts (Tables 1–4). Because of the apparent significant obliteration of smaller craters over much of geologic time (as discussed later), emplacement ages are assigned more confidently from the densities of larger craters (generally >5 km in diameter).

The generalized results of our mapping are shown in Figure 3, which highlights major, extensive units, and in Figure 4, which shows only units interpreted to be of eolian origin. These map units are briefly discussed in this section; detailed discussion of the formational and modification histories of some of the units is presented in following sections.

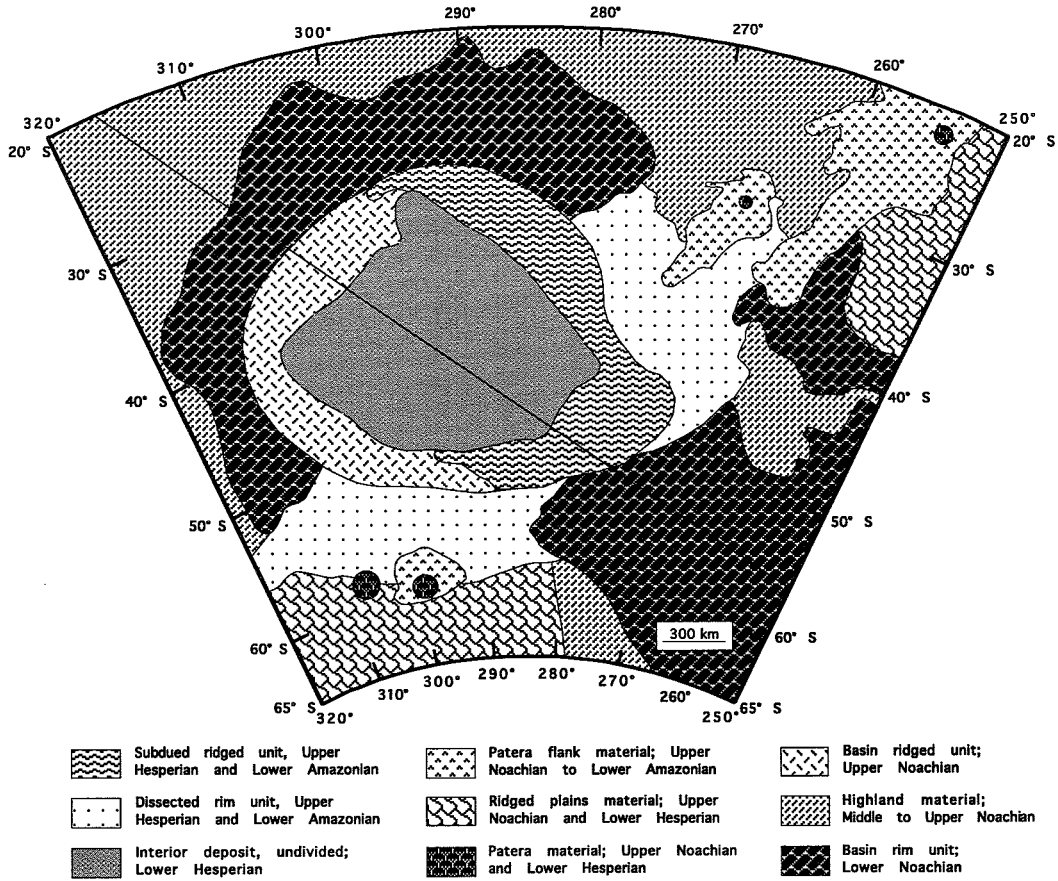


Figure 3. Generalized geologic map of the Hellas region of Mars (from G.J. Leonard and K.L. Tanaka, work in progress, 1995). Line indicates proposed low-angle impact trajectory for the Hellas basin-forming bolide. Note symmetry about trajectory of (1) distribution of basin rim unit and (2) paterae and paterae flank materials.

Table 1. Cumulative Crater Densities and Inferred Surface Ages of Geologic Units in the Hellas Region of Mars

Geologic Unit* or Feature	Crater Density†		Area, km²	Surface Age‡	
	N(2)	N(5)		N(2)	N(5)
Smooth material and interior mantles	275±26	54±12	407,368	EA	EA
Subdued ridged unit	362±35	163±23	301,139	EA	EH
Valles material§	239±72	109±49	45,947	EA	LH/EA
Flank flows of Tyrrhena Patera§	397±70	87±33	80,576	LH/EA	LH/EA
Dissected rim unit, Malea Planum	347±26	133±16	512,364	EA	EH/LH
Dissected rim unit, east rim§	460±49	132±27	188,976	LH	EH/LH
Interior deposit	345±36	139±23	266,498	EA	EH/LH
Degraded interior deposit	250±22	83±13	519,018	EA	LH
Ridged plains material, Malea Planum	384±30	159±19	440,090	LH/EA	EH
Hadriaca Patera§	572±75	175±41	103,089	LH	LN/EH
Hadriaca Patera‖	816±100	223±52	80,895	EH/LH	LN/EH
Tyrrhena Patera‖	666±161	313±111	25,544	EH/LH	LN
Basin ridged unit	406±37	186±25	295,244	LH/EA	LN/EH

*Most units shown in Figures 3 and 4; some from *Crown et al.* [1992] and G.J. Leonard and K.L. Tanaka (work in progress, 1995).

†Crater density N(D) equals number of craters larger than diameter D (in kilometers) per million square kilometers.

‡Relative ages based on *Tanaka* [1986] and reflect both material and modification ages (see text); N, Noachian; H, Hesperian; A, Amazonian; E, Early; M, Middle; L, Late; for material units, equivalent stratigraphic positions of epochs are series, where "Lower" position corresponds to "Early" age and "Upper" to "Late."

§Data from *Crown et al.* [1992].

‖Data derived from Table 1 and Figure 7 of *Plescia and Saunders* [1979].

The oldest (Early Noachian) material in the region consists of high-standing crust of the Hellas rim, known as the basin rim unit. This unit is marked by a high density of craters tens of kilometers in diameter as well as numerous rugged massifs and mountain ranges. Embaying the basin rim unit and surrounding Hellas basin are Middle to Upper Noachian highland material. Locally, this unit is dissected by small valleys. Within Hellas basin is an annulus of Upper Noachian basin ridged unit. On the northeast rim of Hellas, Hadriaca and Tyrrhena Paterae are made up mostly of Upper Noachian patera and patera flank materials. Ridged plains material was emplaced south of Hellas on Malea Planum beginning at the end of the Noachian and on Hesperia Planum at the beginning of the Hesperian.

Volcanism in the Hesperian produced shield and dissected deposits on the south rim of Hellas; these deposits form patera material of the Amphitrites Formation [*Tanaka and Scott*, 1987]. Within the Hellas basin, a voluminous eolian interior deposit was emplaced during the Early Hesperian; where deeply eroded, this unit is mapped as degraded interior deposit (Figure 4). Upper Hesperian and Lower Amazonian rim and interior mantles and smooth material overlie and embay, in places, older material on the Hellas floor and rim (Figure 4). In the northeastern and southern parts of the basin and basin rim, Late Hesperian to Early Amazonian channel dissection formed the dissected rim unit, which includes small valleys as well as a few large, deep channels that emanate from large depressions. Locally, this unit may have been produced by volcano and ground-ice interaction, because late-stage lava flows from Tyrrhena Patera are coeval with and adjacent to source areas of some of the channels. Emplacement of the subdued ridged unit (Figure 3), a plains material that covers and

Table 2. Binned Crater-Density Boundaries for Martian Epochs

Series	Crater Density*						
	N(2–2.828)	N(2.828–4)	N(4–5.657)	N(5.657–8)	N(8–11.32)	N(11.32–16)	N(16)
Late Amazonian	<20	<10	<5				
Middle Amazonian	20–75	10–38	5–19	2.5–9			
Early Amazonian	75–200	38–100	19–50	9–25	5–13		
Late Hesperian	200–375	100–188	50–94	25–47	13–23	6–12	
Early Hesperian	375–600	188–300	94–150	47–75	23–38	12–19	12–19
Noachian	>600	>300	>150	>75	>38	>19	>19

*See Table 1; data determined for specified diameter bins.

Table 3. Binned Crater Densities for Hellas Basin Map Units

Unit	Crater Density*						
	N(2–2.828)	N(2.828–4)	N(4–5.657)	N(5.657–8)	N(8–11.32)	N(11.32–16)	N(16)
Smooth material and interior mantles	125±18	59±12	39±10	25±8	12±5 (N(8))	—	—
Subdued ridged unit	73±16	76±16	66±15	63±14	33±11	17±7	30±8
Interior deposit	79±17	79±17	71±16	30±11	26±10	26±10	34±11
Degraded interior deposit	50±10	83±13	44±9	40±9	21±6	—	—
Dissected rim unit, Malea Planum	64±11	105±14	61±11	47±10	25±7	18±6	27±7
Ridged plains material, Malea Planum	100±15	86±14	52±11	41±10	41±10	30±8	34±9
Basin ridged unit	74±16	85±17	71±16	41±12	41±12	24±9	54±14

*See Table 1; data determined for specified diameter bins.

embays ridges of the eastern outcrops of the Hellas basin ridged unit, is also possibly related to this dissection. Patches of dunes and transient splotches (Figure 4) appear to be relatively recent, and probably active, deposits within the central and northwestern parts of Hellas basin and on various parts of the rim.

3. The Hellas Impact Structure

The topographic expression and distribution of structures around the Hellas region compose the signature of the impact event that produced the basin. These and other related features provide clues for understanding the geometry of the impact, which in turn provides a basis for better analysis of the styles and degrees of geologic and degradational processes that have occurred in the Hellas region. The topographic rim of Hellas is marked in part by many rugged massifs of the basin rim unit that were uplifted as a result of the Hellas impact. Shadow and photoclinometric measurements of isolated massifs yield heights between 1 and 5 km (Figure 5). Also, differential uplifts of the basin rim are bounded by escarpments, particularly along the southeast basin rim, where an extensive region of uplift occurs. Here a lobe of the basin rim unit extends for about 1500 km [*Greeley and Guest,* 1987]. By contrast, the northwest rim has only a narrow, discontinuous band of rim material, which averages about 300 km in width. One explanation for the broad, southeastern region of uplift is that it had been the site of an earlier impact basin and thus was more susceptible to tectonic deformation [*Wichman and Schultz,* 1989]. Recently, *Leonard and Tanaka* [1993] proposed that this region resulted from a southeast-trending (S60°E), low-angle impact of the basin-forming bolide (see Figure 3).

Laboratory studies and observations of impact craters on planetary bodies indicate that oblique impacting projectiles produce features that are symmetric about (normal to) their trajectory path and asymmetric along their range path [*Gault and Wedekind,* 1978; *Moore,* 1969; *Schultz,* 1994]. The Hellas impact structure displays both types of features. Features symmetric across the proposed trajectory path include (1) broad concentric troughs at distances 1800 to 2500 km north and west of the basin (outside of study area), which formed at about the time of the impact; these features have been modeled as resulting from flow of the asthenosphere in

Table 4. Binned Crater-Density Ages for Hellas Basin Map Units

Unit	Crater Density						
	N(2–2.828)	N(2.828–4)	N(4–5.657)	N(5.657–8)	N(8–11.32)	N(11.32–16)	N(16)
Smooth material and interior mantles	EA	EA	EA	LH/EA	LH/EA	—	—
Subdued ridged unit	EA/MA	EA	LH	LN/EH	LN–LH	LN–LH	LN
Interior deposit	EA/MA	EA	LH	LH/EA	EH/LH	LN/EH	LN
Degraded interior deposit	MA	EA	LH/EA	EH/LH	EH/LH	—	—
Dissected rim unit, Malea Planum	MA	LH/EA	LH	EH/LH	EH/LH	LN/EH	LN
Ridged plains material, Malea Planum	EA	EA	LH/EA	EH/LH	LN/EH	LN	LN
Basin ridged unit	EA/MA	LH/EA	LH	EH/LH	LN/EH	LN/EH	LN

Abbreviations defined in Table 1.

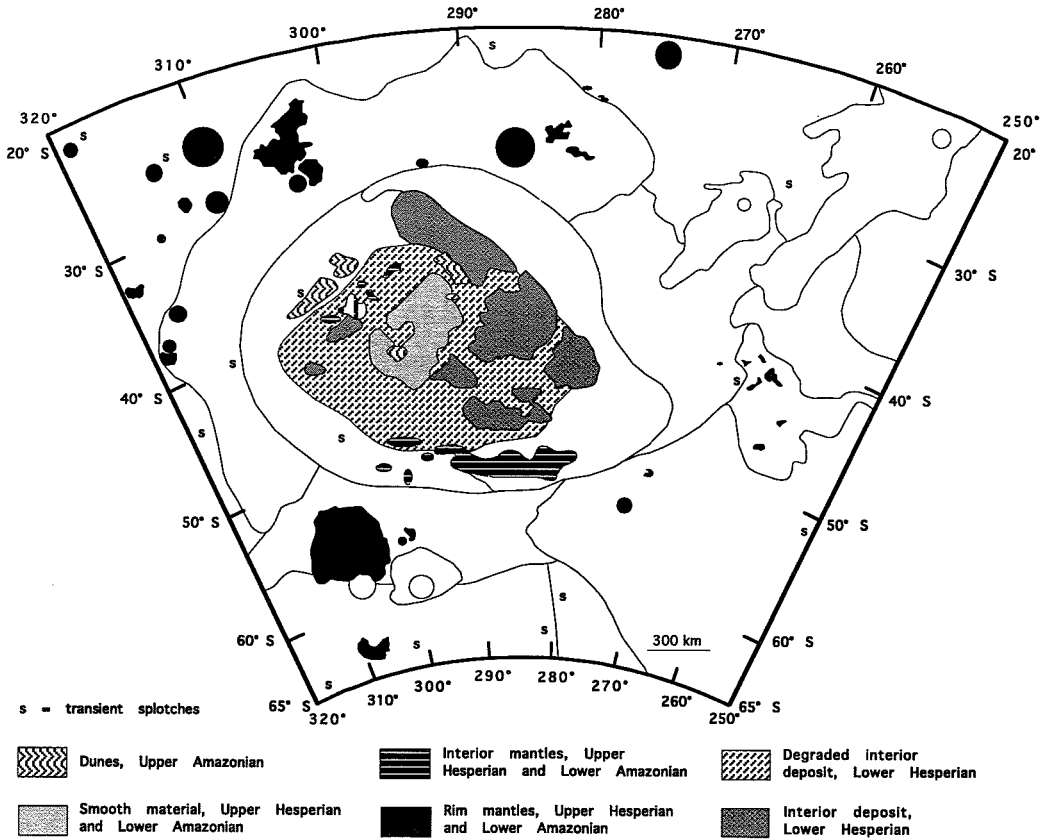


Figure 4. Map of eolian deposits and features in the Hellas region of Mars. The approximate boundary shown for the rim mantle north of Peneus Patera is uncertain due to the lack of high-resolution images necessary to define the mantle's extent, which is likely to be much more than what is shown. Other units discussed in text. Lines show contacts of general geologic units in Figure 3.

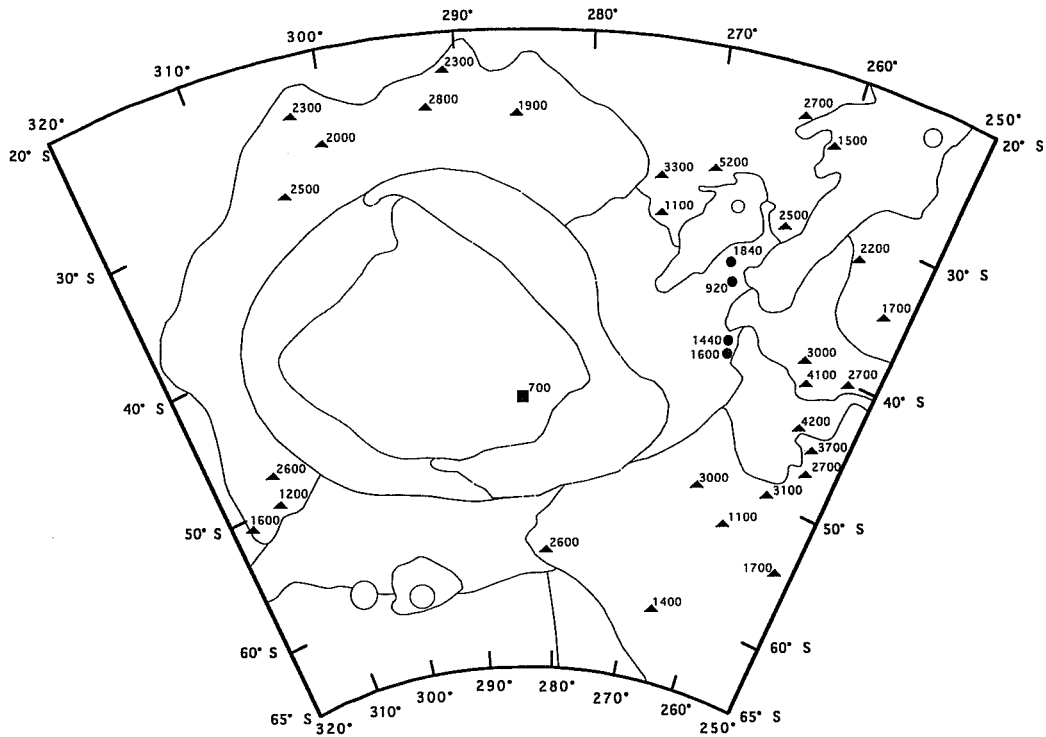


Figure 5. Shadow-height and photoclinometric measurements (in meters) in the Hellas region. Triangles indicate heights of basin-rim massifs; circles show depths of channels; square shows scarp height of basin interior deposit. Some massif heights represent averages of closely spaced massifs that had values that varied by less than 400 m. Measurement precisions vary from <100 m to a few hundred meters depending on pixel size, lighting geometry, and definition of shadow boundary.

response to the transient impact cavity [*Wichman and Schultz,* 1989]; (2) prominent scarps that bound the southeastern region of uplift; areas in front (west) of the scarps appear to be structurally lower (Figure 6); (3) large areas between the narrow and broad parts of the rim (see Figure 3) that show a marked absence of massifs (Figure 5); and (4) large volcanic paterae surrounded by vast volcanic plains that occur just outside the basin rim, also between the narrow and broad parts of the rim (Figure 3). Asymmetric features along the range path include (1) higher (generally by ~1–2 km) and more densely concentrated massifs in the large southeast rim lobe (inferred downrange direction) than in the narrow northwest rim margin (see Figure 5); (2) deeper excavation and steeper walls on the uprange profile (Figure 2); and (3) an uprange zone of subarcuate fractures and rim collapse, particularly the extensional faulting that formed the scarps associated with Hellespontus Montes along the west rim of Hellas.

Although the patterns of ejecta and crater morphology produced by oblique impacts have been simulated by laboratory experiments of projectiles impacting fine-grained materials, deformation of the lithosphere by oblique impacts at the scale of multiring basins has not been simulated by such experiments or predicted by theory. However, on the Moon, where basin ejecta is commonly better preserved than on Mars [*Edgett,* 1989], the asymmetry of ejecta deposits for a dozen basins suggests oblique impact; three of the basins are elongate and two have elevated sectors [*Wilhelms,* 1987, table 4.4]. Also on the Moon, basin-rim and ring-height variations have been interpreted to be caused by pre-impact surface morphology, oblique impact, and post-impact modification [*Whitford-Stark,* 1981]. Unfortunately, because of the differences in preservation of lunar versus Martian impact basins, a clear definition of basin-rim structure for oblique impacts cannot be demonstrated empirically.

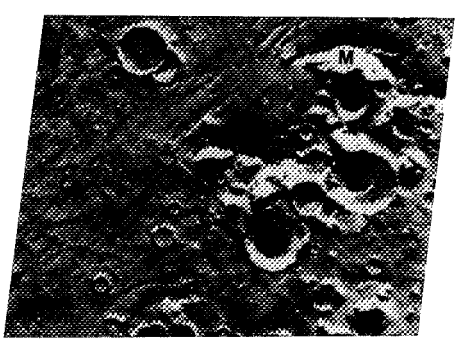


Figure 6. Part of southeast rim of Hellas, showing prominent, cratered massifs of the basin rim unit embayed by dissected rim material. Tallest massif (M) is 2.6 km high (Figure 5). Sparsity and decrease in exposed size of massifs to the west indicatesrelative structural lowering of basin rim. Note Mad Vallis, a fretted channel (F) along edge of basin rim unit. (Viking image 365S03; centered at 56.14°S, 282.67°, 276 m/pixel).

4. Volcanism

Two and possibly three significant volcanic areas occur in the Hellas region as (1) paterae and associated plains on the northeast rim of Hellas basin, (2) paterae and associated plains on the south Hellas rim, and (3) an early floor of Hellas basin [*Potter,* 1976]. The northeast rim volcanics and channels have been described and discussed extensively [*Greeley and Crown,* 1990; *Crown et al.,* 1992; *Crown and Greeley,* 1993]; we will only briefly summarize those results here. The southern volcanics, which are also extensively channeled, have not been analyzed previously, and the possibility that the basin floor is filled by lavas is examined here.

4.1 Northeast Rim Volcanics

This volcanic area includes Hesperia Planum and the broad, low shields of Hadriaca and Tyrrhena Paterae. Crater counts indicate that Tyrrhena Patera formed during the Late Noachian, and Hadriaca during the Late Noachian and Early Hesperian (Table 1). Hesperia Planum ridged plains, by definition, are Lower Hesperian [*Scott and Carr,* 1978; *Tanaka,* 1986] and are thought to be composed of broad, thin, low-viscosity lava flows [*Greeley and Spudis,* 1981]. The morphologies and flank slopes and lengths of flows from the paterae indicate that they consist mainly of early pyroclastic flows followed by late-stage lava flows and modest caldera collapse. This volcanic evolution may be consistent with early volatile-rich magmas followed by volatile-poor magmatism; alternatively, the pyroclastic deposits could have resulted from hydromagmatic eruptions through the near-surface interaction of magma and ground water in a permeable crust [*Greeley and Spudis,* 1981; *Greeley and Crown,* 1990; *Crown and Greeley,* 1993]. Late-stage effusive activity (Late Hesperian/Early Amazonian) at Tyrrhena Patera produced a field of lava flows that cover the plains from Tyrrhena to near the east flank of Hadriaca.

4.2 South Rim Volcanics

Peneus and Amphitrites Paterae are at the crest of the south rim of Hellas and are surrounded by the ridged plains of Malea Planum (Figure 1). North of the paterae, the ridged plains are dissected by narrow channels of Axius Valles. The entire volcanic area covers, in the study area alone, about 10^6 km^2; more ridged plains extend south of the study area [see *Tanaka and Scott,* 1987]. The paterae reach an elevation of about 4 km, and the surrounding plateau slopes away down to 1 km southeast of Amphitrites Patera and to -2 km at the edge of Hellas Planitia (Figure 2).

Peneus Patera (Figures 7a, 7b) is 100 x 120 km across and elongate in a northeastward direction. The patera is made up of a 10- to 30-km-wide zone of concentric normal faults and grabens that encircle a central floor, which is probably hundreds of meters below the surrounding plateau. The southeastern part of the floor is deformed by wrinkle ridges. The high elevation, ring faults, and central depression of the patera within ridged plains material suggest that Peneus is a major volcanic caldera formed by multiple, voluminous eruptions of magma and subsequent collapse of evacuated magma chambers [*Tanaka and Scott,* 1987].

Amphitrites Patera (Figure 7a), centered 250 km east of the center of Peneus, has a significantly different morphology from its neighbor. Unlike Peneus, Amphitrites has only a few

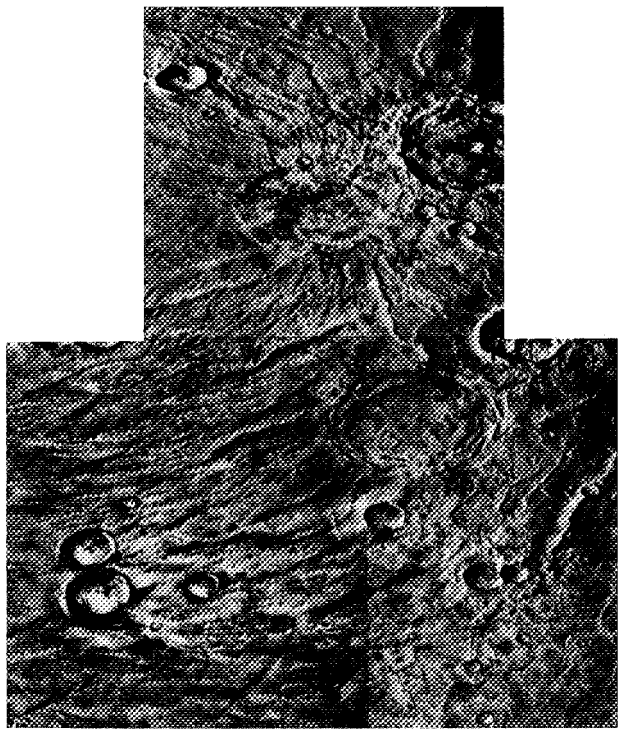


Figure 7a. Oblique view of Peneus (PP) and Amphitrites (AP) Paterae. Peneus Patera, 100–120 km across, is slightly elongate along a northeasterly trend and bounded by a zone of concentric normal faults and grabens. The floor is marked by large wrinkle ridges; one (arrow) may account for prominent bench (B) on southeast side of patera. Amphitrites Patera, 120 km across, forms a ridge and scarp structure surrounded by a 300-km-diameter low shield dissected by narrow channels; it is bordered to the south by crater Barnard. Note possible ring graben (G) and collapse pit (P). A few wrinkle ridges (W) radiate from Amphitrites Patera, thus diverting channels. South and west of Peneus Patera, local sets of similarly trending and spaced wrinkle ridges crosscut one another (mosaic of Viking images 94A74–76; centered at ~57.00°S, 305.00°, 240–250 m/pixel).

ring faults; within the summit region, these faults locally bound a shallow depression. Furthermore, the patera has traditionally been defined by a 120-km-diameter circular ridge and scarp (which is partly buried by the rim of crater Barnard) [e.g., *Tanaka and Scott,* 1987]; however, our mapping has established the outline of a low-relief shield nearly 300 km across that surrounds the circular structure. Channels radiate from the circular structure and terminate at the shield's base. Locally, the channels are highly deformed, degraded, or masked due to wrinkle ridges, mantling deposits, and ejecta from crater Barnard. On the west flank of the shield, a 20-km-long pit may be a collapsed vent structure (Figure 7a). The broad form, low relief, and heavily dissected and degraded morphology of the Amphitrites shield suggest that it is mainly made up of pyroclastic rocks. In contrast to Peneus, eruptions of Amphitrites were probably of much lower volume, leading to shield buildup, local dissection, and little collapse.

The basin-rim slopes of Malea Planum, north of the paterae, are extensively dissected by Axius Valles, a collection of narrow, sinuous channels having very little tributary development (Figure 8). Discontinuous lobate tongues, possibly degraded lava, lahars, or other volcaniclastic deposits, are also common. Because of the highly degraded appearance of the slopes, *Tanaka and Scott* [1987] interpreted the material there to be pyroclastic rocks, similar to those of Hadriaca and Tyrrhena Paterae. Average slopes from Peneus and Amphitrites Paterae across the dissected ridged plains range from about 0.2–0.9° over distances of 400–500 km (the slope values are subject to a high degree of error). These

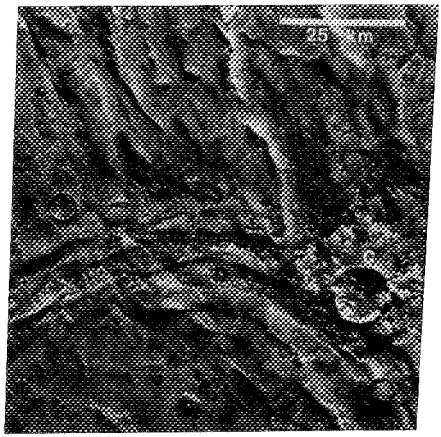


Figure 7b. Northeastern part of Peneus Patera showing modification of ring structures. Flatter areas in patera rim and surrounding ridged plains covered by pitted mantle material, particularly evident on impact crater (C). Narrow ridges, scarps, crater-rim crests, and inner patera floor appear unmantled; however, pedestal craters on inner patera floor indicate former presence of mantle material. At least two mantles present: (1) a lower, widespread mantle marked by small (mostly subkilometer), locally chained pits, and (2) an upper, dark mantle marked by larger (~kilometer size) pits preserved along the patera rim. (Viking image 578B01; centered at 57.13°S, 305.62°, 71 m/pixel).

values for slope and distance, if accurate, are similar to those of Hadriaca and Tyrrhena Paterae and generally within the constraints for runout distance estimated for Martian pyroclastic flows [*Greeley and Crown,* 1990; *Crown and Greeley,* 1993]; however, the distances involved exceed the 250–350 km limit that has been suggested for air-fall deposits on Mars [*Mouginis-Mark et al.,* 1988]. Such pyroclastic eruptions could have been induced by a high volatile content of erupted magmas or by the interaction of magma and ground water.

4.3 Basin Ridged Unit

In Late Noachian time, the basin was partly filled, resulting in smooth plains material marked by wrinkle ridges that make up the basin ridged unit. The unit may have once been largely covered by the interior deposit, which appears to have retreated in the northwestern and southwestern parts of the basin (Figures 9a and 10). Along the south and east edges of the basin, the basin ridged unit is partly buried by younger surficial deposits (e.g., Figures 9–11); locally, only the upper parts of wrinkle ridges indicate the presence of the unit.

We suggest that the basin ridged unit is made up of a succession of relatively smooth, broad lava flows. Although other origins are possible, we favor the lava-flow origin for the following reasons: (1) the weight of the unit may have compressed relatively strong rock (the lava flows) overlying weaker material, thus forming wrinkle ridges [*Watters,* 1993]; (2) the deposit occurs at the lowest elevations on Mars, where magma buoyancy considerations favor effusive volcanism [e.g., *Wilhelms and Baldwin,* 1989]; (3) a few flows appear to emanate from an impact crater along the basin rim (at lat 31° S, long 300.5°), indicating that local effusive eruptions have occurred within the basin; and (4) the unit nowhere appears highly eroded, which would be expected if the unit were made up of a more friable material (as discussed later, deep eolian etching has occurred in the overlying interior deposit). Thickness of the deposit is uncertain, but was probably sufficient to produce the wrinkle ridges.

5. Formation of Wrinkle Ridges

The most common structures at map scale in the Hellas region are wrinkle ridges. Wrinkle ridges are generally interpreted to be compressional fold and (or) thrust-fault structures [e.g., *Golombek et al.,* 1991; *Watters,* 1993]. As

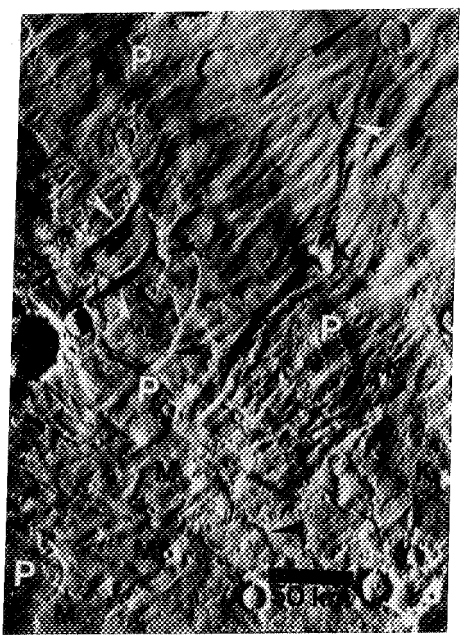


Figure 8. Part of northern Malea Planum, showing complex surface that includes valleys of Axius Valles (white arrows), prominent knobs (K), pedestal craters (P), wrinkle ridges (black arrows), and high-standing mesas (M). The surface evolution appears to be as follows: embayment of basin-rim knobs and large impact craters by pyroclastic deposits and lava flows from Amphitrites Patera area (southwest of image), valley incision, wrinkle-ridge formation, mantle deposition, and mantle erosion (producing pedestal craters and mesas). Some wrinkle ridges here are concentric to Amphitrites Patera, indicating some influence of the patera on the local stress field. (Viking image 95A63; centered at 57.04°S, 288.76°, 241 m/pixel).

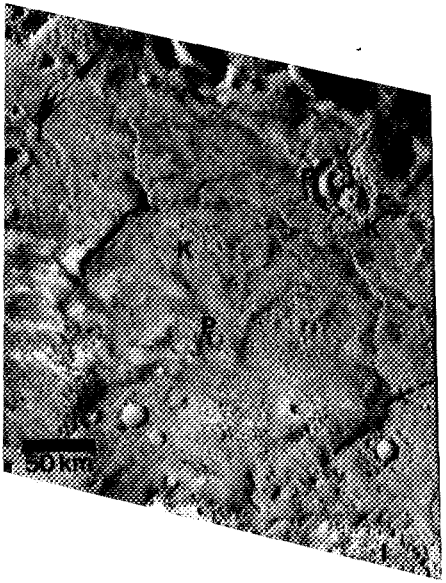


Figure 9a. Northwestern Hellas Planitia, showing basin-ridged unit embaying basin-rim unit (along top of image). Interior deposit superposed on basin ridged unit has largely been removed; traces include crater pedestals (P), some of which are knobby, and knobby outliers (K) below distinctive scarp of interior deposit (I). Connected ridges in lower half of image may be networks of compound crescentic dunes (see terrestrial example in Figure 9b). (Viking image 620A47; centered at 33.82°S, 303.18°, 247 m/pixel).

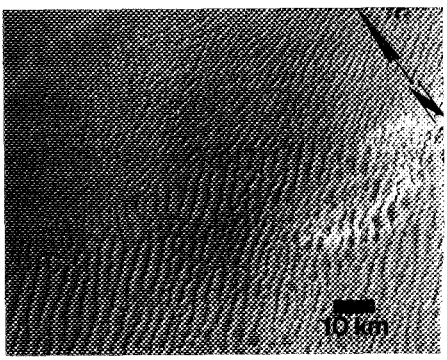


Figure 9b. Top center of image shows network of compound crescentic (transverse) dunes in the Takla Makan Sand Sea, China. Dune sizes and pattern similar to those of ridges in Hellas Planitia (Figure 9a). Dominant wind direction is transverse to longer ridge axes. (Part of Landsat image 1129-04311-7).

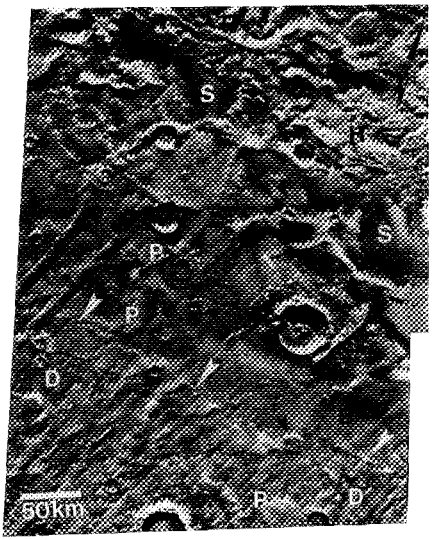


Figure 10. Southern Hellas Planitia, showing rugged, hummocky (H; deeply eroded) areas of the interior deposits and basin ridged material (B); smooth material (S) buries the contact between these large units. A scarp (large white arrows) outlines the edge of the dissected rim unit (D). Although wrinkle ridges do not appear to deform the interior deposit, some scarps in the deposit may be controlled by underlying ridges. Small, pristine channels (small white arrows) are deflected around wrinkle ridges and cut dissected rim unit and interior deposit but are buried by smooth material. Mantling on basin ridged material has partly buried crater rims (black arrows) and resulted in pedestal craters (P). (Viking image 363S04 and 06; centered at ~ 51.50°S, 295.50°, ~245 m/pixel).

discussed previously, wrinkle ridges are common on the basin ridged unit in Hellas Planitia, on the paterae and their associated flank materials, and on Hesperia and Malea Planum.

The Hellas Planitia wrinkle ridges probably date back to the Late Noachian, which is the large-crater age of the basin ridged unit (Table 4). Many of these ridges have been modified by mantling and exhumation (Figures 9–11), which tends to mask some of their secondary morphologic characteristics. Wrinkle ridges were not mapped in Hellas Planitia in global surveys by *Chicarro et al.* [1985, Figure 4] and *Watters* [1993, see Figure 2a and Plate 2], who described ridges in southern Hellas Planitia as nontectonic, erosional "simple ridges." However, *Greeley and Guest* [1987] and our study indicate that wrinkle ridges did form in Hellas Planitia. The ridges in Hellas are made up of relatively resistant material (compared to overlying surficial deposits), and they generally preserve gross, and locally detailed, diagnostic morphologic features of wrinkle ridges. In particular, en echelon, sinuous planforms and local preservation of superposed crenulations (i.e., narrow ridges on top of broader arches) of some of the ridges is

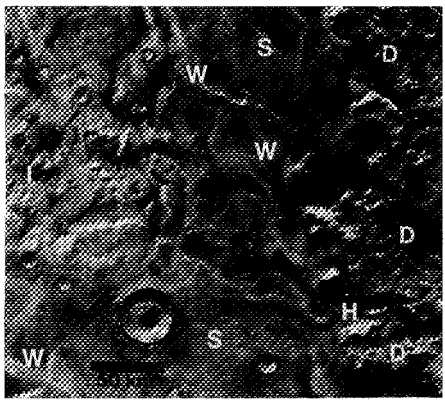


Figure 11. Terminus of Harmakhis Vallis (H) in eastern Hellas Planitia, showing deep dissection of vallis across dissected rim unit (D), anastomosing channel reach across subdued ridged unit (S), and interior deposit (I). Note vallis dissecting interior deposit and subdued appearance of wrinkle ridges (W). (Viking image 363S44; centered at 42.89°S, 276.48°, 256 m/pixel).

evident (Figures 9a and 10). Wrinkle ridges are present in an annulus of material surrounding the interior deposit (Figure 3); this interior unit in places appears to be draped over wrinkle ridges (Figure 10). Thus, the basin ridged unit probably stretches across the entire floor of Hellas, similar to how ridged plains material fills much of Chryse basin on Mars [*Scott and Tanaka,* 1986] and the lunar mare basins [e.g., *Wilhelms,* 1987].

The Hellas Planitia wrinkle ridges have a wide array of orientations; some preference to radial orientation is apparent. Compressional loading stresses would have developed within the basin due to its infilling by lavas (of which the thickness and mass are unknown) [cf. *Solomon and Head,* 1979]. In addition, near-global compressional stress apparently occurred during the Late Noachian [*Tanaka et al.,* 1991]; as a result, large horizontal compressional stress would have developed in the basin ridged unit. Although global compressive stress would tend to preclude ring-graben development, some normal faulting at Hellespontus Montes on the west rim of Hellas occurred during the Late Noachian (based on the crater age of *Wichman and Schultz* [1989]). These ring graben are consistent with a basin-loading stress model produced by at least 1.2 km of basalt [*Wichman and Schultz,* 1989]. Finally, the stress model predicts magma ascent outside the basin load [*Solomon and Head,* 1979], which is consistent with the possible vent along the north edge of the basin ridged unit mentioned previously. This model is favored because it predicts basin-rim normal faulting, basin compressional stresses, and magma ascent along the basin rim, all of which are observed in and around Hellas and may be coeval.

Complex patterns of wrinkle ridges are also well recognized on Hesperia and Malea Plana [*Chicarro et al.,* 1985; *Watters,* 1993]. Many of these wrinkle ridges have circular patterns interpreted as resulting from subsidence over crater structures [*Watters,* 1993]; similarly, Amphitrites Patera is outlined by a circular wrinkle-ridge structure that may be controlled by a buried, circular collapse structure (Figure 7a). In addition, a few systems of wrinkle ridges radiate north-south and east-west from Amphitrites and Peneus Paterae (Figure 7a), and others appear concentric to Amphitrites (Figure 8); these relations suggest some influence of the paterae on the local stress field.

6. Landscape Evolution: Eolian, Fluvial, and Mass-Wasting Modifications

This section discusses observations used to infer past eolian and fluvial activity in the Hellas region of Mars. In some cases, other processes and factors may also account for the observations. (For example, many features in the Hellas region have been interpreted as glacial or periglacial in origin [*Kargel and Strom,* 1992]; this hypothesis is discussed in section 6.8.) Relative ages (Tables 1 and 4) are deduced from crater densities shown in Tables 1 and 3. As mentioned earlier, because of the apparent significant obliteration of smaller craters over much of geologic time, emplacement ages are more confidently assigned from the densities of larger craters (generally >5 km in diameter). Also, because craters superposed on a geologic unit are commonly obliterated up to a particular size range, we have derived a table of density values for crater-size bins (using a $\sqrt{2}$ geometric progression) for each Martian series (based on the scheme of *Tanaka* [1986]) for craters larger than 2 km in diameter (Table 2). (Recall that "series" are time-stratigraphic units and "epochs" are corresponding periods.) The results of this method for our crater counts of Hellas units are shown in Tables 3 and 4.

6.1. Noachian/Hesperian Highland and Patera Modification

The earliest modifications in the Hellas region affected the rugged basin rim unit. The unit is intensely cratered, particularly on the southeast side of the basin (Figure 6), which has one of the highest densities of craters >32 km in diameter on the planet [e.g., *Tanaka,* 1986]. That area, compared to other highland areas, has a relatively low abundance of craters in the few tens of kilometers size range when compared to the abundance of larger craters [*Woronow,* 1977]. This result may be due to post-Hellas impact obliteration and to the rugged topography, which would favor rapid degradation of smaller craters on the flanks of steep massifs and which would provide deep intermassif basins in which smaller craters could be buried by later deposition. Local channeling and eolian mantling (some of which is post-Noachian) contributed to the degradation of the massifs and to the deposition of intermassif plains material throughout the basin rim unit.

Highland material that mantles and embays the basin rim unit also shows significant erosion. In contrast to the craters on the basin rim unit, which mostly have preserved rims, most craters <20 km in diameter in the highland material appear "rimless" (Figures 12–14). The term "rimless" refers to craters that have little to no outer raised rim but have a distinctive inner rim scarp and crater floor that is below the level of the adjacent surface [*Grant and Schultz,* 1993]. In addition, some

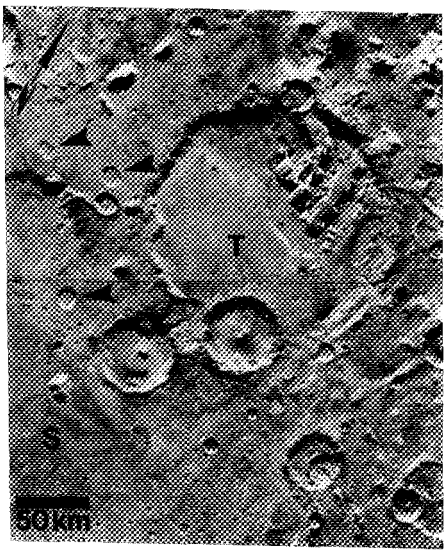


Figure 12. Crater Terby, on north rim of Hellas basin. Subdued ridged unit (S) embays rugged basin rim unit. Floor of crater Terby (T) is covered by a smooth, deeply etched rim mantle. Note rimless craters (arrows) common on relatively smooth, intermassif highland material of basin rim, whereas most craters on massifs and hilly parts of basin rim have well-defined rims. (Viking images 363S54, 56; centered at ~28.75°S, 286.00°, 257 m/pixel).

craters of all sizes on highland material have well-preserved rims and ejecta. These observations appear independent of latitude (and thus the effects of ice-creep [see *Squyres and Carr,* 1986]), but they may indicate that craters on highland material were highly eroded during the Noachian or Early Hesperian [*Tanaka,* 1986; *Wilhelms and Baldwin,* 1989; *Craddock and Maxwell,* 1990, 1993] and that craters on the basin rim unit were better preserved. The latter deduction suggests that highland materials are poorly indurated and fine grained, whereas the basin rim unit is likely made up of more resistant rock. In western Terra Tyrrhena, valleys are rare but mantles are present, which indicates that eolian mantling and deflation may have muted crater rims (see Figures 4 and 13; also compare with rimless craters in northeastern Arabia Terra [*Grant and Schultz,* 1990; *Moore,* 1990]).

Also during the Noachian or Early Hesperian, much of eastern Terra Tyrrhena and a few intercrater basins in Promethei Terra were dissected by arrays of small valleys (Figure 14). The valleys debouch into small intracrater and intercrater basins that appear to be partly filled by deposits. Although rimless craters are common in the dissected material, the valleys appear to postdate crater-rim degradation, because the valleys cut the material that mutes the craters. Where valleys do cross crater rims, they either cut perpendicularly across the rim (thus removing only a small section of rim), or,

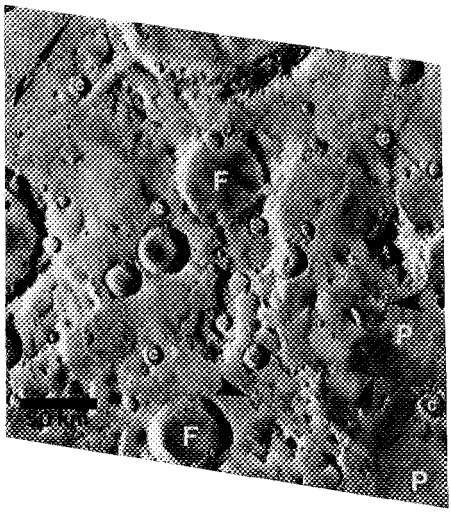


Figure 13. Part of northwest basin rim of Hellas showing smooth mantles in crater floors (F) and on intercrater plains (P). Locally, mantles appear etched (arrows). Mantles appear to contribute to formation of rimless craters (c). (Viking image 620A27; centered at 26.21°S, 307.25°, 235 m/pixel).

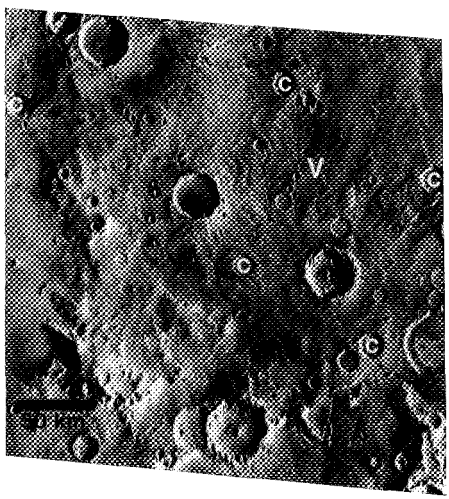


Figure 14. Part of Terra Tyrrhena along northeast rim of Hellas basin. Valley systems (V) cut across broad slopes and crater rims. Rimless craters (c) are common. (Viking image 625A31; centered at 24.99°S, 266.88°, 242 m/pixel).

in rare cases, they cut along the rim to below the level of the floor, which results in a raised crater floor. Most rimless and degraded craters show extensive gully development on the inner rim, suggesting that fluvial erosion contributed to rim degradation [see *Grant and Schultz,* 1994].

Valley erosion is also prevalent on the flanks of the four paterae in the study region. On Hadriaca and Tyrrhena Paterae, deeply etched and exposed layers demonstrate erosion by wind, water, and (or) mass wasting of pyroclastic deposits [*Greeley and Crown,* 1990; *Crown et al.,* 1992]. The valleys on Tyrrhena are distinctly flat-floored, whereas those on Hadriaca are V-shaped. This difference may be controlled by the presence of an upper layer(s) on Tyrrhena that has a relatively weak base along which scarp retreat progresses. Amphitrites Patera (Figure 7a) is surrounded by a low but intensely gullied shield that grades into northern Malea Planum, which is also dissected (by Axius Valles) all the way to Hellas Planitia (Figures 8 and 10). These valleys are more irregular in form than those of Hadriaca and Tyrrhena; this difference may have been caused by the lack of distinctive or continuous layering and by the prevalence of topographic obstacles produced by wrinkle ridges and possible flows, as well as by subsequent eolian mantling and deflation in the Malea Planum area (discussed below). Peneus Patera (Figure 7), on the other hand, does not cap a discernible shield, and valleys originate only on its sloping north edge.

Crater counts indicate that the erosion of the paterae and associated plains has been long lived. At decreasing crater sizes, crater densities show decreasing relative age (Tables 1 and 4), from Late Noachian to Middle Amazonian. At northern Malea Planum, where the largest crater dataset exists, the density of craters >16 km in diameter yields a Late Noachian age. Some of these larger craters appear embayed or dissected, and thus the age represents an intermediate age between the underlying basin rim unit and overlying dissected material, which is probably at least several hundred meters thick (given the embayment relations with large craters and knobs). Craters between 4 and 16 km across yield Hesperian ages, which probably reflect the age of the dissected material. At diameters <4 km, crater densities yield Amazonian ages, which probably date the eolian modification (which likely affected the upper tens to couple hundred meters of the patera units).

6.2. Lower Hesperian Basin Infilling and Modification

On the floor of the Hellas basin, a thick interior deposit covers about 10^6 km^2 (Figure 3). The deposit overlies Upper Noachian to Lower Hesperian ridged plains material, and it is embayed by fluvial and eolian materials of Late Hesperian and Amazonian age. Although its thickness has not been precisely determined, it probably exceeds 1 km over much of its extent. The upper part of one of the tallest erosional scarps within the deposit has a height (determined by shadow measurement) of about 700 m (Figure 5); the lower part of the scarp was not in shadow and the base of the deposit is at some unknown depth. Least eroded areas occur mostly in the eastern part of the deposit (Figure 4), and most of these areas appear armored by crater ejecta (Figure 15). Erosional landforms within the deposit include steep scarps, hummocky and knobby terrains, and deeply etched margins and hollows (Figures 15 and 16).

The crater-density distribution for the interior deposit is

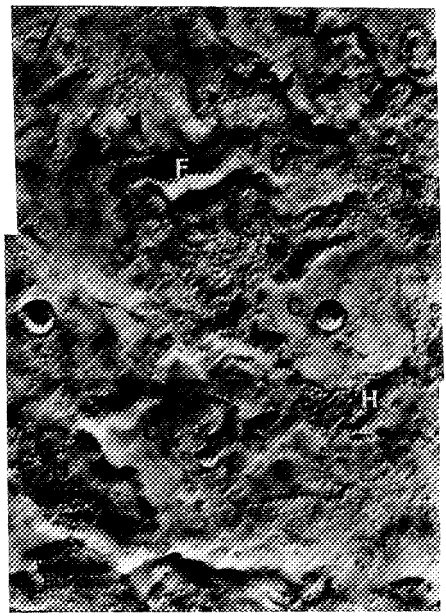


Figure 15. Part of interior deposit, Hellas Planitia, showing relict plateaus armored by crater ejecta (C), hummocky and knobby terrain (H), a smooth flow (F) originating from a collapse depression (D), and small dune fields (arrows). (Viking images 363S28 and 30; centered at ~ 40.50°S, 289.50°, 250 m/pixel).

difficult to interpret because of very deep erosion in most places. In addition, relict surfaces may be preserved by ejecta from larger craters, causing preferential survival of larger craters. These effects result in a low-sloping crater distribution, in which N(16) yields a Late Noachian age and N(2–2.8) an Early to Middle Amazonian age (Table 4). Highly degraded hummocky and knobby areas of the interior deposit yield Hesperian to Middle Amazonian ages, indicating that crater obliteration has been long lived. Based on the Late Noachian age for the underlying basin ridged unit, the preferential preservation of larger craters, and the relative age of younger deposits (as well as modified surfaces of the interior deposit); we assign an Early Hesperian age to the unit (the Late Noachian crater density is suspect because of the erosional preservation of surfaces armored by the ejecta of large craters). The cratering result, along with the observed erosional features, may indicate that initial erosion as much as several hundreds of meters deep during the Early or Late Hesperian was followed by relatively modest erosion (typically, tens of meters deep) through the Late Hesperian and Amazonian.

Erosion of the interior deposit probably occurs primarily by eolian deflation. This conclusion is based on the friable nature of the deposits (which is attested to by deep erosion of

Figure 16. Part of north-central Hellas Planitia showing relatively high albedo interior deposit (I) surrounded by darker basin (B) and subdued (S) ridged units. Interior deposit has complex morphology, including hummocky (H), knobby (K), and relatively smooth (SM) areas; plateau at top capped by west-trending ridges that may be yardangs (Y). Relatively low areas are covered by north- to multi-trending patterns of dunes (D). Western border of interior deposit is marked by prominent scarp (arrows); crater (C) below scarp apparently is filled with outlier of interior deposit. (Viking images 363S32 and 36; centered at ~ 35.75°S, 292.50°, 250 m/pixel).

the deposit except where armored by crater ejecta), the paucity of possible fluvial channels, the intense wind stresses [*Greeley et al.*, 1993], and the dunes (Figures 4, 9, 15, and 16) that are common in Hellas Planitia. The materials of the interior deposit are likely to be fine grained and poorly consolidated, which is in agreement with an interpretation made from Viking IRTM data [*Moore and Edgett*, 1993]. Depletion of fines that likely resulted from eolian deflation probably has led to the development of a sandy, possibly indurated surface lag deposit [*Moore and Edgett*, 1993], which may have contributed to dune deposits in the basin and surrounding highlands. Locally, the interior deposit may be degraded by collapse and mass wasting. A smooth flow deposit originates from a shallow depression (Figure 15), which may have formed due to liquefaction and collapse, similar to outflow channels on the east rim of Hellas and elsewhere on Mars [*Carr and Clow*, 1981; *Nummedal and Prior*, 1981]. Some debris aprons found mostly along steep scarps of the interior deposit may have formed with the assistance of interstitial ice. Thus, ground ice and water probably were present locally in the deposit; water and ice could have been supplied from frost and from the runoff of various channels that debouch onto the basin floor.

The interior deposit is not only deeply eroded, it also appears to have undergone significant lateral retreat in the northern, western, and southern parts of Hellas Planitia, where it may have once covered virtually the entire extent of the basin ridged unit (see Figure 3). Along the edge of proposed retreat, the interior deposit is marked by a fairly continuous prominent scarp, which is only absent along the eastern edge of the deposit. Beneath this scarp, the basin ridged unit is superposed by several pedestal craters (e.g., Figures 9 and 10) and partly exhumed craters that appear to have been once completely buried (vicinity of latitude 33° S., longitude 298°). Dunes, knobby outliers of the interior deposit, and younger mantle material are common below this scarp (Figures 9 and 16). Also, the edge of dissected patera flank material on northern Malea Planum forms a distinctive scarp, perhaps marking the location where the flank material once abutted the preretreat edge of the interior deposit (Figure 10). The total area of retreat may have been as much as 0.3 x 10^6 km^2.

Given the likelihood that the interior deposit is fine grained and poorly indurated, what is its possible origin? Some obvious sources include clastic sediment from valleys that feed into Hellas basin and pyroclastic fines generated by highland paterae on the Hellas rim (see Table 5). However, these sources probably have provided only a small part of the volume of the interior deposit (which at one time amounted to ~1–3 x 10^6 km^3 or more prior to its erosion). Small valleys elsewhere on Mars generally have filled only small highland basins; much more clastic material can be provided by outflow channels. However, the outflow channels that debouch into Hellas postdate the interior deposit and emplaced into the basin only about 5% or less equivalent of sediment volume of the deposit (see Table 5). The highland paterae appear to have generated a similar volume of material, based on the 1–2 km maximum heights of the paterae [*Blasius and Cutts*, 1981; *U.S. Geological Survey*, 1991; *Mouginis-Mark et al.*, 1992] and the mapped extents of their surrounding deposits, which probably are hundreds of meters thick (Figure 3; *Greeley and Crown* [1990]; *Crown and Greeley* [1993]). However, the deposits of Tyrrhena and Hadriaca Paterae, given their distribution, prevailing wind directions, and theoretical eruption dynamics, are more likely to be pyroclastic flows rather than air-fall deposits [*Greeley and Crown*, 1990; *Crown and Greeley*, 1993]; similar arguments can be made for Peneus and Amphitrites Paterae (see above). These materials are deeply dissected, and the dissected material on the south rim of Hellas would have been the largest source for fluvial transport. This dissection may have overlapped in time with formation of the interior deposits. Overall, erosion of rim materials may have contributed only as much as 18% of the original volume of the interior deposit (Table 5).

By default, the majority of the deposit must therefore originate from beyond the rim of Hellas, which requires eolian transport. Eolian deposits of various ages have been proposed in various highland and lowland areas of Mars [e.g., *Schultz and Lutz*, 1988; *Grant and Schultz*, 1990, 1993; *Moore*, 1990]. Possible eolian deposits of similar volume include the Medusae Fossae Formation (~4 x 10^6 km^3; *Scott and Tanaka*

Table 5. Estimated Volumes of Material Eroded From the Hellas Rim

Unit or feature	Area, km^2	Average Depth of Erosion, km	Volume Contributed, km^3
Valles material	50,000*	1.0*	50,000†
Dissected rim unit, Malea Planum	512,364*	0.2‡	102,473
Paterae			
Tyrrhena	25,544§	0.3‖	7,663
Hadriaca	103,089‖	0.3‖	30,927
Amphitrites	110,447*	0.3	33,134
Peneus	12,272*	0.3	3,682
Total			227,879

*This work.

†Note that east rim valles deposits postdate the age of the Hellas interior deposit. Therefore, contributions from observable eroded rim materials are 177,879 km^3, which is only ~18% of volume of the interior deposit assuming minimum volume of deposit of 10^6 km^3.

‡*Johnson et al.* [1991].

§*Plescia and Saunders* [1979].

‖*Crown et al.* [1992].

[1982]), a sequence of friable materials along the highland/lowland boundary between Elysium and Olympus Montes [*Scott and Tanaka,* 1986; *Greeley and Guest,* 1987], and the polar layered deposits (~3 x 10^6 km^3, assuming an average thickness of 1 km and their areal extent as given by *Tanaka et al.* [1988]). Given the ~Early Hesperian age of the interior deposit, the eolian material may have been derived from the fine fraction of eroded highland materials [*Wilhelms and Baldwin,* 1989; *Craddock and Maxwell,* 1993], fretted and knobby terrain development along the highland/lowland boundary [*Tanaka,* 1986], and fines produced by impacts [e.g., *MacKinnon and Tanaka,* 1989].

Apparently Hellas basin provided a sink for these materials. The depth of the Hellas basin is accompanied by elevated surface pressures within the lower parts of the basin, thereby allowing CO_2 to condense at relatively higher temperatures [*James et al.,* 1992]. Presently, CO_2 frost covers the entire basin floor during midwinter in the southern hemisphere [*Briggs et al.,* 1977]. Any available dust particles act as nuclei for condensing CO_2, leading to the precipitation of these dust/frost particles onto the Hellas floor. In this way, large volumes of available dust could have been deposited within Hellas. This depositional mechanism may also contribute a significant fraction of volatiles within the interior deposit.

6.3. Hesperian-Amazonian Rim Mantling and Channeling

During the Late Hesperian/Early Amazonian, many small and a few large channels dissected the eastern and southern rim areas of Hellas basin (Figures 7a, 8, and 17). On the east rim (Figure 17), Niger, Dao, and Harmakhis Valles may have resulted from catastrophic outbreaks related to volcanism or intrusion at Hadriaca Patera and volcanic flows from Tyrrhena Patera [*Squyres et al.,* 1987]. Reull Vallis originates from ridged plains material of southern Hesperia Planum and therefore could also have an origin partly related to volcanism and perhaps controlled by basin structure. Its terminus is abrupt near the head of Harmakhis Vallis, where it appears to be partly buried by an adjacent debris apron (Figure 17), indicating that at one time Reull Vallis may have been continuous into the head of Harmakhis Vallis. In total, the eastern valles cover about 0.05 x 10^6 km^2, and deeper sections locally exceed 1 km (Figure 5). The extensive dissection of the valles may be in part due to the nature of the dissected material, which may include eolian mantles [*Greeley and Guest,* 1987].

All of the valles except Reull terminate in eastern Hellas Planitia, incising into the dissected rim unit and the interior deposit (Figure 11). Much of the deposits of Reull Vallis may make up parts of the dissected rim unit (Figure 17), whereas the other valles would have deposited sediment onto eastern Hellas Planitia. Such deposition in this area would explain the subdued appearance of wrinkle ridges and the density of craters 2–8 km across (Table 4). However, the densities of craters larger than 8 km yield chiefly Late Noachian to Early Hesperian ages, which suggests that the basin ridged unit underlies the subdued ridged unit. Given the 0.3 x 10^6 km^2 extent of the subdued ridged unit and a maximum valles volume of about 0.05 x 10^6 km^3, the average depth of burial is about 170 m, which is consistent with burial of most craters up to a few kilometers in diameter [*Pike and Davis,* 1984] and partial burial of wrinkle ridges (measured elsewhere on Mars at up to 150–200 m in height [*Golombek et al.,* 1991]). This interpretation is also consistent with the lack of fan development at the mouths of the channels, which indicates that deposition was in the form of broad sheets [*Ori,* 1994]. Locally, viscous flow of the deposit is indicated where the subdued ridged unit forms a lobate scarp and embays rim material north of the mouth of Dao Vallis.

The south rim of Hellas also has numerous small channels and at least one large fretted channel, Mad Vallis (Figure 6). The small channels here are generally deeper and less sinuous than their eastern rim counterparts and are locally infilled with relatively higher albedo material that in places choked off

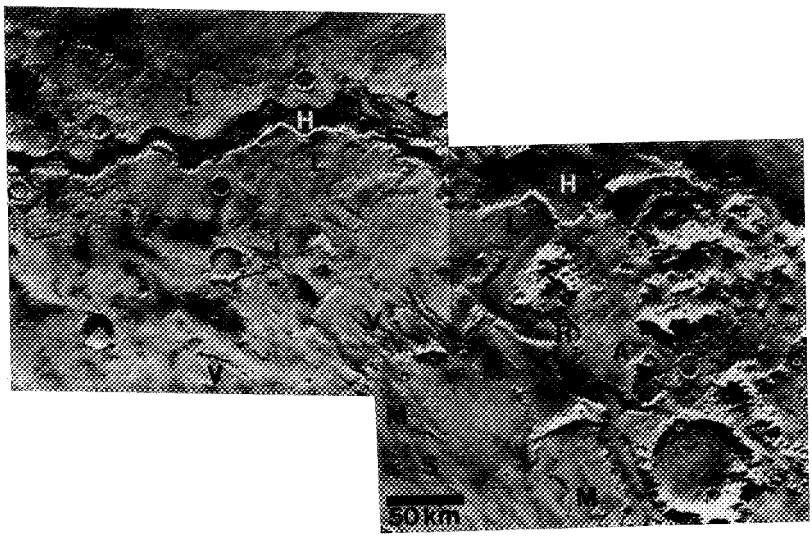


Figure 17. Harmakhis Vallis (H) and lower extremity of Reull Vallis (R) on eastern rim of Hellas. Rim material locally is dissected by degraded-appearing valleys (V) that emerge from smooth plateau deposits (S). These deposits in turn are capped in places by etched appearing mesas (M). Smooth, lobate, channeled material (L) may be deposit of Reull Vallis; note how deposit surrounds and channels cut inliers of rim material (I). Note debris aprons (A) along walls of valles and bases of massifs and layering in south wall of Reull Vallis (arrows). (Viking images 329S25 and 26; centered at ~ 40.00°S, 265.00°, 235 m/pixel).

channels completely. Their proximity to Amphitrites and Peneus Paterae, coupled with an observed paucity of channels on the opposite side of the local basin rim, suggest that they too may have formed by groundwater sapping or catastrophic outbreaks related to local intrusions and volcanism. Surface runoff, by contrast, should have produced channels on both sides of the high south rim.

On the upper slopes and rim of the Hellas basin, smooth, flat, irregularly shaped mantles are locally preserved (Figure 4). These mantles (tens to hundreds of meters thick) occur (1) in shallow topographic traps (such as in highland material as discussed previously (Figure 13)), within Amphitrites and Peneus Paterae, and on smooth plains surrounded by wrinkle ridges; (2) on broad plains, such as northern Malea Planum (Figures 7b and 8) and south of Harmakhis and Reull Valles (Figure 17), (3) within large craters (Figure 12), and (4) as pedestals beneath impact ejecta (Figures 8 and 10). Because pedestal craters of fairly high density occur on Malea Planum and small craters of moderate density are superposed on some of the mantles within craters, at least some of the mantles are probably Hesperian [see *Tanaka and Scott*, 1987]. However, other mantles appear uncratered (at 200–300 m/pixel resolution) and are likely to be Amazonian. These young mantles commonly appear etched, such as the material that thinly blankets the northern parts of Peneus Patera (Figure 7b) [*Zimbelman et al.*, 1989] and the dissected slopes of northern Malea Planum. Some of the crater-fill mantles have pits hundreds of meters deep (Figure 12; see also *Thomas* [1984]). Pits and etched features indicate the poorly indurated, friable nature of these mantles.

6.4. Hesperian-Amazonian Interior Mantling

Along the south edge of the floor of Hellas, in a topographic low between the Lower Hesperian interior deposit and the north edge of Malea Planum, smooth material of relatively low albedo forms a nearly continuous blanket (Figure 10). The blanket is made up of broad (>70 km across), rolling mounds having gently sloping margins that embay or feather into adjacent units. The blanket appears to be covered locally by dunes, and thus it may be made up of thick sequences of dune material in places. Elsewhere, the unit is pitted and etched, suggesting that the deposit is made up of friable material, likely dust or silt. Its crater density indicates a Late Hesperian or Early Amazonian age. In addition, possible eolian mantles that appear wind sculptured and similar in age to the blanket form generally sharp-edged, etched, irregular mesas near the center and the northwest edge of the Lower Hesperian interior deposit.

6.5. Amazonian Debris Aprons

Many steep slopes in the map area have lobate debris aprons that extend for several to more than ten kilometers from their bases. These lobes are on massifs of the basin rim unit (Figure 17), on channel walls of the dissected material (Figure 17), and on a few high mesa walls in the interior

deposit. These features have been interpreted as talus; local pitting suggests that the materials may have been ice-rich [*Crown et al.*, 1992]. The lobate form, steep slopes, and pitting of the aprons indicate emplacement from ice creep, as in terrestrial rock glaciers [*Squyres*, 1979; *Crown et al.*, 1992].

6.6. Recent Wind Features: Dunes, Yardangs, and Splotches

Occurrences of dunes in the Hellas region are indicated by (1) linear, arcuate, and reticulate patterns of low ridges in the central, northern, and western parts of the Hellas floor (Figures 4, 9, and 15), and (2) individual ridges (perhaps climbing dunes) on the top and flanks of some irregular mesas within the Hellas basin (Figures 15 and 16). Also, dark dunes occur within craters west of the Hellas basin [*Thomas*, 1984]. A few large occurrences of reticulate ridge sets interpreted to be dune fields occupy the central and northwestern parts of the basin (Figures 9a and 16). The size and morphology of these features are similar to networks of compound crescentic dunes (also termed akle dunes) observed in the Takla Makan Sand Sea, China (Figure 9b), as well as at many other terrestrial locations [*Breed and Grow*, 1979]. Crescentic dunes include transverse types, which generally infer a specified wind regime based upon slipface or ridge-axis orientation, where ridge axes develop transverse (normal) to the resultant direction of effective winds.

By contrast, yardangs are ridges formed by wind fluting and are prevalent in some areas of Mars, where friable materials and strong winds exist [e.g., *Ward*, 1979]. In north-central Hellas Planitia, possible yardangs cap a large plateau (Figure 16). The wind direction inferred from these features is ambiguous, as no blunt or tapered ends are readily discernible at existing image resolutions. However, most yardangs on Mars are more elongate than those on earth and may be carved into ridges with very low width-to-length ratios [*Ward*, 1985], thereby making it more difficult to detect tapering. Another possible set of yardangs occurs in the center of Hellas Planitia and is marked by numerous small (<1 to 10 km long) knobs, many of which are elongate roughly north-south (Figure 18). Within this set, tapering is readily apparent, with tapered ends predominantly pointing north-northeast.

Some Viking images reveal seasonal, surficial, bright and dark splotches in various areas of the Hellas region, which have been interpreted to be eolian frost and sand deposits [*Thomas et al.*, 1979; *Thomas*, 1984]. High-albedo, high-latitude splotches are likely CO_2 frost streaks. Other features attributable to wind include light and dark albedo streaks that trail from crater rims and other topographic obstacles (Figure 19). Some of these features are elongate and linear, indicating modification by a predominant, directional wind.

In addition to the local transport of sand and dust indicated by this discussion, net dust removal from Hellas and accumulation in equatorial regions is presently occurring [*Moore and Edgett*, 1993; *Christensen*, 1986]. Saltation and dust devils are primary mechanisms that can account for lifting dust into the atmosphere, and the high predicted wind stresses in Hellas favor this activity [*Greeley et al.*, 1992, 1993]. The identification of possible dune fields both within and outside the Hellas basin shows that saltation probably does occur or had previously occurred in this region, which coincides with observations that Hellas is a major source region for regional and global dust storms on Mars [*Martin and Zurek*, 1993]).

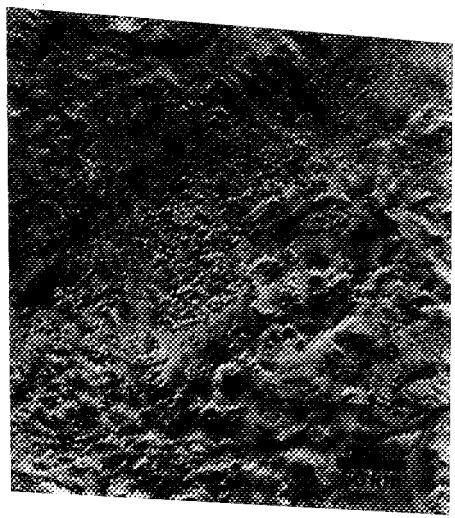


Figure 18. Knobby terrain in central Hellas Planitia represents highly degraded remnants of the Lower Hesperian interior deposit. The edge of a large crater (C) armors the deposit beneath it. Elongate knobs (K) are interpreted as yardangs and show definite tapering to the north-northeast. This orientation is consistent with high wind stresses predicted by a Mars general circulation model [*Greeley et al.*, 1993], as well as with inferred wind directions from nearby dunes (D) and splotches. Note smooth material (S) embaying the knobby material, and the superposed pedestal craters (P). These features indicate mantling and stripping of relatively unconsolidated, likely eolian, material. (Viking image 363S12; centered at 41.50°S, 300.44°, 247 m/pixel).

Depositional cycles of 10^5–10^6 years under present astronomical and climatic conditions may reverse this transfer and result in cyclic formation of dust deposits as much as a few meters thick [*Christensen*, 1986; *Greeley et al.*, 1992].

6.7. Theoretical Considerations Of Modern Sand and Dust Transport in Hellas

Many of the features described in the previous section show orientations indicative of wind direction. In this section we compare these inferred wind directions with those directions predicted by a recent Mars general circulation model (GCM) [*Greeley et al.*, 1993; see also *Skypeck*, 1989]. The *Greeley et al.* [1993] GCM is a three-dimensional model that includes the effects of varying global topography, albedo, thermal inertia, polar CO_2 condensation, and dust opacity. Each cell in the 25 by 40 cell computational grid is 7.5° latitude by 9.0° longitude, and the model runs cover only six brief, discreet L_S (Martian aerocentric longitude) intervals. Although the runs were spaced fairly evenly at ~60° L_S intervals, the model integrations at these L_S intervals covered only ~5° of L_S. Therefore, observed features were compared with the nearest model (seasonal) interval to when the image was taken, which in some cases differed by as much as 20° to 30° of L_S. Inferred

Figure 19. Light-colored streak (S) probably composed of frost or dust, drapes over an outlier of the interior deposit (I). Orientation of streak indicates northward or southward wind flow. Note subdued wrinkle ridges (W) and pedestal craters (P) on the basin ridged unit (B), indicative of subsequent mantling. (Viking image 578B14; centered at 48.14°S, 302.89°, 74 m/pixel).

wind vectors and GCM correlations are shown on Figure 20 and Table 6, respectively.

Transverse (dune) ridge orientations in the northwestern part of Hellas basin correlate very well with highest surface-stress winds predicted for the northern basin by the GCM. The orientations of most other transverse dune fields identified in the Hellas basin also correlate exceptionally well with the highest seasonal winds predicted by the GCM. These surface wind stresses have been attributed to broad slopes, high solar insolation during southern summer, and high atmospheric density within the basin, and they are sufficient to induce saltation of fine sand [*Greeley et al.*, 1993].

The yardang set in north-central Hellas Planitia has orientations that are only slightly askew from the GCM's highest wind vector predicted in this grid cell. The second set of yardangs in central Hellas Planitia also show orientations that are consistent with GCM predictions and with other nearby dune and transient-splotch orientations (Figure 20). Overall, the orientations of Hellas yardangs are in better agreement with GCM predictions than with other yardang orientations elsewhere on Mars [see *Greeley et al.*, 1993], perhaps because (1) the region currently receives the highest predicted wind stresses on Mars and (2) the interior deposit has very little internal structure or other zones of weaknesses to control the erosional forms.

Seasonal, surficial, bright and dark splotches in the Hellas region have varying degrees of correlation with the GCM. High-albedo, high-latitude splotches (likely CO_2 frost streaks) correlate well with predicted southern hemisphere autumn and winter winds coming off the annual polar cap [*Thomas et al.*, 1979;]. This relation fits for frost streaks observed along the south and southeast Hellas rim, as well as a couple of occurrences in the Hellas basin (see Table 6).

Most other ephemeral features correlate only marginally (53% in agreement) with GCM-predicted surface stresses of closest L_S (seasonal) interval; however the degree of agreement varies greatly from place to place (Table 6). In particular, the west rim and basin show excellent agreement between observed and predicted vectors, whereas all other areas have poor to marginal agreement (27% to 50%). The same ephemeral features also show marginal positive correlation (47%) with the GCM as a function of the maximum stress vector (maximum wind-stress direction regardless of season). The west and southeast rim vectors show the best GCM agreement (100% and 67%, respectively) as a function of maximum wind stress; remaining areas show poor to fair correlation.

The marginal correlation of observed transient feature orientations in places with those predicted by the GCM is not surprising and may be mainly due to the limited spatial and temporal resolutions of the model that do not detect rapid, local variations in wind patterns. In contrast to the relatively larger dune and yardang fields in Hellas, most splotches and streaks are much smaller than the spatial resolution of the model, where local factors such as topography, structure, and albedo, as well as seasonal atmospheric changes, can greatly influence winds. Even small-scale structure (including joint sets) can influence local wind direction and speeds [*Ward*, 1979].

The orientations of yardangs and dunes are generally controlled by the directions of the strongest seasonal winds, so the high correlation between the Hellas features and the GCM is not surprising. In addition, we observe an anticorrelation between the landform orientations and GCM predictions for other Martian seasons, even though these stresses are of considerable strength. One possibility is that these features are relict, preserving orientations governed by previous wind regimes. Also, we have not attempted to investigate whether the Hellas dunes show some control by secondary winds, which is entirely possible, particularly where dune patterns are complex (e.g., Figure 9a).

6.8. Discussion of Proposed Glacial Modification

Periods of net deposition and erosion are indicated by our stratigraphic and geomorphologic interpretations of the Hellas region. We attribute this history mainly to eolian and fluvial processes and, to a lesser degree, mass-wasting and ground-ice processes. *Kargel and Strom* [1992], on the other hand, proposed a late (Amazonian) glaciation event in Hellas. Their cited evidence for this event includes (1) erosional scouring and channel formation on northern Malea Planum, (2) east-trending ridges (interpreted as transverse and recessional moraines) and other related features interpreted as ice-stagnation landforms (such as eskers and drumlins) on the floor of Hellas, and (3) possible lacustrine plains that may have been material deposited in glacial meltwater. In addition, *Kargel and Strom* [1992] argue that the overall pattern and morphologic associations of these features attest to a glaciation event. Our detailed study, however, has led us to

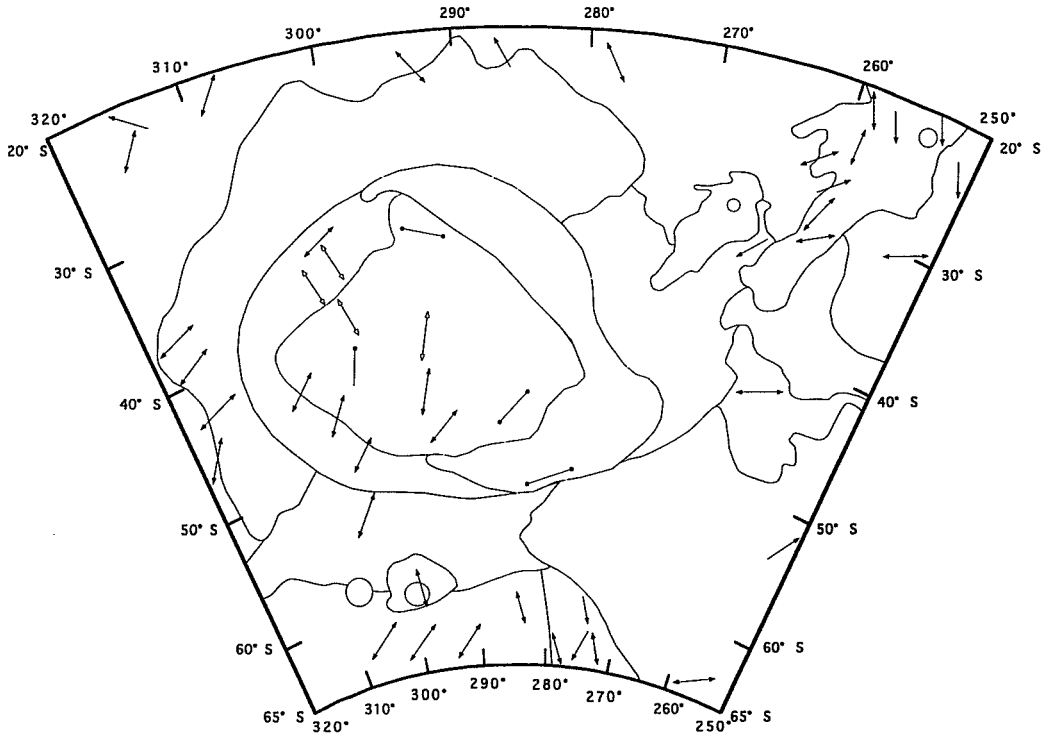


Figure 20. Wind vectors in the Hellas region as inferred from oriented dust and frost splotches and from yardang and dune forms. Arrows indicate downwind direction; bidirectional arrows indicate ambiguous wind directions. Solid arrows, dust and frost streaks; open arrows, dunes; circles, yardangs. Correlation of these wind vectors with GCM predictions is shown in Table 6.

Table 6. Wind Vectors in the Hellas Region of Mars as Inferred From Dust and Frost Streaks, Yardangs, and Dune Forms and Their Correlation With Predicted Wind Stress Vectors From the Mars GCM of *Greeley et al.* [1993].

Viking Image	Location (latitude, longitude)	Feature	L_s,* deg	Wind Vector	GCM Correlation†	
					L_S	Maximum Stress
			North Rim			
620A06	18°S, 315°	dark streak	53.2	N50W	+	−
620A05	22°S, 315°	light streak	53.2	N40E–S40W	+	−
620A09	18°S, 308°	light/dark streak	53.2	N30E–S30W	−	−
395S06	23°S, 294°	light/dark streak	96.3	N35W–S35E	−	+
397S10	20°S, 287°	light streak	97.2	N30W	+	−
397S05	23°S, 279°	light streak	97.2	N30W–S30E	−	+
			Northeast Rim			
365S44	20°S, 259°	light streak	83.1	N20W–S20E	−	−
365S42	23°S, 257°	light streak	83.1	S20E	−	−
365S43	19°S, 255°	light streak	83.1	S20E	−	+
365S39	24°S, 251°	light streak	83.1	S20E	+	−
087A42	29°S, 252°	light/dark streak	123.5	N60E–S60W	−	+
478A18	24°S, 259°	light/dark streak	346.1	NS	+	−
625A36	26°S, 261°	light streak	55.4	N50E–S50W	−	+
478A09	29°S, 260°	light streak	346.1	N20E	+	+
553A05	30°S, 260°	light/dark streak	23.0	N60E–S60W	−	+
625A20	34°S, 264°	dark streak	55.4	S40W	−	−
097A60	43°S, 260°	light/dark streak	127.9	N70E–S70W	−	+
			Southeast Rim			
097A75	53°S, 249°	dark streak	127.9	N30E	−	−
510B26	65°S, 255°	dark streak	31.5	N65E–S65W	+	+
504B51	63°S, 273°	light streak	28.8	N20W–S20E, S20W	−	+
504B50	65°S, 278°	light streak	28.8	N20W–S20E	+	+
504B53	61°S, 274°	light streak	28.8	S20E	−	-
504B35	60°S, 282°	light/dark streak	28.8	N20W–S20E	+	+
			South Rim			
431B41	63°S, 292°	light streak	353.6	N30E–S30W	−	−
094A75	59°S, 297°	light/dark streak	126.5	N10W–S10E	−	+
466B53	65°S, 300°	dark streak	10.9	N45E–S45W	−	−
431B40	61°S, 306°	light/dark streak	353.6	N45E–S45W	−	−
578B10	52°S, 305°	light streak	61.8	N25E–S25W	+	−
			West Rim			
510A54	43°S, 314°	dark streak	2.3	N70E–S70W	+	+
510A50	46°S, 321°	light/dark streak	2.3	N40E–S40W	+	+
584A28	40°S, 315°	dark streak	37.2	N60E–S60W	+	+
547A49	38°S, 318°	light/dark streak	20.2	N70E–S70W	+	+
			Hellas Basin			
620A47	34°S, 303°	light/dark streak	53.2	N60E–S60W	+	−
584A31	43°S, 306°	dark streak	37.2	N40E–S40W	+	−
584A12	46°S, 305°	dark streak	37.2	N30E–S30W	+	−
584A15	48°S, 293°	dark streak	37.2	N45E–S45W	+	+
578B14	48°S, 303°	light streak	61.8	N40E–S40W	+	−
584A16	46°S, 294°	dark streak	37.2	N20E–S20W	+	−
545B36	51°S, 281°	yardangs	47.3	N70E–S70W	−	+
363S24	47°S, 285°	yardangs	82.3	N40E–S40W	−	+
363S34	34°S, 293°	yardangs	82.3	N70W–S70E	−	+

Table 6. (continued)

Viking Image	Location (latitude, longitude)	Feature	L_S,* deg	Wind Vector	GCM Correlation†	
					L_S	Maximum Stress
363S12	41°S, 300°	yardangs	82.3	N20E	–	+
620A45	36°S, 307°	dunes	53.2	N15W–S15E	–	+
620A47	34°S, 303°	dunes	53.2	N15W–S15E	–	+
363S14	39°S, 302°	dunes	82.2	N25W–S25E	–	+
363S11	41°S, 294°	dunes	82.2	N15E–S15W	–	+

*L_S denotes aerocentric longitude of the Sun as measured in a Mars-centered fixed coordinate system. L_S = 90°, 180°, 270° marks beginning of Martian southern hemisphere winter, spring, and summer, respectively.

†Correlations are positive (+) if measured wind vector is within ±20° of that predicted by nearest GCM vector. L_S correlation relates to wind vector for GCM run of L_S closest to that of image; maximum stress correlation compares wind vector with GCM vector of highest magnitude for any season. Note that most wind vectors are bidirectional, which indicates ambiguity in possible wind directions inferred from the feature; for these, correlations are made with wind vector closest to that of GCM.

different conclusions, as stated previously. Here we reassert these conclusions for each point of evidence of the glacial hypothesis, and we provide further arguments that pose serious problems to the glacial interpretation.

Scouring of northern Malea Planum has been complex. *Johnson et al.* [1991] determined a Middle Amazonian population of craters with ejecta blankets, which according to *Kargel and Strom* [1992] represents the minimum age of scouring. In addition, the densities of degraded craters falls below a production distribution at smaller diameters, which *Kargel and Strom* [1992] suggest as evidence for ~200 m of scouring. However, we note that scouring is characteristic of all the volcanic paterae in the Hellas region. Scouring at Hadriaca and Tyrrhena Paterae has been attributed to wind, water (both runoff and sapping), mass-wasting, and lava-channel development [*Greeley and Crown,* 1990; *Gulick and Baker,* 1990; *Crown and Greeley,* 1993]. The scouring may thus be due to the erosion of friable material perhaps sapped by ground water circulated by volcanic heating. On Malea Planum, we also note the widespread evidence of mantling, which has obscured many channels and led to the formation of numerous pedestal craters. These observations indicate that channeling was early and that the fall-off in degraded craters at smaller diameters can be explained by the deposition and removal of thick (100–200 m) eolian mantles that appear in high-resolution images (e.g., Figure 7b).

Ridges in Hellas Planitia interpreted by *Kargel and Strom* [1992] as moraines may have a variety of other origins. The proposed terminal moraine is simply the lobate bounding scarp of the Hellas interior deposit (Figures 9a and 10). We attribute this prominent scarp to eolian backwasting of a fine-grained deposit, which also appears to produce the morphologically similar margin of the south polar layered deposit [*Thomas,* 1982, Figure 4b]. Beyond and within the scarp, dunes and knobby material (some of which occurs in craters) are interpreted as derivatives and outliers of the deposit (Figures 9a and 16); again, polar layered deposits show the same associations [*Thomas,* 1982; *Tanaka and Scott,* 1987]. Furthermore, the orientations of the interior-deposit-bounding scarp are inconsistent with the deposition of moraine material from the proposed direction of glacier flow (from the Hellas south rim moving north into the basin). The southern edge of the interior deposit is marked by a prominent scarp where we would expect the proposed glacial deposits to be more continuous. The eastern side of the interior deposit is the only part of the deposit that does not have a bounding scarp, and thus shows the only definitive continuity between Hellas basin materials and rim-derived materials (which we interpret as emplaced by local valles).

The ridges thought to be recessional moraines, including a string of mesas [*Kargel and Strom,* 1992, Figure 4] that compose a group of plateaus on the Hellas interior deposit, appear to us as erosional remnants of the original surface material of the interior deposit that are locally armored by crater ejecta (e.g., Figure 15). Other ridges may be structural (Figure 10). The proposed "drumlins" in western Hellas are knobs elongate in a northerly direction. An alternative explanation is that the elongation is due to wind fluting, which forms yardangs. The knobs are associated with low mesas and small pedestal craters, some of which have elongate pedestals (Figure 19). Adjacent areas north and east of the knobs appear covered by transverse dunes oriented perpendicular to the knobs (see Viking images 363S12 and 14). Thus, a yardang interpretation is consistent with apparent prevailing wind direction, as inferred by local dune orientation and GCM predictions and the friable nature of the knobby material [see *Ward,* 1979].

We agree with *Kargel and Strom* [1992] that the northeastern ridged plains of Hellas Planitia may have been covered by a temporary lake, for which we have cited evidence in the form of Late Hesperian/Early Amazonian deposition. However, as we have discussed previously, the likely source of the water, based on temporal and spatial relations, were the large valles on the east rim of Hellas. If the water source was from a melting glacier within Hellas Planitia, we might expect to see source channels coming off the interior deposit; none are observed.

Only local areas provide some equivocal geomorphologic evidence for ice sheets. First, a sinuous ridge system about 100 km long in southwestern Hellas Planitia (see Viking image 361S11) may be interpeted as eskers [*Kargel and Strom,* 1992]. However, inverted stream topography produced by evaporite cementation of stream gravels and subsequent exhumation [*Maizels,* 1990; *Rice and Mollard,* 1994] is

another good geomorphologic analogy, because of the apparent importance of armoring in Hellas landform development. Second, a patch of "thumbprint terrain" about 40 km across in central Hellas Planitia (see Viking images 363S11 and 13) may be formed by retreat or removal of ice [e.g., *Rossbacher and Judson,* 1981], although other explanations have been proposed [see *Schaefer,* 1990; *Rice and Mollard,* 1994].

In addition to showing alternative (and less problematic) interpretations to the glacier hypothesis, we find other serious difficulties with that proposal. *Kargel and Strom* [1992] proposed that glacial activity produced a large morainal deposit within Hellas basin during the Middle Amazonian. We have assigned the interior deposit an Early Hesperian age based on crater densities and stratigraphic relations. Also, the glacial hypothesis requires the rough equivalent of the entire estimated Martian water budget [*Kargel and Strom,* 1992].

Overall, the assemblage of landforms and crater chronology in Hellas basin indicate a landscape evolution dominated by eolian processes and punctuated by brief and less important fluvial events. Mass wasting, ground-ice processes, and perhaps glaciation played local, minor roles.

Acknowledgments. The research in this paper was stimulated largely by discussions at the wrap-up, Mars Surface and Atmosphere Through Time workshop in Houston in November 1993, in which planetary geologists, astronomers, and atmospheric scientists expressed a keen interest in understanding the eolian geology of Hellas basin. We appreciate the careful reviews provided by John Grant, Jeff Kargel, Baerbel Lucchitta, and Jeff Moore. Much of our capability to detect key geologic relations and to perform precise crater counts was based on working with digital Viking images on CD-ROM's produced by NASA's Planetary Data System. In addition, Kathy Hoyt (USGS Planetary Data Facility) made available to us digital copies of Viking images that were yet unavailable on CD-ROM. We also are grateful to Carol Breed (USGS), who provided us with expert consultation on dune forms and the Landsat image of the Takla Makan Sand Sea in China. This research was funded by NASA's Mars Surface and Atmosphere Through Time and Planetary Geology and Geophysics programs.

References

Blasius, K.R., and J.A., Cutts, Topography of Martian central volcanoes, *Icarus, 45,* 87-112, 1981.

Breed, C.S., and T. Grow, Morphology and distribution of dunes in sand seas observed by remote sensing, in A Study of Global Sand Seas, edited by E.D. McKee, *U.S. Geol. Surv. Prof. Pap. 1052,* 253–302, 1979.

