Modeling the development of martian sublimation thermokarst landforms

Colin M. Dundas a,*, Shane Byrne b, Alfred S. McEwen b

a Astrogeology Science Center, U.S. Geological Survey, 2255 N. Gemini Dr., Flagstaff, AZ 86001, USA
b Lunar and Planetary Laboratory, University of Arizona, Tucson, AZ 85721, USA

Abstract

Sublimation-thermokarst landforms result from collapse of the surface when ice is lost from the subsurface. On Mars, scalloped landforms with scales of decameters to kilometers are observed in the mid-latitudes and considered likely thermokarst features. We describe a landscape evolution model that couples diffusive mass movement and subsurface ice loss due to sublimation. Over periods of tens of thousands of Mars years under conditions similar to the present, the model produces scallop-like features similar to those on the martian surface, starting from much smaller initial disturbances. The model also indicates crater expansion when impacts occur in surfaces underlain by excess ice to some depth, with morphologies similar to observed landforms on the martian northern plains. In order to produce these landforms by sublimation, substantial quantities of excess ice are required, at least comparable to the vertical extent of the landform, and such ice must remain in adjacent terrain to support the non-deflated surface. We suggest that martian thermokarst features are consistent with formation by sublimation, without melting, and that significant thicknesses of very clean excess ice (up to many tens of meters, the depth of some scalloped depressions) are locally present in the martian mid-latitudes.

1. Introduction

Ground ice on Mars represents a key link between martian geology and climate history, and is a potential resource for future exploration. The theoretical basis for the distribution of ground ice on Mars was established by Leighton and Murray (1966) and has been refined by many subsequent workers (e.g., Mellon et al., 2004; Schorghofer and Aharonson, 2005; Chamberlain and Boynton, 2007). The distribution of near-surface ice is expected to change substantially with variations in Mars’ orbit and climate; at present, ice is thought to be stable at a depth of decimeters in the mid-latitudes, and millimeters to centimeters near the poles (e.g., Mellon and Jakosky, 1995; Head et al., 2003; Chamberlain and Boynton, 2007; Schorghofer, 2007). The presence of ground ice can have a dramatic effect on the geomorphology, resulting in a diverse range of landforms.

Excess ice is ground ice exceeding the natural pore space of unfrozen soil (van Everdingen, 1998). A number of lines of evidence now indicate widespread excess ice on Mars (e.g., Boynton et al., 2002; Byrne et al., 2009; Mouginit et al., 2010; Dundas et al., 2014). Excess ice with only ~1 vol% regolith was also directly excavated by the Phoenix lander (Smith et al., 2009; Mellon et al., 2009; Cull et al., 2010). The distribution and origins of this ice, however, are still not understood. Geomorphology provides another avenue for learning about this ice. Thermokarst landforms develop when ice is lost from ice-rich ground in a spatially heterogeneous way, leading to local surface subsidence and collapse. Therefore, they can serve as indicators of the loss of excess ice and provide insight into the conditions and processes forming and removing the ice. Thermokarst features on Mars were proposed by several authors based on Mariner and Viking imagery (e.g., Sharp, 1973; Anderson et al., 1973; Costard and Kargel, 1995). Most recent analyses of potential thermokarst have focused on concentrations of large scalloped depressions in Utopia Planitia and south of the Hellas basin (Plescia, 2003; Morgenstern et al., 2007; Soare et al., 2007, 2008, 2011; Lefort et al., 2009, 2010; Ulrich et al., 2010; Zanetti et al., 2010; Séjourné et al., 2011, 2012). In their simplest form, these rimless depressions have shallow, equator-facing slopes, somewhat steeper, pole-facing scarps, and are meters to tens of meters deep, although they may merge to produce much more complicated landscapes (Figs. 1–4). Because Mars is a cold and dry planet, most recent authors have suggested that these form through sublimation, although Soare...
et al. (2007, 2008) proposed melting ground ice instead, and argued for stable, ponded surface water. A number of terrestrial thermokarst landforms are known, including alases (depressions with steep sides and flat floors) and oriented thaw lakes (e.g., French, 2007). The formation of terrestrial thermokarst features normally results from melting rather than sublimation, and on Earth, excess ground ice primarily results from processes involving liquid water, such as ice segregation or the growth of ice wedges (e.g., French, 2007). Soare et al. (2011) raised several questions about sublimation models for martian scalloped depressions and considered such “wet” periglacial processes to be a more effective terrestrial analog. Consequently, the formation of martian thermokarst features is of importance not only for understanding the state of ground ice but also the climate of the geologically recent past.