Briggs, G.A., K. Klaasen, T. Thorpe, J. Wellman, and W. Baum, Martian dynamical phenomena during June–November, 1976: Viking Orbiter imaging results, *J. Geophys. Res., 82,* 4121–4149, 1977.

Carr, M.H., and G.D. Clow, Martian channels and valleys: Their characteristics, distribution, and age, *Icarus, 48,* 91–117, 1981.

Chicarro, A.F., P.H. Schultz, and P. Masson, Global and regional ridge patterns on Mars, *Icarus, 63,* 153–174, 1985.

Christensen, P.R., Regional dust deposits on Mars: Physical properties, age, and history, *J. Geophys. Res., 91,* 3533–3545, 1986.

Craddock, R.A., and T.A. Maxwell, Resurfacing of the Martian highlands in the Amenthes and Tyrrhena region, *J. Geophys. Res., 95,* 14,265–14,278, 1990.

Craddock, R.A., and T.A. Maxwell, Geomorphic evolution of the Martian highlands through ancient fluvial processes, *J. Geophys. Res., 98,* 3453–3468, 1993.

Crown, D.A., and R. Greeley, Volcanic geology of Hadriaca Patera and the eastern Hellas region of Mars, *J. Geophys. Res., 98,* 3431–3451, 1993.

Crown, D.A., K.H. Price, and R. Greeley, Geologic evolution of the east rim of the Hellas basin, Mars, *Icarus, 100,* 1–25, 1992.

De Hon, R.A., Geologic map of the Eridania quadrangle of Mars, *U.S. Geol. Surv. Misc. Inv. Ser. Map, I-1008,* 1977.

Edgett, K.S., Ejecta deposits of large martian impact basins: a useful geologic tool and window to early martian history?, in MEVTV Workshop on Early Tectonic and Volcanic Evolution of Mars, edited by H. Frey, pp. 32–34, *LPI Tech. Rep. 89–04,* 1989.

Gault, D.E., and J.A. Wedekind, Experimental studies of oblique impact, *Proc. Lunar Planet. Sci. Conf., 9,* 3843–3875, 1978.

Golombek, M.P., J.B. Plescia, and B.J. Franklin, Faulting and folding in the formation of planetary wrinkle ridges, *Proc. Lunar Planet. Sci. Conf., 21,* 679–693, 1991.

Grant, J.A., and P.H. Schultz, Gradational epochs on Mars: Evidence from west-northwest of Isidis and Electris, *Icarus, 84,* 166–195, 1990.

Grant, J.A., and P.H. Schultz, Gradation of selected terrestrial and martian impact craters, *J. Geophys. Res. 98,* 11,025–11,042, 1993.

Grant, J.A., and P.H. Schultz, Early fluvial degradation in Terra Tyrrhena, Mars: Constraints from styles of crater degradation on the Earth, (abstract), *Lunar Planet. Sci., XXV,* 457–458, 1994.

Greeley, R., and D.A. Crown, Volcanic geology of Tyrrhena Patera, Mars, *J. Geophys. Res., 95,* 7133–7149, 1990.

Greeley, R., and J.E. Guest, Geologic map of the eastern equatorial region of Mars, *U.S. Geol. Surv. Misc. Inv. Ser. Map I-1802-B,* 1987.

Greeley, R., and P.D. Spudis, Volcanism on Mars, *Rev. Geophys., 19,* 13–41, 1981.

Greeley, R., N. Lancaster, S. Lee, and P. Thomas, Martian aeolian processes, sediments, and features, in *Mars,* edited by H. Kieffer, B. Jakosky, C. Snyder, and M. Matthews, pp. 730-766, University of Arizona Press, Tucson, 1992.

Greeley, R., A. Skypeck, and J.B. Pollack, Martian aeolian features and deposits: Comparisons with general circulation model results, *J. Geophys. Res., 98,* 3183–3196, 1993.

Gulick, V.C., and V.R. Baker, Origin and evolution of valleys on Martian volcanoes, *J. Geophys. Res., 95,* 14,325–14,344, 1990.

James, P.B., H.H. Kieffer, and D. A. Paige, The seasonal cycle of carbon dioxide on Mars, in *Mars,* edited by H. Kieffer, B. Jakosky, C. Snyder, and M. Matthews, pp. 934-968, University of Arizona Press, Tucson, 1992.

Johnson, N., J.S. Kargel, R.G. Strom, and C. Knight, Chronology of glaciation in the Hellas region of Mars (abstract), *Lunar Planet. Sci.., XXII,* 651–652, 1991.

Kargel, J.S., and R.G. Strom, Ancient glaciation on Mars, *Geology 20,* 3–7, 1992.

Leonard, G.J., and K.L. Tanaka, Hellas basin, Mars: Formation by oblique impact (abstract), *Lunar Planet. Sci., 24,* 867–868, 1993.

MacKinnon, D.J., and K.L. Tanaka, The impacted Martian crust: Structure, hydrology, and some geologic implications, *J. Geophys. Res., 94,* 17,359–17,370, 1989.

Maizels, J., Raised channel systems as indicators of palaeohydrologic change: a case study from Oman, *Palaeogeogr. Palaeoclimatol. Palaeocol., 76,* 241–277, 1990.

Martin, L.J., and R.W. Zurek, An analysis of the history of dust activity on Mars, *J. Geophys. Res., 98,* 3221–3246, 1993.

Moore, H.J., Subsurface deformation resulting from missile impact, *U.S. Geol. Surv.Prof. Pap., 650-B,* B107–B112, 1969.

Moore, J.M., Nature of the mantling deposit in the heavily cratered terrain of northeastern Arabia, Mars, *J. Geophys. Res., 95,* 14,279–14,289, 1990.

Moore, J.M., and K.S. Edgett, Hellas Planitia, Mars: Site of net dust erosion and implications for the nature of basin floor deposits, *Geophys. Res. Lett., 20,* 1599–1602, 1993.

Mouginis-Mark, P.J., L. Wilson, and J.R. Zimbelman, Polygenic eruptions on Alba Patera, Mars, *Bull. Volcanol., 50,* 361–379, 1988.

Mouginis-Mark, P.J., L. Wilson, and M.T. Zuber, The physical volcanology of Mars, in *Mars,* edited by H. Kieffer, B. Jakosky, C. Snyder, and M. Matthews, pp. 424-452, University of Arizona Press, Tucson, 1992.

Nummedal, D., and D.B. Prior, Generation of Martian chaos and channels by debris flows, *Icarus, 45,* 77–86, 1981.

Ori, G.G., Depositional patterns at the mouths of the Martian outflow channels (Dao, Harmakhis, Maja, and Ares Valles, Mars) (abstract), *Lunar Planet. Sci., XXV,* 1029–1030, 1994.

Peterson, J.E., Geologic map of the Noachis quadrangle of Mars, *U.S. Geol. Surv. Misc. Inv. Ser. Map, I-910,* 1977.

Pike, R.J., and P.A. Davis, Toward a topographic model of martian craters from photoclinometry (abstract), *Lunar Planet. Sci., XV,* 645–646, 1984.

Plescia, J.B., and R.S. Saunders, The chronology of the Martian volcanoes, *Proc. Lunar Planet. Sci. Conf., 10th,* 2841–2859, 1979.

Potter, D.B., Geologic map of the Hellas quadrangle of Mars, *U.S. Geol. Surv. Misc. Inv. Ser. Map I-941,* 1976.

Rice, J.W., Jr., and J.D. Mollard, Analogs and interpretations for the Martian thumbprint terrain and sinuous ridges (abstract), *Lunar Planet. Sci. Conf., XXV,* 1127–1128, 1994.

Rossbacher, L.A., and S. Judson, Ground ice on Mars: Inventory, distribution, and resulting landforms, *Icarus, 45,* 39–59, 1981.

Schaefer, M.W., Karst on Mars? The thumbprint terrain, *Icarus, 83,* 244–247, 1990.

Schultz, P.H., Chicxulub as an oblique impact (abstract), *Lunar Planet. Sci., XXV,* 1211–1212, 1994.

Schultz, P.H., and A.B. Lutz, Polar wandering of Mars, *Icarus, 73,* 91–141, 1988.

Scott, D.H., and M.H. Carr, Geologic map of Mars, *U.S. Geol. Surv. Misc. Inv. Ser. Map I-1083,* 1978.

Scott, D.H., and K.L. Tanaka, Ignimbrites of Amazonis Planitia region of Mars, *J. Geophys. Res., 87,* 1179–1190, 1982.

Scott, D.H., and K.L. Tanaka, Geologic map of the western equatorial region of Mars, *U.S. Geol. Surv. Misc. Inv. Ser. Map I-1802-A,* 1986.

Skypeck, A. P., Comparison of a Mars general circulation model with aeolian features and deposits, M.S. thesis, 79 pp., Ariz. State Univ., Tempe, 1989.

Solomon, S.C., and J.W. Head, Vertical movement in mare basins: Relation to mare emplacement, basin tectonics, and lunar thermal history, *J. Geophys. Res., 84,* 1667–1682, 1979.

Squyres, S.W., The distribution of lobate debris aprons and similar flows on Mars, *J. Geophys. Res., 84,* 8087–8096, 1979.

Squyres, S.W., and M.H. Carr, Geomorphic evidence for the distribution of ground ice on Mars, *Science, 231,* 249–252, 1986.

Squyres, S.W., D.E. Wilhelms, and A.C. Moosman, Large-scale volcano-ground ice interactions on Mars, *Icarus, 70,* 385–408, 1987.

Tanaka, K.L., The stratigraphy of Mars, *Proc. Lunar Planet. Sci. Conf. 17th, Part 1, J. Geophys. Res., 91, Suppl.* E139–E158, 1986.

Tanaka, K.L., and D.H. Scott, Geologic map of the polar regions of Mars, *U.S. Geol. Surv. Misc. Inv. Map I-1802-C,* 1987.

Tanaka, K.L., N.K. Isbell, D.H. Scott, R. Greeley, and J.E. Guest, The resurfacing history of Mars: A synthesis of digitized, Viking-based geology, *Proc. Lunar Planet. Sci. Conf., 18th,* 665–678, 1988.

Tanaka, K.L., M.P. Golombek, and W.B. Banerdt, Reconciliation of stress and structural histories of the Tharsis region of Mars, *J. Geophys. Res., 96,* 15,617–15,633, 1991.

Thomas, P., Present wind activity on Mars: Relation to large latitudinally zoned sediment deposits, *J. Geophys. Res., 87,* 9999–10,008, 1982.

Thomas, P., Martian intracrater splotches: Occurrence, morphology, and colors, *Icarus, 57,* 205–227, 1984.

Thomas, P., J. Veverka, and R. Campos-Marquetti, Frost streaks in the south polar cap of Mars, *J. Geophys. Res., 84,* 4621–4633, 1979.

U.S. Geological Survey, Topographic maps of the polar, western, and eastern regions of Mars, *U.S. Geol. Surv. Misc. Invest. Ser. Map, I-2160,* 1991.

Ward, A.W., Yardangs on Mars, *J. Geophys. Res., 84,* 8147–8166, 1979.

Ward, A.W., Global map of eolian features on Mars, *J. Geophys. Res., 90,* 2038–2056, 1985.

Watters, T.R., Compressional tectonism on Mars, *J. Geophys. Res., 98,* 17,049–17,060, 1993.

Whitford-Stark, J.L., Modification of multi-ring basins—the Imbrium model, in *Multi-Ring Basins, Proc. Lunar Planet. Sci., vol. 12A,* edited by P. Schultz and R. Merrill, pp. 113-124, Pergamon, New York, 1981.

Wichman, R.W., and P.H. Schultz, Sequence and mechanisms of deformation around the Hellas and Isidis impact basins on Mars, *J. Geophys. Res., 94,* 17,333–17,357, 1989.

Wilhelms, D.E., The geologic history of the Moon, *U.S. Geol. Surv. Prof. Pap., 1348,* 1987.

Wilhelms, D.E., and R.J. Baldwin, The role of igneous sills in shaping the Martian uplands, *Proc. Lunar Planet. Sci. Conf., 19th,* 355–365, 1989.

Woronow, A., A size-frequency study of large Martian craters, *J. Geophys. Res., 82,* 5807–5820, 1977.

Zimbelman, J.R., S.M. Clifford, and S.H. Williams, Concentric crater fill on Mars: An aeolian alternative to ice-rich mass wasting, *Proc. Lunar Planet. Sci. Conf., 19th,* 397–407, 1989.

G. J. Leonard and K. L. Tanaka, Branch of Astrogeology, U.S. Geological Survey, 2255 North Gemini Drive, Flagstaff, AZ 86001-1698. (Internet: ktanaka@astrog.span.nasa.gov)

(Received March 29, 1994; revised October 13, 1994; accepted October 27, 1994.)

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5433–5447, MARCH 25, 1995

The circum-Chryse region as a possible example of a hydrologic cycle on Mars: Geologic observations and theoretical evaluation

Jeffrey M. Moore,[1] Gary D. Clow,[2] Wanda L. Davis,[1] Virginia C. Gulick,[3] David R. Janke,[4] Christopher P. McKay,[3] Carol R. Stoker,[3] Aaron P. Zent[1]

Abstract. The transection and superposition relationships among channels, chaos, surface materials units, and other features in the circum-Chryse region of Mars were used to evaluate relative age relationships and evolution of flood events. Channels and chaos in contact (with one another) were treated as single discrete flood-carved systems. Some outflow channel systems form networks and are inferred to have been created by multiple flood events. Within some outflow channel networks, several separate individual channel systems can be traced to a specific chaos which acted as flood-source area to that specific flood channel. Individual flood-carved systems were related to widespread materials units or other surface features that served as stratigraphic horizons. Chryse outflow channels are inferred to have formed over most of the perceivable history of Mars. Outflow channels are inferred to become younger with increasing proximity to the Chryse basin. In addition, outflow channels closer to the basin show a greater diversity in age. The relationship of subsequent outflow channel sources to the sources of earlier floods is inferred to disfavor episodic flooding due to the progressive tapping of a juvenile near-surface water supply. Instead, we propose the circum-Chryse region as a candidate site of past hydrological recycling. The discharge rates necessary to carve the circum-Chryse outflow channels would have inevitably formed temporary standing bodies of H_2O on the Martian surface where the flood-waters stagnated and pooled (the Chryse basin is topographically enclosed). These observations and inferences have led us to formulate and evaluate two hypotheses: (1) large amounts of the sublimated H_2O off the Chryse basin flood lakes precipitated (snowed) onto the flood-source highlands and this H_2O was incorporated into the near surface, recharging the H_2O sources, making possible subsequent deluges; and (2) ponded flood-water in Chryse basin drained back down an anti basinward dipping subsurface layer accessed along the southern edge of the lake, recharging the flood-source aquifers. H_2O not redeposited in the flood-source region was largely lost to the hydrologic cycle. This loss progressively lowered the vitality of the cycle, probably by now killing it. Our numerical evaluations indicate that of the two hypotheses we formulated, the groundwater seep cycle seems by far the more viable. Optimally, ~ 3/4 of the original mass of an ice-covered cylindrical lake (albedo 0.5, 1 km deep, 100-km radius, draining along its rim for one quarter of its circumference into substrata with a permeability of 3000 darcies) can be modeled to have moved underground (on timescales of the order of 10^3 years) before the competing mechanisms of sublimation and freeze down choked off further water removal. Once underground, this water can travel distances equal to the separation between Chryse basin and flood-source sites in geologically short (~10^6 year-scale) times. Conversely, we calculate that optimally only ~40% of the H_2O carried from Chryse can condense at the highlands, and most of the precipitate would either collect at the base of the highlands/lowlands scarp or sublimate at rates greater than it would accumulate over the flood-source sites. Further observations from forthcoming missions may permit the determination of which mechanisms may have operated to recycle the Chryse flood-waters.

[1] Center for Mars Exploration and the SETI Institute, NASA Ames Research Center, Moffett Field, California.

[2] Branch of Astrogeology, U.S. Geological Survey, Menlo Park, California.

[3] Space Science Division, NASA Ames Research Center, Moffett Field, California.

[4] Department of Geography, Arizona State University, Tempe.

Paper number 94JE02805.
0148-0227/95/94JE-02805$05.00

Introduction

There has been a renewed interest in the idea of Martian hydrologic cycles largely brought on by T. S. Parker, V. R. Baker, and their colleagues, who explored the possibility that Mars once had large bodies of standing water and large continental-style glacial ice sheets located in topographically confined lowlands [e.g., *Parker et al.*, 1989; *Baker et al.*, 1991; *Kargel and Strom*, 1992]. Other studies have pointed to evidence for large lacustrine basins on Mars which are postulated to be very young in the martian stratigraphic sequence [*Scott and Chapman*, 1991; *Scott et al.*, 1992].

These investigations imply that large amounts of water were cataclysmically released, late in Martian history, at least in part by mechanisms which restore lake-forming water to its source by some form of a hydrologic cycle. In a separate study focused on the sediments of ancient, sapping channel-fed lakes, *Goldspiel and Squyres* [1991] and *Goldspiel et al.* [1993] concluded that unless the upper several hundred meters of the Martian highlands were composed of very fine-grained noncohesive material, it would be necessary to recharge the associated up-channel aquifers to produce the amount of water needed to form the observed landforms.

The concept of a Martian hydrologic cycle (as envisioned for example by *Parker et al.* [1989] and *Baker et al.* [1991]) has been criticized [*Carr*, 1991] on the grounds that it is very difficult to understand the physical mechanisms that would repeatedly allow vast (ocean-sized) amounts of liquid water to suddenly appear on the martian surface and then just as suddenly disappear several million years later. It is the lack of recognized mechanisms (exclusive of radical degressions from the modern Martain environment) to drive a martian hydrologic cycle that makes many investigators inclined to question the possibility of such a cycle.

Over time a general consensus has developed that Mars has had large flood events in its history (see reviews in *Carr* [1981], *Baker* [1982], *Mars Channel Working Group* [1983]). Of the several channel types observed on Mars, those classified as outflow channels [*Sharp and Malin*, 1975; *Pieri*, 1976; *Mars Channel Working Group*, 1983] are considered the best evidence for catastrophic flooding. Outflow channels are characterized as generally large features that start full-born from localized sources (usually chaotic terrain) [*Sharp and Malin*, 1975]. They are widest and deepest at the head and decrease in size distally. Generally, they have few, if any, tributaries. The interiors and down-flow plains of some channels appear to be scoured. This assemblage of landforms bears a remarkable resemblance to catastrophic flood terrains on Earth [*Baker and Milton*, 1974]. On Mars, outflow channels are thought to have formed by cataclysmic release of groundwater [*Mars Channel Working Group*, 1983] perhaps from confined aquifers [*Carr*, 1979]. Once released, the water would have flowed down the regional slope, locally diverted by topographic barriers, carving the channels much in the way the sudden draining of Pleistocene Lake Missoula carved the Channeled Scabland in the American Northwest [*Baker and Milton*, 1974; *Mars Channel Working Group*, 1983]. Most outflow channels on Mars are located around and disgorge into the topographically enclosed Chryse Planitia basin.

In this paper we present some geologic observations for water cycling in the circum-Chryse region of Mars and evaluate various mechanisms which might play a role in a hydrologic cycle operating under modern Martian conditions. We also propose specific observational tests which can be conducted during forth-coming Mars missions which can verify the observations, inferences, and speculations produced by this investigation.

Observations

Determining the age of individual outflow channels is difficult. The limited area occupied by each channel precludes the presence of a statistically significant number of craters. However, *Carr and Clow* [1981] dated the areally extensive scoured surfaces which surround a few outflow channels to conclude that outflow channel events took place episodically throughout martian history. Another approach was taken by *Janke* [1985] to historically order the circum-Chryse channels. He used stratigraphic relationships among the channel systems and adjacent widespread materials units to determine relative ages of flood events.

Using the method employed by *Janke* [1985], the following approach was used in our work. Channels and chaotic terrain in contact with one another were treated as single discrete flood-carved systems. Individual flood-carved systems were related to widespread materials units that served as stratigraphic horizons. The two materials units used as stratigraphic horizons were the intercrater plains and the Chryse smooth-plains. The intercrater plains material is actually a collection of different materials deposited throughout the early history of Mars [e.g., *Carr*, 1981]. Collectively it can be characterized as the extensive smooth-to-undulatory surface that exists between widely spaced, ancient degraded craters >20 km across. Intercrater plains materials are placed in the oldest of the martian time-stratigraphic systems (the Noachian System) [*Scott and Carr*, 1978; *Scott and Tanaka*, 1986], which dates from the end of the period of heavy bombardment [*Barlow*, 1988]. *Tanaka* [1986] used crater statistics and various inferred cratering rates to date the end of the Noachian between 3.5 and 3.8 x 10^9 years. The second materials unit used as a stratigraphic horizon in this study was the Chryse smooth plains. The Chryse smooth plains are those materials which cover the floor of the Chryse Planitia. Like the intercrater plains materials, the Chryse smooth plains materials are subdivided by post-Viking mappers [e.g., *Scott and Tanaka*, 1986] and straddle the boundary between the two younger martian time-stratigraphic systems, the Hesperian and the Amazonian. *Tanaka* [1986] placed the timing of this boundary between 1.8 and 3.55 x 10^9 years. Circum-Chryse outflow channels are inferred to have formed over most of the perceivable history of Mars. The oldest channels antedate the intercrater plains and the youngest channels postdate the Chryse smooth plains.

Although determining the relative age of the youngest channels is clear, the unambiguous relative age of the older channels is more difficult. For example, although Margaritifer Vallis, one of the oldest channels, appears to have been exhumed from beneath intercrater plains material (Figure 1a). However, it is also possible that Margaritifer Vallis formed contemporaneously with the surrounding terrain (T.S. Parker, personal communication, 1994). Ravi Vallis (Figure 2b) has clearly formed in intercrater plains material. Though its lower reach appears to be buried by the Chryse smooth plains material, the possibility that the distal end of Ravi Vallis is truncated by Hydraotes Chaos cannot be ruled out. However, while detailed stratigraphic relations of the older channels to the intercrater plains and the Chryse smooth plains is not clear in some cases, what is clear is that some channels postdate and some channels predate the Chryse smooth plains. Tiu Vallis (Figure 1c) is an example of a younger channel that has eroded into the Chryse smooth plains and therefore postdates this surface.

Figure 2 shows the results of our relative age dating of outflow channels located in the Chryse drainage basin region of Mars. From these results, two trends become apparent from the placement of the circum- Chryse outflow channels into three time-stratigraphic groups. (1) The relative age of outflow channels tends to decrease towards the north and with

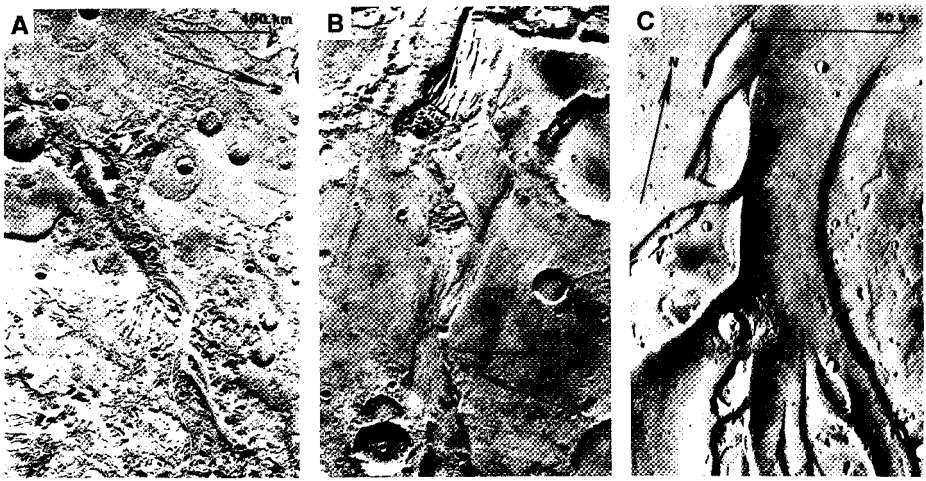


Figure 1. The transection and superposition relationships among channels, chaos, surface materials and other features were used to evaluate the relative ages of the flooding events in the circum-Chryse region. (a) Channel (Margaritifer Vallis) being exhumed from beneath, and thus predates, intercrater plains material, which is the older of the two stratigraphic horizons used in this study. (b) Channel (Ravi Vallis) eroded into intercrater plains material but whose distal end is buried under Chryse smooth-plains material. Occurring between the two stratigraphic horizons, it is identified as an intermediate-aged flood feature. (c) Channel (Tiu Vallis) that has eroded the Chryse smooth-plains material and this thus classified among the youngest flood features.

decreasing elevation. (2) Outflow channels in the north tend to show a greater diversity in age. These trends do not change even if the outflow channels are divided into two groups: those that predate and those that postdate the formation of the Chryse smooth plains.

Another observation from this study is that several outflow channels contain multiple flood source sites within their boundaries. Although multiple episodes of flooding within the outflow channels have been noted by previous studies [e.g., *Chapman and Scott*, 1989], it is not possible to determine unambiguously whether these episodes represent several discrete cataclysmic flooding events within a channel separated by significant geological periods or whether these episodes reflect the waning stages of a single flood event and the progressive dewatering of a groundwater reservoir. If these episodes are separated by geologically significant periods, then recharge of aquifers is probably necessary for subsequent flood events. An example will illustrate this point. Ravi Vallis/Aromatum Chaos system is an outflow channel system that may have experienced multiple episodes of fluid flow (Figure 3), because it contains three regions of chaotic terrain within its boundaries. Chaotic terrain forms where groundwater is cataclysmically released from the subsurface. In the first episode, fluid from Aromatum Chaos carved the main channel of Ravi Vallis. In succeeding events, fluid derived from two additional chaos sites reoccupied portions of the inner channels and modified previously collapsed areas. The areas of chaotic terrain associated with subsequent events are smaller. The first preserved event that carved Ravi Vallis was much larger than any succeeding event within the channel system. This relationship is consistent with all other well-preserved channel systems. If groundwater needed to form later floods was not present during the oldest observed flood events, then recharging of groundwater reservoirs was needed for the younger floods. The older and larger floods may have wiped out evidence for any previous smaller flood events. However, it is not inconceivable that multiple and subsequently smaller flood source sites within an outflow channel represent progressive tapping of a large groundwater reservoir or series of aquifers located at different elevations. If this was indeed the case, then a hydrological cycle may not have been required for each flood-source site within a given channel.

A final significant observation derived from image analysis is that the morphology of the circum-Chryse region is inconsistent with rainfall as a significant contributor of water to surface modification. Rainfall and its runoff act to intensely dissect the landscape into dense dendritic drainage networks. Nothing of this sort is observed in the circum-Chryse region. However, water transported to the surface of the study area in the form of snow cannot be ruled out.

Two major inferences are formulated from the observations. Water may have been repeatedly resupplied to the near-surface reservoirs, as evidenced by the extended history of channeling in the region where subsequent channels tend to start interior to the sites of earlier flood-source sites. The discharge rates necessary to have carved the circum-Chryse outflow channels, which have been estimated on occasion to exceed 10^6 m^3 s^{-1},

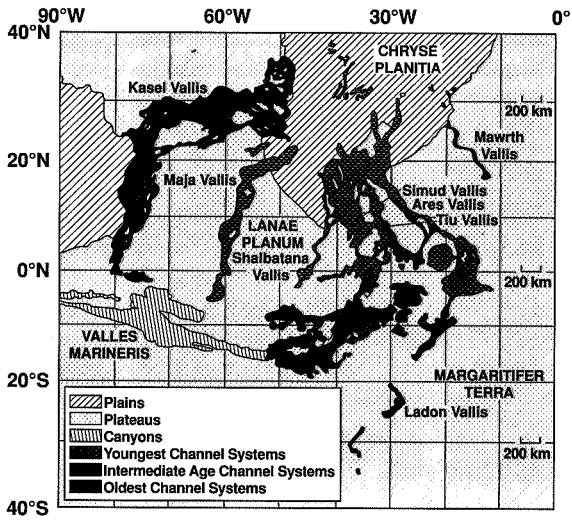


Figure 2a. A sketch map of the circum-Chryse region illustrating the relative age groups assigned the outflow channels within this area.

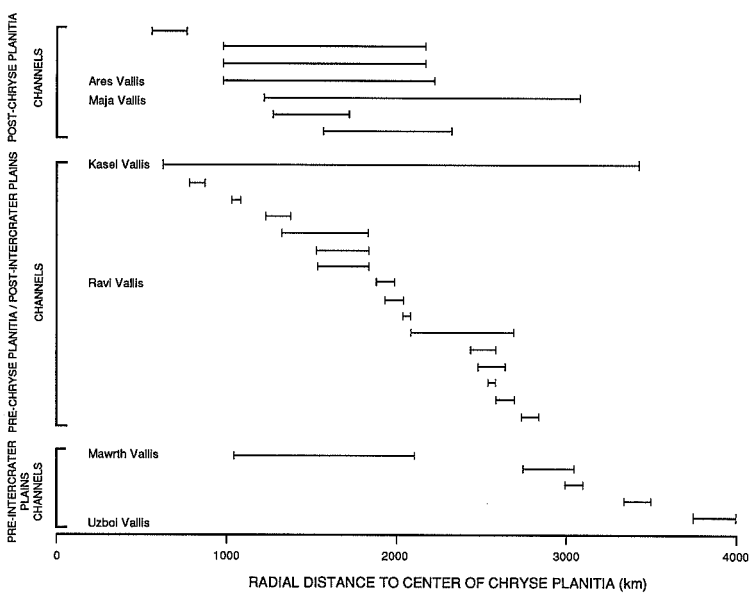


Figure 2b. Plot of outflow channels as a function of age group along the vertical axis. Along the horizontal axis is plotted the source and terminus of each channel as a function of radial distance from the center of Chryse Planitia. This plot illustrates that outflow channels become younger with increasing proximity to the Chryse basin and that outflow channels closer to Chryse show a greater diversity in age.

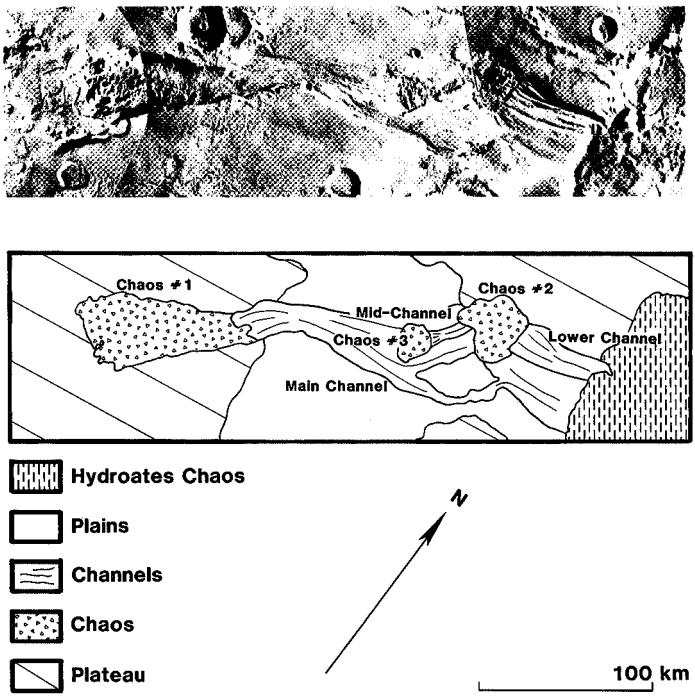


Figure 3. A major observation of this study is that the relationship of subsequent outflow channel sources (chaos) to the sources of earlier floods is inconsistent with the idea that flooding is due to the progressive tapping of a juvenile water supply. In this figure, the channel (Ravi Vallis) can be observed to have experienced at least three episodes of fluid flow. The oldest flood source (Chaos 1) is Aromatum Chaos. Note the presence of the younger flood-source chaos on the floor of an older flood channel. Chaos 3 is younger, and Chaos 2 is youngest. That these younger flood sources were untapped during the formation of the oldest flood channel implies that the water for the younger floods was not present at the time of this oldest flood.

and upwards to 10^9 m^3 s^{-1} [e.g., *Baker*, 1982; *Baker et al.*, 1991], would have formed temporary standing bodies of water where floodwaters stagnated and pooled. The site of this pooling would have been Chryse basin, which is topographically enclosed and lowest along its boundary with the highlands to the south [*U.S. Geological Survey*, 1989]. Such ponding in Chryse and elsewhere on the northern plains (and the consequences of such ponding) has been discussed by *Lucchitta et al.* [1986], *Parker et al.* [1989], *Rotto and Tanaka* [1991], and *Baker et al.* [1992]. The possible fates of this water is discussed in the next section.

Hypotheses

These observations and inferences have led us to propose the following two alternative hypotheses, which will be referred to as the airborne precipitation cycling hypothesis, and the groundwater seep cycling hypothesis. Throughout this study we conservatively assume that the past Martian environment was similar to that of the present. Both hypotheses require some mechanism to mobilize the H_2O stored in the flood-source region aquifer, making possible its discharge as a new Chryse-bound flood. A good candidate mechanism is regional geothermal heat associated with the episodic suffusion of magma [*Marsursky et al.*, 1977; *Baker*, 1982]. *Baker et al.*[1991] proposed that catastrophic flooding was triggered by massive magmatic emplacement events, such as those associated with Tharsis volcanism. These events would have produced enormous hydrothermal systems capable of drawing in water from deep and adjacent supplies toward the thermal anomalies. *Baker et al.* [1991] explained that sudden cataclysmic outflows of groundwater were driven by hydrothermal convection, and likely assisted by rapid exsolution of CO_2. In our study, we hypothesize that H_2O not

redeposited in the flood-source region was largely lost to the hydrologic cycle. This loss progressively lowered the vitality of the cycle, perhaps by now killing it.

The airborne precipitation cycling hypothesis postulates that large amounts of the H_2O that sublimated off of the temporary postflood lakes within the Chryse basin precipitated (snowed) onto the flood-source highlands. This scenario differs from the canonical view that the H_2O went without pause to the polar sinks. In the airborne precipitation cycling hypothesis, strong northerly winds (which are inferred from modern wind streak orientation) transported the lake-derived H_2O back up on the highlands to the south. The >2 km relief of the flood-source highlands forced orographic cooling and precipitation in the air mass as it was carried southward. The precipitated H_2O was incorporated into the subsurface (perhaps by geothermal processes), recharging the H_2O sources, making possible subsequent deluges. The high albedo and low thermal inertia of the "snow" relative to the lake ice would allow the "snow" to persist on the highlands much longer than the ice on the lake, thus providing the time needed to incorporate it into the subsurface. The decay of the heat pulse thought to have brought on the cycle-initiating flood progressively lowered the isotherms and moved the H_2O down from the snowpack to the subsurface reservoir. Subsequent pulses of anomalous geothermal activity would initiate a new episode of flooding.

Alternatively, the groundwater seep cycling hypothesis holds that liquid lake water, under its ice cover, drained into a permeable subsurface layer. Concurrent erosion of the highlands/lowlands boundary along the southern periphery of the lake by the floodwaters associated with a given flood could have exposed the 0°C isotherm, which may be only ~1 km from the equilibrium surface interface [e.g., *Rossbacher and Judson*, 1981]. In order for water to be transported toward the flood-source sites, a permeable layer(s) must dip away from the basin so that the underground flood-source reservoirs are downgradient from the Chryse lake sites (Figure 4). *Schultz et al.*, [1982] proposed that Chryse basin was formed by a very large impact. Such an impact would have reoriented preexisting horizontal layers so that they would dip away from the impact site for many hundreds of kilometers beyond the basin rim. The dip of the reoriented stratigraphic layers would be steepest near the rim and become increasingly gentle with distance. Subsequent surficial erosion and modification of the circum-Chryse region would have created a regional topographic slope toward the basin while underground layers remained oppositely oriented. A possible indication of an antibasinward dip to the bedding layers composing the circum-Chryse highlands may be the unusual reverse (though very low) gradients of some of the outflow channels (e.g., Simud and Tiu Valles), reported by *Lucchitta and Ferguson* [1983].

The flow of groundwater can be confined to natural conduits and may travel large distances without much dispersion lateral to the direction of flow [e.g., *White*, 1969]. As long as the water can travel in a permeable zone that lies beneath the 0°C isotherm and is sealed against vapor diffusion loss through the material overlying the permeable (phreatic) zone, then this water will move back under the highlands until either the slope of the subsurface layer flattens or the water flow is interrupted and ponded by a subsurface obstruction, such as a fault plane. Large faults coincide with several flood-source chaotic terrains and probably played a role in the storage of significant quantities of groundwater. These faults have traces that are circumferential about Chryse and are probably the manifestation of basin tectonics [*Schultz et al.*, 1982], and thus reveal the extent that the subsurface has been disturbed by the formation of Chryse basin. Episodic geothermal heat raised the water from depth up to the surface, where it would burst out to do new flooding, perhaps in the manner suggested by *Marsursky et al.* [1977]. Simply stated, in this hypothesis, ponded floodwater in Chryse basin drained down an antibasinward dipping subsurface layer, ultimately reaching and recharging the flood-source aquifer.

Discussion

We will now discuss the evaluations conducted to determine which of these hypotheses can better explain the apparent recycling of floodwater in the circum-Chryse region. If the airborne precipitation cycling hypothesis is viable, then it may be generally applicable to other portions of Mars and serve as an explanation for other examples of water recycling

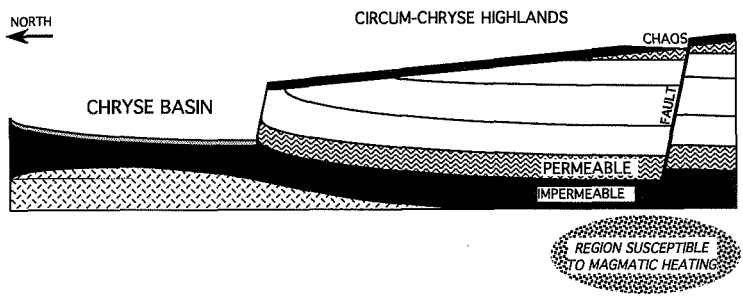


Figure 4. Idealized and simplified hypothetical cross section of the Chryse region illustrating the concepts described in the groundwater seep hypothesis. The dip of strata underlying the circum-Chryse highlands away from the basin is a structural imprint of the impact that created Chryse basin. The schematically shown fault is a product of basin ring tectonics. Note that the modern surface of circum-Chryse highlands slopes toward the basin as a consequence of erosion. Considerable vertical exaggeration is shown.

[e.g., *Goldspiel and Squyres*, 1991; *Goldspiel et al.*, 1993]. Alternatively, if the groundwater seep cycling hypothesis is more viable, then perhaps it is the unusual structural (subsurface) geology of the region that makes it one of the few (if not only) sites where water can be recycled on Mars on a large regional scale. This may explain why this region contains most of the outflow channels on the planet. These evaluations are grouped into three categories: those which are applicable to the airborne precipitation cycling hypothesis; those which are particular to the groundwater seep cycling hypothesis; and those phenomena and issues that are germane to both hypotheses. We will begin with the common issues first.

Lake Lifetime and Evaporative Loss Rate

Evaporation Rates

The water vapor flux E_o from the lake's surface to the atmosphere is governed by the water-vapor density gradient ($\partial\rho_w/\partial z$) within the atmospheric boundary layer (ABL),

$$E_o = -K_w \frac{\partial \rho_w}{\partial z}; \tag{1}$$

K_w is the eddy diffusivity for water vapor. Integrating across the "interfacial sublayer" to a point z within the "surface sublayer" (see e.g., *Brutsaert* [1982] for a description of the various sublayers within the ABL), (1) can be re-expressed in the following form:

$$E_o = -\frac{u_*[\rho_w(z) - \rho_w(0)]}{F_w(0,z)} \tag{2}$$

where $u*$ is the "friction velocity" and

$$F_w(0,z) = \int_0^z \frac{u_*}{K_w} dz. \tag{3}$$

The resistance function F_w embodies the dependence of the mass transfer rate on molecular gas properties and atmospheric stability. F_w can be further broken into two components

$$F_w(0, z) = F_w(0, z) + F_w(z_l, z), \tag{4}$$

corresponding to separate integrations across the interfacial and surface sublayers (first and second terms on the right-hand side, respectively); z_l is the height of the interfacial sublayer.

Within the interfacial sublayer, heat and water vapor transfer involve nonsteady molecular diffusion into Kolmogorov-scale eddies that penetrate down from the fully turbulent surface sublayer. Turbulent airflow within this zone can take one of two forms, "aerodynamically smooth" or "aerodynamically rough," depending on the nature of the surface and the viscous length scale ($\upsilon/u*$), where υ is the kinematic viscosity of the CO_2/H_2O gas mixture. For aerodynamically smooth flow, the height of the interfacial sublayer is

$$z_l = 30\left(\frac{\nu}{u_*}\right), \tag{5}$$

and it is

$$z_l = d_o + 6.5 z_o, \tag{6}$$

for aerodynamically rough flow, where z_o is known as the "surface roughness length" and d_o is the zero-plane displacement, used to shift the coordinate system to the base of the roughness elements on the surface. *Brutsaert* [1975] effectively determined the function $F_w(z_l, z)$ for both flow regimes. In the present notation, he found

$$F_w(0, z_l) = \frac{1}{C_s} S_c^{2/3} \tag{7}$$

for aerodynamically smooth flow and

$$F_w(0, z_l) = \frac{1}{C_r} S_c^{1/2} \left(z_{o+}\right)^{1/4} \tag{8}$$

for aerodynamically rough flow; $S_c = (\upsilon/\kappa_w)$ is the dimensionless Schmidt number, κ_w is the molecular diffusivity for water molecules through a CO_2 gas, $z_{o+} = (u* z_o/\upsilon)$ is the "roughness Reynold's number," $C_r = 1/7.3$, and $C_s = 1/13.6$.

Within the surface sublayer, the eddy diffusivity can be described in terms of the dimensionless water-vapor gradient $\phi_w(\zeta)$,

$$K_w = \frac{k(z - d_o)}{\phi_w(\zeta)}, \tag{9}$$

where k is von Karman's constant and ζ is the Monin-Obukhov stability parameter. Because heat and water vapor are transferred through the surface sublayer by the same eddy processes, $\phi_w(\zeta)$ is generally assumed to be equal to the dimensionless temperature gradient $\phi_w(\zeta)$ for this layer. Several parameterizations exist for the $\phi_w(\zeta)$, and hence for $\phi_w(\zeta)$. Given K_w, the resistance function $F_w(z_l, z)$ can be determined.

We have evaluated the water vapor flux from a large Martian lake assuming a 7-mbar CO_2 atmosphere and temperatures in the range 170 to 240 K. Over this temperature range, the Schmidt number $S_c \approx 0.48$-0.49 [*Clow and Haberle*, 1990a] and the kinematic viscosity υ varies from 4.0×10^{-4} to 8.0×10^{-4} $m^2\ s^{-1}$. The air mass blowing onto the lake is assumed to be perfectly dry (relative humidity = 0%). For the dimensionless gradients $\phi_h(\zeta)$ and $\phi_w(\zeta)$, we utilize the *Businger et al.* [1971] parameterization. Because solar radiation is absorbed at depth within the ice, diurnal surface temperature variations are expected to be only a few degrees. Hence, we assume the atmospheric boundary layer is neutrally stable ($\zeta = 0$) above the lake (this assumption would be invalid for a small lake whose dimension is comparable to the thickness of the ABL). The friction velocity was specified using the relation,

$$u_* = \frac{U(z)}{F_m(z_{om}, z)}, \tag{10}$$

where $U(z)$ is the wind speed and $F_m(z_{om}, z)$ is the momentum resistance function, analogous to $F_w(0, z)$. Since we have assumed neutral atmospheric stability, F_m reduces simply to

$$F_m(z_{om}, z) = \frac{1}{k}\left[\ln\left(\frac{z - d_o}{z_{om}}\right)\right], \tag{11}$$

where the parameter z_{om} is the momentum roughness length, equal to $0.135\ \upsilon/u*$ for aerodynamically smooth flow and to the surface roughness length, z_o, for aerodynamically rough flow. The surface roughness length z_o for a Martian lake is obviously unknown. We consider two values, 0.03 cm and 1 cm. Measurements over polar snow surfaces, which typically include sastrugi (wind-eroded ridges < 5 cm high on snow surfaces), yield roughness lengths between 0.003 and 0.1 cm [*Inoue*, 1989]. The 0.03-cm value is taken to be representative of these surfaces, while the 1-cm value is used

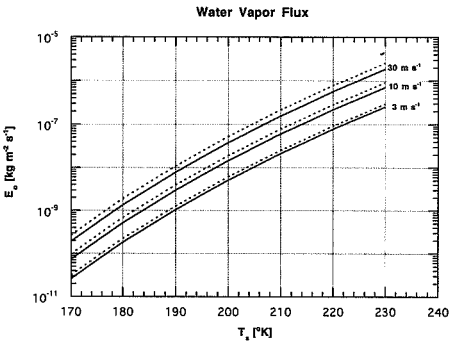


Figure 5. Water vapor flux as a function of surface temperature T_s and the wind speed at height of 1000 m above the surface. The solid lines are the results for a smooth surface (z_o = 0.03 cm) and the dashed lines are for a relatively rough surface (z_o = 1 cm).

to simulate a fairly rough surface. For the wind speeds and pressures considered in this paper, airflow within the interfacial sublayer is found to be aerodynamically smooth whenever z_o = 0.03 cm and aerodynamically rough when z_o = 1 cm. Free convection [*Clow and Haberle*, 1990b] is not found to occur even under the calmest conditions discussed in this paper.

Calculated water vapor fluxes are shown in Figure 5 as a function of surface temperature T_s for three sample wind speeds specified at a height 1000 m above the lake's surface. As expected, the mass flux E_o is very sensitive to the surface temperature and moderately sensitive to the wind speed. The sensitivity to the surface roughness length z_o is small by comparison. Thus, it is very important to establish the range of surface temperatures to properly evaluate E_o for any particular simulation. We regard wind speeds (100 m above the surface) of the order of 10 m s^{-1} as being the most likely for the Chryse region, although western boundary currents during the southern winter solstice could lead to winds as strong as 30 m s^{-1} [*Joshi et al.*, 1994].

Lake Lifetime

Once Lake Chryse formed, its lifetime would have depended on its depth and rate of water loss via sublimation from the surface and possibly from groundwater seepage (discussed in a later section). The maximum depth of the lake was limited to ~2000 m by the geometry of the Chryse basin, and the most likely initial depth was considerably less than this. As we have seen, the sublimation rate is very sensitive to the surface temperature of the ice, which in turn depends on its albedo.

Using present orbital parameters and solar luminosity, minimum-seasonal, mean-annual, and maximum-seasonal surface temperatures were calculated for ice at the latitude of Chryse (~25°N) by solving the surface energy balance,

$$\sigma T_s^4 = IR^{\downarrow} + \kappa \frac{\partial T}{\partial z} \tag{12}$$

where σ is the Stefan-Boltzman constant and κ is the thermal conductivity of ice; the sensible and latent heat fluxes over the lake would have been small in comparison to the three terms in (12) and hence have were ignored in the calculation of T_s. At equilibrium, the conductive heat flux in the ice immediately beneath the surface (z = 0) is equivalent to the absorbed solar energy,

$$\kappa \frac{\partial T}{\partial z} = (1 - A) S_o \tag{13}$$

where A is the albedo and S_o is the incident solar flux. Equilibrium can be assumed to hold for the mean-annual calculations and is approximately correct at the minimum and maximum points in the seasonal cycle. We have also assumed an atmospheric optical depth of 0.4 and have taken the downwelling infrared flux values, $IR^{\downarrow}$, from *Haberle and Jakosky* [1991].

Although the albedo of the lake ice is unknown, the mostly likely range is 0.3-0.6. The integrated albedo of glacial ice and thick multiyear sea ice is rarely higher than 0.6 unless covered with a fresh blanket of snow. For multiyear lake ice in Antarctica, *McKay et al.* [1994] measured albedo values of 0.3-0.4 before the seasonal melt period began. For the 0.3-0.6 albedo range, mean-annual surface temperatures are expected to be in the range 200-220 K (Table 1). Seasonal temperature variations are expected to be less than ±10 K. Due to the nonlinearity of the water vapor flux with temperature, the use of the seasonal temperature cycle to calculate the annual water vapor flux will lead to an enhancement of 16-23% over that which would be determined using the mean-annual temperature.

Figure 6 shows the expected lifetime of a lake initially 1 km deep in Chryse basin as a function of albedo and wind speed. These results assume water is lost from the lake solely by sublimation and that no additional water is input to the lake during this time. The lifetime of a lake with a different initial depth can be found by simply scaling the values in Figure 6.

Table 1. Minimum Seasonal, Mean-Annual, and Maximum Seasonal Surface Temperatures (°K) for Ice Using Present Orbital Parameters and Solar Luminosity

	Latitude 25°N			Latitude 0°N		
Albedo	Minimum	Mean	Maximum	Minimum	Mean	Maximum
0.3	210.6	**220.6**	227.3	215.3	**225.4**	232.5
0.4	205.1	**214.4**	220.7	209.5	**218.9**	225.6
0.5	199.2	**207.7**	213.5	203.2	**211.8**	218.0
0.6	192.7	**200.3**	205.4	196.3	**204.0**	209.6
0.7	185.3	**191.9**	196.3	188.5	**195.1**	199.9
0.8	177.4	**182.3**	185.8	179.7	**184.8**	188.6

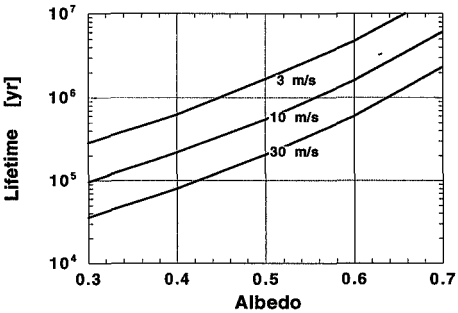


Figure 6. Lifetime of a lake initially 1 km deep in Chryse basin for wind speeds of 3, 10, and 30 m s^{-1}.

For the most likely range of albedo (0.3-0.6) and wind speed (~10 m s^{-1}), a 100-m-deep lake is expected to last 10-160 kyr before completely subliming away, while a 1000-m-deep lake would last 96-1610 kyr. We note that if the lake formed early in Martian history when the solar luminosity was perhaps only 70% of present, surface temperatures would have been roughly 10 K less (Table 2). The corresponding water vapor fluxes would have been reduced by a factor of 5, extending the lifetime of the lake by the same factor.

It is interesting to estimate how long liquid water might have existed in such a lake. Once a perennially ice-covered lake achieves equilibrium, its ice thickness is given by

$$z_{eq} = \frac{b\,\ln(T_m/T_s) + c(T_s - T_m) - h\left[1 - \exp\left(\frac{-z_{eq}}{h}\right)\right](1-A)\,S_o}{F_l + F_g} \qquad (14)$$

where T_m is the melting point of water, h is the extinction pathlength for solar radiation in ice, $F_l = \rho\, v\, L_f$ is the latent heat flux due to freezing at the base of the ice cover, ρ is the ice density, v is the rate of ice formation, L_f is the latent heat of fusion, F_g is the geothermal heat flux, and the constants b and c are 780 W m^{-1} and 0.615 W m^{-1} K, respectively [*McKay et al.*, 1985]. Once equilibrium is reached, the rate of ice formation at the base of the ice cover equals the mass loss from the surface due to sublimation, $v = E_o/\rho$. Table 3 lists the equilibrium ice thicknesses for a lake in Chryse basin again utilizing current orbital parameters and solar luminosity. The extinction path length h was assumed to be 1 m, similar to Antarctic lakes, and the geothermal flux was assumed to be 30 mW m^{-2}. The calculated equilibrium ice thicknesses are listed in Table 3. For most combinations of albedo and wind speed, Z_{eq} exceeds the depth of the basin, indicating the lake could not equilibrate even if it filled the basin and water somehow recharged the lake at a rate v. Even for the low-albedo cases where Z_{eq} is less than the depth of the basin, the large values in Table 3 indicate the lake would have to be hundreds of meters deep in order to stably exist even if the lake was somehow recharged from below. Without a recharge mechanism, equilibrium obviously can never be achieved and the lake eventually freezes solid and subsequently sublimates away.

How long will liquid water exist in the lake once it initially forms? This can be addressed using the solution to the well-known Stefan problem as long as the freezing rate v is much greater than $F_g/(\rho L_f)$. If this condition is satisfied, the freezing rate is

$$v = \lambda\sqrt{\frac{\kappa}{t}} \qquad (15)$$

[*Turcotte and Shubert*, 1982] where κ is the thermal diffusivity of ice, t is time, and λ is a constant that is determined by the condition

$$\frac{\exp(-\lambda^2)}{\lambda\,\mathrm{erf}(\lambda)} = \frac{L_f\sqrt{\pi}}{C_p(T_m - T_o)}. \qquad (16)$$

L_f is the latent heat of fusion and C_p is the specific heat. T_o is near-surface temperature taken at a depth below that at which the absorption of solar radiation affects the temperature profile (nominally 5 m). The total amount of water that has been converted to ice at the base of the ice cover is shown in Figure 7 as a function of time. For a lake initially 100 m deep, the complete freeze-up occurs on the order of 1 kyr from the time of formation, and the lake completely sublimes away in 10-160 kyr. For a 1-km-deep lake, complete freeze-up occurs in roughly 100 kyr, while the total lifetime is limited to 100-1600 kyr. For ease of comparison, the complete freeze-up times and total lifetimes are listed in Tables 4 and 5 for initial water depths of 100, 200, 400, and 800 m.

The Airborne Precipitation Cycling Hypothesis

In this section we consider how moisture from a large body of ice-covered water could provide a source of meteoric water for an accumulation region at high elevation. The system we

Table 2. Minimum Seasonal, Mean-Annual, and Maximum Seasonal Surface Temperatures (°K) for Ice at 70% Present Solar Luminosity

	Latitude 25°N			Latitude 0°N		
Albedo	Minimum	Mean	Maximum	Minimum	Mean	Maximum
0.3	198.6	**207.1**	212.7	202.6	**211.1**	217.2
0.4	194.1	**201.8**	207.1	197.8	**205.6**	211.3
0.5	189.2	**196.2**	201.0	192.5	**199.7**	204.9
0.6	184.0	**190.1**	194.3	186.9	**193.2**	197.8
0.7	178.2	**183.3**	186.9	180.6	**185.9**	189.8
0.8	171.9	**175.7**	178.4	173.7	**177.6**	180.6

Table 3. Equilibrium Ice Thickness (Meters)

	Wind Speed at 1000-m Altitude, m s^{-1}		
Albedo	3	10	30
0.3	510	245	105
0.4	1430	855	410
0.5	2800	2060	1250
0.6	4340	3830	2930
0.7	5960	5750	5280

are considering is shown schematically in Figure 8. A large body of ice-covered water is assumed to exist at the lower elevation, taken to be $z = 0$. Air moving over this water reaches 100% humidity at the temperature of the lake surface. We assume that the air initially contains negligible water and that the flow is sufficient to carry all the water that sublimates from the lake surface into the atmosphere, i.e., the flux of water is determined solely by sublimation from the lake. The moistened air then moves upslope by some height $z = Z$ and cools by adiabatic expansion. Only when the air parcel reaches the upper layer does the moisture in excess of the saturation at the lower temperature condense onto the surface. This condensation represents the source of H_2O in the accumulation region. We also consider sublimation in the upslope area as a loss of H_2O from the accumulation zone.

In this model we assume that winds with an initial relative humidity of 0% move across a frozen lake at altitude $z = 0$. Water ice sublimates, and a moist parcel with a relative humidity of 100% forms. It is assumed that this parcel of mass ascends a height Z. With the ascent, Δz, the temperature decreases which causes a fraction (F) of the parcel to condense. Note that for the purpose of this evaluation, we assume that all of this fraction of condensate is deposited as snow in the highlands at height Z. In addition to the evaporation from the lake, we assume evaporation from the snow field in the highlands at altitude Z. The rate (R) of snow accumulation in the highland snow field is a function of (1) the rate of production of the moist parcel of mass, (2) the amount of loss of this mass due to the ascent of the parcel, and (3) the rate of sublimation from the snow field. The key controlling variable in our model is the rate of sublimation from the ice-covered surface of the lake. The evaporation rate is computed as a function of wind speed and temperature for a 7-mbar atmosphere of CO_2 using the method of *Clow et al.* [1988]. In these calculations we have assumed that the planetary boundary layer is isothermal and that the air is completely dry at $z = 100$ m. It can be seen from these curves that the sublimation rate is approximately exponential with temperature and thus dominated by the high-temperature period:

$$E \cong E_o \exp\left(\frac{T}{a}\right) \tag{17}$$

where a = slope^{-1} for respective wind speeds. We chose $a \cong$ 6.56 K, reflecting a nominal wind speed of 10 m s^{-1} for our initial calculation.

The global mean-annual temperature is used to characterize the moisture content of the parcel, assuming a simple sinusoidal variation in temperature during a Martian year (temperatures were derived from the model of *Kieffer et al.* [1977]). There is an enhanced period of evaporation due to peak temperatures greater than the mean-annual temperature. For simplification, we treat the temperature T as equal to the mean-annual temperature plus a sinusoidal variation about the mean that has an amplitude of ΔT. It can be shown that for a sinusoidal temperature variation for which the amplitude $\Delta T \leq 5a$, the average sublimation over the period can be approximated by

$$\frac{1}{P}\int_0^P E\,dt = E(T_o)\frac{1}{P}\int_0^P \exp\left(\frac{\Delta T \sin\left(\frac{2\pi t}{P}\right)}{a}\right)dt \simeq E(T_o)\left[1+\frac{1}{5}\left(\exp\left(\frac{\Delta T}{a}\right)-1\right)\right] \tag{18}$$

Where P is the period, representing 1 year. The term $[1 + {}^1/_5 (e^{\Delta T/a} - 1)]$ represents the enhancement over the mean sublimation due to the nonlinear nature of the exponential function.

We now consider the trajectory of a parcel of air beginning at the surface of the lake and moving up to the accumulation zone. At the lake surface the amount of water in the parcel, w_i, is determined by the relative humidity, RH_i, times the saturation vapor pressure of water vapor, P_{sat}. Over the temperature range considered here (190-260 K), we can approximate $P_{sat} \cong P_o e^{-T/\beta}$, where $\beta = 7.215$ K, and $P_o = 3.8 \times 10^{-16}$ mbar [*Marti and Mauersberger*, 1993]. Similarly, the amount of water in the parcel once it has reached the accumulation zone, w_f, is determined by the relative humidity, RH_f, times the saturation vapor pressure and the temperature there.

The fraction of water condensing over the highlands is given by $F = (w_i - w_f)/w_i$, assuming that the relative humidity above the lake and the highlands are the same. Therefore,

$$F \simeq \frac{RH_i\left[P_o \exp\left(\frac{T_s}{\beta}\right)\right] - RH_f\left[P_o \exp\left(\frac{T_s+\Gamma\Delta z}{\beta}\right)\right]}{RH_i\left[P_o \exp\left(\frac{T_s}{\beta}\right)\right]} = 1-\exp\left(\frac{\Gamma\Delta z}{\beta}\right) = 1-\exp\left(\frac{\Delta T_z}{\beta}\right) \tag{19}$$

where the lapse rate, $\Gamma = \Delta T_z / \Delta z$, and the change in temperature from the upper elevation to the lower elevation is given by $\Delta T_z = T_z - T_s$. To a good approximation, the amount of vapor that condenses during a nominal ascent is not

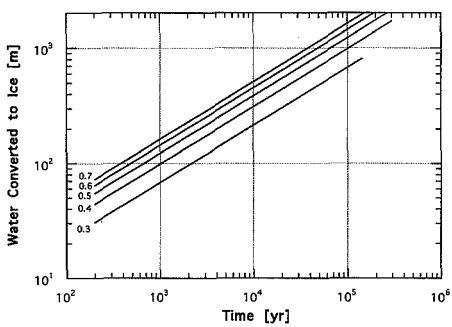


Figure 7. Total amount of liquid water converted to ice by freezing at the base of the ice cover for various albedo values between 0.3 and 0.7.

Table 4. Times for Complete Freeze-Up (Kiloyears)

	Initial Water Depth, m			
Albedo	100	200	400	800
0.3	2.2	8.6	34.7	-
0.4	1.0	4.1	16.3	66.0
0.5	0.67	2.6	10.7	42.9
0.6	0.48	1.9	7.7	30.9

sufficient to alter the moist adiabat very much from the dry adiabat. Thus

$$T_z - T_g - \frac{g}{C_p}\Delta z = T_g + \Gamma \Delta z = T_g - \Delta T_z \qquad (20)$$

where T_z is the temperature of the parcel at the scarp top (i.e., where it snows), T_g is the temperature at the lake surface, $\Gamma = -g/C_p$, and Δz represents the height of the ascent. On Mars, $g \cong 373$ cm s^{-2}, C_p for $CO_2 = 8.3 \times 10^6$ cm^2 s^{-2} K^{-1}, implying that $-g/C_p = \Gamma = -4.45 \times 10^{-5}$ K cm^{-1} $= -4.45$ K km^{-1}. Therefore, $\Delta T_z = \Gamma \Delta z = -8.9$ K.

As previously mentioned, the sublimation from the lake that takes place at mean-annual temperatures, $E(T_o)$, is enhanced due to peak temperatures, ΔT_{sp}, greater than mean-annual temperature during the summer time. Thus,

$$\text{Production} = F \times E(T_o)\left[1 + \frac{1}{5}\left(\exp\left(\frac{\Delta T_{sp}}{a}\right) - 1\right)\right] \qquad (21)$$

Substituting for F gives

$$\text{Production} = \left[1 \pm \exp\left(\frac{\Delta T_z}{\beta}\right)\right] E(T_o)\left[1 + \frac{1}{5}\left(\exp\left(\frac{\Delta T_{sp}}{a}\right) \pm 1\right)\right] \qquad (22)$$

The areal loss rate of sublimation from the snow accumulation zone is less than the loss rate from the lake surface because of the decrease in temperature with the increase in altitude. Thus,

$$\text{Loss} = \pm E(T_o) \exp\left(\frac{\Delta T_z}{a}\right)\left[1 + \frac{1}{5}\left(\exp\left(\frac{\Delta T_{sp}}{a}\right) \pm 1\right)\right] \qquad (23)$$

where ΔT_{sp} represents the increase in temperature due to peak temperatures at the snowfield during the summer. The net accumulation of snow in the highland snowfield is the difference between the production rate and the loss rate.

We now apply this theoretical framework to a specific case. We assume values that characterize the Chryse basin region of present Mars. For this region, the current annual average temperature is 208 K and the peak summertime temperature is 218 K. The scarp top is 2 km higher than the lake surface. We calculate that the fraction (F) of H_2O that condenses over the highlands from the original amount in the 100% RH parcel just above the lake in Chryse is 0.39. We compared the area of the lake to the "equilibrium" area of the highlands snowfield. The snowfield's "equilibrium" area is defined by the area the snow would have to be spread so that its subsequent sublimation rate would exactly match its precipitation rate. The calculated ratio of "equilibrium" highlands snowfield area to Chryse lake area is 3.7, a value which is independent of temperature.

The Groundwater Seep Cycling Hypothesis

Several phenomena are integral to the groundwater seep cycling hypothesis. For this mechanism to work, there must be a permeable layer(s) along which water can travel back toward the flood-source sites. Assuming that antibasinward dipping subsurface layers exist beneath the Chryse highlands region, we evaluate the ability of water, beneath an ice-covered flood lake in Chryse Planitia, to seep into these layers.

To make this evaluation, we constructed a simple model (Figure 9). We began with a lake containing 2×10^{12} m^3 liquid water beneath its ice cover. We assumed the liquid portion of the lake is shaped like a circular disk with vertical sides, 200 km in diameter and initially 200 m deep. The lake is allowed to drain through a time-varying cross-sectional area $A(t)$ equal to one-quarter of its circumference (31×10^6 m) times the depth of the lake d, which varies with time (Figure 9).

$$A(t) = \frac{2\pi r}{4} d(t) \qquad (24)$$

We assume that the source region for the floodwaters is located 1000 km away from the perimeter of the lake and lies 1 km deeper than the lake's base. The time required to drain the lake completely was calculated for several values of average subsurface permeabilities. Megaregolith, which is thought to comprise much of the subsurface of Mars, has a permeability estimated by *Carr* [1979] of approximately 3000 darcies. In comparison, terrestrial basalts typically have permeabilities ranging from 10^{-1} to 10^3 darcies [*Davis and DeWiest*, 1966].

Table 5. Lake Lifetimes (kiloyears) for $U = 10$ m s^{-1}

	Initial Water Depth, m			
Albedo	100	200	400	800
0.3	9.5	19.1	38.2	76.4
0.4	21.8	43.7	87.4	175
0.5	54.4	109	218	435
0.6	161	322	643	1290

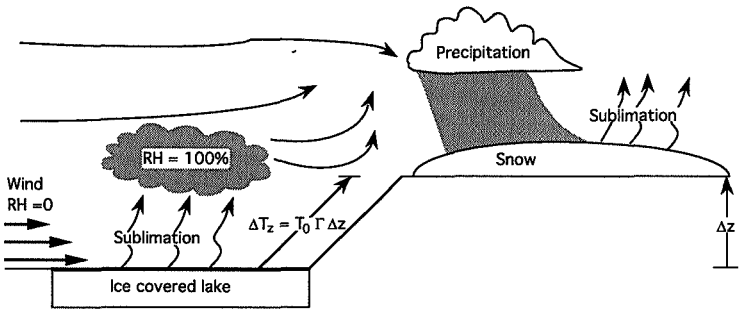


Figure 8. Schematic diagram illustrating the parameters used for the lake and the uplands condensation region in the airborne hypothesis model.

To determine the total time it takes to drain the lake, we need to calculate the discharge as a function of time $Q(t)$ through the permeable zone. The integrated discharge is equal to the volume of the lake (V_{lake})

$$\int_{t_0}^{t_{empty}} Q(t)\,dt = V_{lake} \tag{25}$$

where

$$Q(t) = A(t)q(t) \tag{26}$$

and

$$q(t) = \frac{k\rho g}{\mu}\frac{\Delta h(t}{L} \tag{27}$$

The average velocity of the water q through the permeable zone varies with time because the difference in the hydraulic head decreases as the lake drains. In this equation, which is a form of Darcy's law, k is the average permeability of the subsurface, ρ is the water density, g is the gravitational acceleration for Mars, μ is the viscosity of the water, and L is the total distance the water travels through the permeable zone. We numerically integrated the lake volume to solve for the time it takes to completely drain the lake (t_{empty}) having set $\Delta h(0)$ = 1.2 km. The change in the hydraulic head as a function of time ($\Delta h(t)$) is given by

$$\Delta(t) = \left[z_{lake} + d(t)\right] - h \tag{28}$$

where z_{lake} is the elevation of the lake bed above the source-site deep storage region (1 km) and d is the lake depth which varies with time. The hydraulic head at the source h_s is set equal to zero because water is not allowed to pool. If the initial change in hydraulic head is 1.2 km over a distance of 1000 km and if we assume a permeability of 1000 darcies, then the lake will completely drain in approximately 3300 years. This calculation takes into account the fact that both the lake volume and the wetted surface area of the permeable zone

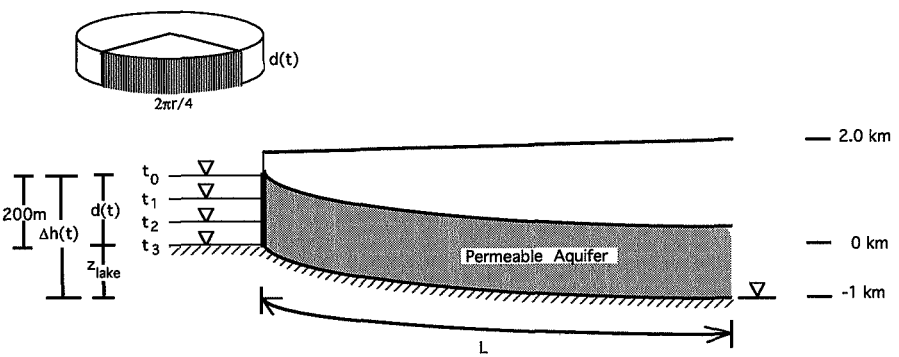


Figure 9. Illustration of the groundwater seep model. Lake is assumed to be disk-shaped with an initial depth equal to 200 m. The cross-sectional area, through which the lake is allowed to drain is equal to one-quarter the circumference of the disk times the depth of the lake at time t_i. All lake water drains through the lake margin and into the permeable aquifer for a distance equal to 1000 km beyond the lake's edge. An impermeable zone underlies the lake and aquifer. The area subsurface storage at the floodsource site lies 1 km deeper than the base of the lake and is assumed to be much larger than the lake. Therefore, we assume that the hydraulic head at the flood source does not change.

decrease with time, so initially draining is quite rapid. Half of the water will drain into the subsurface in 923 years, and 3/4 of the water will be removed by 1717 years. The time required to drain the lake is inversely proportional to the average permeability of the subsurface. Therefore, if the average subsurface permeability is actually 100 darcies then it will take an order of magnitude longer to completely drain the lake. Ultimately, the time constraint for removal of water by seepage is the lifetime of the lake as defined by the rate its ice cover thickens until the lake is frozen solid and by the sublimation rate of H_2O from the top of the lake's surface.

Another issue to address is how long does it take for the entire lake volume to reach the source region of the floodwaters. This can easily be calculated by dividing the total distance L the water has to travel by the velocity of the water through the permeable zone q at t_{empty} when the lake has fully drained. Because Δh within the permeable zone is equal to the elevation of the lake bed at t_{empty}, we can use equation (27) to calculate q at t_{empty}. Assuming $k = 10^3$ darcies, the entire volume of water contained in the lake will have reached the source region in 8466 years (Table 6).

We can now calculate the time required for a column of liquid to be vertically transported to the surface, where it is again available for discharge. The transport mechanism is no longer liquid flow, but vertical vapor diffusion from the surface of the deep ponded water to the cold, near-surface regoltih. We assume an initially dry regolith column, bounded at the top by the surface, and at the bottom by standing water. The surface temperature is assumed to be the temperature at the annually averaged insolation, calculated from equation (12). We assume that the surface albedo is 0.214, and the lower boundary condition is

$$dt/dz = -Q/K \tag{29}$$

where Q is the heat flow and K is the thermal conductivity of the regolith at the bottom of the column. We assume a background heat flow of 0.03 J m^{-2} s^{-1} [*Fanale et al.*, 1982] and a thermal conductivity of 0.6 J s^{-1} m^{-1} K^{-1} [see *Clifford*, 1993].

As long as the thermal gradient is steeper than 0.02 K m^{-1}, the 0°C melting isotherm will lie above the water level (assumed to be at a depth of 3 km). The vapor pressure at the base of the column is the vapor pressure of water at the appropriate temperature. The upper boundary condition is assumed to be the atmospheric H_2O abundance, although in practice, the gradient is soon controlled by the precipitation vertically throughout the column. The vertical flux of H_2O through the column to a near-surface cold-trap is controlled by

$$\Phi_v = -\frac{\beta D_{eff} P_{H_2O} L_v}{\rho R_v^2 T^3} \frac{dT}{dz} \tag{30}$$

where D_{eff} is the effective diffusion coefficient of H_2O in CO_2, P_{H_2O} is the saturation vapor pressure of H_2O at temperature T of the regolith, L_v is the latent heat of vaporization, and β is a dimensionless empirical factor which has the value of 1.83 [*Jury and Letey*, 1979]. The term ρ is the density of liquid water, and R_v is the gas constant.

The calculation of the total H_2O deposited above the 0°C isotherm is simplified because the regolith is saturated throughout the column. In roughly the upper kilometer, the H_2O precipitates as ice. At greater depths, water can precipitate as a liquid if the confining pressure is high enough, and will ultimately drain back down to the water table. From equation (30), we can calculate directly the flux of water passing through the 0°C isotherm as ~3 x10^{-12} kg m^{-2} s^{-1}. Such a flux can transport ~100 column-meters of ice to the near-surface cryosphere in 10^6 years. If a subsequent heat pulse raises the 0°C isotherm into an ice-charged cryosphere, melting can take place, forming and bursting a charged, confined aquifer. For confined aquifers we invoke perched, near-surface groundwater resting on aquatards, which while permeable to the upward infiltration of H_2O vapor over a 10^6-year period, are sufficiently impermeable to liquid water so that "rapid" melting and downslope drainage at the surface depletes the aquifers much more quickly than would vertically downward seepage. We conclude that given the circumstances invoked to initiate a new episode of flooding, ordinary thermal vapor diffusion is adequate to sustain the system.

Conclusions

Our analysis of the airborne precipitation cycling hypothesis indicates that this mechanism seems unlikely to effectively operate under current conditions on Mars. Our modeling was idealized to show the most robust case for this hypothesis. We calculate that less than half (~40%) of the H_2O carried from Chryse can condense over the highlands. Another result of our analysis is that snowfields with areas larger than 3.7 times the lake area would sublimate faster than they accumulate. It is possible that local topography might cause local "heavy" precipitation within a confined area. The most likely place for this to occur in the Chryse region is directly over the lowlands/highlands scarp. Unfortunately, precipitation at the scarp would dump snow back on the southern edge of the lake, leaving very little to travel on to the flood-source sites.

The particle size range of snow in the Martian atmosphere varies at least by roughly a factor of 30, from ~0.4 μm to ~12 μm [*Kahn*, 1990]. Small variations in particle size result in very different settling rates. *Murphy et al.* [1990] calculated that it would take ~28 hours for a 10-μm particle to fall from 1

Table 6. Time Required to Drain the Lake and the Time Required for the Entire Lake Volume to Reach the Highland Source Region as a Function of Average Subsurface Permeability

Permeability k ,darcies	Total Time to Drain Lake, years	Time Required for Entire Lake Volume to Reach Source Region, years
10^3	3.30 x 10^3	8466
10^2	3.30 x 10^4	8.466 x 10^4
10^1	3.30 x 10^5	8.446 x 10^5
10^{-1}	3.30 x 10^7	8.446 x 10^7

km above the surface, and ~139 hours for a 1-μm particle to fall from the same height. Parcels of air in which precipitation is taking place traveling as slow as 1 m s^{-1} would cover a distance of ~400 km in the time it took particles between 1 and 10 μm to fall 1 km to the ground. Assuming that the parcel initially contains all the H_2O it collected over the lake, that all the condensate forms snow particles that are within 1-10 μm, and that there is no lateral spreading or upward mixing of the parcel, precipitation would take place over an area approximately equal to our model "equilibrium" snowfield. Any change toward more realistic conditions would make retention of snow on the highlands unlikely. Moreover, we assumed, for modeling purposes, that 100% of the H_2O that does condense over the highlands actually collects on the ground as snow. It is very likely that much, maybe most, of the condensates formed under present Martian conditions would simply stay suspended until the air parcel is mixed with enough dry gas to cause their sublimation before ever settling to the ground [*Kahn*, 1990; *Michelangeli et al.*, 1993].

Given all these normally unfavorable conditions, it is difficult to imagine that any more than a few percent of the H_2O leaving a lake in Chryse could collect and remain in the circum-Chryse highlands. Even if several percent H_2O were incorporated into an aquifer, such efficiency would only marginally qualify as a "cycle."

The results of our analysis of the groundwater seep cycling hypothesis are more promising. Optimally, ~ 3/4 of the original mass of an ice-covered cylindrical lake (albedo 0.5, 1 km deep, 100-km radius, draining along its rim for one quarter of its circumference into substrata with a permeability of 1000 darcies) can be modeled to have moved underground before the competing mechanisms of sublimation and freeze-down choked off further water removal. Another encouraging result of our modeling is that floodwater will travel distances equal to the separation between Chryse basin and floodwater-source sites in less than 10^8 years for our "worst case" (k = 0.10 darcies) and 10^4 years for "best case" (k = 1000 darcies) permeabilities. Like the evaluation of the airborne precipitation cycling hypothesis, our modeling of groundwater seep cycling was idealized to make the most robust case. Combinations of shallow lakes and relatively impermeable strata will prevent significant amounts of H_2O from returning by this route to its source sites. Also, the hypothesis is absolutely dependent on the reality of the structural situation we have proposed.

For the groundwater seep hypothesis to be viable, the water under its ice cover must have extensive access to a ground interface that is ≥ 0°C. This is because subfreezing ground has sufficient thermal inertia to seal itself by freezing the H_2O that seeps into its pores. Something on the order of a kilometer of material must be removed to uncover the 0°C isotherm. We propose that access to warmer ground could be accomplished by backwasting of the highlands/lowlands boundary scarp and gouging of the basin floor in and around where the flood disgorges into Chryse basin. Certainly extensive examples of this type of erosion are observed in association with the Chryse outflow channels [e.g., *Baker*, 1982; *Tanaka and Chapman*, 1992].

If the dip of the strata in the circum-Chryse region is as we have hypothesized, and these layers are reasonably permeable and accessible, then perhaps the majority of the water in a Chryse flood lake could have returned to the highlands reservoirs and participated in subsequent flood events. Certainly of the two hypotheses we formulated and evaluated, the groundwater seep cycle seems by far the more viable. Of course, there is the possibility that there has never been a floodwater hydrologic cycle operating in the Chryse region. While observations of landform associations seem to imply that such a cycle has existed, they do not absolutely mandate this explanation. Further observations may permit the determination of which mechanism, if either or any, may have operated to recycle the Chryse floodwaters. A principal conclusion of this study is that the supposition of water recycling, while not conclusively demonstrated, is at least reasonable, even under the stringent conditions of the modern Martian environment.

We have identified several tests that hopefully could be conducted with data from forthcoming missions, such as Mars 96 and the Mars Global Surveyor (MGS) spacecraft. First and foremost, for the groundwater seep cycling hypothesis to gain greater acceptance, substantiating evidence for the existence and nature of the putative permeable, antibasinward dipping, subsurface layer must be obtained. The MGS Mars orbiting laser altimeter (MOLA) can provide the topographic information to determine the gradients of the outflow channels and relative elevations of sites where dipping stratigraphic marker horizons might outcrop. The Mars orbiter camera (MOC) and thermal emission spectrometer, both which fly on the MGS, could be used to recognize such outcrops where they might occur at widely scattered locations. High-resolution images of the highlands/lowlands scarp where it borders the Chryse basin acquired from MOC and the Mars 96 camera images (particularly in stereo) would provide a more quantitative characterization of the nature and extent of flood-induced erosion along the southern margin of the Chryse basin. Such imaging would allow a detailed evaluation of the efficacy of floods to access the 0°C isotherm. A high-resolution altimeter (e.g., MOLA) could also give a much better determination of the size, shape, and depth of Chryse basin, which bears on the location, areal extent, and depth of the postflood lakes.

Acknowledgments. We wish to express our gratitude to Vic Baker and Tim Parker for their thorough reviews of our manuscript. We would like to thank Bob Haberle, Jim Murphy, and Ray Reynolds for their valuable comments and insight to various issues regarding this study. This research was conducted while J. M. Moore and V. C. Gulick were Associate Research Fellows of the National Research Council; D. R. Janke was supported under NASA grant NAGW 1; G. D. Clow, W. L. Davis, and C. P. McKay were funded under NASA grant RTOP 154-10-80-48; C. R. Stoker under NASA grant RTOP 154-60-80-45; and A. P. Zent was supported by NASA grant RTOP 151-01-60-11.

References

Baker, V.R., *The Channels of Mars*, 198 pp., University of Texas Press, Austin, 1982.

Baker, V.R., and D.J. Milton, Erosion by catastrophic floods on Mars and Earth, *Icarus*, *23*, 27-41, 1974.

Baker, V.R., R.G. Strom, V.C. Gulick, J.S. Kargel, G. Komatsu, and V.S. Kale, Ancient oceans, ice sheets and the hydrological cycle on Mars, *Nature*, *352*, 589-594, 1991.

Baker, V.R., M.H. Carr, V.C. Gulick, C.R. Williams, and M.S. Marley, Channels and valley networks, in *Mars*, edited by H.H. Kieffer et al., p.p. 493-522, University of Arizona Press, Tucson, 1992.

Barlow, N.G., Crater size-frequency distribution and a revised martian chronology, *Icarus*, *75*, 285-305, 1988.

Brutsaert, W., A theory for local evaporation (or heat transfer) from rough and smooth surfaces at ground level, *Water Resour. Res., 11*, 543-550, 1975.

Brutsaert, W., *Evaporation into the Atmosphere*, 299 pp., D. Reidel, Norwell, Mass., 1982.

Businger, J.A., J.C. Wyngaard, Y. Izumi, and E.F. Bradley, Flux-profile relationships in the atmospheric surface layer, *J. Atmos. Sci., 28*, 181-189, 1971.

Carr, M.H., Formation of martian flood features by release of water from confined aquifers, *J. Geophys. Res., 84*, 2995-3007, 1979.

Carr, M.H., *The Surface of Mars*, 232 pp., Yale University Press, New Haven, Conn., 1981.

Carr, M.H., The ancient ocean hypothesis: A critique (abstract), *Bull. Am. Astron. Soc., 23*, 1206, 1991.

Carr, M.H., and G.D. Clow, Martian channels and valleys: Their characteristics, distribution, and age, *Icarus, 48*, 91-117, 1981.

Chapman, M.G., and D.H. Scott, Geology and hydrology of the North Kasei Valles Area, Mars, *Proc. Lunar Planet. Sci. Conf., 19th*, 367-375, 1989.

Clifford, S.M., A model for the hydrologic and climatic behavior of water on Mars, *J. Geophys. Res., 98*, 10,973-11,016, 1993.

Clow, G.D., C.P. McKay, G.M. Simmons Jr., and R.A. Wharton Jr., Climatological observations and predicted sublimation rates at Lake Hoare, Antarctica, *J. Clim., 1*, 715-728, 1988.

Clow, G.D., and R.M. Haberle, Characteristics of the martian atmospheric surface layer (abstract), *Lunar Planet. Sci., 21*, 210-211, 1990a.

Clow, G.D., and R.M. Haberle, Free-convection in the martian atmosphere, *Lunar Planet Sci. 21*, Lunar and Planetary Institute, Houston, 210-211, 1990b.

Davis S., and R. DeWiest, *Hydrogeology*, 463 pp., John Wiley and Sons, New York, 1966.

Fanale, F.P., J.R. Salvail, W.B. Banerdt, and R.S. Saunders, Mars: The regolith-atmosphere-cap system and climate change, *Icarus, 50*, 381-407, 1982.

Goldspiel, J.M., and S.W. Squyres, Ancient aqueous sedimentation on Mars, *Icarus, 89*, 392-410, 1991.

Goldspiel, J.M., S.W. Squyres, M.A. Slade, R.F. Jurgens, and S.H. Zisk, Radar-derived topography of low southern latitudes of Mars, *Icarus, 106*, 346-364, 1993.

Haberle, R.M., and B.M. Jakosky, Atmospheric effects on the remote determination of thermal inertia on Mars, *Icarus, 90*, 187-204, 1991.

Inoue, J., Surface drag over snow surfaces in the Antarctic Plateau 2. Seasonal change of surface drag in the katabatic wind region, *J. Geophys. Res., 94*, 2219-2224, 1989.

Janke, D.R., A relative time scale for channeling events in the Oxia Palus/Margaritifer Sinus region, Mars (abstract), *Lunar Planet. Sci., 16*, 398-399, 1985.

Joshi, M.M., S.R. Lewis, P.L. Read, and D.C. Catling, Western boundary currents in the atmosphere of Mars, *Nature, 367*, 561-562, 1994.

Jury, W.A., and J. Letey Jr., Water vapor moment in soil: Reconciliation of theory and experiment, *Soil Sci. Soc. Am. J., 43*, 823-827, 1979.

Kahn, R., Ice haze, snow, and the Mars water cycle, *J. Geophys. Res., 95*, 14,677-14,693, 1990.

Kargel, J.S., and R.G. Strom, Ancient glaciation on Mars, *Geology, 20*, 3-7, 1992.

Kieffer, H.H., T.Z. Martin, A.R. Peterfreund, B.M. Jakowsky, E.D. Miner, and F.D. Palluconi, Thermal and albedo mapping of Mars during the Viking primary mission, *J. Geophys. Res., 82*, 4249-4292, 1977.

Lucchitta, B.K., and H.M. Ferguson, Chryse basin channels: Low-gradients and ponded flows, *Proc. Lunar Planet. Sci. Conf. 13th*, Part 2, *J. Geophys. Res., 88*, suppl., A553-A568, 1983.

Lucchitta, B.K., H.M. Ferguson, and C. Summers, Sedimentry deposits in the northern lowland plains, Mars, *Proc. Lunar Planet. Sci. Conf. 17th*, Part 1, *J. Geophys. Res., 91*, suppl., E166-E174, 1986.

Mars Channel Working Group, Channels and valleys on Mars, *Geol. Soc. Am. Bull., 94*, 1035-1054, 1983.

Marti, J., and K. Mauersberger, A survey and new measurements of ice vapor pressure at temperatures between 170 and 250 K, *Geophys. Res. Lett., 20*, 363-366, 1993.

Marsursky, H, J.M. Boyce, A.L. Dail Jr., G.G. Schaber, and M.E. Strobell, Classification and time of formation of Martian channels based on Viking data, *J. Geophys. Res., 82*, 4016-4037, 1977.

McKay, C.P., G.D. Clow, R.A. Wharton, and S.W. Squyres, Thickness of ice on perennially frozen lakes, *Nature, 313*, 561-562, 1985.

McKay, C.P., G.D. Clow, D.T. Andersen, and R.A. Wharton Light transmission and reflection in perennially ice-covered Lake Hoare, Antarctica, *J. Geophys. Res., 99*, 20,427-20,444, 1994.

Michelangeli, D.V., O.B. Toon, R.M. Haberle, and J.B. Pollack, Numerical simulations of the formation and evolution of water ice clouds in the martian atmosphere, *Icarus, 100*, 261-285, 1993.

Murphy, J.R., O.B. Toon, R.M. Haberle, and J.B. Pollack, Numerical simulations of the decay of global dust storms, *J. Geophys. Res., 95*, 14,629-14,648, 1990.

Parker, T.S., R.S. Saunders, and D.M. Schneeberger, Transitional morphology in the west Deuteronilus Mensae region of Mars: Implications for modification of the lowland/upland boundary, *Icarus, 82*, 111-145, 1989.

Pieri, D.C., Martian channels: Distribution of small channels on the martian surface, *Icarus, 27*, 25-50, 1976.

Rossbacher, L.A., and S. Judson, Ground ice on Mars: Inventory, distribution, and resulting landforms, *Icarus, 45*, 39-59, 1981.

Rotto, S.L., and K.L. Tanaka, Chryse Planitia region, Mars: Channeling history, flood-volume estimates, and scenarios for bodies of water in the northern plains (abstract), in *Workshop on the Martian Surface and Atmosphere Through Time (MSATT)*, pp. 111-112, Lunar and Planetary Institute, Houston, Texas, 1991.

Schultz, P.H., R.A. Schultz, and J.R. Rogers, The structure and evolution of ancient impact basins on Mars, *J. Geophys. Res., 87*, 9803-9820, 1982.

Scott, D.H., and M.H. Carr, Geologic map of Mars, *U.S. Geol. Surv. Misc. Invest. Map, I-1083*, 1978.

Scott, D.H., and K.L. Tanaka, Geologic map of the western equatorial region of Mars, *U.S. Geol. Surv. Misc. Invest. Map, I-1802-A*, 1986.

Scott, D.H., and M.G. Chapman, Mars Elysium basin: Geologic/volumetric analysis of a young lake and exobiologic implications, *Proc. Lunar Planet. Sci., Vol. 21*, pp. 669-677, 1991.

Scott, D.H., M.G. Chapman, J.W. Rice, and J.M. Dohm, New evidence of lacustrine basins on Mars: Amazonis and Utopia Planitiae, *Proc. Lunar Planet. Sci., 22*, 53-62, 1992.

Sharp, R.P., and M.C. Malin, Channels on Mars, *Geol. Soc. Amer. Bull., 86*, 593-609, 1975.

Tanaka, K.L., The stratigraphy of Mars, *Proc. Lunar Planet. Sci. Conf. 17th*, Part 1, *J. Geophys. Res., 91*, suppl., E139-E158, 1986.

Tanaka, K.L., and M.G. Chapman, Kasei Valles, Mars: interpretation of canyon materials and flood sources, *Proc. Lunar Planet. Sci. Conf., 22*, 73-83, 1992.

Turcotte, D.L., and G. Schubert, *Geodynamics*, 450 pp., John Wiley, New York, 1982.

U.S. Geological Survey, Topographic maps of the western and eastern equatorial and polar regions of Mars, *U.S. Geol. Surv. Misc. Invest. Map, I-2020*, 1989.

White, W.B., Conceptual models for carbonate aquifers, *Groundwater, 7*, 15-71, 1969.

G. D. Clow, U.S. Geological Survey, 345 Middlefield Road, Menlo Park, CA 94025.

W. L. Davis, J. M. Moore, A. P. Zent, Center for Mars Exploration and the SETI Institute, NASA Ames Research Center, MS 245-3, Moffett Field, CA 94035.

V. C. Gulick, C. P. McKay, C. R. Stoker, NASA Ames Research Center, MS 245-3, Moffett Field, CA 94035.

D. R. Janke, Department of Geography, Arizona State University, Tempe, AZ 85287.

(Received March 8, 1994; revised October 24, 1994;
accepted October 27, 1994.)

Page intentionally left blank

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5449–5455, MARCH 25, 1995

A search for nitrates in Martian meteorites

Monica M. Grady[1]
Department of Mineralogy, The Natural History Museum, London, England, United Kingdom

I. P. Wright and C. T. Pillinger
Department of Earth Sciences, The Open University, Milton Keynes, England, United Kingdom

Abstract. Martian atmospheric nitrogen is highly enriched in ^{15}N; nitrates formed by interaction of the atmosphere with the Martian regolith should therefore also be characterized by an elevated $\delta^{15}N$ value. A search has been made for nitrates in two Martian meteorites, in order to determine the extent of possible regolith-atmosphere interaction. Shock-produced glass from the Elephant Moraine (EET) A79001 shergottite (E1,149) and a water-soluble extract from Nakhla were analyzed by Fourier transform infrared (FTIR) spectroscopy and stepped combustion-stable isotope mass spectrometry. FTIR of both meteorites had features at 1375 cm^{-1} and 1630 cm^{-1}, consistent with nitrates. On account of their low thermal stability, nitrates break down at temperatures below 600°C; in this temperature range, E1,149 yielded ~1250 ppb nitrogen with $\delta^{15}N$ -8±5‰. If this nitrogen is from a nitrate, then it cannot be distinguished from terrestrial salts by its isotopic composition. The water-soluble extract from Nakhla also released nitrogen at low temperatures, approximately 17 ppb with $\delta^{15}N$ approximately -11±4‰. Since Nakhla is an observed "fall", this is unlikely to be a terrestrial weathering product. Nitrates apparently occur in E1,149 and Nakhla, but in very low abundance, and their origin is unclear. The isotopic composition of the salts, which is within the range of that proposed for Martian magmatic volatiles, is far removed from that of nitrogen in the present-day Martian atmosphere. If the nitrates are Martian in origin, they did not form in recent times from reactions involving atmospheric gases. Rather, the nitrates could be the result of an earlier episode of atmospheric interaction with the regolith, or with implantation of magmatic volatiles introduced during degassing.

Introduction and Summary

The surface of Mars has had an active history: images taken by Mariner 9 and the Viking orbiters testify to episodes of fluvial and aeolian activity, volcanism and cratering. A chronology can be established by relating craters on Mars with events recorded on the lunar surface [e.g., *Tanaka*, 1986]. Timing of surface features in relation to the history of fluvial activity has implications for Martian evolution: liquid water was presumably an important medium for long periods of Mars' history, suggesting that the atmosphere in the past was more dense than that at present (the current surface pressure being ~ 6 mbar). Under conditions of higher palaeoatmospheric pressures, there is an increased probability of interactions between atmospheric and surficial phases, with the formation of salt deposits (e.g., carbonates, sulphates, and nitrates). Thermal emission spectroscopy of the surface of Mars has yielded spectra which have been identified as HCO_3^-, CO_3^{2-}, SO_4^{2-}, etc. [*Pollack et al.*, 1990]. Although nitrates have not yet been recorded by spectroscopic techniques, they are still regarded as a potential component of the Martian soil [*Pollack et al.*, 1990; *Owen*, 1992]. Typically, NO_3^- anions are thought to have been incorporated into the regolith from atmospheric species. Thus, if present, nitrates might have an isotopic composition related to that of nitrogen in the Martian atmosphere (i.e., ^{15}N-enriched).

We have searched for nitrates in Martian meteorites, in a shock-produced glass from the Elephant Moraine (EET) A79001 shergottite (E1,149) and in a water-soluble extract from Nakhla. On the basis of Fourier transform infrared (FTIR) spectroscopy there is some evidence for nitrates in the glass sample. However, FTIR does not distinguish between discrete nitrate minerals, or NO_3^- dissolved in the glass (either from trapped Martian atmospheric gases or original magmatic volatiles). E1,149 was therefore also analyzed by stepped combustion; a large portion of the total nitrogen in the sample (1250 ppb) was found to be released at low temperatures, consistent with the breakdown of nitrate minerals. The best estimate of the $\delta^{15}N$ value of this species is about -8±5‰. EET A79001 is a meteorite collected from Antarctica; hence the nitrates may have been produced by terrestrial weathering. The second meteorite studied, Nakhla, was observed to fall, and thus should not have suffered from any obvious terrestrial weathering. A water extraction of Nakhla was prepared, to concentrate any water-soluble salts present. Results indicate ~ 17 ppb nitrogen as nitrates with $\delta^{15}N$ approximately -11±4‰.

Assuming that the low-temperature material is from Martian nitrates, the $\delta^{15}N$ data are inconsistent with formation by direct incorporation of present-day Martian atmospheric species (N_2 or NO_x) into the regolith. Any minor nitrates present in SNC meteorites may therefore have resulted from the redeposition of degassed magmatic species within the Martian crust. The mechanisms of salt formation in Martian meteorites are far from well understood. The carbonate to nitrate ratio of the two meteorites ranges from ~ 10 to 2000. The model of *Yung et al.* [1977] pre-

[1]Also at Department of Earth Sciences, The Open University, Milton Keynes, England, United Kingdom.

Paper number 94JE02803.
0148-0227/95/94JE-02803$05.00

dicts a layer of nitrates 0.3 cm thick over the surface of Mars, from which the presence of carbonate deposits 3 to 600 cm thick might be inferred. This calculation is in agreement with the prediction by *Kahn* [1985] that up to 280 cm of carbonates are stored in the Martian regolith.

Background

Nitrogen, currently present at a level of ~ 2.7%, is the second most abundant constituent of the Martian atmosphere, after CO_2. The nitrogen isotopic composition of the present-day Martian atmosphere has been measured directly, by the Viking landers, and indirectly, by analysis of species trapped within pockets of shock-produced glass (lithology C) in the Martian meteorite EET A79001. Nitrogen isotopic composition ($\delta^{15}N$) is given as a parts per thousand (per mil; ‰) deviation from that of nitrogen in the terrestrial atmosphere (AIR, the international reference point):

$$\delta^{15}N = \left(\frac{(^{15}N/^{14}N)_{SAMPLE} - (^{15}N/^{14}N)_{AIR}}{(^{15}N/^{14}N)_{AIR}} \right) \times 1000$$

Both measurements have found the atmosphere to be enriched in ^{15}N, with $\delta^{15}N$ calculated as approximately $+600^{+155}_{-130}$‰ from Viking results [*Nier et al.*, 1976] and about +320‰ from the meteorite [*Becker and Pepin*, 1984; *Wiens*, 1988]; the meteorite data are a lower limit, due to admixture of isotopically lighter indigenous nitrogen. It is widely held that the enrichment in ^{15}N is a direct consequence of gravitational separation within the atmosphere, coupled with loss of nitrogen atoms from above the exobase [e.g., *McElroy et al.*, 1977]. There are several possible mechanisms by which nitrogen might be removed [e.g., *Pepin*, 1991]; however, the present-day level of ^{15}N enrichment is less than anticipated [*Wallis*, 1989], an effect which has been variously explained as the result of continued outgassing of nitrogen from the interior acting to dilute the enrichment in the atmosphere [*Wallis*, 1989], or the existence of an early, dense atmosphere which suppressed the operation of loss processes for the first part of Martian history [e.g., *Fox*, 1993]. Other possibilities include addition of nitrogen from meteoroid or comet impacts, thermal release of regolith volatiles, or breakdown of surficial salts.

Nitrate deposits in the regolith might also provide an important sink for atmospheric nitrogen [e.g., *Fanale*, 1976]. If nitrates do occur on the Martian surface, and their $^{14}N/^{15}N$ ratio can be determined, an important constraint on the operation of a "nitrogen cycle" on Mars would be provided. Model calculations have predicted that, via the formation of NO_x, HNO_2 and HNO_3 in the lower layers of the Martian atmosphere, the regolith might trap nitrite and nitrate anions, leading to the buildup of salts. Integrated over 4.5×10^9 years, such a mechanism would contribute the equivalent of a "layer" of nitrates approximately 0.3 cm thick across the whole of the Martian surface [e.g., *Yung et al.*, 1977]. This "layer" would be even thicker if the nitrogen content of the atmosphere had been higher in the past.

For terrestrial geochemical studies, it is possible to deduce past surficial processes by analyzing samples of the Earth's atmosphere, in conjunction with precipitated salts and magmatic volatiles. Until material is recovered directly from Mars' surface by a sample return mission, the only way in which an analogous study of the planet might be undertaken is by analyzing the constituents of SNC meteorites, a group of rocks which almost certainly originate from Mars (see *McSween* [1994] for a review). It is believed that during their history the meteorites resided sufficiently close to the surface of Mars to have been affected by secondary alteration [e.g., *Gooding*, 1992]. Furthermore, the impact events which removed the meteorites from the Martian surface have, in some cases, caused atmospheric gases to become trapped within the ejecta [e.g., *Bogard and Johnson*, 1983; *Becker and Pepin*, 1984; *Carr et al.*, 1985]. *Wright et al.* [1990] deduced a "carbon cycle" for Mars on the basis of the abundance and isotopic composition of carbon-bearing components in SNC meteorites. This paper is an attempt to produce a similar relationship for nitrogen on Mars. An impression of the possible complexities associated with the interactions of nitrogen-bearing components at the Martian surface is shown in Figure 1.

The nitrogen-bearing components that might occur in Martian meteorites include trapped terrestrial atmospheric gas, organic material (i.e., terrestrial contamination), inorganic weathering products (of terrestrial or Martian origin), magmatic nitrogen, and

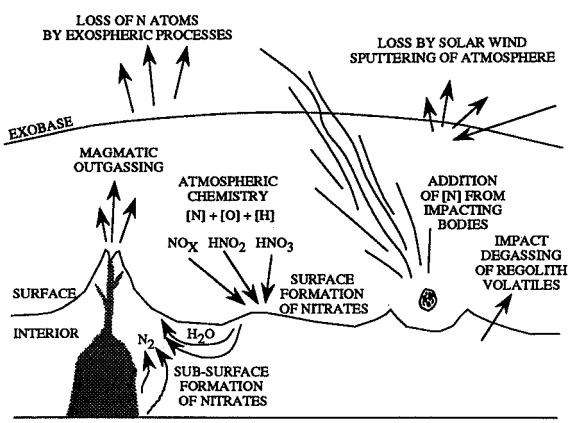


Figure 1. Schematic diagram of some of the possible interactions of nitrogen and its components at the Martian surface.

Martian atmospheric gas trapped in impact glass. Nitrates are unstable, decomposing at low temperatures and readily soluble in water. They might be removed from Martian meteorites during ejection from the planet's surface, passage to Earth or by weathering during the sample's terrestrial history. However, stability may be imparted to nitrates by the formation of mixed salts. The nitrate anion (NO_3^-) is in some cases isostructural with the carbonate anion (CO_3^{2-}), e.g., KNO_3 with calcite, $CaCO_3$ [*Hill,* 1981]. Acid sulphate-nitrate hydrates form when nitrogen oxides are distilled onto hydrated sulphates such as $MgSO_4.4H_2O$ [*Plumb et al.*, 1987]. Carbonates and sulphates occur in SNC's [e.g.,*Carr et al.*, 1985; *Gooding et al.*, 1988; *Burgess et al.* 1989]; thus nitrates might form stable mixed salts with these species.

Two meteorites were selected for analysis, the shergottite EET A79001 and Nakhla. Two subfractions from EET A79001 were hand-picked from the bulk meteorite at the Antarctic Meteorite Curatorial Facility, Houston, Texas. The first is E1,149, a sample of lithology C (shock-produced glass), in which nitrates might be "protected" (i.e., enclosed in glass), like some of the carbonates. The second fraction, E1,239, is a carbonate-rich sample of druse derived from the interior of a vug in lithology A [*Gooding et al.*, 1988; *Wright et al.*, 1988b]. Since E1 was collected from Antarctica, and has a terrestrial residence age of 12,000 years [*Jull and Donahue,* 1988], salt formation may have taken place, or the salts modified, on Earth. A non-Antarctic SNC meteorite was therefore also studied in addition. Nakhla is a meteorite retrieved from a temperate location soon after it was observed to fall and which has therefore suffered only minimal terrestrial weathering. Nakhla is also known to contain carbonates, of presumed Martian origin [*Carr et al.*, 1985].

Analytical Methods

The analytical techniques utilized, namely FTIR and stepped combustion-stable isotope mass spectrometry, allow discrete phases to be recognized within a whole-rock sample without physical separation of individual components. A reservoir of whole-rock Nakhla (BM 1911,369) was prepared by gently crushing an interior fragment (approximately 1 g) to a homogeneous powder; aliquots of bulk sample were drawn from this reservoir for analysis. In an attempt to concentrate water-soluble salts, such as nitrates, an extract was prepared from Nakhla by suspending ~ 500 mg of powdered sample in 1 mL of triply distilled, deionized water for 20 hours. The resulting solution (which contained fine mineral grains from the meteorite, in addition to any water-soluble salts) was evaporated to dryness, yielding a poorly crystalline residue.

The instrument used for FTIR was a Philips PU 9800 infrared spectrophotometer, wavelength range from 225-4400 cm^{-1}, continuously purged with CO_2-free dry air. Fifty scans in transmission mode at a resolution of 2 cm^{-1} were acquired for both samples and standards. All specimens were finely powdered in a matrix of spectroscopic-grade KBr, dried, then cold-pressed into a disc ~ 1 cm in diameter. Results from the sample are compared directly with that from a KBr-only disc, to allow background to be subtracted. Several nitrate species (both analytical-grade chemicals and natural minerals) were also subjected to FTIR, in order to assist calibration.

Stepped combustion is the incremental oxidation of nitrogen-bearing components in discrete temperature intervals from room temperature upwards [*Boyd et al.*, 1988; *Wright et al.*, 1988a]. All gaseous nitrogen products of the combustion are converted to N_2 gas using a combination of heated Cu/CuO and Pt. The nitrogen is then separated cryogenically from other products of combustion (predominantly CO_2, SO_2 and H_2O), prior to admission to a triple-collector noble gas-type mass spectrometer, operated in static mode, for determination of nitrogen isotopic composition. Nitrogen yield is gained from calibration of the *m/z* 28 ion beam current. The instrument currently used (known colloquially as FINESSE) can measure the $\delta^{15}N$ of samples as small as 100 pg of nitrogen to a precision of ±1‰. As a by-product of the gas clean-up procedure, carbon yields are also obtained, by measurement of the CO_2 gas on a capacitance manometer, to a precision of ±0.1 ng, enabling determination of C/N ratios of the various components.

During stepped combustion, different nitrogen-bearing species exhibit a variety of quantifiable thermal breakdown characteristics. Table 1 summarizes the measured combustion temperature ranges of nitrogen-bearing species anticipated present in Martian meteorites. Data for nitrates were determined by stepped combustion of analytical-grade reagents $Mg(NO_3)_2.6H_2O$ (nitromagnesite) and $Ca(NO_3)_2.4H_2O$ (nitrocalcite) [*Grady et al.*, 1993]. Nitromagnesite decomposed across a single temperature step between 400 and 500°C, but the release of nitrogen from nitrocalcite was more complex, showing a bimodal release with peaks at 200-300°C and 400-450°C.

Results

FTIR

Nitrates and carbonates have strong, characteristic absorptions in the infrared. From the work carried out here, the most prominent feature in all spectra measured for the standard nitrates was a strong singlet peak at approximately 1375 cm^{-1}, and a second, minor singlet at ~ 1630 cm^{-1} (Figure 2). Comparison of data from inorganic nitrates with those from carbonates indicated that the two types of salt are resolved using FTIR spectroscopy (the strongest carbonate feature is at ~ 1425-1450 cm^{-1} [*Jones and Jackson,* 1993]).

Analysis of powdered E1,149 glass yielded a spectrum with several features (Figure 3): singlets at 1375 and 1630 cm^{-1}, indicative of nitrates, and a doublet between 1400 and 1450 cm^{-1}, which suggests carbonates. FTIR is not able to give an indication of the relative concentrations of the two salt components, and so all that can be concluded is that E1,149 appears to contain carbonates and nitrates. The absorptions at 1375 cm^{-1} and 1630 cm^{-1} represent the first evidence that nitrates might occur in Martian meteorites. However, it is not clear from the FTIR data whether the nitrates exist as isolated anions dissolved in the glass, or discrete mineral phases.

The spectra obtained from a powdered whole-rock sample of Nakhla (~ 2 mg) and the water-extract of the same meteorite (500 mg) are also shown in Figure 3. For the whole-rock sample, the FTIR data are dominated by broad features below 1250 cm^{-1},

Table 1. Possible Nitrogen-Bearing Components Released at Low Temperatures From Martian Meteorites

Species	Temperature, °C
Absorbed terrestrial atmospheric N_2	<100
Organic materials	200-450
Nitrocalcite [$Ca(NO_3)_2.4H_2O$]	200-300, 400-450
Nitromagnesite [$Mg(NO_3)_2.6H_2O$]	400-500

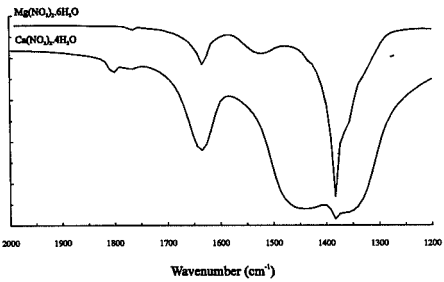


Figure 2. FTIR data for two pure nitrate minerals (nitrocalcite and nitromagnesite). Note the characteristic features in both traces at 1375 and 1630 cm^{-1}.

from the silicate minerals, with only minor features around 1375 and 1630 cm^{-1}. No carbonate absorptions were observed, even though calcite is known to be present within Nakhla at a level of approximately 20-30 ppm [*Carr et al.*, 1985]. However, previous determinations of carbonate in Nakhla were carried out using sample sizes in the range 15-150 mg. Thus the FTIR data here may demonstrate that Nakhla is heterogeneous with respect to carbonate on the scale of 2 mg.

Silicate features also occurred in the water extract, but at reduced intensity. Broad absorptions, observed from 1325-1475 cm^{-1} and 1550-1700 cm^{-1}, were absent from the whole-rock spectrum. Since nitrates are soluble, and would concentrate in a water extract, it is possible that these features are from nitrates. The absence of carbonate absorptions from the water extract of Nakhla is commensurate with the fact that, under the conditions of the extraction, calcium carbonate would be insoluble.

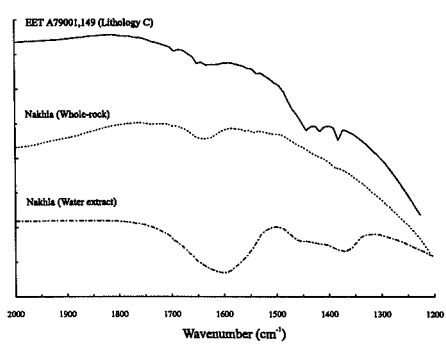


Figure 3. FTIR data for a glass sample of EET A79001 (E1,149), and a whole-rock and water extract from the Nakhla meteorite. The data for E1,149 show evidence for the presence of nitrates and also carbonates. The results for the water extract of Nakhla show a very broad feature in the region where nitrates and carbonates are expected; it is difficult to draw firm conclusions regarding the presence of these minerals in this sample.

Stepped Combustion

Results for the stepped combustion from room temperature to 700°C of a single chip (5.842 mg) of E1,149 glass are shown in Figure 4. The yield histogram may be interpreted as indicating either the presence of two components of differing thermal stability or the bimodal release of nitrogen from a single species. The first release of nitrogen (~ 1000 ppb) occurs below ~ 320°C. Unfortunately, due to analytical problems, isotopic data for these steps were not all recorded; thus the isotopic composition of the lowest temperature material remains poorly defined, with $\delta^{15}N$ between -8‰ and +30‰. The second release of nitrogen (~ 250 ppb) is between 350°C and 400°C, with $\delta^{15}N$ approximately -14‰. If this is from a nitrate, its isotopic composition is not sufficiently distinct from that of Antarctic nitrates [*Wada et al.*, 1981] to assign a specific origin to the material. Above 400°C, nitrogen is released in gradually decreasing quantities with generally increasing $\delta^{15}N$, up to a maximum of +13.2‰ at 650°C. C/N ratios range from ~ 30 to 100; interpretation of these values is ambiguous, since at low temperatures the sources of carbon and nitrogen are decoupled: carbon derives from contaminant organic materials of terrestrial biogenic origin, whereas nitrogen can arise from organics, nitrates, or trapped terrestrial atmospheric gases.

Nitrogen yields and $\delta^{15}N$ values for the water-soluble material from Nakhla are shown in Figure 5. The residue comprises approximately 600 ppm of whole-rock Nakhla; hence the nitrogen content of the water-soluble material represents a whole-rock nitrogen concentration of 17 ppb, most of which is liberated between 200°C and 500°C. As in the case of E1,149, the water extract of Nakhla shows a bimodal release of nitrogen at low temperatures, with $\delta^{15}N$ between about -20‰ and -5‰. C/N ratios exhibit the same range as the whole-rock material, and again fail to constrain the likely nature of the nitrogen-bearing species.

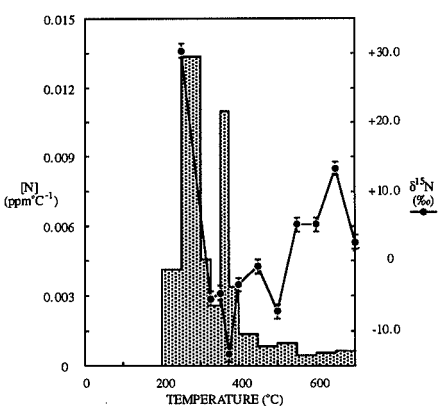


Figure 4. Nitrogen released by stepped combustion (room temperature to 700°C) of E1,149 (5.482 mg). The yield of nitrogen released at each step of the experiment, in ppm°C^{-1}, is given as a histogram; the corresponding isotopic composition of nitrogen ($\delta^{15}N$), is plotted as dots joined by a line. Errors in $\delta^{15}N$ are less than the size of the symbol, unless shown otherwise. Data are corrected for a system blank of < 150 pg per step.

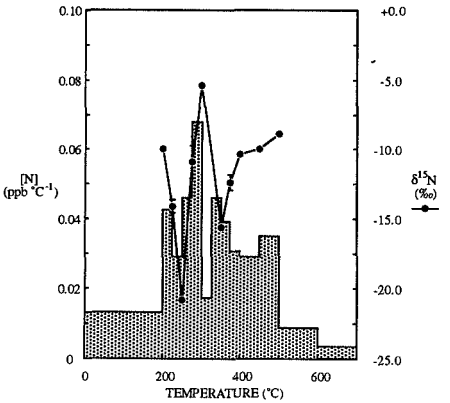


Figure 5. Nitrogen released by stepped combustion (room temperature to 700°C) of water-soluble material from Nakhla (0.142 mg). The extract comprised ~ 600 ppm whole-rock Nakhla. Yield (normalized to whole-rock values) and $\delta^{15}N$ as in Figure 4.

Douglas et al. [1993] have shown that individual pure salts, such as nesquehonite, can exhibit a bimodal release of CO_2 during stepped heating. Nitrocalcite gave a bimodal release of nitrogen; implying that the release pattern of nitrogen from the water-soluble material from Nakhla, and also in E1,149, might arise from a single species: 17 ppb with $\delta^{15}N$ approximately -11±4‰ in Nakhla and 1250 ppb with $\delta^{15}N$ approximately -8±5‰ in E1,149.

To explore further the possible occurrence of nitrates in Martian meteorites, data from two additional samples were also considered: E1,239 [*Wright et al.*, 1988b] and whole-rock Nakhla [*Wright et al.*, 1992]. For these samples, the combination of temperature increments utilized (50°C and 100°C) precludes the observation of a low-temperature bimodal release of nitrogen from a nitrate source. Thus, in order to facilitate comparison of data among the four separate analyses of the two meteorites, all results have been normalized to a consistent heating program. Nitrogen yields are compared in Figure 6a. For each analysis the highest nitrogen yield and lowest C/N ratio is from the 200-300°C step, following which there is a general decline in yield with increasing temperature. E1,239, known to be enriched in minerals with a low-temperature origin [*Gooding et al.*, 1988], has the highest nitrogen content. In Nakhla, if the low-temperature nitrogen in the whole-rock sample were all from a nitrate source, it should have been concentrated in the water extraction. Figure 6a indicates that this did not occur; nitrogen in the water-soluble material is approximately 3 orders of magnitude less abundant than in the whole-rock. One possible explanation is that any nitrates in Nakhla are present within the carbonate lattice, and hence insoluble in water. However, this seems unlikely, since carbonates generally decrepitate on stepped heating between 400°C and 700°C, and so coexisting nitrates would be liberated over the same temperature range, whereas most of the nitrogen in whole-rock Nakhla is released below 400°C.

A further constraint on the identity of the low-temperature components of the different fractions of E1 and Nakhla is provided by the nitrogen isotopic data. E1,149 and E1,239 show some similarities to each other (Figure 6b): data from E1,239 indicate that there are at least two nitrogen-bearing components released below 700°C. The first is liberated below 300°C ($\delta^{15}N$ about -14‰), comprising some 75% of the low-temperature nitrogen. A second, isotopically heavier, component is released above 300°C, with $\delta^{15}N$ between -6‰ and -1‰ (which, in itself, might be two separate nitrogen-bearing entities, as described by *Wright et al.* [1988b]). The similarities between E1,239 and E1,149 lead to the conclusion that if this component is a nitrate, then distribution of the material within E1 is quite widespread.

Whole-rock Nakhla contains low-temperature nitrogen with $\delta^{15}N$ in the range -7.6‰ to +7.1‰ (Figure 6c), again indicative of at least two components. The water-soluble material, on the other hand, releases nitrogen with an approximately constant isotopic composition, $\delta^{15}N$ about -11‰, implying that only one component is present. It is possible (indeed, probable) that the whole-rock sample contains some terrestrial contaminants (e.g., atmospheric nitrogen with $\delta^{15}N$ about 0‰ and organic materials with $\delta^{15}N$ about -5‰ to +10‰); these are presumably not transferred to the water extract. Thus, only indigenous nitrogen, i.e., nitrates with $\delta^{15}N$ about -11‰, are present in the water extract. On the basis of Figure 6a, perhaps as little as 0.1% of the total nitrogen released below 700°C in Nakhla whole-rock derives from indigenous nitrates.

Discussion and Implications

On Earth, three stages have been identified in the evolution of the nitrogen cycle [*Mancinelli and McKay*, 1988]. The second two stages occur after biological processes have become established, and are therefore probably not relevant to Mars. However, the first, prebiotic stage, might well apply: NO_x species are formed in the atmosphere by electrical discharge, cosmic ray interactions and by ultraviolet radiation, and subsequently fixed at the surface by implantation and exchange. Similar processes have been suggested for the formation of contemporary inorganic nitrate deposits in arid regions on Earth, such as Antarctica and Chile [e.g., *Claridge and Campbell*, 1968; *Mueller*, 1968; *Wada et al.*, 1981]. In these instances, however, the anions formed by discharge are washed out of the atmosphere during precipitation, then deposited by sublimation or evaporation as nitrates, rather than directly implanted as NO_xspecies. If nitrates were formed by similar mechanisms on Mars, then their isotopic composition should reflect that of the atmosphere from which they originated. It is apparent that very low levels of nitrate minerals occur in SNC's (17-1250 ppb nitrogen) with $\delta^{15}N$ in the range -8 to -14‰, isotopic compositions that do not reflect the present-day Martian atmosphere. Assuming that nitrates in E1 and Nakhla are not terrestrial contaminants, it is clear that they could not have formed on the Martian surface in recent times by incorporation of atmospheric species. Alternatively, the nitrates might have formed more recently by reactions of degassed magmatic volatiles ($\delta^{15}N$ about -20 to -10‰ [*Wright et al.*, 1992]) within the regolith.

Since nitrates are highly soluble salts and Mars has obviously experienced the effect of mobile water at various stages in its history, it is perhaps not surprising that nitrates are largely absent from specimens that came from the near-surface environment. It is possible that the salts dissolved in liquid water during a period when fluids percolated through the pore spaces of the Martian regolith. Any nitrates might therefore have been mobilized and either redeposited in a subsurface "hard-pan" layer, or become fro-

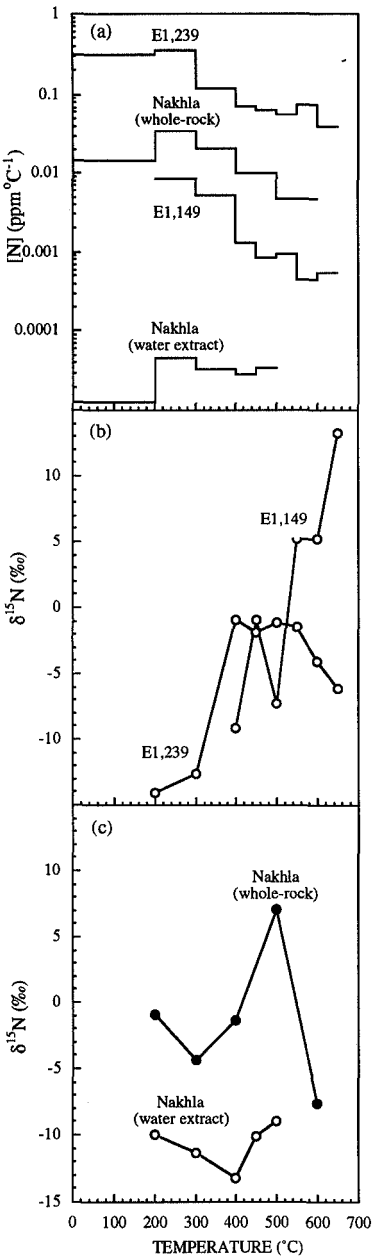


Figure 6. Stepped combustion data from room temperature to 700°C for the samples described herein, as well as another subsample of EET A79001 (E1,239) and Nakhla whole-rock. To facilitate comparisons, the data from each extraction have been calculated to a consistent temperature program. (a) Nitrogen yields for all four samples, (b) $\delta^{15}N$ for the two E1 samples, and (c) $\delta^{15}N$ for the two Nakhla samples.

zen within permafrost. Although such mechanisms are plausible, they do not account for the presence of other soluble species in SNC meteorites. Halite (NaCl), sylvite (KCl), and the soluble sulphate epsomite ($MgSO_4.7H_2O$) have been detected in Nakhla [*Gooding*, 1992]. Herein, halite was the only species identified by X ray diffraction of the water extract of Nakhla.

Under the current partial pressure of NO calculated for Mars' atmosphere, nitrates are predicted to be more stable than coexisting carbonate species [*Zolotov and Siderov*, 1986]. From the data acquired in this study, along with results from *Wright et al.* [1992] and *Wright and Pillinger* [1994], the relative proportion of carbonates to nitrates in E1 and Nakhla are calculated as~ 10 and 2200, respectively. The apparent deficit of nitrates in Nakhla might be due to the soluble salts having being washed out of the samples by the action of fluids; in E1 the same processes may have acted to concentrate the salts. Alternatively, the differences may reflect the admixture of a terrestrial carbonate in the Antarctic sample. Regardless of the difficulties of interpretation, it may be assumed that the carbonate/nitrate ratios are representative of the surface reservoirs in general. If, as has been modeled by *Yung et al.* [1977], there were a nitrate layer 0.3 cm thick across the surface of Mars, then the corresponding carbonate deposits could be anything from 3 to 600 cm thick. For comparison, *Kahn* [1985] has suggested up to 140 mbar of CO_2, equivalent to 280 cm of carbonate, stored in the regolith on Mars.

Acknowledgments. The Antarctic Meteorite Working Group is gratefully acknowledged for provision of the fractions from EET A79001. G. Jones and J. Gibson are thanked for assistance with analytical techniques. The paper benefited by reviews from R. H. Becker and A. J. T. Jull. This project was undertaken as part of the Cosmic Mineralogy Programme of the Natural History Museum (M.M.G.). Financial support to I.P.W. and C.T.P. was provided by the Science and Engineering Research Council (now PPARC).

References

Becker, R.H., and R.O. Pepin, The case for a Martian origin of the shergottites: nitrogen and noble gases in EETA 79001, *Earth Planet. Sci. Lett., 69*, 225-242, 1984.

Bogard, D.D., and P. Johnson, Martian atmospheric gases in an Antarctic meteorite?, *Science, 221*, 651-654, 1983.

Boyd, S.R., I.P. Wright, I.A. Franchi, and C.T. Pillinger, Preparation of sub-nanomole quantities of nitrogen gas for stable isotope analysis, *J. Phys. E Sci. Instrum., 21*, 876-885, 1988.

Burgess, R., I.P. Wright, and C.T. Pillinger, Distribution of sulphides and oxidised sulphur components in SNC meteorites, *Earth Planet. Sci. Lett., 93*, 314-320, 1989.

Carr, R.H., M.M. Grady, I.P. Wright, and C.T. Pillinger, Martian atmospheric carbon dioxide and weathering products in SNC meteorites, *Nature, 314*, 248-250, 1985.

Claridge, G.G.C., and I.B. Campbell, Origin of nitrate deposits, *Nature, 217*, 428-430, 1968.

Douglas, C., I.P. Wright, M.M. Grady, C.S. Romanek, and C.T. Pillinger, Carbon isotopic measurements of third-generation salts from LEW 85320, *Meteoritics, 28*, 341-342, 1993.

Fanale, F.P., Martian volatiles: the degassing history and geochemical fate, *Icarus, 28*, 179-202, 1976.

Fox, J.L., The production and escape of nitrogen atoms on Mars, *J. Geophys. Res., 98*, 3297-3310, 1993.

Gooding, J.L., Soil mineralogy and chemistry on Mars: possible clues from salts and clays in SNC meteorites, *Icarus, 99*, 28-41, 1992.

Gooding, J.L., S.J. Wentworth, and M.E. Zolensky, Calcium carbonate and sulphate of possible extraterrestrial origin in the EETA 79001 meteorite, *Geochim. Cosmochim. Acta, 52*, 905-915, 1988.

Grady, M.M., I.P. Wright, I.A. Franchi, and C.T. Pillinger, Nitrates in SNCs: implications for the nitrogen cycle on Mars, (abstract), *Lunar Planet. Sci. Conf., 24*, 553-554, 1993.

Hill, C.A., Mineralogy of cave nitrates, *NSS Bull., 43*, 127-132, 1981.

Jones, G.C., and B. Jackson, *Infrared Transmission Spectra of Carbonate Minerals*, 150 pp., Chapman and Hall, London, 1993.

Jull, A.J.T., and D.J. Donahue, Terrestrial ^{14}C age of the Antarctic shergottite, EETA 79001, *Geochim. Cosmochim. Acta, 52*, 1309-1311, 1988.

Kahn, R., The evolution of CO_2 on Mars, *Icarus, 62*, 175-190, 1985.

Mancinelli, R.L., and C.P. McKay, The evolution of nitrogen cycling, *MBL Lect. Biol., 18*, 311-325, 1988.

McElroy, M.B., T.Y. Kong, and Y.L. Yung, Photochemistry and evolution of Mars' atmosphere: a Viking perspective, *J. Geophys. Res., 82*, 4379-4388, 1977.

McSween, H.Y., What we have learned about Mars from SNC meteorites, *Meteoritics, 29*, 757-779, 1994.

Mueller, G., Genetic histories of nitrate deposits from Antarctica and Chile, *Nature, 219*, 1131-1134, 1968.

Nier, A.O., M.B. McElroy, and Y.L. Yung, Isotopic composition of the Martian atmosphere, *Science, 194*, 68-70, 1976.

Owen, T., The composition and early history of the atmosphere of Mars, in *Mars*, edited by H.H. Kieffer, B.M. Jakosky, C.W. Snyder and M.S. Matthews, pp. 818-834, University of Arizona Press, Tucson, 1992.

Pepin, R.O., On the origin and early evolution of terrestrial planet atmospheres and meteoritic volatiles, *Icarus, 92*, 2-79, 1991.

Plumb, R., R.F. Thivierge Jr., and W.W. Xu, Identifying acid salts of magnesium, *J. Phys. Chem., 91*, 6074-6076, 1987.

Pollack, J.B., T. Roush, F. Witteborn, J. Bregman, D. Wooden, C. Stoker, O.B. Toon, D. Rank, B. Dalton, and R. Freedman, Thermal emission spectra of Mars (5.4-10.5μm): evidence for sulfates, carbonates, and hydrates, *J. Geophys. Res., 95*, 14,595-14,627, 1990.

Tanaka, K.L., The stratigraphy of Mars, *Proc. Lunar Planet. Sci. Conf., 17th, Part 1, J. Geophys. Res., 91, suppl.*, E139-E158, 1986.

Wada, E., R. Shibata, and T. Torii, ^{15}N abundance in Antarctica: origin of soil nitrogen and ecological implications, *Nature, 292*, 327-329, 1981.

Wallis, M.K., C, N, O isotope fractionation on Mars: implications for crustal H_2O and SNC meteorites, *Earth Planet. Sci. Lett., 93*, 321-324, 1989.

Wiens, R.C., Noble gases released by vacuum crushing of EETA 79001 glass, *Earth Planet. Sci. Lett., 93*, 55-65, 1988.

Wright, I.P., and C.T. Pillinger, On the isotopic chemistry of carbon at the Martian surface, *Philos. Trans. R. Soc. London* (in press), 1994.

Wright, I.P., S.R. Boyd, I.A. Franchi, and C.T. Pillinger, High precision determination of nitrogen stable isotope ratios at the subnanomole level, *J. Phys. E Sci. Instrum., 21*, 865-875, 1988a.

Wright, I.P., M.M. Grady, and C.T. Pillinger, Carbon, oxygen and nitrogen isotopic compositions of possible Martian weathering products in EETA 79001, *Geochim. Cosmochim. Acta, 52*, 917-924, 1988b.

Wright, I.P., M.M. Grady, and C.T. Pillinger, The evolution of atmospheric CO_2 on Mars: the perspective from carbon isotope measurements, *J. Geophys. Res., 95*, 14,789-14,794, 1990.

Wright, I.P., M.M. Grady, and C.T. Pillinger, Chassigny and the nakhlites carbon-bearing components and their relationship to Martian environmental conditions, *Geochim. Cosmochim. Acta, 56*, 817-826, 1992.

Yung, Y.L., D.F. Strobel, T.Y. Kong, and M.B. McElroy, Photochemistry of nitrogen in the Martian atmosphere, *Icarus, 30*, 26-41, 1977.

Zolotov, M.Y., and Y.I. Siderov, Nitrates in Martian soil? (abstract), *Lunar Planet. Sci. Conf., 17*, 975-976, 1986.

M. M. Grady, Department of Mineralogy, The Natural History Museum, Cromwell Road, London, England, United Kingdom. (e-mail: M.Grady@nhm.ac.uk)

C. T. Pillinger and I. P. Wright, Department of Earth Sciences, The Open University, Walton Hall, Milton Keynes, MK7 6AA, England, United Kingdom. (e-mail: I.P.Wright@open.ac.uk)

(Received April 4, 1994; revised October 10, 1994; accepted October 27, 1994.)

Page intentionally left blank

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5457-5464, MARCH 25, 1995

A slightly more massive young Sun as an explanation for warm temperatures on early Mars

Daniel P. Whitmire,[1] Laurance R. Doyle,[2] Ray T. Reynolds,[3] and John J. Matese [1]

Abstract. The valley network channels on the heavily cratered ancient surface of Mars suggest the presence of liquid water ~3.8 Gyr ago. However, the implied warm climate is difficult to explain in the context of the standard solar model, even allowing for the maximum CO_2 greenhouse heating. In this paper we investigate the astronomical and planetary implications of a nonstandard solar model in which the zero-age, main-sequence Sun had a mass of 1.05 ± 0.02 $M_\odot$. The excess mass was subsequently lost in a solar wind during the first $1.2^{-0.2}_{+0.4}$ Gyr of the Sun's main sequence phase. The implied mass-loss rate of $4^{+3}_{-2} \times 10^{-11}$ $M_\odot yr^{-1}$, or about $10^3\times$ that of the current Sun, may be detectable in several nearby young solar type stars.

1. Introduction and Summary

Because of the evidence that early Mars was warmer than can be easily explained in the context of the standard solar model, we have investigated several astronomical models that could in principle increase the solar flux at early Mars. The most plausible of these is one in which the early main sequence Sun was somewhat more massive than today's Sun. (Other, less promising, astronomical mechanisms for heating early Mars are considered briefly in the appendix.) Because the solar flux at a planet is extremely sensitive to mass ($F \propto M^{6.75}$), the required minimum mass of the young Sun is only 1.03 $M_\odot$. An upper limit of 1.07 $M_\odot$ is necessary to prevent the loss of water from Earth. We first discuss the constant mass-loss model and the climatic constraints imposed on it. We then consider the nonclimatic astronomical and planetary implications of the model and find that it is consistent with existing planetary and astronomical data. However, a convincing case for the model can be made only if similar mass-loss rates ($\geq 2 \times 10^{-11} M_\odot yr^{-1}$) are found to be a common feature of other young solar-type stars, or by the discovery (or reinterpretation) of supporting meteoritic and lunar implantation data. The recent detection [*Güdel et al.*, 1994] of continuous microwave emission from four nearby G-type dwarfs would, if interpreted in terms of free-free emission from an optically thick wind [*Mullan et al.*, 1992], imply mass-loss rates of $\sim 10^{-9}$ $M_\odot yr^{-1}$. This interpretation can be tested by additional measurements at millimeter wavelengths.

2. Background

Assuming the standard solar model, *Kasting* [1991] has shown that the maximum CO_2 greenhouse warming is insufficient to raise the early Martian mean surface temperature to near 273 K, although CO_2 can resolve the faint young Sun paradox on Earth [*Kasting and Grinspoon*, 1991]. Models employing the greenhouse gases methane or ammonia have generally been unsuccessful, as these gases have short photodissociation lifetimes. However, a shielding particulate haze similar to that observed on Titan may mitigate this objection [*Chyba and Sagan*, 1994]. Some geothermal heating is expected [*Squyres*, 1993], but this is not likely to completely explain temperatures required for surface liquid water. Even if the channels were formed by subsurface sapping of groundwater, Martian surface temperatures must have been significantly higher than today to allow groundwater to be mobile near the surface and to recharge the aquifer sources. Controversial evidence of glacial markings would, if substantiated, also indicate high temperatures since precipitation is implied. The issue of warm temperatures on early Mars is also important for the determination of the size of circumstellar habitable zones [*Kasting et al.*, 1993; *Whitmire and Reynolds*, 1995] and therefore the broader topic of life in the universe.

[1]Department of Physics, University of Southwestern Louisiana, Lafayette.
[2]SETI Institute, Mountain View, California.
[3]Theoretical Studies Branch, NASA-Ames, Moffett Field, California.

Paper number 94JE03080.
0148-0227/95/94JE-03080$05.00

3. Climatic Constraints on the Mass of the Early Sun

In standard constant-mass stellar models the mass-luminosity relation for stars near 1 $M_\odot$ is given by *Iben* [1967] as

$$L \propto M^{4.75}$$

A more recent evolutionary calculation of a mass losing 1.1 $M_\odot$ star gives an effective index of about 4.5 [*Boothroyd et al.*, 1991]. However, observations of G stars in detached binary systems give an index of 4.8 near 1.0 $M_\odot$ [*Popper et al.*, 1986]. (The mass-luminosity index has its maximum value near 1.0 $M_\odot$.) Recognizing that the precise value of the index is likely to depend on composition and age [*Henry and McCarthy*, 1993], we adopt a value of 4.75 and assume that this is also an adequate approximation when the star's mass is slowly decreasing.

In the standard model the evolution of the luminosity is given as an analytic function of time up to the present by *Gough* [1981]. Using this function and the mass-luminosity relation we model the luminosity of a mass-losing Sun as [*Whitmire et al.*, 1995]

$$L(t) = \frac{L_o}{1 - \frac{2t}{7t_\odot}} \left[\frac{M(t)}{M_o}\right]^{4.75}$$

where $L_o = 0.71(M_o/M_\odot)^{4.75}L_\odot$ is the luminosity at time $t = 0$, M_o is the mass at $t = 0$ and $t_\odot = 4.6$ Gyr. We model the solar mass $M(t)$ as a linear function of time (i.e., constant mass loss) given by

$$M(t) = M_o(1 - \Delta\tfrac{t}{\tau}) \quad 0 \le t \le \tau$$

$$M(t) = M_\odot \quad \tau \le t \le t_\odot$$

where τ is the mass-loss timescale and $\Delta = (M_o - M_\odot)/M_o$ is the mass-loss fraction. The relation between a planet's semimajor axis $a(t)$ and the solar mass is obtained from conservation of angular momentum (or its equivalent, the first adiabatic invariant equation for J_3 discussed below). The result is $a(t) \propto M^{-1}$, and therefore the flux at a planet $F(t) = L(t)/4\pi a^2(t)$ is

$$F(t) = \frac{F_o}{1 - \frac{2t}{7t_\odot}}(1 - \Delta\tfrac{t}{\tau})^{6.75} \quad 0 \le t \le \tau$$

$$F(t) = \frac{F_o}{1 - \frac{2t}{7t_\odot}}(1 - \Delta)^{6.75} \quad \tau \le t \le t_\odot$$

where $F_o = 0.71[1\mathrm{AU}/a(t_\odot)]^2[M_o/M_\odot]^{6.75}$ is the flux at the planet at $t = 0$, located today at $a(t_\odot)$, in units of the present solar flux at Earth.

We now impose two independent planetary climate constraints on the flux model. First, we require that the flux at Earth at $t \approx 0$ (when it is a maximum during the first 4.6 Gyr) not be so great that all planetary H_2O would be lost. This would occur through the moist greenhouse effect if the flux exceeded 1.1, in units of the flux at Earth today [*Kasting*, 1988]. A moist greenhouse occurs when the stratosphere becomes wet and H_2O is lost through UV dissociation and the subsequent loss of the hydrogen to space. In the climate model used to calculate this flux, clouds were not fully taken into account (only their albedo was modeled) and the actual critical flux could be somewhat greater than this, but probably not as high as the runaway greenhouse critical flux of 1.4, for which the oceans evaporate entirely. A critical flux of 1.1 limits the Sun's zero-age, main-sequence (ZAMS) mass to ≤ 1.07 $M_\odot$, independent of the mass-loss timescale τ. This upper mass limit is fairly insensitive to the precise value of the critical flux, being proportional to (critical flux)$^{0.148}$.

The suggestion by *Guzik et al.* [1987] that the Sun may have had twice its present mass is inconsistent with the climate history of Earth (predicting a t = 0 flux of $0.71 \times 2^{6.75} = 76$) unless the mass was all lost before the terrestrial planets accreted (i.e., prior to the ZAMS). *Graedel et al.* [1991] have suggested that an initial solar mass of 1.1-1.2 $M_\odot$ could explain the faint young Sun paradox [*Sagan and Mullen*, 1972] for Earth. However, this mass range is beyond the moist greenhouse limit. Warm temperatures on early Earth can be explained by an expected large CO_2 greenhouse [*Kasting and Ackerman*, 1986; *Kasting and Grinspoon*, 1991] and/or the enhanced flux of an early Sun of mass $\le 1.07M_\odot$.

The second climate constraint is based on the assumption that annual average surface temperatures on Mars should have been at least 273 K at the end of the late heavy bombardment about 3.8 Gyr ago ($t \approx 0.8$ Gyr). The drainage channels and valley networks, which indicate the flow of liquid water, are found primarily in the highly cratered southern highlands and are therefore assumed to date to the end of the period of the late heavy bombardment [*Pollack et al.*, 1987]. This translates to a flux that is at least 13% greater ($= 1.13 \times 0.32$) than predicted by the standard solar model at $t \approx 0.8$ Gyr [*Kasting, 1991; Kasting et al.*, 1993]. From the flux equation $F(t)$, it can be shown that a 13% increase in flux at $t = 0.8$ Gyr requires an initial solar mass of 1.03 $M_\odot$ if we assume that the maximum value of τ is 1.6 Gyr. This timescale limit is based on the apparent absence of any ion implantation evidence of a strong wind more recent than ~3 Gyr ago. Inserting the above upper mass limit of 1.07 $M_\odot$ (determined from Earth's climate history) into the flux equation (as applied to Mars) yields a minimum value of $\tau = 1.0$ Gyr.

We conclude that a ZAMS solar mass of 1.05 ± 0.02 $M_\odot$, corresponding to mass-loss timescales of $1.2^{-0.2}_{+0.4}$ Gyr, can explain the existence of liquid water on early Mars and avoid the loss of water from Earth. The corresponding mass-loss rates are $4^{+3}_{-2} \times 10^{-11} M_\odot \mathrm{yr}^{-1}$, or about $10^3\times$ that of the current Sun. These results are not very sensitive to uncertainties in the critical cli-

matic fluxes assumed. If mass loss is assumed to be a monotonically decreasing function of time the lower mass limit would be increased but the upper limit would be unaffected. Next we consider other, nonclimatic implications of the model.

4. Nonclimatic Implications of a More Massive Early Sun

4.1. Astronomical

4.1.1. Observed stellar mass loss from main sequence dwarfs. The direct observation of mass-loss rates less than about $10^{-9}M_{\odot}yr^{-1}$ is difficult in solar-type dwarfs [*Brown et al.*, 1990]. However, because of the presence of a UV continuum from a white dwarf companion, significant mass loss has been detected by *Mullan et al.* [1989] from the K2V dwarf in the eclipsing detached binary V471 Tauri in the Hyades cluster. The measured mass-loss rate was $\sim 10^3$ times that of the present Sun. If this wind has been constant or decreasing during the age of the Hyades, the implied mass loss to date would be ≥ 0.01 $M_{\odot}$. The K star's spin should be tidally synchronized with its orbital period of 0.5 day. This rotational period is in the range of those observed in young solar-type stars [*Stauffer and Soderblom*, 1991], suggesting that for isolated main sequence stars, significant mass loss is associated with rapid rotation. Observations show that by the age of the Hyades ($\sim$0.6 Gyr) this initial very fast rotation has been reduced, though typically still $\sim 5 \times$ that of the Sun today.

Mullan et al. [1992] have reported winds $\geq 10^{-10}$ $M_{\odot}$ yr^{-1} from several M dwarfs. Evidence of mass loss was based on measurements over a wide range of radio, millimeter, and infrared wavelengths and the fact that free-free emission from an optically thick wind has a well established characteristic spectrum in which the flux is proportional to $\nu^{\frac{2}{3}}$, where ν is the frequency. There are several nearby young GO stars that may be good candidates for similar observations. The age of HD39587 is determined from its membership in the Ursa Majoris cluster ($\leq$200 m.y.). This star is known to have one or two low-mass companions. The apparently single star HD72905 is also a good candidate. However, although its rapid rotation (period 4.7 days) and CaII H and K emission line flux indicate an age of about 700 m.y., the errors on its age determination are rather large. The stars HD115383 (rotation period 3.5 days) and HD206860 ($V \sin i = 10.2$ km s^{-1}) are also within the potential detection range but again more stringent age constraints are needed.

If significant mass loss is occurring in these young stars, there will be an excess flux given in mJy by [*Leitherer and Robert, 1991; Wright and Barlow, 1975; Panagia and Felli, 1975*]

$$S_{\nu} = 2.32 \times 10^{10} (\dot{M}Z)^{4/3} (\gamma g_{\nu} \nu)^{2/3} (v_{\infty}\mu)^{-4/3} d^{-2}$$

where $\dot{M}$ is the stellar mass-loss rate in $M_{\odot}yr^{-1}$, Z is the rms ionic charge of the wind particles, γ is the number of electrons per ion, g_{ν} is the free-free Gaunt factor, ν is the frequency in Hz, v_{∞} is the stellar wind velocity in km s^{-1}, μ is the mean molecular weight of wind particles, and d is the distance to the star in parsecs (1 parsec = 3.2 light years). For a solar-type stellar wind, $Z \approx \gamma \approx \mu \approx 1$. The Gaunt factor is given by $g_{\nu} = 9.77[1 + 0.13 \log(T_e^{3/2}/Z\nu)]$, where T_e is the stellar wind temperature in degrees Kelvin [*Leitherer and Robert, 1991; Abbott et al.*, 1986]. In the derivation of S_{ν}, it was assumed that the wind was isothermal, spherically symmetric, and the outflow was stationary with constant velocity. The dependence of the flux on frequency was found not to be very sensitive to changes in spherical symmetry and wind gradients in temperature, velocity, and density [*Leitherer and Robert*, 1991].

In preliminary work, we have made long integration-time observations of two of the four candidate stars (HD39587 and HD72905) using the National Radio Astronomy Observatory very large array in Socorro, New Mexico. These observations showed that mass loss could be taking place in the HD39587 system but, unfortunately, the flux levels were confused by the unresolved nearby smaller-mass companions, and further observations are required to ensure that the source of the mass loss is the G star. HD72905 shows significant flux at 8.0 GHz, but very little at 4.8 GHz. It appears that the slope of this flux spectrum is too steep to be indicative of mass loss, but further deconvolution is needed. As noted above, the age of this star is relatively uncertain. We have not observed HD115383 or HD206860 long enough to draw any conclusions regarding mass-loss rates less than about 2×10^{-9} $M_{\odot}$ yr^{-1}.

Recently, *Güdel et al.* [1994] have reported the discovery of 8.5-GHz microwave emission from four nearby "solar-twin" G dwarfs selected from a set of 15 G dwarfs with X ray luminosity 100-800 times greater than the quiet Sun. The largest radio flux (0.338 mJy; 1Jy = 10^{-26} Wm^{-2} Hz^{-1}) and luminosity were found in the young GOV star HD129333 (EK Draconis), which is believed to be a member of the Pleiades cluster (age $\approx$70 m.y.) and is often studied as an infant-Sun proxy [*Dorren and Guinan*, 1994]. A slightly lower radio luminosity was found for Gl 97.0 (age 2 Gyr) and the other two G stars, one of which (Gl 755) had no age estimate and the other (HD 225239) was tentatively dated at a surprising $\sim$10 Gyr. The emission mechanism is uncertain, since spectra are not yet available. If we interpret these observations in terms of mass loss, the implied mass-loss rates are $\sim 10^{-9}$ $M_{\odot}$ yr^{-1}.

The theoretical arguments (e.g., high coronae temperatures) given by *Mullan et al.* [1992] as motivation for searching for mass loss in M dwarfs also apply to these four solar-like G stars. It appears that the mass-loss model can be tested in these stars by measurements made at millimeter wavelengths. These measurements would test the predicted $\nu^{2/3}$ dependence of the flux

over a large frequency interval from the existing 8.5 GHz to ≈300 GHz. The predicted 1-mm flux for HD129333 is $\sim(300/8.5)^{2/3} \times 0.338$ mJy ~3.6 mJy. The M star fluxes measured at 1 mm by Mullan et al. were ~10 mJy.

4.1.2. Lithium depletion. The element lithium is known to be depleted by two orders of magnitude in the photosphere of the Sun relative to its primordial abundance. *Weymann and Sears* [1965] suggested that this depletion could be explained if the Sun was originally somewhat more massive than 1.0 $M_\odot$. This would allow the lithium to be convectively mixed into regions with temperatures high enough to destroy it. This idea was independently rediscovered and extended by *Hobbs et. al* [1989], who found that an initial mass of 1.04 $M_\odot$ would result in the removal of 99% of the initial lithium. Using the latest nuclear reaction rates and opacities, *Boothroyd et al.* [1991] computed detailed solar models with an initial mass of 1.1-1.2 $M_\odot$. They found that this mass resulted in the desired depletion of lithium while at the same time not over depleting the element beryllium. Further indirect support for this explanation of lithium depletion in the Sun has recently been reported [*Guzik and Cox*, 1993]. *Soderblom et al.*, [1990] found that if mass loss is the explanation for lithium depletion in F and G dwarfs in the Hyades cluster, then the implied mass loss would be ≈0.07 $M_\odot$ and the timescale for mass loss would be ≈0.6 Gyr.

On the other hand, *Swenson et al.* [1994] find that, although mass loss can be a marginal explanation for the solar lithium problem, it cannot be a significant contributor to the depletion of lithium in Hyades G dwarfs. Rotational mixing associated with magnetic braking on the main sequence has been proposed [*Belcher and MacGregor*, 1976] as an alternative model to explain the lithium depletion. Other mechanisms have also been proposed [e.g., *Swenson et al.*, 1994; D'Antona and Mazzitelli, 1984] and are further discussed in the work by *Boothroyd et al.* [1991].

4.1.3. Angular momentum loss. Magnetic braking, which requires an associated stellar wind, is believed to be the explanation for the observed spin down of young solar-type stars after they have arrived on the main sequence [*Stauffer, 1995; Soderblom et al., 1990; Belcher and MacGregor, 1976*]. However, in many models, the required change in angular momentum is insensitive to the mass loss [*Pinsonneault et al.*, 1989], depending instead almost entirely on the magnetic field and stellar wind velocity. An exception is the model of *Bohigas et al.* [1986], in which the total mass loss was $\sim 10^{-3}$ $M_\odot$, assuming that the convective envelope was decoupled from the interior of the star. As noted by *Graedel et al.* [1991], helioseismic observations suggest that the Sun rotates more nearly like a solid body and, consequently, the implied mass loss should be increased by a factor of 20 (= moment of inertia ratio), giving a value roughly consistent with the present model.

Assuming the ZAMS Sun rotated at a rate typical of solar-type stars in the Pleiades cluster, the loss of rotational energy suggests a minimum fractional mass loss of $\sim k(V_o/V_{sw})^2$, where V_o is the ZAMS solar equatorial velocity, ~20-100 km s^{-1} [*Stauffer and Soderblom*, 1991], V_{sw} is the average early wind velocity at infinity, and k is the moment of inertia constant. Unfortunately, V_{sw} is unknown, but may have been lower than the current average value of ≈450 km s^{-1} [*Geiss and Bochsler*, 1991]. Observed stellar wind velocities range from tens of km s^{-1} for red giants to greater than $\sim 10^3$ km s^{-1} for early-type stars.

4.2. Planetary

4.2.1. Planet and planetesimal migration. The effect of slow radial mass loss on planet orbits can be deduced from the adiabatic invariance of the action variables,

$$\begin{aligned} J_3 &= 2\pi m\sqrt{GMa} \\ J_2 &= J_3\sqrt{1-e^2} \\ J_1 &= J_2 \cos i \end{aligned}$$

where $M(m)$ is the solar (planet) mass and a is the semimajor axis. It can be seen that the planetary eccentricities, e, and inclinations, i, are constant during adiabatic mass loss but, as noted earlier, the semimajor axes increase inversely with stellar mass, assuming planetary accretion is negligible.

During the Sun's mass loss phase, any remaining unaccreted planetesimals would migrate at rates different than the planets due to dynamical drag. To model the planetesimal migration process we assume that the early solar wind was radial and isotropic and carried no significant angular momentum. A planetesimal of radius r would accrete mass from the solar wind at the rate $-(dM/dt)(\pi r^2/4\pi a^2)$. The composition of the solar wind is mostly hydrogen and helium, so almost all of the accreted mass would rapidly volatilize at the same rate and escape from the orbiting planetesimal with negligible thermal speeds; consequently, we take the planetesimal mass and radius to be constant. With these assumptions a straightforward angular momentum analysis results in the differential equation [*Whitmire et al.*, 1991]

$$\frac{da^2}{dM} = -\left(\frac{2a^2}{M} - \frac{r^2}{m}\right)$$

the solution of which is

$$a^2 = a_o^2\left(\frac{M_o}{M}\right)^2 + \frac{M}{4\pi\rho r}\left[1-\left(\frac{M_o}{M}\right)^3\right]$$

or, when the final mass $M = M_\odot$

$$\frac{a}{a_o} \approx 1 + \Delta\left(1 - \frac{1}{\rho r a_o^2}\right)$$

to lowest order in the mass-loss fraction Δ. In the latter

expression ρ is the planetesimal density in g cm^{-3}, r is the planetesimal radius in units of 10 km, and a_o is the initial semimajor axis in units of AU.

In our inelastic-collision/volatilization model the migration is independent of the solar wind velocity and the functional form of $M(t)$, to first order in Δ. Different assumptions about the angular momentum content of the solar wind and the planetesimal accretion/emission mass rates will lead to different migration formulae. For example, if elastic collisions or sputtering are assumed to be important, the migration formulae will be somewhat different. However, independent of the model, there will, in general, be differential migration between planetesimals and planets.

It is evident from the above migration equation that the drag term is negligible for planets, but not necessarily for planetesimals. The outward migrating planets would perturb any remaining unaccreted planetesimals lying within an original radial band that can be determined by applying the migration equation to both planets and planetesimals. The differential migration between a planet and planetesimal near a_o is then found to be

$$\delta a \approx \frac{\Delta}{\rho r a_o} \quad \text{AU}$$

Assuming a density of 3.0 (0.5) g cm^{-3} for terrestrial (Jovian) planetesimals, and taking $\Delta \sim 0.1$, the total band cleared by all planets is $\sim$0.2 AU for planetesimals of radius 10 km. Resonant pumping of planetesimal eccentricities may substantially multiply the total cleared bandwidth. The resulting equilibrium eccentricity in our model would be much greater than that due to drag in the solar nebula [*Weidenschilling and Davis*, 1985] because of our much smaller drag (or effective density). If the mass outflow stopped in less than $\sim$200 m.y., the result might resemble the late heavy bombardment on the Moon [e.g., *Wetherill*, 1975]. Although the crater counts are consistent with the occurrence of a unique event about 3.8 Gyr ago, they are also consistent with the relatively sudden end of an extended post accretion period of enhanced impacts. It remains to be shown whether the predicted crater size distribution is consistent with the observed distributions on the Moon and outer solar system satellites.

If warmer temperatures on early Mars were the result of a more massive Sun, and if the period of the late heavy bombardment was related to mass loss and migration, then the association on Mars between the heavily cratered ancient surface and the valley network channels would be expected. If we make this association, then the mass-loss timescale τ is further restricted to $\sim$1 Gyr. For consistency this requires the initial solar mass to be near the upper limit of 1.07 $M_\odot$. The heavy bombardment should have continued somewhat beyond the end of the warm climate period. Assuming $\tau \sim 1$ Gyr, the late heavy bombardment would have continued for some $\sim$0.2 Gyr after the creation of most of the valley network channels. We emphasize that our model does not require an association with the late heavy bombardment.

4.2.2. Ion implantation. The ion implantation record of the solar wind and of the higher-energy flare nuclei is primarily limited to epochs more recent than $\sim$3 Gyr ago, during which time span there is evidence for as much as an order of magnitude higher average particle flux than at present [e.g., *Geiss and Bochsler*, 1991]. In our model, the rate at earlier times would need to be $\sim 10^3\times$ the present rate of $\sim 3 \times 10^{-14}$ $M_\odot$ yr^{-1}. This is not necessarily incompatible with the ion implantation record, however, since the measured quantity is fluence and not flux. In a given meteorite or lunar rock it is generally not possible to experimentally determine the exposure time to the solar wind, or to its higher-energy component. Some older, solar-flare irradiated grains from independent meteorites imply an early flare activity $\sim 10^3\times$ that of the present Sun [*Caffee et al.*, 1987], assuming that these grains were exposed for $\sim 10^5$ years on the meteorite's parent body regolith prior to protective compaction (burial). The associated solar wind may also have been enhanced by a comparable amount. Caffee et al. interpret this finding in terms of pre-main sequence T Tauri flare activity. However, their results may also be consistent with early main sequence flare activity if final compaction inside the meteorites' parent bodies occurred later than 4.6 Gyr ago. The short range of solar flare heavy ions ($\leq$1 mm) and wind-implanted ions (≤ 1 μm) in meteorites show that the grains were exposed at a time when there was very little intervening material between the Sun and the grains [*Walter and Barry*, 1991]. This is most consistent with post-T Tauri irradiation. Further, *Papanastassiou et al.* [1974] studied two igneous clasts in Kapoeta (the breccia on which Caffee et al. based their argument) and found them to have relatively young Rb-Sr isochron ages (3.89±0.05 and 3.63±0.08 Gyr). They concluded that implanted solar wind gases and tracks in the breccia must record events during the first 1 Gyr after the formation of the solar system, and not early solar system formation processes [*Wood and Pellas*, 1991]. Finally, we note that if the early main sequence mass loss was bipolar, as observed in many pre-main sequence T Tauri stars, there may be no meteoritic or lunar implantation record of it.

Appendix: Other Astronomical Mechanisms

Here we comment briefly on several other astronomical mechanisms for heating early Mars in increasing order of plausibility [*Whitmire et al.*, 1995].

A.1. Interstellar Cloud Accretion

Accretion of interstellar matter onto the Sun would increase the solar luminosity due to the release of gravi-

tational potential energy. However, this seems to be an unlikely possibility, since for a $\geq$1% increase in luminosity, the required interstellar cloud hydrogen number density is $\geq 10^4$ cm^{-3} for a relative velocity $\leq$10 km s^{-1} [*McCrea*, 1975]. The probability of simultaneously satisfying both of these requirements in 4.6 Gyr is small. Further, for any plausible relative velocity, the resulting transient time through such a dense cloud is likely to be less than the time required to form the valley network channels on Mars.

A.2. Perturbation by a Third Body

The flux at the surface of a planet averaged over an orbital period is

$$\overline{F} \propto \overline{\frac{1}{r^2}} \propto \frac{1}{a^2\sqrt{1-e^2}}$$

where a is the semimajor axis and e is the eccentricity. The eccentricity of Mars' orbit can vary from today's value of 0.093 to 0.2 [*Laskar*, 1994]. The maximum increase in average flux due to this variation is only ~2%. The variation in flux between perihelion and aphelion in a single orbit is proportional to $[(1+e)/(1-e)]^2$, which could be as large as 2.2. The degree to which the thermal inertia of a dense early Martian atmosphere and/or oceans would mitigate this large annual increase in solar flux remains to be investigated.

Although the average flux is not sensitive to eccentricity, it is sensitive to the semimajor axis. The required flux enhancement of $\geq$13% (more than the standard solar model value) corresponds to a semimajor axis $\leq$1.43 AU. If Mars accreted at this distance from the Sun and was subsequently perturbed to its present orbit by another body (or through solar system nonlinear dynamical effects), its present orbit would be expected to overlap its original orbit. Today, Mars' perihelion distance (1.38 AU) is consistent with this expectation. Although probably untestable, the occurrence of such a perturbation is not totally ad hoc. The main objection is that the required close encounter would have had to occurred so long after planetary accretion was over.

A.3. Variations in Solar Luminosity

In young solar proxy stars, photometric brightness and magnetic changes are anticorrelated [*Baliunas, 1991; Radick et al.*, 1990]. Although an increase in photometric brightness does not necessarily imply an increase in bolometric luminosity, it is unlikely that there is no correlation. The near cancellation of the decrease in luminosity due to dark spots by the increase in luminosity due to faculae in the Sun is not representative of younger and more magnetically active solar-type stars [*Foukal*, 1994]. For these stars it is likely that spots result in net luminosity variations 10-100 times larger than the current solar variation [*Radick*, 1991]. Young solar-type stars are expected to be heavily spotted, up to several tens-of-percent coverage. In equilibrium, the star's luminosity is fixed by core nuclear reactions, and so the temperature of the radiating surface is higher than it would be without spots. If the spots disappeared on a timescale $\leq 10^5$ years the newly exposed area would result in an increase in total luminosity given by $L = L_o(1+f)$, where L_o is the core luminosity and f is the fraction of the star's surface previously covered by spots [*Spruit*, 1982]. The luminosity would return to the core value on the convective envelope (thermal) timescale of $\sim 10^5$ years in the case of a G2 dwarf like the Sun. However, this timescale (like that for accretion) is probably too short to explain the Martian valley networks.

A.4. Ecliptic Flux Focusing

A related and more promising spot model is that of ecliptic focusing [*Whitmire et al.*, 1993]. Unlike the current Sun, photometric modeling of young solar-type stars in RS CVn systems, in particular, show a definite tendency for spots to cluster toward northern latitudes [e.g., *Rodonò et al.*, 1986]. The spotted stars in these binary systems are often young G or K dwarfs. In many cases these stars remain young in the sense that rapid rotation is enforced by spin-orbit synchronization. If spots tend to cluster in the polar regions over timescales of ~10^8 years, the luminosity is unaffected but the distribution of stellar flux emitted is altered. For young heavily spotted stars the temperature and emitted flux from the spotted regions decreases while that in the equatorial regions increases proportionately (this is not the case for the present Sun). We have investigated a simple model in which both poles are capped down to a stellar latitude b. The average enhancement in equatorial flux is then given by

$$\frac{F}{F_o} = \frac{2}{\pi}\left(\frac{\sin b \cos b + b}{\sin b}\right)$$

The theoretical maximum enhancement is a factor of $4/\pi$ for 100% coverage (i.e., $b = 0$, corresponding to an infinitesimal equatorial ring). To obtain the early Mars minimum enhancement of 1.13 would require a coverage of ≈30%. Taking into account limb darkening would reduce the required coverage somewhat. Additional photometric observations of single G dwarfs at the age of the Hyades are needed to see if the amount of coverage at this age is consistent with the focusing model.

Acknowledgments. We thank Dana Backman, Dermont Mullan and Pat Whitman for valuable suggestions. D.P.W. and J.J.M. acknowledge the Louisiana Educational Quality Support Fund for partial support of this work.

References

Abbott, D.C., J.H. Bieging, E. Churchwell and A.V. Torres, Radio emission from galactic Wolf-Rayet stars and the structure of Wolf-Rayet winds, *Astrophys. J.*, 303, 239-261, 1986.

Baliunas, S., The past, present and future of solar magnetism: Stellar magnetic activity, in *The Sun in Time*, edited by C.P. Sonett, M.S. Giampapa and M.S. Matthews, pp. 809-831, University Arizona Press, Tucson, 1991.

Belcher, J.W. and K.B. MacGregor, Magnetic acceleration of winds from solar-type stars, *Astrophys. J.*, 210, 498-507, 1976.

Bohigas, J., L. Carrasco, C. Torres, and G. Quast, Rotational braking of late-type main sequence stars, *Astron. Astrophys.*, 157, 278-292, 1986.

Boothroyd, A.I., I.J. Sackmann, and W.A. Fowler, Our sun. II. Early mass loss of 0.1 $M_\odot$ and the case of the missing lithium, *Astrophys. J.*, 377, 318-329, 1991.

Brown, A., A. Vealé, P. Judge, J. Bookbinder, and I. Hubeny, Stringent limits on the ionized mass loss from A and F dwarfs, *Astrophys. J.*, 361, 220-224, 1990.

Caffee M., C. Hohenberg, and T. Swindle, Evidence in meteorites for an active early sun, *Astrophys. J.*, 313, L31-L35, 1987.

Chyba, C. and C. Sagan, in *Proceedings of the First International Conference on Circumstellar Habitable Zones*, edited by L. Doyle, Travis House Pub., Menlo Park, CA, in press, 1995.

D'Antona, F. and I. Mazzitelli, Lithium depletion in stars. Pre-main sequence burning and extra-mixing, *Astron. Astrophys.*, 138, 431-442, 1984.

Dorren, J.D. and E.F. Guinan, HD129333: The Sun in its infancy, *Astrophys. J.*, 428, 805-818, 1994.

Foukal, P., Stellar luminosity variations and global warming, *Science*, 264, 238-239, 1994.

Geiss, J. and P. Bochsler, Long time variations in solar wind properties: Possible causes versus observations, in *The Sun in Time*, edited by C.P. Sonett, M.S. Giampapa and M.S. Matthews, pp. 98-117, University Arizona Press, Tucson, 1991.

Gough, D.O., Solar interior structure and luminosity variations, *Solar Phys.*, 74, 21-34, 1981.

Graedel, T., I. Sackmann, and A. Boothroyd, Early solar mass loss: A potential solution to the weak sun paradox, *Geophys. Res. Lett.*, 18, 1881-1884, 1991.

Güdel, M., J.H.M.M. Schmitt, and A.O. Benz, Discovery of microwave emission from four nearby solar-type G stars *Science*, 265, 933-935, 1994.

Guzik, J.A., L.A. Wilson, and W.M. Brunish, A comparison between mass-losing and standard solar models, *Astrophys. J.*, 319, 957-965, 1987.

Guzik, J.A. and A.N. Cox, Using solar p-modes to determine the convection zone depth and constrain diffusion-produced composition gradients, *Astrophys. J.*, 411, 394-401, 1993.

Henry, T.J. and D.W. McCarthy, The mass-luminosity relation for stars of mass 1.0 to 0.08 $M_\odot$, *Astron. J.*, 106, 773-789, 1993.

Hobbs, L.M., I. Iben, and C. Pilachowski, On lithium removal from G dwarfs, *Astrophys. J.*, 347, 817-820, 1989.

Iben, I., Stellar evolution within and off the main sequence, *Annu. Rev. Astron. Astrophys.*, 5, 571-626, 1967.

Kasting, J.F. and T.P. Ackerman, Climatic consequences of very high CO_2 levels in the Earth's early atmosphere *Science*, 234, 1383-1385, 1986.

Kasting, J.F., Runaway and moist greenhouse atmospheres and the evolution of Earth and Venus, *Icarus*, 74, 472-494, 1988.

Kasting, J.F., CO_2 condensation and the climate of early Mars, *Icarus*, 94, 1-13, 1991.

Kasting, J.F. and D.H. Grinspoon, The faint young Sun problem, in *The Sun in Time*, edited by C.P. Sonett, M.S. Giampapa and M.S. Matthews, pp. 447-462, University Arizona Press, Tucson, 1991.

Kasting, J.F., D.P. Whitmire, and R.T. Reynolds, *Icarus*, 101, 108-128, 1993.

Laskar, J., Large-scale chaos in the solar system, *Astron. Astrophys.*, 287, L9-L12, 1994.

Leitherer, C. and C. Robert, Observations of stellar winds from hot stars at 1.3 millimeters, *Astrophys. J.*, 377, 629-638, 1991.

McCrea, W.H., Ice ages and the Galaxy, *Nature*, 255, 607-609, 1975.

Mullan, D., E. Sion, F. Bruhweiler, and K. Carpenter, Evidence for a cool wind from the K2 dwarf in the detached binary V471 Tauri, *Astrophys. J.*, 339, L33-L36, 1989.

Mullan, D.J., J.G. Doyle, R.O. Redman, and M. Mathioudakis, Limits on detectability of mass loss from cool dwarfs, *Astrophys. J.*, 397, 225-231, 1992.

Panagia, N. and M. Felli, The spectrum of the free-free radiation from extended envelopes, *Astron. and Astrophys.*, 39, 1-5, 1975.

Papanastassiou, D.A., R.S. Rajan, J.C. Hunecke, and G.L. Wasserburg, *Lunar Sci.*, 583, 585, 1974.

Pinsonneault, M.H., S.D. Kawaler, S. Sofia, and P. Demarque, Evolutionary models of the rotating Sun, *Astrophys. J.*, 338, 424-452, 1989.

Pollack, J.B., J.F. Kasting, S.M. Richardson, and K. Poliakoff, The case for a wet, warm climate on early Mars, *Icarus*, 71, 203-224, 1987.

Popper, D.M., C.H. Lacy, M.L. Frueh, and A.E. Turner, Properties of main-sequence eclipsing binaries: Into the G stars with HS Aurigae, FL Lyrae, and EW Orionis, *Astron. J.*, 91, 383-404, 1986.

Radick, R.R., G.W. Lockwood, and S.L. Baliunas, Stellar activity and brightness variations: A glimpse at the Sun's history, *Science*, 247, 39-44, 1990.

Radick, R.R., The luminosity variability of solar-type stars, in *The Sun in Time*, edited by C.P. Sonett, M.S. Giampapa and M.S. Matthews, pp. 787-808, University Arizona Press, Tucson, 1991.

Rodonò, M. *et al.* Rotational modulation and flares on RS CVn and BY Dra-type stars, *Astron. and Astrophys.*, 165, 135-156, 1986.

Sagan, C. and G. Mullen, Earth and Mars: Evolution of atmospheres and surface temperatures, *Science*, 177, 52-56, 1972.

Soderblom, D.R., M.S. Oey, D.R.H. Johnson, and R.P.S. Stone, The evolution of the lithium abundances of solar-type stars. I. The Hyades and Coma Berenices clusters, *Astron. J.*, 99, 595-607, 1990.

Spruit, H.C., Effects of spots on a star's radius and luminosity, *Astron. Astrophys.*, 108, 348-355, 1982.

Squyres, S. (Ed.), MSATT LPI Workshop on Early Mars: How Warm and How Wet?, *LPI Tech. Rep. 93-03*, Part 1, 1993.

Stauffer, J. R., Angular momentum evolution of young main sequence stars, in *8th Cambridge Workshop on Cool Stars*, edited by J.-P. Caillault, Pub. PASP, San Francisco, in press, 1995.

Stauffer, J.R. and D.R. Soderblom, The evolution of angular momentum in solar mass stars, in *The Sun in Time*, edited by C.P. Sonett, M.S. Giampapa and M.S. Matthews, pp. 832-847, University Arizona Press, Tucson, 1991.

Swenson, F.J. and J. Faulkner, F.J. Rogers and C.A. Iglesias, The Hyades lithium problem revisited, *Astrophys. J.*, 425, 268-302, 1994.

Walter, F.M. and D.C. Barry, Pre- and main-sequence evolution of solar activity, in *The Sun in Time*, edited by C.P. Sonett, M.S. Giampapa and M.S. Matthews, pp. 633-657, University Arizona Press, Tucson, 1991.

Weidenschilling, S.J. and D.R. Davis, Orbital resonances in the solar nebula: Implications for planetary accretion, *Icarus*, 62, 16-29, 1985.

Wetherill, G.W., *Proc. Lunar Sci. Conf. 6*, 1539, 1975.

Weymann, R. and R.L.Sears, The depth of the convection envelope on the lower main sequence and the depletion of lithium, *Ap.J.*, 142, 174-181, 1965.

Whitmire, D.P., J.J. Matese, L.R. Doyle, and R.T. Reynolds, Planetological implications of mass loss from the early Sun, in *International Conference on Asteroids, Comets, Meteors*, p.238, (abstract), LPI Contribution 765, Houston, 1991.

Whitmire, D.P., L.R. Doyle, and R.T. Reynolds, and P.G. Whitman, Mars and the early Sun, in MSATT LPI Workshop on Early Mars: How Warm and How Wet? *LPI Tech. Rep. 93-03*, Part 1, 23, 1993.

Whitmire, D.P., L.R. Doyle, R.T. Reynolds, and J.J. Matese, Circumstellar habitable zones and mass loss from young solar-type stars. I. Theory, in *Fourth international bioastronomy symposium: Progress in the search for extraterrestrial life*, edited by S. Shostak, PASP Conf. Proc, San Francisco, in press, 1995.

Whitmire, D.P. and R.T. Reynolds, Circumstellar habitable zones: Astronomical considerations, in *First International Conference on Circumstellar Habitable Zones*, edited by L. Doyle, Travis House Pub., Menlo Park, CA, in press, 1995.

Wood, J.A. and P. Pellas, What heated the parent meteorite planets?, in *The Sun in Time*, edited by C.P. Sonett, M.S. Giampapa and M.S. Matthews, pp. 740-760, University Arizona Press, Tucson, 1991.

Wright, A.E. and M.F. Barlow, The radio and infrared spectrum of early-type stars undergoing mass loss, *Mon. Not. R. Astron. Soc.*, 170, 41-51, 1975.

L.R. Doyle, SETI Institute, 2035 Landing Drive, Mountain View, CA 94043. (e-mail: doyle@galileo.arc.nasa.gov)

J.J. Matese and D.P. Whitmire, Department of Physics, University of Southwestern Louisiana, Lafayette, LA 70504-4210. (e-mail: matese@usl.edu; whitmire@usl.edu)

R. T. Reynolds, MS-245, Theoretical Studies Branch, NASA-Ames Research Center, Moffett Field, CA 94035. (e-mail: reynolds@cosmic.arc.nasa.gov)

(Received March 9, 1994; revised October 24, 1994; accepted November 28, 1994.)

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5465–5484, MARCH 25, 1995

Diagnostic calculations of the circulation in the Martian atmosphere

Michelle L. Santee[1]
Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena

David Crisp
Earth and Space Sciences Division, Jet Propulsion Laboratory, California Institute of Technology Pasadena

Abstract. The circulation of the Martian atmosphere during late southern summer is derived from atmospheric temperature and dust distributions retrieved from a subset of the Mariner 9 infrared interferometer spectrometer (IRIS) thermal emission spectra (L_S = 343°–348°) (Santee and Crisp, 1993). Zonal-mean zonal winds are calculated by assuming gradient wind balance and zero surface zonal wind. Both hemispheres have intense midlatitude westerly jets with velocities of 80–90 m/s near 50 km; in the southern tropics the winds are easterly with velocities of 40 m/s near 50 km. The net effect of the zonal-mean meridional circulation and large-scale waves can be approximated by the diabatic circulation, which is defined from the atmospheric thermal structure and net radiative heating rates. The radiative transfer model described by Crisp (1990) and Santee (1993) is used to compute solar heating and thermal cooling rates from diurnal averages of the retrieved IRIS temperature and dust distributions. At pressures below 4 mbar, there are large net radiative heating rates (up to 5 K/d) in the equatorial region and large net radiative cooling rates (up to 12 K/d) in the polar regions. These net radiative heating rates are used in a diagnostic stream function model which solves for the meridional and vertical components of the diabatic circulation simultaneously. We find a two-cell circulation, with rising motion over the equator, poleward flow in both hemispheres, sinking motion over both polar regions, and return flow in the lowest atmospheric levels. The maximum poleward velocity is 3 m/s in the tropics at ~55 km altitude, and the maximum vertical velocity is 2.5 cm/s downward over the north pole at ~60 km altitude. If these large transport rates are sustained for an entire season, the Martian atmosphere above the 1-mbar level is overturned in about 38 days. This diabatic circulation is qualitatively similar to the terrestrial diabatic circulation at the comparable season, but is more vigorous.

1. Introduction

Atmospheric transport may be a crucial element in the seasonal cycles of carbon dioxide, dust, and water on Mars, but the extent of its contribution remains uncertain. The only in situ wind measurements were made by the meteorology experiments onboard the Viking landers [*Hess et al.*, 1977]. They acquired detailed temporal records of wind speed and direction, but only for two locations on the planet and only at a fixed elevation (~1.6 m) above the surface. Several authors [e.g., *Ryan et al.*, 1978; *Ryan and Henry*, 1979; *Tillman et al.*, 1979; *Barnes* 1980, 1981] have analyzed these data to provide constraints on larger-scale processes and the general circulation.

Recently, *Lellouch et al.* [1991] have reported the first measurements of winds in the middle atmosphere of Mars (40–70 km), obtained from Doppler shifts of the $J = 2 - 1$ ^{12}CO rotational line. Taken during the 1988 opposition, these observations provide extensive coverage of the southern hemisphere near southern summer solstice (L_S = 279°). Their results are consistent with the thermal structure of the Martian middle atmosphere derived from infrared heterodyne spectroscopy [*Deming et al.*, 1986; *Rothermel et al.*, 1988]. They are also in general agreement with the modeling results of *Barnes* [1990, 1991], *Pollack et al.* [1990], and *Haberle et al.*

[1]Now at Earth and Space Sciences Division, Jet Propulsion Laboratory, California Institute of Technology, Pasadena.

Paper number 94JE03207.
0148-0227/95/94JE-03207$05.00

[1993]. However, the Lellouch et al. measurements cannot be used to infer wind fields below 40 km.

There have been no direct systematic measurements of wind speed or direction at altitudes between the surface and approximately 40 km. In the absence of direct wind measurements, exhaustive compilations of cloud structures, including their spatial and temporal variations, have been undertaken for both Mariner 9 and Viking Orbiter images [*Leovy et al.*, 1972; *Briggs and Leovy*, 1974; *Briggs et al.*, 1977; *French et al.*, 1981; *Kahn*, 1983]. Although studies of cloud motions and morphology have yielded wind directions, and in some cases, estimates of wind speeds, they suffer from a lack of systematic global coverage. Because regional winds arise from a variety of factors, including the presence of slopes or strong horizontal surface temperature variations (e.g., at the edges of the polar caps or in regions of large albedo or thermal inertia contrasts), these observed winds may not be representative of the mean flow.

Winds in the intermediate region between the surface and the upper portion of the middle atmosphere can be determined indirectly from the observed atmospheric temperature distribution. The thermal structure of the Martian atmosphere has been investigated by remote-sensing instruments on several different spacecraft. The best existing data for studying the vertical temperature distribution on a global scale are thermal emission spectra recorded by the Mariner 9 infrared interferometer spectrometer (IRIS). Using a new technique for the simultaneous retrieval of atmospheric temperatures and airborne dust abundances from IRIS spectra, we derived a global description of the thermal structure and dust loading of the Martian atmosphere for a relatively clear period after a severe dust storm in late southern summer [*Santee and Crisp*, 1993] (hereinafter Paper I). In Paper I we examined a subset of the IRIS data consisting of approximately 2400 spectra in a 12-day interval extending between $L_S = 343°$ and 348°. Temperature and dust optical depth profiles were retrieved from each spectrum, and global maps were constructed as functions of latitude (±90°), altitude (~0–60 km), and Mars local time.

One of the principal conclusions from Paper I is that both dayside and nightside atmospheric temperatures at altitudes above about 40 km are warmer over the winter (north) polar regions than over the equator or the summer (south) polar regions. These anomalous temperatures are consistent with the thermal structure in the middle atmosphere observed by *Deming et al.* [1986] and *Rothermel et al.* [1988], and they indicate that the atmosphere is not in radiative equilibrium. A similar reversed temperature gradient is present in the terrestrial mesosphere and is known to be associated with the meridional circulation, which produces enough adiabatic heating (or cooling) to overcome the radiative forcing [*Leovy*, 1964].

In this paper we describe methods for deriving the atmospheric circulation from observed temperatures. Zonal-mean zonal winds are calculated from the observed meridional gradients of the zonally averaged temperatures using the gradient thermal wind equation (section 2.1). Both the zonal-mean meridional circulation and large-scale waves contribute to the north–south atmospheric transport. *Dunkerton* [1978] shows that their net effect can be approximated by the diabatic circulation, i.e., that circulation needed to maintain the observed temperature distribution (warm winter pole, cool tropics) in the presence of the radiative drive. The Transformed Eulerian-Mean (TEM) formulation for the diabatic circulation is described in section 2.2. In section 2.3, we present a stream function model which solves for the meridional and vertical components of the diabatic circulation simultaneously. This model requires values for the net radiative heating rates, which we derive from a radiative transfer model [*Crisp*, 1990; *Santee*, 1993] that evaluates the diurnally averaged solar heating and thermal cooling rates for the retrieved atmospheric temperature and dust distributions.

The results of the zonal wind and diabatic circulation calculations are discussed in section 3. Although similar calculations of the zonal-mean zonal wind have been reported previously [*Conrath*, 1981; *Leovy*, 1982; *Michelangeli et al.*, 1987], they were based on Mariner 9 meridional cross sections of temperature which were not global in coverage. In addition, although numerical simulations of the Martian atmosphere have been conducted using a full three-dimensional general circulation model [*Pollack et al.*, 1981, 1990; *Haberle et al.*, 1993] and a two-dimensional coupled chemical, radiative, and dynamical model [*Moreau et al.*, 1991], our results constitute the first estimation of the diabatic circulation of the Martian atmosphere from actual temperature observations.

2. Diagnostic Calculations of the Atmospheric Circulation

2.1. The Gradient Thermal Winds

In describing the large-scale circulation of the atmosphere, it is common to mathematically distinguish between the zonally symmetric flow and the azonal disturbances (usually called "eddies" or "waves") superimposed on it. Any field variable can be written as the sum of the zonal mean part, denoted by an overbar, and the deviation from the mean, denoted by a prime. Using the zonal wind u for illustration,

$$u(\lambda, \phi, z, t) = \bar{u}(\phi, z, t) + u'(\lambda, \phi, z, t) , \qquad (1)$$

where t is time, λ is longitude, and ϕ is latitude. The vertical coordinate z used throughout this paper is not the geometric height but rather the log-pressure altitude: $z = -H \ln(p/p_s)$, where p is the atmospheric pressure and the subscript s denotes a value at the surface. H is the mean atmospheric scale height ($H = RT_0/g$) calculated by assuming a constant reference temperature T_0 and a mean gas constant R; g is the gravitational acceleration (assumed constant).

One definition of the zonal averaging operator is

$$\overline{u}(\phi, z, t) = \frac{1}{2\pi} \int_0^{2\pi} u(\lambda, \phi, z, t)\, d\lambda . \qquad (2)$$

This average is an Eulerian mean, taken at fixed values of the coordinates ϕ, z, and t. Applying the expansion of equation (1) and the zonal averaging operator of equation (2) to the primitive equations of atmospheric motion leads to a reformulated set of equations in which the zonal circulation and the eddy terms have been separated [e.g., *Andrews et al.*, 1987].

The zonal–mean zonal winds can be estimated directly from the observed thermal structure of the atmosphere by combining the meridional momentum equation with the hydrostatic equation to yield

$$\left[f + \frac{2\overline{u} \tan\phi}{a}\right] \frac{\partial \overline{u}}{\partial z} = -\frac{R}{aH} \frac{\partial \overline{T}}{\partial \phi} . \qquad (3)$$

Equation (3), the gradient thermal wind equation, specifies the vertical wind shear due to meridional temperature gradients; a is the mean radius of the planet, $f = 2\Omega \sin\phi$ is the Coriolis parameter, and Ω is the rotation rate. The length of the Martian day is 24.6 hours, the rotation rate $\Omega = 7.1 \times 10^{-5}\,\mathrm{s}^{-1}$, and the Coriolis forces are as important on Mars as they are on Earth. This equation describes a three–way balance between the Coriolis force, the centrifugal force, and the horizontal pressure gradient force. If the centrifugal acceleration term is small compared to the Coriolis term, it can be omitted to yield the geostrophic thermal wind equation:

$$f \frac{\partial \overline{u}}{\partial z} = -\frac{R}{aH} \frac{\partial \overline{T}}{\partial \phi} . \qquad (4)$$

Equation (4) may not always be valid for Mars. Both equations (3) and (4) highlight the fact that the wind field is strongly coupled to the atmospheric temperature distribution.

To solve equation (3) for the zonal–mean zonal wind $\overline{u}$, the derivatives with respect to ϕ and z are approximated by finite differences [e.g., *Haltiner and Williams*, 1980]. We divide the ϕ, z space into a grid of $M \times N$ points, with 10° latitude bins and 30 vertical levels between 6.0 mbar (taken to be the "surface") and 10^{-3} mbar. The grid points are evenly spaced in both the horizontal and vertical directions, and any point on the grid is uniquely identified by the indices (m, n). We use centered finite difference formulas for the derivatives at the point (m, n) with respect to ϕ. However, because the quantity we are solving for appears in both the $2\overline{u}\tan\phi/a$ and $\partial\overline{u}/\partial z$ terms, a forward difference formula is more appropriate for the derivatives with respect to z. Applying the finite difference formulas to equation (3) and grouping like terms together yields a quadratic equation for $\overline{u}_{m,n}$:

$$\left[\frac{2/a \tan\phi_m}{z_{n+1} - z_n}\right] \overline{u}_{m,n}^2 - \left[\frac{2/a \tan\phi_m\, \overline{u}_{m,n+1} - f_m}{z_{n+1} - z_n}\right] \overline{u}_{m,n}$$

$$- \left[\frac{R}{aH} \frac{\overline{T}_{m+1,n} - \overline{T}_{m-1,n}}{\phi_{m+1} - \phi_{m-1}} + \frac{f_m\, \overline{u}_{m,n+1}}{z_{n+1} - z_n}\right] = 0 . \qquad (5)$$

Because the solution for $\overline{u}_{m,n}$ involves the value of $\overline{u}_{m,n+1}$, the winds at some reference level must be known. Typically, $\overline{u}$ is assumed to be zero at the surface. Evidence that surface winds on Mars are generally light (less than 10 m/s) comes from the meteorology experiments onboard the Viking landers [*Hess et al.*, 1977; *Ryan et al.*, 1978; *Ryan and Henry*, 1979; *Barnes*, 1980, 1981] and from constraints imposed by observations of resonant gravity waves in the winter polar region [*Briggs and Leovy*, 1974].

There are circumstances under which surface winds are not negligible. Viking lander observations revealed that surface winds rose steadily with the onset of global dust storms, with gusts reaching 30 m/s [*Ryan and Henry*, 1979]. Numerical simulations of global dust storms predict that (up to a point) higher dust opacities are accompanied by stronger surface winds [*Haberle et al.*, 1982]. Other modeling studies [*Haberle et al.*, 1979] indicate that the magnitude of the surface wind is substantially enhanced (up to 30 m/s) when the condensation rate onto, or the sublimation rate from, the seasonal polar caps is at a maximum. Because this data subset represents a relatively clear period after a global dust storm has subsided (Paper I), and because the polar cap at this season is relatively stable (section 2.2), we assume negligible surface winds for the purposes of calculating the vertical wind shear.

2.2. The Diabatic Circulation

The evolution in the understanding of wave–mean flow interactions in the terrestrial atmosphere has been discussed in depth by *Dunkerton* [1978], *Dunkerton* [1980], *Hsu* [1980], *Matsuno* [1980], *Mahlman et al.* [1984], and others. These concepts were reviewed and applied to Mars by *Santee* [1993], and will only be briefly summarized here. The splitting of the flow into eddy fluxes and a mean meridional circulation is an artificial distinction imposed by the Eulerian–mean approach. In actuality, these two processes are highly interrelated. One of the earliest investigations to use a numerical general circulation model to study quantitatively the relative importance of the mean meridional circulation and the eddies in transporting trace substances in the terrestrial stratosphere was that of *Hunt and Manabe* [1968]. Their work demonstrated that eddy fluxes due to planetary waves in the stratosphere are almost counteracted by advection due to the mean meridional circulation. This near cancellation of the two transport processes is a consequence of the fact that stratospheric planetary waves approximately satisfy the "noninteraction" theorem. This theorem, which was established by several authors and generalized by *Andrews and McIntyre* [1976, 1978], states that waves which are linear, steady (constant amplitude), and conservative (not being internally forced or dissipated) induce a mean meridional cell which acts to identically

cancel the eddy fluxes of the waves themselves. Under these conditions, the waves cause no changes to the basic flow. Because the eddy heat and momentum fluxes tend to force a compensating mean meridional circulation, in the Eulerian–mean framework the net transport is generally the small residual of two large terms of opposite sign.

The near cancellation of the eddy and zonal–mean components suggests that the conventional Eulerian mean is not well suited for defining the transport circulation. An alternative is the Lagrangian mean. Instead of taking a simple zonal average along a latitude circle, the Lagrangian mean involves averaging with respect to a curved material tube consisting of a set of fluid parcels [e.g., *Hsu*, 1980; *Matsuno*, 1980]. In the absence of waves, the material tube coincides with a latitude circle; under the influence of waves, it is distorted into a wavy shape. The motion of the center of mass of the tube defines the Lagrangian–mean motion. The Lagrangian–mean and Eulerian–mean circulations are quite distinct in general and may, in fact, be of opposite sign. However, net mass transport is achieved entirely by the Lagrangian–mean flow [*Dunkerton*, 1978].

Although conceptually Lagrangian means are more appropriate for transport calculations, in practice the direct application of the Lagrangian–mean formalism to atmospheric data encounters serious technical difficulties. Eulerian quantities are much easier to calculate. An improvement over the traditional Eulerian formulation was suggested by *Andrews and McIntyre* [1976]. They introduced a "residual" mean meridional circulation $(\overline{v}^*, \overline{w}^*)$ defined through the transformation

$$\overline{v}^* \equiv \overline{v} - \frac{1}{\rho}\frac{\partial}{\partial z}\left(\frac{\rho\,\overline{v'T'}}{\frac{\partial\overline{T}}{\partial z}+\frac{g}{c_p}}\right) \tag{6a}$$

$$\overline{w}^* \equiv \overline{w} + \frac{1}{a\cos\phi}\frac{\partial}{\partial\phi}\left(\frac{\overline{v'T'}\cos\phi}{\frac{\partial\overline{T}}{\partial z}+\frac{g}{c_p}}\right), \tag{6b}$$

where $\overline{v}$ and $\overline{w}$ are the zonal–mean velocity components in the meridional and vertical directions, respectively, ρ is the atmospheric density, the $\overline{v'T'}$ term represents an eddy heat flux, and c_p is the specific heat at constant pressure, which for a pure CO_2 atmosphere can be obtained from an expression of the form [*Touloukian and Makita*, 1970]

$$c_p(T) = 443.15 + 1.688\,T - 1.269\times10^{-3}T^2 + 3.470\times10^{-7}T^3 . \tag{7}$$

The term g/c_p is the adiabatic lapse rate.

After using the definitions in equation (6) to substitute in for $\overline{v}$ and $\overline{w}$ in the Eulerian–mean version of the primitive equations and performing many algebraic manipulations, a set of TEM equations is obtained. For completeness, the full set of TEM equations is given here:

$$\frac{\partial\overline{u}}{\partial t} + \frac{\overline{v}^*}{a\cos\phi}\frac{\partial}{\partial\phi}\left(\overline{u}\cos\phi\right) - f\overline{v}^* + \overline{w}^*\frac{\partial\overline{u}}{\partial z} = \overline{X} + \frac{1}{\rho a\cos\phi}\nabla\cdot\mathcal{F} \tag{8a}$$

$$\frac{\overline{u}^2\tan\phi}{a} + f\overline{u} + \frac{1}{a}\frac{\partial\overline{\Phi}}{\partial\phi} = \mathcal{G} \tag{8b}$$

$$\frac{\partial\overline{\Phi}}{\partial z} - \frac{R\overline{T}}{H} = 0 \tag{8c}$$

$$\frac{1}{a\cos\phi}\frac{\partial}{\partial\phi}\left(\overline{v}^*\cos\phi\right) + \frac{1}{\rho}\frac{\partial}{\partial z}\left(\rho\,\overline{w}^*\right) = \frac{\partial\rho}{\partial t} \approx 0 \tag{8d}$$

$$\frac{\partial\overline{T}}{\partial t} + \frac{\overline{v}^*}{a}\frac{\partial\overline{T}}{\partial\phi} + \overline{w}^*\left[\frac{\partial\overline{T}}{\partial z}+\frac{g}{c_p}\right] - \overline{Q} = -\frac{1}{\rho}\frac{\partial}{\partial z}\left[\frac{\rho\,\overline{v'T'}\,\frac{\partial\overline{T}}{\partial\phi}}{a\left(\frac{\partial\overline{T}}{\partial z}+\frac{g}{c_p}\right)} + \rho\,\overline{w'T'}\right]. \tag{8e}$$

The first two equations express momentum balance in the zonal and meridional directions, respectively. In equation (8a), X represents the zonal component of friction or other unspecified nonconservative forces. The vector $\mathcal{F}$ is called the Eliassen–Palm (EP) flux. The divergence of the EP flux is a direct measure of the degree of interaction between eddies and the mean flow at each latitude and height, and is therefore of prime importance in determining the total wave forcing of the zonal–mean state [*Edmon et al.*, 1980]. Assuming steady state conditions and neglecting the drag term, equation (8a) reduces to a balance between the force per unit mass exerted on $\overline{u}$ by the mean meridional circulation and the force per unit mass produced by the waves. Thus a nonzero EP flux divergence implies the existence of a nonzero residual mean circulation. In equation (8b), $\mathcal{G}$ represents all the terms that disrupt the gradient wind balance between $\overline{u}$ and the geopotential $\overline{\Phi}$. In most meteorological applications, $\mathcal{G}$ is small [*Andrews et al.*, 1987]. The third equation, which is written in terms of the geopotential, expresses hydrostatic balance in the vertical.

Equation (8d), the continuity equation, describes the conservation of mass. The source and sink terms, represented by $\partial\rho/\partial t$, are formally included in this equation to account for the condensation and sublimation of a significant fraction of the CO_2 atmosphere at high latitudes over the course of a year. Approximately 25% of the atmospheric mass is cycled into and out of the seasonal polar caps [*Hess et al.*, 1979], but these terms are most important during the fall, when the cap is growing, and during the spring, when it is receding. A corresponding large–amplitude seasonal oscillation in the daily–average atmospheric pressure was observed at both Viking lander sites [*Hess et al.*, 1979]. There was a deep pressure minimum near $L_S = 150°$, signaling maximum CO_2 accumulation onto the south polar cap at the end of southern winter. A secondary, much shallower minimum occurred near $L_S = 350°$, signaling maximum CO_2 accumulation onto the north polar cap at the end of northern winter. The daily–average pressure at both sites remained relatively constant from $L_S = 340°$ to 360°, suggesting that the north polar cap was stable during this period (at least during the Viking year). It is therefore probable that the north polar cap was neither

growing nor retreating in any significant way during the time this data subset was recorded ($L_S \simeq 343°$–$348°$). Furthermore, based on a systematic review of wind direction information obtained from a variety of indicators, *Kahn* [1983] found that recession of the polar cap in the north produces much smaller mass fluxes than in the south, and hence has very little influence on wind directions during early northern spring. He concluded that during the late winter and early spring seasons in the north, wind directions are determined by the temperature field, not cap flows. In addition, *Pollack et al.* [1990] used a three–dimensional general circulation model of the Martian atmosphere to conduct numerical simulations of polar processes for different seasons and dust loadings. Simulations at $L_S = 342°$ (with a visible dust optical depth $\tau = 1$) indicated that the importance of heat transport by the condensation flow (mass flow toward or away from polar regions where CO_2 is condensing or subliming) is relatively small at this season, except at the very center of the cap. Consequently, we neglect the $\partial p/\partial t$ term for these calculations.

Finally, equation (8e) is the thermodynamic energy equation. This equation relates the time rate of change of the temperature following the fluid motion to the diabatic heating Q. The term "diabatic" refers to a change in temperature involving the actual addition or removal of heat (as opposed to adiabatic heating). Diabatic processes include absorption of solar radiation, absorption and emission of infrared radiation, and latent heat release. Also, mixing of different air masses within a convective layer leads to diabatic heating through conduction, friction, and diffusion [*Hantel and Baader*, 1978; *Wallace and Hobbs*, 1977]; however, for most meteorological applications these contributions are neglected. We also neglect the latent heat component because temperatures in our data subset rarely approach the CO_2 condensation temperature. Therefore, the diabatic heating reduces to the net radiative heating.

As discussed above, in the Eulerian–mean framework the net transport is the small residual between nearly equal, but opposite, eddy and mean circulation terms. In the transformed mean formalism, the problem has been formulated directly in terms of the remainder. One of the principal advantages to the TEM equations is that $\overline{v}^*$ and $\overline{w}^*$ represent the velocity components whose adiabatic heating and/or cooling is not balanced by eddy heat fluxes, and they are therefore the part of the mean circulation which is relevant for mass transport. A second advantage is that the residual mean circulation ($\overline{v}^*$, $\overline{w}^*$) provides a description of atmospheric motions that is close to the actual time–averaged behavior of fluid parcels.

Dunkerton [1978] argued that under certain approximations the residual (hence mass transport) circulation can be inferred solely from the Eulerian–mean net diabatic heating. The resulting $\overline{v}^*$ and $\overline{w}^*$ are called the "diabatic circulation." The main approximation involves the omission of the "wave heating" term on the right–hand side of equation (8e). The generalized noninteraction theorem can be used to prove that for steady, linear, conservative disturbances, the EP flux divergence and the wave heating both vanish [*Andrews and McIntyre*, 1976]. *Andrews et al.* [1987] state that the wave–heating term is negligible for quasi–geostrophic motions (in fact, it drops out of the quasi–geostrophic form of the TEM equations) and for steady, conservative gravity waves. *Hitchman and Leovy* [1986] estimated this term for the terrestrial middle atmosphere from satellite data and found it to be much smaller than the other terms in the thermodynamic energy equation. Similar assumptions have routinely been made by other researchers [e.g., *Holton and Wehrbein*, 1980; *Crisp*, 1983; *Solomon et al.*, 1986; *Rosenfield et al.*, 1987; *Gille et al.*, 1987; *Shine*, 1989]. With these precedents, and in the absence of information about the atmosphere of Mars, we will also neglect the wave heating.

In actuality, because the observed climatological temperature field is far from the radiative equilibrium structure, the EP flux divergence must be nonzero and the waves must be transient, nonlinear, or nonconservative. Thus, the diabatic circulation approximation introduces an essential inconsistency: the wave heating is negligible, but the EP flux divergence is not. That is, eddy transience and dissipation are neglected in the thermodynamic energy equation, but not in the momentum equation. This inconsistency can be justified because the wave heating is most likely much smaller than the radiative heating, whereas the EP flux divergence is most likely much larger than the drag. It is believed that the "breaking" of vertically propagating gravity waves provides the majority of the forcing of a nonzero EP flux divergence in the terrestrial mesosphere [*Holton*, 1983; *Andrews et al.*, 1987]. *Barnes* [1990, 1991] has suggested that breaking topographically forced gravity waves could explain the large deviations from radiative equilibrium that have been observed [*Deming et al.*, 1986; *Rothermel et al.*, 1988] in the middle atmosphere of Mars.

The terms in equation (8e) associated with the meridional velocity ($\overline{v}^*$) and seasonal change ($\partial/\partial t$) are also customarily ignored [e.g., *Dunkerton*, 1978; *Rosenfield et al.*, 1987; *Gille et al.*, 1987]. This leads to a very simplified balance equation:

$$\overline{w}^* \left[\frac{\partial \overline{T}}{\partial z} + \frac{g}{c_p} \right] \approx \overline{Q} \,. \tag{9}$$

Equation (9) states that the net vertical transport can be obtained from the diabatic heating and the static stability alone. Equation (9) is often combined with the continuity equation (equation (8d)) and used in an iterative procedure to estimate the residual meridional circulation ($\overline{v}^*$, $\overline{w}^*$). Although this strategy provides an adequate first approximation to the diabatic circulation, it is predicated on an inherent discrepancy, and it depends strongly on a proper normalization of the net radiative heating rate to ensure that global mass balance constraints are satisfied (i.e., nonzero net heating rates indicate net mass transport into or out of the system; see, e.g., the review by *Shine* [1989]). Therefore, we chose not to use a procedure based on equation (9).

2.3. The Stream Function Model

In this paper we adopt an approach that solves for $\overline{v}^*$ and $\overline{w}^*$ simultaneously. The meridional advection term is retained in the thermodynamic energy equation, but the $\partial\overline{T}/\partial t$ term is of little consequence due to the limited seasonal extent (only 12 days) of the data subset analyzed in Paper I. Neglecting the $\partial\overline{T}/\partial t$ and wave–heating terms, equation (8e) reduces to

$$\frac{\overline{v}^*}{a}\frac{\partial\overline{T}}{\partial\phi} + \overline{w}^*\left[\frac{\partial\overline{T}}{\partial z} + \frac{g}{c_p}\right] = \overline{Q}\,. \tag{10}$$

This equation is coupled to the continuity equation through the definition of a zonal–mean stream function, ψ. The form of ψ was adopted from *Shia et al.* [1989]:

$$\overline{v}^* = -\frac{1}{\cos\phi}e^{\frac{z}{H}}\frac{\partial}{\partial z}\left(e^{-\frac{z}{H}}\psi\right) = \frac{1}{\cos\phi}\left[\frac{\psi}{H} - \frac{\partial\psi}{\partial z}\right] \tag{11a}$$

$$\overline{w}^* = \frac{1}{a\cos\phi}\frac{\partial\psi}{\partial\phi}\,. \tag{11b}$$

Substituting these expressions into equation (10) yields

$$\frac{1}{a\cos\phi}\left[\frac{\psi}{H} - \frac{\partial\psi}{\partial z}\right]\frac{\partial\overline{T}}{\partial\phi} + \frac{1}{a\cos\phi}\left[\frac{\partial\overline{T}}{\partial z} + \frac{g}{c_p}\right]\frac{\partial\psi}{\partial\phi} = \overline{Q}\,. \tag{12}$$

Straightforward integration of this equation is precluded by numerical instabilities. R.-L. Shia (personal communication, 1990) suggested taking the derivative of the entire equation with respect to latitude (ϕ). Applying $\partial/\partial\phi$ to equation (12) and grouping ψ terms together leads to

$$\begin{aligned}
&\frac{1}{aH\cos\phi}\left[\frac{\partial^2\overline{T}}{\partial\phi^2} + \tan\phi\,\frac{\partial\overline{T}}{\partial\phi}\right]\psi \\
&- \frac{1}{a\cos\phi}\left[\frac{\partial^2\overline{T}}{\partial\phi^2} + \tan\phi\,\frac{\partial\overline{T}}{\partial\phi}\right]\frac{\partial\psi}{\partial z} \\
&+ \frac{1}{a\cos\phi}\left[\frac{1}{H}\frac{\partial\overline{T}}{\partial\phi} + \left(\frac{\partial\overline{T}}{\partial z} + \frac{g}{c_p}\right)\tan\phi + \frac{\partial}{\partial\phi}\left(\frac{\partial\overline{T}}{\partial z} + \frac{g}{c_p}\right)\right]\frac{\partial\psi}{\partial\phi} \\
&+ \frac{1}{a\cos\phi}\left(\frac{\partial\overline{T}}{\partial z} + \frac{g}{c_p}\right)\frac{\partial^2\psi}{\partial\phi^2} \\
&- \frac{1}{a\cos\phi}\frac{\partial\overline{T}}{\partial\phi}\frac{\partial^2\psi}{\partial z\partial\phi} = \frac{\partial\overline{Q}}{\partial\phi}\,.
\end{aligned} \tag{13}$$

This mathematical "trick" changes the character of the equation and allows it to be integrated numerically. An additional benefit is that in this form of the equation, the diabatic heating appears only as a derivative with respect to latitude, eliminating the need for a heating rate normalization.

Our approach for solving equation (13) is described in Appendix A. The solution depends on the boundary conditions that are applied. We assume that $\psi \equiv 0$ on the boundaries of the domain, which are the south pole, the north pole, the surface, and the top of the atmosphere:

$$\psi_{1,n} = \psi_{M,n} = \psi_{m,1} = \psi_{m,N} = 0\,. \tag{14}$$

This requires that there is no mass flow through the boundaries, and assumes nonacceleration conditions. These assumptions may introduce some error near the surface, where eddy heat fluxes can be large. With these boundary conditions, we only need to solve for interior grid points, resulting in a reduction in the total dimensions of the grid such that $M \rightarrow M-2$, $N \rightarrow N-2$.

The stream function model was validated by computing the diabatic circulation of the terrestrial atmosphere for four representative months (January, April, July, October). The meridional temperature cross sections for these calculations were obtained from an 8–year average of National Meteorological Center (NMC) data (1979–1986, inclusive), and the ozone profiles were obtained from a 7–year average of stratospheric aerosol and gas experiment II (SAGE II) data (M. Gerstell, personal communication, 1992). The circulation above the 20–km level deduced from the NMC temperatures using the stream function model was compared to a number of other recent diagnostic calculations. These studies, which employ a variety of radiative transfer schemes and input temperature and ozone data, include *Hitchman and Leovy* [1986], *Solomon et al.* [1986], *Rosenfield et al.* [1987], *Gille et al.* [1987], *Shine* [1989], and *Yang et al.* [1990]. Despite the differences in both data and models, all these studies produce generally similar fields, and our results compare favorably to them in both overall structure and actual velocities.

3. Results and Discussion

We use the diagnostic methods described in section 2 to define two aspects of Martian atmospheric dynamics: the zonal winds, and the net meridional circulation. These calculations require knowledge of the atmospheric temperature structure. The retrieved 0200 LT and 1400 LT meridional cross sections of atmospheric temperature presented in Paper I were averaged to produce a diurnal–mean temperature distribution, which is shown in Figure 1a. In section 3.1 the zonal–mean zonal winds are discussed. Although the principal results are obtained from the gradient thermal wind equation, winds obtained from the geostrophic thermal wind equation are also included to facilitate comparisons with previous studies.

Solar heating rates and thermal cooling rates are calculated from the temperature (Figure 1a) and dust op-

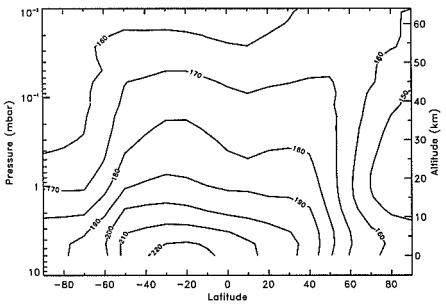


Figure 1a. Diurnal-mean temperatures obtained by averaging the 1400 LT and 0200 LT temperature maps retrieved from the Mariner 9 IRIS spectra in Paper I. For consistency with the results of Paper I, the vertical coordinate in this and all subsequent figures is the atmospheric pressure p. The approximate altitude z corresponding to a given pressure level is also included in these figures (see section 2.1). This correlation of altitude with pressure is based on the following values: $p_s = 6.0$ mbar, $R = 191.0$ J K^{-1} kg^{-1}, $g = 3.74$ m/s^2, and $T_0 = 196$ K (the globally and diurnally averaged value of the atmospheric temperature at the surface), resulting in a mean scale height $H = 10$ km.

tical depth (Figure 1b) distributions retrieved from the IRIS data. The radiative transfer model is described by *Crisp* [1990] and in the appendix of *Santee* [1993]. The radiative heating rate calculation is very sensitive to the assumed dust optical properties and particle size distribution. Our assumptions for these parameters are described in detail in Appendix B. The contributions from convective heating and cooling are ignored in the radiative transfer calculations, leading to some inaccuracies in the radiative heating rates in the lowest half scale height of the atmosphere. Net radiative heating rates, found by summing the solar heating and thermal cooling rates at each latitude and atmospheric level, are shown in section 3.2.

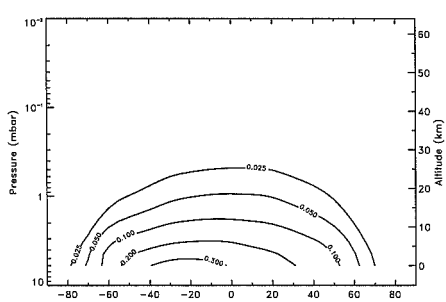


Figure 1b. Dust optical depth distribution obtained from the 1400 LT opacity map retrieved from Mariner 9 IRIS spectra in Paper I.

Finally, in section 3.3 we present the vertical and meridional components of the diabatic circulation diagnosed from the net heating rates using the stream function model, along with the EP flux divergence calculated from the zonal momentum equation. In addition, the diabatic circulation is compared to results from previous modeling simulations, and some implications of the strength of the circulation are discussed.

3.1. Zonal–Mean Zonal Winds

Zonal winds computed from the gradient thermal wind equation (equation (3)) assuming zero wind velocity at the surface are shown in Figure 2a. One of the characteristics of the thermal wind equation is that the zonal wind increases with altitude in the presence of equator-to-pole temperature gradients. Because the horizontal temperature gradients are steeper in the northern (winter) midlatitudes (see Figure 1a), the eastward (westerly) winds aloft are slightly more intense over the northern regions, although both hemispheres have strong midlatitude eastward jets (with velocities of 80–90 m/s near 50 km). Winds in the equatorial regions, where the Coriolis parameter approaches zero, cannot be determined from gradient wind balance and are omitted from this figure. In the southern tropics the winds are westward (easterly), but there is only a small corresponding region of easterlies at the top of the domain in the northern hemisphere.

A second calculation was performed with the same temperatures but using the simpler geostrophic thermal wind equation (equation (4)). The overall morphology of the geostrophic thermal wind field (shown in Figure 2b) is consistent with that of the gradient thermal wind field (Figure 2a), but the geostrophic upper level westerly jets are up to 30 m/s stronger. A similar overestimation of the zonal-mean westerlies has been found in calculations of the terrestrial zonal winds. In a comparison of gradient and geostrophic mean zonal winds, *Marks* [1989] notes that their spatial structures are very

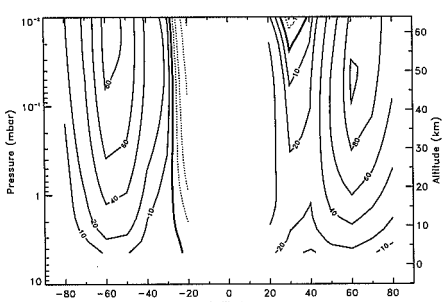


Figure 2a. Zonal-mean zonal winds, in m/s, calculated from the gradient thermal wind equation. Positive values represent eastward winds, negative contours are dashed, the zero contour is thicker, and the contour interval is nonuniform.

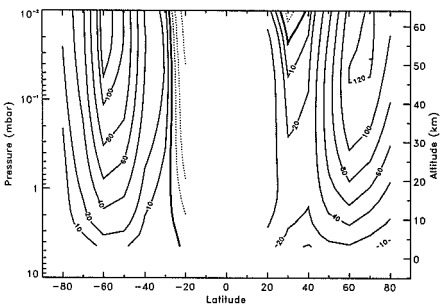


Figure 2b. Zonal–mean zonal winds, in m/s, calculated from the geostrophic thermal wind equation. Positive values represent eastward winds, negative contours are dashed, the zero contour is thicker, and the contour interval is nonuniform.

similar but that the maximum values of the positive gradient winds are up to 20% less than the corresponding geostrophic values, and that the differences between them are greatest in the regions of strong winds. This arises from the fact that the centrifugal acceleration term, absent in the geostrophic wind equation, is proportional to the zonal wind speed.

Several different calculations of the zonal winds based on Mariner 9 data have been reported previously, but none of these are global in coverage. *Briggs and Leovy* [1974] find evidence for resonant gravity waves in the middle latitudes in Mariner 9 television images taken during northern winter. Using gravity–wave theory, along with temperatures in the 45°–50°N latitude range obtained from inversion of IRIS spectra recorded on one orbit near $L_S = 320°$ and a wavelength for the lee waves of 30 km, Briggs and Leovy estimate the surface zonal wind to be 10 m/s. From the geostrophic approximation they calculate a profile of the zonal wind up to 30–km altitude. Their results show the wind speed increasing up the column to a maximum value at the top of the domain of nearly 120 m/s.

Pollack et al. [1981] present geostrophic zonal winds computed from Mariner 9 IRIS data as one means of validating the NASA/Ames Mars general circulation model (GCM). These cross sections encompass the area between 50°S and 70°N, from the surface to the 0.3–mbar level, and include spectra spanning $L_S = 43°$–54°. This season is roughly midway between the spring equinox and the summer solstice in the northern hemisphere. The zonal wind field is characterized by a strong eastward jet (with velocities exceeding 120 m/s) between 1.0 mbar and 0.5 mbar at 50°S latitude, much weaker eastward winds (with velocities of 20–30 m/s) at that altitude between 60°N and 80°N latitude, and westward winds (of unspecified magnitude) in the region between the equator and 40°N. The occurrence of westward winds in the northern tropics, rather than the southern tropics as in Figure 2, and the much weaker eastward winds in the northern high latitudes, are consequences of the fact that *Pollack et al.* [1981] analyzed IRIS spectra representing a different season from those of Paper I (late southern summer).

Geostrophic zonal winds calculated from Mariner 9 IRIS spectra at the same season as those of Paper I are reported by *Conrath* [1981]. Conrath constructs a northern hemisphere mean meridional temperature cross section by averaging temperature profiles retrieved from spectra recorded during $L_S = 330°$–350°. He finds an eastward jet of 160 m/s at about 0.2 mbar and 60°N. *Michelangeli et al.* [1987] use the *Conrath* [1981] temperature distribution, but assume gradient wind balance to compute the zonal winds. They present a cross section of the zonal–mean gradient wind between 15°N–85°N and 5.0–0.2 mbar which exhibits the same features as Figure 2a. The ageostrophic correction to the zonal–mean zonal wind reduces the maximum eastward velocity from 160 m/s to 100 m/s, which is in better agreement with the results displayed in Figure 2a.

Haberle et al. [1993] use the NASA/Ames Mars GCM to simulate the response of the Martian atmospheric circulation to variations in season and dust loading. Comparisons with these GCM results provide insight into the behavior depicted in Figure 2. Because the interval represented by our data subset ($L_S = 343°$–348°) falls between solstice and equinox, we will examine the GCM results at both of these seasons. GCM experiments for northern winter solstice ($L_S = 255°$–286°) under relatively clear conditions reveal zonal winds which are easterly in the southern hemisphere (with peak velocities of 60 m/s) and westerly in the northern hemisphere (with peak velocities of 120 m/s). The intense westerly jet is centered near 50°N–60°N and 35 to 40–km altitude, whereas the easterly jet in the summer hemisphere is much broader, covering the entire tropical region at 40–km altitude. The GCM zonal wind field for a period just past northern spring equinox ($L_S = 0°$–23°) is very symmetric with respect to the equator. Westerlies appear in both hemispheres, but these westerly jets (peak velocities of 80 m/s at 50°S and 40 m/s at 60°N) are weaker than those present during solstice conditions due to weaker latitudinal temperature gradients. Easterlies have formed in the tropical regions of both hemispheres above 30–km altitude and near the surface, although not in the intermediate region of the atmosphere. Thus, Figure 2 represents a transitional period from a season in which the winds are easterly throughout the southern hemisphere and westerly throughout the northern hemisphere, to a season in which the winds are westerly at the higher latitudes of both hemispheres and easterly in the tropics.

Zonally averaged zonal wind and temperature fields from the NASA/Ames Mars GCM are presented by *Barnes et al.* [1993] for relatively low–dust conditions ($\tau \sim 0.3$) in late northern winter ($L_S \sim 332°$–348°). Because this experiment approximates well the conditions of our data subset in terms of both season and dust loading, it provides an opportunity to compare winds inferred from data and those simulated by the GCM.

Although westerlies appear in both hemispheres in the simulation, there is a much more pronounced interhemispheric difference in the strength of the jets than is seen in Figure 2. The peak velocity of the GCM northern hemisphere jet (centered near 50°N–60°N and 35 to 40–km altitude) exceeds 100 m/s, whereas the peak velocity of the GCM southern hemisphere jet (centered near 60°S and 40–km altitude) is only about 30 m/s. This difference in peak intensity of the westerly jets is reduced to a factor of two when GCM fields for a more limited seasonal range ($L_S \sim 343°$–$346°$) are examined (J. Barnes, personal communication, 1994), but in either case the GCM late southern summer southern hemisphere jet is considerably weaker than that obtained from analysis of Mariner 9 data. Sensitivity tests indicate that a major portion of this difference is attributable to the fact that GCM temperatures at high southern latitudes, above the near–surface region, are roughly 10 K warmer than those in Figure 1a. One possible explanation for warmer simulated temperatures in this region is that the GCM horizontal dust distribution is uniform (except for topography effects), whereas the retrieved horizontal dust distribution falls off rapidly poleward of 40°S (see Figure 1b). Larger airborne dust abundances at high southern latitudes in the model would produce increased solar heating and warmer temperatures in this region, leading to weaker latitudinal temperature gradients and less intense westerlies.

3.2. Net Radiative Heating Rates

Net radiative heating rates for relatively low–dust conditions ($\tau(9\mu m)) \sim 0.2$) are shown in Figure 3. At pressures below 4.0 mbar, there are large net heating rates (up to 5 K/d) in the equatorial region and large net cooling rates (up to 12 K/d) in the polar regions. The heating and cooling rates are approximately symmetric about the equator, although the polar cooling is much weaker in the southern hemisphere (where the sun is shining). Because no large temperature changes occur during the 12 days represented by this subset of IRIS data (see Paper I), these results imply the existence of a nonradiative mechanism which acts to cool the tropics (by up to 5 K/d) and warm the polar regions (by up to 12 K/d) in order to maintain the observed temperature structure in the presence of this radiative forcing. For example, upwelling over the equator would produce expansional cooling to offset the net radiative heating, while downwelling over the poles would produce compressional heating to offset the net radiative cooling.

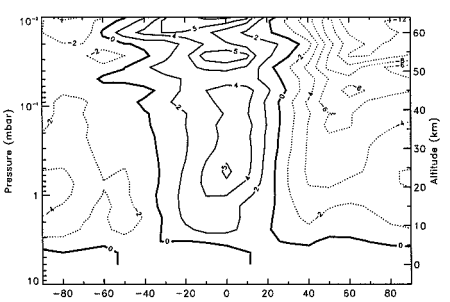


Figure 3. Net radiative heating rates (K/d) calculated from retrieved IRIS temperatures and dust abundances using the radiative transfer model described by *Crisp* [1990] and in the appendix of *Santee* [1993]. Negative contours are dashed, the zero contour is thicker, and the contour interval is nonuniform.

3.3. The Diabatic Circulation

3.3.1. Meridional and vertical velocities. The dynamical heat transport in the meridional plane that balances the net radiative heating is called the diabatic circulation. In the absence of wave transience or dissipation, the diabatic circulation represents the true net motion of air parcels. The diagnostic stream function model described in section 2.3 is used to simultaneously solve for the meridional and vertical components of the diabatic circulation. This model uses the diurnal–mean temperatures of Figure 1a and the net radiative heating rates of Figure 3. The resulting meridional and vertical velocities are depicted in Figures 4a and 4b, respectively. The mass–weighted stream function can be defined [*Reiter*, 1975]

$$\Psi_m = 2\pi a \rho_s e^{\frac{z}{H}} \psi \,. \tag{15}$$

The density of the atmosphere at the surface is $\rho_s = p_s/gH = 1.6 \times 10^{-2}$ kg/m^3. Equation (15) was used to construct the streamlines of mass flow, which are shown in Figure 4c.

Figures 4a–4c depict a two–cell diabatic circulation, with rising motion in the equatorial region and sinking motion over both poles. Figures 4a and 4c show transport from the equator toward the poles throughout most of the atmosphere, with return flow in the low-

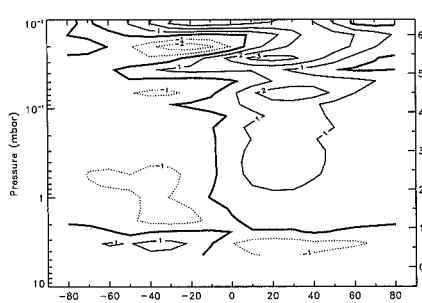


Figure 4a. Diabatic meridional velocity, in m/s. Positive values represent northward winds, negative contours are dashed, and the zero contour is thicker.

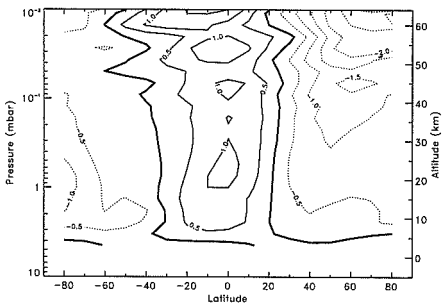


Figure 4b. Diabatic vertical velocity, in cm/s. Positive values represent upward winds, negative contours are dashed, and the zero contour is thicker.

est levels. The vertical velocity field is fairly symmetric (Figure 4b), although the southern hemisphere exhibits a bias toward upward motion, while the northern hemisphere exhibits a bias toward downward motion. This data set represents a transitional period (L_s = 343°–348°) between solstitial and equinoctial conditions. It is likely that the vertical velocity field becomes even more symmetric as the equinox approaches. Because the circulation is considerably weaker at equinox than at solstice [*Haberle et al.*, 1993], the proximity to equinox of this period suggests that the magnitudes of the diabatic velocities are near the lower limits reached over the course of this particular Mars year. However, although this period represents non–dust–storm conditions, the dust optical depth is still slightly above the background value, and model simulations indicate that the circulation intensifies with increased dust loading [*Haberle et al.*, 1982, 1993]. Therefore, the derived velocities may be larger than those for a more typical year at the same season.

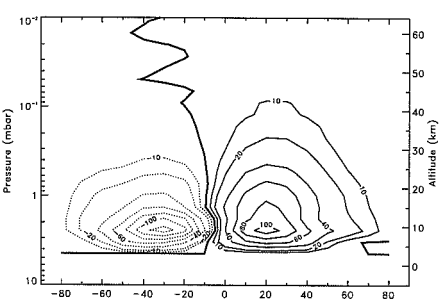


Figure 4c. Mass–weighted stream function, Ψ_m, in units of 10^7 kg/s. Positive values represent clockwise flow, negative contours are dashed, the zero contour is thicker, and the contour interval is nonuniform.

Comparison of Figure 3 and Figure 4b shows that the vertical velocity structure strongly reflects the diabatic heating distribution. The strength of the meridional circulation is directly proportional to the strength of the radiative drive because this circulation provides the necessary adiabatic heating and/or cooling compensation. The diabatic circulation presented in Figures 4a–4c is qualitatively similar to the terrestrial diabatic circulation at the comparable season, but it is more vigorous. Comparisons with the terrestrial diabatic circulation diagnosed for March or April by *Solomon et al.* [1986], *Rosenfield et al.* [1987], *Gille et al.* [1987], *Shine* [1989] and *Yang et al.* [1990] reveal that the late northern winter/late southern summer Martian circulation is roughly five times stronger. A similar factor is obtained from GCM results by *Haberle et al.* [1993].

The results of Figure 4 can be compared to those of *Moreau et al.* [1991], which were derived from a two–dimensional coupled chemical, radiative, and dynamical model of the dust–free Martian atmosphere. Moreau et al. obtain the meridional distribution of temperature at $L_S = 351°$ by explicitly including radiative transfer and dynamical heat transport in the solution of the thermodynamic energy equation. Their transport model describes a transformed Eulerian mean meridional circulation which is forced by gravity wave dissipation, but neglects the effects of atmospheric tides. Perhaps the most serious shortcoming of their model is that it completely neglects the radiative impact of airborne dust, which dominates the solar heating even during relatively "clear" periods like that examined here. Their radiative transfer model includes other limitations as well. For example, they use an empirical parameterization for the CO_2 absorption at near–infrared wavelengths which does not include Voigt effects, and consequently is not valid for the low–pressure Martian atmosphere. Their thermal cooling rate model also neglects Voigt effects, and above 40–km altitude employs a Newtonian cooling formulation. These simplifications result in modeled temperatures which are about 20 K cooler than those of Figure 1a throughout most of the atmosphere. They present a plot of the stream function, the general character of which agrees well with Figure 4c. Although they do not include plots of the velocity fields, they quote values for the maximum meridional velocity of 1.6 m/s at 50 km and 45°N latitude, and the maximum vertical velocity downward over the polar regions of 1.3 cm/s. The diabatic circulation computed from IRIS observations (Figure 4) is stronger than the circulation derived by *Moreau et al.* [1991]. We find a maximum meridional velocity of 3 m/s at about 55 km at 20°N latitude, and a maximum downward vertical velocity of 2.5 cm/s at 60 km over the north pole.

The Martian atmosphere contains substantial quantities of dust even during non–dust–storm conditions, and we account for the consequent enhancement of the radiative forcing in our diagnostic models. Because the thermal structure the atmosphere would have in the absence of dust is not an observable quantity, our diagnostic model cannot be used to infer the dust–free cir-

culation. However, model simulations [*Haberle et al.*, 1982, 1993] indicate that the introduction of dust into the atmosphere causes a significant strengthening of the circulation.

The stream function model results can also be compared to the Mars GCM results of *Haberle et al.* [1993] (see the discussion of this model in section 3.1). Haberle et al. present plots of the mass stream function and meridional and vertical velocities for the northern spring equinox ($L_S = 0°$–$23°$). This circulation consists of two cells of fairly comparable strength whose common rising branch is centered just south of the equator. The peak mass flux associated with these cells (10^9 kg/s) is in good agreement with the maximum contour of Figure 4c. The overall pattern of the vertical velocity conforms well to that of Figure 4b, although the GCM–generated field exhibits substantially more structure and smaller values than the stream function model results. The GCM meridional velocities are also smaller than results from the stream function model, and they tend to be negative over a larger range of latitudes.

3.3.2. EP flux divergence. We can now calculate the EP flux divergence from equation (8a) by assuming steady state conditions and neglecting drag. Figure 5 shows the quantity $(\rho a \cos\phi)^{-1}\nabla \cdot \mathcal{F}$, which is a mean zonal force per unit mass, scaled to units of ms^{-1}/d. At most levels of the atmosphere in both hemispheres, the EP flux divergence is negative. In these regions, the waves exert a negative force (i.e., a drag) on the zonal–mean flow. This forcing is consistent with an induced residual circulation having a poleward meridional velocity at high altitudes and a downward vertical velocity over the poles. Examination of the magnitudes of the individual terms on the left–hand side of equation (8a) reveals that the force produced by the mean meridional circulation is dominated by the Coriolis torque ($f\overline{v}^*$), although the contributions from the other terms are not negligible. For this reason, the $\nabla \cdot \mathcal{F}$ and $\overline{v}^*$ fields are strongly correlated. The magnitude of the EP flux divergence is also correlated with the values of $\overline{w}^*$ (cf. Figure 4b and Figure 5).

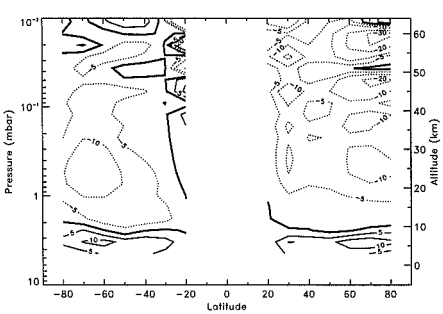


Figure 5. Eliassen–Palm flux divergence ($(\rho a \cos\phi)^{-1}\nabla \cdot \mathcal{F}$) in units of ms^{-1}/d. Negative contours are dashed, the zero contour is thicker, and the contour interval is nonuniform.

The EP flux divergence identifies regions where the noninteraction theorem is being violated. As discussed in section 2.2, in regions where the $\nabla \cdot \mathcal{F}$ values are large, the waves must be strongly transient, nonconservative, or nonlinear. In Figure 5, the largest (negative) values of $\nabla \cdot \mathcal{F}$ occur at high latitudes in the winter hemisphere at altitudes above about 40 km (~0.1 mbar). This coincides with the area distinguished by temperatures far in excess of the radiative equilibrium values (see Paper I; also *Deming et al.* [1986] and *Rothermel et al.* [1988]). The results of Figure 5 are consistent with those of *Barnes* [1990], who uses a simplified dynamical model of the zonal–mean flow to investigate the effects of breaking topographically forced internal gravity waves on the circulation in the middle Martian atmosphere. His model results show that breaking gravity waves induce little high–latitude warming at levels below ~40 km; however, they have a significant impact in the 40 to 80–km region, generating a strong mean meridional circulation that acts to produce very warm winter polar temperatures.

3.3.3. Overturning time. There is no consensus on the nomenclature for the different altitude regions of the Martian atmosphere. For the sake of analogy with terrestrial conventions, in this section we will call the region of the atmosphere with $p > 1$ mbar the "troposphere" (although this designation is sometimes given to the region from the ground to 45 km [*Zurek*, 1992]), and the region with $p < 1$ mbar the "stratosphere" (although it could also be called the mesosphere [*Zurek*, 1992]). Thus, the net meridional circulation described by Figures 4a–4c provides a mechanism for exchange between "tropospheric" and "stratospheric" air masses. The tropical branch of the circulation carries significant amounts of tropospheric air into the stratosphere, which, for reasons of mass continuity, must return to the troposphere at high latitudes. If advection by the net meridional circulation is the most important process involved in troposphere–stratosphere exchange, then the characteristic advective exchange time, or residence time, for parcels in the stratosphere is defined [*Reiter*, 1975; *Chamberlain and Hunten*, 1987]:

$$t_A^{-1} = -\frac{1}{\mathcal{M}}\frac{d\mathcal{M}}{dt} \tag{16}$$

where $\mathcal{M}$ is the total stratospheric air mass. We estimate the mass of the Martian "stratosphere" (the atmosphere above 1 mbar, or approximately 20 km) to be $\mathcal{M} = 4\pi(a + 20\text{ km})^2 \int_{20\text{ km}}^{\infty} \rho(z)dz \approx 3.9 \times 10^{15}$ kg.

The $d\mathcal{M}/dt$ term in equation (16) is estimated from the mass–weighted stream function Ψ_m. The physical meaning of Ψ_m can be illuminated by substituting equation (15) into equations (11a) and (11b) and rearranging:

$$\Psi_m(\phi, z) = -2\pi a \rho_s \cos\phi \int_0^z e^{\frac{-z'}{H}} \overline{v}^* \, dz' \tag{17a}$$

$$\Psi_m(\phi, z) = 2\pi a^2 \rho_s \, e^{\frac{z}{H}} \int_{-\pi/2}^{\phi} \cos\phi' \, \overline{w}^* \, d\phi' \,. \quad (17b)$$

The first expression represents the vertically integrated mass flux in the region from 0 to z across latitude ϕ, and the second expression represents the horizontally integrated mass flux in the region from $-\frac{\pi}{2}$ to ϕ across a pressure surface, denoted by the log-pressure coordinate z. The calculation of the characteristic timescale for troposphere–stratosphere exchange requires a value for the integrated mass flux across the 1-mbar pressure surface in the region of equatorial upwelling where the diabatic vertical velocity $\overline{w}^*$ is positive. Therefore, from Figures 4b and 4c and equation (17b), the term $d\mathcal{M}/dt$ can be approximated by $\Psi_m(20°\mathrm{N}, 1\,\mathrm{mbar}) - \Psi_m(40°\mathrm{S}, 1\,\mathrm{mbar}) \approx 1.2 \times 10^9$ kg/s. This value of $d\mathcal{M}/dt$ leads to a residence time $t_A = 3.2 \times 10^6$ s $\sim$ 38 days.

Because fluid parcels are constrained to travel around closed streamlines, to satisfy mass continuity, the horizontal and vertical transport times must be equal. If these large transport rates are sustained for an entire season, the Martian atmosphere above the 1-mbar level is ventilated in about 38 days. This value of the exchange time is roughly comparable to modeling results obtained by *Barnes* [1990]. Barnes computes vertical velocities of several centimeters per second, and concludes that the exchange time between the lower portion of the middle atmosphere (10–40 km) and the upper portion of the middle atmosphere (40–100 km) is 5–10 days. In sharp contrast, *Shia et al.* [1989] show that it takes about 1.8 years for the terrestrial stratosphere to be completely overturned.

3.3.4. Eddy diffusion coefficient. The simplest approach to modeling meridional and vertical transport in the Martian atmosphere is to parameterize these processes with eddy diffusion coefficients. Then the characteristic vertical diffusion time, the dynamical time associated with vertical mixing, is given by [*Chamberlain and Hunten*, 1987]

$$t_D = \frac{H^2}{K_{zz}} \,, \quad (18)$$

where K_{zz} is the vertical eddy diffusion coefficient. K_{zz} has been estimated by several authors using different methods. *Conrath* [1975] deduces the dust optical depth during the dissipation phase of the 1971 dust storm from Mariner 9 IRIS temperature profiles and a simplified atmospheric heating model. To explain the comparable exponential decay rates of the dust opacity at two different atmospheric levels, he requires a vertical eddy diffusion coefficient $K_{zz} \geq 10^3$ m^2/s between the surface and 50 km. *Zurek* [1976] estimates the eddy diffusion coefficient for the turbulence generated by the diurnal atmospheric tide in the middle Martian atmosphere (30–80 km) to be $K_{zz} \sim 5 \times 10^3$ m^2/s. *Kong and McElroy* [1977] investigate a variety of photochemical models and also conclude that a mixing coefficient of about $K_{zz} \sim 5 \times 10^3$ m^2/s at 40 km is required to match the observed abundances of various atmospheric constituents. *Toon et al.* [1977] use a multiple scattering radiation code to find the best fit synthetic spectra to Mariner 9 IRIS observations. They conclude that the size distribution of the dust particles did not change appreciably during the dissipation of the dust storm, and they invoke vertical mixing on the order of 10^3 m^2/s to explain the constancy of the distribution. *Anderson and Leovy* [1978] examine Mariner 9 television reflectance profiles which crossed the limb of Mars. From an analysis of the decay of dust scale heights they determine the vertical eddy mixing coefficient to be $K_{zz} \sim 10^3$ m^2/s in the 40 to 60-km height range.

Using a value of $K_{zz} = 10^3$ m^2/s in equation (18) yields $t_D = 10^5$ s ~ 1.2 days. This estimate of the transport timescale due to diffusion is an order of magnitude less than the estimate of the advective timescale. If correct, this would imply that vertical transport in the Martian atmosphere is primarily diffusive, rather than advective, in nature. In contrast, based on an upper limit of $K_{zz} = 1$ m^2/s for the stratosphere of the Earth from *Shia et al.* [1989], the characteristic diffusion time for the terrestrial stratosphere is 1.5 years, comparable to the timescale due to advection, and both processes are important for tracer transport.

In actuality, the temperature and net heating rate fields calculated from the Mariner 9 spectra provide strong evidence for the importance of advection in the Martian atmosphere because they require the upgradient meridional transport of heat at high altitudes. The determinations of the eddy diffusion coefficient outlined above parameterized all contributions to the vertical transport. The derived values of the eddy diffusion coefficient therefore include the effects of large-scale organized motions, such as the mean meridional circulation, and are consequently much larger than they would be if they represented only the small-scale eddy mixing processes. Although these large values of K_{zz} may be valid for the dustiest conditions or in limited regions, e.g., within the boundary layer, they are not appropriate for application on a global scale.

Recent photochemical models do not appear to require as strong vertical mixing in the lowest region of the atmosphere as the earlier calculations indicated [*Nair et al.*, 1994, and references therein]. *Moreau et al.* [1991] calculate a meridional cross section of K_{zz} for $L_S = 351°$ using a two-dimensional model which accounts for both advection due to the residual mean meridional circulation and eddy mixing associated with absorption of waves. As discussed in section 3.3.1, several simplifications in their radiative transfer model and wave parameterization scheme lead to a residual circulation that is weaker than that of Figure 4. They find that a value of $K_{zz} = 30$ m^2/s is representative of the majority of the atmosphere below 50-km altitude. This value of K_{zz} yields $t_D = 3.3 \times 10^6$ s $\sim$ 38 days, in agreement with the timescale for vertical advection calculated above. Using dynamical arguments, *Plumb and Mahlman* [1987] maintain that advective and dif-

fusive transport effects have comparable magnitudes in the terrestrial atmosphere (globally, but not necessarily locally). Although the information currently available is insufficient to prove conclusively whether either advective or diffusive processes dominate in the Martian atmosphere, the assumption that their transport timescales are equivalent leads to a vertical eddy diffusion coefficient two orders of magnitude smaller than the commonly used value.

3.3.5. Transport issues. The influence of the meridional circulation on a chemical species can be evaluated by comparing the dynamical time constant to the photochemical time constant, which is calculated for each species from its rate of loss. One of the most radiatively important trace gases in the Martian atmosphere is ozone. O_3 molecules are temporarily removed from the atmosphere by photolysis, but its overall abundance is not affected because of the rapid recombination of O and O_2. The four main reactions leading to the destruction of odd oxygen ($O + O_3$) in the Martian atmosphere are

$$HO_2 + O \rightarrow O_2 + OH \quad (19a)$$

$$O + OH \rightarrow O_2 + H \quad (19b)$$

$$O_3 + H \rightarrow O_2 + OH \quad (19c)$$

$$NO_2 + O \rightarrow NO + O_2 . \quad (19d)$$

The chemical lifetime for odd oxygen is then given by

$$t_C^{O_x} = \frac{[O_3] + [O]}{k_1[O][HO_2] + k_2[O][OH] + k_3[O_3][H] + k_4[NO_2][O]} \quad (20)$$

where the square brackets signify a concentration in molecules/cm^3, and k_1–k_4 are the rate constants for reactions (19a)–(19d), respectively.

These reactions are evaluated for midlatitude conditions using the Caltech/JPL one–dimensional photochemical model of the Martian atmosphere [*Yung et al.*, 1988; *Nair et al.*, 1994]. *Nair et al.* [1994] find that good agreement between modeled and observed CO, O_2, and O_3 can be achieved by varying the rate constants for a few key reactions within their published recommended uncertainties. All model results presented here are obtained using the modified rate constants. Two different water vapor profiles are considered: a low–water case with an integrated column abundance of 3.0 precipitable micrometers (pr μm), and a high–water case with an integrated column abundance of 8.0 pr μm. The water vapor is assumed to be well mixed up to the altitude where condensation first occurs, above which it roughly follows the saturation curve. Results from the two water cases are sufficiently similar for our purposes that only the high–water values are included here. Further details of the photochemical model initialization and constraints can be found in *Nair et al.* [1994]. Model results at 40–km altitude (H. Nair, personal communication, 1994) indicate that the chemical time constant for odd oxygen is about $t_C^{O_x} = 4200$ s = 1.2 hours. (For comparison, the photochemical lifetime of ozone based solely on photolysis is 625 s.) Because the photochemical time constant is much shorter than the dynamical time constant, ozone is in a state of photochemical equilibrium. Although the circulation may affect the ozone distribution indirectly, through the effects of temperature or by altering the concentration of water vapor (to which ozone is strongly coupled), it does not have a direct impact.

There are atmospheric constituents, e.g., CO, whose distributions are controlled by dynamics. The only significant loss mechanism for CO in the Martian atmosphere is

$$CO + OH \rightarrow CO_2 + H . \quad (21)$$

The chemical lifetime for CO $t_C^{CO} = 1.5 \times 10^7$ s $\sim$ 175 days (H. Nair, personal communication, 1994). CO should therefore be sufficiently long–lived to be thoroughly mixed by atmospheric motions. This conclusion is supported by the work of *Clancy et al.* [1990], who show that ground–based microwave spectra of CO in the atmosphere of Mars can be fit well by assuming a CO mixing ratio which is constant in altitude and latitude. However, photochemical model results [*Nair et al.*, 1994] indicate that the CO mixing ratio actually increases slightly with altitude for a K_{zz} profile varying exponentially from 10 to 1000 m^2/s between the surface and 40 km.

A similar comparison can be undertaken for airborne dust, although in this case the only removal process is gravitational settling. Following *Conrath* [1975], the terminal velocity w_T of the dust particles is given by the Stokes–Cunningham relation for spherical particles for the case where the molecular mean free path λ is large compared to the particle diameter d:

$$w_T = \left[\frac{\rho_d\, g\, d^2}{18\eta}\right]\left[1 + \frac{2}{d}\left(\lambda_r \frac{p_r}{p}\right)\right] . \quad (22)$$

The following values are assumed: dust particle density $\rho_d = 3 \times 10^3$ kg/m^3, molecular viscosity $\eta = 1.5 \times 10^{-5}$ kg/m/s, and reference pressure $p_r = 25$ mbar at which $\lambda_r = 2.2 \times 10^{-6}$ m. Once the fall velocity has been calculated for given values of d and p, the characteristic settling time is found by

$$t_S = \frac{H}{w_T} . \quad (23)$$

The characteristic settling time for a 1–μm–diameter particle at 1 mbar is $t_S \sim 25$ days. The fact that the settling time is comparable to the timescale for vertical advection implies that the dust is well mixed. Any 1–μm dust particles entrained by local dust storms at 1 mbar ($\sim$20 km, which is the upper limit on the local dust cloud height inferred by *Peterfreund and Kieffer* [1979]) have a residence time sufficiently long to allow for horizontal transport. Assuming the dust remains suspended for 25 days and 1 m/s is a representative meridional velocity, the dust particles would be trans-

ported 2100 km, or approximately 40% of the distance between the equator and the pole. In contrast, the characteristic settling time for a 5-μm-diameter particle varies from $t_S \sim 5$ days at 1 mbar, to $t_S \sim 20$ days at 5 mbar ($\sim$2 km). The ascending motion in the nonperturbed atmosphere, after the dissipation of the dust storm, is not intense enough to support dust particles of this size over large vertical extent.

4. Summary

We presented diagnostic methods for defining the circulation of the atmosphere from a knowledge of its thermal structure. We used diurnal averages of meridional cross sections of atmospheric temperature and dust optical depth retrieved from Mariner 9 IRIS thermal emission spectra for a 12-day interval in late southern summer ($L_S = 343°$–$348°$). The retrieval methods and the resulting global maps of temperature and dust optical depth are described by *Santee and Crisp* [1993].

Zonal-mean zonal winds were derived from the meridional gradients of the observed temperatures assuming gradient wind balance and zero surface zonal wind. Outside of the tropics in both hemispheres, the zonal winds are westerly, with intense midlatitude jets near 50-km altitude. In the southern tropics the winds are easterly at all altitudes, but there is only a small corresponding region of easterlies near 60 km in the northern tropics.

Solar heating and thermal cooling rates were computed from the diurnally averaged IRIS temperature and dust distributions using the radiative transfer model described by *Crisp* [1990] and in the appendix of *Santee* [1993]. At pressures below 4 mbar, there are large net radiative heating rates (up to 5 K/d) in the equatorial region and large net radiative cooling rates (up to 12 K/d) in the polar regions.

The diabatic meridional circulation was calculated from the net radiative heating rates using a diagnostic stream function model which solves for the vertical and meridional velocities simultaneously. At this season, the circulation consists of two cells, with rising motion over the equatorial region, poleward flow in both hemispheres, sinking motion over both polar regions, and return flow in the lowest atmospheric levels. The maximum meridional velocity is 3 m/s at about 55 km at 20°N latitude. The vertical velocity field is fairly symmetric, although the winds in the winter hemisphere are stronger, with a maximum downward vertical velocity of 2.5 cm/s at 60 km over the north pole. If these large transport rates are sustained for an entire season, the Martian atmosphere above the 1-mbar level is ventilated in about 38 days. The assumption that the advective and diffusive transport timescales should be comparable leads to a vertical eddy diffusion coefficient of about 30 m^2/s, two orders of magnitude smaller than the commonly used value.

The EP flux divergence was calculated from the zonal momentum equation. This quantity identifies regions where the noninteraction theorem is being violated. At most levels of the atmosphere in both hemispheres, the EP flux divergence is negative, indicating that the waves exert a drag on the zonal-mean flow. The largest values occur at high latitudes in the winter hemisphere at altitudes above 40 km, which coincides with the region of anomalously warm temperatures described by *Santee and Crisp* [1993].

This paper presents the first global estimates of the meridional circulation in the Martian atmosphere diagnosed from actual temperature observations. This diabatic circulation is qualitatively similar to that of Earth's atmosphere at the corresponding season, but is more vigorous. It is also more vigorous than the residual mean meridional circulation derived previously from a two-dimensional tracer transport model of the Martian atmosphere [*Moreau et al.*, 1991]. This diabatic meridional circulation places constraints on the atmospheric transport of dust in late southern summer, during non-dust-storm conditions. A comparison of characteristic settling and meridional transport times reveals that 1-μm dust particles can be transported over global scales, whereas 5-μm particles do not remain suspended in the atmosphere for periods sufficient for long-range transport.

Although this study has yielded insight into the magnitude and direction of atmospheric transport processes on Mars, it does not provide a complete picture of the circulation patterns. We have investigated one 12-day interval in late southern summer during a relatively clear period after a severe dust storm. It is uncertain how well the diagnosed circulation describes that of other seasons or other years; the proximity to the equinox of the study period suggests that the derived velocities may be smaller than average over the course of this particular Mars year, but the fact that the dust optical depth is still slightly above the background value suggests that the derived velocities may be larger than those for a more typical year at the same season. Even if all available spectra had been processed, the entire Mariner 9 data set encompasses only two Martian seasons, and for much of that time the IRIS data are too sparse to differentiate diurnal and longitudinal variations for the construction of reliable zonal means. A more comprehensive examination of the meridional circulation over an annual cycle must await new observations.

Appendix A: Solution to the Stream Function Equation

The solution to equation (13) is described here in some detail. To simplify the notation, A, B, C, D, and E are introduced for the coefficients of the ψ terms, and F is used for the right-hand side of equation (13):

$$A\psi + B\frac{\partial\psi}{\partial z} + C\frac{\partial\psi}{\partial\phi} + D\frac{\partial^2\psi}{\partial\phi^2} + E\frac{\partial^2\psi}{\partial z\partial\phi} = F \quad \text{(A1)}$$

The simplest approach to solving equation (A1) is to use finite differences to approximate the derivatives on a grid of points in ϕ and z. Centered finite difference

formulas are applied to both the ψ terms and their coefficients in equation (A1). The coefficients in their expanded finite difference forms are written as follows:

$$A_{m,n} = \frac{1}{aH\cos\phi_m}\left[\frac{\overline{T}_{m+1,n} - 2\overline{T}_{m,n} + \overline{T}_{m-1,n}}{0.25\,(\phi_{m+1} - \phi_{m-1})^2} + \tan\phi_m\left(\frac{\overline{T}_{m+1,n} - \overline{T}_{m-1,n}}{\phi_{m+1} - \phi_{m-1}}\right)\right] \quad \text{(A2a)}$$

$$B_{m,n} = -H\,A_{m,n} \quad \text{(A2b)}$$

$$C_{m,n} = \frac{1}{a\cos\phi_m}\left[\frac{1}{H}\frac{\overline{T}_{m+1,n} - \overline{T}_{m-1,n}}{\phi_{m+1} - \phi_{m-1}} + g\left(\frac{c_{p_{m+1,n}}^{-1} - c_{p_{m-1,n}}^{-1}}{\phi_{m+1} - \phi_{m-1}}\right) + \frac{\overline{T}_{m+1,n+1} - \overline{T}_{m-1,n+1} - \overline{T}_{m+1,n-1} + \overline{T}_{m-1,n-1}}{(z_{n+1} - z_{n-1})\,(\phi_{m+1} - \phi_{m-1})} + \tan\phi_m\left(\frac{\overline{T}_{m,n+1} - \overline{T}_{m,n-1}}{z_{n+1} - z_{n-1}} + \frac{g}{c_{p_{m,n}}}\right)\right] \quad \text{(A2c)}$$

$$D_{m,n} = \frac{1}{a\cos\phi_m}\left[\frac{\overline{T}_{m,n+1} - \overline{T}_{m,n-1}}{z_{n+1} - z_{n-1}} + \frac{g}{c_{p_{m,n}}}\right] \quad \text{(A2d)}$$

$$E_{m,n} = -\frac{1}{a\cos\phi_m}\left[\frac{\overline{T}_{m+1,n} - \overline{T}_{m-1,n}}{\phi_{m+1} - \phi_{m-1}}\right] \quad \text{(A2e)}$$

$$F_{m,n} = \frac{\overline{Q}_{m+1,n} - \overline{Q}_{m-1,n}}{\phi_{m+1} - \phi_{m-1}}\,. \quad \text{(A2f)}$$

Expanding the ψ terms in equation (A1) leads to

$$A_{m,n}\,\psi_{m,n} + B_{m,n}\left[\frac{\psi_{m,n+1} - \psi_{m,n-1}}{z_{n+1} - z_{n-1}}\right] + C_{m,n}\left[\frac{\psi_{m+1,n} - \psi_{m-1,n}}{\phi_{m+1} - \phi_{m-1}}\right] + D_{m,n}\left[\frac{\psi_{m+1,n} - 2\psi_{m,n} + \psi_{m-1,n}}{0.25(\phi_{m+1} - \phi_{m-1})^2}\right] + E_{m,n}\left[\frac{\psi_{m+1,n+1} - \psi_{m+1,n-1}}{(z_{n+1} - z_{n-1})(\phi_{m+1} - \phi_{m-1})} - \frac{\psi_{m-1,n+1} - \psi_{m-1,n-1}}{(z_{n+1} - z_{n-1})(\phi_{m+1} - \phi_{m-1})}\right] = F_{m,n}\,. \quad \text{(A3)}$$

Rearranging this equation so that like ψ terms are grouped together and defining $\Delta z \equiv z_{n+1} - z_{n-1}$ and $\Delta\phi \equiv \phi_{m+1} - \phi_{m-1}$ results in

$$\frac{E_{m,n}}{\Delta z\Delta\phi}\psi_{m-1,n-1} - \frac{B_{m,n}}{\Delta z}\psi_{m,n-1} - \frac{E_{m,n}}{\Delta z\Delta\phi}\psi_{m+1,n-1} + \left[\frac{4D_{m,n}}{\Delta\phi^2} - \frac{C_{m,n}}{\Delta\phi}\right]\psi_{m-1,n} + \left[A_{m,n} - \frac{8D_{m,n}}{\Delta\phi^2}\right]\psi_{m,n} + \left[\frac{C_{m,n}}{\Delta\phi} + \frac{4D_{m,n}}{\Delta\phi^2}\right]\psi_{m+1,n} - \frac{E_{m,n}}{\Delta z\Delta\phi}\psi_{m-1,n+1} + \frac{B_{m,n}}{\Delta z}\psi_{m,n+1} + \frac{E_{m,n}}{\Delta z\Delta\phi}\psi_{m+1,n+1} = F_{m,n}\,. \quad \text{(A4)}$$

The introduction of a second set of coefficients will again simplify the notation:

$$a_{m,n}\,\psi_{m-1,n-1} + b_{m,n}\,\psi_{m,n-1} + c_{m,n}\,\psi_{m+1,n-1} + d_{m,n}\,\psi_{m-1,n} + e_{m,n}\,\psi_{m,n} + f_{m,n}\,\psi_{m+1,n} + g_{m,n}\,\psi_{m-1,n+1} + h_{m,n}\,\psi_{m,n+1} + i_{m,n}\,\psi_{m+1,n+1} = F_{m,n}\,. \quad \text{(A5)}$$

Equation (A5) represents a system of simultaneous equations for the ψ that can be compactly expressed using matrix notation:

$$\mathbf{A}_n\,\vec{\psi}_{n-1} + \mathbf{B}_n\,\vec{\psi}_n + \mathbf{C}_n\,\vec{\psi}_{n+1} = \vec{\mathbf{f}}_n \quad \text{(A6)}$$

where $\vec{\psi}_n$ and $\vec{\mathbf{f}}_n$ are $M \times 1$ column vectors composed of the individual ψ and f:

$$\vec{\psi}_n = \begin{bmatrix}\psi_{1,n}\\ \psi_{2,n}\\ \vdots\\ \psi_{M,n}\end{bmatrix}, \qquad \vec{\mathbf{f}}_n = \begin{bmatrix}f_{1,n}\\ f_{2,n}\\ \vdots\\ f_{M,n}\end{bmatrix} \quad \text{(A7)}$$

and $\mathbf{A}_n$, $\mathbf{B}_n$, and $\mathbf{C}_n$ are $M \times M$ tridiagonal matrices composed of the coefficients of the ψ terms:

$\mathbf{A}_n =$

$$\begin{bmatrix} b_{1,n} & c_{1,n} & 0 & \cdots & \cdots & 0\\ a_{2,n} & b_{2,n} & c_{2,n} & 0 & \cdots & 0\\ 0 & a_{3,n} & b_{3,n} & c_{3,n} & \cdots & 0\\ \vdots & 0 & \ddots & \ddots & \ddots & \vdots\\ \vdots & \vdots & \cdots & a_{M-1,n} & b_{M-1,n} & c_{M-1,n}\\ 0 & 0 & 0 & \cdots & a_{M,n} & b_{M,n}\end{bmatrix} \quad \text{(A8a)}$$

$\mathbf{B}_n =$

$$\begin{bmatrix} e_{1,n} & f_{1,n} & 0 & \cdots & \cdots & 0\\ d_{2,n} & e_{2,n} & f_{2,n} & 0 & \cdots & 0\\ 0 & d_{3,n} & e_{3,n} & f_{3,n} & \cdots & 0\\ \vdots & 0 & \ddots & \ddots & \ddots & \vdots\\ \vdots & \vdots & \cdots & d_{M-1,n} & e_{M-1,n} & f_{M-1,n}\\ 0 & 0 & 0 & \cdots & d_{M,n} & e_{M,n}\end{bmatrix} \quad \text{(A8b)}$$

$$\mathbf{C}_n = \begin{bmatrix} h_{1,n} & i_{1,n} & 0 & \dots & \dots & 0 \\ g_{2,n} & h_{2,n} & i_{2,n} & 0 & \dots & 0 \\ 0 & g_{3,n} & h_{3,n} & i_{3,n} & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \dots & g_{M-1,n} & h_{M-1,n} & i_{M-1,n} \\ 0 & 0 & 0 & \dots & g_{M,n} & h_{M,n} \end{bmatrix} \tag{A8c}$$

Of course, equations (A6)–(A8) hold for all N grid levels, and are therefore also more compactly expressed in matrix form:

$$\begin{bmatrix} \mathbf{B}_1 & \mathbf{C}_1 & 0 & \dots & \dots & 0 \\ \mathbf{A}_2 & \mathbf{B}_2 & \mathbf{C}_2 & 0 & \dots & 0 \\ 0 & \mathbf{A}_3 & \mathbf{B}_3 & \mathbf{C}_3 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \dots & \mathbf{A}_{N-1} & \mathbf{B}_{N-1} & \mathbf{C}_{N-1} \\ 0 & 0 & 0 & \dots & \mathbf{A}_N & \mathbf{B}_N \end{bmatrix} \begin{bmatrix} \vec{\psi}_1 \\ \vec{\psi}_2 \\ \vec{\psi}_3 \\ \vdots \\ \vec{\psi}_{N-1} \\ \vec{\psi}_N \end{bmatrix} = \begin{bmatrix} \vec{\mathbf{f}}_1 \\ \vec{\mathbf{f}}_2 \\ \vec{\mathbf{f}}_3 \\ \vdots \\ \vec{\mathbf{f}}_{N-1} \\ \vec{\mathbf{f}}_N \end{bmatrix} \tag{A9}$$

The supermatrix on the left-hand side is an $N \times N$ tridiagonal matrix whose elements are themselves tridiagonal matrices. Because this is an ill-conditioned matrix, special precautions must be taken in inverting it to solve for the ψ. Y. Yung (personal communication, 1991) provided a routine for solving this set of equations; its structure is similar to that of programs designed to invert tridiagonal matrices with scalar elements [*Press et al.*, 1989].

Once the tridiagonal solver has yielded values for the ψ, the definition of the stream function is used to calculate the residual circulation:

$$\begin{aligned} \overline{v}^*{}_{m,n} &= \frac{1}{\cos\phi_m}\left[\frac{\psi_{m,n}}{H} - \frac{\psi_{m,n+1} - \psi_{m,n-1}}{z_{n+1} - z_{n-1}}\right] \\ \overline{w}^*{}_{m,n} &= \frac{1}{a\cos\phi_m}\left[\frac{\psi_{m+1,n} - \psi_{m-1,n}}{\phi_{m+1} - \phi_{m-1}}\right] \end{aligned} \tag{A10}$$

Appendix B: Optical Properties of Dust in the Martian Atmosphere

Mariner 9 IRIS observations provide the best existing constraints on the wavelength dependence of the dust optical properties at thermal infrared wavelengths, but no quantitative information about the dust optical properties at solar wavelengths. Subsequent spacecraft observations of the dusty Martian atmosphere have yielded additional insights into these properties, but only in a relatively narrow range of visible wavelengths [*Pollack et al.*, 1977; *Pollack*, 1982; *Drossart et al.*, 1991; *Clancy and Lee*, 1991]. Much less is known about the dust properties at other wavelengths. The minimum spectral information needed to assess the impact of dust on the net radiative forcing of the Martian atmosphere is the ratio of the broadband solar and thermal optical depths, but even this ratio is uncertain by about a factor of two [cf. *Zurek*, 1982; *Martin*, 1986; *Toon et al.*, 1977; *Clancy et al.*, this issue].

For this study, we combine data from a variety of sources to construct a more complete wavelength-dependent description of the dust optical properties. At wavelengths longer than 5 μm, we assume that palagonite is an acceptable spectral analog for Mars dust, and we adopt the refractive index measurements of *Roush et al.* [1991]. These data are used in a Mie scattering model [*Crisp*, 1986] to derive the wavelength-dependent extinction efficiencies, single scattering albedos, and phase functions for the particle size distributions published by *Toon et al.* [1977], *Drossart et al.* [1991], and *Clancy and Lee* [1991].

This approach is not possible at shorter wavelengths, where refractive index data for palagonite are not available. We therefore derive these quantities from observations of the albedo of the Martian atmosphere during dust-storm conditions, including ground-based 0.3 to 1.0-μm spectrophotometric measurements [*McCord et al.*, 1977] and Viking Orbiter infrared thermal mapper (IRTM) visible channel measurements [*Christensen*, 1988]. At wavelengths where no dust-storm observations are available, we assume that the albedo of the dusty Martian atmosphere is similar to that of a relatively bright region on the Martian surface [*McCord et al.*, 1977]. The dusty-atmosphere albedo spectra compiled from these data are shown in Figure A1.

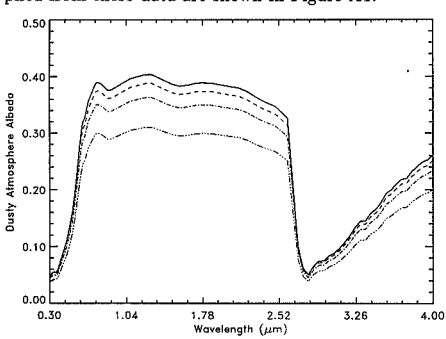


Figure A1. Albedos of a dusty Martian atmosphere used to derive dust optical properties. The nominal albedos (solid line) are near 0.4 at solar wavelengths, consistent with Viking IRTM observations. The other curves are scaled by constant factors to yield solar albedos near 0.375 (dashed line), 0.35 (dash–dot line), and 0.3 (dash–triple–dot line).

To derive the wavelength–dependent optical constants (n, k) from these measurements, we developed a simple retrieval algorithm consisting of the Mie scattering model, a multiple scattering model [*Stamnes et al.*, 1988; *Bell and Crisp*, 1993], and an iteration scheme based on Newton's method. This approach produces physically reasonable single scattering optical properties for each of the dust size distributions listed above. The wavelength–dependent extinction efficiency Q_{eff}, single scattering albedo ω_o, and particle scattering asymmetry parameter g are shown in Figures A2–A4.

The absorption coefficients corresponding to the imaginary index values for the *Toon et al.* [1977] size distribution provide the best agreement with available laboratory measurements of terrestrial palagonites [*Clark et al.*, 1990]. Surface radiance calculations using the optical properties derived for this size distribution also produce the best fit to the radiation field measured by the Viking Lander cameras [*Pollack et al.*, 1977]. Although these methods cannot yield a truly unique description of the Mars dust, they should be adequate for estimating the dust solar forcing, since they are consistent with the available observational constraints both at the top of the Martian atmosphere and at the surface.

These dust optical properties and the atmospheric temperatures and dust distributions derived from Mariner 9 IRIS observations (Figure 1) are used in the radiative transfer model [*Crisp*, 1990; *Santee*, 1993] to generate a two–dimensional (latitude/altitude) description of the solar heating rates, the thermal cooling rates, and the net radiative heating rates for late southern summer on Mars. The IRIS observations for this season indicate 9–μm dust optical depths between 0.2 and 0.3 above 6 mbar (Figure 1b). Because the optical depths are specified at 9 μm in our calculations, the thermal cooling rates are relatively insensitive to the size distribution. However, the solar heating rates are strongly dependent on the size distribution, because this property determines the ratio of the solar and thermal optical depths. For the optical constants and size distributions described above, 9–μm optical depths near 0.2 correspond to solar optical depths between 0.35 and 0.6. Size distributions consisting primarily of smaller particles [e.g., *Clancy and Lee*, 1991; *Drossart et al.*, 1991] produce the largest solar/thermal ratios, and the largest solar heating rates. Size distributions with larger particles [e.g., *Toon et al.*, 1977] produce much smaller solar/thermal ratios and solar heating rates for the observed 9–μm dust amounts.

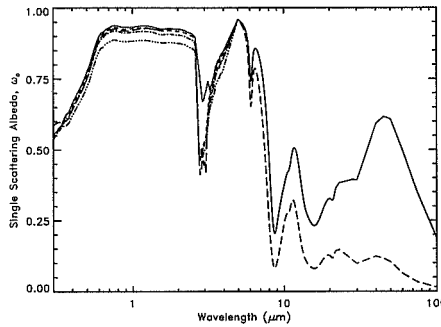


Figure A3. Single scattering albedos for the *Toon et al.* [1977] size distribution and mean solar albedos near 0.4 (solid line), 0.375 (short–dash line), 0.35 (dash–dot line), and 0.3 (dash–triple–dot line). Values for the smaller *Drossart et al.* [1991] distribution and a mean visible albedo of 0.4 are also shown (long–dash line).

No large variations in the globally averaged temperatures are observed during the late southern summer period studied here. The globally averaged solar heating rates and thermal cooling rates should therefore be roughly equal. For the smaller size distributions, we calculate globally averaged solar heating rates that exceed the globally averaged thermal cooling rates by almost

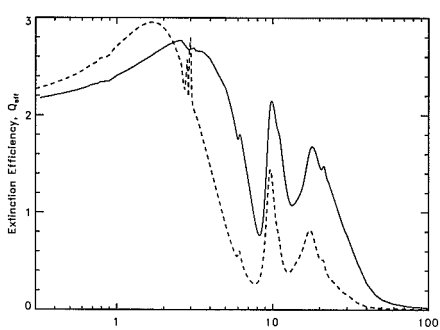


Figure A2. Dust extinction efficiencies for size distributions with large (solid line; *Toon et al.* [1977]) and small (dashed line; *Drossart et al.* [1991]) particles.

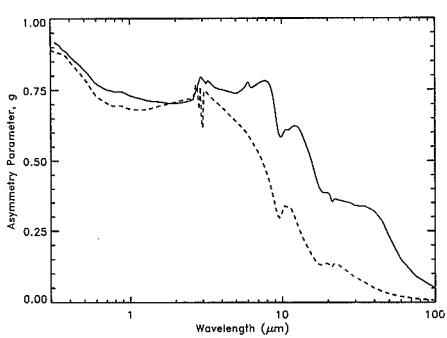


Figure A4. Scattering asymmetry parameter for size distributions with large (solid line; *Toon et al.* [1977]) and small (dashed line; *Drossart et al.* [1991]) particles.

a factor of two. Only simulations that use dust optical properties derived for the highest visible albedos and largest particles roughly satisfy global balance. We therefore adopt the optical properties derived for dusty-atmosphere visible albedos near 0.4 and the *Toon et al.* [1977] size distribution as our nominal values.

Acknowledgments. We would like to thank Y. Yung for providing the tridiagonal solver and H. Nair for discussions concerning the CIT/JPL photochemical model results. We are especially grateful to R.-L. Shia for suggesting how to solve the stream function equation. Insightful and constructive reviews by J. Barnes and C. Leovy were greatly appreciated. M.L.S. gratefully acknowledges support from the NASA Graduate Student Researchers Program. Some of this work was funded by grants from the NASA Planetary Atmospheres Program to the Jet Propulsion Laboratory of the California Institute of Technology. We also acknowledge computer support from the Hubble Space Telescope WF/PC2 Program. This work represents partial fulfillment of the requirements for a Ph.D. degree in Planetary Science at Caltech. Division of Geological and Planetary Sciences, California Institute of Technology, contribution 5383.

References

Anderson, E., and C. Leovy, Mariner 9 television limb observations of dust and ice hazes on Mars, *J. Atmos. Sci.*, *35*, 723–734, 1978.

Andrews, D. G., and M. E. McIntyre, Planetary waves in horizontal and vertical shear: The generalized Eliassen–Palm relation and the mean zonal acceleration, *J. Atmos. Sci.*, *33*, 2031–2048, 1976.

Andrews, D. G., and M. E. McIntyre, Generalized Eliassen–Palm and Charney–Drazin theorems for waves on axisymmetric mean flows in compressible atmospheres, *J. Atmos. Sci.*, *35*, 175–185, 1978.

Andrews, D. G., J. R. Holton, and C. B. Leovy, *Middle Atmosphere Dynamics*, 489 pp., Academic, San Diego, Calif., 1987.

Barnes, J. R., Time spectral analysis of midlatitude disturbances in the Martian atmosphere, *J. Atmos. Sci.*, *37*, 2002–2015, 1980.

Barnes, J. R., Midlatitude disturbances in the Martian atmosphere: a second Mars year, *J. Atmos. Sci.*, *38*, 225–234, 1981.

Barnes, J. R., Possible effects of breaking gravity waves on the circulation of the middle atmosphere of Mars, *J. Geophys. Res.*, *95*, 1401–1421, 1990.

Barnes, J. R., A simple nearly analytic model of a gravity wave driven middle atmosphere circulation, *J. Atmos. Sci.*, *48*, 225–235, 1991.

Barnes, J. R., J. B. Pollack, R. M. Haberle, C. B. Leovy, R. W. Zurek, H. Lee, and J. Schaeffer, Mars atmospheric dynamics as simulated by the NASA Ames General Circulation Model, II, Transient baroclinic eddies, *J. Geophys. Res.*, *98*, 3125–3148, 1993.

Bell, J. F., III, and D. Crisp, Ground-based imaging spectroscopy in the near infrared: Preliminary results, *Icarus*, *104*, 2–19, 1993.

Briggs, G. A., and C. B. Leovy, Mariner 9 observations of the Mars north polar hood, *Bull. Am. Meteorol. Soc.*, *55*, 278–296, 1974.

Briggs, G. A., K. Klaasen, T. Thorpe, J. Wellman, and W. Baum, Martian dynamical phenomena during June–November 1976: Viking Orbiter imaging results, *J. Geophys. Res.*, *82*, 4121–4149, 1977.

Chamberlain, J. W., and D. M. Hunten, *Theory of Planetary Atmospheres: An Introduction to their Physics and Chemistry*, 481 pp., Academic, San Diego, Calif., 1987.

Christensen, P. R., Global albedo variations on Mars: Implications for active aeolian transport, *J. Geophys. Res.*, *93*, 7611–7624, 1988.

Clancy, R. T., and S. W. Lee, A new look at dust and clouds in the Mars atmosphere: Analysis of emission–phase–function sequences from global Viking IRTM observations, *Icarus*, *93*, 135–158, 1991.

Clancy, R. T., D. O. Muhleman, and G. L. Berge, Global changes in the 0–70 km thermal structure of the Mars atmosphere derived from 1975 to 1989 microwave CO spectra, *J. Geophys. Res.*, *95*, 14,543–14,554, 1990.

Clancy, R. T., S. W. Lee, G. R. Gladstone, W. W. McMillan, and T. Roush, A new model for Mars atmospheric dust based upon analysis of ultraviolet through infrared observations from Mariner 9, Viking, and Phobos, *J. Geophys. Res.*, this issue.

Clark, R. N., G. A. Swayze, R. B. Singer, and J. B. Pollack, High-resolution reflectance spectra of Mars in the 2.3 μm region: Evidence for the mineral Scapolite, *J. Geophys. Res.*, *95*, 14,463–14,480, 1990.

Conrath, B. J., Thermal structure of the Martian atmosphere during the dissipation of the dust storm of 1971, *Icarus*, *24*, 36–46, 1975.

Conrath, B. J., Planetary-scale wave structure in the Martian atmosphere, *Icarus*, *48*, 246–255, 1981.

Crisp, D., Radiative forcing of the Venus mesosphere, Ph.D. thesis, Princeton Univ., Princeton, N. J., 1983.

Crisp, D., Radiative forcing of the Venus mesosphere, I, solar fluxes and heating rates, *Icarus*, *67*, 484–514, 1986.

Crisp, D., Infrared radiative transfer in the dust-free Martian atmosphere, *J. Geophys. Res.*, *95*, 14,577–14,588, 1990.

Deming, D., M. J. Mumma, F. Espenak, and T. Kostiuk, Polar warming in the middle atmosphere of Mars, *Icarus*, *66*, 366–379, 1986.

Drossart, P., J. Rosenqvist, S. Erard, Y. Langevin, J.-P. Bibring, and M. Combes, Martian aerosol properties from the Phobos ISM experiment, *Ann. Geophys.*, *9*, 754–760, 1991.

Dunkerton, T., On the mean meridional mass motions of the stratosphere and mesosphere, *J. Atmos. Sci.*, *35*, 2325–2333, 1978.

Dunkerton, T., A Lagrangian mean theory of wave, mean-flow interaction with applications to nonacceleration and its breakdown, *Rev. Geophys.*, *18*, 387–400, 1980.

Edmon, H. J., B. J. Hoskins, and M. E. McIntyre, Eliassen–Palm cross sections for the troposphere, *J. Atmos. Sci.*, *37*, 2600–2616, 1980.

French, R. G., P. J. Gierasch, B. D. Popp, and R. J. Yerdon, Global patterns in cloud forms on Mars, *Icarus*, *45*, 468–493, 1981.

Gille, J. C., L. V. Lyjak, and A. K. Smith, The global residual mean circulation in the middle atmosphere for the northern winter period, *J. Atmos. Sci.*, *44*, 1437–1452, 1987.

Haberle, R. M., C. B. Leovy, and J. B. Pollack, A numerical model of the Martian polar cap winds, *Icarus*, *39*, 151–183, 1979.

Haberle, R. M., C. B. Leovy, and J. B. Pollack, Some effects of global dust storms on the atmospheric circulation of Mars, *Icarus*, *50*, 322–367, 1982.

Haberle, R. M., J. B. Pollack, J. R. Barnes, R. W. Zurek, C. B. Leovy, J. R. Murphy, H. Lee, and J. Schaeffer, Mars atmospheric dynamics as simulated by the NASA/Ames general circulation model, 1, The zonal-mean circulation, *J. Geophys. Res.*, *98*, 3093–3123, 1993.

Haltiner, G. J., and R. T. Williams, *Numerical Prediction and Dynamic Meteorology*, 477 pp., John Wiley, New York, 1980.

Hantel, M., and H.-R. Baader, Diabatic heating climatology of the zonal atmosphere, *J. Atmos. Sci.*, *35*, 1180–1189, 1978.

Hess, S. L., R. M. Henry, C. B. Leovy, J. A. Ryan, and J. E. Tillman, Meteorological results from the surface of Mars: Viking 1 and 2, *J. Geophys. Res.*, *82*, 4559–4574, 1977.

Hess, S. L., R. M. Henry, and J. E. Tillman, The seasonal variation of atmospheric pressure on Mars as affected by the south polar cap, *J. Geophys. Res.*, *84*, 2923–2927, 1979.

Hitchman, M. H., and C. B. Leovy, Evolution of the zonal mean state in the equatorial middle atmosphere during October 1978 – May 1979, *J. Atmos. Sci.*, *43*, 3159–3176, 1986.

Holton, J. R., The influence of gravity wave breaking on the general circulation of the middle atmosphere, *J. Atmos. Sci.*, *40*, 2497–2507, 1983.

Holton, J. R., and W. M. Wehrbein, A numerical model of the zonal mean circulation of the middle atmosphere, *Pure Appl. Geophys.*, *118*, 285–306, 1980.

Hsu, C.-P., Air parcel motions during a numerically simulated sudden stratospheric warming, *J. Atmos. Sci.*, *37*, 2768–2792, 1980.

Hunt, B. G., and S. Manabe, Experiments with a stratospheric general circulation model, II, Large–scale diffusion of tracers in the stratosphere, *Mon. Weather Rev.*, *96*, 503–539, 1968.

Kahn, R., Some observational constraints on the global–scale wind systems of Mars, *J. Geophys. Res.*, *88*, 10,189–10,209, 1983.

Kong, T. Y., and M. B. McElroy, Photochemistry of the Martian atmosphere, *Icarus*, *32*, 168–189, 1977.

Lellouch, E., J. J. Goldstein, S. W. Bougher, G. Paubert, and J. Rosenqvist, First absolute wind measurements in the middle atmosphere of Mars, *Astrophys. J.*, *383*, 401–406, 1991.

Leovy, C. B., Simple models of thermally driven mesospheric circulation, *J. Atmos. Sci.*, *21*, 327–341, 1964.

Leovy, C. B., Martian meteorological variability, *Adv. Space Res.*, *2*, 19–44, 1982.

Leovy, C. B., G. A. Briggs, A. T. Young, B. A. Smith, J. B. Pollack, E. N. Shipley, and R. L. Wildey, The Martian atmosphere: Mariner 9 television experiment progress report, *Icarus*, *17*, 373–393, 1972.

Mahlman, J. D., D. G. Andrews, D. L. Hartmann, T. Matsuno, and R. G. Murgatroyd, Transport of trace constituents in the stratosphere, in *Dynamics of the Middle Atmosphere*, edited by J. R. Holton and T. Matsuno, pp. 387–416, Terra Scientific Pub. Co., Boston, 1984.

Marks, C. J., Some features of the climatology of the middle atmosphere revealed by Nimbus 5 and 6, *J. Atmos. Sci.*, *46*, 2485–2508, 1989.

Martin, T. Z., Thermal infrared opacity of the Martian atmosphere, *Icarus*, *66*, 2–21, 1986.

Matsuno, T., Lagrangian motion of air parcels in the stratosphere in the presence of planetary waves, *Pure Appl. Geophys.*, *118*, 188–216, 1980.

McCord, T. B., R. L. Huguenin, D. Mink, and C. Pieters, Spectral reflectance of Martian areas during the 1973 opposition: Photoelectric filter photometry 0.33–1.10 μm, *Icarus*, *31*, 25–39, 1977.

Michelangeli, D. V., R. W. Zurek, and L. S. Elson, Barotropic instability of midlatitude zonal jets on Mars, Earth, and Venus, *J. Atmos. Sci.*, *44*, 2031–2041, 1987.

Moreau, D., L. W. Esposito, and G. Brasseur, The chemical composition of the dust–free Martian atmosphere: Preliminary results of a two–dimensional model, *J. Geophys. Res.*, *96*, 7933–7945, 1991.

Nair, H., M. Allen, A. D. Anbar, Y. L. Yung, and R. T. Clancy, A photochemical model of the Martian atmosphere, *Icarus*, *111*, 124–150, 1994.

Peterfreund, A. R., and H. H. Kieffer, Thermal infrared properties of the Martian atmosphere, 3, Local dust clouds, *J. Geophys. Res.*, *84*, 2853–2863, 1979.

Plumb, R. A., and J. D. Mahlman, The zonally averaged transport characteristics of the GFDL general circulation/transport model, *J. Atmos. Sci.*, *44*, 298–327, 1987.

Pollack, J. B., Properties of dust in the Martian atmosphere and its effect on temperature structure, *Adv. Space Res.*, *2*, 45–56, 1982.

Pollack, J. B., D. Colburn, R. Kahn, J. Hunter, W. Van Camp, C. E. Carlston, and M. R. Wolf, Properties of aerosols in the Martian atmosphere, as inferred from Viking lander imaging data, *J. Geophys. Res.*, *82*, 4479–4495, 1977.

Pollack, J. B., C. B. Leovy, P. W. Greiman, and Y. Mintz, A Martian general circulation experiment with large topography, *J. Atmos. Sci.*, *38*, 3–29, 1981.

Pollack, J. B., R. M. Haberle, J. Schaeffer, and H. Lee, Simulations of the general circulation of the Martian atmosphere, 1, Polar processes, *J. Geophys. Res.*, *95*, 1447–1473, 1990.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing*, 702 pp., Cambridge University Press, New York, 1989.

Reiter, E. R., Stratospheric–tropospheric exchange processes, *Rev. Geophys.*, *13*, 459–474, 1975.

Rosenfield, J. E., M. R. Schoeberl, and M. A. Geller, A computation of the stratospheric diabatic circulation using an accurate radiative transfer model, *J. Atmos. Sci.*, *44*, 859–876, 1987.

Rothermel, H., H. U. Käufl, U. Schrey, and S. Drapatz, Thermal structure of the Martian mesosphere, *Astron. Astrophys.*, *196*, 296–300, 1988.

Roush, T., J. Pollack, and J. Orenberg, Derivation of mid-infrared (5–25 μm) optical constants of some silicates and palagonite, *Icarus*, *94*, 191–208, 1991.

Ryan, J. A., and R. M. Henry, Mars atmospheric phenomena during major dust storms, as measured at surface, *J. Geophys. Res.*, *84*, 2821–2829, 1979.

Ryan, J. A., R. M. Henry, S. L. Hess, C. B. Leovy, J. E. Tillman, and C. Walcek, Mars meteorology: Three seasons at the surface, *Geophys. Res. Lett.*, *5*, 715–718, 1978.

Santee, M., The thermal structure, dust loading, and meridional transport in the Martian atmosphere during late southern summer, Ph.D. thesis, Calif. Inst. of Technol., Pasadena, 1993.

Santee, M., and D. Crisp, The thermal structure and dust loading of the Martian atmosphere during late southern summer: Mariner 9 revisited, *J. Geophys. Res.*, *98*, 3261–3279, 1993.

Shia, R.-L., Y. L. Yung, M. Allen, R. W. Zurek, and D. Crisp, Sensitivity study of advection and diffusion coefficients in a two-dimensional stratospheric model using excess carbon 14 data, *J. Geophys. Res.*, *94*, 18,467–18,484, 1989.

Shine, K., Sources and sinks of zonal momentum in the middle atmosphere diagnosed using the diabatic circulation, *Q. J. R. Meteorol. Soc.*, *115*, 265–292, 1989.

Solomon, S., J. T. Kiehl, R. R. Garcia, and W. Grose, Tracer transport by the diabatic circulation deduced from satellite observations, *J. Atmos. Sci.*, *43*, 1603–1617, 1986.

Stamnes, K., S. C. Tsay, W. Wiscombe, and K. Jayaweera,

Numerically stable algorithm for discrete–ordinate–method radiative transfer in multiple scattering and emitting layered media, *Appl. Opt.*, *27*, 2502–2509, 1988.

Tillman, J. E., R. M. Henry, and S. L. Hess, Frontal systems during passage of the Martian north polar hood over the Viking Lander 2 site prior to the first 1977 dust storm, *J. Geophys. Res.*, *84*, 2947–2955, 1979.

Toon, O. B., J. B. Pollack, and C. Sagan, Physical properties of the particles composing the Martian dust storm of 1971–1972, *Icarus*, *30*, 663–696, 1977.

Touloukian, Y. S., and T. Makita, Specific heat: Nonmetallic liquids and gases, in *Thermophysical Properties of Matter 6*, edited by Y. S. Touloukian, pp. 143–151, IFI/Plenum, New York, 1970.

Wallace, J. M., and P. V. Hobbs, *Atmospheric Science: An Introductory Survey*, 467 pp., Academic, San Diego, Calif., 1977.

Yang, H., K. K. Tung, and E. Olaguer, Nongeostrophic theory of zonally averaged circulation, II, Eliassen–Palm flux divergence and isentropic mixing coefficient, *J. Atmos. Sci.*, *47*, 215–241, 1990.

Yung, Y. L., J.–S. Wen, J. P. Pinto, M. Allen, K. K. Pierce, and S. Paulson, HDO in the Martian atmosphere: Implications for the abundance of crustal water, *Icarus*, *76*, 146–159, 1988.

Zurek, R. W., Diurnal tide in the Martian atmosphere, *J. Atmos. Sci.*, *33*, 321–337, 1976.

Zurek, R. W., Martian great dust storms: An update, *Icarus*, *50*, 288–310, 1982.

Zurek, R. W., Comparative aspects of the climate of Mars: An introduction to the current atmosphere, in *Mars*, edited by H. H. Kieffer, B. M. Jakosky, C. W. Snyder, and M. S. Matthews, pp. 799–817, Univ. of Arizona, Tucson, 1992.

D. Crisp, Jet Propulsion Laboratory, Mail Stop 169–237, 4800 Oak Grove Drive, Pasadena, CA 91109.

M. Santee, Jet Propulsion Laboratory, Mail Stop 183–701, 4800 Oak Grove Drive, Pasadena, CA 91109.

(Received March 8, 1994; revised November 20, 1994; accepted December 3, 1994.)

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5485–5500, MARCH 25, 1995

Western boundary currents in the Martian atmosphere: Numerical simulations and observational evidence

M. M. Joshi, S. R. Lewis, P. L. Read and D. C. Catling

Atmospheric, Oceanic and Planetary Physics, Department of Physics, Oxford University, Oxford, England

Abstract. Western boundary currents (hereafter WBCs) are an intensification of meridional fluid flow along the eastward facing flank of a boundary. They occur in the Earth's oceans, the best known instance being the Gulf Stream, and in the terrestrial atmosphere, an example of which is the East African Jet. We have investigated WBCs in numerical simulations of the Martian atmosphere, where they occur in the presence of large longitudinal topographical gradients, combined with the β effect. We suggest that WBCs have already been simulated by other Martian atmospheric models, although not identified as such. The intensity of these currents is dependent on basic model parameters, notably surface drag. We show that for physically reasonable values of drag, frictional forces dominate the WBCs' behavior. The requirement of zero net cross-equatorial time-mean mass flow produces a "return flow" in simulations with relatively low surface drag, which acts in opposition to the Tharsis WBC. We also show that the intense flow associated with WBCs advects potential vorticity across the equator, although the presence of strong radiative damping inhibits low-level inertial instabilities that might result from this. Slope winds are found to have a profound effect on WBC structure, especially where they have a component parallel to the jet. In these cases, slope winds can cancel out or reinforce WBCs, depending on their direction. Enhanced low-level winds associated with WBCs may be a factor that influences the locations of dust storm generation. The high vertical shear associated with WBCs may also play a role in the maintenance and subsequent decay of these dust storms. Some observational evidence may exist for WBCs in the form of the alignment of bright depositional dust streaks, wave clouds, and in wind measurements taken during the parachute descent phase of Viking Lander 1 through the planetary boundary layer.

Introduction

One of the most prominent phenomena observed in the oceans is that of western intensification. This occurs close to western boundaries of oceans where the flow reaches far higher speeds than in their interior. An example of this is the Gulf Stream, which is a narrow and intense current off the eastern coast of North America. This feature, and others like it, have been explained in terms of western boundary currents (WBCs), which are caused by a combination of the presence of the meridional boundary and the variation of the strength of the Coriolis force with latitude [*Gill*, 1980]. WBCs can also exist in the terrestrial troposphere, although in this case the "western boundary" is not truly a boundary, but a meridionally aligned mountain range which reaches heights of the order of 3-4 km. The intense flow that occurs on the eastern flank of the East African mountain chain has been interpreted as a WBC [*Anderson*, 1976]. This jet significantly affects the South-East Asia monsoon circulation by controlling the cross-equatorial flow between its high- and low-pressure systems. Indeed, recent simulations of this system suggest that without the effect of topography the monsoon circulation would not take place (M. J. Rodwell and B. J. Hoskins, A model of the Asian summer monsoon, II, Cross-equatorial flow and PV behavior, submitted to Journal of Atmospheric Science, 1994; hereinafter referred to as submitted manuscript).

Prompted by these studies, and the fact that equatorial topography on Mars is of very high amplitude ($O($ 1 scale height or 10 km)), and has large longitudinal gradients associated with it, we looked for WBCs in numerical simulations of the Martian atmosphere. As on Earth, the zonally averaged circulation of Mars is thought to be dominated at most times of the year by a cross-equatorial Hadley cell [*Lindzen and Hou*, 1988], whose lower branch, according to Martian global circulation models (GCMs), is confined in the vertical be-

Paper number 94JE02716.
0148-0227/95/94JE-02716$05.00

tween the ground and about 10-km altitude [*Haberle et al.*, 1993b]. Since this vertical scale is the same as that associated with the topography near the equator, the cross-equatorial flow associated with the Hadley cell should be strongly affected by the presence of this topography at low latitudes, especially in places where large longitudinal topographical gradients occur. One such location is on the eastern flank of the Tharsis Plateau, a large volcanic outcrop that straddles the equator and lies between 160°W and 90°W. This rises about 10 km above the mean geoid [*Wu*, 1978].

In an earlier paper [*Joshi et al.*, 1994] (henceforth paper 1), we first reported some preliminary model results indicating the likely existence of WBCs in the Martian atmosphere. It was noted that the nature of the flows was affected by surface drag, and that WBCs may influence the locations of dust storms. Some possible observational evidence was also provided. In this paper, more detailed results on the structure of WBCs given differing model parameters are discussed. Some calculations have been carried out to determine whether frictional or inertial forces control the simulated WBCs, and these results are then compared with the scale analyses performed in paper 1. The advection of potential vorticity by the jet is also examined, since studies of WBCs in the terrestrial atmosphere suggest that cross-equatorial transport of this quantity can cause inertial instabilities in subtropical latitudes (M. J. Rodwell and B. J. Hoskins, submitted manuscript, 1994). Results are then presented from an enhanced model with more realistic representations of forcing and dissipation processes, and similar analyses performed. The influence of WBCs on the generation, maintenance, and decay of dust storms is discussed, and a more quantitative description of possible observational evidence for the existence of WBCs from Viking lander entry data is presented. Finally, because of the sparsity of data available for Mars, we suggest some future observational methods to test our theoretical predictions of WBCs.

Basic Theory of WBCs and Scale Analysis

Barotropic Representations

WBCs can be simulated in a two-dimensional barotropic fluid using the shallow water equations to describe a fluid confined by two meridional boundaries and forced, e.g., by a surface wind stress which varies sinusoidally with latitude in one hemisphere, representing the transition from tropical easterlies to midlatitudinal westerlies [*Gill*, 1980]. The circulation spins up to a steady state everywhere except at the western boundary, where in the absence of frictional forces, the current becomes stronger with time but narrower, eventually reaching an infinitesimal width. In reality, other physical processes not represented above, such as friction and/or nonlinear inertial terms, become significant before this happens, causing the WBC to reach a finite magnitude in the steady state. Where nonlinear inertial terms dominate, the width of an inertial WBC will tend to $(u/\beta)^{1/2}$, where u is a typical wind speed and β is the variation of Coriolis force with latitude, while the width of a frictionally controlled WBC will depend on the friction parameterization itself. In the case of linear Rayleigh friction, represented schematically as $\partial \mathbf{u}/\partial t \approx -\mathbf{u}/\tau$, where τ is the timescale over which the friction acts, the WBC width tends to $W \simeq 1/\tau\beta$ [*Gill*, 1980].

This scaling argument can be generalized to show how the width of a WBC might vary using a more complicated frictional parameterization. We consider here a quadratic drag parameterization of the form

$$(\partial \mathbf{u}/\partial t)_{\text{drag}} = -c_d \partial(|\mathbf{u}|^2/H)/\partial\sigma, \tag{1}$$

where c_d is the nondimensional drag coefficient, H is the scale height, and σ is a terrain-following vertical coordinate p/p_*, where p_* is surface pressure, as commonly used in GCM studies for the Earth and Mars. In this case, we evaluate the derivative shown above and approximate it across the interface between the interior flow and the ground by finite differences

$$(\partial \mathbf{u}/\partial t)_{\text{drag}} = -(2c_d|\mathbf{u}|\Delta|\mathbf{u}|)/(H\Delta\sigma). \tag{2}$$

Assuming that $\mathbf{u}$ is zero at the ground, $\Delta|\mathbf{u}| \simeq |\mathbf{u}|_1 - |\mathbf{u}|_0$ where $|\mathbf{u}|_0 = 0$. Hence the frictional acceleration can be approximated by

$$(\partial \mathbf{u}/\partial t)_{\text{drag}} = -2c_d(|\mathbf{u}|^2/H)/\Delta\sigma, \tag{3}$$

where $\Delta\sigma$ in reality is the distance in σ between the ground and the lowest model level. This leads naturally to a frictional drag timescale $\tau_d \simeq H\Delta\sigma/(2|\mathbf{u}|c_d)$, suggesting that the width of a frictionally controlled boundary current will tend to

$$W \simeq 1/\tau_d\beta \approx 2c_d|\mathbf{u}|/(\beta H\Delta\sigma). \tag{4}$$

Baroclinic Representations

The barotropic representation discussed above describes the formation of an idealized WBC adjacent to a western boundary of effectively infinite height (so no flow takes place over it). In an atmosphere, however, topography is never sufficiently high to completely block the flow at all levels, and WBCs will consequently only extend over a limited range in altitude. The proper description of atmospheric WBCs therefore requires a more complex three-dimensional (baroclinic) model to capture their vertical and horizontal structure.

Despite this, the considerations outlined above should also apply in determining the horizontal width of WBCs close to the topographic features which control them in relation to the parameterization of friction and/or inertial terms. However, the vertical structure and extent of a WBC remains to be determined.

If the horizontal scale derived above may be taken to be comparable to the relevant radius of deformation (either ND/f at midlatitudes or $(ND/\beta)^{\frac{1}{2}}$ near the

equator), this would lead naturally to a vertical scale $D \simeq fW/N$ at midlatitudes or $D \simeq \beta W^2/N$ in the tropics. Thus we might expect inertially controlled WBCs to penetrate over a characteristic depth in the atmosphere, and to adopt a particular aspect ratio which depends on the ratio N/f or N/β. Substituting values for W, N, and β typical of the simulations of Martian WBCs discussed below ($W = 1000$ km, $N = 10^{-2}$ s^{-1}, $\beta = 4 \times 10^{-11}$ m^{-1}s^{-1}) leads to estimates for their vertical scale of around 2-4 km, comparable to the amplitude of the surface topography itself.

Numerical Models

The Simplified Three-Dimensional Model

The simple general circulation model (henceforth SGCM), is a fully nonlinear, multilevel primitive equation model, based on the model originally developed by *Hoskins and Simmons* [1975]. Model fields are represented in the horizontal by spherical harmonics, and a finite difference formulation is used in the vertical. The convention adopted to describe the model resolution is as follows: TnLk denotes a run having model fields truncated at total horizontal wavenumber n, with k levels in the vertical (T21 resolution is approximately equivalent to a grid-point resolution of 6°x 6°, while T10 corresponds to 12°x 12°and T42 to 3°x 3°.).

The coordinate system is (λ, ϕ, σ) where λ is longitude, ϕ is latitude, and σ is as defined previously. This vertical coordinate has the advantage that it follows the terrain, thus providing an accurate representation of orographic forcing. The topography data set used is the U.S. Geological Survey digital terrain model (DTM) [*Wu*, 1978], smoothed to the relevant resolution by a box-averaging process, and filtered in spectral space by the method of *Hoskins* [1980] in order to remove structure at resolution scale.

The only forcing and dissipation processes present are (1) Newtonian relaxation towards a prescribed zonal-mean state, (2) hyperdiffusion of temperature and momentum, and (3) Rayleigh friction, which are represented by the four terms on the right-hand side of the following thermodynamic and momentum equations:

$$\frac{DT}{Dt} = \frac{T - \overline{T}_{\text{eq}}}{\tau_N(\sigma)} + k\nabla^6 T \tag{5}$$

$$\frac{D[U,V]}{Dt} - f[V,-U] + \nabla\Phi = -\frac{[U,V]}{\tau(\sigma)} + k\nabla^6[U,V] \tag{6}$$

where τ_N varies with σ, and $\overline{T}_{\text{eq}}$ represents an approximation to a radiative equilibrium state. It is important to note that the Newtonian relaxation acts on T fields interpolated onto pressure surfaces. The interpolation therefore isolates the purely mechanical effect of topography, and does not model the thermal effect. One important consequence of this is that the SGCM does not represent slope winds [*Savijärvi and Siili*, 1993].

SGCM runs were carried out with 20 levels in the vertical domain reaching from 1 km above the ground to 80-km altitude, with a spacing of approximately 1.5 km near the ground, rising to around 6 km at the model top. Parameter τ_N was set to 3 sols (a sol is a Martian day) up to 40-km altitude, then linearly decreased with altitude above that level to reach 0.5 sols at 80 km. This also acts as a "sponge layer", damping upward propagating eddies produced by the model, and attenuating their reflection from the top of the model. Rayleigh friction was set to vary with height, from 20 sols at a height of 40 km linearly to 1 sol at the model top, and effectively acted at model levels above 50 km as a simplified parameterization of gravity wave drag. It also functioned as a parameterization of surface drag at the lowest model level. The sensitivity tests detailed below were carried out by changing τ in the lowest model level, while keeping it constant everywhere else.

The SGCM has two limitations when trying to represent the subtropical Hadley circulation. First, it has no convective adjustment scheme to simulate vertical mixing stimulated by surface heating. Because such a scheme will tend locally to enhance vertical motion via the thermodynamic equation, this leads to a Hadley cell far smaller in magnitude in the SGCM than that produced by more realistic GCMs [e.g., *Haberle et al.*, 1993b]. A consequence of this is smaller cross-equatorial velocities and hence weaker WBCs. Second, more realistic GCM simulations suggest that the Hadley cell in the southern summer is typically a factor of 2 stronger than its northern summer counterpart [*Haberle et al.*, 1993b].

The first problem can be resolved by running a more complex model that includes a parameterization of vertical energy transfer by convection (see below). The second, however, can only be addressed by a full GCM that explicitly represents the seasonal variation of insolation, which is beyond the scope of the present work.

The Intermediate Three-Dimensional Model

The intermediate global circulation model (henceforth IGCM) is structured in a very similar way to the SGCM as described above, except that some of the main physical processes are handled in a more realistic manner. Friction is parameterized by a quadratic drag law in the lowest model level only. The momentum tendency in this level due to friction is shown in equation (2). The temperature tendency in the lowest model level is given by

$$\left(\frac{\partial T}{\partial t}\right)_{\text{heating from surface}} = \frac{g}{\rho_*}\frac{\partial H_s}{\partial \sigma} \tag{7}$$

where H_s is the sensible heat, which in stable conditions is given by

$$H_s = \rho_* c_{\text{dh}} |\mathbf{u}| \Delta\theta \tag{8}$$

and in unstable conditions by

$$H_s = \rho_* c_{\text{dh}} \left(|\mathbf{u}| + A\frac{\Delta\theta}{\theta}^{1/2} \right) \Delta\theta \tag{9}$$

where ρ_* is the surface density, c_{dh} is the drag coefficient for heat transfer and is equal to $c_d/5$, and A is a stability parameter set to 500 m s^{-1}. The surface potential temperature θ is specified as varying with latitude only. These processes mix momentum and heat in the vertical far more efficiently than linear processes such as Ekman or Rayleigh friction, and so their effect on WBCs may be substantial. A convective adjustment scheme is included which adjusts the vertical model temperature profile from a convectively unstable state to a stable one (determined from observations), thus transferring heat upwards through the model. This scheme does not mix momentum in the vertical. For a fuller description of the IGCM, see *Collins* [1993].

Infrared cooling in the atmosphere is represented by a uniform cooling of 10 K sol^{-1}, which is an approximation to cooling rates produced in GCMs [*Hourdin*, 1992]. The IGCM was run with 10 levels in the vertical equispaced in σ, such that the top of the model was at $\sigma = 0.05 \simeq 30$ km.

The IGCM, therefore, while being only slightly more computationally expensive than the SGCM, is only different from a full Martian GCM in two main respects: it has a very simple representation of surface heating and radiative cooling, and does not represent the meridional source-sink flow caused by condensation and sublimation of CO_2 at the poles. This model can therefore serve as a useful intermediate step between the SGCM above and a full GCM.

It should be noted at this point that the SGCM only models the mechanical forcing effects of topography because equilibrium temperature in the lowest model level varies with elevation. In the IGCM, however, surface temperature does not vary with elevation. The IGCM will therefore simulate both the thermal as well as the mechanical forcing effects of topography. The latter case is the more realistic one for Mars [*Webster*, 1977]. A consequence of this is that the IGCM simulates upslope winds, which exist in balance with upward motion above elevated regions. These winds occur in the same places as WBCs, have the same horizontal scales, as well as similar magnitudes. The along-slope component of these winds is small at the equator as the strength of the Coriolis force there is small [*Savijärvi and Siili*, 1993]. However, as is shown later, the presence of significant meridional topographic slopes leads to significant interactions between these winds and WBCs.

Experiments and Mean States

In all of the simulations described here, both the SGCM and IGCM were typically run for 60 sols. The 15-sol time-mean fields were averaged in time from sols 45-60, representing the statistically stationary state reached by the model simulation.

Most SGCM experiments were performed at T21L20 resolution, though some cases described below were done at T10 and T42 resolution. All these cases employed temperature relaxation states representative of either of the two solstice seasons. The southern winter temperature state is shown in Figure 1, along with the resulting time-mean state. The former is an approximation to a radiative-convective equilibrium state produced by *Haberle et al.*, [1982], and the latter is the *resultant time-mean temperature state. The relaxation* state is warmer at the pole than GCM simulations of the same season [*Haberle et al.*, 1993b]. However, this offset makes no difference to the results presented here. The main differences between the two are due to the advection of temperature at high altitudes into the win-

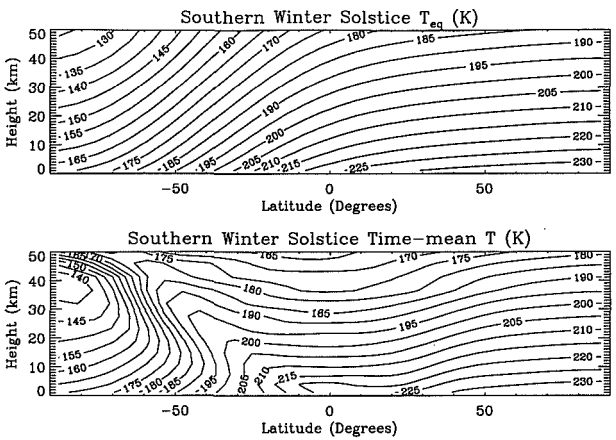


Figure 1. (Top) $\overline{T}_{eq}$ (K), an approximation to the radiative equilibrium temperature, as used for the SGCM southern winter solstice runs and (Bottom) resulting 15 sols time mean temperature $\overline{T}$.

ter midlatitudes, and are largely consistent with other models [*Haberle et al.*, 1993b].

In the southern winter case, τ at the surface was varied between 0.2 and 10 sols, while in the northern winter case it took the values 1 sol or 10 sols. A few SGCM runs were performed either at T10L20 with $\tau = 10$ sols or at T42L20 with 0.2 sols $\leq \tau \leq$ 10 sols to test the effects of varying the horizontal resolution.

The IGCM experiments described here were performed (1) for southern winter, with surface friction drag coefficient c_d having the values either 0.001 and 0.005, and (2) for one northern winter run, with c_d = 0.001. These values are consistent with values of c_d used in GCM studies of Mars for frost-free surfaces, such as 0.001 [*Haberle et al.*, 1993b] and 0.002 [*Hourdin et al.*, 1993].

The time-mean temperature state produced by the IGCM in southern winter is shown in Figure 2. This resolves the sharp latitudinal temperature gradient at 50°S better than the SGCM (see Figure 1) because in the IGCM, the steep surface meridional temperature gradient is explicitly specified.

SGCM Results

An Idealized Equatorial WBC

To illustrate the formation of WBCs more clearly prior to investigating more realistic simulations, an "idealized" run was performed in which the only topographic feature consisted of a uniformly shaped mountain. This had the shape

$$h(\lambda, \phi) = A_0 \exp\left[-(\lambda - \lambda_0)(\phi - \phi_0)/W_\lambda W_\phi\right]^4 \quad (10)$$

where the peak amplitude A_0 = 7 km, and the widths in longitude and latitude are $W_\lambda = 30°$ and $W_\phi = 30°$, respectively (representing an idealization of the Tharsis plateau). Figure 3 shows 15 sol time-mean velocity vectors at the end of a 60 sol run, at the lowest model level ($\sigma = 0.9 \simeq 1$ km above the surface), with topography contours overlaid. τ = 1 sol in all cases.

In the absence of topography, control integrations with no surface orographic features show no longitudinal variation of v. In the idealized mountain case, however, the cross-equatorial flow is confined to the eastern flank of the Gaussian mountain at a longitude of 50°W. Since this is the only significant area of cross-equatorial flow at the lowest model level, we deduce that on Mars this intense flow is the response of the Hadley cell to the presence of the large longitudinal topographical gradient at 50°W.

This phenomenon is not a mesoscale along-slope wind of the type investigated, e.g., by *Savijärvi and Siili* [1993], as it is present only on the eastern side of the idealized Tharsis. Moreover, as shown in paper 1, the variation of jet width with varying surface drag is consistent with simple theories of WBCs. We conclude therefore that the area of intense meridional flow at 50°W can be interpreted as a WBC.

The response of the Hadley cell to more realistic orography, while being broadly of the same form, is expected to have more complex structure than the above idealized case. Results are now presented for the full Martian topography, mainly at T21 resolution.

Solstitial Simulations With Full Topography

Low-level winds from two southern winter runs are shown in Figure 4, which shows horizontal velocity vectors at $\sigma = 0.9$, superimposed on contours of topography, for two different values of τ. As expected, winds are stronger in the low-drag case than in the high-drag case, by about a factor of 2. However, in both cases, the positions of the WBCs produced are similar. One WBC lies on the eastern flank of the Tharsis Plateau, in Chryse Planitia at a longitude of about 50°W, in association with the large longitudinal gradient of topography there. The other lies on the eastern slopes of Syrtis Major, around longitude 80°E, although the ratio of its intensity to that of the Tharsis WBC is smaller in the northern winter case (see below). The lower branch of the Hadley cell has therefore been funnelled into two relatively narrow channels, each about 15-30° wide. This funnelling is expected to have a profound effect on the cross-equatorial transport of heat, momentum, and constituents.

WBCs exist in the same places during northern winter simulations, as shown in Figure 5, although they are weaker at their cores than their southern winter counterparts. This is despite a variation of no more than 20% in the strengths of the Hadley cells in the four "real" topography simulations discussed so far. This might be due to the fact that in northern winter, fluid

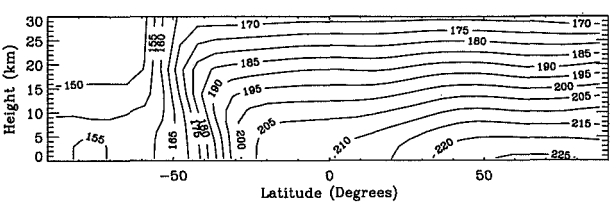


Figure 2. Southern winter 15 sols time-mean temperature $\overline{T}$ (K) in the IGCM. Note that the vertical scale is different from that shown in Figure 1 due to the different vertical model domains.

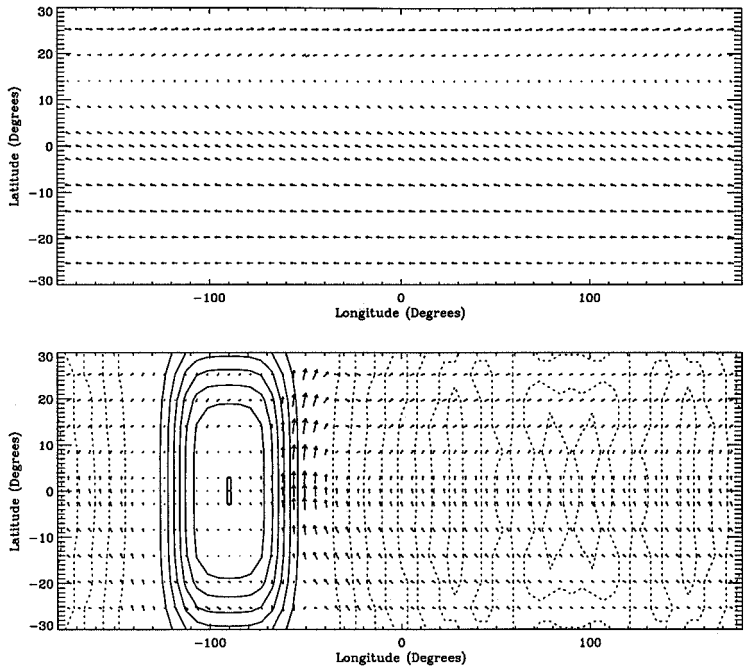


Figure 3. Time-mean velocity vectors at the lowest model level ($\sigma = 0.9$, equivalent to about 1 km above the surface) for two runs at T21 resolution with $\tau = 1$ sol. The top diagram is from a simulation with no topography and exhibits no latitudinal variation of v at the equator. The bottom diagram is from a simulation with an "idealized" Tharsis and clearly shows a WBC at 50°W. The topography contours are at 1 km intervals. The dotted contours are an artifact of the spectral transform method.

crossing the equator has to flow slightly uphill (see contours on Figure 4) and will therefore be relatively reduced in intensity. This result is repeated in the IGCM (see below).

The jet structure can be more clearly seen by examining longitude-height sections of time-mean meridional wind v at the equator. Two examples of these are shown in Figure 6, for southern winter simulations performed with $\tau = 1$ sol and 10 sols. The Tharsis WBC in the $\tau = 10$ sol case is wider than the $\tau = 1$ sol case, consistent with the barotropic theory of inertial WBCs (see text). The wider WBC also has the deeper structure, consistent with the simple vertical scaling arguments presented above. A comparison of the Syrtis WBC structure between the $\tau = 1$ sol and $\tau = 10$ sols cases also yields the same result.

Cross-Equatorial Mass Flow

The strength of the Hadley circulation, as measured by the maximum value of the time-mean mass stream function, in the presence of varying topography and surface drag, is summarized in Table 1. In the case with no topography, peak transport is diminished when drag is reduced, consistent with the arguments outlined in paper 1. The opposite happens in the presence of topography, where the increased strength of the WBCs counteracts the lessening drag, leading to greater mass transport. The strength of the Hadley cell is principally determined by the large-scale thermal structure, and so is relatively unaffected by WBCs. However, the largest mass transport of all four cases occurs with $\tau = 10$ sols, in the presence of topography.

In the $\tau = 1$ sol case, the core of the Tharsis WBC lies about 2 km below the Syrtis WBC, while in the $\tau = 10$ sols case, it lies at about the same level. This, allied with the fact that the Tharsis WBC is the more intense of the two, suggests that more of the cross-equatorial mass transport in the Hadley cell takes place in the Tharsis WBC, although the proportion of mass carried by each will again be dependent on surface drag.

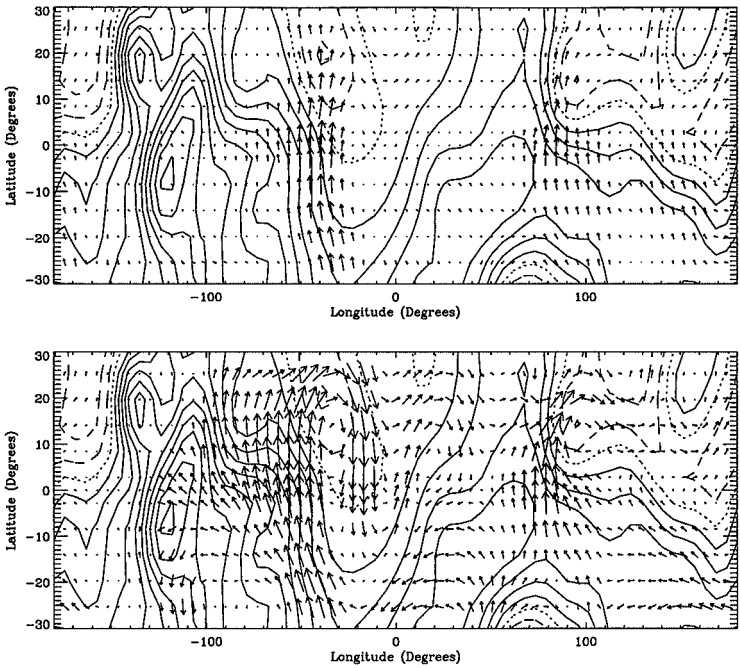


Figure 4. Time-mean velocity vectors at the lowest model level for two southern winter simulations with T21 topography. The top diagram is from a simulation with $\tau = 1$ sol and the bottom from one with $\tau = 10$ sols. The contours show orography contours at 1-km intervals.

The total mass flow below 10 km (i.e., the lower branch only) was analyzed in order to determine the fraction typically concentrated into WBCs. In both northern winter and southern winter simulations, the results were similar, and are shown in Table 2. In the τ=1 sol case, the two jets transport 65% of the total mass across about 80° of longitude. In the τ=10 sols case, when the effects of the return flow are included, about 70% of the total mass of fluid is transported across these longitudes. This shows that WBCs do indeed transport the bulk of the fluid carried in the lower branch of the Hadley cell.

One curious result, only observed in the lowest drag ($\tau = 10$ sols) simulations, is a "return flow". This is an area of high winds directed away from the WBC and back toward the equator that can be observed at 20°W in Figure 4. One explanation for this phenomenon is that there is such a large amount of mass transported across the equator by the intense WBCs in the $\tau = 10$ sols case, that it cannot be returned across the equator solely by the upper branch of the SGCM Hadley circulation. A consequence of this is the development of a low-level return flow, which can be thought of as an overshoot of the Tharsis WBC, and is depicted in Figure 7. This shows the time-averaged cross-equatorial mass transport in the southern winter $\tau = 10$ sols run, integrated over height from the lowest model level to the model top. The major areas of northward mass transport coincide with WBCs identified earlier (cf. Figure 4; lower panel), and the southward return flow is evident again at 20°W. As there is no condensation flow in the SGCM, the time-mean globally integrated value of cross-equatorial mass transport must be zero. This criterion can apparently only be satisfied by the existence of the return flow.

Potential Vorticity

Topography is important at middle latitudes in enabling net meridional flow to satisfy the requirement of geostrophy. However, in the tropics, this geostrophic constraint is of lesser importance, and the primary constraint on fluid crossing the equator arises from the necessity to conserve potential vorticity (PV). In the absence of diabatic and frictional effects, fluid which crosses the equator will inevitably find itself having PV of opposite sign to f. Such a situation satisfies a nec-

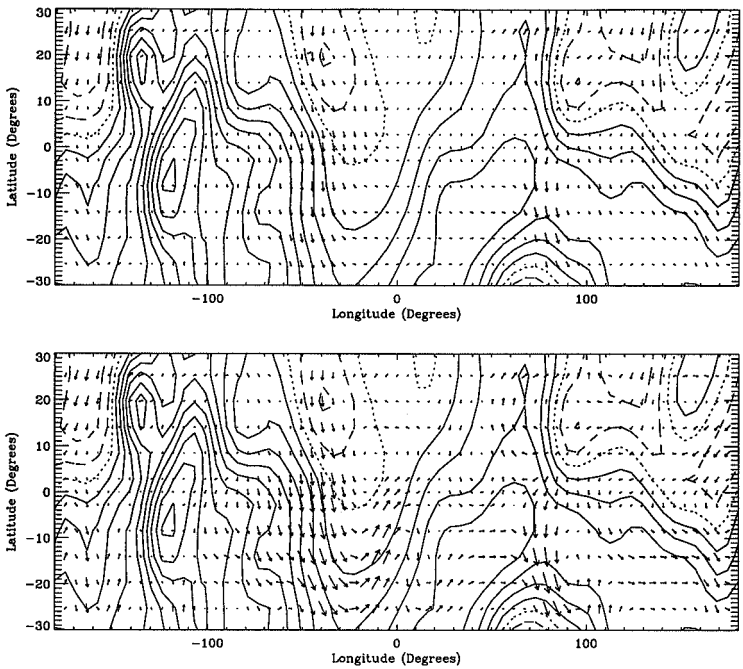


Figure 5. Time-mean velocity vectors at the lowest model level for two northern winter simulations with T21 topography. The top diagram is from a simulation with $\tau = 1$ sol, and the bottom from one with $\tau = 10$ sols. The contours show orography contours at 1-km intervals.

essary (though not sufficient) condition for inertial instability, which would result in significant modification of the flow. In practice, however, friction and/or diabatic effects are not negligible and tend to balance the effects of cross-equatorial PV advection. Such effects are therefore crucial in enabling the Hadley circulation to cross the equator.

WBCs will tend to advect PV across the equator, which could lead to the possible development of inertial instability. Denoting Ertel's PV by P, it is defined as

$$P = \zeta \cdot \nabla\theta/\rho \tag{11}$$

where ζ is absolute vorticity, θ is potential temperature, and ρ is density. A necessary (but not sufficient) condition for inertial instability is that $fP < 0$. In a region of strong cross-equatorial flow, such a condition might well be caused by the advection of positive (negative) P into the southern (northern) hemisphere. P was analyzed at an isentropic level coinciding with the jet to verify this hypothesis. The 232 K potential temperature surface was chosen for this analysis.

It was found that areas of negative fP only existed in the simulations with $\tau = 10$ sols. Since these values of drag are probably not physically realistic for Mars, it is concluded that inertial instability will not result from the cross-equatorial flow associated with WBCs. This contrasts with similar work done for the East African Jet by M. J. Rodwell and B. J. Hoskins (submitted manuscript, 1994), who suggested that inertial instability might result from high winds associated with the jet. It is suggested that the strong Newtonian relaxation in the SGCM does not allow anomalies of fP to build up, as they might on Earth.

In order to fully understand the behavior of P in the Tharsis WBC, a P budget analysis, such as that carried out by M. J. Rodwell and B. J. Hoskins (submitted manuscript, 1994) for the Somali jet and monsoon circulation on Earth, should ideally be performed. This would indicate the relative importance of terms such as advection, diabatic heating, and friction. However, work such as this is beyond the scope of this paper.

One point to note when trying to analyze isentropic P in the Martian tropics is that, in southern winter, θ surfaces tend to slope downward toward the north at the equator as $T_\phi < 0$. This slope is in the same direction as the north-south topography gradient present at this

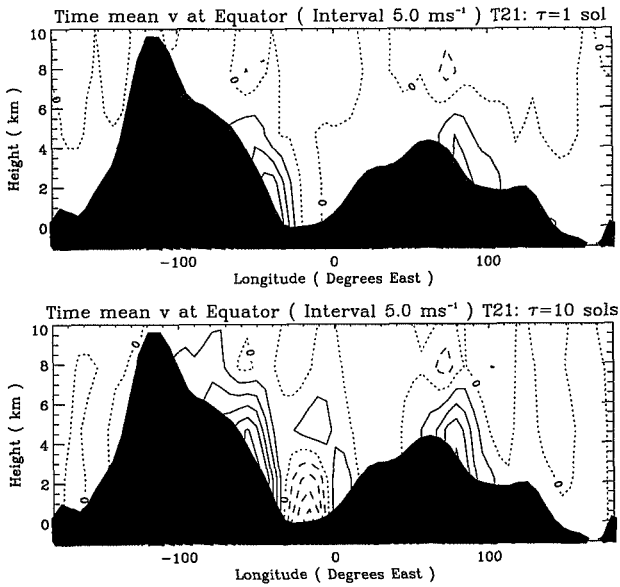


Figure 6. Two longitude-height plots showing v at the equator in simulations of southern winter with (top) $\tau = 1$ sol and (bottom) $\tau = 10$ sols. Solid lines indicate northward flow, and dashed lines vice versa. Dotted lines are zero contours. The equatorial topography has been masked out. Tharsis is shown at 100°W-50°W, and Syrtis Major is shown at 100°E.

location (see Figure 4), implying that the surface has an (approximately) constant potential temperature. This is not the case in northern winter, where θ contours intersect the ground, as can be seen by comparing the top and bottom panels of Figure 8. Therefore, in the northern winter PV1 case, there is relatively more advection of P across θ surfaces, rather than along θ surfaces, as is the case in southern winter, and indeed typically in the East African Jet as well.

Sensitivity to Model Parameters

It was shown in paper 1 that the variation of jet width was consistent with the barotropic theories outlined above. Frictional forces dominated in the Tharsis WBC for values of $\tau < 1$ sol. We now explicitly find the ratio of frictional to inertial forces by calculating the term

$$R \equiv (|\mathbf{u}|/\tau)/(\mathbf{u} \cdot \nabla \mathbf{u}) \tag{12}$$

in the two jets. For $\tau = 1$ sol, $R \approx$ 2-3 in the jet core, while for $\tau = 10$ sols, $R \approx 0.2$. These results hold for both southern and northern winter simulations. It therefore appears that WBCs are frictionally controlled for all but the most inviscid cases.

As shown in paper 1, the effect of higher horizontal resolution (T42) on the structure of the jet was minimal. This result is perhaps surprising as the minimum longitudinal width of the WBC in any of the simulations was about 15°, which might not be expected to be fully resolved at T21. However, the core of the jet had the same strength in both cases, indicating that T21 res-

Table 1. Peak Amplitude of the Meridional Mass Stream Function, Averaged Over a Period of 15 Sols in Units of 10^9 kg s^{-1} for SGCM Simulations at T21 Resolution[1]

Model Run	$\tau = 0.2$ sol	$\tau = 1$ sol	$\tau = 10$ sols
SH, topog.	1.99	2.18	2.21
SH, flat	2.09	2.05	1.90
NH, topog.		1.97	2.06
NH, flat		2.05	1.90

[1]Southern winter solstice runs (SH) have $\tau = 0.2$, 1, and 10 sols, and northern winter solstice runs (NH) have $\tau = 1$ and 10 sols.

Table 2. Mass Flow Fraction Carried by the Tharsis WBC (20°W-70°W), the Return Flow (0W-20°W) and the Syrtis WBC (60°E-100°E)

τ, sols	Tharsis	Return	Syrtis
1	0.30	–	0.25
10	0.60	-0.15	0.25

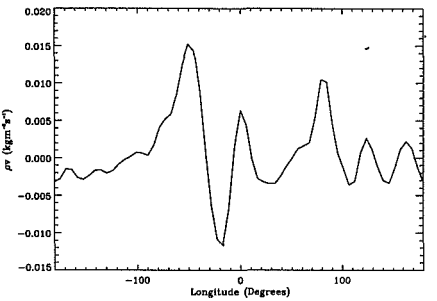


Figure 7. Time-mean vertically averaged mass transport ρv across the equator in the southern winter simulation with $\tau = 10$ sols (see Figure 4, bottom panel). $v > 0$ indicates northward flow.

olution was sufficient to fully resolve WBCs on Mars. It was also shown that T42 topography had little effect on WBC structure on all but the smallest scales (see Figure 9; lower panel). The reason for this is that the main differences between topography resolved at T21 and T42 at the Martian equator occur in the altitude range 10-15 km above the 6.1-mbar datum, which is the transition zone between the upper and lower branches of the Hadley cell, where meridional wind speeds are very low.

A southern winter run at T10L20 was carried out with τ set to 10 days, and the result compared with that of the corresponding T21L20 run. This is shown as a longitude-height contour plot of v in Figure 9. The winds are much weaker when compared to Figure 6(lower panel), because T10 resolution is not high enough to resolve the jet core. T21 resolution ($6° \times 6°$) is therefore approximately the lower limit for the adequate resolution of Martian WBCs.

IGCM Results

We now discuss results from the IGCM. As explained above, this model better represents the zonally averaged Hadley circulation, and simulates slope winds. The basic structure of WBCs and the forces that control them were examined and compared with their counterparts in the SGCM. WBC interaction with slope winds (a factor introduced in the IGCM) is also discussed.

WBC Structure

Latitude-longitude plots of WBCs in the IGCM are shown in Figure 10. As in the SGCM, the cross-equatorial flow is concentrated into two jets. However, in this case, the slope winds interact with the WBCs. The Tharsis WBC is directed toward the west, a consequence of the upslope winds in this region. The effect on the Syrtis WBC is more profound. In southern winter solstice, this jet seems to almost completely disappear.

A longitude-height plot of time-mean meridional velocity v in the IGCM, with c_d set to 0.001, is shown in Figure 11. In the northern winter case, meridional winds in the Tharsis WBC are $O(20$ m s$^{-1})$, and are stronger than in southern winter, where they only reach $O(15$ m s$^{-1})$. This is largely as expected due to the stronger northern winter Hadley cell. In both cases, v is not at its most intense in the lowest model level, as it is in the SGCM, indicating the presence of much stronger drag at the lowest model level of the IGCM than in the corresponding level of the SGCM.

There is a distinct seasonal asymmetry in Hadley circulation strength and structure. The peak meridional stream function obtained in the IGCM is 6×10^9 kg s^{-1} in a simulation of northern winter solstice, with $c_d = 0.001$, as compared to a peak of 4×10^9 kg s^{-1} in the corresponding southern winter solstice simulation. The northern winter Hadley cell is therefore $\approx$1.5 times as strong its southern winter counterpart.

Full GCMs show that the intensity of the Hadley circulation during northern winter is about 2 times greater than during southern winter. The subsolar insolation is 43% greater during northern winter than in southern winter, and so cannot fully account for this asymmetry. *Zurek et al.* [1992] suggested the influence of other factors such as orography and wave activity.

Our IGCM simulations show that even with no variation in insolation due to orbital eccentricity, i.e., a

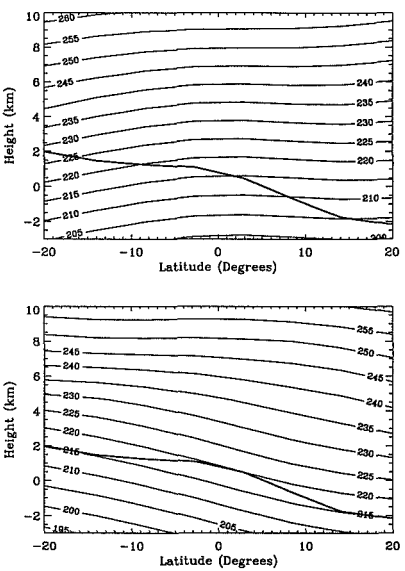


Figure 8. Potential temperature θ contours and topography at 45°W, for (top) northern winter and (bottom) southern winter.

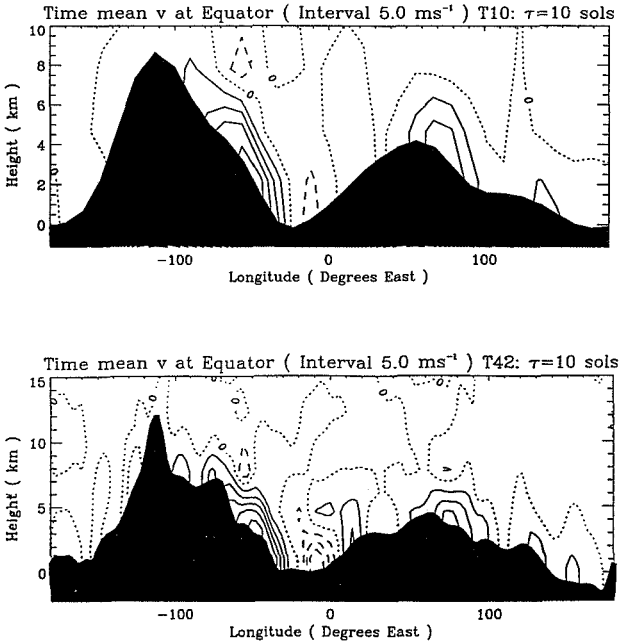


Figure 9. A longitude-height plot showing v at the equator in a (top) T10 simulation and (bottom) T42. The equatorial topography, resolved to the appropriate resolution, has been masked out in the same way as Figure 6.

zonal-mean temperature state reflected between hemispheres for each solstice, the intensity of the Hadley cell is a factor of 1.5 greater during northern winter solstice. We therefore suggest that a major influence on increased Hadley circulation intensity during northern winter is that its rising branch is in an area of elevated terrain relative to the northern hemisphere. The relatively low surface pressure associated with this leads to an enhanced Hadley circulation at this time.

The northern winter Tharsis WBC is only about 1.2 times as strong as its southern winter counterpart. This can be partly explained as before in terms of fluid flowing uphill (southwards) and therefore flowing with relatively lower intensity than otherwise might be expected. The northern winter Syrtis WBC is a factor of 4 stronger than its southern winter counterpart. This large asymmetry can be explained in terms of slope winds (see next subsection).

The ratio of frictional to inertial forces in the jets (R; see equation (12)), is greater than unity for all simulations, indicating that frictional forces dominate WBCs in this model and are therefore likely to do so on Mars.

Influence of Slope Winds

As shown earlier, the IGCM will represent upslope winds as a consequence of the lack of any surface temperature variation with elevation. On the eastern side of the Tharsis Plateau, although there is some meridional topographical gradient present, the slope is primarily directed from the west downward to the east. Upslope winds will therefore flow E to W. This causes a veering of the jet toward the west, but no significant change in cross-equatorial flow.

On the eastern side of Syrtis (in Isidis), however, the slope at the equator is actually downward toward the northeast (see Figure 4). A component of the upslope winds therefore flows across the equator from north to south. It therefore appears that in northern winter, upslope winds reinforce the Syrtis WBC, and cancel it out in southern winter. Precisely the opposite result might be expected for downslope winds, which reach their peak strength just before dawn [*Savijärvi and Siili*, 1993], although a model which represents diurnal variation is needed to test this hypothesis.

WBCs and Dust

WBCs may play an important role in the generation of global and great dust storms, since dust raising is critically dependent on surface stress, which is a function of wind speed [*Greeley et al.*, 1993], and this, in turn, is

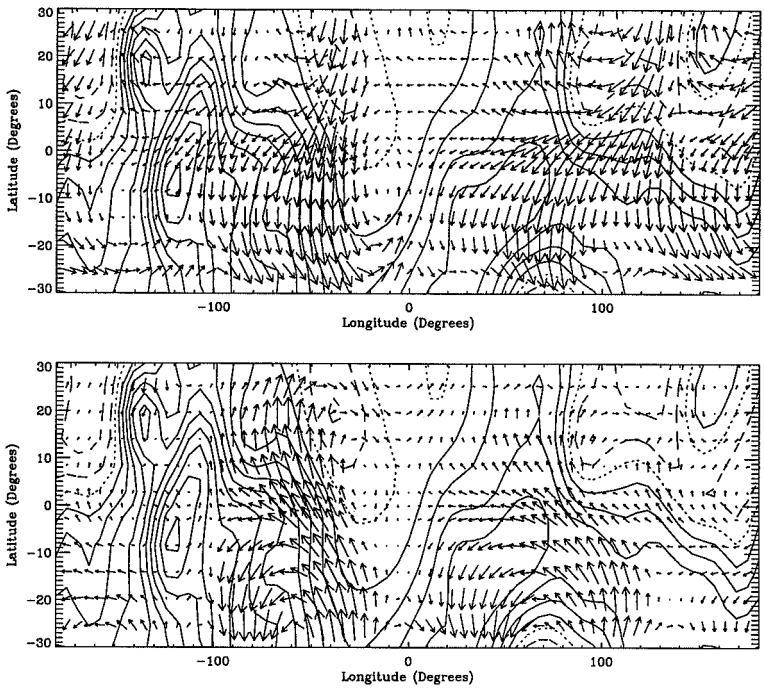


Figure 10. Time-mean velocity vectors at the lowest model level for two IGCM simulations with T21 topography. The top diagram represents northern winter solstice, and the bottom southern winter. The contours show orography contours at 1-km intervals.

the result of the superposition of winds caused by different forcing mechanisms [*Leovy et al.*, 1973]. Time-mean winds, when added to winds resulting from transient processes such as tides, can cause locally high winds at the surface during the day, when the turbulent boundary layer couples these winds to winds at the surface [*Haberle et al.*, 1993a].

Large eastward facing slopes in low latitudes are locations where global dust storms have been seen to originate, especially in Hellas (60°E, -40°S), Chryse Planitia (40°W, 0°N) and to a lesser extent Isidis (90°E, 10°N) [*Martin and Zurek*, 1993]. The latter two are coincident with the locations of WBCs in the SGCM and IGCM, and in addition, areas of high surface wind stresses produced by GCMs also correspond to locations of WBCs [see *Greeley et al.*, 1993, Figure 2e].

WBCs will have an effect on all phases of global dust storms. As shown above, they can help to cause dust raising. Second, there is a positive feedback mechanism present: the more dust that is raised, the more intense the Hadley cell will be. This suggests that, since WBCs are the response of the Hadley cell in the presence of topography, WBCs will be even more intense at this time. They will therefore provide even higher wind stresses than in relatively dust-free times and will consequently help dust storm maintenance. The third phase is that of dust storm decay, which is thought to occur when large amounts of airborne dust decouple the surface winds from winds above the planetary boundary layer (PBL) due to the increased static stability [*Leovy et al.*, 1973]. However, the criterion for boundary layer stability is that the Richardson number $Ri = N^2/|\mathbf{u}|_z^2$ be less than a critical value $Ri_c \simeq 0.2$. This decay mechanism will therefore be opposed, and possibly even nullified, by the intense "dusty state" low-level winds acting to enhance u_z, and therefore lower Ri. Model results support this conclusion (J. Murphy, personal communication, 1994). Areas of high near-surface vertical shear, such as those found in WBCs, may therefore locally arrest the decay of global dust storms.

Historically, more dust activity has been observed during southern summer not only because this is the season of maximum solar insolation, but also because the southern hemisphere is most easily seen from Earth

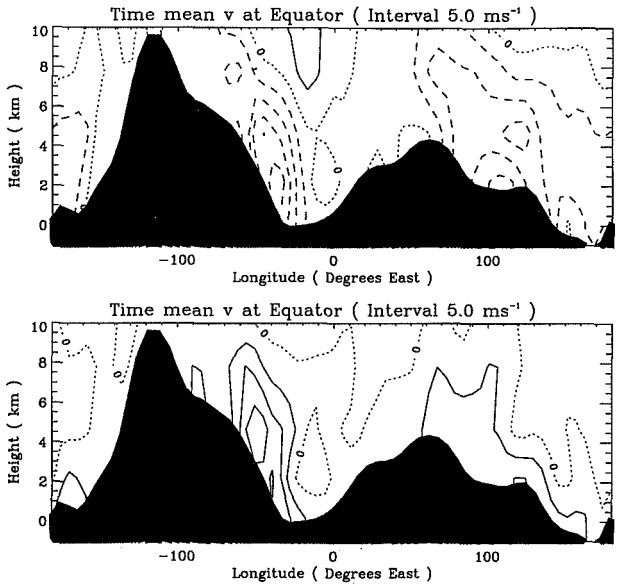


Figure 11. A longitude-height section at the equator of v in a southern winter T21L10 IGCM simulation. T21 topography has been masked out.

when the two planets are closest to each other. However, some dust activity has been observed on Mars during southern winter, especially on the eastern and southern flanks of the Tharsis plateau, e.g., in early 1984 [*Martin and Zurek*, 1993], which is consistent with the presence of intense WBCs in that season, and also with the SGCM results shown above.

Observations of WBCs

Current Observational Evidence for WBCs

There are three types of observational evidence: evidence of low-level direction, such as dust streaks or lee waves; evidence of high wind strength, such as dust storm activity as presented in the previous section; and evidence that sheds light on both of these physical aspects, such as wave clouds. We now present observations which, while not conclusively proving the existence of WBCs, are entirely consistent with their presence.

On Mars, bright depositional streaks observed in subtropical latitudes are believed to form in late southern summer and autumn [*Greeley et al.*, 1992], and are thought to be representative of global-scale winds at these times [*Greeley et al.*, 1993]. Bright streaks are associated with times of high static stability, when dust in the atmosphere is deposited on the ground in the lee of small-scale topographical features such as craters. The directions of the streaks then show the direction of the surface wind. GCM simulations show that during late southern summer, a cross-equatorial Hadley cell is present, although it is lower in intensity than during southern summer solstice [*Haberle et al.*, 1993b]. This implies that the Tharsis WBC will decay in strength as southern summer progresses. The orientation of bright streaks, when averaged over 5°x 5°cells, infer N-S wind flow and are especially coherent at 40-50°W, which is coincident with the longitude and direction of the Tharsis WBC [*Kahn et al.*, 1992].

A possible signature of WBCs can be found by comparing SGCM and IGCM results with those obtained from a one-dimensional (1-D) Martian boundary layer model [*Haberle et al.*, 1993a], as well as with winds obtained during the parachute descent of Viking Lander 1 (henceforth VL1) [*Seiff*, 1993]. When the 1-D model winds were compared with values of u and v obtained during VL1 entry, the modeled values were 6 times less than the observed ones and rotated clockwise with increasing height, which was opposite to the sense in which the VL1 winds rotated. The magnitude and rotation of the observed winds could be accounted for by significant advection of cold fluid from the southeast [*Haberle et al.*, 1993a]. One mechanism that would

explain this anomalous advection is the Tharsis WBC, which could lead to large winds at the VL1 site. This can be seen in Figure 4, at 48°W, 22°N (the coordinates of the VL1 site). There is significant meridional advection caused by the WBC which does not occur anywhere else at this latitude. Moreover, the decrease in meridional wind strength with altitude in the lowest 4 km of the SGCM and IGCM (see Figures 6 and 11) would cause an anticlockwise turning of the wind with height.

VL1 winds can be compared with time-mean values obtained from the SGCM during southern winter with τ set to 10 sols and 1 sol, and with the IGCM, with c_d set to 0.001. For reference, a "control" simulation having no topography, with low drag ($\tau = 10$ sols), is also diagnosed, to show the effect of the Hadley cell on a "flat" Mars. While it is not expected that the SGCM can properly simulate winds in the Martian boundary layer, the effect of the WBCs on winds at the VL1 site can be gauged qualitatively. Figure 12 shows a comparison of the 1-D PBL model, inferred VL1 winds at 1.45 km above the surface at the landing site, and model winds at the 6.0-mbar level, which corresponds to 1.45 km above the VL1 site.

The control run (arrow c) does not correctly simulate the direction or magnitude of the observed winds (arrow a). The runs with topography do show significant meridional winds, consistent with (arrow a), although their zonal components are wrong. This might be explained in the low drag case (arrow e) by the fact that the return flow causes the westerly turning of the WBC at this location, which might not happen in reality. Case d has wind vectors in the right direction, but a factor of 2 too small. The IGCM case (case f) exhibits significant meridional winds, albeit weaker than those observed. To conclude, the presence of topography is necessary to simulate observed meridional winds at VL1, which is consistent with the large values of v being produced by the Tharsis WBC.

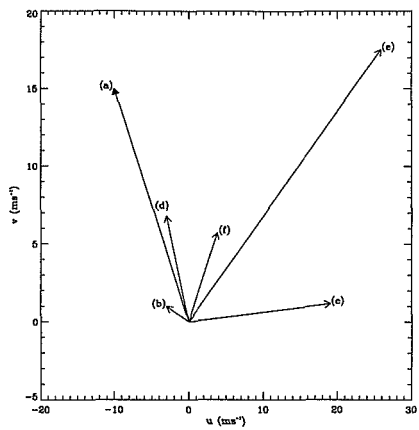


Figure 12. Wind vectors from the SGCM, the 1-D model, and entry data from VL1. (arrow a) VL1 **u** at 1.45 km above the landing site, (arrow b) **u** at 1.4 km height from the 1-D model, (arrow c) **u** in the "flat" SGCM run, $\tau = 10$ sols, (arrow d) **u** in SGCM with topography, $\tau = 1$ sol, (arrow e) **u** in SGCM with topography, $\tau = 10$ sols, (arrow f) **u** in IGCM, $c_d = 0.001$.

Future Observational Tests

WBCs may be indirectly observed in the future by measuring the near-surface temperature signal associated with advection of cold fluid across the equator by the WBCs. The SGCM produces a strong thermal signal, as shown in paper 1, although its magnitude may be significantly different from that actually observed in reality because the SGCM has such a simplified parameterization of surface drag. A problem with this method is that not only is the vertical scale of the signal small (approximately 3 km), but it is also very near the surface, and so would be difficult to measure by conventional remote sounding instruments. However, as mentioned in paper 1, it might be possible for a network of surface weather stations to measure this thermal signal, as well as other WBC signatures such as wind, if some stations were deployed along WBC longitudes.

Wave clouds are produced by shear instabilities near the surface. They take the form of periodic cloud cylinders, with the axes of the cylinders being aligned perpendicular to the mean wind. These clouds have been observed around the Tharsis Plateau [*Kahn*, 1984], although their directions are almost randomly distributed. Given more observations of these phenomena, more supporting evidence for WBCs might be accumulated.

Visual evidence in the form of dust storms and condensate clouds will shed light on low-level wind speed and direction, and so might help to provide evidence for the presence of WBCs.

Balloons have the unique capability to determine wind directly in both the boundary layer and the free atmosphere. They can also provide coverage of other meteorological variables such as temperature and pressure. Tethered balloons, which rise and sink daily, would be particularly useful for determining vertical profiles through the boundary layer at WBC locations.

Conclusions

This paper has shown that WBCs are the manifestation of the lower branch of the Martian cross-equatorial Hadley cell in the presence of the large equatorial topography on Mars, and the latitudinal gradient of planetary vorticity. They have been simulated previously in numerical models, but have not been investigated in detail until now. The most robust WBC produced lies on the eastern flank of the Tharsis Plateau, due to the large longitudinal topographical gradients there. Most of the mass transported across the equator by the

lower branch of the Hadley cell is within this WBC. The peak strength of the Hadley cell is increased by approximately 10% by the presence of intense WBCs in the low-drag case, despite the fact that WBCs only have a relatively localized effect on the atmosphere's thermal structure.

The character of the WBCs produced was found to be dependent on the strength of the surface drag applied, ranging from being weak flows with a large longitudinal extent ($\simeq$40°) in the high drag cases to narrow and intense flows ($\simeq$10°) in simulations with low drag. Moreover, their width changed in a manner roughly consistent with simple barotropic theory, reaching a minimum at about $\tau = 1$ sol, and then increasing as $v^{1/2}$. SGCM simulations with low drag resulted in a "return flow" consisting of fluid flow across the equator in the opposite sense to the WBC. It is thought that this phenomenon is due to the constraint of the model having to satisfy the continuity equation requirement of no net time-mean cross-equatorial flow. For values of surface drag thought to be "reasonable" for Mars, WBCs were shown to be controlled by frictional forces. This implies that the $\tau = 10$ sols simulation probably has unrealistically low surface drag, so that although the return flow is an interesting phenomenon, it is unlikely to occur in reality.

There is significant advection of Ertel's PV across the equator, which when combined with the presence of high winds in the low-drag simulation, suggests that low-level inertial instabilities should not be present in the equatorial regions. This is probably a consequence of the strong diabatic heating present on Mars. If these sorts of instabilities do occur on Mars, they might be important in the context of dust storm generation, as they would lead to higher than average near-surface transient winds. It is suggested that future work might consider explicitly calculating the PV sources and sinks present, as done for the Earth by M. J. Rodwell and B. J. Hoskins (submitted manuscript,1994).

The differences between model resolution at T21 and T42 were found to be small. This was a surprising result, since the minimum width of WBCs produced in the models was 15°, equivalent to about 2.5 grid spaces at T21, or 5 at T42. As shown in paper 1, T21 resolution was sufficient to resolve WBC cores. However, T10 simulations carried out in this paper show that T10 is not sufficient to resolve the core of the Tharsis WBC, suggesting that T21 resolution is about the minimum model resolution needed to simulate WBCs.

The IGCM, with its more realistic treatment of surface fluxes of heat and momentum, produces similar WBCs to the SGCM. However, in this case it was found that surface drag dominated in the lowest model level. This result is a consequence of having a representation of surface drag that mixes momentum in the vertical very efficiently.

Upslope winds are found to drastically affect the character of the Syrtis WBC, by reinforcing it during northern winter and cancelling it during southern winter. Downslope winds might be expected to have the opposite effect. The Tharsis WBC is more robust, not having its intensity significantly affected by slope winds because they flow east to west in this case.

The PBL varies quite dramatically over the course of one sol. Modeling studies show that the PBL can reach heights of 5 km during late afternoon [*Haberle et al.*, 1993a; *Savijärvi and Siili*, 1993] and observationally the PBL depth for VL1 entry during late afternoon has been estimated as 6.5 km above the ground [*Seiff*, 1977]. During the night, however, it "collapses", decoupling the winds in the atmosphere from the surface. This implies that momentum transfer through the PBL at this time will be relatively inefficient. Hence, the structure of WBCs might be expected to change over the course of a day. However, the variability is not easy to predict. First, the high winds associated with WBCs might act to destabilize the PBL (see below). Second, as WBCs are set up by propagating Rossby waves over timescales $>1/\omega$, it is not clear whether they can react to dynamical changes having timescales less than a sol. Again, the only way to test this hypothesis is to analyze data from GCMs.

The top of the PBL is defined as the height at which the Richardson number, Ri, first exceeds the critical Richardson number Ri_c, as vigorous heat and momentum exchange in the vertical due to sub grid-scale processes such as turbulent mixing are only assumed to take place when $Ri < Ri_c$. The above results suggest that during the day, WBCs are expected to lie within the PBL. Since both of the models used in this study cannot represent the diurnally varying PBL, full GCMs which parameterize turbulent exchange between different model levels must be used to model the behavior of WBCs at this time.

WBCs may be an important factor in the raising of dust and hence in the generation of dust storms. The eastern side of Tharsis has historically been known as an area of frequent dust activity, at all times of the Martian year. This is consistent with WBCs acting to give a time-mean nonzero near-surface $|\mathbf{u}|$ field, which, when added to $|\mathbf{u}|$ produced from other sources, e.g., tides, might raise the total $|\mathbf{u}|$ above that needed for significant dust raising. Fluctuations in WBC strength, possibly as a result of inertial oscillations near the equator, might raise the probability of dust raising even higher. Near-surface winds may act to stop the decay of global dust storms because one of the effects of dust storms is a more intense Hadley cell, which would lead to stronger WBCs and surface summer westerlies, thereby counteracting the stabilizing effect of dust on the PBL.

Some evidence for WBCs may already exist, in the form of meridionally aligned depositional dust streaks as well as in VL1 entry data, and a means of measuring temperature fluxes caused by WBCs, albeit difficult, has been suggested.

It is finally concluded that although we have demonstrated that it is extremely likely that WBCs of some sort exist in the Martian atmosphere, our knowledge of their general character, intensity, and effect on the circulation remains incomplete, and will do so until future measurements of the near-surface Martian circulation are made.

Acknowledgments. The authors would like to thank M. Collins, R. Haberle, J. Murphy, the anonymous reviewers of paper 1 as well as the reviewers of this paper for their helpful comments. This work formed part of a D.Phil. project in the Department of Physics, Oxford University. M. M. Joshi was supported by a studentship from the Gassiot Grants Committee of the UK Meteorological Office. S. R. Lewis and D. C. Catling were supported by the UK Science and Engineering Research Council.

References

Anderson, D. L. T., The low-level jet as a western boundary current, *Mon. Weather Rev.*, *104*, 907–921, 1976.

Collins, M., The Meteorology of Mars, Ph.D. thesis, University of Reading, Reading, England, 1993.

Gill, A.E., *Atmosphere-Ocean Dynamics*, pp. 511–512, Academic, San Diego, Calif., 1980.

Greeley, R., N. Lancaster, S. Lee, and P. Thomas, Martian aeolian processes, sediments and features, in *Mars*, edited by H. H. Kieffer *et al.*, chap. 22, University of Arizona Press, Tucson, 1992.

Greeley, R., A. Skypeck, and J. B. Pollack, Martian aeolian features and deposits: Comparisons with general circulation models, *J. Geophys. Res.*, *98*, 3183–3196, 1993.

Haberle, R. M., C. B. Leovy, and J. B. Pollack, Some effects of global dust storms on the atmospheric circulation of Mars, *Icarus*, *50*, 322–367, 1982.

Haberle, R. M., H. C. Houben, R. Hertenstein, and T. Herdtle, A boundary-layer model for Mars: Comparison with Viking lander and entry data, *J. Atmos. Sci.*, *50*, 1544–1559, 1993a.

Haberle, R. M., J. B. Pollack, J. R. Barnes, R. W. Zurek, C. B. Leovy, J. R. Murphy, H. Lee, and J. Schaeffer, Mars atmospheric dynamics as simulated by the NASA/Ames general circulation model 1, The zonal mean circulation, *J. Geophys. Res.*, *98*, 3093–3124, 1993b.

Hoskins, B. J., Representation of the Earth's topography using spherical harmonics, *Mon. Weather Rev.*, **108**, 111–115, 1980.

Hoskins, B. J., and A. J. Simmons, A multi-layer model and the semi-implicit method, *Q. J. R. Meteorol. Soc.*, *101*, 637–655, 1975.

Hourdin, F., A New Representation of the Absorption by the CO_2 15 μm Band for a Martian General Circulation Model, *J. Geophys. Res.*, *97*, 18319–18335, 1992.

Hourdin, F., P. Le Van, F. Forget, and O. Talagrand, Meteorological variability and the annual surface pressure cycle on Mars, *J. Atmos. Sci.*, *50*, 3625–3640, 1993.

Joshi, M. M., S. R. Lewis, P. L. Read, and D. C. Catling, Western boundary currents in the atmosphere of Mars, *Nature*, *367*, 548–551, 1994.

Kahn, R. A., The spatial and seasonal distribution of Martian clouds and some meteorological implications, *J. Geophys. Res.*, *89*, 6671–6688, 1984.

Kahn, R. A., T. Z. Martin, R. W. Zurek, and S. W. Lee, The Martian dust cycle, in *Mars*, edited by H. H. Kieffer *et al.*, pp. 1042–1043, University of Arizona Press, Tucson, 1992.

Leovy, C. B., R. W. Zurek, and J. B. Pollack, Mechanisms for Martian dust storms, *J. Atmos. Sci.*, *30*, 749–762, 1973.

Lindzen, R. S., and A. Y. Hou, Hadley circulations for zonally-averaged heating centered off the equator, *J. Atmos. Sci.*, *45*, 2416–2427, 1988.

Martin, L. J., and R. W. Zurek, An analysis of the history of dust activity on Mars, *J. Geophys. Res.*, *98*, 3221–3246, 1993.

Savijärvi, H., and T. Siili, The Martian slope winds and the nocturnal PBL jet, *J. Atmos. Sci.*, *50*, 77–88, 1993.

Seiff, A., Thermal structure of Mars atmosphere from Viking entry measurements, in *Symposium on Planetary Atmospheres*, edited by A. Vallance Jones, pp. 135–140, Royal Society of Canada, Ottawa, Ontario, 1977.

Seiff, A., Mars atmospheric winds indicated by motion of the Viking landers during parachute descent, *J. Geophys. Res.*, *98*, 7461–7474, 1993.

Webster, P. J., The low-latitude circulation of Mars, *Icarus*, *30*, 626–649, 1977.

Wu, S. S. C., Mars synthetic topographic mapping, *Icarus*, *33*, 417–440, 1978.

Zurek, R. W., J. R. Barnes, R. M. Haberle, J. B. Pollack, J. E. Tillman, and C. B. Leovy, in *Mars*, edited by H. H. Kieffer *et al.*, p. 887, University of Arizona Press, Tucson, 1992.

D. C. Catling, M. M. Joshi, S. R. Lewis, and P. L. Read, Atmospheric, Oceanic and Planetary Physics, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, England. (e-mail: catling@oxpln3.atm.ox.ac.uk; joshi@atm.ox.ac.uk; lewis@atm.ox.ac.uk; read@gordon.atm.ox.ac.uk)

(Received March 24, 1994; revised September 27, 1994; accepted October 14, 1994.)

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5501–5523, MARCH 25, 1995

The sensitivity of the Martian surface pressure and atmospheric mass budget to various parameters: A comparison between numerical simulations and Viking observations

Frédéric Hourdin, François Forget, and Olivier Talagrand

Laboratoire de Météorologie Dynamique du Centre National de la Recherche Scientifique, Ecole Normale Supérieure, Paris, France

Abstract. The sensitivity of the Martian atmospheric circulation to a number of poorly known or strongly varying parameters (surface roughness length, atmospheric optical depth, CO_2 ice albedo, and thermal emissivity) is investigated through experiments performed with the Martian version of the atmospheric general circulation model of Laboratoire de Météorologie Dynamique, with a rather coarse horizontal resolution (a grid with 32 points in longitude and 24 points in latitude). The results are evaluated primarily on the basis of comparisons with the surface pressure records of the Viking mission. To that end, the records are decomposed into long-period seasonal variations due to mass exchange with the polar caps and latitudinal redistribution of mass, and short-period variations due to transient longitudinally propagating waves. The sensitivity experiments include a 5-year control simulation and shorter simulations (a little longer than 1 year) performed with "perturbed" parameter values. The main conclusions are that (1) a change of horizontal resolution (twice as many points in each direction) mostly affects the transient waves, (2) surface roughness lengths have a significant impact on the near-surface wind and, as a matter of consequence, on the latitudinal redistribution of mass, (3) atmospheric dust optical depth has a significant impact on radiative balance and dynamics, and (4) CO_2 ice albedo and thermal emissivity strongly influence mass exchange between the atmosphere and the polar caps. In view of this last conclusion, an automatic procedure is implemented through which the albedo and emissivity of each of the two polar caps are determined, together with the total (i.e., including the caps) atmospheric CO_2 content, in such a way as to get the closest fit of the model to the Viking pressure measurements.

1. Introduction

The meteorology of the planet Mars is probably, after that of Earth, the one which has been most studied and is best understood. This is due, first, to the large number of observations acquired in the seventies on the occasion of several spacecraft missions (the Soviet Mars missions and the American Mariner and Viking missions). In particular, the two Viking landers have recorded near-surface temperature, pressure, and winds over several Martian years. Such records are unique for extraterrestrial planets. The Viking measurements, especially the pressure measurements, contain very instructive information on the global atmospheric circulation. The large-amplitude seasonal oscillations of the pressure are due to the variations of the atmospheric mass (which result from condensation-sublimation of a substantial fraction of the atmospheric carbon dioxide in the polar caps), but also to internal latitudinal mass redistribution. The more rapid oscillations of the surface pressure, with periods of 2 to 5 sols (sols are Martian solar days), are signatures of the transient planetary waves which are present, at least in the northern hemisphere, during autumn and winter. Our relatively good knowledge of the Martian meteorology is also due to the results obtained with atmospheric general circulation models (GCM's). The usefulness of GCM's in this particular instance is certainly due in great part to the similarity of the atmospheric circulations of Earth and Mars. As early as the late sixties, the then recently developed terrestrial GCM of the University of California, Los Angeles, was successfully adapted to Mars [*Leovy and Mintz*, 1969]. The same model, on which various improvements were continuously performed, was later used both for the preparation of the Viking mission and for many studies of the Martian meteorology performed at NASA/Ames Research Cen-

Paper number 94JE03079.
0148-0227/95/94JE-03079$05.00

ter by J. B. Pollack and collaborators [*Pollack et al.*, 1981, 1990, 1993; *Haberle et al.*, 1993; *Barnes et al.*, 1993].

More recently, a new version of the terrestrial GCM of Laboratoire de Météorologie Dynamique (LMD) has been adapted to Mars [*Hourdin*, 1992; *Hourdin et al.*, 1993] . The LMD GCM was the first self-consistent model to simulate the Martian atmospheric circulation over more than a year. Despite the somewhat arbitrary use of an a priori set of model parameters, some of which are poorly known from observation, the model was able to reproduce rather accurately many observed features of the Martian atmosphere, such as the thermal structure derived from Mariner 9 temperature retrievals, and the long- and short-period oscillations of the surface pressure observed by the Viking landers. Moreover, those simulations have shown that the large seasonal pressure variations were not only due, as usually thought until then, to mass exchange between the atmosphere and the polar caps, but were also strongly influenced by internal mass redistribution associated to the atmospheric circulation (*Talagrand et al.* [1991], *Hourdin et al.* [1993], referred to as paper 1 hereinafter, and *Pollack et al.* [1993]).

The present paper, which is a continuation of paper 1, is mainly devoted to sensitivity studies of the LMD model to some very uncertain or strongly varying parameters (surface roughness length, atmospheric optical depth, CO_2 ice albedo, and thermal emissivity) with emphasis on their impact on the atmospheric mass budget and on the meteorological contribution to the annual pressure cycle. Similar studies have already been performed by several authors. *Paige and Ingersoll* [1985], *Wood and Paige* [1992], and *Paige and Wood* [1992] have concentrated on the mass budget with a model of the energetics of the polar caps. *Pollack et al.* [1993] performed a more complete study, since they included in their energy budget model the effect of dynamics (heat transport and meteorological contribution to the pressure variations) using the results of off-line simulations performed with the NASA/Ames GCM on a series of selected short periods. The present work is new in that the sensitivity experiments described here have been performed with a self-consistent GCM allowing for complex feedback between dynamics and diabatic processes.

We try here to answer the following questions: (1) What is the uncertainty in the model resulting from the uncertainties in model parameters or parameterizations? (2) Can the model parameters be constrained from the available observations (and especially from the Viking surface pressure measurements)? (3) What information can be derived from the GCM on the polar processes which dominate the mass cycle?

To that end, we propose simple diagnostics for the Viking pressure measurements, well adapted to the validation of GCMs and to the study of the annual pressure cycle. Those diagnostics are the subject of section 2. We then give in section 3 a brief description of the atmospheric model (already described in paper 1) and present in section 4 a 5-year simulation used afterward as control for the sensitivity experiments. In section 5 we test the uncertainties related to the most uncertain model parameters and, in section 6, we study the impact of the variations of the dust amount in the atmosphere. We take advantage of the high sensitivity of the atmospheric mass budget to the CO_2 ice radiative properties (visible albedo and thermal emissivity) to derive an automatic best fit procedure to the Viking pressure. This procedure and the best fit results are presented in section 7.

2. Analysis of the Viking Surface Pressure Measurements

The records of the surface pressure, acquired over almost 2 Martian years by the Viking 2 lander, and over more than 3 years by the Viking 1 lander [*Tillman and Guest*, 1987] , have been widely used for studies of the Martian transient planetary waves [*Tillman*, 1988; *Barnes*, 1980, 1981], for comparisons with the results of seasonal energy balance models [*Pollack et al.*, 1993; *Paige and Ingersoll*, 1985; *Wood and Paige*, 1992; *Paige and Wood*, 1992] , and, in paper 1, for direct comparisons with GCM simulations. We present here simple diagnostics, well suited to GCM validations, based on a decomposition of the observed variations between a seasonal component (which is itself decomposed into a total atmospheric mass contribution and a meteorological contribution) and a shorter term transient component. In the present paper, which does not discuss diurnal cycle and the associated thermal tides (known to be very strong on Mars), we use directly the diurnally averaged pressure data.

Seasonal Variations

The large-amplitude seasonal variations of the surface pressure can easily be isolated, as proposed by *Tillman et al.* [1993], by retaining the first harmonics of the annual cycle. Following the same authors, we retain only the observations performed after the 1977-B dust storm (clear-sky conditions), but we use solar longitude L_S instead of real time t as the temporal variable of our Fourier analysis. L_S is a more appropriate coordinate for the seasonal cycle, since both the inclination of the planet, $\sin\delta = \sin L_S \sin\delta_0$ (where $\delta_0 \sim 24^o$ is the Martian obliquity), and the distance to the sun, $r = p/(1 + e\cos L_S) \sim p(1 - e\cos L_S)$ (where $p = 1.51$ UA and $e = 0.093$), are almost first harmonics of a year in terms of solar longitude (and not in terms of real time).

The values of the amplitude and phase lag of the first eight harmonics are given in Table 1 for the two Viking landers. The quality of the fit (measured from the root mean square of the difference between the data and the synthetic spectra) reaches a minimum value for $N \geq 5$, the residual value corresponding to the mean amplitude of the transient oscillations (about 0.04 mbar for Viking 1 and 0.1 mbar for Viking 1). Somewhat arbitrarily, the eight-harmonic fit (compared to Viking data

Table 1. Amplitude ($p^{(k)}$) and Phase Lag ($L_S^{(k)}$, Solar Longitude of the Maximum Pressure) of the First Eight Harmonics of the Seasonal Pressure Variations Observed by the Viking Landers (Clear-Sky Conditions)

Site	Mode k	$p^{(k)}$, mbar	$L_S^{(k)}$, deg	Mean Quadratic Error ϵ, mbar
VL1	0	7.9623		
	1	0.6887	310.97	
	2	0.5453	66.76	0.0501
	3	0.0154	20.87	0.0512
	4	0.0535	68.53	0.0389
	5	0.0105	33.36	0.0383
	6	0.0053	12.02	0.0378
	7	0.0058	19.85	0.0376
	8	0.0025	13.11	0.0376
VL2	0	8.6889		
	1	0.8320	308.07	
	2	0.6061	66.62	0.1139
	3	0.0422	25.48	0.1120
	4	0.0231	71.05	0.1104
	5	0.0355	65.98	0.1086
	7	0.0119	46.37	0.1081
	8	0.0122	28.59	0.1081

ϵ is the mean quadratic difference between the Viking individual data and the synthetic values reconstructed from the first N harmonics:

$$p^{[N]}(n) = \sum_{k=0}^{k=N} p^{(k)} \cos\left[k \times \left(L_S - L_S^{(k)}\right)\right]$$

in Figure 1) is retained as a reference seasonal pressure cycle in the present study and used for direct comparison with simulated pressures. However, for both landing sites, the first two harmonics explain the main part of the seasonal variations. Note that the two-harmonic fit is much better when solar longitude is used instead of real time as a spectral coordinate. For the time coordinate, for instance, the amplitude of the third harmonic at Viking 1 is of the order of 0.1 mbar, and the root mean square of the difference between the two-harmonic fit and the data is of the order of 0.075 mbar.

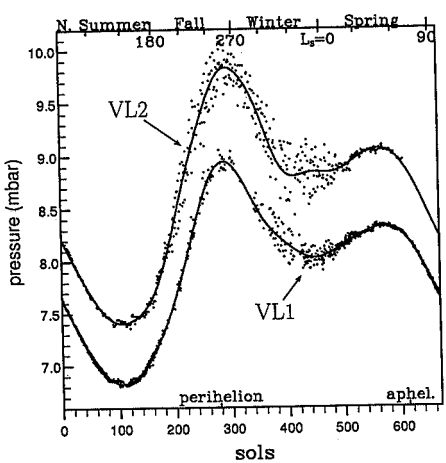


Figure 1. Diurnally averaged Viking pressure observations for clear-sky conditions (points) are compared to the eight-harmonic L_S fit (solid curve).

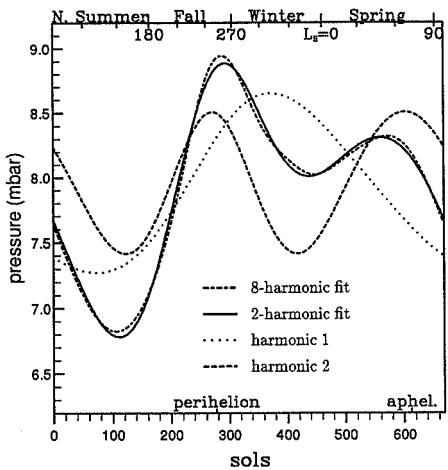


Figure 2. Two- and eight-harmonic fits of the time evolution of the surface pressure observations at Viking 1. The reference eight-harmonic fit is already shown in Figure 1.

As illustrated in Figure 2 and Table 1, harmonic 2 really corresponds to the seasonal cycle with two minima during the two winters, lagging the solstices by about 66°. The minimum of insolation on one hemisphere at winter solstice corresponds to the maximum condensation rate, the maximum mass of the corresponding cap occurring much later. Harmonic 1 reflects the asymmetry between both winters: mainly forced by orbital eccentricity, it can also be affected by hemispheric asymmetries in orography, in the dust content, and in cloud coverage over caps or amplified by albedo feedbacks on the polar caps. The maximum pressure of the first harmonic also lags the maximum planetary insolation, at perihelion, by about 60°.

Various Contributions to the Local Seasonal Pressure Variations

As discussed above, the seasonal pressure variations are due in part to the variations of the total atmospheric mass $M_{\rm atm}$ resulting from the mass exchange between the polar caps and the atmosphere. It is convenient to introduce the planetary mean of the surface pressure $p_{\rm atm} = gM_{\rm atm}/(4\pi a^2)$, where a is the planetary radius and g the gravity, and to define in the same way an equivalent surface pressure for the northern and southern polar caps, p_N and p_S respectively, and a total

pressure $p_{tot} = p_{atm} + p_N + p_S$, which is independent of time.

However, as first pointed out by *Talagrand et al.* [1991], and studied in greater details in paper 1 and by *Pollack et al.* [1993], the seasonal evolution of the surface pressure at a particular point of the planet can differ significantly, under the effect of latitudinal mass redistribution, from the evolution of the mean surface pressure p_{atm}. It is convenient to introduce here the ratio α of the local pressure to the mean atmospheric surface pressure. This ratio will be called the "meteorological factor" hereafter. As shown in paper 1, the horizontal variations of α can be accurately decomposed into the sum of an orographic and a dynamical component. The orographic component (accounting for the horizontal pressure variations along slopes) varies in time because of the strong seasonal variations of temperature. The dynamical component is due to the geostrophic balance between the surface pressure and the near-surface horizontal winds. It is particularly responsible for strong latitudinal pressure variations in balance with the strong Martian zonal winds. Those zonal winds are produced themselves by latitudinal angular momentum redistribution by the condensation flow and thermally driven atmospheric circulation.

The absolute meteorological contribution cannot be deduced from Viking observations. However, the relative meteorological factor between the two Viking sites, defined as

$$\alpha_{VL2/1} = \frac{p_{VL2} - p_{VL1}}{p_{VL2} + p_{VL1}} = \frac{\alpha_{VL2} - \alpha_{VL1}}{\alpha_{VL2} + \alpha_{VL1}} \qquad (1)$$

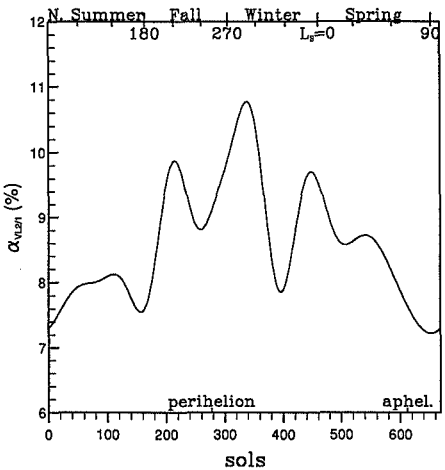


Figure 3. Meteorological contribution to the relative difference between the Viking 2 and Viking 1 pressures $\alpha_{VL2/1} = (p_{VL2} - p_{VL1})/(p_{VL2} + p_{VL1})$ computed from the eight-harmonic fits.

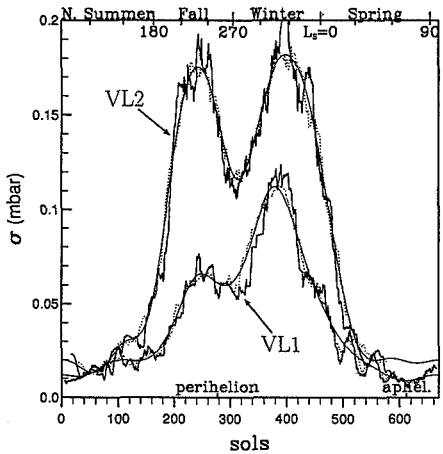


Figure 4. Seasonal evolution of the amplitude of the transient pressure oscillations at both Viking sites computed as the running quadratic mean over 25 (thick solid curve) and 50 (dotted curve) consecutive sols of the difference between the instantaneous (diurnally averaged) Viking pressure and the eight-harmonic fit. The thin solid curve is the eight-harmonic L_S fit to the dotted curve.

can be used for model validation. Figure 3 shows the relative meteorological factor as computed from the eight-harmonic fits to the Viking 1 and Viking 2 data. The maximum of the curve near northern winter solstice can be explained by the very cold temperatures which strongly enhance the pressure in the lowest Viking 2 site. This effect is partly compensated by the presence of the strong winter jet which should correspond by geostrophic balance to a northward decreasing pressure. As explained in paper 1, the disappearance of the near-surface westerlies during the great 1977-B dust storm ($L_S \sim 280$) was responsible for a sudden increase of the surface pressure recorded by the Viking 2 lander.

It is noteworthy that models must be used in addition to observations of the surface pressure in order to deduce the evolution of the atmospheric mass which requires the knowledge of the absolute meteorological factor. The relative factor $\alpha_{VL2/1}$ can only help for the model validation.

Transient Eddies

Another important feature of the observed pressure variations is the presence of short-period fluctuations linked to transient planetary waves. Without going into a complete study of these transient waves, which would require the use of more sophisticated analysis techniques, we can easily determine the seasonal evolution of the amplitude of the transient oscillations by retaining the running quadratic mean, σ_N, over N con-

secutive data, of the difference between the pressure observation for a given sol and the corresponding value computed from the eight-harmonic reference pressure cycle. In order that the seasonal amplitude of the transient waves, but nothing else, be retained in the running mean, the value of N must be significantly larger than the period of the transient waves, and significantly smaller than the length of seasons. As already stated, the analysis is based on diurnally averaged data and thus the transient eddies do not account for the thermal tides.

In Figure 4 we show the time evolution of σ_{25} (thick solid curve) and σ_{50} (dotted curve), which appear to be very similar. For diagnostics, we will retain the eight-harmonic fit to σ_{50} (thin solid line).

Transient eddies occur from northern autumn equinox to early northern spring. The amplitude of the perturbations at the Viking 2 site is about twice that at Viking 1 and shows a two-peak structure with a relative minimum at northern winter solstice.

3. Model

General Description

The LMD Martian GCM has been described in paper 1. The dynamical part is based on a finite difference formulation of the "primitive equations" of meteorology developed for the Terrestrial GCM of LMD [*Sadourny and Laval*, 1984] . We use a more recent formulation (developed by R. Sadourny and P. Le Van and described by F. Hourdin and O. Talagrand, Superrotation of planetary atmospheres: A numerical study, submitted to *Journal of Atmospheric Science*, 1994) which allows us to change automatically the distribution of grid points in longitude and/or latitude. The formulation exactly conserves (1) the total atmospheric mass (in absence of CO_2 condensation), (2) the potential enthalpy and its square for adiabatic motions, (3) the potential enstrophy for barotropic flows, and (4) the total angular momentum for axisymmetric flows.

The physical part has already been presented in detail in paper 1: (1) the radiative transfer is computed using a code adapted by *Hourdin* [1992] from the code currently used at the European Center for Medium-Range Weather Forecasts (ECMWF) [*Fouquart and Bonnel*, 1980; *Morcrette et al.*, 1986] ; it includes absorption and emission by carbon dioxide and dust in the thermal infrared and absorption and scattering by dust in the visible; (2) a simple formulation is used for the vertical turbulent mixing in the whole atmosphere; (3) vertically unstable profiles are prevented by applying a vertical convective adjustment which conserves energy; momentum is also mixed in the unstable layer, the intensity of the mixing depending on the intensity of the vertical instability; (4) the temperature of the surface is computed using an 11-level model of thermal conduction in the soil which correctly simulates the response for oscillatory forcing with periods going from a tenth of day up to a few years; (5) condensation and sublimation of carbon dioxide are computed in a simple energy and mass conserving manner. If somewhere in the atmosphere, the temperature falls below the condensation temperature T_{CO_2}, condensation occurs in an amount appropriate to restore T_{CO_2} by latent heat release. All the condensed CO_2 is assumed to precipitate instantly to the ground, the surface pressure being consequently modified. At the surface, the temperature of the frost is kept at the condensation value either by condensing atmospheric CO_2 or by sublimating CO_2 ice.

The approximation for the pressure-vapor curve used in paper 1 is replaced by a more accurate relationship based on the Clausius-Clapeyron relation for perfect gas. Assuming that the latent heat of sublimation L is independent of temperature, the vapor pressure curve reduces to

$$\ln(p/p_0) = \frac{L}{R}\left(\frac{1}{T_0} - \frac{1}{T_{CO_2}}\right) \tag{2}$$

(where R is the gas constant and T_0 is the condensation temperature corresponding to the pressure chosen as a reference, here $p_0 = 1$ mbar). $L = 5.9 \times 10^5$ J kg^{-1} and $T_0 = 136.3$ K are fixed in the range of experimental values [e.g., *James et al.*, 1992] . The change of the pressure-vapor curve was found to have a minor effect on the atmospheric mass budget.

The only significant change with respect to paper 1 is linked to the diurnal cycle. It was intentionally turned off in the previous study. Here, the radiative forcing is computed about 40 times per day. It then becomes necessary to introduce the dependency of the surface drag coefficient (a constant in paper 1) with both the magnitude of wind and vertical stability above the surface which are known to vary significantly between night and day. In the new version, we use the formulation developed by *Louis* [1979] for the terrestrial planetary boundary layer, in the form which has been implemented at ECMWF.

The drag coefficient is given by

$$C = \left(\frac{k}{\ln\frac{z}{z_0}}\right)^2 f\left(R_i, \frac{z}{z_0}\right) \tag{3}$$

where z_0 is the roughness length, z is the height of the middle of the first atmospheric layer, $k = 0.4$ is the von Karman constant, and

$$R_i = \frac{gz(\theta - T_s)}{\theta ||V||^2} \tag{4}$$

is the Richardson number, where $\theta = T(p_s/p)^\kappa$ and $||V||$ are the potential temperature and wind velocity in the first atmospheric layer and T_s the surface temperature. Two different functions f are used for momentum (f_m) and potential temperature (f_θ). In either case, a different formulation is used, depending on whether the atmosphere is stable ($R_i > 0$),

$$f_m = \frac{1}{1 + 2bR_i/\sqrt{1 + dR_i}} \tag{5}$$

$$f_\theta = \frac{1}{1 + 3bR_i\sqrt{1 + dR_i}} \quad (6)$$

or unstable ($R_i < 0$),

$$f_m = 1 - \frac{3bR_i}{1 + 3bC\left[k/\ln(z/z_0)\right]^2 \sqrt{-R_i\,(1 + z/z_0)}} \quad (7)$$

$$f_\theta = 1 - \frac{3bR_i}{1 + 3bC\left[k/\ln(z/z_0)\right]^2 \sqrt{-R_i z/z_0}} \quad (8)$$

As in the original terrestrial version, we use $C = b = d = 5$.

Grid

As in paper 1, we use two different horizontal resolutions in the present study: a low horizontal resolution based on a grid with 24 points in latitude and 32 points in longitude and a high resolution with 48 points in latitude and 64 points in longitude. In the vertical, the atmosphere is discretized using 15 σ-levels (σ is the pressure normalized by its surface value $\sigma = p/p_s$), the distribution of which is given in paper 1. The middle of the first layer is located at about 40 m above the surface and the 15th at about 60 km (the pressure is zero at the top of this layer). As pointed out in paper 1, even a coarse horizontal resolution (the low resolution) is sufficient to reproduce, at least qualitatively, the main features of the available observations of the Martian atmospheric circulation. The low resolution was generally retained for sensitivity studies.

Surface Conditions

In addition to the model parameters, three maps are introduced for the surface conditions: orography, visible albedo, and thermal inertia. These maps correspond to the consortium data set, completed by more recent estimation of the surface thermal inertia and albedo in the polar regions [*Keegan et al.*, 1991; *Paige and Keegan*, 1991], which were provided to us by J. B. Pollack.

4. The Control Simulation

Description

In the present section, we describe a first simulation, which is afterward used as a reference for the sensitivity study. In this "control simulation," the dust visible optical depth is a constant over time and space. Its value is set to 0.2, which corresponds to typical clear-sky conditions [*Martin*, 1986] . Vertically, the dust mixing ratio is almost constant from the surface up to the 0.3-mbar level and then strongly decreases, as proposed by *Pollack et al.* [1990]. The mixing length for vertical turbulent mixing is fixed to 35 m as in the LMD terrestrial GCM. The surface roughness is set to 1 cm everywhere, a value typical of near Viking lander conditions [*Sutton et al.*, 1978] . The emissivity of the bare surface (without ice) is fixed to 0.95, as suggested by *Santee and Crisp* [1993]. The CO_2-ice thermal emissivity ϵ and visible albedo A are set to 0.7 and 0.5, respectively (*Pollack et al.* [1993] have found that those values give reasonable fits to the Viking pressure observations). The dust optical properties are fixed, following *Pollack* [1982], with single-scattering albedo and asymmetry factor in the visible of 0.86 and 0.79, respectively, and a ratio between the absorption efficiency in the thermal infrared $Q_{\mathrm{abs}}(\mathrm{IR})$ and extinction efficiency in the visible $Q_{\mathrm{ext}}(\mathrm{Vis})$ of 0.2.

As mentioned in paper 1, the interaction between explicitly resolved and unresolved motions is parameterized using an iterated Laplacian acting on the potential temperature and wind components. The operator depends upon one time constant, τ_{diss}, which corresponds to the time of dissipation of the smallest resolved scales. The Laplacian is iterated in order to be more selective in the smaller horizontal scales not acting directly on the large-scale dynamics. Here, the number of iterations n_{diss} is fixed to 2, and τ_{diss} is set to 20,000 s between the surface and the 0.1-mbar level, with rapid decrease to $\tau_{\mathrm{diss}} = 10,000$ s above. With such values (which were tuned in order to ensure model stability), the characteristic timescale for the parameterized horizontal dissipation is of the order of 0.2 sol for zonal wave numbers $n \sim 12$, of a few sols for $n \sim 6$, and of more than a few tens of sols for $n \leq 3$.

The initial state was taken from an old simulation and, after a phase of stabilization (shorter than one Martian year) due to the modification of model parameters, we simulated 5 consecutive Martian years in order to estimate the interannual variability of the model results, which is of course, crucial for sensitivity studies. For comparison with Viking observations, the simulated surface pressure is interpolated at the two Viking sites from the four closest grid points. The Viking 1 point is somewhat shifted in longitude (from 47.9W to 42.5W) in order to match the currently assumed altitude of the landing site, 1.5 km below the reference altitude of Mars. For Viking 2 (which is generally assumed to be 1 km below Viking 1), since none of the the four closest grid points is below -2 km, the surface pressure is extrapolated downward based on the hydrostatic relationship using the near surface simulated air temperature. As for the Viking observations, the surface pressures are afterward decomposed in terms of seasonal variations and transient eddies. For the seasonal evolution, instead of comparing directly the observed and simulated pressure at the two Viking sites, we present on the one hand, direct comparisons for the Viking 1 site only and use, on the other, the relative pressure difference between Viking 2 and 1 in order to compare the observed and simulated meteorological component.

Viking 1 Pressure

The eight-harmonic fit to the simulated pressure at Viking 1 is reported in Figure 5 (thin solid curves). The interannual variability of the simulated seasonal pressure variations appears to be very weak. The amplitude of the seasonal variations is comparable to Viking 1 observations (dashed curve). This amplitude was some-

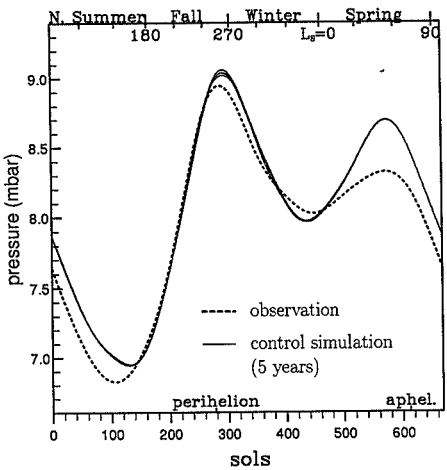


Figure 5. Eight-harmonic L_S fits to the seasonal evolution of the surface pressure at Viking 1. The thin solid curves, almost superimposed, correspond to the five consecutive years of the control simulation. The thick dashed curve (shown in Figure 1) corresponds to Viking 1 observations.

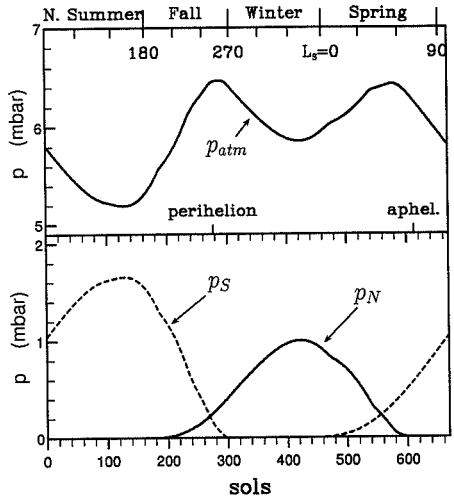


Figure 6. Time evolution of the mean atmospheric surface pressure $p_{\rm atm}$ (upper panel) and that of the equivalent surface pressure of the northern (p_N) and southern cap (p_S) for the first year of the control simulation. The sum of the three components $p_{\rm tot}$ (not shown) does not evolve along seasons and is equal to 6.854 mbar for the control experiment.

what overestimated in the simulations of paper 1. The difference is mainly due to the change of the albedo and emissivity of the polar ice, whereas the introduction of the diurnal cycle (which was tested independently) appears to have a minor effect on the seasonal pressure variations. In the present simulation, the minima and maxima occur approximately at the right seasons. However, a significant discrepancy exists between observations and model, the second pressure maximum being too high in the simulation. The amplitude of the second harmonic is also too large if compared to that of the first harmonic (compare Table 2 for the simulated pressure with Table 1) resulting from an insufficient asymmetry between the two hemispheres.

Planetary Mean of the Surface Pressure

The planetary mean of the surface pressure, $p_{\rm atm}$, is shown in Figure 6 for the first year of the control simulation, as well as the equivalent pressure of the CO_2 trapped in each cap, p_N and p_S. First, it must be noticed that there is no permanent polar cap in our numerical experiments (this is a common point of most numerical studies of the atmospheric mass budget on Mars). Since the maximum mass of a given cap coincides with a bare pole in the other hemisphere, each pressure minimum really coincides with the maximum mass of the winter cap. The pressure maxima correspond to periods where the two caps are present but very small. The interannual variability of those three curves is very weak and would not be visible on the graph. Thus the small variability in the Viking 1 pressure (Figure 5) is caused almost exclusively by changes in the internal mass redistribution (meteorological component).

Table 2. Phase Lag (Solar Longitude of the Maximum Pressure) and Amplitude of the First Two Harmonics of the Pressure Seasonal Cycle as Simulated at Viking 1 for the First Year of the Control Simulation

Mode k	$p^{(k)}$, mbar	$L_S^{(k)}$, deg
0	8.063	
1	0.613	320.19
2	0.628	69.87

Relative Pressure at the Two Sites and Interannual Variability

Figure 7 shows the simulated relative pressure difference between the two landing sites for the five years. The amplitude is in good agreement with Viking observations with a maximum near northern winter solstice and a minimum at southern winter solstice, but high-frequency variations are weaker in the simulation.

The simulation shows year-to-year variability but smaller than the discrepancy from Viking data. These interannual variations are due to modifications in the global latitudinal pressure variations. We will concen-

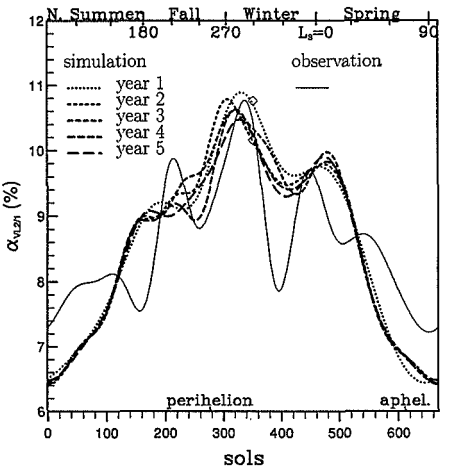


Figure 7. The relative pressure difference between Viking 2 and Viking 1. The thin solid curve (shown in Figure 3) is the observation, and the thick curves correspond to the five successive years of the control simulation.

trate our analysis on the difference between the first and second year of the simulation at $L_S \sim 300$. The corresponding values of $\alpha_{VL2/1}$ have been marked with diamonds in Figure 7. Relative to the Viking 1 pressure, the Viking 2 pressure is about 0.55% lower in the second year. This is directly related to a change in the zonal wind shown in Figure 8. The upper panel shows the mean zonal wind for the first two years for $L_S = 295$ to 305. We recognize (1) the strong eastward jet in the northern middle and high latitude in geostrophic balance with the strong latitudinal gradient at the edge of the polar cap; (2) the near-surface eastward jet in the southern tropics created, as the Indian monsoon winds on Earth, by transport of angular momentum across the equator in the lower branch of the Hadley cell; and (3) rather weak winds in the rest of the southern hemisphere. The difference between year 1 and year 2 (lower panel) can be interpreted as an equatorward displacement of the winter jet.

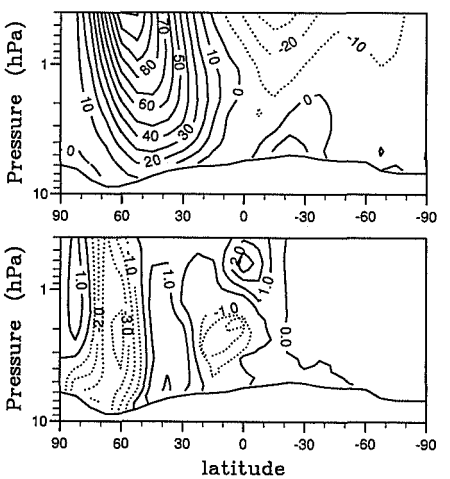


Figure 8. Zonally averaged zonal wind (m s^{-1}) in the control experiment, just after northern solstice (L_S = 295 to 305). The upper panel shows the mean zonal wind for the first two years of the simulation, and the lower panel the difference between year 1 and 2 (year 2 − year 1).

The wind in the third layer of the GCM ($p/p_s = 0.966$ and $z \sim 300$ m) is shown for the two years in the upper panel of Figure 9 and the latitudinal pressure variation corresponding to geostrophic balance with this near surface zonal wind is reported in the middle panel of the same figure. The relative difference between year 1 and year 2 of that dynamical component is the solid curve of

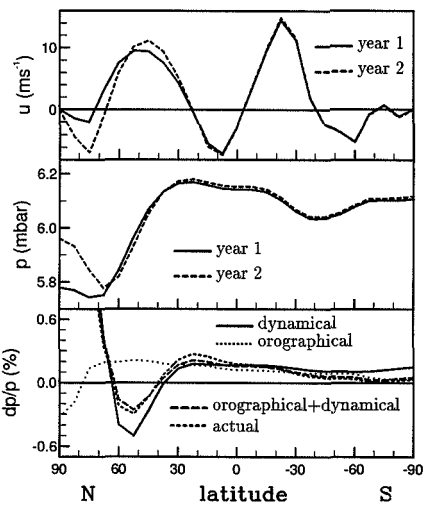


Figure 9. Diagnostic of the pressure differences between year 1 and year 2 of the control simulation (see paper 1 for details). **(Upper panel)** Zonally averaged zonal wind in the third atmospheric layer ($p/p_s = 0.966$ and $z \sim 300$ m) for the first two years. **(Middle panel)** Latitudinal pressure profiles in geostrophic balance with the zonal winds of the upper panel. **(Lower panel)** Various contributions to the pressure modification between year 1 and 2 (percent): dynamical (computed as the relative difference between the two curves of the middle panel) and orographical contributions and their combined effect (computed as the sum of the latitudinal derivative of the two contributions) as well as the actual pressure change.

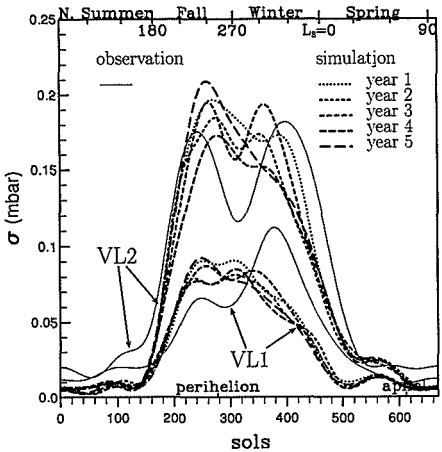


Figure 10. Amplitude of the transient pressure variations at Viking sites. The thin solid curve (shown in Figure 4) is the observation, and the thick curves correspond to the five successive years of the control simulation.

the lower panel. The corresponding pressure changes at Viking latitudes (22N for Viking 1 and 48N for Viking 2) are compatible with the change in the pressure simulated at Viking sites, which means that these pressure changes are almost totally explained by the modification of the mean zonal winds. Note also the good agreement between the actually simulated pressure modification and that reconstructed from the orographical and dynamical contributions.

Transients

Finally, the seasonal evolution of the amplitude of the simulated transient eddies is compared to Viking data in Figure 10. The agreement with Viking data is rather good in view of the very coarse horizontal resolution of this simulation. This is probably the consequence of the importance of transient waves with small wave numbers, which can be well simulated, even with such a coarse resolution. Once again, there is an interannual variability, but less than the discrepancy with observations. The model does not simulate the decrease of the amplitude of the transient eddies near northern winter solstice, and the decrease of this amplitude at the end of the northern winter is somewhat anticipated.

5. Sensitivity to Model Parameters for Clear-Sky Conditions

We present a series of numerical experiments performed by changing some model parameters, starting at $L_S = 330°$ (from a state of the control simulation) just before the formation of the southern polar cap in order to minimize the memory of the climatic system. In all the simulations the northern cap disappears near $L_S = 90$, after which the seasonal mass cycle adjusts in the sense that the total atmospheric mass is almost unchanged after one more year of simulation. Each sensitivity experiment thus includes an adjusting period going from $L_S = 330°$ to $L_S = 98°$ (date of the Viking 1 landing) followed by a 1-year simulation used for comparisons with the control simulation and Viking observations.

In this first set of experiments, the dust optical depth is kept to its "clear-sky" value: $\tau = 0.2$.

Horizontal Resolution

A simulation was performed at high horizontal resolution (64 longitude points and 48 latitude points) using exactly the same set of parameters as for the control simulation except for the near-surface horizontal diffusion time constant, which was set to $\tau_{\rm diss} = 12,000$ s. This corresponds to a horizontal dissipation about 8 times weaker than that of the control simulation at a given spatial scale.

The seasonal variations of the surface pressure are compared to the control experiment and to the observations in Figure 11. In terms of the Viking 1 seasonal pressure variations, a simple quadratic error analysis would suggest that the agreement with observations is worse with the high resolution. However, it must be kept in mind that the total CO_2 mass is an adjustable parameter which directly influences the mean value of the simulated pressure. In fact, the results are somewhat better for the high resolution, since the differ-

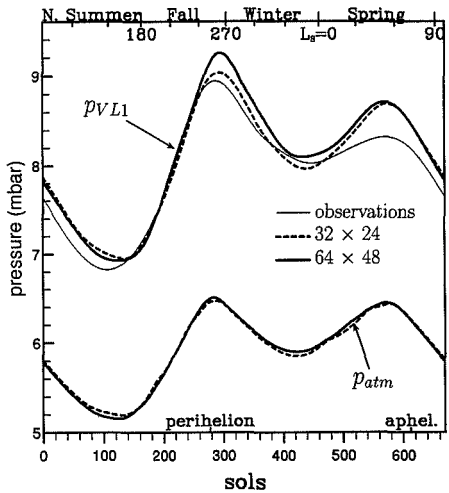


Figure 11. Viking 1 pressure, $p_{\rm VL1}$, as simulated with the low (32 longitude points and 24 latitude points) and high (64 longitude points and 48 latitude points) horizontal resolutions, Viking observations and the simulated mean atmospheric pressure, $p_{\rm atm}$.

ence between the two pressure maxima is somewhat enhanced. The change does not come from the atmospheric mass budget, which is almost unmodified (see the lower curves in Figure 11), but from the meteorological contribution α_{VL1}. The representation of the differential pressure between Viking 1 and 2 is also closer to observations as shown in Figure 12, with a better amplitude and stronger rapid oscillations (Figure 7) than for the low resolution. However, the improvement of the amplitude must be taken rather carefully: it can be the consequence of changes in the representation of the winds but also of the modification of the representation of the orography at Viking sites.

For transient eddies (Figure 13), the main improvement, with respect to the control simulation (see Figure 10), is their longer persistence during late winter. Note that the main discrepancy with Viking observations is an overestimation of the amplitude of the transients near northern winter solstice.

In all the sensitivity experiments performed herein, we use the low horizontal resolution (in order to be able to perform a large number of experiments at a reasonable numerical cost), although a finer horizontal resolution would have been preferable to correctly represent the atmospheric dynamics.

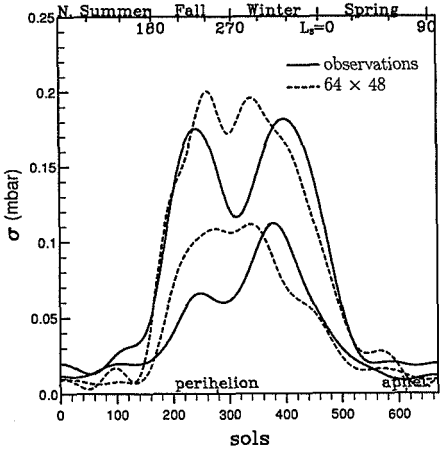


Figure 13. Amplitude of the transient pressure variations at Viking sites computed from observations and from the results of the high-resolution simulation (64 longitude points and 48 latitude points).

Surface Roughness

Sutton et al. [1978] deduced values of 0.1-1 cm for the surface roughness length, based on the camera images of the surface around the Viking landers. These values, which were derived for local studies of the planetary boundary layer, are not representative of a mean Martian surface and may also not be appropriate for GCM's. It must be noted that even for terrestrial GCM's, a good determination of the roughness length is an open question. At the same time, this surface roughness strongly influences the near-surface wind and, in turn, the dynamical component of the horizontal pressure variations. Some recent experiments performed with the LMD terrestrial climate GCM, have shown, for instance, that a decrease by a factor of 2 of the surface drag coefficient leads to latitudinal pressure gradients 25% steeper in the southern midlatitudes (30-60S), where because of the absence of a significant orography, the horizontal pressure variations are completely dominated by the geostrophic balance with near-surface winds (dynamical component).

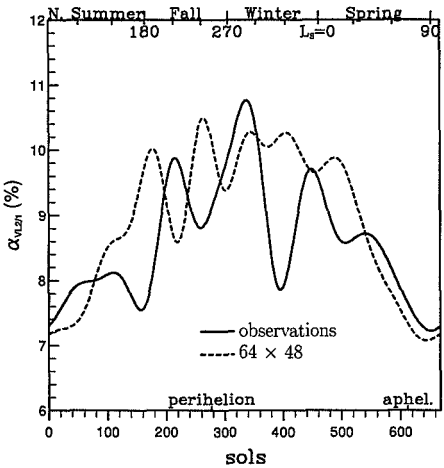


Figure 12. The $\alpha_{VL2/1}$ computed from observations and from the results of the high-resolution simulation (64 longitude points and 48 latitude points).

Four simulations were conducted for z_0 = 0.1 mm, 1 mm, 10 cm and 1 m (the control value was 1 cm). The simulated atmospheric pressure p_{atm} is almost identical for the five experiments (including control), but there is significant effect on the relative pressure difference between Viking 2 and 1 (Figure 14): for the period L_S = 300 to 307, for instance, $\alpha_{VL2/1}$ = 9.7% for z_0 = 1 mm and $\alpha_{VL2/1}$ = 10.7% for z_0 = 10 cm (diamonds in Figure 14). As for the interannual variations of the control experiment, the local pressure changes at Viking sites are representative of the modification of the zonally averaged surface pressure (a 0.5% decrease at 22N and a 0.6% increase at 48N, lower panel of Figure 15) and are a direct consequence of the reduction of the intensity of the near-surface zonal wind (upper panel). Note that the change does not affect significantly (in comparison to the interannual variations) the global zonal wind distribution, but selectively, the near-surface winds. It is also noticeable that such large

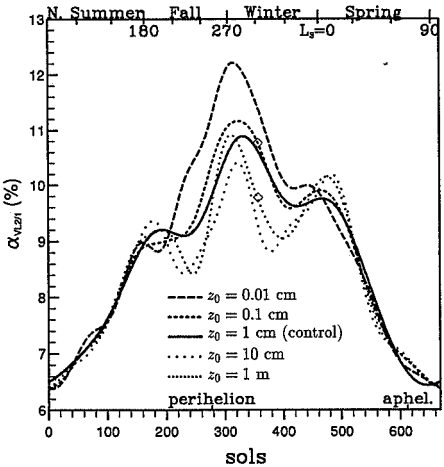


Figure 14. Relative meteorological factor between Viking 2 and 1 for various values of the surface roughness length z_0. The two diamonds (from the $z_0 = 0.1$ cm and $z_0 = 10$ cm experiments) are analyzed in detail in the text.

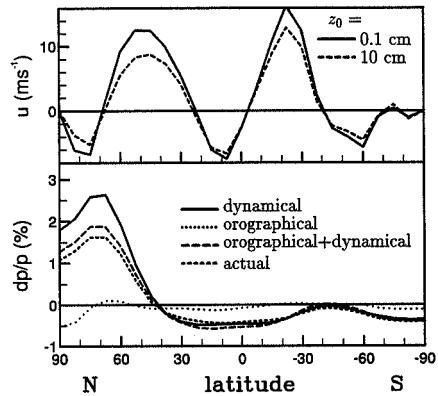


Figure 15. (Upper panel) Mean zonal wind (m s^{-1}) in the third atmospheric layer ($p/p_s = 0.966$ and $z \sim 300$ m) for the $z_0 = 0.1$ cm and $z_0 = 10$ cm sensitivity experiments for the period $L_S = 295 - 305$. **(Lower panel)** Various contributions to the pressure modification between $z_0 = 0.1$ cm and $z_0 = 10$ cm (percent): dynamical and orographical contributions and their combined effect (computed as the sum of the latitudinal derivative of the two contributions) as well as the actual pressure change.

changes in the surface drag do not affect significantly the amplitude of the transient pressure variations (not shown).

Caps Albedo and Emissivity

Four experiments were performed by changing the ice visible albedo A or thermal emissivity ϵ simultaneously on the two caps. The four experiments correspond to $A = 0.4$, $A = 0.6$, $\epsilon = 0.6$, and $\epsilon = 0.8$. Albedo and emissivity have a direct impact on the energy budget on the caps and therefore strongly affect the mass budget. This is illustrated in Figure 16 and Figure 17, which show the time evolution of the equivalent pressure of the two caps. As expected, a reduction of the ice albedo (Figure 16) increases the absorption of the solar radiation, and reduces the mass of the polar cap. Emissivity acts in the same direction (Figure 17): less thermal emission by the cap results in smaller polar caps. However, the effect of the change in albedo and that of emissivity are not completely equivalent: a change in the ice emissivity has a rather constant effect during both the formation and recession phase, whereas albedo changes have a much weaker effect during the cap formation which mainly occurs in the polar night.

The dynamical component, $\alpha_{VL2/1}$, and the transient pressure variations are not significantly affected by the modifications of the ice properties.

Other Uncertain Parameters

The choice of the parameters which have been varied involve very different physical processes and may give, all together, an overview of the sensitivity of the simulated atmospheric circulation to the values of the model parameters: the changes of horizontal resolution and dissipation time constants directly influence

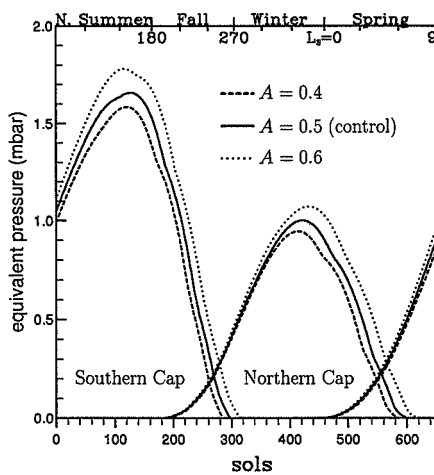


Figure 16. Time evolution of the equivalent pressure of both polar caps for various values of the ice albedo: $A = 0.4$, 0.5 (control), and 0.6.

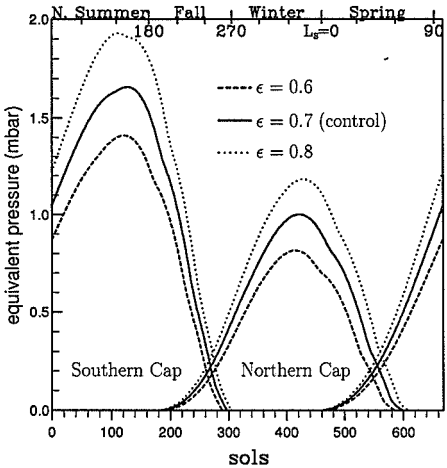


Figure 17. Time evolution of the equivalent pressure of both polar caps for various values of the ice emissivity: $\epsilon = 0.6$, 0.7 (control), and 0.8.

the global atmospheric dynamics and, in particular, the representation of the transient and stationary eddies; the surface roughness has a rather selective impact on the meteorological pressure variations, not affecting significantly the general dynamics and energy and mass budget, whereas the emissivity and albedo changes selectively affect the energy and mass balance of the polar regions. However, some other very uncertain parameters can significantly influence our results. *Wood and Paige* [1992] have shown, for instance, that the seasonal pressure variations, simulated with an energy balance model, are very sensitive to the soil thermal inertia. The thermal inertia we use [*Keegan et al.*, 1991; *Paige and Keegan*, 1991] corresponds to the skin surface and may not be adapted for seasonal variations. The Martian orography is also rather poorly known. It has a direct influence on the meteorological contribution as well as on the emissivity of the polar caps since their temperature is directly related to the surface pressure through the solid-vapor pressure curve.

6. Dusty Conditions

Another series of sensitivity experiments was performed for different values of the dust opacity.

Dust optical depths show a very large spatial and temporal variability on Mars, large enough to significantly affect the atmospheric energy and mass budget [e.g., *Pollack et al.*, 1990, 1993]. Unfortunately, there is no complete climatology of the Martian dust optical depth, at least for a year without great dust storms, which could be used to constrain numerical simulations (the best temporal coverage corresponds to the Viking infrared thermal mapper (IRTM) data analyzed by *Martin and Richardson* [1993]). In this section, we first present simple sensitivity experiments performed with various values of the global dust optical depth: $\tau = 0.2$, 0.5, 1, and 2.5. A stronger horizontal dissipation in the upper atmosphere was needed for model stability when increasing the dust optical depth, probably due to shorter radiative time constants. For $\tau = 0.2$, 0.5, and 1, we use the same time constant as for the control experiment in the lower atmosphere ($\tau_{\text{diss}} = 20,000$ s), but $\tau_{\text{diss}} = 4000$ s above 0.15 mbar. For $\tau = 2.5$, we set $\tau_{\text{diss}} = 15,000$ s in the lower atmosphere and 2000 s in the upper atmosphere. The $\tau = 0.2$ experiment (same value as for the control experiment) was performed in order to make the three first experiments ($\tau = 0.2$, 0.5, and 1) identical except for the dust optical depth, and to check the impact of changing the horizontal dissipation parameters for the $\tau = 0.2$ case.

Effect on the Mass Cycle

The first effect of an increase of the dust opacity in our model is a global increase of the amount of carbon dioxide trapped in the caps, as shown in Figure 18, somewhat in contradiction with the results obtained by *Pollack et al.* [1990], which showed a small decrease of the condensation rate at $L_S = 279$ for increasing dust optical depths. In fact, the reduction of the atmospheric mass, in our simulations, is mainly due to a reduction of the sublimation rate during cap recession, especially near spring equinox, the formation phase being much less affected. The extension of the caps is also modified (Figure 19): for larger values of τ, the caps are smaller during their formation but larger at the end of the recession.

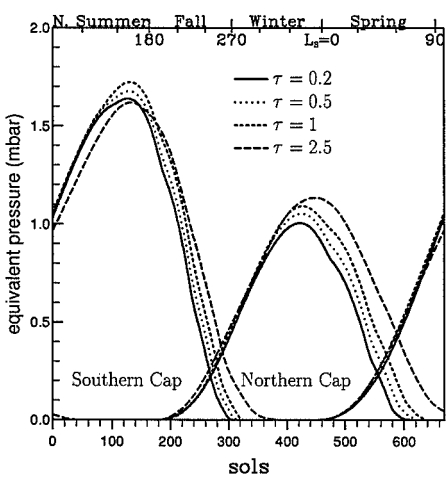


Figure 18. Time evolution of the equivalent polar cap pressures p_N and p_S for the various dust sensitivity experiments ($\tau = 0.2$, 0.5, 1, and 2.5).

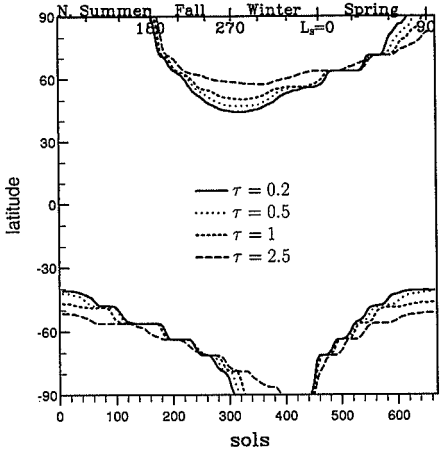


Figure 19. Time evolution of the mean latitude of the edge of the polar caps for the dust sensitivity experiments ($\tau = 0.2$, 0.5, 1, and 2.5).

The sensitivity of the climate to the dust optical depths involves radiative as well as dynamical mechanisms. Here, we center our analyses on the $\tau = 0.2$ and $\tau = 1$ experiments (which are representative of the seasonal variability of dust for years without global dust storm) and on a period just before northern solstice ($L_S = 253 - 259$) when the northern cap is forming while the southern cap is still subliming.

Modifications of the Radiative Budget

The solar radiation (noted SW hereafter) reaching the surface is much smaller in the $\tau = 1$ than in the $\tau = 0.2$ experiment, by about 16% in midlatitude and 40% (40 W m^{-2}) on the southern cap (upper panel in Figure 20; note that the insolation is null from the northern pole down to 67°N during that period). The total planetary albedo is also somewhat reduced in the $\tau = 1$ experiment (see SW$^{\uparrow}$, second panel of Figure 20), the increase of the atmospheric albedo, resulting from dust scattering, being smaller than the reduction of the visible light reflected by the surface, especially on the highly reflecting polar ice. As a result, for $\tau = 1$, more visible radiation is absorbed by the global system (atmosphere plus solid planet), and much more by the atmosphere alone, especially in the summer hemisphere, which increases the latitudinal thermal contrasts.

The surface temperature is almost identical in the two simulations, as can be seen from the surface thermal emission ($\epsilon_s \sigma T_s^4$, Figure 20), which is only significantly affected near the edge of the southern cap due to a larger ice cover in the 70-80S region at that particular season (also responsible for the larger planetary albedo for $\tau = 1$ in this latitude range). The atmospheric thermal radiation absorbed by the surface ϵ_s LW$_s^{\downarrow}$ is larger in the $\tau = 1$ experiment, as expected from the larger atmospheric emissivity. This flux is also affected by the modifications of the atmospheric temperature shown in Figure 21. The atmosphere is globally warmer in the $\tau = 1$ experiment, as a consequence of the larger direct absorption of solar radiation, but colder by about 5 K over the forming cap, up to the 3-mbar level. On the north pole itself, the temperature is not affected and corresponds in both cases to a condensing atmosphere

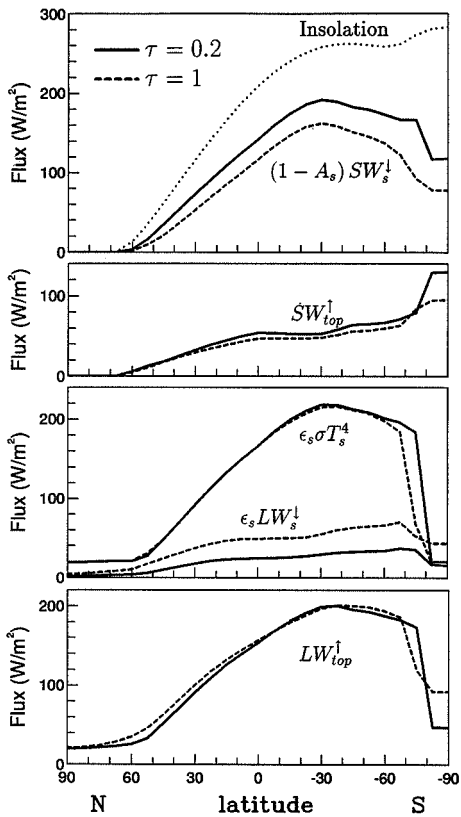


Figure 20. Radiative fluxes at the surface and top of the atmosphere computed for the $\tau = 0.2$ (solid curves) and $\tau = 1$ (dashed curves) sensitivity experiments just before northern winter solstice, $L_S = 253 - 259$. From top to bottom: (1) the insolation at the top of the atmosphere and solar radiation absorbed by the surface; (2) the outgoing shortwave radiation at the top of the atmosphere; (3) the surface thermal emission and the atmospheric thermal radiation absorbed by the surface; (4) the total thermal emission to space.

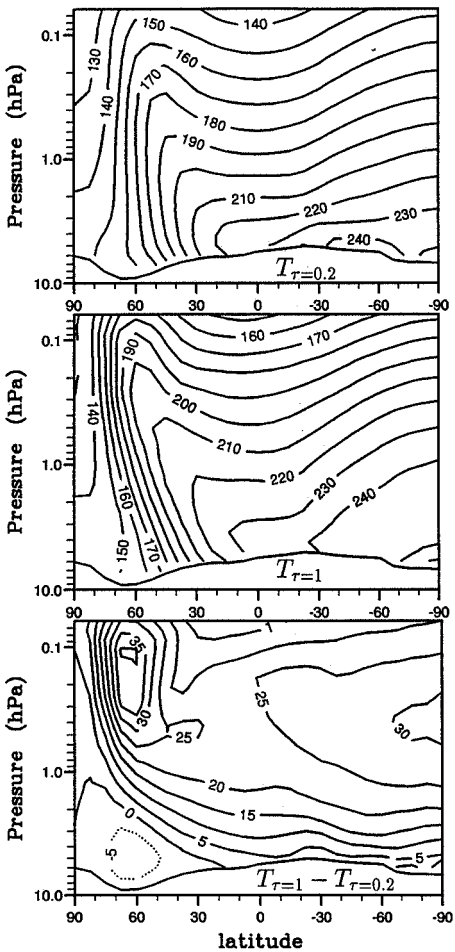


Figure 21. Zonally averaged temperature simulated just before northern winter solstice ($L_S = 253-259$) for the $\tau = 0.2$ (upper panel) and the $\tau = 1$ (middle panel) experiment as well as the difference between both (lower panel).

up to very low pressures. Note that the atmospheric condensation extends to much lower latitudes for the $\tau = 0.2$ than for the $\tau = 1$ simulation (not shown).

On the northern pole surprisingly, although the air temperature decreases with altitude, the thermal emission to space, $LW^{\uparrow}_{top}$, is (1) somewhat larger than the surface thermal emission for both simulations and (2) larger in the $\tau = 1$ than in the $\tau = 0.2$ experiment. Both results are the consequence of the low thermal inertia of the icy surface, which minimizes the role of the emission emitted by the surface and enhances the role of the thermal emission by the atmosphere (in particular through its reflection on the surface) as can be checked with a simple analytical one-layer radiative transfer model.

Note that for a ice emissivity $\epsilon = 1$, *Pollack et al.* [1990] found a decrease of the total condensation rate at this season for increasing dust amounts.

Finally in our simulations, the net effect, in terms of atmospheric radiative budget at the top of the atmosphere, of an increase of τ from 0.2 to 1 are (Figure 22, upper panel): (1) an increase of the hemispheric asymmetry with a larger gain of energy in the summer hemisphere and larger loss in the winter hemisphere (including the polar region); and (2) a decrease of the strong energetic gain on the subliming cap (except at the edge of the cap due to the different cap extensions of the two simulations).

Energy and Mass Budget in the Polar Regions

In fact, there is a great difference between the processes controlling the formation and the recession of the polar caps. During recession, the mass exchange is controlled by the surface heat balance (compare upper and

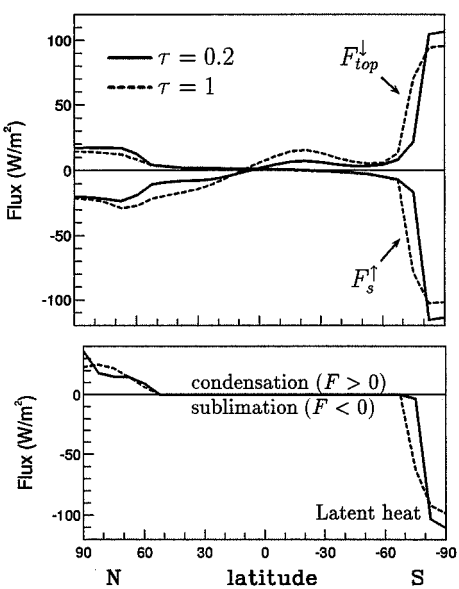


Figure 22. Net radiative fluxes at the top of the atmosphere and at the surface (upper panel, positive for a gain of energy by the atmosphere) and latent heat release (lower panel) computed for the $\tau = 0.2$ (solid curves) and $\tau = 1$ (dashed curves) sensitivity experiments just before northern winter solstice, $L_S = 253 - 259$. The net surface flux includes the turbulent flux.

lower panel of Figure 22) which is itself very close to the radiative balance at the top of the atmosphere, due to the weak role of the thermally driven circulation, shown in the upper panel of Figure 23. To the contrary, during winter, condensation occurs both in the atmosphere and on the surface, and since the radiative fluxes are much smaller, the latitudinal heat advection plays a significant role in the model sensitivity.

What we call thermally driven circulation hereafter is the total circulation minus the condensation flow (second panel of Figure 23). As proposed by *Pollack et al.* [1990], the condensation flow is defined at a given latitude as the zonally and vertically averaged meridional mass flux. The condensation flow accounts for the latitudinal mass redistribution and includes, in fact, in addition to condensation, the meteorological latitudinal mass redistribution. On the subliming cap, the condensation flow is the principal component of the latitudinal energy transport and just corresponds to the transport toward lower latitudes of the energy the atmosphere has gained by increasing its mass.

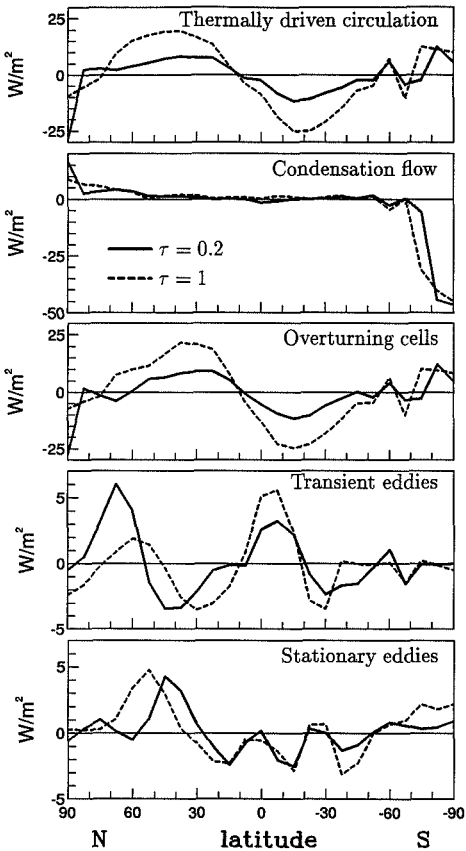


Figure 23. Atmospheric heating (in W m^{-2}) computed for various components of the atmospheric circulation and the $\tau = 0.2$ (solid curves) and $\tau = 1$ (dashed curves) sensitivity experiments just before northern winter solstice, $L_S = 253 - 259$. The thermally driven circulation is the difference of the total circulation and condensation flow. The overturning cell is the difference between the mean meridional circulation and condensation flow.

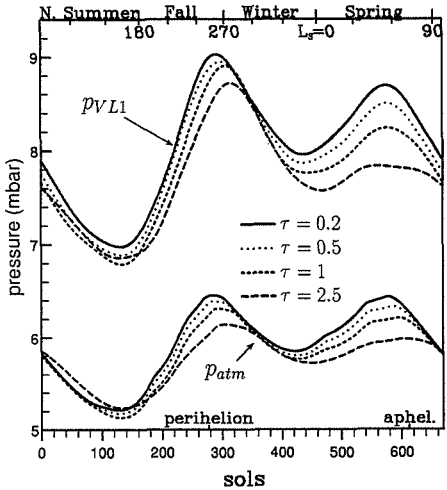


Figure 24. Simulated Viking 1 pressure and planetary averaged surface pressure p_{atm} for the dust sensitivity experiments ($\tau = 0.2$, 0.5, 1, and 2.5).

The cap recession is thus essentially controlled by the radiative balance, giving a simple explanation for the major difference between the $\tau = 0.2$ and $\tau = 1$ simulations. In late winter, when the latitudinal extension of the winter cap is maximum and the sublimation rate is already high in the midlatitudes, the smaller incident flux on the surface plus smaller extent of the cap in the $\tau = 1$ simulation explains the much slower cap recession at this season. Since the recession is slower, the cap finally becomes larger than in the $\tau = 0.2$ case, after which the smaller radiative flux in the $\tau = 1$ experiment is partially compensated by the larger surface of sublimation and leads to almost identical recession rates, in terms of mass, in late spring.

During the cap formation, the slightly larger atmospheric thermal emission to space in the $\tau = 1$ experiment tends to reduce the total (atmosphere plus surface) condensation. Heat advection by overturning cells (mean meridional circulation minus condensation flow) is much stronger in midlatitudes than for the $\tau = 0.2$ simulation, while the transient eddies, which transport energy from middle to high latitudes in the winter hemi-

sphere (as expected from baroclinic eddies), are strongly reduced.

The increased heating by the mean meridional circulation is mainly sensitive near the edge of the cap (50-70N) responsible for a weaker condensation in that region for $\tau = 1$ (lower panel of Figure 22) and for the smaller cap extension during the formation phase (Figure 19). To the contrary, the reduction of the poleward energy transport by transient eddies has a major impact in very high latitudes (70-85N), where it is responsible for a larger condensation rate in the $\tau = 1$ experiment. As a consequence, the mass of the cap is more concentrated near the pole in the $\tau = 1$ experiment.

Seasonal Pressure Variations

The change in the condensation-sublimation affects directly the mean atmospheric pressure $p_{\rm atm}$ (lower curves in Figure 24). For increasing dust opacities, the Viking 1 pressures (upper curves) are closer to observations, in that the second pressure maximum is more reduced than the first one. This is due in part to the modification of the meteorological component $\alpha_{\rm VL1}$ shown in Figure 25. It must be noticed however, that $\alpha_{\rm VL1}$ is mainly sensitive to the dust optical depth near northern summer solstice, which is in reality, the period of minimum atmospheric dust content [*Martin*, 1986].

The effect on the differential pressure between Viking 2 and 1 (Figure 26) is very different, in that it is large only during northern autumn and winter. It corresponds to a strong increase of the Viking 2 pressure for increasing dust opacities. As for the control and roughness sensitivity experiments, the modification of $\alpha_{\rm VL2/1}$ (by about 2% for $L_S = 295 - 300$) is representative of the modification of the zonally averaged surface pressure (see Figure 27): the main effect is a strong pressure

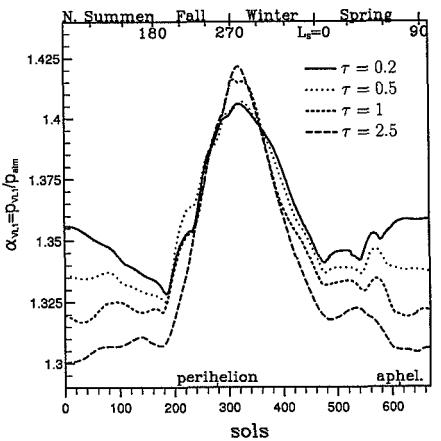


Figure 25. The $\alpha_{\rm VL1} = p_{\rm VL1}/p_{\rm atm}$ for the dust sensitivity experiments ($\tau = 0.2$, 0.5, 1, and 2.5).

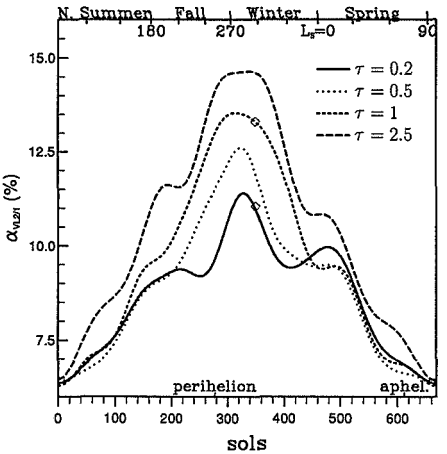


Figure 26. Relative pressure difference $\alpha_{\rm VL2/1}$ between the two Viking sites for the dust sensitivity experiments ($\tau = 0.2$, 0.5, 1, and 2.5).

increase in the northern middle and high latitudes (by about 4% at 60N), corresponding to a weakening of the circumpolar depression due to the reduction of the near-surface eastward winds in the 30N-60N region. In the present case, modifications of the near-surface winds

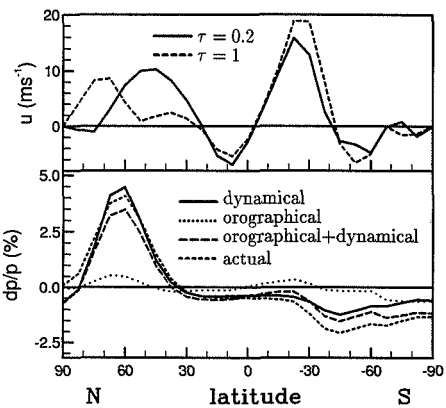


Figure 27. **(Upper panel)** Mean zonal wind (m s^{-1}) in the third atmospheric layer ($p/p_s = 0.966$ and $z \sim 300$ m) for the $\tau = 0.2$ and $\tau = 1$ experiments for the period $L_S = 295 - 300$. **(Lower panel)** Various contributions to the pressure modification between $\tau = 0.2$ and $\tau = 1$ (percent): dynamical and orographical contributions and their combined effect (computed as the sum of the latitudinal derivative of the two contributions) as well as the actual pressure change.

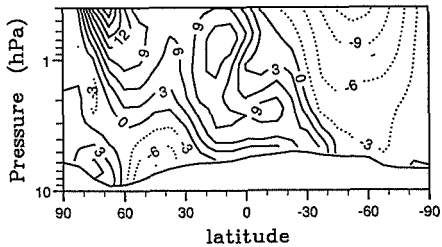


Figure 28. Zonal wind (m s^{-1}) modifications resulting from an increase of the dust visible optical depth from $\tau = 0.2$ to $\tau = 1$ ($L_S = 285$ to 300).

are the signature of large changes in the global wind structure (Figure 28) mainly due to an increase of the mean meridional advection of angular momentum (not shown).

Transient Eddies

The modification of the amplitude of the transient eddies (reported in Figure 29 for Viking 2) is also very instructive. First, the amplitude for the $\tau = 0.2$ experiment is in the range of the variability of the control simulation despite the stronger horizontal dissipation in the upper atmosphere. Note also the general but not systematic decrease of the amplitude of the transients for increasing optical depths, coinciding with the decrease of the latitudinal energy transport mentioned previously. More remarkable is the fact that the reduction is much more effective near northern winter solstice. This result may explain both (1) the two-peak structure found in the amplitude of the transient eddies for the Viking observations (the northern winter solstice corresponding to the maximum dust optical depth both for year with and without global dust storms) and (2) the fact that the amplitude of the transient eddies was much less reduced during the first global dust storm 1977-A, $L_S \sim 210 - 230$, than during the 1977-B dust storm, $L_S \sim 280 - 300$. Note also that the variance of the transient pressure variations in the $\tau = 2.5$ experiment for the period $L_S = 280 - 300$ (about 0.03-0.04 mbar) is of the order of that observed by Viking 2 during the 1977-B dust storm. The explanation for this seasonal behavior would require much more sophisticated diagnostics and is beyond the scope of the present paper, but it is worthwhile noticing that Martian GCM's seem to be well suited for further studies of the Martian transients (see *Barnes* [1980, 1981], and *Barnes et al.* [1993] for more complete studies).

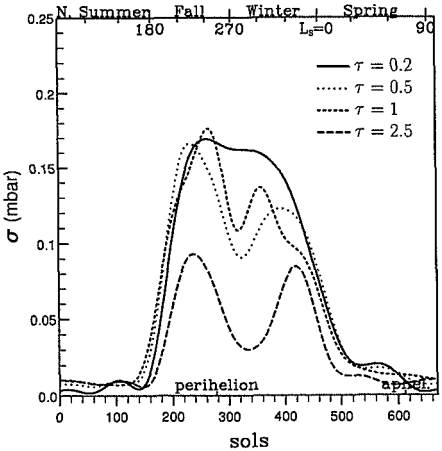


Figure 29. Amplitude of the transient eddies at Viking 2 for the dust sensitivity experiments ($\tau = 0.2$, 0.5, 1, and 2.5).

Time-Varying Dust Opacity

We also performed a sensitivity experiment varying the dust opacity as a simple cosine function of solar longitude ($\tau = 0.6 + 0.3\cos(L_S - L_{S0})$, with $L_{S0} = 280$), which qualitatively matches the seasonal evolution deduced for the Viking years, except for the period of dust storms [*Pollack*, 1982; *Martin*, 1986] .

This experiment is compared to the previous experiments for constant optical depths, in terms of the two first harmonics of the pressure simulated at Viking 1 (Table 3). The improvement of the simulation with increasing optical depth (described previously) is very clear in the two-harmonic decomposition: as τ increases, the ratio between the first and second harmonic increases. The experiment with variable optical depth (VAR1) does not show any significant improvement, even when compared to the $\tau = 0.2$ experiment.

Uncertainties Arising From Dust Optical Properties

Some of the results presented above are rather sensitive to the numerical values adopted for the dust optical properties. In particular, we use in the present paper a ratio $Q_{abs}(IR)/Q_{ext}(Vis)$ of 0.2 taken from *Pollack* [1982], whereas the radiative code of the Ames GCM, which accounts for scattering in the thermal infrared,

Table 3. Phase Lag (Solar Longitude of the Maximum Pressure) and Amplitude of the First Harmonics of the Pressure Seasonal Cycle as Simulated at Viking 1 for the Various Dust Sensitivity Experiments

Experiment	$p^{(0)}$, mbar	$p^{(1)}$, mbar	$L_S^{(1)}$, deg	$p^{(2)}$, mbar	$L_S^{(2)}$, deg
$\tau = 0.2$	8.067	0.592	320.36	0.623	69.99
$\tau = 0.5$	7.965	0.610	316.05	0.602	71.21
$\tau = 1$	7.841	0.626	313.34	0.573	75.10
$\tau = 2.5$	7.703	0.636	311.49	0.506	82.64
VAR1	7.968	0.586	314.22	0.621	73.10
VAR2	7.924	0.601	314.75	0.624	75.14

corresponds to a significantly higher value (J. M. Murphy, personal communication, 1994). The impact of this choice was tested numerically by rerunning the $\tau = 0.2$ and $\tau = 1$ experiments for $Q_{abs}(IR)/Q_{ext}(Vis) = 0.4$ and $Q_{abs}(IR)/Q_{ext}(Vis) = 1$. In both cases, contrary to the $Q_{abs}(IR)/Q_{ext}(Vis) = 0.2$ case, the surface temperature is globally warmer in the $\tau = 1$ experiment, the enhancement of the greenhouse effect being larger than the diminution of the amount of solar energy reaching the surface. Note also that for increasing optical depths, the total condensation on the forming cap increases for $Q_{abs}(IR)/Q_{ext}(Vis) = 0.2$, is almost unmodified for $Q_{abs}(IR)/Q_{ext}(Vis) = 0.4$, and decreases for $Q_{abs}(IR)/Q_{ext}(Vis) = 1$, which provides an additional explanation for the difference between the present results and that published by *Pollack et al.* [1990].

The simulation with time-varying dust opacity was also rerun for $Q_{abs}(IR)/Q_{ext}(Vis) = 0.4$, which did not produce any significant improvement in terms of comparison with Viking seasonal pressure variations (experiment VAR2 in Table 3).

7. Best Fit Simulations

Method

The method we propose in order to fit Viking pressure data is based on the decomposition of the local surface pressure presented above: for Viking 1, for instance,

$$p_{VL1} = \alpha_{VL1} \times p_{atm} = \alpha_{VL1} \times (p_{tot} - p_N - p_S) \quad (9)$$

(in that section, we always refer to the eight-harmonic fit instead of the pressure variations themselves). Hopefully, the time evolution of α_{VL1}, which cannot be deduced from observation, is much less sensitive to the model parameters than $\alpha_{VL2/1}$. For all the low-resolution simulations with $\tau = 0.2$, this factor does not vary by more than 3%. Even the experiments with varying dust opacities are within this range, the effect of dust opacity on the value of α_{VL1} being maximum near northern summer solstice (Figure 25) when the atmosphere is clearest. Finally, for our best fit simulation, α_{VL1} is just taken from the simulation used as a basis for the best fit procedure.

We then take advantage of the high sensitivity of the atmospheric pressure to the ice albedo A and emissivity ϵ to fit the Viking 1 pressure variations by allowing, for the two caps, independent values of both parameters, (A_N, ϵ_N) and (A_S, ϵ_S) for the northern and southern cap, respectively. We also allow variations of the total amount (ice plus gas) of carbon dioxide p_{tot}. Assuming that the CO_2 trapped in one particular cap is to a first order independent of the total CO_2 amount p_{tot} and of the ice properties of the other cap, the mean surface pressure p_{atm} can be written as

$$p_{atm} = p_{tot} - p_N(A_N, \epsilon_N) - p_S(A_S, \epsilon_S) \quad (10)$$

The change δp_{atm} of the mean atmospheric pressure, resulting from small changes of these five parameters, can be written formally as

$$\delta p_{atm} = \delta p_{tot} - \delta A_N \frac{\partial p_N}{\partial A_N} - \delta\epsilon_N \frac{\partial p_N}{\partial \epsilon_N} - \delta A_S \frac{\partial p_S}{\partial A_S} - \delta\epsilon_S \frac{\partial p_S}{\partial \epsilon_S}. \quad (11)$$

The sensitivity functions, $\partial p_N/\partial A_N$, $\partial p_N/\partial \epsilon_N$, $\partial p_S/\partial A_S$, and $\partial p_S/\partial \epsilon_S$, computed from the results of the sensitivity experiments performed with $A = 0.4$–0.6 and $\epsilon = 0.6$–0.8 are shown in Figure 30. The functions themselves (thin curves) show rapid oscillations. This may be partly explained by the fact that the ice cover is not fractional: either a mesh is completely icy, or it is bare. If, due to a change in the ice parameters, the icing of a mesh occurs later or sooner, it makes a discontinuous change in the sensitivity. We retained smoothed functions (thick curves) for the best fit procedure.

The differences between the response to emissivity and albedo changes mentioned previously appears very clearly in Figure 30: at the beginning of the cap formation, an increase of the ice emissivity increases the mass of the cap about three times more than an equivalent increase of the ice albedo, whereas albedo and emissivity have an equivalent effect in the late recession period. This is especially important for our present purpose since, by acting independently on the four ice parameters, we act on functions which are strongly uncorrelated, with four maxima rather regularly distributed within a year.

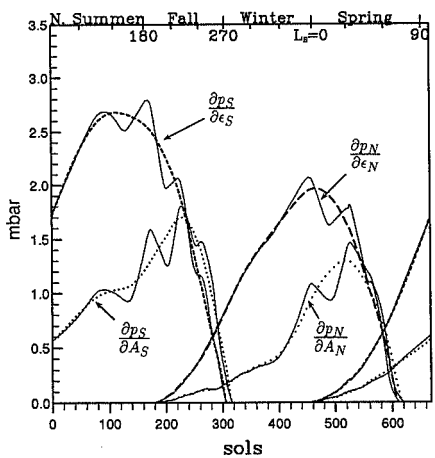


Figure 30. Sensitivity of the equivalent pressure of a given polar cap to a change of its albedo or emissivity, computed using the results of the sensitivity experiments presented previously. For instance, $\partial p_S/\partial A_S$ is computed as $[p_S(A_S = 0.6) - p_S(A_S = 0.4)]/0.2$. The thin lines correspond to the direct computation, and the thick (dotted or dashed) curves correspond to the smoothed functions used for the best fit procedure.

Finally, the best fit values of the parameters can be obtained by minimizing the following cost function measuring the difference between the simulated and observed pressures:

$$J\left(p_{\rm tot}, A_N, \epsilon_N, A_S, \epsilon_S\right) =$$

$$\sum_{\rm sol=1}^{668} \left\{ p_{\rm VL1} - \alpha_{\rm VL1} \left(p_{\rm tot} - p_N - p_S - \delta A_N \frac{\partial p_N}{\partial A_N} \right.\right.$$

$$\left.\left. -\delta\epsilon_N \frac{\partial p_N}{\partial \epsilon_N} - \delta A_S \frac{\partial p_S}{\partial A_S} - \delta\epsilon_S \frac{\partial p_S}{\partial \epsilon_S} \right) \right\}^2 \qquad (12)$$

(where the sum is done with one value per day on one Martian year). The cost function J can be easily minimized by cancellation of its derivative which is equivalent to inverting a five-equation linear system. However, such a direct minimization produces extreme values of the emissivity and albedo. In fact, good fits could also be obtained by imposing one of the ice parameters. We thus introduce a supplementary constraint: we look for the fit which requires the minimum artificial asymmetry between the two hemispheres by adding to the cost function J a second term

$$\xi \left[\left(A_S - A_N\right)^2 + \left(\epsilon_S - \epsilon_N\right)^2 \right] \qquad (13)$$

accounting for hemispheric asymmetries, where ξ represents the relative importance given to this asymmetry with respect to the departure from Viking 1 data (in fact, ξ has been normalized in order to have a transition region around $\xi = 1$). For very high values of ξ, we obtain the best fit with the same ice properties for both caps. For $\xi = 0$, we obtain the best unconstrained fit.

Once we have obtained, for a given value of ξ, a set of best fit parameters $p_{\rm tot}$, A_N, ϵ_N, A_S and ϵ_S, it is possible to compute a synthetic pressure curve at Viking 1 as

$$p = \alpha_{\rm VL1} \left(p_{\rm tot} - p_N - p_S - \delta A_N \frac{\partial p_N}{\partial A_N} \right. \qquad (14)$$

$$\left. -\delta\epsilon_N \frac{\partial p_N}{\partial \epsilon_N} - \delta A_S \frac{\partial p_S}{\partial A_S} - \delta\epsilon_S \frac{\partial p_S}{\partial \epsilon_S} \right)$$

Best Fit Results

We first present results of the best fit procedure applied to the control simulation. Figure 31 shows, for values of ξ varying from 10^{-3} to 10^4, the values of the best fit albedos and emissivities as well as the mean quadratic error between the synthetic and observed Viking 1 pressure. For $\xi < 1$, the best fit values are very different for the two caps, especially for albedo with $A_S = 0.3$ and $A_N = 0.65$. On the contrary, for large ξ, we find values of the ice parameters rather close the control values, with $\epsilon_S = \epsilon_N = 0.64$ and $A_S = A_N = 0.5$. Figure 32 shows the synthetic pressure at Viking 1 as obtained for $\xi = 0$, $\xi = 23$, and $\xi = 10^6$. Of course, the best agreement with Viking data (thin solid curve)

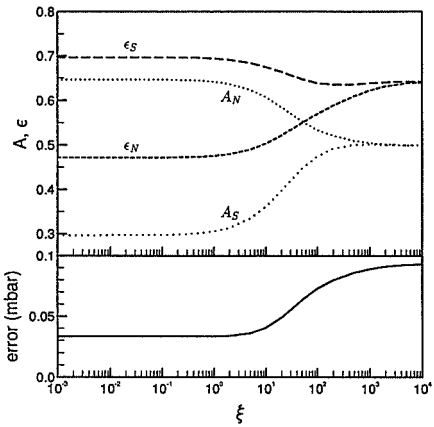


Figure 31. Best fit values of the ice albedo and emissivity computed for various values of the constraint coefficient ξ and the corresponding mean quadratic difference between the synthetic and observed Viking 1 pressure.

is obtained for $\xi = 0$, but the agreement is still satisfactory for $\xi = 23$. The corresponding values of the ice parameters are given in Table 4. The difference $A_N - A_S$ is twice smaller for $\xi = 23$ than for $\xi = 0$. As expected also, it is mainly the first harmonic of the seasonal cycle which is sensitive to the value of ξ.

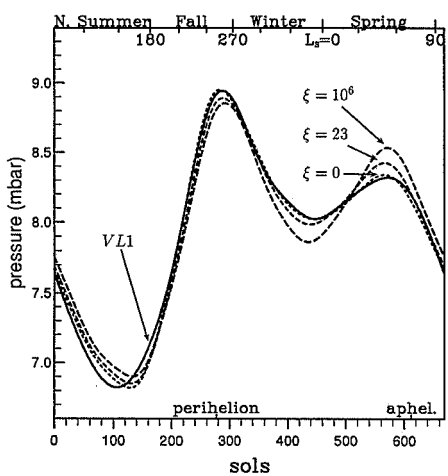


Figure 32. Three synthetic best fits for three different values of the constraint parameter ξ (10^6, 23, and 0). The pressure at Viking 1 is also shown (solid curve). All the curves correspond to eight-harmonic fits.

Table 4. Best Fit Parameters for $\xi = 0$, $\xi = 23$ and $\xi = 10^6$, and Mean Quadratic Error Based on the Low-Resolution Control Simulation

ξ	p_{tot}	A_N	ϵ_N	A_S	ϵ_S	Error, mbar
0	6.58	0.64	0.47	0.40	0.70	0.033
23	6.60	0.58	0.53	0.41	0.66	0.051
10^6	6.67	0.50	0.64	0.50	0.64	0.094

The best fit procedure was validated, a posteriori, by performing a new simulation analogous to the control simulation except that the ice parameters and total atmospheric mass were changed to the best fit values of the $\xi = 23$ case. The mean quadratic difference between the synthetic (deduced from the best fit algorithm) and simulated seasonal pressure evolutions is 0.035 mbar, which is surprisingly good in view of the simplicity of the best fit procedure.

The same approach was applied to the high-resolution simulation. Since the annual mass cycle is very close in the control and high-resolution simulations, we used the sensitivity functions computed from the low-resolution experiments. After some tests, we chose the $\xi = 10$ case ($p_{tot} = 6.152$ mbar, $A_N = 0.57$, $\epsilon_N = 0.53$, $A_S = 0.41$, and $\epsilon_S = 0.62$), which gives a mean quadratic error of 0.042 mbar, to perform a new high resolution simulation. The simulated pressure at Viking sites are presented in Figure 33.

Note that when we do not allow differences between the northern and southern ice properties, the best fit algorithm gives a mean quadratic error of 0.064 mbar only (much lower than for the low resolution) for $p_{tot} = 6.537$ mbar, $A = 0.51$, and $\epsilon = 0.58$.

The best fit procedure was also applied for other dust scenarios, some of which (like the $\tau = 1$ experiment) were closer to the Viking observations in terms of the two first harmonics of the seasonal cycle than for the control experiment. But at the same time, with more aerosols, the atmospheric mass budget is less sensitive to albedo or emissivity changes. As a consequence, all the best fits with high or time-varying dust opacities led to more extreme values of the ice emissivity and albedo.

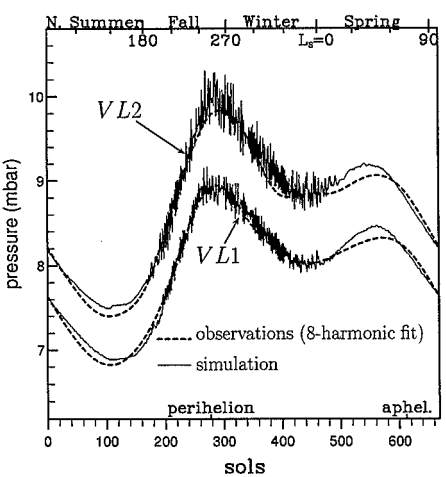


Figure 33. High-resolution best fit simulation ($p_{tot} = 6.152$ mbar, $A_N = 0.57$, $\epsilon_N = 0.53$, $A_S = 0.41$, and $\epsilon_S = 0.62$) of the pressure at the Viking sites and the observations smoothed by retaining the eight first harmonics of the seasonal cycle.

8. Concluding Remarks

Best Fit Parameters and Polar Processes

We have derived an algorithm which provides best fits to the Viking pressure seasonal cycle by varying independently the ice albedos and emissivities of the two polar caps. Thanks to the automatic procedure we use, we obtain values of the mean quadratic difference with Viking observations $\sigma \sim 0.04$–0.05 mbar when north-south asymmetries are allowed and $\sigma \sim 0.1$ mbar when identical values are imposed for the two caps. For comparison, the best fit published by *Pollack et al.* [1993] corresponds to $\sigma \sim 0.13$.

As *Wood and Paige* [1992] and *Pollack et al.* [1993] results, our study confirms the necessity for low values of the polar cap emissivity (of the order of 0.7). It also suggests the necessity for an asymmetry between the southern and the northern polar caps with a lower emissivity and larger albedo in the north. Note that, except for the $\tau = 1$ simulation, our best fit values of the cap albedo lie within the range of observations [see *Pollack et al.*, 1993, Table 4].

As already discussed, for instance, by *Pollack et al.* [1993], the necessity for low emissivity can be interpreted either in terms of real CO_2 ice properties or in terms of physical processes not accounted for in the model. Those low emissivity values are, at first approximation, compatible with the low values of the thermal emission recorded by the IRTM instrument aboard Viking orbiter, which have been recently analyzed extensively by F. Forget et al. (Low brightness temperatures of Martian Polar caps: CO_2 clouds or low emissivity?, manuscript in preparation, 1995).

If not due to a real low emissivity of the icy surface, this low thermal emission could be due for instance to opaque CO_2 clouds colder than the surface [*Pollack et al.*, 1990, 1993]. An alternative or additional explanation could arise from our rather poor knowledge of the polar orography. In fact, on the ice caps, there is a direct link, through the Clausius-Clapeyron law (2) and hydrostatic balance, between surface temperature and altitude variations:

$$\delta z = -\frac{L}{g}\frac{dT}{T} \sim -1.6 \times 10^5 \frac{dT}{T}. \qquad (15)$$

An underestimation by 1.6 km of the elevation of the

surface orography would correspond to an overestimation of the surface temperature in the model by about 1.4 K (the cap temperature being of the order of 140 K), which should be compensated, in order to reproduce good global condensation rates, by a decrease of the surface emissivity by 4%.

Improvements in our understanding of the polar atmospheric energy and mass budget on the polar caps will probably require more sophisticated parametrizations of the surface properties and atmospheric condensation, including the representation of CO_2–ice clouds, a careful comparison of the numerical results with the observations of the thermal emission by the Viking IRTM instruments as well as a better knowledge of the actual orography in the polar regions.

Interannual Variability and Dust

One intriguing result concerning the seasonal pressure variations recorded by the Viking landers is the very weak interannual variability. In particular, the pressure cycle was not significantly affected by the two global dust storms of the first Viking year, the temporary increase of the surface presure at Viking 2 during the second dust storm being mainly due to modifications of the dynamical component [*Pollack et al.*, 1993; paper 1].

For a given set of model parameters, the simulated interannual variability of the seasonal mass cycle is also very weak (see Figure 5) and compatible with observations, but, on the other hand, this cycle depends strongly on the atmospheric dust content, in a complex manner involving modifications of the radiative budget and atmospheric dynamics.

The link with observation is premature at this stage and would require a more realistic representation of the temporal and spatial variations of the atmospheric dust content. This could be achieved either by introducing in the model a climatology of the atmospheric dust content or by modeling directly the atmospheric dust transport in the GCM. It must also be kept in mind that the uncertainties on the dust optical properties can strongly affect both the seasonal pressure cycle and its sensitivity to changes of the atmospheric dust content.

Dynamical Contribution

The decomposition of the surface pressure variations into various components, already proposed in paper 1, has proved to be a powerful diagnostic tool for GCM simulations. The surface pressure appears to be especially sensitive to modifications of the dynamical component arising from modifications of the mean zonal winds following changes of the atmospheric angular momentum budget. This dynamical component is particularly sensitive to surface roughness and atmospheric dust content. Note that landing probes at 30° and 60° latitude in each hemisphere (the latitudes which are the most affected by the dynamical component) plus one at the equator would probably be optimum to constrain the dynamical component and, in turn, some model parameters and the annual mass cycle.

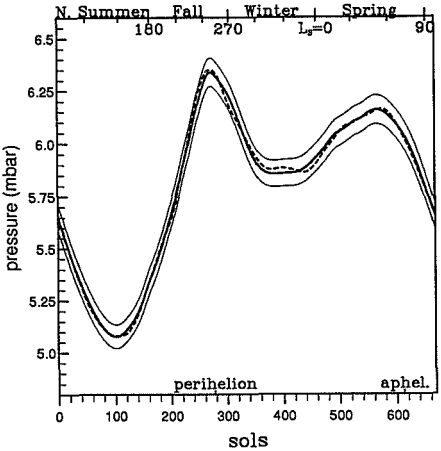


Figure 34. Planetary averaged surface pressure $p_{atm} = p_{VL1} \times \alpha_{VL1}$, where p_{VL1} is the eight-harmonic fit to the seasonal pressure variations observed by Viking 1 and α_{VL1} is taken from the high-resolution sensitivity experiment (dashed curve) and high-resolution best fit simulation (thick solid curve). The two thin solid curves have been computed by adding ±1.5% to the meteorological factor of the high-resolution best fit simulation.

One output of the present study is also the determination of the time evolution of the atmospheric mass from the Viking observations. Figure 34 shows the seasonal evolution of the planetary averaged surface pressure reconstructed from the Viking 1 pressure observations scaled by the meteorological factor α_{VL1} taken from the high-resolution sensitivity experiment (thick dashed curve) and best fit experiment (thick solid curve). The two thin curves, computed by adding ±1.5% to the meteorological factor of the best fit simulation, correspond to the maximum uncertainty estimated from the maximum variations of the meteorological factor in all the sensitivity experiments using a constant $\tau = 0.2$ or a time-varying dust optical depth.

Note also that the agreement of the simulated and observed relative pressure between the two Viking landers can be obtained only for a rather short range of values for the altitude difference between the two sites. The dashed curve in Figure 12 was computed assuming that Viking 2 was 980 m below Viking 1. A change by 50 m of this altitude difference would have shifted $\alpha_{VL2/1}$ by about 0.5%, resulting in a rather bad fit. This gives an estimation of the altitude difference: $\delta z = -980 \pm 50$ m.

Transient Eddies

The amplitude of the transient eddies in the clear-sky high-resolution simulation is in good agreement with Viking observations, except near northern winter sol-

stice, when it is strongly overestimated. The sensitivity experiments suggest that the increase of the dust optical depth during winter is responsible for the reduction of the transient activity in that season. In good agreement with Viking observations, this reduction is found to be stronger near northern winter solstice. Note also that the amplitude of the transient waves is much smaller in the southern hemisphere (not shown), as also found with the NASA/Ames GCM [*Barnes et al.*, 1993].

Implications for Spatial Exploration

We have mentioned above how GCM's can be used for the preparation of spacecraft missions, for instance, for the selection of landing sites which would lead to a better constraint of the atmospheric mass cycle and dynamical component.

It also appears clearly from the present study that one priority of the future Martian missions must be the determination of the Martian orography. An accurate knowledge of the orography is of prime importance both (1) for estimation of the energy budget over the caps, and (2) for fully efficient use of local pressure measurements, such as the Viking measurements.

This study underlines more generally the usefulness of using self-consistent models such as GCM's in conjunction with spacecraft observations. This approach could be made much more systematic by using the "data assimilation techniques" developed for the purpose of operational weather forecasting. For that goal, we are currently developing at LMD the adjoint of our Martian GCM, which we intend to use for four-dimensional variational assimilation [e.g., *Talagrand and Courtier*, 1987; *Thépaut and Courtier*, 1991] of the data of future Martian missions.

Acknowledgments. The computational means were provided by Institut du Développement et des Resources en Informatique Scientifique of Centre National de la Recherche Scientifique. The Viking pressure measurements were produced by *Tillman and Guest* [1987] and *Tillman* [1988] and provided to us by Steven Lee. They are available in digital form through the Planetary Data System. For information on ordering the data set(s) or related documentation, contact the PDS Planetary Atmospheres Discipline Node (303-492-5348; e-mail - Internet: lee@syrtis.colorado.edu, SPAN: ORION::LEE; Steven Lee, Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, CO 80309) or the PDS Central Node (818-306-6026; e-mail - SPAN: JPLPDS::PDS_OPERATOR; PDS Operator, Jet Propulsion Laboratory, MS 525-3610, 4800 Oak Grove Drive, Pasadena, CA 91109). We are also grateful to James Murphy and Robert Haberle for their useful comments on the first manuscript.

References

Barnes, J. R., Time spectral analysis of the midlatitude disturbances in the Martian atmosphere, *J. Atmos. Sci.*, *37*, 2002–2015, 1980.

Barnes, J. R., Midlatitude disturbances in the Martian atmosphere: A second Mars year, *J. Atmos. Sci.*, *38*(2), 225–234, 1981.

Barnes, J. R., J. B. Pollack, R. M. Haberle, R. W. Zurek, C. B. Leovy, H. Lee, and J. Schaeffer, Mars atmospheric dynamics as simulated by the NASA/Ames general circulation model, 2, Transient baroclinic eddies, *J. Geophys. Res.*, *98*(E2), 3125–3148, 1993.

Fouquart, Y., and B. Bonnel, Computations of solar heating of the Earth's atmosphere: A new parametrization, *Contrib. Atmos. Phys.*, *53*, 35–62, 1980.

Haberle, R. M., J. B. Pollack, J. R. Barnes, R. W. Zurek, C. B. Leovy, J. R. Murphy, H. Lee, and J. Schaeffer, Mars atmospheric dynamics as simulated by the NASA/Ames general circulation model, 1, The zonal-mean circulation, *J. Geophys. Res.*, *98*(E2), 3093–3124, 1993.

Hourdin, F., A new representation of the CO_2 15-μm band for a Martian general circulation model, *J. Geophys. Res.*, *97*(E11), 18,319–18,335, 1992.

Hourdin, F., P. Le Van, F. Forget, and O. Talagrand, Meteorological variability and the annual surface pressure cycle on Mars, *J. Atmos. Sci.*, *50*, 3625–3640, 1993.

James, P. B., H. H. Kieffer, and D. A. Paige, The seasonal cycle of carbon dioxide on Mars, in *Mars*, edited by H.H. Kieffer, B.M. Jakosky, C.W. Snyder and M.S. Matthews, pp. 934–968, University of Arizona Press, Tucson, 1992.

Keegan, K. D., J. E. Bachman, and D. A. Paige, Thermal and albedo mapping of the north polar region of Mars (abstract), *Lunar Planet Sci.*, *XXII*, 701–702, 1991.

Leovy, C., and Y. Mintz, Numerical simulation of the atmospheric circulation and climate of Mars, *J. Atmos. Sci.*, *26*, 1167–1190, 1969.

Louis, J.-F., A parametric model of vertical eddy fluxes in the atmosphere., *Boundary Layer Meteorol.*, *17*, 187–202, 1979.

Martin, T. Z., Thermal infrared opacity of the Mars atmosphere, *Icarus*, *66*, 2–21, 1986.

Martin, T. Z., and M. I. Richardson, New dust opacity mapping from Viking infrared thermal mapper data, *J. Geophys. Res.*, *98*(E6), 10,941–10,949, 1993.

Morcrette, J. J., L. Smith, and Y. Fouquart, Pressure and temperature dependence of the absorption in longwave radiation parametrizations, *Contrib. Atmos. Phys.*, *59*(4), 455–469, 1986.

Paige, D. A., and A. P. Ingersoll, Annual heat balance of Martian polar caps: Viking observations, *Science*, *228*, 1160–1168, 1985.

Paige, D. A., and K. D. Keegan, Thermal and albedo mapping of the south polar region of Mars (abstract), *Lunar Planet Sci.*, *XXII*, 1013–1014, 1991.

Paige, D. A., and S. E. Wood, Modeling the Martian seasonal CO_2 cycle, 2, Interannual variability, *Icarus*, *99*, 15–27, 1992.

Pollack, J. B., Properties and effects of Martian atmospheric dust, *Adv. Space. Res.*, *2*, 43–56, 1982.

Pollack, J. B., C. B. Leovy, P. W. Greiman, and Y. Mintz, A Martian general circulation model experiment with large topography, *J. Atmos. Sci.*, *38*, 3–29, 1981.

Pollack, J. B., R. M. Haberle, J. Schaeffer, and H. Lee, Simulations of the general circulation of the Martian atmosphere, 1, Polar processes, *J. Geophys. Res.*, *95*, 1447–1473, 1990.

Pollack, J. B., R. M. Haberle, J. R. Murphy, J. Shaeffer, and H. Lee, Simulation of the general circulation of the Martian atmosphere, 2, Seasonal pressure variations, *J. Geophys. Res.*, *98*(E2), 3149–3181, 1993.

Sadourny, R., and K. Laval, January and July performance of the LMD general circulation model, in *New Perspectives in Climate Modeling*, edited by A.Berger and C.Nicolis, pp. 173–197, Elsevier, Amsterdam, 1984.

Santee, M., and D. Crisp, Thermal structure and dust loading of the Martian atmosphere during late southern sum-

mer: Mariner 9 revisited, *J. Geophys. Res.*, *98*(E2), 3261–3279, 1993.

Sutton, J. L., C. B. Leovy, and J. E. Tillman, Diurnal variations of the Martian surface layer meteorological parameters during the first 45 sols at two Viking lander sites, *J. Atmos. Sci.*, *35*, 2346–2355, 1978.

Talagrand, O., and P. Courtier, Variational assimilation of meteorological observations with the adjoint vorticity equation, 1, Theory, *Q. J. R. Meteorol. Soc.*, *113*, 1331–1328, 1987.

Talagrand, O., F. Hourdin, and F. Forget, The LMD Martian general circulation model: Results about the annual pressure cycle, *Bull. Am. Astron. Soc.*, *23*, 1217, 1991.

Thépaut, J.-N., and P. Courtier, Four-dimensional variational data assimilation using the adjoint of a multilevel primitive-equation model, *Q. J. R. Meteorol. Soc.*, *117*, 1225–1254, 1991.

Tillman, J. E., Mars global atmospheric oscillations: Annually synchronized transient normal-mode oscillations and the triggering of global dust storms, *J. Geophys. Res.*, *93*(D8), 9433–9451, 1988.

Tillman, J. E., and W. Guest, Atmospheric pressure, point by point values: VL 1 sols 1-2245 and VL 2 sols 1-1050, technical report, National Space Science Data Center, Greenbelt, Md., 1987.

Tillman, J. E., N. C. Johnson, P. Guttorp, and D. B. Percival, The Martian annual atmospheric pressure cycle: Years without great dust storms, *J. Geophys. Res.*, *98*(E6), 10,963–10,971, 1993.

Wood, S. E., and D. A. Paige, Modeling the Martian seasonal CO_2 cycle: Fitting the Viking lander pressure curves, *Icarus*, *99*, 1–14, 1992.

F. Forget, F. Hourdin, and O. Talagrand, Laboratoire de Météorologie Dynamique du Centre National de la Recherche Scientifique, Ecole Normale Supérieure, 24 rue Lhomond, 75231 Paris cedex 05, France. (e-mail: hourdin@lmd.ens.fr, forget@lmd.ens.fr, talagran@lmd.ens.fr)

(Received March 14, 1994; revised October 21, 1994; accepted November 28, 1994.)

Page intentionally left blank

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5525–5539, MARCH 25, 1995

Post-Phobos model for the altitude and size distribution of dust in the low Martian atmosphere

Eric Chassefière
Service d'Aéronomie, Centre National de la Recherche Scientifique, Verrières-le-Buisson, France

P. Drossart
Département de Recherche Spatiale, Observatoire de Paris-Meudon, Meudon, France

O. Korablev[1]
Space Research Institute, Moscow, Russia

Abstract. Four experiments flown on board Phobos 2 provided information on the characteristics of the dust particles suspended in the Martian atmosphere : Auguste (UV-visible-IR spectrometer working in solar occultation geometry), ISM (IR spectrometer measuring the light of the Sun reflected by the planet), Termoskan (scanning radiometer mapping the planetary thermal radiation), KRFM (UV-visible multiphotometer providing limb-to-limb profiles). These experiments, which sounded equatorial regions (20°S-20°N) near the northern spring equinox (L_S=0-20°), are shown to yield a reasonably consistent picture of the dust distribution over the whole altitude range from the ground level, or just above, outside the boundary layer, up to ≈25 km. The vertical profiles of particle volume mixing ratio and effective (projected area-weighted) radius deduced from Auguste measurements, performed in the 15-25 km altitude range, are extrapolated down to the ground by using a simple, physical parameterization of the altitude dependence of dust mixing ratio and radius. This parameterization, which must be understood as describing the vertical distribution of dust in the zonal average and on the mesoscale in latitude, assumes that gravitational settling and eddy diffusion are the only two processes driving vertical dust transport. The vertically averaged effective radius and optical depth of dust particles, as well as vertical profiles of related quantities, are obtained. Optical depth at 1.9-μm wavelength is found to be 0.2 on average, with a typical variation of ±0.1 with time and space. This result is similar to that obtained from ISM spectra analysis. It is also consistent with the Termoskan and KRFM measurements, which yield near-infrared optical depths of 0.12-0.26 and 0.12-0.24, respectively. The particle number density near the surface, as derived from extrapolation of solar occultation profiles, is in the range 1-3 cm-3, in good agreement with Termoskan results (1-2 cm-3). The scale height of the dust volume mixing ratio just above the surface is ≈8-9 km on average, that is, of the same order as the background atmospheric scale height. The vertically averaged effective radius of dust particles is found to lie in the range 1.7±0.2 μm, possibly ≈2 μm in the case of a large effective variance of 0.4. The most likely ISM value is 1.2 μm, with a rather large uncertainty of ±0.4 μm, mainly due to the fact that the spectral dependence of the Minnaert coefficient is not well known. Because ISM data used in the present work were obtained on the Tharsis plateau, at a mean altitude of ≈7 km, the ISM radius must be compared to the Auguste vertically averaged radius for z>7 km, that is, ≈1.5±0.2 μm. Auguste and ISM radii are therefore consistent at the 1-σ level. Three typical vertical profiles of the dust particle radius and number density, obtained by averaging all solar occultation profiles, including their extrapolated parts below ≈15 km, are proposed as reference models, for three selected values of the effective variance of the particle size distribution (0.10, 0.25, and 0.40).

Introduction

It has been known since the Mariner 9 and Viking missions that a persistent dust opacity, varying from a few tenths to nearly 5 during global dust storms at visible wavelengths, characterizes the Martian atmosphere. The dust cycle is expected to play an important role in determining the climate of Mars through, notably, its influence on the atmospheric thermal structure, the ground albedo, and the interhemispheric transport of water and carbon dioxide ices. It was recently pointed out by *Murphy et al.* [1993] that the temporal evolution of the dust opacity during global dust storms cannot be well reproduced by numerical simulations. One of the most critical aspects is that the size distribution of particles suspended in the atmosphere is not preserved over space and time, with a rapid depletion of large particles through gravitational settling. The question of the size distribution of the suspended dust particles, including its

[1]Also at Service d'Aéronomie du CNRS, Verrières-le-Buisson, France.

Paper number 94JE03363.
0148-0227/95/94JE-03363$05.00

dependence with altitude, is therefore of crucial importance for a better understanding of the Martian dust cycle. First results on the composition and size of the dust particles were obtained from Mariner 9 UV and IR spectrometer data [*Conrath*, 1975; *Pang and Ajello*, 1977; *Toon et al*, 1977]. Information was also obtained with TV experiments [*Anderson and Leovy*, 1978]. Later observations of aerosols from Viking Orbiter and Lander cameras, as discussed respectively by *Clancy and Lee* [1991] and *Pollack et al.* [1979], allowed the picture previously obtained to be completed and refined. The effective (projected area-weighted) radius of the dust particles, as seen from the Martian surface, is ≈2.5 μm, although a dust particle size 5-10 times smaller cannot be ruled out [*Clancy and Lee*, 1991]. Within the large measurement uncertainties and the interdependent sensitivities on composition, particle size, and optical depth, the measurements suggest that the optical parameters of the particles (single scattering albedo, asymmetry factor) do not vary with place and time. The size distribution of dust is constant at subsolar latitudes. A compilation of all results obtained from spacecraft observations up to the Viking mission is given by *Murphy et al.* [1993].

More recently, the Martian aerosols were observed by different instruments onboard the Phobos 2 spacecraft. The vertical profiles of the mixing ratio and size of the dust particles in the 15-25 km altitude range were deduced from solar occultation measurements performed by the Auguste instrument [*Korablev et al.*, 1993; *Chassefière et al.*, 1992]. These measurements were made in a narrow band of latitude in equatorial regions (0-20°N) near northern spring equinox (L_S=0-20°). The effective radius of particles was found to vary from ≈0.8 μm at 25 km up to ≈1.6 μm at 15 km. Their number density is ≈0.3 cm^{-3} in the same altitude range. The effective variance of the distribution is ≈0.25. The vertical optical depth of dust at 1.9-μm wavelength is ≈0.2, that is, about the minimum value observed from Viking landers, suggesting a somewhat clearer year even at this season of minimum dust loading. A small effective radius of ≈1.25 μm was inferred from the analysis of the dust component observed by the ISM infrared spectrometer [*Drossart et al.*, 1991]. Spectrophotometry of Mars by the KRFM instrument yielded a small value of the dust optical depth, ≈0.1-0.2 [*Moroz et al.*, 1993]. Finally, the thermal infrared emission of the limb of Mars, as measured by the scanning radiometer Termoskan, suggested a small visible optical depth of ≈0.13, with a number density of dust particles at the ground level of 1-2 cm^{-3} [*Moroz et al.*, 1995]. All pre-cited Phobos measurements were made in equatorial regions (20°S-20°N).

The aim of this paper is to reconcile these various estimates and thereby obtain a consistent view of Phobos results. Whereas solar occultation measurements provided for the first time the vertical structure of the dust in the 15-25 km altitude range, infrared imaging allowed the dust in the low atmosphere to be characterized, typically for the first atmospheric scale height above the ground. It is tempting to merge these two kinds of measurements in order to infer the complete profile of the dust from the ground up to ≈25 km altitude. Additional information provided by other instruments may be used to check the global consistency of the results. The rather marked similarity between the nine vertical profiles of dust extinction coefficient obtained by solar occultation at various longitudes shows that atmospheric conditions do not change much in space and time at these low latitudes during this early northern spring period. This result is consistent with the pre-Phobos picture, although the dust amount indicated by Phobos observations is quite low, as compared to the higher opacities observed by Viking during the same Mars season. It suggests, together with the small optical depth of dust, a typical, stable equatorial atmosphere. Other arguments in this sense are developed by *Rosenqvist and Chassefière* [this issue] from a more general analysis of Phobos results. In this context, it seems interesting, and useful for modelers of the dust cycle, to examine the different results obtained by the Phobos mission with the main objective to propose a single synthetic vertical profile of the dust (size, number density). Solar occultation profiles are extrapolated from the 15-25 km altitude range down to the ground, that is, the zero reference altitude (altitude of the reference ellipsoid), by using a specific parametrization of the altitude dependence of the volume of particles per unit atmospheric volume, which will be called volume mixing ratio in the following, and the particle effective (projected-area weighted) radius. Vertically averaged quantities obtained in this way are compared to those observed by the infared spectrometer and other instruments in order to check the validity of the extrapolation procedure. The parameterization used for extrapolation is based on the simplified steady state model used by *Chassefière et al.* [1992]. Gravitational settling is assumed to be balanced by eddy diffusion, parameterized by an empirical eddy diffusion coefficient K. The validity of the steady state model will be reassessed in the concluding section. In particular, the fact that no marked variation of the dust column density with altitude is measured on the flank of the volcanoes by ISM suggests additional mechanisms at low atmospheric levels [*Drossart et al.*, 1991].

The Data

The main characteritics of the instrument are listed in Table 1.

Table 1. Instrument Characteristics

Instrument	Wavelength Range, μm	Wavelengths Used for Dust Retrieval, μm	Observation Mode	Altitude Range
Auguste	0.21-0.33 μm 0.76/0.94, 1.9/3.7	1.9, 3.7 μm	solar occultation	15-25 km at limb
ISM	0.7-3.2	1-2	surface mapping	column
KRFM	0.3-0.6	1-2	limb-to-limb scan	column
Termoskan	7-13	1-2	limb scan	column scale height

Solar Occultation Data

The infrared spectrometer on board the Phobos spacecraft has been designed for solar occultation measurements of the Martian atmosphere. It is a part of a more complete instrument analyzing the spectrum of the solar radiation in different wavelength ranges. The principle of the Auguste experiment is to measure from the Phobos orbit the spectrum of the Sun modified by atmospheric extinction during sunset. The infrared spectrometer provides two spectra of 20 and 16 elements in the ranges 5292-5372 cm^{-1} (≈1.9 μm) and 2707-2740 cm^{-1} (≈3.7 μm) with a resolution of 1200. A detailed description of the instrument is given by *Blamont et al.* [1989] and *Krasnopolsky et al.* [1989]. Because of the large apparent diameter of the Sun projected on the occultation plane (≈60 km in circular orbit), it is necessary to select a small part of the Sun. The IR field has an angular diameter of 1.5 arc min, or ≈4 km vertical resolution, and is located near the edge of the solar disk.

The instrument operated from February 8 to March 26, 1989 (the Martian equinox occurred on February 17). Results reported here were made when the spacecraft was in a circular orbit (planetocentric radius, 9800 km; period, 8.1 hours). The data set examined in this paper consists of nine successful IR occultations obtained during ingress at sunset between February 20 and March 26. The two IR signals are averaged over the spectral range of each channel. These occultations, which provide information on the vertical profile of the dust volume mixing ratio and particle radius, are described by *Korablev et al.* [1993]. The latitude of the intersection of Sun-spacecraft axis with the surface of Mars varied from -3° on February 20 to + 20° on March 26. The distance between the spacecraft and the limb is 10^4 km. Each session is characterized by seven numbers indicating the time period from the launch (day, hour, minutes). The characteristic numbers, dates, and areocentric solar longitudes for the nine available occultations are given in Table 2.

Because of the large tangential optical depth of dust particles below 15 km, the lowest altitude sounded by the instrument during the nine IR occultations is in the range 10-15 km. The signal obtained below 15 km is not reliable due to several instrumental problems, which were described by *Blamont et al.* [1991]. The altitude of the line of sight was deduced from the combined analyses of the pointing sequence and of a weak CO_2 absorption band [*Chassefière et al.*, 1992], and a specific treatment has been applied to the data in order to take into account instrumental problems and the solar limb darkening effect when the view axis moves over the solar disk [*Korablev et al.*, 1993]. The altitude ranges sounded by the instrument for the nine occultations, together with corresponding uncertainties, are given in Table 2. The Δz is the uncertainty on the line-of-sight altitude resulting from uncertainties on CO_2 absorption and atmospheric profile. The coordinates (longitude, latitude) of the tangent point are shown in Figure 1.

Infrared Spectrometer Data

The observations taken by the infrared spectrometer ISM, on board the Phobos 2 spacecraft, are described by *Bibring et al.* [1989] and *Combes et al.* [1991]. This spectrometer is operating close to opposition, with phase angles typically ranging from 3° to 10°. Therefore, the contribution of dust opacity is seen in a different viewing geometry as compared to the Auguste experiment. However, the main difficulty in these observations is to extract the dust component from the contribution of the surface. A procedure has been developed by *Drossart et al.* [1991] to extract the backscattered dust contribution from an average reflectance spectrum, obtained in the region of Pavonis Mons (outside the high-altitude ranges close to the volcano), recorded on March 14, 1989, with a phase angle of 6°. In this paper, we use the dust spectrum retrieved in Figure 4 of this Drossart et al., rescaled by using an improved calibration procedure for ISM (S. Erard, personal communication, 1994). This spectrum is recovered under the assumption of small optical depth for suspended dust, and uniform surface properties along the sequence of observations. Both hypotheses are consistently verified along the retrieval procedure. The reflectance of the ISM spectra can be shown to verify

$$R = a_s \mu_0^K \mu^{K-1} + \frac{\omega_0 \tau_0}{4\mu} H(\pi-\alpha) \qquad (1)$$

where H is the vertically averaged phase function of the aerosols, ω_0 the vertically averaged single scattering

Table 2. Times and Altitude Ranges of the Nine Solar Occultation Measurements of the Dust (All for Sunset Observations)

Occultation	Date in 1989	Occultation Time, UT	L_S, deg	z_{max}, km	z_{min}, km	Δz, km
2212158	Feb. 2	1636:57	1.9	24	15	+3, -2
2360639	March 7	0115:28	8.9	26	15	+4, -2
2362241	March 7	1717:14	9.2	25	20	+6, -3
2421504	March 13	0940:01	12.0	27	16	+2, -2
2440712	March 15	0148:22	12.8	26	15	+2, -3
2441514	March 15	0949:58	12.9	24	13	+2,-2
2450717	March 16	0153:19	13.3	23	16	+2,-2
2510743	March 22	0219:08	16.1	28	17	+4, -2
2560312	March 26	2147:43	18.4	24	17	+2,-3

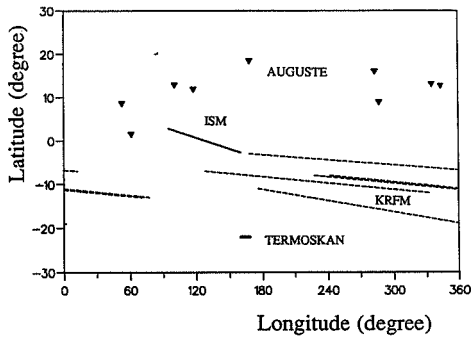


Figure 1. Coordinates of regions observed by Auguste (triangles), ISM (thin solid segment), Termoskan (thick solid segment), and KRFM (dashed segments).

albedo, τ_0 the integrated optical depth, and a_S the surface albedo. The main uncertainty in the data reduction comes from the estimate of the photometric properties of the surface described by the Minnaert coefficient K; it has been estimated that the Minnaert coefficient for the surface alone lies between 0.7 and 0.9, which gives the error bars in the retrieved dust spectrum. Therefore a dust spectrum is obtained between 1 and 2 μm (for calibration purposes, and to avoid atmospheric gaseous absorptions, the full ISM spectrum from 0.7 to 3.2 μm had to be limited), whose main characteristics is a decreasing slope. As will be shown in the following, this property can be related to the aerosol characteristics, and can be checked against the measurements of other Phobos instruments recorded at the same period.

Other Experiments

A series of five limb-to-limb photometric profiles was obtained by the KRFM instrument at low latitudes in the southern hemisphere [*Moroz et al.*, 1993]. The traces of the scanning line of sight are shown in Figure 1 for each one of the five observations. All profiles were obtained in eight selected spectral bands in the 315-550 nm spectral range. The spectral and spatial resolutions of the spectrometer are ≈10 ($\lambda/\Delta\lambda$) and ≈30 km, respectively. Thanks to an appropiate modeling of the brightness of the Martian disk, it was possible to determine several important characteristics of dust and cloud particles. In particular, the visible optical depth of the dust component on the morning limb in equatorial regions was found to lie in the range 0.1-0.2 at the time of the measurements (L_S=5°-16°). Because opacities derived from solar occultation measurements are obtained at 1.9-μm wavelength, it is necessary to scale the KRFM value. By using the optical index given in the appendix, one can roughly estimate that a multiplicative factor of 1.2 must be applied to transform the visible opacity into an infrared opacity [e.g., *Hansen and Travis*, 1974]. The resulting KRFM estimate at 1.9 μm is ≈0.12-0.24.

The thermal radiation of Mars in the 7-13 μm spectral range has been measured by the Termoskan instrument [*Moroz et al.*, 1994]. In particular, information was obtained near the limb of the planet where the atmosphere contributes significantly to the planetary thermal radiation. Termoskan is a scanning radiometer able to measure the spatial distribution of brightness with a high spatial resolution (≈2 km). Eight adjacent, regularly spaced (≈20 km intervals) cross-limb profiles of brightness were extracted from one image obtained on March 26, 1989, at ≈20° S latitude (Figure 1). A visible optical depth of 0.13 (+0.09, -0.03) was inferred for the dust component. The corresponding infrared value (at 1.9 μm), calculated in the way already described, is 0.16 (+0.10, -0.04). The dust scale height was estimated to be of the order of the gaseous atmosphere scale height, and the number density of dust particles near the surface was found to be in the range 1-2 cm^{-3}.

Extrapolation of Solar Occultation Profiles

Treatment of Auguste Data

The corrections applied by *Korablev et al.* [1993] to the Auguste data take into account the nonlinearity of the detector and the change of the signal due to both the solar limb darkening and the partial loss of the solar disk when the pointing is disturbed. These corrections generally improve the aspect of the vertical profiles of the extinction coefficients $\Sigma^b(z)$ and $\Sigma^r(z)$ at 1.9 μm and 3.7 μm, respectively. Nevertheless, they induce small oscillations of the ratio $\Sigma^b(z)/\Sigma^r(z)$, and therefore of the inferred radius of the particles. Because these oscillations occur on a vertical scale smaller than the projected size of the field of view (≈4 km), they may be unambiguously attributed to an instrumental bias in the corrective algorithms. As shown hereafter, the vertical profile of the radius $r(z)$ is extrapolated from ≈15 km down to the ground by parameterizing the vertical variation of its "scale height" $h(z)=[d \log (r)/dz]^{-1}$, and even small oscillations of $r(z)$ imply well-marked variations of $h(z)$. It is emphasized that the radius obtained by comparing the uncorrected extinction coefficients is near the radius obtained by using the corrected values but does not present the small oscillations resulting from the instrumental bias. On the other hand, the accuracy of the

extinction coefficient profiles is improved when using Korablev's corrections. For this reason, these corrections are used only for retrieving the extinction coefficient profiles, and therefore the volume mixing ratio of particles, but not for the ratio. Corrected and uncorrected profiles at 1.9 μm and 3.7 μm are shown in Figure 2 for occultation 2212158 (February 20). Although the general shape of the profiles is not changed by instrumental corrections, the irregularities of the uncorrected profiles, which cannot be from atmospheric origin due to the low vertical resolution (≈4 km), are removed by the corrective algorithm.

The Method

The objective of the present section is to explain how the vertical profiles of particle radius and dust volume mixing ratio are extrapolated down to the ground in order to derive the ranges of the possible values of the effective radius and the optical depth, as seen in backscattering viewing geometry by the ISM instrument. As previously explained, nine vertical profiles of the extinction coefficients $\Sigma^b(z)$ and $\Sigma^r(z)$ at 1.9 and 3.7 μm are obtained in an average altitude range of 15-25 km (see Table 2). Let us denote by a and b the effective radius and effective variance of the dust size distribution. The shape of the size distribution of particles is assumed to be [*Hansen and Travis*, 1974]

$$\frac{dn}{dr} = k\, r^{\frac{1-3b}{b}}\, e^{-\frac{r}{ab}} \qquad (2)$$

The ratios $\Sigma^b(z)/\Sigma^r(z)$ and σ^b/σ^r, where σ^b and σ^r are the extinction cross sections calculated using the same optical index as *Pollack et al.* [1979], which depend on a, b, and a roughness parameter ρ (see appendix), are denoted by $F(z)$ and $f(a,b,\rho)$, respectively. For each couple (b,ρ), the value $a_{b,\rho}(z)$ of the effective radius such as $f(a,b,\rho) = F(z)$ may be derived at any altitude z. The number density of the particles may be deduced:

$$n_{b,\rho}(z) = \frac{\Sigma^b(z)}{\sigma^b[a_{b,\rho}(z)]} \qquad (3)$$

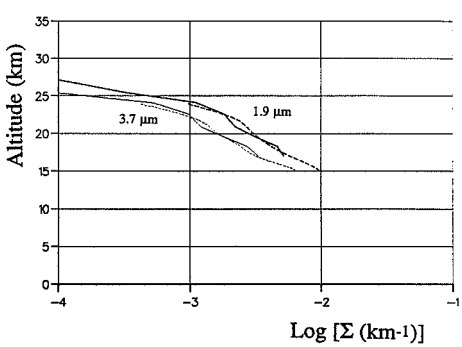


Figure 2. The extinction coefficients at 1.9 μm (thick lines) and 3.7 μm (thin lines) are plotted as a function of the altitude for occultation 2212128. The profiles obtained by applying the instrumental corrections of *Korablev et al.*[1993] are indicated with dashed lines.

Let us denote by $v_{b,\rho}(a)$ the average volume of one particle, obtained by averaging $4/3\pi r^3$ over the size distribution given by (2). The volume mixing ratio of dust may be deduced:

$$V(z)=n(z)v(a(z)) \qquad (4)$$

The subscripts b,ρ are omitted to simplify notations. The vertical profiles of $V(z)$ and $a(z)$ derived from observations may be extrapolated down to the ground. As shown hereafter, a simple linear extrapolation would be unphysical. Realistic profiles, derived from basic physical laws governing the motions of particles, will be used. In this way, the vertical optical depth τ at 1.9 μm and the vertically averaged effective radius A may be calculated:

$$\tau = \int_0^\infty n(z)\, \sigma^b(z)\, dz \qquad (5)$$

and

$$A = \frac{\int_0^\infty n(z)\, s(z)\, a(z)\, dz}{\int_0^\infty n(z)\, s(z)\, dz} \qquad (6)$$

where $s(z)$ is the particle projected area, πr^2, averaged over the distribution given by (2). Practically, the integration may be stopped at the highest altitude where some extinction is detected, that is, $z_{max} \approx 25$ km (see Table 2). At this altitude, the extinction coefficient is lower than 10^{-3} km^{-1} and the residual optical depth ($z>z_{max}$) is not in excess of 10^{-3} x 10, that is, 0.01, using an atmospheric scale height as the characteristic vertical scale. This value is 1 order of magnitude below the uncertainty which will be finally derived. In the same way, the contribution of dust particles located above z_{max} to the vertically averaged effective radius (6) is weak. The values of τ and A given by (5) and (6) depend on the assumed values of the effective variance b and roughness factor ρ, and the associated derived value of the effective radius a.

Extrapolation of the Dust Volume Mixing Ratio

The scale height of the volume mixing ratio $V(z)$ of dust particles may be written [see *Chassefière et al.*, 1992]

$$H' \approx \frac{H}{1+V_{set}/V_{edd}} \qquad (7)$$

where H is the scale height of the gaseous CO_2 atmosphere, V_{set} the settling velocity of dust particles, and V_{edd} their eddy diffusion velocity defined as the ratio K/H, where K is the eddy diffusion coefficient. Values of K found in this way are of the order of $\approx 10^6$ cm^2 s^{-1} and physical implications of this result have been extensively discussed in previous papers [*Korablev et al.*, 1993; *Chassefière et al.*, 1992]. Equation (7) is a direct consequence of the assumed balance between the eddy diffusion and settling fluxes of the dust. V_{set} varies as the product of the particle radius by the inverse of the atmospheric density in relevant conditions.

By neglecting in first approximation the vertical variation of the radius, which is slow compared to the exponential

variation of the atmospheric density, the settling velocity V_{set} may be assumed to vary as the inverse of the atmospheric density. Equation (7) may be written -

$$H'(z) \approx \frac{H}{1 + c\ \exp(z/H)} \qquad (8)$$

where c is assumed to be constant with altitude. In this simplified parameterization, it is assumed that the vertical variation of the ratio V_{set}/V_{edd} is dominated by the exponential component of the settling velocity and any rapid variation of K over the first 25 km above the ground is ruled out. This assumption is a priori justified by the fact that K seems to remain about constant between 15 km and 65 km [*Rosenqvist and Chassefière*, this issue] as derived from solar occultation data obtained at sunset. In addition, the average error bar on the total volume mixing ratio of dust deduced from the present study is about a factor ≈2 on either side: to introduce a reasonable variation of K between the ground and 20 km altitude (typically a factor ≈2) would not change results by a factor larger than the error bar. It must be noted that the values of the ratio V_{set}/V_{edd} and scale height H' at ≈20 km are 1.5-2 and ≈4 km, respectively, which implies H≈10-12 km from (7), that is, a quite reasonable value. At the ground level, H' becomes of the order of 10 km, and the dust is uniformly mixed. A typical atmospheric scale height H of 10 km is assumed for the calculations. The only free parameter considered in the present study is c.

The principle of the extrapolation is to fit the parametrized profile (8) to the observed profile of the volume mixing ratio scale height in the region where this profile is observed (≈15-25 km). To this end, the observed $H'(z)$ profile is averaged over the first 4 km above the lowest sounded altitude z_{min} ($\langle H' \rangle$) and its standard deviation $\sigma_{H'}$ around $\langle H' \rangle$ calculated at the corresponding average altitude $\langle z \rangle$ ($\approx z_{min}$ + 2 km). The choice of a recovering zone of 4-km thickness is motivated by two independent arguments. First, the vertical resolution of the IR spectrometer is of the order of 4-km, and no major information is lost when averaging over 4-km altitude. Second, a slight change of the slope of the dust volume mixing ratio profile above ≈20 km altitude was noted by *Korablev et al.* [1993], and it seems better to remove points above ≈20 km altitude to connect theoretical and observed profiles. By adjusting the parameterized value of the dust scale height (8) to the mean ($\langle H' \rangle$) and extreme ($\langle H' \rangle \pm \sigma_{H'}$) values at altitude $\langle z \rangle$, it is possible to define a mean and two extreme values of c at the 1-σ level, therefore of the $H'(z)$ profile. The extrapolation of $V(z)$ below z_{min} is performed by using the different possible profiles of $H'(z)$. The limits of the possible extrapolated profiles of $H'(z)$ and $V(z)$ are shown in Figure 3 for occultation 2212158 and 2360639 (February 20 and March 7 respectively). At the 1-σ level, the uncertainty on V at the ground level is around a factor of ≈2. The dust scale height H' averaged over the first atmospheric scale height above the ground is given in Table 3 for the nine available occultations. It is in the range 7.5-9.5 km.

Extrapolation of the Dust Effective Radius

The following differential equation has been obtained by *Chassefière et al.* [1992] for the effective radius a:

$$H \frac{d \ln (a)}{dz} \approx - b \frac{V_{set}}{V_{edd}} \qquad (9)$$

where b is the effective variance of the dust size distribution. It has been analytically shown that b is expected to remain constant with altitude [*Chassefière et al.*, 1992]. These results have been obtained by solving the balance equation (eddy diffusion/settling) for each radius range independently from each other, which is justified by the fact that particles are coupled to the fluid atmosphere over characteristic times much smaller than the coagulation time, and by deriving the dependence of the size distribution with altitude. The scale

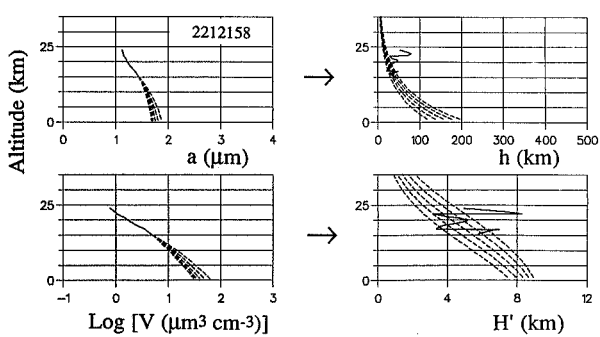


Figure 3a. The vertical profile of the effective radius scale height is shown in the right upper corner for occultation 2212158. The five extrapolated profiles (dashed lines) are obtained by adjusting the parameterized profile to the observed scale height $\langle H' \rangle$, and to scale heights shifted by one half ($\langle H' \rangle \pm \sigma H'/2$) and one full ($\langle H' \rangle \pm \sigma H'$) standard deviation. The vertical profile of the effective radius, including extrapolated profiles, is shown in the left upper corner. The corresponding observed and extrapolated profiles for the dust volume mixing ratio and its scale height are plotted respectively in the left and right lower corners of the figure, respectively.

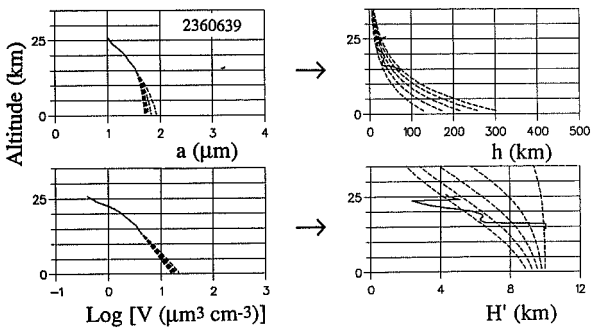


Figure 3b. Same as Figure 3a, but for occultation 2360639.

height h of the effective radius $a(z)$, defined as the height over which $a(z)$ decreases by a factor exp(1), is therefore

$$h \approx \frac{H}{b}\frac{V_{edd}}{V_{set}} \quad (10)$$

and varies nearly as the atmospheric density:

$$h \approx h_0 \exp(-z/H) \quad (11)$$

As for deriving the vertical profile from the dust volume mixing ratio, and for the same reasons, K is assumed to not vary significantly between the ground and 25-km altitude. As shown hereafter, the vertical variation of $h(z)$ over the altitude range where dust is observed agrees approximately with the law given by (11) in seven cases out of nine, with a definite decrease of $h(z)$ between 15-km and 25-km. This variation must be taken into account when extrapolating $h(z)$ down to the ground. The only free parameter is the radius scale height h_0 at the ground level. As in the case for the dust volume mixing ratio, the average value of $h(z)$ over a thickness of 4 km above z_{min} and the associated standard deviation are calculated, allowing a mean and two extreme values of h_0 at the 1-σ level to be defined. The effective radius a is extrapolated by using the different possible extrapolated profiles of h. Occultation 2212158 (Figure 3a) is one of the two previously mentioned occultations for which the scale height of the radius above 20-km altitude does not follow well the parameterized law. The case of occultation 2360639 (Figure 3b), for which the scale height of the radius remains inside the envelope over the whole observed altitude range, is more representative of the general case. A direct examination of the radius profile in Figure 3b shows a well-marked curvature between 15-km and 25-km altitude, and a linear extrapolation would be clearly less realistic than our parameterized approach.

Treatment of ISM Data

To model the ISM spectra, the backscattering component of the dust has been simulated, in the single scattering approximation. According to (1), the backscattered dust component is completely defined from the vertically averaged ω_0, H and from the total optical depth τ_0. These quantities are calculated using a Mie calculation, modified to take into account the scattering by irregular particles [*Drossart*, 1990], with the introduction of a roughness factor which takes into account the irregularities of the scatterers. The parameters of the model are therefore a size distribution (we use the same form as (2) used for the solar occultation data), with an effective radius a and a variance b. The optical constants are taken from *Pollack et al.* [1973], interpolated to each wavelength. In the calculations, a roughness factor of 0.067 was applied. It can be shown that this factor influences mostly the short-wavelength part of the spectrum, and is not predominent in the 1-2 μm range. The simulation is also not very sensitive to the variance b, which has been taken equal to 0.25 in the calculations, according to the solar occultation data analysis. Given an effective radius, the optical depth τ_0 is measured from the comparison with ISM dust spectrum (Figure 4). It can be seen that the largest radius corresponds to the flatter dust spectrum. We conclude from this analysis that an effective radius between 0.8 and 1.6 μm provides a good fit to the dust spectrum, with larger radii giving too flat spectra, and smaller radii too small backscattering at 2 μm. An even broader range of radius on the small side is consistent with error bars due to the surface Minnaert coefficient. Large radii give a poorer fit, and the size distribution of *Toon et al.* [1977] cannot account for the dust spectral shape between 1 and 2 μm. Nevertheless, it must be emphasized that we cannot exclude from ISM spectra bimodal distributions, with a large particle component ($a > 2$ μm), giving a flat dust backscattering component in addition to the decreasing slope due to the scattering by smaller particles.

Results

The possible values of the vertically averaged effective radius A (6) and optical depth τ at 1.9 μm (5) are shown in Figure 5 for the nine occultations. The effective variance b is supposed to be 0.25, which is the most likely value derived from previous analyses (b=0.25±0.15 [*Chassefière et al.*, 1992]; b=0.2±0.1 [*Korablev et al.*, 1993]), and the roughness factor ρ is 0.067. Due to the uncertainty on the observed dust volume mixing ratio and effective radius in the altitude range where the theoretical profiles are constrained

Table 3. Characteristics of the Dust as Inferred From Phobos 2 Measurements

Instrument	L_S, deg	Longitude, deg	Latitude, deg	τ_{IR}	H', km	n, cm^{-3}	r_{eff}, μm
Auguste	1.9	60.9	1.6	0.44	7.5	3.3	1.6
				(+0.20,-0.09)	(±1)	(x2.7,:1.9)	(±0.1)
	8.9	52.6	8.7	0.21	9	1.4	1.7
				(+0.09,-0.03)	(±1)	(x2.6,:1.8)	(±0.1)
	9.2	286.6	9.0	0.60	8	2.4	2.3
				(+0.90,-0.22)	(±1.5)	(x1.8,:1.3)	(±0.4)
	12.0	117.4	12.0	0.23	9	1.9	1.5
				(+0.06,-0.04)	(±1. 5)	(x2.9,:2.1)	(±0.1)
	12.8	343.3	12.8	0.22	9	1.7	1.6
				(+0.06,-0.05)	(±1)	(x2.8,:2.0)	(±0.1)
	12.9	100.5	13.0	0.18	9	0.9	1.8
				(+0.04,-0.03)	(±1)	(x2.3,:1.6)	(±0.2)
	13.3	334.8	13.2	0.22	8.5	1.9	1.5
				(+0.08,-0.05)	(±1.5)	(x2.9,:2.1)	(±0.1)
	16.1	282.9	16.1	0.16	9.5	0.9	1.7
				(+0.06,-0.02)	(±0.5)	(x2.4,:1.7)	(±0.1)
	18.4	168.3	18.5	0.14	8.5	1.7	1.9
				(+0.06,-0.04)	(±1.5)	(x2.6,:1.7)	(±0.2)
ISM	12.4	95/160	2.8/-2.7	0.2±0.1	-	-	1.2
							(±0.4)
Termoskan	18.0	168/160	-22	0.16	≈11	≈1-2	-
				(+0.10,-0.04)			
KRFM	8.4	168/12	-3/-7				
	11.3	128/-28	-7/-12				
	11.8	242/77	-8/-13	0.12-0.24	-	-	-
	12.3	228/70	-8/-13				
	15.6	176/2	-11/-19				

by the observed profiles, there is not a single value of the couple (A, τ) but an extended range which is represented by a box (solid line). In order to include the uncertainty on the altitude (see Table 2), the same calculations are made for the two extreme altitudes at 1 σ, and the corresponding boxes are indicated with dashed lines. Except in the case of occultation 2362241, for which the estimates of optical depth and radius are inaccurate, there is a rather good agreement among the different observations. The large uncertainties for occultation 2362241 are linked to the fact that the observation is restricted to a small altitude range of 20-25 km, due to instrumental problems (see Table 2). As previously mentioned, the scale height of the dust volume mixing ratio tends to decrease above 20 km and, due to the fact that the parameterized profile is constrained by the observed profile in the range 20-25 km instead of 15-20 km in the other cases, the optical depth and radius in the lower atmosphere could be overestimated in this particular case.

Uncertainties resulting, on the one hand, from the extrapolation process and, on the other hand, from the not well-known altitude are combined (by adding variances, not standard deviations) for these results in Figure 6, with error bars for both IR optical depth τ and effective radius r_{eff}. Corresponding values of τ and r_{eff} are given in Table 3. No correlation between optical depth and radius is exhibited. In addition, these two quantities are not correlated to the position (longitude, latitude, see Table 3) of the observed region on Mars. Apart from occultations 2212158 and 2362241, the optical depth at 1.9 μm is concentrated around a modal value of 0.2 and the effective radius of particles is in the range 1.5-1.8 μm. Histograms of optical depth and effective radius are also shown in Figure 6. The Gaussian

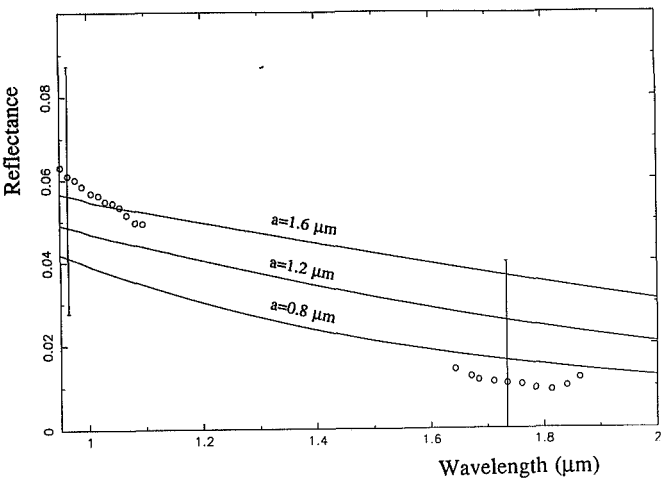


Figure 4. Atmospheric reflectances retrieved from ISM observations (circles) on Pavonis Mons (March 14, 1989) compared with simulated radiances according to the procedure described in the text. Error bars correspond to the uncertainties in the (unknown) Minnaert surface coefficient. The solid lines correspond to three calculations with an optical depth of 0.2, with a= 0.8 (bottom), 1.2 (middle) and 1.6 μm (top).

contributions of the nine occultations are cumulated in unit ranges of 0.2 μm for the radius and 0.1 for the optical depth. The contribution of occultation 2362241 is spread over wide intervals of radius and optical depth due to the large uncertainties. Well-marked maxima are obtained around 1.7 μm (±0.3 μm) for the radius and 0.2 (±0.1) for the optical depth. The peak of the effective radius is narrower for a small effective variance (b=0.10, Figure 7) and wider for a large effective variance (b=0.40, Figure 7), but the range is not changed. The assumption of spherical particles (ρ=0) does not change the picture, but an increase in ρ (ρ=0.2) yields a small decrease of the modal effective radius (Figure 7). The nine vertical profiles of dust particle effective radius, extinction coefficient, volume mixing ratio, and number density have been averaged, and corresponding mean profiles are shown in Figures 8, 9, 10 and 11. Both average instrumental and geographical error bars at the 1-σ level are represented by couples of extreme profiles. The instrumental

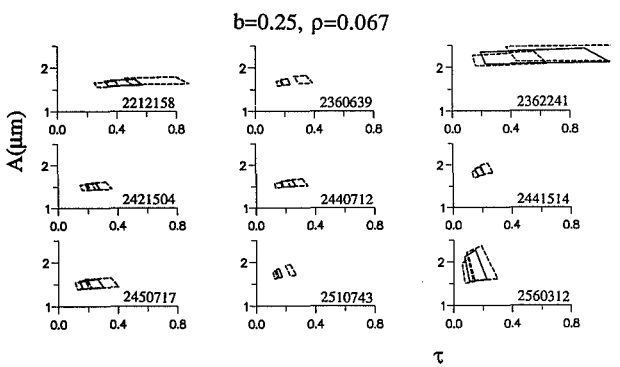


Figure 5. Domains of the possible values of the vertically averaged effective radius A and vertical optical depth τ at 1.9 μm for the nine IR occultations. The three domains indicated for each occultation correspond to the most likely (solid line) and extreme (dashed lines) values of the lowest sounded altitude.

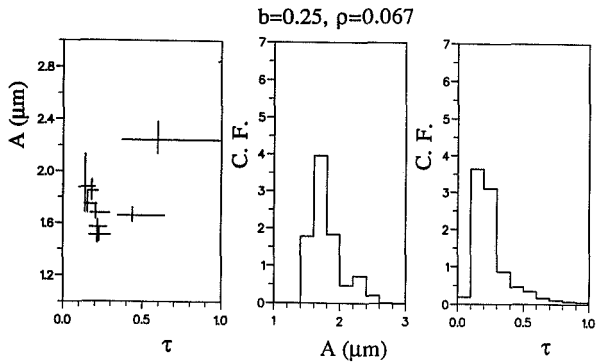


Figure 6. The values of the vertically averaged effective radius A and optical depth τ at 1.9 μm are indicated for the nine IR occultations with error bars at the 1-σ level in the left panel. Corresponding histograms of the radius and optical depths are plotted in the right panel. The effective variance b and the roughness parameter ρ are 0.25 and 0.067, respectively.

error bar is, for one given occultation, the error resulting from the different uncertainties linked to the measurement (instrumental noise, uncertain altitude, error in the retrieval of size and number density). The geographical error bar is the empirical standard deviation obtained from a statistical treatment of the results of the nine occultations at each altitude level. It is an indication of the natural variability of profiles with time and space. The geographical error bar is generally larger than the instrumental one in these data, except for the dust volume mixing ratio (Figure 10). The mixing ratio of the dust suspended in the atmosphere seems therefore to be almost constant with space and time, as expected if transport dominates. The number density, radius, and optical depth of particles are more variable. It should be noted that the volume mixing ratio of the dust is obtained without any assumption on the radius, whereas number density of particles depends on both mixing ratio and radius. The retrieval of number density, radius, and optical depth is therefore expected to be more model dependent. Nevertheless, we think that the geographical variability of dust size, and derived quantities at constant volume mixing ratio, is real.

The particle number density averaged over the first atmospheric scale height is given for the nine available occultations in Table 3. The major source of uncertainty in this case is the effective variance b of the particle size distribution. This may be seen in Figure 12, where the vertical profiles of effective radius and number density, averaged over the nine occultations, are plotted as a function of altitude for various values of b and ρ. The radius and number density profiles corresponding to the mean (b =0.25) and extreme (b =0.10, 0.40) cases are given in Table 4 for a plausible roughness parameter of 0.067. It should be noted that a simple exponential extrapolation of optical depths measured by occultation would lead to a larger value of the total optical depth (≈0.25 rather than 0.2), although the vertical profile of number density would remain rather similar to those proposed in Table 4.

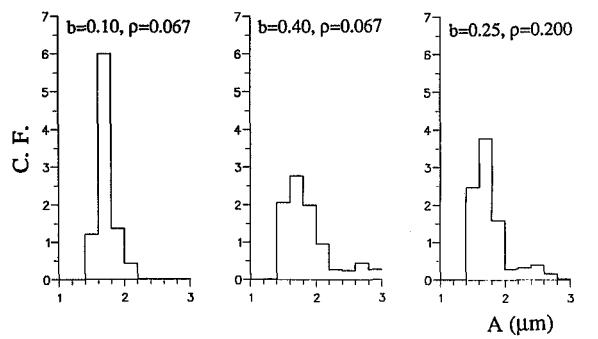


Figure 7. Histogram of the radius for the three following cases: (left) b=0.10, ρ=0.067; (middle) b=0.40, ρ=0.067; (right) b=0.25, ρ=0.2.

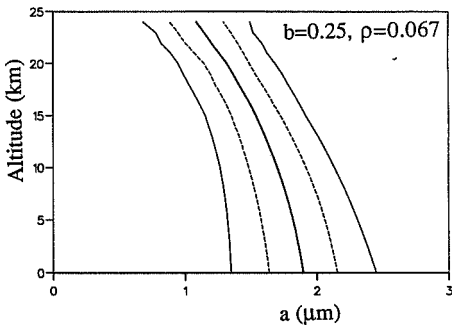


Figure 8. The mean vertical profile of the dust effective radius is shown with a thick solid line. Extreme profiles at the 1-σ level corresponding to the averaged instrumental error bar and geographical variability are plotted with dashed and thin solid lines, respectively.

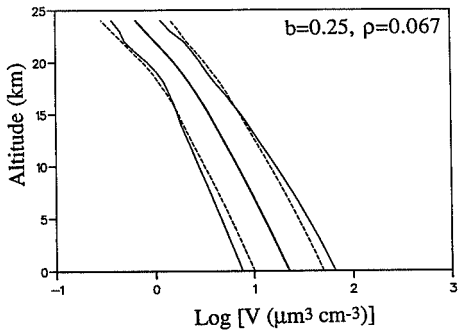


Figure 10. Same as Figure 8 for the dust volume mixing ratio.

The dust optical depth derived from the present study (≈0.1-0.3, Figure 6) is in good agreement with KRFM and Termoskan ranges, ≈0.12-0.24 and 0.12-0.26 (Table 3), respectively. It is similar to the ISM value (Table 3). A detailed examination of Table 3 and Figure 7 shows that there is a certain variability of optical depth with space and time. This is proven by Figure 9, where it can be seen that the instrumental error bar is a factor of 2 smaller than the geographical one. The dust scale height, and more significantly the number density of dust particles near the surface, as deduced from Termoskan measurements, are in good agreement with Auguste values (Table 3). Seven of the nine values of number density obtained from solar occultation measurements are in the Termoskan range (1-2 cm^{-3}). One of the most interesting points is the relatively small value of the effective radius deduced from both Auguste (1.7±0.2 μm) and ISM (1.2±0.4 μm), as compared to the generally accepted pre-Phobos value (≈2.5 μm [*Toon et al.*, 1977; *Pollack et al.*, 1979]). *Clancy and Lee* [1991] have argued that the Viking data are consistent with a dust particle size 5-10 times smaller than previously admitted, but there is no confirmation of such a small radius in Phobos data. It is emphasized that the variability in space and time of the effective radius, as deduced from solar occultation measurements, is small. There is one exception (occultation 2362241, Tables 2, 3) but, as already mentioned, this particular case is less reliable than the others. The particle effective radius, derived from Auguste, is never smaller than 1.5 μm. It is nevertheless emphasized that the ISM observation selected in the present work was performed on the Tharsis plateau, westward of Pavonis Mons, and that the footprint of the field of view covered a wide altitude range from ≈0 km up to ≈10 km, with a mean altitude of the sounded region of ≈7 km. The averaged value of 1.7 μm, obtained at zero reference altitude, must be reduced by about ≈0.2 μm, which put Auguste and ISM data in good agreement at the 1-σ level. Cross validation of data is therefore satisfactory, and the extrapolation used to extend solar occultation vertical profiles from ≈15 km, which is the lowest sounded altitude in occultation, down to zero reference altitude may therefore be considered as realistic.

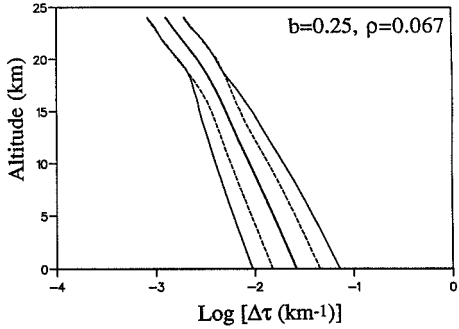


Figure 9. Same as Figure 8 for the dust extinction coefficient.

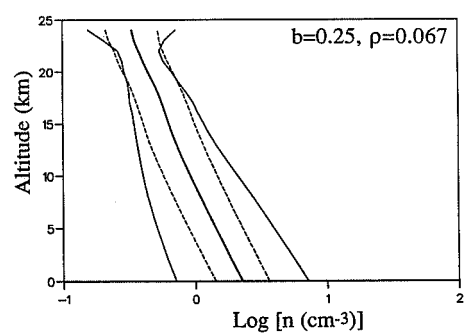


Figure 11. Same as Figure 8 for the dust number density.

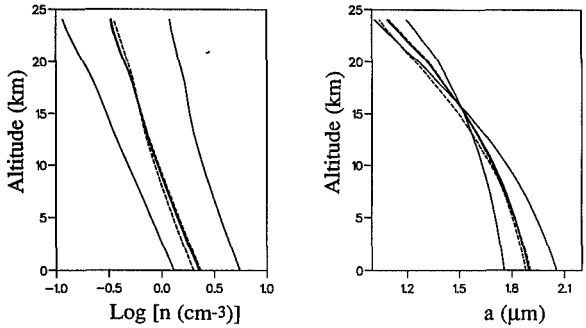


Figure 12. Most likely vertical profiles of the number density and effective radius for the nominal case (b=0.25, ρ=0.067, thick solid lines), extreme effective variances (b=0.10/0.40, ρ=0.067, thin solid lines) and extreme roughness parameters (b=0.25, ρ=0/0.2, dashed lines).

Discussion

The approach which was used to extrapolate vertical profiles of dust volume mixing ratio and particle radius from the range where they are measured (15-25 km) down to the ground is based on the assumption that vertical transport may be treated in terms of eddy diffusion. Prior to any assessment of the validity of this concept, it is necessary to estimate the relevant temporal and spatial scales. The timescale for eddy mixing is equal to ≈10 days by assuming a value of the eddy diffusion coefficient of 10^6 cm^2 s^{-1} [*Korablev et al.*, 1993]. The typical time for a dust particle of radius ≈2 μm to fall one atmospheric scale height at the 2mbar level is also ≈10 days [*Kahn et al.*, 1992]. These times are much smaller than the mean coagulation time (≈10^8 s) and particles may be supposed to not interact. On the other hand, they are definitely greater than 1 sol (1 Martian solar day) and the vertical structure of dust is not expected to follow possible diurnal variations of atmospheric conditions, at least outside the boundary layer (z>≈4 km). Spatial scales are obtained by multiplying timescales by the wind velocities in the relevant altitude range (0-25 km). By assuming typical values of the zonal and meridional wind velocities of ≈10 m s^{-1} and ≈1 m s^{-1}, respectively, zonal and meridional scales are ≈10^4 km and ≈10^3 km, respectively. The coefficient K used in the steady state model must therefore be considered as a zonally averaged empirical quantity including both vertical mixing and meridional advective transport on the mesoscale. Although it is not appropriate in a rigorous way, this kind of assumption is currently used in terrestrial photochemical models. Because there is no dramatic change of the observed dust profile and inferred quantities for varying time and longitude of observation, the use of a spatially averaged value of K is justified.

It is emphasized that the steady state model is not used to fit the data, but only to parameterize the vertical variation law of some dust characteristics. In particular, the possibility that gravitational settling of particles should be counteracted by upwelling, with typical ascending velocities of some mm s^{-1} in equatorial regions, does not invalidate the extrapolation laws used in the present work (8) and (11). Although it might introduce a small error due to the variation of the vertical wind velocity over the relevant altitude range (z<25 km), the possible presence of upwelling is implicitly taken into account when fitting data to the model in the sounded altitude range. In a more general way, the range of extrapolated profiles obtained for each observed profile, as shown in Figure 3, does not seem to be very dependent on possible large-scale disturbing effects. Moreover, the validity of the extrapolation method is supported by the lack of any slope breaking around 15-km altitude on profiles presented in Figures 8-12.

There are several ways dust particles may be lifted from the surface into the atmosphere. The recent model of *Haberle et al.* [1993 b] shows a strong diurnal variation of the height of the boundary layer. During the morning and the beginning of the afternoon, the surface is heated by sunlight. The overlying atmosphere is warmed by ground thermal radiation, with a regular increase of the height of the convective layer from 100 m up to ≈4 km at 16 hours, followed by a rapid collapse due to the fact that the ground temperature falls below the atmospheric temperature. It is thought that direct heating of the ground might give rise to strong upward transport of dust by convection, with subsequent formation of sporadic dust-devils or convective vortices [*Kahn et al.*, 1992]. These kinds of structures have been observed, with dust columns of 1 km in width reaching 6 km in altitude, near the top of the boundary layer. It was also argued that, in addition to direct solar heating, regional winds might generate the initial shear stress required to lift surface dust. Although no spiral dust cloud larger than a few kilometers across has been observed, the role of local storms in redistributing the dust could be of prime importance. The fact that no marked variation of the dust column density with altitude was measured from Phobos 2 on the flank of the volcanoes [*Drossart et al.*, 1991] is perhaps the consequence of such a phenomenon. It was argued by the authors that strong slope winds, induced by the solar heating of the slopes, should be responsible for transport of dust and result in the complete redistribution of the suspended dust over the whole area of volcanoes. In this way, the layer of dust would follow the relief instead of being confined in low regions,

Table 4. Three Reference Models for the Vertical Distribution of the Number Density and Effective Radius of Dust Particles

	Case 1: b=0.25		Case 2: b=0.10		Case 3: b=0.40	
Altitude, km	n, cm^{-3}	r_{eff} μm	n, cm^{-3}	r_{eff}, μm	n, cm^{-3}	r_{eff}, μm
0	2.26	1.90	1.31	1.76	5.56	2.05
1	2.06	1.88	1.18	1.75	5.09	2.03
2	1.88	1.87	1.07	1.74	4.67	2.01
3	1.71	1.85	0.97	1.73	4.28	1.99
4	1.56	1.84	0.87	1.72	3.93	1.97
5	1.43	1.82	0.79	1.71	3.62	1.94
6	1.30	1.80	0.72	1.70	3.34	1.91
7	1.19	1.78	0.65	1.68	3.08	1.88
8	1.09	1.75	0.59	1.67	2.86	1.85
9	1.00	1.73	0.53	1.65	2.65	1.81
10	0.92	1.70	0.48	1.64	2.47	1.78
11	0.85	1.67	0.43	1.62	2.31	1.74
12	0.78	1.64	0.39	1.60	2.16	1.69
13	0.72	1.61	0.36	1.58	2.04	1.65
14	0.68	1.57	0.33	1.55	1.94	1.60
15	0.64	1.54	0.30	1.53	1.86	1.55
16	0.60	1.49	0.27	1.50	1.79	1.50
17	0.56	1.45	0.25	1.47	1.72	1.44
18	0.52	1.41	0.23	1.44	1.65	1.38
19	0.48	1.36	0.20	1.41	1.56	1.33
20	0.44	1.32	0.18	1.38	1.47	1.27
21	0.40	1.26	0.16	1.34	1.40	1.20
22	0.38	1.20	0.14	1.29	1.32	1.14
23	0.35	1.15	0.13	1.25	1.25	1.08
24	0.33	1.09	0.12	1.20	1.20	1.02

and no vertical variation of the optical depth should be seen on a small scale or even mesoscale.

If this were the case on a global scale, a correlation between the optical depths derived from solar occultation measurements, which are calculated with respect to the common zero reference altitude (altitude of the reference ellipsoid), and the actual altitude of the sounded region would be expected. It is definitely not the case. Small optical depths of ≈0.15-0.20 are obtained indifferently above high regions, like the flanks of Ascraeus Mons (occultation 2441514), and low regions like Amazonis Planitia (occultation 2510743). Any redistribution of dust by convection in the low atmosphere does not propagate to higher atmospheric levels, where solar occultation measurements were made (15-25 km). Since the timescale for vertical mixing is ≈10 days, horizontal homogenization must occur over a typical scale of 10^3 km (by assuming a near-surface wind velocity of ≈1 m s^{-1}), which suggests that convective redistribution of dust must be regional rather than global. It must be noted that dust optical depths at 1.9-μm wavelength, as shown in Table 3, must be reduced by ≈30% on average when stopping the integration of the extinction coefficient at the ground level, instead of the zero reference altitude. Profiles presented in Figures 8-12 must be considered as averages over the low-latitude reference geoid. They should be cut at the ground altitude in any modeling using them as input data. In summary, the Phobos observational results are self-consistent. However,

uncertainties in dust properties remain large, particularly near the surface.

The sporadic presence of expanding low dust clouds may result in an excess of large particles in the first half scale height above the ground on the mesoscale. Nevertheless, particles whose radius is larger than ≈10 μm and therefore whose settling time between ≈5 km altitude and the ground is lower than ≈1 sol (typical period of one diurnal cycle of the boundary layer) cannot be efficiently levitated. Typical mean profiles presented in Figures 8-12 might be locally modified below ≈5 km altitude, with a possible additional component of large particles, whose radius should be in the range 2-10 μm. There is no clear evidence for such a phenomenon in the Phobos data.

The vertically averaged effective radius derived from Phobos data peaks around a modal value of ≈1.7 μm (Figures 6, 7; Table 3), which is smaller than the value derived from Viking lander measurements [*Pollack et al.*, 1979], ≈2.5 μm. The Phobos value of the radius at the ground level is ≈1.9 μm (Figure 8). This becomes nearly 2.1 μm if a large effective variance of 0.4 is assumed (Figure 12). It must be noted that the Viking value is consistent at 1-σ, considering geographical variability, with Phobos values. This is mainly due to occultation 2362241, which provides a rather large value of the radius (2.3 μm). Although this particular occultation, which is responsible for the high-radius wing (2 to 3 μm) of the histograms shown in Figures 6 and 7, is less reliable than others, for reasons already explained, no strong disagreement appears between Viking and Phobos data. A relatively small difference (<50%) in eddy mixing coefficient or vertical winds in the low atmosphere might account for such a disagreement. The variability of the eddy diffusion coefficient in space and time, as derived from Phobos data, is a factor of ≈2 on either side below 35-km altitude, increasing up to 1 order of magnitude above this level [*Rosenqvist and Chassefière*, this issue]. It is emphasized that the atmosphere appears to be significantly clearer during the Phobos observations than during most of the Viking mission. Under these conditions, where dust has presumably settled out, one might expect to find smaller particle sizes for the remaining dust. It seems in any case possible to use present results in a wide range of situations. Synthetic profiles given in Table 4 can be modified by any linear scaling for modeling purposes. There is nevertheless some uncertainty due to the fact that the variance of the size distribution of particles is not accurately known.

Appendix : Calculation of Scattering Extinction Cross Sections for Irregular Silicate Particles

The physical parameters used to calculate the scattering extinction cross sections for a given size distribution are as follows.

1. The real and imaginary optical index [*Pollack et al.*, 1973] at 1.9 μm: N=1.51-0.0026 i, and at 3.7 μm: N=1.48-0.0040 i.

2. The roughness parameter ρ [*Drossart*, 1990], which parameterizes the scale of irregularities, as compared to the size of the scatterer.

3. The two parameters of the size distribution, according to the analytical form used in (2): the effective (projected area-weighted) radius of particles and the effective variance of the particle size distribution.

A tabulation is obtained for the scattering cross section for a set of parameters in the following ranges: $0.1<a<4.0$ μm, $0.1<b<0.5$, $0<\rho<0.2$. The roughness parameter has been estimated from comparison to both ISM spectra and Viking dust properties [*Pollack et al.*, 1979]. A most likely value of 0.067 is adopted for calculations. Practically, the extinction cross section may be shown to be nearly insensitive to ρ for effective radii larger than ≈1 μm. The value of ρ has therefore very little effect on the results of this study, implying differences of no more than a few percent, since the effective radius of dust particles is found to lie in the range 1-2 μm. The sensitivity of results to ρ is tested.

Acknowledgments. We thank R. Zurek and J. Schofield for their very constructive referee reports. We are indebted to N. Rey for her help in the preparation of the camera-ready manuscript.

References

Anderson, E., and C. Leovy, Mariner 9 television limb observations of dust and ice hazes on Mars, *J. Atmos. Sci.*, *35*, 723-734, 1978.

Bibring, J.P., et al., Results from the ISM experiment, *Nature*, *341*, 591-593, 1989.

Blamont, J.E., E. Chassefière, J.P. Goutail, B. Mège, M. Nunes-Pinharanda, G. Souchon, V.A. Krasnopolsky, A.A. Krysko, and V.I. Moroz, Vertical profiles of dust and ozone in the Martian atmosphere deduced from solar occultation measurements, *Nature*, *341*, 600-603, 1989.

Blamont, J.E., E. Chassefière, J.P. Goutail, B. Mège, M. Nunes-Pinharanda, G. Souchon, V.A. Krasnopolsky, A.A. Krysko, and V.I. Moroz, Vertical profiles of dust and ozone in the Martian atmosphere deduced from solar occultation measurements, *Planet. Space Sci.*, *39*, 175-187, 1991.

Chassefière, E., J.E. Blamont, V.A. Krasnopolsky, O.I. Korablev, S.K. Atreya, and R.A. West, Vertical structure and size distributions of Martian aerosols from solar occultation measurements, *Icarus*, *97*, 46-69, 1992.

Clancy, R.T., and S.W. Lee, A new look at dust and clouds in the Mars atmosphere: Analysis of emission-phase function sequences from global Viking IRTM observations, *Icarus*, *93*, 135-158, 1991.

Combes, M., et al., Martian atmosphere studies from the ISM experiment, *Planet. Space Sci.*, *39*, 189-197, 1991.

Conrath, B.J., Thermal structure of the Martian atmosphere during the dissipation of the dust storm of 1971, *Icarus*, *24*, 36-46, 1975.

Drossart, P., A statistical model for the scattering by irregular particles, *Astrophys. J.*, *361*, L29-L32, 1990.

Drossart, P., J. Rosenqvist, S. Erard, Y. Langevin, J.P. Bibring, and M. Combes, Martian aerosol properties from the Phobos/ISM experiment, *Ann. Geophysicae*, *9*, 754-760, 1991.

Haberle, R.M., J.B. Pollack, J.R. Barnes, R.W. Zurek, C.B. Leovy, J.R. Murphy, H. Lee, and J. Schaeffer, Mars atmospheric dynamics as simulated by the NASA Ames general circulation model, 1, The zonal-mean circulation, *J. Geophys. Res.*, *98*, 3093-3124, 1993a.

Haberle, R.M., H.C. Houben, R. Hertenstein, and T. Herdtle, A boundary layer model for Mars: Comparison with Viking lander and entry data, *J. Atmos. Sci.*, *50*, 1544-1559, 1993b.

Hansen, J.E., and L.M. Travis, Light scattering in planetary atmospheres, *Space Sci. Rev.*, *16*, 527-610, 1974.

Kahn, R.A., T.Z. Martin, R.W. Zurek, and S.W. Lee, The Martian dust cycle, in *Mars*, pp. 1017-1053, University of Arizona Press, Tucson, 1992.

Korablev, O.I., V.A. Krasnopolsky, A.V. Rodin, and E. Chassefière, Vertical structure of the Martian dust measured by the solar infrared occultations from the Phobos spacecraft, *Icarus*, *102*, 76-87, 1993.

Krasnopolsky, V.A., V.I. Moroz, A.A. Krysko, O.I. Korablev, V.S. Zhegulev, A.V. Grigoriev, A.Yu. Tkachuk, V.A. Parshev, J.E. Blamont and J.P. Goutail, Solar occultation spectroscopic

measurements of the Martian atmosphere at 1.9 and 3.7 μm, *Nature, 341*, 603-606, 1989.

Moroz, V.I., E.V. Petrova, and L.V. Ksanfomality, Spectrophotometry of Mars in the KRFM experiment of the Phobos mission: Some properties of the particles of atmospheric aerosols and the surface, *Planet. Space Sci., 41*, 569-585, 1993.

Moroz, V.I., Y.M. Getkin, M.K. Naraeva, A.S. Selivanov, and D.V. Titov, Aerosol vertical profile on Mars from the measurements of the thermal radiation on limb, *Planet. Space Sci.*, in press, 1995.

Murphy, J.R., R.M. Haberle, O.B. Toon, and J.B. Pollack, Martian global dust storms: Zonally symmetric numerical simulations including size-dependent particle transport, *J. Geophys. Res., 98*, 3197-3220, 1993.

Pang, K., and J.M. Ajello, Complex refractive index of Martian dust: Wavelength dependence and composition, *Icarus, 30*, 63-74, 1977.

Pollack, J.B., O.B. Toon, and B.N. Khare, Optical properties of some terrestrial rocks and glasses, *Icarus, 19*, 372-389, 1973.

Pollack, J.B., D.S. Colburn, F.M. Flasar, R. Kahn, C.E. Carlston, and D. Pideck, Properties and effects of dust suspended in the Martian atmosphere, *J. Geophys. Res., 84*, 2929-2945, 1979.

Rosenqvist, J., and E. Chassefière, A reexamination of the relationship between eddy mixing and O_2 in the Martian middle atmosphere, J. Geophys. Res., this issue.

Toon, O.B., J.B. Pollack, and C. Sagan, Physical properties of the particles composing the Martian great dust storm of 1971-1972, *Icarus, 30*, 663-696, 1977.

E. Chassefière, Service d'Aéronomie du CNRS, BP 3, 91371 Verrières-le-Buisson Cedex, France. (e-mail: eric.chassefiere@aerov.jussieu.fr)

P. Drossart, DESPA, Observatoire de Paris-Meudon, 92195 Meudon, France.

O. Korablev, Space Research Institute, Profsoyuznaya 84/32; 117810 Moscow, GSP-7, Russia.

(Received February 22, 1994; revised December 22, 1994; accepted December 27, 1994.)

Page intentionally left blank

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5541–5551, MARCH 25, 1995

A reexamination of the relationship between eddy mixing and O_2 in the Martian middle atmosphere

Jan Rosenqvist
Département de Recherche Spatiale, Observatoire de Paris-Meudon, Meudon, France

E.Chassefière
Service d'Aéronomie du Centre National de la Recherche Scientifique, Verrières-le-Buisson, France

Abstract. A value of the eddy diffusion coefficient K of $\approx 1.5 \times 10^6$ cm^2 s^{-1} in the middle atmosphere of Mars was obtained from Phobos 2 solar occultation measurements of dust, ozone, and clouds at low latitude. The aim of the present study is to complete this picture by using a steady state photochemical one-dimensional model. The main regulation mechanism of O_2 is the reaction of O with itself, whose rate depends on the value of K in the middle atmosphere. By comparing calculated and observed values of the O_2 abundance, an upper limit of $\approx 2 \times 10^7$ cm^{-2} s^{-1} on K is inferred. By including an additional constraint provided by H_2 balance, a lower limit of $\approx 4 \times 10^6$ cm^{-2} s^{-1} may be placed. It results from the present analysis that the most realistic value of K to be used in works resorting to one-dimensional modeling (long-term evolution, escape, surface/atmosphere exchanges) is $\approx 10^7$ cm^2 s^{-1} rather than $\approx 10^6$ cm^2 s^{-1}. The difference between theoretical and observational values might be due to the regular occurrence of global dust storms, whose effect should be to increase the yearly average value of K. The present study suggests less than 3 precipitable micrometer (pr-μm) of the yearly averaged water vapor column, unless H_2O is confined in a layer near the ground. Although the first possibility seems more probable, the second hypothesis cannot be ruled out. It could reflect a continuous supply of H2O from the regolith to the atmosphere on a seasonal scale.The loss to production ratio of CO, which has a lifetime of ≈5 years, is shown to depart from unity by no more than ≈10% over a wide range of atmospheric conditions. The stability of the Martian atmosphere is therefore realized in the classical frame of homogeneous chemistry.

Introduction

Recent observations of dust, ozone, and clouds in the middle atmosphere of Mars from the Phobos 2 spacecraft by the solar occultation technique have stimulated a modelling effort, partly focused on the estimate of the eddy diffusion coefficient in different altitude ranges from ≈15 km up to nearly ≈60 km. The specific interest of Phobos solar occultation measurements is that they cover a wide range of longitudes, although limited to a rather narrow latitude band (0°-20°N) and a small interval of solar areocentric longitude (L_S=0°-20°), at the time of northern spring equinox. The small variability of the dust distribution and inferred parameters with space and time shows that atmospheric conditions in the low and middle equatorial atmosphere of Mars are zonally homogeneous. These data provide constraints on eddy mixing strength and some other atmospheric parameters (humidity level, temperature) in three different altitude ranges, i.e., ≈15-30 km [*Korablev et al.*, 1993], ≈37-55 km [*Blamont and Chassefière*, 1993], and ≈45-60 km [*Chassefière et al.*, 1992].

The vertical structure and size distribution of the dust component between 15 km and 30 km altitude were estimated by *Korablev et al.* [1993]. By using a one-dimensional model involving gravitational settling and eddy diffusion as vertical transport processes, the value of the eddy diffusion coefficient K below 30 km altitude was found to be $\approx 10^6$ cm^2 s^{-1} within a factor of 2. Ozone was observed in the 37 to 55-km altitude range [*Blamont and Chassefière*, 1993]. By using a simple one-dimensional photochemical model, a good fit of model to data was obtained for $K \approx 2\ 10^6$ cm^2 s^{-1} within a factor of 2. The fit was improved by assuming a steep increase of K for increasing altitude around 40-45 km, from $\approx 5 \times 10^5$ cm^2 s^{-1} below 40 km up to $\approx 10^7$ cm^2 s^{-1} above 45 km. The analysis and modeling of cloud layers observed in the 45 to 60-km altitude range [*Chassefière et al.*, 1992] suggested a most likely value of K of $\approx 10^6$ cm^2 s^{-1} within 1 order of magnitude. Apart from one of these clouds, for which a complete confidence interval was obtained (K=4 10^5-4 10^6 cm^2 s^{-1}), only upper limits on K were derived. A synthetic view of all previous results is given in Figure 1. The large scatter on K found in the last case presumably reflects an increased variability at higher atmospheric levels. It is emphasized that a common moderate value of the eddy diffusion coefficient ($K \approx 1.5 \times 10^6$ cm^2 s^{-1}) accounts for all observed phenomena.

The fact that no reliable information exists in the 60 to 100-km range was already mentioned by *Krasnopolsky* [1993a]. A lower limit on K of 2×10^7 cm^{-2} s^{-1} at 50 to 70-km altitude was deduced from the analysis of a high cloud observed in the Viking orbiter images [*Kahn*, 1990; *Chassefière et al.*, 1992]. This cloud was detected at

Paper number 95JE00324.
0148-0227/95/95JE-00324$05.00

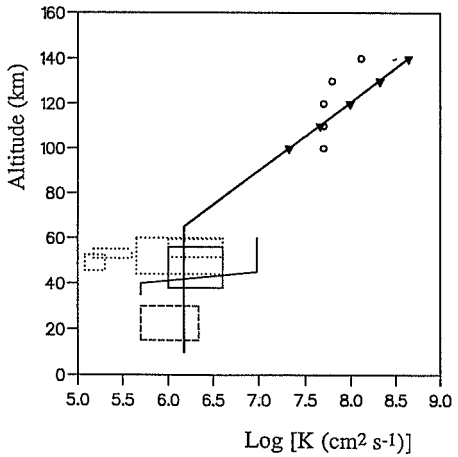


Figure 1. Values of the eddy diffusion coefficient as inferred from Phobos 2 observations (dust: box with dashed lines; ozone: box with solid lines; clouds: boxes with dotted lines) and Viking upper atmosphere measurements (V2, triangles, V1, circles). A mean vertical profile of K consistent with observations is shown with a thick solid line. An alternate profile, showing a steep increase around 40-45 km, is plotted with a thin solid line.

midlatitudes in early northern winter (L_S=139°). It is emphasized that K is expected to be highly variable in this altitude range due to the necessarily sporadic character of turbulence generated by internal gravity waves, whose breaking level is extremely sensitive to atmospheric thermal profile and wave characteristics [*Théodore et al.*, 1993]. Above 100 km, the vertical profile of K has been derived from N_2, Ar, CO, and O_2 measurements by the two Viking entry probes [*Nier and McElroy*, 1977], leading to values increasing from 2-5x10^7 cm^{-2} s^{-1} at ≈100 km up to 1-4x10^9 cm^{-2} s^{-1} at ≈170 km (Figure 1).

Steady state models used for the different analyses previously mentioned were supposed to describe the behavior of the atmosphere on short timescales, typically 1 day to 1 week. Moreover, despite the fact that phenomena cannot be considered as local but rather acting on mesoscale, the vertical profiles of background quantities (T, H_2O, CO, O_2, H_2) were fixed in an external way, either by using simultaneous measurements (H_2O), or from global-scale Earth-based observations for long-lifetime species (CO, O_2). In other words, the abundances of background species were not equilibrated by using the model on a global scale, due to the fact that relevant temporal and spatial scales are much smaller than 1 year and 1 planetary radius, respectively .

As explained by *Blamont and Chassefière* [1993], the vertical structure of O_3 in the middle atmosphere crucially depends on the altitude profile of the eddy diffusion coefficient K. The value of K in the middle atmosphere similarly influences the O_2 mixing ratio on the long-term. Indeed, processes involved in the short-term balance of O_3 and long-term balance of O_2 are basically the same: the main two reactions yielding O_2 (O+O+CO_2, O+HO_2) deplete O_3. The balance of O_2 is therefore expected to be sensitive to K. This idea was implicit in the work of *Kong and McElroy.* [1977]. The first goal of this paper is to reexamine the relationship between K and O_2 in the Martian middle atmosphere by using up-to-date values of reaction rates and Phobos 2 observations. New elements are proposed concerning the sensitivity of the results to the water vapor vertical distribution.

Photochemical Modeling

The Photochemical Model

The photochemical model is basically the same as the one used by *Chassefière* [1991]. It is a one-dimensional model including the by-products of CO_2 and H_2O and using the photochemical scheme proposed by *Parkinson and Hunten* [1972]. It is described in detail by *Blamont and Chassefière* [1993]. The chemical rates are those used by *Moreau et al.* [1991]. The reactions are listed in Table 1. All other reactions used in more complete models may be proven to be of very little importance, with rates lower by at least 3 orders of magnitude than the rates of the corresponding main reactions. Condensation is taken into account for H_2O and H_2O_2. The condensation law of H_2O_2 is the one given by *Krasnopolsky* [1986]. Dust opacity, as derived from Phobos

Table 1. Photochemical reactions in the Martian Atmosphere

Number	Reaction
J_1	$CO_2 + h\nu \rightarrow CO + O$
J_2	$O_2 + h\nu \rightarrow O + O$
J_3	$O_3 + h\nu \rightarrow O_2 + O(^1D)$
J_4	$H_2O + h\nu \rightarrow OH + H$
J_5	$HO_2 + h\nu \rightarrow OH + O$
J_6	$H_2O_2 + h\nu \rightarrow OH + OH$
R_1	$O(^1D) + CO_2 \rightarrow O + CO_2$
R_2	$O(^1D) + H_2O \rightarrow OH + OH$
R_3	$O(^1D) + H_2 \rightarrow OH + H$
R_4	$H + O_2 + CO_2 \rightarrow HO_2 + CO_2$
R_5	$O + O + CO_2 \rightarrow O_2 + CO_2$
R_6	$O + O_2 + CO_2 \rightarrow O_3 + CO_2$
R_7	$O + HO_2 \rightarrow OH + O_2$
R_8	$O_3 + H \rightarrow OH + O_2$
R_9	$O + OH \rightarrow O_2 + H$
R_{10}	$CO + OH \rightarrow CO_2 + H$
R_{11}	$H + HO_2 \rightarrow OH + OH$
R_{12}	$H + HO_2 \rightarrow H_2 + O$
R_{13}	$H + HO_2 \rightarrow H_2O + O$
R_{14}	$HO_2 + HO_2 \rightarrow H_2O_2 + O_2$
R_{15}	$HO_2 + OH \rightarrow H_2O + O_2$
R_{16}	$H_2O_2 + OH \rightarrow H_2O + HO_2$

2 solar occultation data [*Chassefière et al.*, 1992], is included in the calculation of UV fluxes. The thermal profile is basically taken from *Seiff* [1982] for the lower atmosphere. At higher levels, we have also considered the infrared thermal mapper (IRTM) data [*Martin et al.*, 1979]. The pressure at the surface has been taken from *Hess et al.* [1980], with a value of 7 mbar at 0 km ground altitude for L_S≈0-20°. Several tests corresponding to different Martian seasons have been performed to estimate the influence of temperature and pressure profiles on the O_2 equilibrium abundance. The results show that the influence of pressure and temperature is quantitatively weak as compared to that of K and water vapor vertical distribution. Thus, in the following, we fix the surface pressure at a value of 7 mbar and use a unique standard thermal profile corresponding to an approximate average on the year, the season, and the day: T=205 K at 0 km altitude, 175 K at 25 km, and 145 K at 50 km. It is emphasized that this profile fits well the one proposed by *Krasnopolsky et al.* [1991] from solar occultation measurements of water vapor.

In the present work, the total hydrogen escape flux $\Phi(H)$ is parameterized in the following way:

$$\Phi(H) = \Phi(H)_0 \frac{\rho(H_2)}{\rho(H_2)_0} \quad (1)$$

where $\Phi(H)_0$ is the present value of the hydrogen escape flux, $\rho(H_2)$ is the equilibrium value of the H_2 mixing ratio, as given by the model, and $\rho(H_2)_0$ is the present value of the H_2 mixing ratio. In this study, $\rho(H_2)_0$ is fixed at 8.3×10^{-6} [*Kong and McElroy*, 1977]. The present value of the total escape flux of hydrogen $\Phi(H)_0$ is assumed to be 2×10^8 $cm^{-2}s^{-1}$ (see Table 1 of *Krasnopolsky* [1993b]). The oxygen escape flux $\Phi(O)$ is assumed to be twice as small as $\Phi(H)$ [*Liu and Donahue*, 1976].

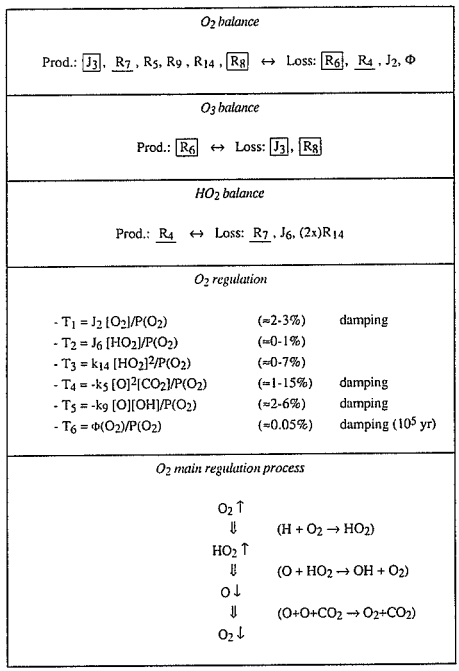


Figure 2. O_2 balance and related regulation mechanisms.

Photochemical Regulation of O_2

Production and loss reactions of O_2 are summarized in Figure 2. The production of O_2 by R_{12} and R_{15}, which are of very minor importance, are omitted for simplicity. The main O_2 production process is photodissociation of ozone (J_3), which represents 40-75% of the total production. Since reaction of O with O_2 (R_6), which is the only one yielding O_3, is the major loss process for molecular oxygen (45-75%); O_3 balance ensures more than one half of the O_2 balance. In the same way, the balance between R_7 and R_4 is roughly maintained through HO_2 balance, which ensures one third of the O_2 balance. The ratio of the loss to the production rate (LPR) of O_2 may be formally written:

$$LPR = \frac{L(O_2)}{P(O_2)} = 1 + \frac{[L(O_2)-P(O_3)-P(HO_2)] - [P(O_2)-L(O_3) - L(HO_2)]}{P(O_2)} \quad (2)$$

Neglecting terms with weights smaller than 1%, one obtains

$$LPR \approx 1 + \sum_{i=1}^{i=6} T_i \quad (3)$$

where the analytical expressions and the orders of magnitude of terms T_i are given in Figure 2. Positive and negative sums (respectively $T_1+T_2+T_3+T_6$ and T_4+T_5) are in the range between 2% and 20% of the total LPR. The loss to production ratio of O_2 varies from 0.9 to 1.05 as the eddy diffusion coefficient K varies from 10^5 to 10^8 cm^2 s^{-1}. This means that the balance between O_2 production and loss rates is realized within 10% through the balances of short-lived O_3 and HO_2. A first step of the O_2 equilibrium is reached in less than 1 hour through the balance of O_3. In a second step, the balance of HO_2 is reached in about 1 hour, ensuring a rough equilibrium between the production and loss rates of molecular oxygen. Fine adjustment is realized only when the sum of the six terms in (3) becomes equal to zero, over a time-scale of the order of the O_2 lifetime (≈10^2 years), or even more in the case of escape ($[O_2]/\Phi(O_2)$≈10^5 years).

A brief examination of the regulation terms is necessary at this step of the work. Figure 3 shows the reaction rates of the 22 photochemical reactions involved in the model as a function of altitude (Figure 3a) and the associated vertical profiles of the number densities of the constituents for standard conditions (Figure 3b). The effect of the condensation of H_2O_2 may be seen on profiles at altitudes between 30 and 40 km. This figure shows that the atmosphere can be roughly divided in four layers, which are [0-20 km], [20-40 km], [40-60 km], and altitudes above 60 km. The dominant balance reactions of OH, O, HO_2, H_2O_2, and H* (odd hydrogen family) in these four layers are given in Table 2.

The term of photodissociation (T_1), whose value is weak (≈2-3%), is a damping term, since it is directly proportional to $[O_2]$. Any relative variation of O_2 will be translated in a similar relative variation of the O_2 photodissociation rate

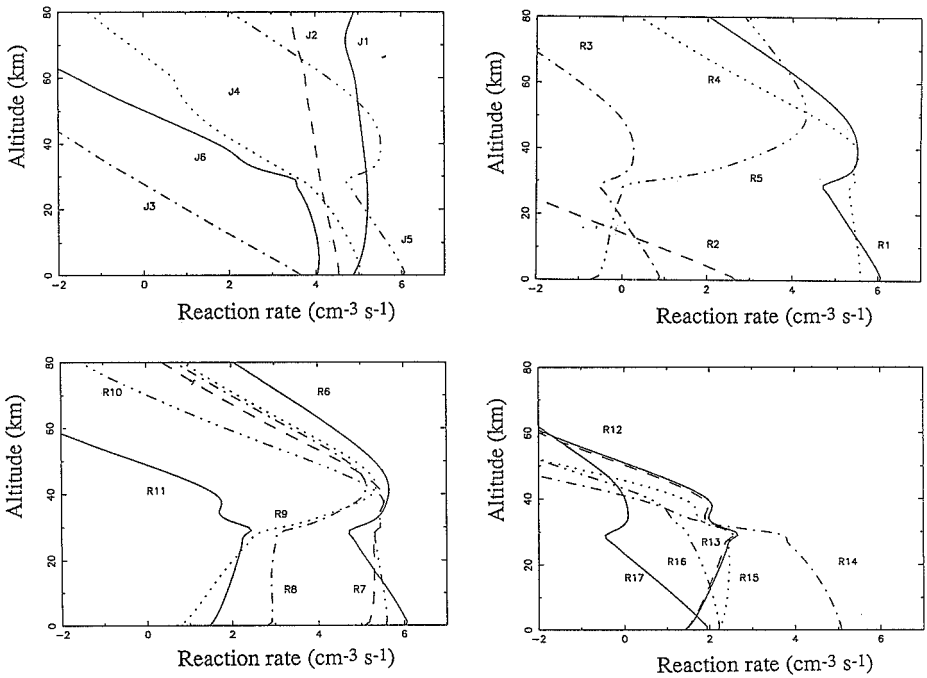


Figure 3a. Reaction rates as a function of altitude for the 22 photochemical reactions listed in Table 1 for standard conditions: $K = 10^7$ cm^2 s^{-1}, $[H_2O]/[CO_2] = 5.6\ 10^{-4}$, c(water vapor column amount)=11 pr-μm, h(H_2O mixing ratio scale height)=3 km.

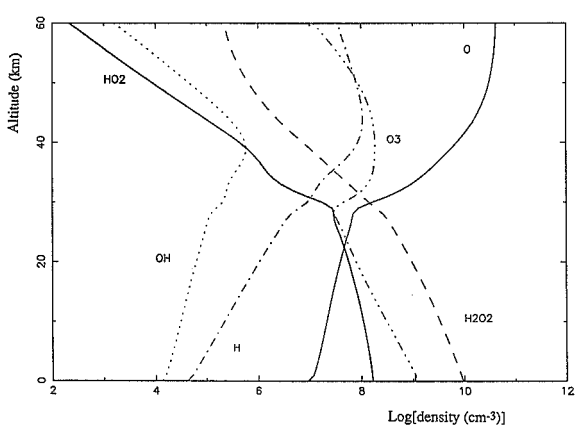


Figure 3b. Vertical profiles of the number densities of the constituents.

because the penetrating solar ultraviolet flux is very little affected by a small variation of O_2. T_1 acts over a typical time-scale of the order of $1/J_2$, which is about the photochemical lifetime of O_2. The escape term ($T_6 \approx 0.05\%$) is also a damping term, since the escape flux is expected to be correlated with the global abundance of O_2. This question was discussed in detail by *Liu and Donahue* [1976].

The effect of an increase of the O_2 mixing ratio is to stimulate the reaction of H with O_2 in the presence of CO_2, which forms HO_2, that is, to modify the internal equilibrium of the H* family (H, OH, HO_2, plus H_2O_2) in such a way that there is relatively less H and more HO_2. In the general case, the population of the H* family does not change enough for this picture being modified when considering absolute amounts. Nevertheless, it must be mentioned that the opposite variation, that is, a decrease of HO_2 for increasing O_2, may be observed for very dry conditions ($[H_2O]/[CO_2] < 10^{-5}$). This phenomenon is due to a marked decrease of H*, which may be linked to a relatively strong loss of H* at high altitude (>25 km), where H dominates the H* family. Due to the relative decrease of H inside the H* family, a related decrease of H* occurs, which is stronger in absolute value than the relative increase of HO_2 inside the H* family, with a subsequent decrease of the absolute HO_2 amount. It may be concluded that T_2 and T_3, which both depend on the HO_2 amount, are damping terms in mean conditions, although acting negatively at some particular times and places, for H_2O column densities lower than 1 pr μm. T_2 and T_3 act over a typical timescale of the order of the H* lifetime (≈1 week).

T_4 and T_5 are indirectly connected to the O_2 amount through the abundance of O atoms. As previously mentioned, an increase of O_2 yields an increase of HO_2. This increase is particularly significant above ≈25 km altitude. Because up to 50 km, OH is regulated by CO [*Chassefière*, 1991], and therefore roughly invariable with O_2, O and O_2 are anticorrelated (see Figure 2) between 25 and 50-km altitude. This is confirmed by calculations. T_4 and T_5 are damping terms, since O is principally concentrated above 25-km altitude, and they act over a typical timescale of ≈1 week. It is emphasized that T_4 is the largest regulating term (see Figure 2). A simplified flow chart of the corresponding regulation chemical chain is given in Figure 2. The production of O_2 by reaction of O with HO_2 (R_7), which is the dominant sink of odd oxygen O* below ≈50 km altitude, does not enter the regulation equation (3). O_2 is mainly regulated through T_4 by the reaction of O with itself (R_5), which takes place above 50 km and constitutes the major sink of odd oxygen at high altitude. For this reason, the equilibrium value of $[O_2]$ is expected to be very sensitive to the value of the eddy diffusion coefficient K, especially in the middle atmosphere. For a low value of K, odd oxygen, formed mainly by photodissociation of CO_2, accumulates in the high atmosphere, where odd hydrogen species are not in sufficient amount to destroy it, and recombination of O with itself, whose rate is a quadratic function of [O], is quite efficient. Conversely, in the case of a strong vertical mixing, there is a rapid downward diffusion of newly created odd oxygen at high altitude and O is therefore depleted, yielding a lower self-recombination rate.

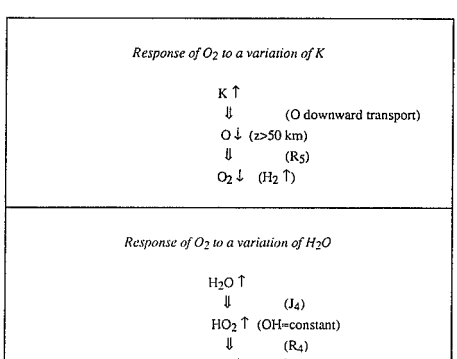


Figure 4. Response of the equilibrium abundance of O_2 to variations of K and H_2O concentration.

Sensitivity of the O_2 Mixing Radio

A simplified flow chart showing the response of O_2 to a variation of K is given in Figure 4. As suggested by *Kong and McElroy* [1977], there is a definite anticorrelation between K and $[O_2]$. The variation of O_2 with H_2O has now to be made precise. If the water vapor amount is large, the formation of H* is also large. Then, in order to ensure the balance of H*, the loss rate of H* must increase. This simple consideration suggests, as confirmed by calculations, that the total amounts of H and HO_2 must globally increase. At this step, the equilibrium of H atoms over the whole atmosphere can be approximately written as follows, by keeping only major terms:

$$\int k_{10}[CO][OH]\,dz \approx \int k_4[O_2][H][CO_2]\,dz \qquad (4)$$

The left side of the equation is nearly invariable, as may be deduced from the study by *Chassefière* [1991]. Consequently, the product $[H][O_2]$ is expected to be roughly constant for varying H_2O. Thus, as water vapor amount increases, $[O_2]$ decreases (see flow chart in Table 2). There is a clear anticorrelation between the equilibrium value of O_2 and the water vapor concentration.

Case of Molecular Hydrogen

The only process leading to the formation of H_2 is the reaction of H with HO_2 (R_{12}). The main loss processes are dissociation by $O(^1D)$ (R_3) and thermal escape of hydrogen atoms from the upper atmosphere. Any variation of H_2 therefore yields a variation of the same sign of the H_2 loss rate, with a slow damping over a timescale of the order of the H_2 lifetime ($\approx 10^4$ years). The amplitude of this variation depends on the relative weights of escape and chemical (R3) terms, which are of the same order of magnitude. It is emphasized that the photochemical behavior of H_2 differs from that of CO and O_2. H_2 is strongly regulated by both dissociation by $O(^1D)$ and thermal escape, and its equilibrium level is well defined by the balance between production and loss, contrary to the cases of CO and O_2, which are only

Table 2. Main Balance Reactions of Short-Lived Constituents as a Function of Altitude

Constituent	0-20 km	20-40 km	40-60 km	>60 km
OH	$2J_4+k_7 \leftrightarrow k_{10}$	$k_7 \leftrightarrow k_{10}$	$k_7+k_8 \leftrightarrow k_9$	
O	$J_1+2J_2 \leftrightarrow k_7$	$J_1 \leftrightarrow k_7$	$J_1 \leftrightarrow k_7+k_8+k_7$	$J_1 \leftrightarrow k_5$
HO_2	$k_4 \leftrightarrow k_7+2k_{14}$	$k_4 \leftrightarrow k_7$	$k_4 \leftrightarrow k_7$	
H_2O_2	$k_{14} \leftrightarrow J_4$	$k_{14} \leftrightarrow J_4$	$k_{14} \leftrightarrow J_4$	
H*	$J_3+k_2 \leftrightarrow k_{12}+k_{13}+k_{15}$	$J_3 \leftrightarrow k_{12}+k_{13}+k_{14}$	$J_3 \leftrightarrow k_{12}+k_{13}+k_{15}$	$J_3 \leftrightarrow k_{12}+k_{13}+k_{15}$

Density terms are systematically omitted in these systems of algebric equations for convenience.

weakly regulated at their equilibrium values. Because there is a well-established anticorrelation between O_2 and H_2 in the Martian atmosphere, the equilibrium value of H_2 is expected to depend on K and H_2O, in the same way as O_2. The H_2 equilibrium abundance is expected to be very sensitive to the vertical distribution of water vapor. Indeed, the formation of odd hydrogen species like H and HO_2, which react together to form H_2, is principally due to the photodissociation of H_2O.

Results

The observed mixing ratio of molecular oxygen is assumed to be $1.3\pm0.3\times10^{-3}$ for the present study [*Barker*, 1972; *Carleton and Traub*, 1972; *Trauger and Lunine*, 1983]. *Anderson* [1974] proposed an H_2 mixing ratio of 2×10^{-5}, consistent with the observed H density profile in the high atmosphere. A reanalysis of these data by *Kong and McElroy* [1977] yields a value twice as small of $\approx8.3\times10^{-6}$. Values obtained by *Shimazaki* [1989], as presented in his Table 7, are in the range 7×10^{-6}-1.5×10^{-5}. We therefore adopt a range of 7×10^{-6}-2×10^{-5}, which accounts well for the major part of the existing estimates.

The CO mixing ratio is fixed at its observed value (8×10^{-4}).The equilibrium mixing ratios of O_2 and H_2 are plotted as a function of K in Figure 5. Two cases are considered.

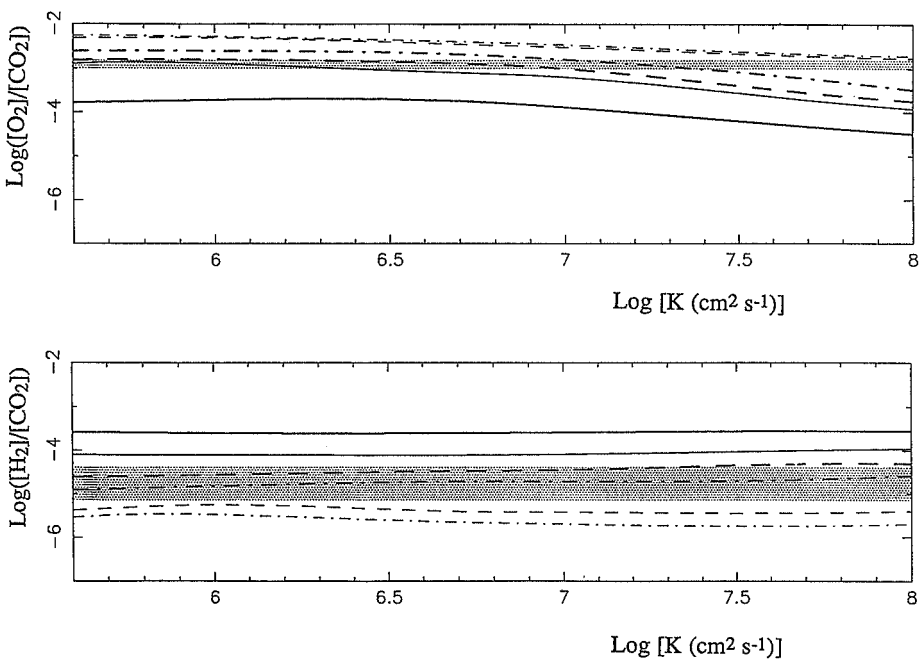


Figure 5. Variations of the O_2 and H_2 equilibrium mixing ratios as a function of K for different H_2O vertical profiles. The gray areas indicate the expected ranges of the O_2 ($1.3\pm0.3\times10^{-3}$) and H_2 (7×10^{-6} to -2×10^{-5}) mixing ratios. Thick and thin lines are associated to column abundances of 11 pr-μm and 2 pr-μm, respectively. Solid, dashed, and dot-dashed lines are associated to H_2O mixing ratio scale heights h of resp. ∞ (uniform), 8, and 3 km, respectively.

Infrared measurements of reflected sunlight performed from Phobos 2 show that the dayside water vapor column above the ground in equatorial regions is ≈11 pr-μm [*Rosenqvist et al.*, 1992]. The quantity of H_2O exchanged between the ground and the atmosphere during the diurnal cycle of the boundary layer is expected to be very small, ≈1 pr-μm [*Zent et al.*, 1993], and Phobos values may therefore be considered as day/night averaged values. The equilibrium O_2 mixing ratio obtained for 11 pr-μm of H_2O uniformly mixed with altitude is too small (Figure 5). By the way, we must emphasize that all water vapor profiles used in the present study, uniformly mixed or not, are systematically constrained by saturation. A more consistent picture is obtained by decreasing the scale height h of the H_2O mixing ratio from infinity (constant mixing ratio) down to 8 km and, further, to 3 km (Figure 5). Limb spectra recently obtained by *Clancy et al.* [1992] in similar seasonal conditions (L_S=340-350°) yield a relatively small equatorial water vapor abundance of 4.5 pr-μm, decreasing to 3 pr-μm at midlatitudes, although they were obtained one Martian year later (December 1990). These measurements suggest a nearly uniformly mixed water vapor profile up to 45-km altitude. A value of the water vapor amount of 1.3 pr-μm was deduced at the same period from millimeter wave measurements of *Encrenaz et al.* [1994]. The case of a uniformly mixed water vapor profile of column 2 pr-μm is shown in Figure 5.

Phobos ISM measurements of water vapor total amount [*Rosenqvist et al.*, 1992] coincide rather well with the globally averaged value of the H_2O abundance, as derived from Viking water vapor mapping [e.g. *Jakosky and Barker*, 1984]. Nevertheless, small amounts measured by Earth-based spectroscopy in 1991 near northern spring equinox (≈2-3 pr-μm), suggest that the mean water vapor amount estimated from both Viking and Phobos observations could be overestimated. An upper limit on K of ≈2×10^7 cm^2 s^{-1} may be derived from Figure 5. Unrealistic cases corresponding to a weak column density of 2 pr-μm with small scale heights of 8 km and 3 km are also shown in Figure 5.

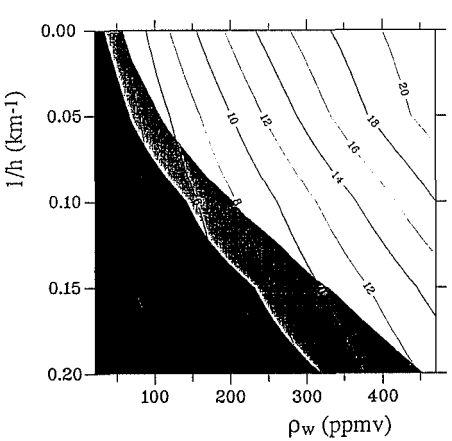


Figure 6. Permitted domain of the near surface water vapor mixing ratio ρ_W and 1/h (h: in the scale height of the water vapor mixing ratio), as derived from the analysis of O_2, for K=5x10^6 cm^2 s^{-1}. The gray area corresponds to expected values of the O_2 mixing ratio (1.3±0.3x10^{-3}). The white area corresponds to an O_2 mixing ratio larger than 1.6x10^{-3} and the black area to an O_2 mixing ratio lower than 10^{-3}. Lines of constant water vapor column density (in pr-μm) are also plotted.

As may be seen in Figure 5, a uniformly mixed water vapor profile with a column amount of 11 pr-μm yields too small values of the O_2 mixing ratio. Figure 6 shows the range of admissible H_2O mixing ratio ρ_W at the surface and inverse 1/h of the scale height h of ρ_W, that is, resulting in an admissible equilibrium value of O_2. The gray area shows the domain for which the calculated O_2 mixing ratio is consistent with the observational range of 1.3±0.3x10^{-3}. The permitted band crosses the isoline corresponding to a water vapor total column c of ≈10 pr-μm for values of 1/h larger than 0.15 km^{-1}, that is, values of h in the range 5-6 km. The scale height of the water vapor absolute concentration is therefore ≈3 km, by assuming an atmospheric scale height of ≈10 km. 90% of the water vapor molecules are concentrated in the first 5 km above the ground, that is, about one half scale height. The low required scale height may be attributed to the fact that H_2O molecules located below ≈5 km altitude are protected from photodissociation by the overlying CO_2 atmosphere. This effect appears clearly in Figure 7, where the total photodissociation rate is plotted versus the total column for various mixing ratio scale heights. For example, for a fixed column of 10 pr-μm, the photodissociation rate is divided by 3 when decreasing h from infinity (constant mixing ratio) down to 3 km. It may be seen from Figure 6 that ρ_W cannot be constant with altitude except for a very low value of c, smaller than ≈4 pr-μm.

The equilibrium mixing ratio of H_2 does not vary much with K (Figure 5). On the other hand, the H_2 equilibrium value is expected to be still more sensitive to the H_2O vertical profile than the O_2 equilibrium value, as already mentioned. Figure 8 shows the range of admissible ρ_W and 1/h derived from H_2 analysis. As for O_2, ρ_W cannot be constant with altitude except for a very low value of c, smaller than 3 pr-μm. By assuming an averaged water vapor column larger than 9 pr-μm, a systematic exploration of parameters shows that h must be in the range 5-6 km, increasing up to 5-9 km for c=6-9 pr-μm, and 5-50 km for c=3-6 pr-μm. Any excess of water with respect to a nominal uniformly mixed profile of column ≈3 pr-μm must be concentrated in the first half atmospheric scale height above the ground.

The intersection between the domains of admissible couples (ρ_W, 1/h) derived, on the one hand from O_2 analysis and, on the other hand from H_2 analysis, defines a third domain, inside of which both O_2 and H_2 equilibrium values are within their expected ranges. Consistency between O_2 and H_2 is obtained only in a narrow band (Figure 9). This band crosses the line of constant column c=8 pr-μm for h≈6 km. It becomes parallel to lines of constant column in the region of the plane corresponding to h>20 km, c=3±0.5 pr-μm. Therefore, the vertical profile of H_2O above the first half scale height must coincide with a uniformly mixed profile of column 3 pr-μm, any additional water vapor being concentrated in the first half scale height above the ground. Different tests have been performed, by increasing or decreasing K with respect to the value of 5x10^6 cm^2 s^{-1} used for calculations presented in Figure 9. It may be shown that K cannot be decreased below ≈4x10^6 cm^2 s^{-1}; otherwise, there

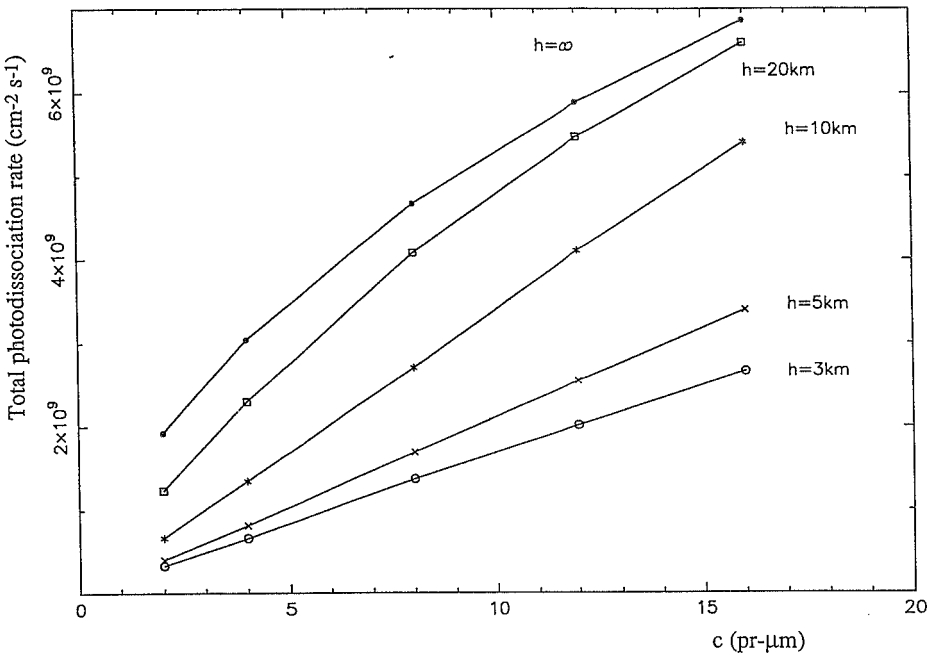


Figure 7. Integrated photodissociation rate of H_2O as a function of the water vapor column for different scale heights of the water vapor mixing ratio in the 3-∞ km range.

should be no consistent field of solutions. As previously mentioned, values larger than $\approx 2 \times 10^7$ cm^2 s^{-1} are excluded, since, in this case, no solution exists for O_2. The widest range of solutions is obtained for $K \approx 5 \times 10^6$ cm^2 s^{-1}. By fixing K below 35 km to a nominal value of 1.5×10^6 cm^2 s^{-1}, it is also possible to analyze the effect of a change of K restricted to the middle atmosphere. The widest range of solutions is obtained for a slightly larger value of K ($\approx 7 \times 10^6$ cm^2 s^{-1}). There is no more upper limit in this case. The relation between h and c is not significantly modified. Results are summarized in Table 3.

Concerning CO, whose mixing ratio is fixed at its observed value (8×10^{-4}), it is emphasized that the loss to production ratio is in the range ≈0.93-1.17 in the whole domain of K and c considered in the previous study. The upper value of 1.17 becomes 1.09 by excluding very low values of the eddy mixing coefficient (2×10^5 cm^2 s^{-1}). The balance of CO, whose lifetime is 5±1 years, is realized with an accuracy better than 10%. This behavior is expected from a theoretical analysis of processes regulating CO [*Chassefière*, 1991]. Isolines corresponding to different values of the relative difference between the loss to production ratio of CO and unity are plotted in the plane [K,c] (Figure 10). The range of K for which the balance of CO is realized with an accuracy better than 5% is $\approx 10^6$-10^8 cm^2 s^{-1} with a most likely value of 10^7 cm^2 s^{-1} (isoline at 1%). Although no strong constraint may be derived on K or c from the necessity to account for an accurate CO balance, a consistent picture of

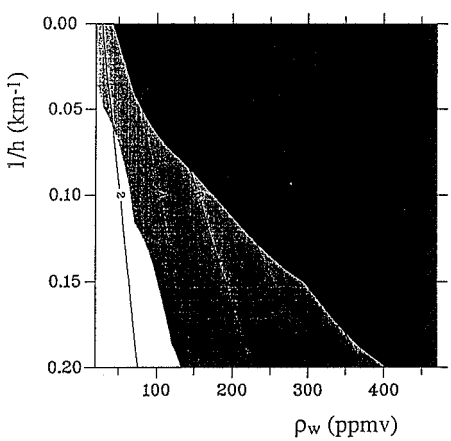


Figure 8. Same as Figure 6 for H_2. The gray area corresponds to expected values of the H_2 mixing ratio (7×10^{-6} to -2×10^{-5}). The white area corresponds to an H_2 mixing ratio lower than 7×10^{-6} and the dark area to an H_2 mixing ratio larger than 2×10^{-5}.

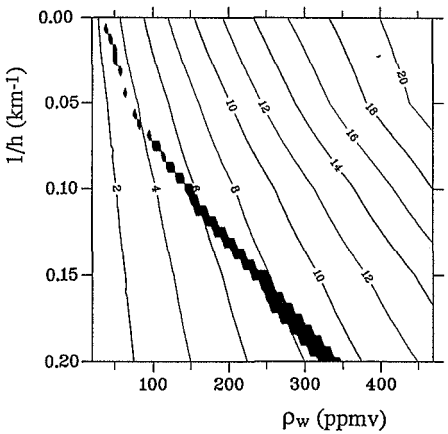


Figure 9. Intersection of domains obtained separately for O_2 and H_2.

balances governing long-lifetime constituents (CO, O_2, H_2) may be found in one-dimensional photochemical models by assuming a small average water vapor column of ≈3 pr-μm. By the way, it must be noted that present results differ in a significant way from those obtained by *Atreya and Gu* [1994] which suggested that heterogeneous chemistry is needed for the stability of the Martian atmosphere. However, the presence of heterogeneous chemistry such as adsorption of HO_x on aerosols [*Anbar et al.*, 1993], or sticking of molecules on suspended dust particles [*Rosenqvist et al.*, 1992], could play a significant role for the variability of the horizontal and vertical distributions of some atmospheric constituents.

Discussion

The eddy diffusion coefficient and its variability are summarized in Table 4. The mean observational value is somewhat smaller than the one inferred from the model. A possible reason should be the regular occurrence of global dust storms. Suspended dust particles reach heights in excess of 40 km [*Leovy et al.*, 1972; *Haberle et al.*, 1982] in a few days. Then the storm covers much of the planet for weeks. The mean vertical velocity of particles is thus of the order of 10-20 cm s^{-1} during storms as compared to 1-2 cm s^{-1} during still conditions. The prestorm clear conditions are typically restored after several months. Consequently, the regular occurrence of global and local dust storms should increase the yearly diffusion by a factor of at least five as compared to K for clear conditions only, such as during the Phobos observations.

A direct consequence of the present work is that the best value of K to be used in all works resorting to one-dimensional modeling (long-term evolution, escape, surface/atmosphere exchanges) is 10^7 cm^2 s^{-1} rather than 1.5×10^6 cm^2 s^{-1}.

Another interesting consequence is the possibility that water vapor should be confined in the first half atmospheric scale height above the ground. There are actually two possibilities: either the globally averaged water vapor column is around ≈3 pr-μm, that is, definitely lower than values derived from Viking and Phobos measurements (≈10 pr-μm), or it is larger than 3 pr-μm. However, in the latter case, the water vapor in excess must be confined in the first half scale height. As shown in the present study, water vapor molecules present in the low atmosphere (z<5 km) are not efficiently dissociated by solar photons. The water vapor content required for global photochemical balance is therefore ≈1-2 pr-μm above ≈5 km. A larger mixing ratio below this level, in the boundary layer, cannot be ruled out a priori. It was proposed by *Hess* [1976] that the diurnal sublimation-condensation cycle of water could prevent water vapor from being uniformly distributed up to the condensation level, but mainly concentrated in the first scale height of the atmosphere, implying some diurnal variability of H_2O. It is likely not the case, as shown by the recent calculations of *Zent et al.* [1993]. No marked diurnal variation of the water vapor mixing ratio in the boundary layer (z<4 km) is expected from a detailed one-dimensional model of the exchange of H_2O between the atmosphere and the regolith, with no more than ≈1 pr-μm exchanged per sol (Martian solar day).

The comparison between Phobos solar occultation and infrared imaging data, which suggests a value of the water vapor scale height of ≈8 km, does not seem either to support the idea that water vapor could be confined in the first half

Table 3. Permitted Ranges of the H_2O Mixing Ratio Scale Height h as a function of $K(z)$

K, cm^2 s^{-1}			H_2O Column		
0-35km	35-80 km	>80km	[3-6]pr-μm	[6-9] pr-μm	>9pr-μm
	4×10^6 $\forall z$		h<50 km	h<9 km	h<6 km
	1×10^7 $\forall z$		h<40 km	h<8.5 km	-
	2×10^7 $\forall z$		h<20 km	-	-
1.5×10^6	3×10^6	1×10^7	h<50 km	h<9.5 km	h<6 km
1.5×10^6	7×10^6	1×10^7	h<40 km	h<8.5 km	h<6 km
1.5×10^6	1×10^7	1×10^7	h<40 km	h<8.0 km	-
1.5×10^6	5×10^7	1×10^7	h<20 km	h<7.5 km	-

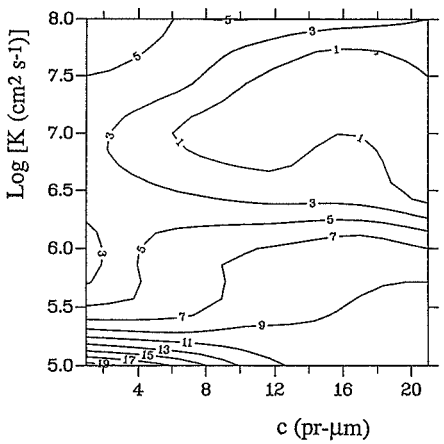


Figure 10. Isolines of constant departure (in percent) of the loss to production ratio of CO from unity as a function of the eddy diffusion coefficient and the water vapor column amount (h=∞).

scale height. It is emphasized that this is precisely the dust scale height in the low atmosphere [*Chassefière et al.*, this issue], which agrees with results of *Davies* [1979]. There was no proof for a low-altitude water vapor enhancement in millimetric data of *Clancy et al.* [1992]. Water vapor was proven to be well mixed, and no significant diurnal variation of the water vapor column was found.

The simplest way to reduce the averaged column above 5 km should be to assume a low mean saturation level (≈5-10 km), in strong disagreement with Mariner 9 and Viking data [*Jakosky and Haberle*, 1992]. Present results may perhaps be connected to the apparent lowering of the altitude level where water vapor is concentrated with time, from northern winter solstice to spring and summer seasons, as derived from Viking water vapor mapping spectrometer data [*Jakosky and Haberle*, 1992]. This phenomenon is not well understood. It might be due to a continuous supply of water from the regolith to the atmosphere during spring and summer. By assuming a typical eddy mixing coefficient of 1.5×10^6 cm^2 s^{-1} (Table 4) in the low atmosphere, the characteristic eddy mixing time is found to be 1 week. In order that a marked vertical gradient of the H_2O mixing ratio would be maintained over one half Martian year, a total amount of ≈500 pr-μm must therefore be transferred from the regolith to the atmosphere in northern hemisphere during spring and summer time. By assuming, with *Zent et al.* [1993], a typical concentration of regolith $H_2O \approx 2$ kg m^{-3}, it is equivalent to the H_2O quantity present in the first ≈25 cm below the surface. The corresponding seasonal flux is ≈1.5 10^{-3} kg m^{-2} day^{-1}, that is, a value of the same order as diurnal fluxes calculated by *Zent et al.* [1993] in their Figure 7. Removal of water molecules by meridional diffusion or transport of water ice clouds toward polar regions must occur on a global scale in order to balance the H_2O input flux.

Table 4. Eddy Diffusion Coefficient as a Function of Altitude

Altitude Range, km	Scale	Mean Value, 10^6 cm^{-2} s^{-1}	Uncertainty (or Variability)
≈10-35		1.5	Factor 2
≈35-65	meso	1.5	Factor 10
>≈65	Equat/Equinox	1.5exp[(z-65)/13.25]	unknown
0-80	global	10	Factor 2-3

Acknowledgments. This research was supported by the Programme National de Planétologie of Institut National des Sciences de l'Univers. We are indebted to N. Rey for her help in the preparation of the camera-ready manuscript.

References

Anbar, A.D., M.T. Leu, H.A. Nair, and Y.L. Yung, Adsorption of HOx on aerosol surfaces: Implications for the atmosphere of Mars, *J. Geophys. Res.*, *98* 10,933-10,940, 1993.

Anderson, D.E., Mariner 6, 7, and 9 ultraviolet spectrometer experiment: Analysis of hydrogen Lyman alpha data, *J. Geophys. Res.*, *79*, 1513-1518, 1974.

Atreya, S.K., and Z.G. Gu, Stability of the Martian atmosphere: Is heterogeneous catalysis essential ? *J. Geophys. Res.*, *99*, 13,133-13,145, 1994.

Barker, E.S., Detection of molecular oxygen in the Martian atmosphere, *Nature*, *238*, 447-448, 1972.

Blamont, J.E., and E. Chassefière, First detection of ozone in the middle atmosphere of Mars from solar occultation measurements, *Icarus*, *104*, 324-336, 1993.

Carleton, N.B., and W.A. Traub, Detection of molecular oxygen on Mars, *Science*, *177*, 988-992, 1972.

Chassefière, E., Photochemical regulation of CO on Mars, *Geophys. Res. Lett.*, *18*, 1055-1058, 1991.

Chassefière, E., J.E. Blamont, V.A. Krasnopolsky, O.I. Korablev, S.K. Atreya, and R.A. West, Vertical structure and size distributions of Martian aerosols from solar occultation measurements, *Icarus*, *97*, 46-69, 1992.

Chassefière, E., P. Drossart, and O. Korablev, Post-Phobos model for the altitude and size distribution of dust in the low Martian atmosphere, *J. Geophys. Res.*, this issue.

Clancy, R.T., A.W. Grossman, and D.O. Muhleman, Mapping of Mars water vapor with the Very Large Array, *Icarus*, *100*, 48-59, 1992.

Davies, D.W., The vertical distribution of Mars water vapor, *J. Geophys. Res.*, *84*, 2875-2879, 1979.

Encrenaz, Th., E. Lellouch, J. Cernicharo, G. Paubert; and S. Gulkis, A tentative detection of the 183 GHz water vapor line in the Martian atmosphere: constraints upon the H_2O abundance and **vertical distribution, *Icarus*, in press, 1994.**

Haberle, R.M., C.B. Leovy, and J.B. Pollack, Some effects of global dust storms on the atmospheric circulation of Mars, *Icarus*, *50, 322--367*, 1982.

Hess, S.L., The vertical distribution of water vapor in the atmosphere of Mars, *Icarus*, *28*, 269-278, 1976.

Hess, S.L., J.A. Ryan, J.E. Tillman, R.M. Henry, and C.B. Leovy, The annual cycle of pressure on Mars measured by Viking landers 1 and 2, *Geophys. Res. Lett.*, *7*, 197-200, 1980.

Hunten, D.M., Aeronomy of the Lower Atmosphere of Mars, *Rev. of Geophys.* **529-535, 1974.**, *12*, 1974

Jakosky, B.M., and E.S. Barker, Comparison of ground-based and Viking orbiter measurements of Martian water vapor: Variability of the seasonal cycle, *Icarus*, *57*, 322-334, 1984.

Jakosky, B.M., and R.M. Haberle, The seasonal behavior of water on

Mars, in *Mars*, edited by H.H. Kiefer et al., pp. 969-1016, University of Arizona Press, Tucson, 1992.

Kahn, R., Ice haze, snow, and the Mars water cycle, *J. Geophys. Res.*, *95*, 14,677-14,693, 1990.

Kong, T.Y., and M.B. McElroy, Photochemistry of the Martian atmosphere, *Icarus*, *32*, 168-189, 1977.

Korablev, O.I., V.A. Krasnopolsky, A.V. Rodin, and E. Chassefière, Vertical structure of the Martian dust measured by the solar infrared occultations from the Phobos spacecraft, *Icarus*, *102*, 76-87, 1993.

Krasnopolsky, V., *Photochemistry of the Atmospheres of Mars and Venus*, Springer Verlag, New York, 1986.

Krasnopolsky, V.A., Solar cycle variations of the hydrogen escape rate and the CO mixing ratio on Mars, *Icarus*, *101*, 33-41, 1993a.

Krasnopolsky, V.A., Photochemistry of the Martian atmosphere (mean conditions), *Icarus*, *101*, 313-332, 1993b.

Krasnopolsky, V.A., O.I. Korablev, V.I. Moroz, A.A. Krysko, J.E. Blamont, and E. Chassefière, Infrared solar occultation sounding of the Martian atmosphere by the Phobos spacecraft, *Icarus*, *94*, 32-44, 1991.

Leovy, C.B., G.A. Briggs, A.T. Young, B.A. Smith, J.B. pollack, E.N. Shipley, and R.L. Wildey, The Martian atmosphere: Mariner 9 television experiment progress report, *Icarus*, *17*, *373-393*, 1972.

Liu, S.C., and T.M. Donahue, The regulation of hydrogen and oxygen escape from Mars, *Icarus*, *28*, 231-246, 1976.

Martin, T.Z., A.R. Peterfreund, E.D. Miner, H.H. Kieffer, and G.E. Hunt, Thermal infrared properties of the Martian atmosphere, 1 Global behavior at 7, 9, 11 and 20 μm, *J. Geophys. Res.*, *84*, 2830-2842, 1979.

Moreau, D., L.W. Esposito, and G. Brasseur, The chemical composition of the dust-free Martian atmosphere: Preliminary results of a two-dimensional model, *J. Geophys. Res.*, *96*, 7933-7945, 1991.

Nier, A.O., and M.B. McElroy, Composition and structure of Mars' upper atmosphere: results from the neutral mass spectrometers on Viking 1 and 2, *J. Geophys. Res.*, *82*, 4341-4349, 1977.

Parkinson, T.D., and D.M. Hunten, Spectroscopy and aeronomy of O_2 on Mars, *J. Atmos. Sci.*, *29*, 1380-1390, 1972.

Rosenqvist, J., P. Drossart, M. Combes, T. Encrenaz, E. Lellouch, J.P. Bibring, S. Erard, Y. Langevin, and E. Chassefière, Minor constituents in the Martian atmosphere from the ISM/Phobos experiment, *Icarus*, *98*, 254-270, 1992.

Seiff, A., Post-Viking models for the structure of the summer atmosphere of Mars, *Adv. Space Res.*, 2, 3-17, 1982.

Shimazaki, T., Photochemical stability of CO_2 in the Martian atmosphere: Reevaluation of the eddy diffusion coefficient and the role of water vapor, *J. Geomagn. Geoelectr.*, *41*, 273-301, 1989.

Théodore, B., E. Lellouch, E. Chassefière, and A. Hauchecorne, Solsticial temperature inversions in the Martian middle atmosphere: Observational clues and 2-D modelling, *Icarus*, *105*, 512-528, 1993.

Trauger, J.T., and J.L. Lunine, Spectroscopy of molecular oxygen in the atmospheres of Venus and Mars, *Icarus*, *55*, 272-281, 1993.

Zent, A.P., R.M. Haberle, H.C. Houben, and B.M. Jakosky, A coupled subsurface-boundary layer model of water on Mars, *J. Geophys. Res.*, *98*, 3319-3337, 1993.

E. Chassefière, Service d'Aéronomie du CNRS, BP 3, 91371 Verrières-le-Buisson Cedex, France

J. Rosenqvist, DESPA, Observatoire de Paris--Meudon, 92195 Meudon Cedex, France (e-mail: rosenqvist@mesiob.obspm.fr)

(Received February 24, 1994; revised November 18, 1994; accepted January 25, 1995.)

Page intentionally left blank

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5553–5563, MARCH 25, 1995

A case for ancient evaporite basins on Mars

Randall D. Forsythe
Department of Geography and Earth Sciences, University of North Carolina at Charlotte, Charlotte

James R. Zimbelman
Center for Earth and Planetary Studies, National Air and Space Museum, Smithsonian Institution, Washington, DC

Abstract. Observations indicate that a Martian analog to the Earth's salt pans and saline lakes of arid regions may have existed in crater-basins during Mars' early (Noachian) epoch. Terraced and channelized crater-basins point to ponding of surface water as well as possible prolonged and evolving base levels. In addition, supportive (evaporite basin) analogs are offered for three other morphologic features of Martian crater-basins. An evaporite basin model for crater-basins on Mars has major implications for the mechanical, chemical, and even biological processes that potentially have operated in Mars' past, and represent a spectrum of potential mineral resources.

Introduction

There is irrefutable evidence for the presence of water ice on Mars. While the surface pressure of the Martian atmosphere is only 0.6% that of Earth, it contains a precipitable 15 μm (in a global layer) of water vapor [*Farmer et al.*, 1977]. The polar deposits today are estimated to include 2 x 10^6 km^3 of water ice [*Carr*, 1986], and variable lines of evidence suggest the presence of ice-rich frozen ground at middle to high latitudes [*Fanale*, 1976; *Squyres*, 1989]. In addition to the possibility of near-surface ground ice today, there is a growing body of work that supports an active role for water in the shaping of Martian landforms during its past epochs [*Squyres*, 1989]. However, since present atmospheric conditions preclude the surface stability of either liquid water at all latitudes or ice at low latitudes, there have emerged two general points of view in attempts to explain the origin of fluvial and other water-related landforms.

One set of interpretations attempts to minimize the need for dramatically different climatic conditions than those found today. The ancient valley networks, documented to exist within the highlands of Mars' southern hemisphere, are seen as most likely related to sapping from groundwater systems [*Pieri*, 1980]. Erosional features in the old southern highlands are typically much too small to be dated accurately by impact crater density, but there is no evidence to suggest that the erosion in the highlands greatly postdates the formation of the cratered terrain during the Noachian epoch on Mars [*Tanaka*, 1986]. The major outflow channels fringing the boundary between the southern hemisphere's highlands and the lowlands of the northern hemisphere can be related to catastrophic flooding [*Baker and Milton*, 1974; *Baker*, 1982] of water possibly released over a prolonged period of time from permafrost-confined groundwater systems [*Carr*, 1979], and surface water lakes, if they existed, were potentially coated with thick layers of ice [*Nedell et al.*, 1987]. Common to these models is the existence of an early Martian geologic epoch with ground water sources (of uncertain origin) that over time have become inactive in low-latitude areas.

From a contrasting perspective, other interpretations require past atmospheric conditions to vary dramatically from those of today, and potentially include the stability of water and ice to equatorial latitudes [*Masursky et al.*, 1977; *Pollack et al.*, 1987; *Baker et al.*, 1991; *Parker et al.*, 1993]. These views which generally beg warmer and wetter climatic models explicitly (or implicitly) invoke mechanisms for recharge directly to the Martian surface, and imply a complex hydrologic cycle once operated on Mars. Ancient valley networks, particularly the degraded ones, are interpreted as runoff channels requiring maintenance of high water tables [*Baker and Partridge*, 1986; Brackenridge, 1990], the outflow channels are visualized as potentially debouching into ancient seas or oceans [*Gulick and Baker*, 1990] and an active hydrologic cycle is seen to potentially persist well beyond Mars' earliest (Noachian) epoch.

Here new observations and arguments are presented, based on morphologic comparisons of Martian crater-basins to terrestrial drainage basins in arid and semiarid areas, and used to advance further the hypothesis for a complex hydrologic cycle in Mars' past that may have included open hydraulic conditions between surface water and groundwater and mechanisms for recharging these reservoirs.

A Case for Base Level and Groundwater-Controlled Drainage on Mars

In the absence of recharge to an unconfined groundwater system, the upper surface of groundwater (i.e., the water table) relaxes to a global minimum equipotential surface, which on Earth is approximated as mean sea level [*Freeze and Cherry*, 1979]. However, the hydrologic processes of precipitation, infiltration, and base flow permit on Earth a dynamic equilibrium between surface topography and the underlying water table. The equilibrium is established by the manner in which the minimum potential of hydraulic head (sea level) is also the base level, below which fluvial erosion cannot occur [*Bloom*, 1978], and inversely, the manner in which local streams, which are responsible for eroding and shaping topography, limit the rise of the water table during recharge periods. On a more regional

Paper number 95JE00325.
0148-0227/95/95JE-00325$05.00

level, the same boundary conditions are applicable to internally drained depressions [*Bloom*, 1978] ,wherein a local base level for fluvial erosion is represented by the corresponding minimum of the water table within the basin.

Partially because of the role of groundwater in determining base levels, streams undergo systematic changes in their channel morphologies from headwater ephemeral regimes (generally without significant floodplains) to base level perennial regimes (with lower gradients and floodplains) [*Leopold et al.*, 1964]. In the latter, periods of climatic and watershed stability (e.g., tectonic quiescence) permit streams to be "graded" or in a state of quasi-equilibrium [*Davis*, 1902; *Mackin*, 1948; *Strahler*, 1952], and it is common during these periods of stability for processes responsible for the lateral migration of channels to be more efficient than the processes of downcutting. The result is the formation of floodplains. Floodplains formed in this fashion, after being subjected to a period of downcutting, are left as terraces or benches along the margins of the valley. Alternative (i.e., nonterrestroid) mechanisms for the formation of fluvial valley terraces on Mars remain possible, perhaps (hypothetically speaking) related in some fashion to changing discharges over an ice-rich permafrost. However, if the ancient valley networks of Mars' southern highlands were produced by fluvial erosion occurring under Earthlike conditions, then similar constraints and conditions can be postulated. One condition of such an analog is that Martian drainage systems would also represent an equilibrium between surface and groundwater systems. The important point being an Earthlike hydraulic coupling between ground water and surface water systems in Mars' past is a plausible hypothesis if geomorphic evidence exists for fluvial drainage under steady base level conditions. Below we describe one ancient Mars drainage basin for which there is evidence in support of fluvial and lacustrine erosion under base level conditions that was sufficiently long-lived in nature to produce what are here interpreted as fluvial and lacustrine terraces.

The middle latitude to equatorial region of the southern highlands has been one of the primary areas from which evidence has emerged for Noachian fluvial erosion [e.g., *Baker and Partridge*, 1986; *Pieri*, 1980] and extensive resurfacing [*Jones*, 1974; *Chapman and Jones*, 1977; *Craddock and Maxwell*, 1990, 1993]. Within the Memnonia Northwest quadrangle a segment of a north trending early drainage system, located approximately 200 km south of the highlands/lowlands boundary at 174.6°W, 14.6°S, has been preserved that contains both fluvial and lacustrine features that are interpreted here on geomorphic grounds to support the existence of base level-controlled fluvial

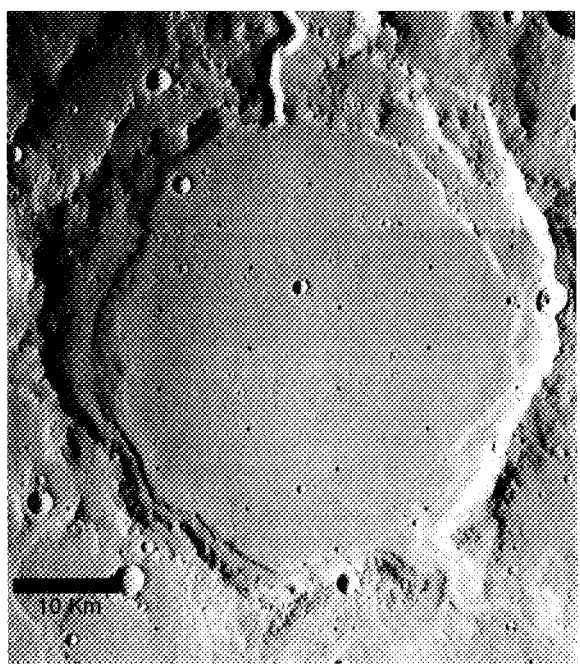


Figure 1a. A terraced crater-basin south of Mars' dichotomy boundary with both an inflowing (lower right) and outflowing (top center) channel. Note the trace of the main terrace up the inflow channel, and the lower level of the outflow channel. Scale bar is 10 km. (Viking Orbiter frame 438S12, centered at 174.6°W, 14.6°S).

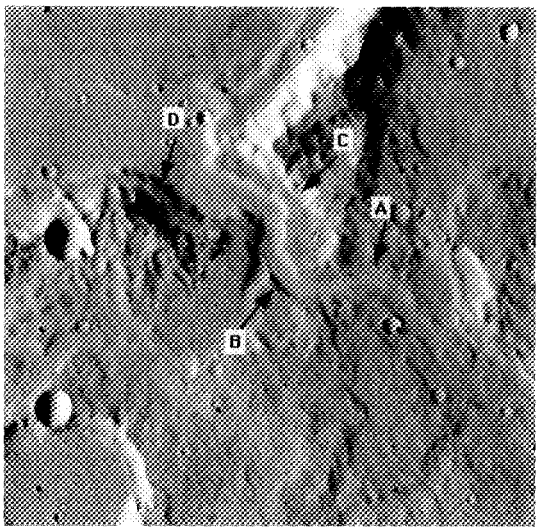


Figure 1b. Closeup of the area of terrace development in the inflowing channel (Viking Orbiter frame 437S15). Localities A - D mark areas near meanders with differential terrace development (see text).

and lacustrine systems. The main fluvial channel can be seen to drain into, and away from, a crater basin. Moderate resolution (63 m/pixel) Viking images of this basin (Figure 1a) shows it to have a polygonal ground plan with an average diameter of approximately 35 km, and a crater rim that, from shadow length measurements, is estimated to rise 0.7 to 1 km above the floor. A prominent terrace or bench, at a height of 200-220 m above the basin floor, is traceable around the entire basin, and extends up into the inflowing channel (Figures 1a and 1b). In addition, an upper smaller bench is traceable just above the main terrace, and two additional, but poorly preserved, erosional benches may exist along the lower slopes (Figure 1c).

The origins of these terraces are of paramount importance to the interpretation that this channel and basin represent a base level-controlled drainage system akin to terrestrial watersheds. There are two questions related to potential ambiguities to be addressed: (1) Are, in fact, the benches of fluvial and lacustrine origin such that their elevations can be interpreted as former channel floodplain levels, and lake stands, respectively; and (2) does the existence and the levels of the postulated lake benches reflect stands of water controlled by a regional groundwater table? These two questions are addressed separately below.

Erosional benches or terraces can only be produced by two sets of alternative mechanisms. Either the terraces represent levels of lithologic differences within the Martian crust that have, in turn, controlled erosion by whatever means (e.g. eolian, mass wasting, or fluvial), or they have been formed by a water-related mechanism [*Bloom*, 1978]. Fluvial terraces within river valleys can be either degradational (strath terrace) or aggradational (fill terrace) but, irregardless, represent uplifted or undercut floodplains [*Leopold et al.*, 1964].

If the terraces seen in Figure 1 are due to nonfluvial or nonlacustrine processes, then these processes should also operate in surrounding areas free from possible fluvial or lacustrine influences. *Craddock and Maxwell* [1993] have demonstrated that pervasive resurfacing has occurred within a marginal (equatorial) province of the highlands, affecting the Martian crust to a depth of 290 to 2300 m. Yet, despite the evidence for extensive erosion, terraces in the Memnonia NW quadrangle (Figure 1) are restricted exclusively to fluvial valleys and degraded crater basins.

A fluvial or lacustrine origin for the terraces seen in Figure 1 is also supported by three specific attributes of the terraces, which would not be easily explained under other mechanisms. First, the main terrace, which can be traced into, and around, the crater basin, ends abruptly at the outlet, despite the continuing relief along the walls of the valley (Figure 1d, location A). A fluvial or lacustrine origin for the terraces explains this asymmetry, since graded fluvial systems are expected to develop preferentially upstream from a base level. Also note that the outlet channel, which is deeply incised at the outlet, opens up to a broad valley floor (i.e., flood plain) at point B (Figure 1d); again a relationship consistent with the existence of local base level influences. Second, if a river terrace represents a former flood plain, then its width should vary in a consistent manner with its precursory flood plain. The latter would broaden as the channel approaches the local base level. Figure 1b shows a closeup of the inflowing channel. In this image one can see that both the channel floor and terrace levels become systematically narrower upstream. Third, the relationships of the valley terrace to channel meanders suggests that they were formed early on in the development of the valley. At locality B (Figure 1b) one can see that the terrace

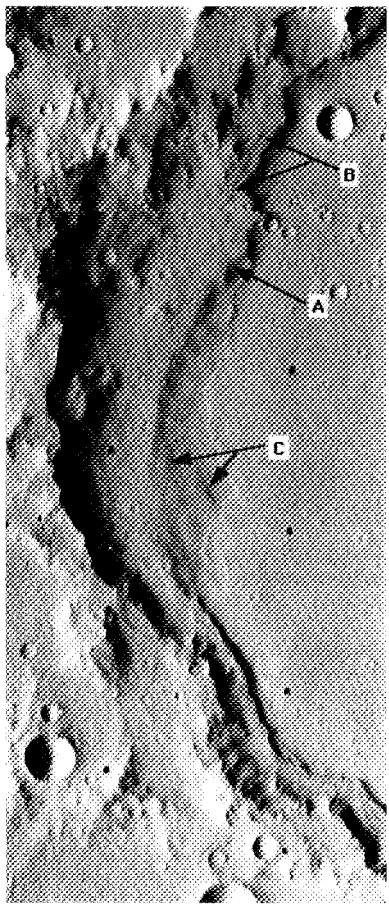


Figure 1c. Closeup of the area of terrace development along the west margins of the basin (Viking Orbiter frame 438S12). Locality A is the uppermost terrace; locality (B) the main terrace, and (C) two other possible lower benches.

does not exist on the outside margin of the meander. Similarly at the next meander, locality C (Figure 1b), the terrace is narrower than it is upstream or downstream. These relationships are expected among river terraces which are incised by later river downcutting. Thus, the terraces described within this system have a number of specific attributes which support a fluvial and lucustrine origin.

The second concern is the possibility that, rather than having an erosional base level defined by a local groundwater depression, it may have been controlled by the existence of a hypothetically impermeable permafrost layer, as envisaged by *Clifford* [1993]. This possibility cannot be ruled out with existing data. However, there are three reasons that support the former hypothesis as being the more plausible alternative in the Memnonia region. There is little doubt that a permafrost zone has existed on Mars, given the abundance and distribution of presumed ice-related geomorphic features [e.g., *Squyres and Carr*, 1986; *Squyres*, 1989] and the estimates of current and past stability of ground ice [*Leighton and Murray*, 1966; *Farmer and Doms*, 1979; *Fanale et al.*, 1986; *Pollack et al.*, 1987b, 1990; *Mellon and Jakosky*, 1993]. However, the hypothesis that a thick permafrost zone could have perched Noachian drainage systems such as the one described above is potentially problematic on three grounds.

First, if permafrost did indeed exist within the southern highlands during the creation of the valley networks, then one should be able to identify features characteristic of geomorphic processes operating in regions with ground ice. As noted above, Mars preserves many geomorphic features in its current polar and midlatitude regions that are interpreted to be related to ground ice (e.g., "terrane softening" of *Squyres and Carr* [1986]). However, ice-free alternatives do exist for some of these geomorphic features [*Zimbelman et al.*, 1989]. The preservation of fluvial features throughout much of the highlands indicates that a sufficient record of past surface processes is available. As pointed out by *Craddock and Maxwell* [1993], the only potential evidence for ice-related processes in the equatorial area are postcrater degradation features (interpreted to be glacial in origin) suggested to be present in some of the more elevated regions by *Baker et al.* [1991] and *Kargel et al.* [1992]. Rampart craters have been cited as possible indicators of subsurface volatiles on Mars, and these craters are present at all latitudes (see *Squyres et al.* [1992] for a detailed discussion). The precise role of volatiles in the rampart crater formation process remains ambiguous [*Squyres et al.*, 1992], but the equatorial rampart craters have been interpreted to indicate that water or ice is present in the subsurface, although at a much greater depth than at midlatitudes [Kuzmin et al., 1988].

Second, while the permafrost zone is commonly described as impermeable, no where on Earth has it been demonstrated that it operates as a pure aquifuge [*Freeze and Cherry*, 1979; *Williams* 1970; *Williams and Smith*, 1989]. While the hydrogeologic literature for permafrost has presented the permafrost zone as impermeable when considering water budgets on timescales pertinent to our understanding of water resources [e.g., *Williams*, 1970], estimates in the laboratory, as well as from field studies, indicate that it is best characterized as an aquitard, or aquiclude, of very low hydraulic conductivity on geologic timescales [*Williams and Smith*, 1989]. The 45 - 50° K difference in the mean annual surface temperatures between the coldest permafrost regions on Earth (-5° to 10°C) and the warmest (equatorial) latitudes on Mars (approximately -55°C) suggest that the hydraulic conductivities of permafrost on Mars under ice-rich conditions would be somewhat lower than that observed for terrestrial conditions. However, the in situ large-scale transmissivities of hydrogeologic units are influenced by both diffuse and nondiffuse flow. Laboratory measurements of the hydraulic conductivity of ice-rich permafrost speak to only the former component.

Third, when surface water persists over permafrost it will have a dramatic affect on the underlying permafrost, producing an extensive zone of thawing both at the top and base of the permafrost zone [*Lauchenbruch*, 1957]. While thermal modelling of these effects for Martian conditions is beyond the scope of this

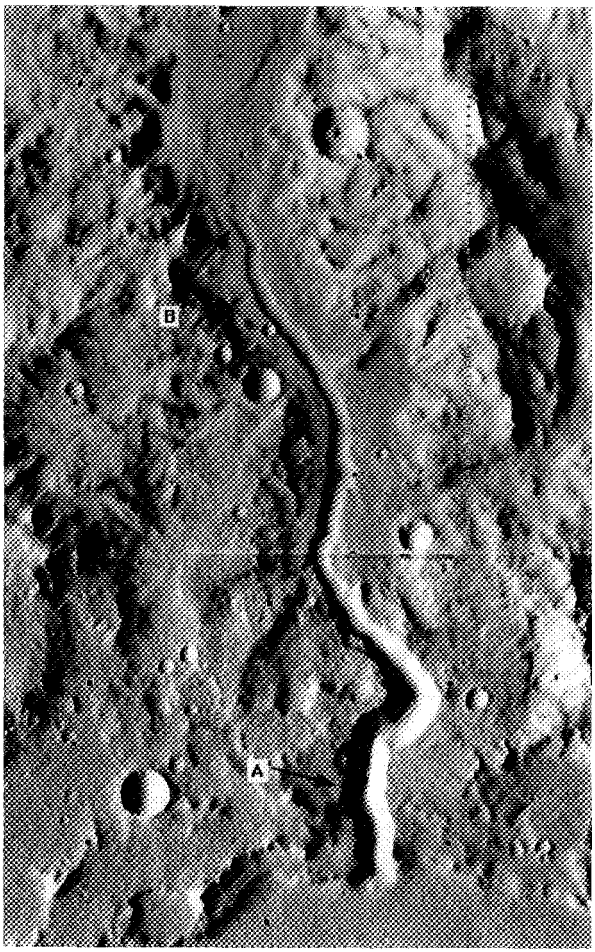


Figure 1d. Closeup of the outflowing channel where no terrace has been developed (location A), and the incised channel opens up downstream to a floodplain (location B) (Viking Orbiter frames 438S11,-12 and -13).

paper, it should be noted that even within the continuous permafrost regime of Alaska (with permafrost thicknesses of several hundred meters), permafrost is either very deep or absent under lakes and large rivers [*Hopkins et al.*, 1955]. Along the latter there are also frequent, but sporadic, areas of groundwater discharge that produce characteristic "aufeis" or icings [*Williams*, 1970] that may reflect the development of nondiffuse flow conditions within the underlying or surrounding permafrost. The colder mean annual surface temperatures on Mars again could have a mitigating effect on comparisons drawn here to terrestrial permafrost regions. Thus, while the case for a permafrost-controlled base level within the drainage basin described above cannot be unequivocally ruled out, it faces a significant number of challenges to proceed with the case for an interacting groundwater and suface water drainage system in the Memnonia region. Here we prefer to leave the valley network drainage systems in direct contact with a Martian regolith, for which porosity estimates at the surface range from 20 to 50% [*Clifford*, 1993], and permeability estimates range from 10 to 103 darcies [*Carr*, 1979; *MacKinnon and Tanaka*, 1989; *Squyres et al.*, 1992]. Under these conditions a hydraulic connection should exist between surface and ground water systems. In summary,

our preferred explanation for the terraces along the drainage system is that they represent erosional base level changes over time associated with a warmer, wetter, early climate, begging a terrestrial watershed analog for such a Martian drainage system.

Morphologic Features of Crater Basins Consistent With Terrestrial Evaporite Basins

The basin described above is not an isolated occurrence; many other basins exist in the southern highlands near the equatorial boundary that may also have been part of comparable surface drainage systems. Some of these have both inflow and outflow channels (e.g. that described above); others have inflow and no outflow channel [*Grant and Schultz*, 1993, figures 8 and 11] and still others have terraces with neither obvious stream inflow or outflow [Figure 2]. In the latter case, water inputs could be purely by base flow, and losses by evaporation and/or sublimation. These conditions are very common in the Atacama desert salars (discussed below). Given that one accepts the arguments presented above for a fluvial or lacustrine interpretation of these terraces, a steady base level must have operated for some significant geomorphic time period for these terraces to form. To maintain these base levels, in the face of either periodic or steady inputs, requires commensurate outputs, either through baseflow or evaporation (or sublimation), to balance the water budget (i.e. conserve mass). But as argued above, maintenance of a base level given the probable hydrogeologic characteristics of the highlands' regolith is problematic without that base level also being the water table. Given that the terraces exist in topographic depressions, groundwater could have provided seepage to these basins, leaving evaporative or sublimative losses as the most likely means of balancing the water exchanges (implied by the geomorphic features). Losses to the atmosphere by evaporation (or perhaps sublimation) within closed basins necessarily leads to saline conditions [*Papke*, 1986; *Stoertz and Eriksen*, 1974]. Thus, many of these crater basins may be akin to the salt flats of the western United States in that their deposits could record a history of progressive desiccation from saline lake to salt pan conditions in time [*Papke*, 1986]. An additional consideration that may influence lake or groundwater salinities under Martian conditions is the exclusion of solutes through freezing.

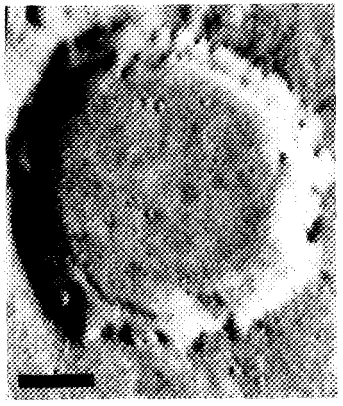


Figure 2. Crater with a prominent internal terrace and no apparent inflow or outflow channel. Scale bar is 5 km. (Viking Orbiter frame 436S15, centered at 176.5°W, 16.0°S).

In addition to terraced crater-basins, all basins of the highlands with depths beneath a local base level would, like terrestrial basins of arid and semiarid areas, accumulate admixtures of wind, stream sediment, colluvium, and evaporites. On Earth these basins are strongly zoned, with evaporite/loess cores, and marginal facies dominated by salt-cemented clastic units [*Papke*, 1986; *Stoertz and Ericksen*, 1974]. Sharp contrasts in albedo are commonly found at the basin margins, where the albedos of salts or wetted salts stand in sharp contrast to that of the clastic materials composing the fringing alluvial or colluvial fans, bahadas, and piedmonts (Figure 3). Several basins on Mars exhibit a narrow curvilinear marginal fringe zonation involving linear zones with higher and/or lower relative albedos than that observed in upland or central basin floor areas (Figure 4). This suggests the presence of a very narrow basin margin deposit with unique spectral characteristics (e.g., mineralogy; see Table 1). Their analogs on Earth (e.g., Death Valley, Mohave Desert, Salar de Atacama, Atacama Desert) typically mark the location of groundwater seepage and/or paleoshorelines.

A further distinctive feature of Salares (salt pans) in the Chilean parts of the Atacama Desert are marginal salt fingers or channels, developed leeward of marginal seeps or springs toward the prevailing wind direction. Winds passing across the marginal zone, where fresh salt is often found, become charged with salt-laden aerosols and etch linear trails in the prevailing wind direction across the salares (Figure 3 and 5). One Martian crater-basin located at 167.15° longitude, -10.26° latitude, in the Memnonia northwest quadrangle exhibits long and narrow high albedo tracks extending into the basin floor deposit (Figure 5). The features appear morphologically distinct from the array of Martian wind features previously described within craters [*Ward*, 1979; *Peterfreund*, 1981], and have aspect ratios much greater than typical Martian slide features. These linear tracks are good candidates for salt fingers.

A Martian analog to terrestrial salares or sebkhas would predict that crater-basins within the Martian highlands that sat at levels near the regional base level would have filled with admixtures of evaporite, eolian, alluvial, and colluvial materials. Two Martian inselbergs formed of partially exhumed crater-basin deposits exist which have appropriate characteristics for an evaporite facies deposit: "White Rock" [*Ward*, 1979; *Williams and Zimbelman*, 1994], and a similar deposit in the Bequeral crater [*Christensen*, 1983]. Both of these Martian inselbergs have relatively high albedos (see Figure 6), but not as high or white as some of the filtered or processed images might lead one to conclude [*Evans*, 1979]. It is, however, fallacious to suggest that because modern perennially flooded salt flats, such as the Bonnevale Salt Flat of Utah, are white that evaporite facies deposits, in general, should also be white [*Clark and Van Hart*, 1981]. Most of the salt playas on Earth that do not receive a seasonal flooding of surface water have reduced albedos (relative to those seasonally flooded). The driest parts of the Atacama Desert, Altiplano, and Puna, show bright, medium, and dark salar surfaces (Figure 3), and in space shuttle color photography display reddish and brownish hues. Since an exhumed salar evaporite deposit is likely to have clastic, eolian, and salt components, its albedo and color would in part be controlled by the nature of its mixed eolian, clastic, and evaporitic facies, as

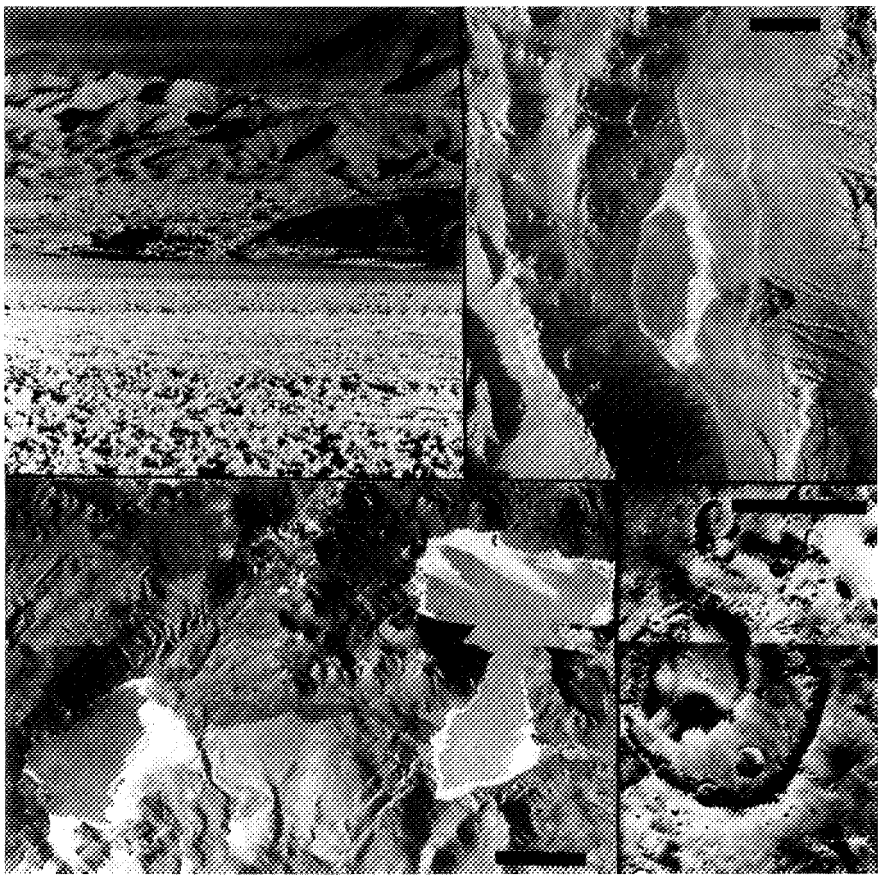


Figure 3. (Top left) Ground photo of the brown to white salt-encrusted surfaces of the Salar de Atacama, northern Chile (68.3°W, 23.5°S), Oasis of Paine in middle ground, Altiplano in background (RDF, taken July 1982). (Top right) Hand-held space shuttle view of the Salar de Punta Negra, Chile (69°W, 24.5°S, NASA 61C-45-018), with bright outer rim of salts and core of dark salt facies. Scale bar is 10 km. (Lower left) Hand-held space shuttle view of evaporite basins in northern Chile with dark central salt facies and bright rims (68.8°W, 25.7°S, NASA STS-23-38-050). (Lower right) Martian crater basin with flat lying dark internal basin fill with a surrounding pediment apron (Viking frames 618A33 and -A35, centered at 322.0°W, 9.0°S). Scale bar is approximately 100 km.

well as the superimposed effects of secondary erosion and weathering of the exhumed deposit. Radiometric data for White Rock indicate higher than Mars-average albedo [*McEwen*, 1987], with red filter reflectivities higher than blue [*Evans*, 1979], consistent with that of many terrestrial evaporite facies deposits.

Terrestrial (nonmarine) evaporite basin deposits are generally well bedded. As such, aerial and space imagery of partially exhumed deposits (Quidham Basin, western China; Tanerouft Basin, southcentral Sahara; Salt Ranges, Atacama Desert) exhibit a fine concentric to convoluted structure produced by the erosional sculpturing of the internal bedding. A similar fine and convoluted layer structure can also be observed in the Bequeral deposit (Figure 6, lower right). Stereo images of White Rock also reveal that the upper erosional surface has a vertically stratified structure [*Williams and Zimbelman*, 1994]. The variation of bedding facies in a nonmarine terrestrial evaporite basin can be used to explain an additional attribute of the Martian inselbergs. Both are relict mounds of basinfill deposits isolated

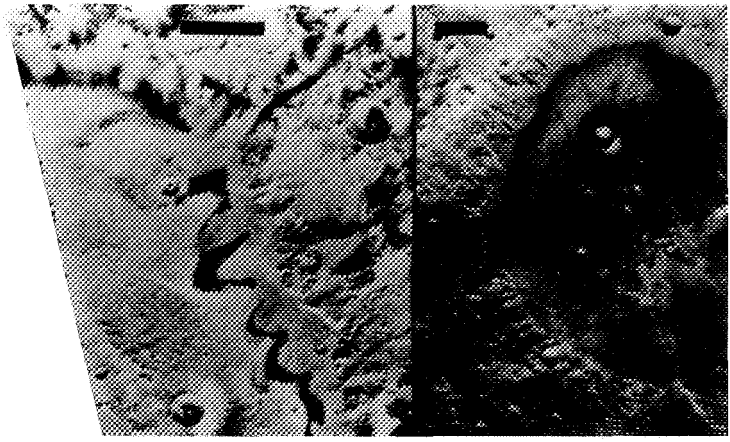


Figure 4. (Left) Margin of a Martian crater-basin with sharp albedo contrasts in the transition zone from the flat lying central facies to the marginal colluvial pediments. Scale bar is 25 km (Viking frame 321A38, centered at 341.8°W, 5.8°S). (Right) Martian crater-basin with a fine bright fringe developed approximately 1 km inward from the basin floor/crater wall margin. Scale bar is 5 km (Viking frame 212A05, centered at 10.2°W, 19°N).

by erosion from the crater-basin margins. Salt-cemented rock (salcrete), while unstable under humid conditions, becomes very stable under arid conditions, and for example, in the hyperarid areas of the Atacama Desert have been used for buildings and roads. Central parts of basins have cores of salt-rich evaporite facies strata originally surrounded by a salt-poor, and clastic-rich, marginal evaporite basin facies, which could have been preferentially left behind as winds deflated these basins over time.

Discussion

An evaporite basin analog for ancient crater-basins on Mars has several implications for future research and exploration efforts. First, from a climatic perspective, the existence of paleolakes and possible large accumulations of evaporite deposits in small basins (e.g., White Rock) reinforces the interpretations that an equilibrium may have existed between surface, groundwater, and atmospheric water reservoirs. This is

Table 1. Authigenic Minerals of Nevada Playas and Atacama Salars [*Papke*, 1976; *Stoertz and Ericksen*, 1974]

	Mineral	Chemical Formula		Mineral	Chemical Formula
Chlorides	Halite	$NaCl$		Natron	$Na_2CO_3{\cdot}10(H_2O)$
	Sylvite	KCl		Thermonatrite	$Na_2CO_3{\cdot}H_2O$
Sulfates	Gypsum	$CaSO_4{\cdot}2(H_2O)$		Pirssonite	$Na_2Ca(CO_3)_2{\cdot}2(H_2O)$
	Basanite	$CaSO_4{\cdot}1/2(H_2O)$		Strontianite	$SrCO_3$
	Anhydrite	$CaSO_4$		Nahcolite	$NaHCO_3$
	Mirabilite	$Na_2SO_4{\cdot}10(SO_4)$	Borates	Borax	$Na_2B_4O_7{\cdot}10(H_2O)$
	Thenardite	Na_2SO_4		Tincalconite	$Na_2B_4O_7{\cdot}5(H_2O)$
	Glauberite	$Na_2Ca(SO_4)_2$		Ulexite	$NaCaB_5O_9{\cdot}8(H_2O)$
	Epsomite	$MgSO_4{\cdot}7(H_2O)$		Colemanite	$Ca_2B_6O_{11}{\cdot}5(H_2O)$
	Aphthitalite	$NaKSO_4$		Searlesite	$NaBSi_2O_6{\cdot}(H_2O)$
	Bleodite	$Na_2Mg(SO_4)_2{\cdot}4(H_2O)$	Nitrates	Soda Niter	$NaNO_3$
	Celestite	$SrSO_4$	Others	Fluorite	CaF_2
Carbonates	Calcite	$CaCO_3$		Phillipsite (other zeolites)	
	Aragonite	$CaCO_3$		PotassiumFeldspar	$KAlSi_3O_8$
	Dolomite	$CaMg(CO_3)_2$		Burkeite	$Na_6CO_3(SO_4)_2$
	Gaylussite	$Na_2Ca(CO_3)_2{\cdot}5(H_2O$		Hanksite	$Na_22KCl(CO_3)_2(SO_4)_9$
	Magnesite	$MgCO_3$			
	Trona Na3	$H(CO_3)_2{\cdot}2(H_2O)$			

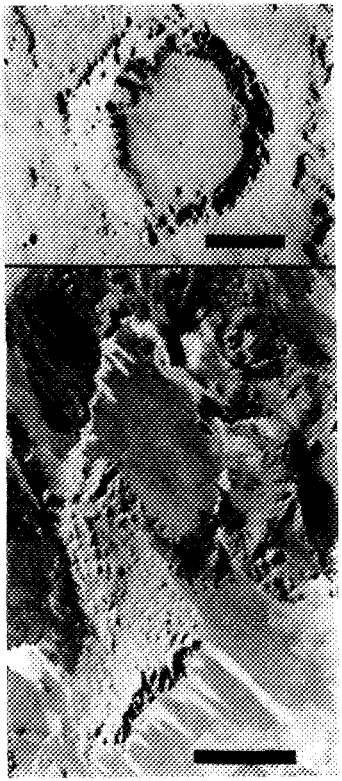


Figure 5. (Top) Mosaic of Viking frames 442S02-4 (centered at 167.15°W, 10.27°S). Along left margin of basin floor bright linear trails of salt are developed inward from the wall. Scale bar is 10 km. (Bottom) Hand-held Space Shuttle photo shows similar linear trails in the Atacama salt basins (67.6°W, 24.3°S), where field work [*Stoertz and Ericksen*, 1974] demonstrates their formation by saltetching. Scale bar is approximately 10 km.

consistent with the interpretations drawn from studies of the Noachian degraded valley networks [*Baker and Partridge*, 1986; *Gulick and Baker*, 1990].

The potential presence of evaporite deposits on Mars would suggest that, at least locally, groundwater conditions reached saline conditions [*Zent and Fanale*, 1986]. The possible existence of saline groundwater has major chemical and mechanical consequences on the surface processes that have operated in the past. The densities of salt-saturated brines can be approximately 20% heavier, and their dynamic viscosities twofold higher than that of distilled water. The mechanical impact is an increase in the driving forces for mass wasting of slopes, and an increase in the internal bulk dynamic viscosity of "flows". While the effects of salts on surface and shallow subsurface processes have been discussed [e.g. *Clark and Van Hart*, 1981; *Malin*, 1974; *Zent and Fanale*, 1986], these have generally not acknowledged the possible presence of localized evaporite facies deposits. As such, a salt karst alternative has not been included along with discussions of thermo- and carbonate-karst alternatives for many of the sapping-related features and terrains.

The evaporite basin hypothesis can be tested by several means, both with existing data and with improved data from future missions. XRF results from the Viking landers are interpreted to indicate the presence of various chloride salts on Mars (*Banin et al.*, 1992), and the present work would suggest that a salt signature might be strongest in the Memnonia region. Even with the limited spatial resolution available to Earth-based instruments, the great spectral resolution now available might be able to detect regional variations in some evaporite concentrations. Viking image data of the cratered highlands outside of Memnonia need to be searched systematically to assess the global distribution of possible fluvial terraces on Mars.

Future missions, particularly the Mars Global Surveyor spacecraft which will refly a large portion of the Mars Observor instruments to Mars, hold great potential for providing further tests to the hypothesis. The 1.4-m resolution of the Mars Observor camera [*Malin et al.*, 1992] may detect deflation pits associated with the "salt finger" weathering, as is observed on evaporite basins in South America, or layering associated with

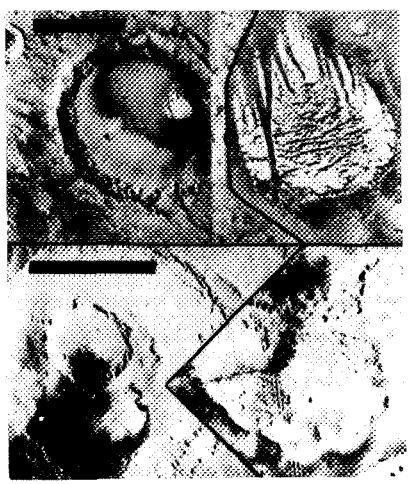


Figure 6. (Top) Views of "White Rock," a relatively high albedo deposit in an unnamed crater southeast of Schalaparelli Impact basin in the Sinus Sabaeus region of Mars. Scale bar is approximately 50 km. Viking Orbiter frames 618A11 (top left) and 826A34-36 (top right), located at 335°W, 8°S. (Bottom) Views of the residual inselberg of a crater-basin deposit within the Bequerel Basin. A convoluted pattern of lines within its surface suggests differential erosion of thinly bedded deposits. Scale bar is approximately 50 km. Viking Orbiter frames 829A37 (bottom left) and 209A07, 9 (bottom right), located at 7.0°W, 21.3°N.

the evaporite deposits themselves. The combined spatial and spectral resolution of the thermal emission spectrometer [*Christensen et al.*, 1992] may be able to detect evaporites at basin margins where eolian erosion has exposed variable albedo materials. Finally, the greatly improved topography which should result from the Mars Observor laser altimeter (*Zuber et al.*, 1992) should be able to determine if terraces within the large crater basins conform to a geopotential surface, as would be expected for a fluvial origin. These new data are crucial to testing a variety of hypotheses related to the history of Mars, such as the evaporite basin hypothesis presented here.

The possible existence of evaporite deposits may at some point in the future have major practical importance for a long-term manned base on Mars. On Earth these deposits contain wide arrays of authigenic sulfates, chlorides, carbonates, borates, and nitrates (Table 1) and are economic sources of sodium chloride, borax, sodium carbonate, sodium sulfate, and lithium, and are potential sources of potash, magnesium, calcium chloride, fluorine, tungsten, uranium, zeolites, nitrates and clays. They may also form basic construction materials for buildings or roads under a hyperarid climatic condition. Finally, in our future efforts to search for past life on Mars, the postulated evaporite basin deposits would be a natural target for paleontologic study.

Acknowledgements. Research was supported by a Smithsonian Senior Fellowship to the first author and aided by resources made available at the Center for Earth and Planetary Studies, National Air and Space Museum. We thank Steve Clifford and an anonymous reviewer for their constructive reviews.

References

Baker, V. R., and D. Milton, Erosion by catastrophic floods on Mars and Earth, *Icarus, 23,* 27-41, 1974.

Baker, V. R., *The Channels of Mars*, University of Texas Press, Austin, 1982.

Baker, V. R. and J. B. Partridge, Small Martian valleys: Pristine and degraded morphology, *J. Geophys. Res., 91,* 3561-3572, 1986.

Baker, V. R., R. G. Strom, V. C. Gulick, J. S. Kargel, G. Komatsu, and V. S. Kale, Ancient oceans, ice sheets and the hydrologic cycle on Mars, *Nature, 352,* 589-594, 1991.

Banin, A., B. C. Clark, and H. Wanke, Surface chemistry and mineralogy, in *Mars* edited by H. H. Kieffer et al., pp. 594-625, University of Arizona Press, Tucson, 1992.

Bloom, A. L., *Geomorphology, A Systematic Analysis of Late Cenozoic Landforms*, 510 pp., Prentice-Hall, Englewood Cliffs, N.J., 1978.

Brackenridge, G. R., The origin of fluvial valleys and early geologic history, Aeolis quadrangle, Mars, *J. Geophys. Res., 95,* 17,289-17,308, 1990.

Carr, M. H., Formation of Martian flood features by release of water from confined aquifers, *J. Geophys. Res., 84,* 2995-3007, 1979.

Carr, M. H. Mars: A water-rich planet, *Icarus, 68,* 187-216, 1986.

Chapman, C. R., and K. L., Jones, Cratering and obliteration history of Mars, *Annu. Rev. Earth Planet. Sci., 5,* 515-540, 1977.

Christensen, P.R., Eolian intracrater deposits on Mars: Physical properties and global distributions, *Icarus, 59,* 496-518, 1983.

Christensen, P. R., et al., Thermal emission spectrometer experiment: Mars Observer mission, *J. Geophys. Res., 97,* 7719-7734, 1992.

Clark, B.C. and D.C. Van Hart, The salts of Mars, *Icarus, 45,* 370-378, 1981.

Clifford, S. M., A model for the hydrologic and climatic behavior of water on Mars, *J. Geophys. Res., 98,* 10,973-11,016, 1993.

Craddock, R. A., and T. A. Maxwell, Resurfacing of the Martian highlands in the Amenthes and Tyrrhena region, *J. Geophys. Res., 95,* 14,265-14,278, 1990.

Craddock, R. A., and T. A. Maxwell, Geomorphic evolution of the Martian highlands through ancient fluvial processes, *J. Geophys. Res., 98,* 3453-3468, 1993.

Davis, W. M., Base-level, grade, and peneplain, *J. Geol. 10,* 77-111, 1902.

Evans, N., White Rock-An Illusion?, Rpt. Planetary Geology Program, 1978-1979, *NASA Tech. Memo., TM-80339,* 28-29, 1979.

Fanale, F. P., Martian volatiles: Their degassing history and geochemical fate, *Icarus, 28,* 179-202, 1976.

Fanale, F. P., J. R. Salvail, A. P. Zent, and S. E. Postawko, Global distribution and migration of subsurface ice on Mars, *Icarus, 67,* 1-18, 1986.

Farmer, C. B., and P. E. Doms, Global and seasonal variation of water vapor on Mars and the implications for permafrost, *J. Geophys. Res., 84,* 2881-2888, 1979.

Farmer, C. B., D. W. Davis, A. L. Holland and P. E. Downs, Water vapor observations from the Viking Orbiters, *J. Geophys. Res., 82,* 4225-4248, 1977.

Freeze, R. A., and J. A. Cherry, *Ground Water*, 604 pp., Prentice-Hall, Englewood Cliffs, N.J., 1979.

Grant, J.A., and P. H. Schultz, Degradation of selected terrestrial and Martian impact craters, *J. Geophys. Res., 98,* 11,025-11,042, 1993.

Gulick, V. G. and V. R. Baker, Valley development on Mars: A global perspective (abstract), Lunar Planet. Sci. Conf., XXI, 443-444, 1990.

Hopkins, D. M., T. N. V. Karstrom, R. F. Black, J. R. Williams, T. L. Pewe, A. T. Fernold, and E. H. Muller, Permafrost and ground water in Alaska, *U.S. Geol. Surv. Prof. Pap. 264-F,* 113-146, 1955.

Jones, K. L., Evidence for an episode of crater obliteration intermediate in Martian history, *J. Geophys. Res., 79,* 3917-3931, 1974.

Kargel, J. S., and R. G. Strom, Ancient glaciation on Mars, *Geology, 20,* 3-7, 1992.

Kuzmin, R. O., N. N. Bodina, E. V. Zabalueva, and V. P. Shashkina, Structure inhomogeneities of the Martian cryolithosphere, *Solar Syst. Res., 22,* 195-212, 1988.

Lauchenbruch, A. H., Thermal effects of the ocean on permafrost, *Geol. Soc. Am. Bull., 68,* 1515-1530, 1957.

Leighton, R. B., and B. C. Murray, Behavior of CO_2 and other volatiles on Mars, *Science, 153,* 136-144, 1966.

Leopold, L. B., M. G. Wolman, and J. B. Miller, *Fluvial Process in Geomorphology*, 522 pp., Freeman, Cooper, San Francisco, Calif., 1964.

Mackin, J. H., Concept of a graded river, *Geol. Soc. Am. Bull., 48,* 813-894, 1948.

MacKinnon, D. J., and K. L. Tanaka, The impacted Martian crust: Structure, hydrology, and some geologic implication, *J. Geophys. Res., 94,* 17359-17370, 1989.

Malin, M. C., Salt weathering on Mars, *J. Geophys. Res., 79,* 3888-3894, 1974.

Malin, M. C., G. E. Danielsen, A. P. Ingersoll, H. Masursky, J. Veverka, M. A. Ravine, and T. A. Soulanille, Mars Observer camera, *J. Geophys. Res., 97,* 7699-7718, 1992.

Masursky, H., J. M. Boyce, A. L. Dial, G.G. Schaber, and M. E. Stobell, Classification and time formation of Martian channels based on Viking data, *J. Geophys. Res., 82,* 4016-4038, 1977.

McEwen, A.S., Mars as a Planet, (abstract), Lunar Planet. Sci. Conf. XVIII, 612-613, 1987.

Mellon, M. T., and B. M. Jakosky, Geographic variations in the thermal and diffusive stability of ground ice on Mars, *J. Geophys. Res., 98,* 3345-3364, 1993.

Nedell, S.S., S.W. Squires, and D.W. Anderson, Origin and evolution of the layered deposits in the Valles Marineris, Mars, *Icarus, 70,* 409-441, 1987.

Papke, K.G., Evaporites and brines in Nevada Ppayas, *Bull. Nev. Bur. Mines Geol., 87,* 35, 1986.

Parker, T. J., D. S. Gorsline, R. S. Saunders, D. C. Pieri, and D. M. Scheneeberger, Coastal geomorphology of the Martian northern plains, *J. Geophys. Res., 98,* 11,061-11,078, 1993.

Peterfreund, A.R., Visual and infrared observations of wind streaks on Mars, *Icarus, 45,* 447-467, 1981.

Pieri, D., Maritan valleys: Morphology, distribution, age and origin, *Science, 210,* 895-897, 1980.

Pollack, J. B., J. F. Kastings, S. M. Richardson, and K. Poliqkoff, The case for a wet, warm climate on early Mars, *Icarus, 71*, 203-224, 1987.

Pollack, J. B., R. M. Haberle, J. R. Murphy, J. Schaeffer, and H. Lee, Simulations of the general circulation of the Martian atmosphere, 1, Polar processes, *J. Geophys. Res., 95,* 1447-1473, 1990.

Pollack, J. B., R. M. Haberle, J. R. Murphy, J. Schaeffer and H. Lee, Simulations of the general circulation of the Martian atmosphere, 2, Seasonal Pressure Variations, *J. Geophys. Res., 98,* 3149-3181. 1993.

Squyres, S., Urey prize lecture: Water on Mars, *Icarus, 79,* 229-288, 1989.

Squyres, S. W., and M. H. Carr, Geomorphic evidence for the distribution of ground ice on Mars, *Science, 231*, 249-252, 1986.

Squyres, S. W., S. M. Clifford, R. O. Kuzmin, J. R. Zimbelman, and F. M. Costard, Ice in the Martian regolith, in *Mars* edited by H.H. Kiefer et al., pp. 523-554, University of Arizona Press, Tucson, 1992.

Stoertz, G.E. and G.E. Ericksen, Geology of salars in northern Chile, *U.S. Geol. Surv. Prof. Pap. 811*, 65, 1974.

Strahler, A. N., Dynamic basis for geomorphology, *Geol. Soc. Am. Bull., 63,* 923-938, 1952.

Tanaka, K. L., The stratigraphy of Mars, *Proc. Lunar Planet. Sci. Conf. 17th, Part 1, J. Geophys. Res., 91, suppl.,* E139-E158, 1986.

Ward, A.W., Yardangs on Mars: Evidence for recent wind erosion, *J. Geophys. Res., 84,* 8147-8166, 1979.

Williams, P. J., and M. W. Smith, *The Frozen Earth, Fundamentals of Geocryology,* 306 pp., Cambridge University Press, New York, 1989.

Williams, J. R., Ground water in the permafrost regions of Alaska, *U.S. Geol. Surv. Prof. Pap., 696,* 83 pp., 1970.

Williams, S. H. and J. R. Zimbelman, "White rock:" An eroded Martian lacustrine deposit(?), *Geology, 22,* 107-110, 1994.

Zent, A.P., and F.P. Fanale, Possible Mars brines: Equilibrium and kinetic considerations. Proc. Lunar Planet. Sci. Conf. 16th, Part 2, *J. Geophys. Res. 91*, D439-D445, 1986.

Zimbelman, J. R., S. M. Clifford, and S. H. Williams, Concentric crater fill on Mars: An aeolian alternative to ice-rich mass wasting, Proc. Lunar Planet. Sci. Conf., 19th, 397-407, 1989.

Zuber, M. T., D. E. Smith, S. C. Solomon, D. O. Muhleman, J. W. Head, J. B. Garvin, J. B. Abshire, and J. L. Bufton, The Mars Observer laser altimeter investigation, *J. Geophys. Res., 97,* 7781-7798, 1992.

R. D. Forsythe, Department of Geography and Earth Sciences, University of North Carolina at Charlotte, Charlotte, NC 28223 (email: rdforsyth@email.uncc.edu)

J. R. Zimbelman, Center for Earth and Planetary Studies, National Air and Space Museum, Smithsonian Institution, Washington, DC 20560 (email: jrz@ceps.nasm.edu)

(Received March 1, 1994; revised January 25, 1995; accepted January 25, 1995.)

Page intentionally left blank

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. E3, PAGES 5565–5566, MARCH 25, 1995

Journal of Geophysical Research

Information for Contributors

Original scientific contributions on the physics and chemistry of the Earth, its environments, and the solar system will be considered for publication in the *Journal of Geophysical Research*. Every part of a paper submitted for publication should convey the author's findings precisely and immediately to the reader. Authors are urged to have their papers reviewed critically by colleagues for scientific accuracy and clarity of presentation.

Typescript. All parts of the paper must be typed double or, preferably, triple spaced on good quality white paper 8½ by 11 inches (21.5 x 28 cm) with at least 1-inch (2.5 cm) margins at top, bottom, and sides. Erasable bond, which smudges easily, and tissue paper are not acceptable. Each page of the typescript should be numbered in the top right-hand corner.

Authors are expected to supply neat, clean copy and to use correct spelling, punctuation, grammar, and syntax. Spelling and hyphenation of compound words follow the unabridged *Webster's Third New International Dictionary*. The metric system must be used throughout; use of appropriate SI units is encouraged. Following recommended style and usage expedites processing and reduces the chance of error in the typesetting.

Because text footnotes are expensive to set and distracting to the reader, they should be incorporated into the text or be eliminated completely.

The typescript should be arranged as follows: (1) title page including authors' names and affiliations, (2) abstract, (3) text, (4) reference list, (5) figure legends, and (6) tables.

Abstract. The abstract should in a single paragraph (250 words or fewer) state the nature of the investigation and summarize its important conclusions. Listing the contents in terms such as "this paper describes" or "the paper presents" should be avoided. Use of the passive voice often indicates that the author is merely describing the procedure rather than presenting conclusions. References should not be cited in the abstract.

The abstract should be suitable for separate publication in an abstract journal and be adequate for indexing.

When a paper is accepted for publication, the author will be requested to prepare camera-ready copy of the abstract for inclusion in *Geophysical Abstracts in Press* (GAP), AGU's alerting service about accepted papers. At the same time, the author will select appropriate index terms for the paper.

Mathematics. All characters available on a standard typewriter must be typewritten in equations as well as in text. The letter l and the numeral 1 and the letter O and the numeral 0, which are identical on most typewriters, should be identified throughout the paper to prevent errors in typesetting. Any symbols that must be drawn by hand should be identified by a note in the margin. Authors will be charged for all corrections in galley proof for material that was handwritten in the typescript. Allow ample space (3/4 inch (2 cm) above and below) around equations so that type can be marked for the printer.

Alignment of symbols must be unambiguous. Superscripts and subscripts should clearly be in superior or inferior position. Fraction bars should extend under the entire numerator.

Barred and accented characters that are available for machine typesetting may be used. Symbols that are not available and therefore must be avoided are triple dots, any accents (other than bars) that extend over more than one character, and double accents (e.g., a circumflex over a bar). Accents over characters can be eliminated by the use of such symbols as ', *, and † set as superscripts.

If an accent or underscore has been used to designate a special typeface (e.g., boldface for vectors, script for transforms, sans serif for tensors), the type should be specified by a note in the margin.

If the argument of an exponential is complicated or lengthy, "exp" rather than e should be used. Awkward fractional composition can be avoided by the proper introduction of negative powers. In text, solidus fractions ($1/r$) should be used, and enough enclosures should be included to avoid ambiguity. According to the accepted convention, parentheses, brackets, and braces are in the order { [( )] }. Displayed equations should be numbered consecutively throughout the paper; the number (in parentheses) should be to the right of the equation.

Notation. The notation is a list of symbols used in the text as an aid to the reader. It should be set up in this form:

c rate of soil accumulation, m/yr.
d median grain size of water-deposited material, μm.
D distance of the locus of points, m.
h elevation of the rock stream channel at a particular time t_i, meters above base level.
Hf maximum vertical displacement of the fault associated with an earthquake, m.
Hf' amount of vertical displacement of the mountain-bounding fault, m.
z vertical coordinate of the model grid system.
β parameter computed by $\beta = b/\log_{10} e$.
γ mean peak flow rate above base flow of stream, m^3/s.

References. A complete and accurate reference list is of major importance. Omissions, discrepancies in the spelling of names, errors in titles, and incorrect dates make citations annoying, if not worthless, to the reader and cast doubt on the reliability of the author as well.

Only works cited in the text should be included in the reference list. References are cited in the text by the last name of the author and the year: [*Jones*, 1990]. If the author's name is part of the sentence, only the year is bracketed. Personal communications and unpublished data or reports are not included in the reference list; they should be shown parenthetically in text: (F. S. Jones, unpublished data, 1990).

References are arranged alphabetically by the last names of authors. Multiple entries for a single author are arranged chronologically. Two or more publications by the same author in the same year are distinguished by a, b, c after the year.

For laboratory, company, or government reports, information should be included on where the report can be obtained. For Ph.D. and M.S. theses the institution granting the degree and its location should be given.

References to papers delivered at meetings should include title of paper, full name of meeting, sponsor, meeting site, and date. Citations of papers presented at meetings have been complicated by the practice of collecting manuscripts from participants, reproducing the manuscripts, and distributing the collections to people attending the meeting. Such collections should not be cited as published works.

References to books should include the page numbers of the material being cited.

Names of periodicals should be either written out in full or abbreviated according to the system employed by the Chemical Abstracts Service. One-word titles should always be given in full: *Science*. It is permissible to give only the initial page number of a paper but preferable to give the range of pages. Samples:

Beal, R. C., The Seasat SAR wind and ocean wave monitoring capabilities, *Rep. JHU/APL SIR79U-019*, 56 pp., Appl. Phys. Lab., Johns Hopkins Univ., Laurel, Md., 1980.

Hartle, R. E., and J. M. Grebowsky, Upward ion flow in the nightside ionosphere of Venus (abstract), *Eos Trans. AGU, 71*, 1431, 1990.

Macdonald, G. A., A. T. Abbot, and F. L. Peterson (Eds.), *Volcanoes in the Sea: The Geology of Hawaii*, 2nd ed., 517 pp., University of Hawaii Press, Honolulu, 1983.

Passey, Q. R., Viscosity structure of the lithospheres of Ganymede, Callisto, and Enceladus, and of the Earth's mantle, Ph.D. thesis, Calif. Inst. of Technol., Pasadena, 1982.

Shaw, H. R., and D. A. Swanson, Eruption and flow rates of flood basalts, in *Proceedings of the Second Columbia River Basalt Symposium*, pp. 271-299, Eastern Washington State College Press, Cheney, 1970.

Skalsky, A., R. Gerard, S. Klimov, C. Nairn, J. G. Trotignon, and K. Schwingenschuh, Martian bow shock: Topological features of the upstream region, paper presented at Chapman Conference on Venus and Mars, AGU, Balatonfured, Hungary, June 3-8, 1990.

Spera, F. J., A. Borgia, J. Strimple, and M. Feigenson, Rheology of melts and magmatic suspensions, 1, Design and calibration, *J. Geophys. Res., 93*, 10,273-10,294, 1988.

Squyres, S. W., and S. K. Croft, The tectonics of icy satellites, in *Satellites*, edited by J. Burns and M. S. Matthews, pp. 193-341, University of Arizona Press, Tucson, 1986.

Refer to recent issues of the *Journal of Geophysical Research* for more examples.

Tables. Tables should be typed as authors expect them to look in print. Every table must have a title, and all columns must have headings. Column headings must be arranged so that their relation to the data is clear. Footnotes should be indicated by reference marks (*, †, ‡, §) or by lowercase letters typed as superiors. Each table must be cited in text. Authors are urged to make supporting data available as supplements, as described below.

Illustrations. One original or glossy set of illustrations suitable for reproduction and three clear photocopies for the reviewers should accompany the typescript. Illustrations should be original drawings in black ink or, preferably, sharply focused glossy prints. Lettering should be mechanical, not hand drawn. When reduced for printing, lettering should be at least 1.5 mm high: the largest lettering should not exceed 3 mm. The following limiting dimensions apply: page height, 24.1 cm; page width, 17.3 cm, column width, 8.4 cm.

Figures not cluttered with information that could be placed in the legend look neater and are easier to read. All details on the figures should be checked carefully because errors found in galley proof stage (e.g., misspelled words, incorrect values, omitted symbols) must be corrected by the author, and the cost of preparing the new figure for printing will be charged to the author.

Each figure must be cited in numerical order in text and must have a figure legend.

Color figures and foldouts can be accommodated, but the additional costs for these special figures must be borne by the author. The author should retain an identical print for checking color proofs. The AGU Publications office should be consulted for current costs and other special instructions.

If figures that have been published under copyright are to be reproduced in JGR, AGU must have permission from the copyright holders. Copies of letters of permission to republish should accompany the typescript.

Supplementary material. Authors are encouraged to submit papers that are as concise as possible. Supporting material, such as tables of data, computer printouts, appendices, additional graphs, lengthy mathematical derivations, and extended background discussions may be made available through videotape, electronic mechanisms, or microfiche. All supplements will be subject to the same peer review as the printed material. Summaries of articles and key figures may be submitted for publication in the journal, the detailed article being published on microfiche. All material published on microfiche is included in the microform editions of JGR (and therefore archived in libraries) and is available to individuals on order. Contact the editors or AGU Publications office for more information on this service.