Studies of martian thermokarst have been hindered by the qualitative nature of formation models. Landscape evolution modeling offers the ability to assess the way in which a landform will evolve, and early applications to Mars have led to insights into a variety of landforms (e.g., Byrne and Ingersoll, 2003; Forsberg-Taylor et al., 2004; Pelletier, 2004; Howard, 2007). In this paper we report on results of landscape evolution modeling of martian sublimation-thermokarst processes that allow us to better understand the three-dimensional evolution of a landform. In Section 2, we summarize observations of scalloped depressions and models of their origins, and then give a brief description of evidence for thermokarstic crater expansion on the martian northern plains. The landscape evolution model is described in Section 3, and the scenarios we consider in Section 4. Section 5 gives the model results, and Section 6 discusses their broader significance for martian geomorphology and ground ice.
The difficulty of melting subsurface ice under recent martian climate conditions (e.g., Mellon and Phillips, 2001) has led most researchers to focus on sublimation as the likely cause. Morgenstern et al. (2007), Zanetti et al. (2010) suggested that higher temperatures on Sun-facing slopes led to greater sublimation and retreat of those slopes. Lefort et al. (2009) suggested both enhanced sublimation on equator-facing slopes and retreat of pole-facing scarps by mass wasting, particularly at high obliquity, along with eolian erosion to remove material from a developing protective lag. Ulrich et al. (2010), Séjourné et al. (2011) both proposed that the major growth process was enhanced sublimation and retreat of the pole-facing scarps by high obliquity.

2.2. Crater expansion

Figs. 5 and 6 show examples of unusual craters seen on the northern plains, particularly in and near Arcadia Planitia (Viola et al., 2015). The basic morphology is crudely funnel-shaped, with a sub-circular pit inside a broader depression with somewhat shallower slopes. The broad depression may be symmetric or asymmetric, but in local areas the craters usually have similar shapes. Smaller craters in the vicinity tend to be highly degraded, and the depressions surrounding adjacent craters may merge (Fig. 5c). In some locations, expanded morphologies are confined to higher-standing terrain, while craters on the lower-lying surface may have typical bowl-shaped morphologies (Fig. 5d). The primary origin is inferred to be impact because this is the process most likely to create tens of thousands of decameter-scale craters, and because some occur in radial patterns extending from well-preserved (5–30 km) craters (Fig. 5e; Viola et al., 2015), implying that dense fields of such craters originated from secondary impacts.

Asymmetric surrounding depressions could be explained by oblique impacts, but some depressions have very irregular shapes uncharacteristic of impact craters (Fig. 5c). The most likely alternative explanation is that the observed features are the result of impact into a layered target. Such layering can produce terraced or flat-floored craters (e.g., Melosh, 1989). We consider this unlikely to be the sole explanation for several reasons. Some simple craters are seen in the immediate vicinity of apparently expanded examples (Fig. 5b); layered-target impacts should develop their characteristic morphology at the time of impact, so these craters should be tiered unless the layers are laterally discontinuous. (In some, but not all, cases, the non-expanded craters appear to be in lower-lying, texturally distinct material.) Small craters in an ice-poor layered target should be largely unaffected, but instead appear to be very heavily modified (Fig. 5c). Asymmetric surrounding depressions could be explained by oblique impacts, but some depressions have very irregular shapes uncharacteristic of impact craters (Fig. 5c). Finally, the commonly observed funnel-like profile is somewhat different from the concentric-crater morphology expected of a layered target, and larger craters do not transition to the flat-floor or central-mound morphology expected for cases where the layer thickness is less than ~0.5 x the crater diameter (Melosh, 1989). (A handful of such terraced craters are observed in Arcadia Planitia with considerably larger diameters (Branson et al., 2014).) It is possible, however, that the target surfaces are layered and this may be an additional influence on crater morphology in some cases.

Eolian modification of impact craters is unlikely to be the main driver of expansion. It is unlikely that wind patterns would be so balanced as to produce essentially symmetric expansion, and it is not clear how wind would produce the observed breaks in slope. Such expansion morphologies are not seen in the Medusae Fossae Formation, which is thoroughly sculpted by the wind. Eolian processes, however, may influence thermokarst evolution by depositing sand or dust in depressions, or by removing lag deposits and accelerating subsequent sublimation.
3. Landscape evolution model

We examine sublimation thermokarst by modeling the landscape evolution of a martian surface undergoing loss of ground ice. The model makes use of the same vapor-diffusion process that determines the depth to stable ice in models used to determine the global ground ice distribution (e.g., Leighton and Murray, 1966; Mellon et al., 2004; Schorghofer and Aharonson, 2005; Chamberlain and Boynton, 2007). Ground ice is stable when the mean vapor pressure in the atmosphere matches that of the saturated gas at the ice table. The occurrence of a lag damps the amplitude of temperature variations in the ice and thus reduces the mean vapor pressure, because of the exponential dependence of vapor pressure on temperature. Consequently, at mid and high latitudes ground ice is stable under a lag of some thickness, dependent on latitude and thermophysical properties. If the ice is too shallow, the average vapor pressure over the course of the year is higher than in the atmosphere, resulting in a net vapor pressure

Fig. 5. (A) Example of a roughly symmetric crater showing apparent expansion, with textural change possibly due to eolian materials in the inner cavity (HiRISE image ESP_016108_2305). (B) Non-expanded crater among expanded craters (HiRISE image ESP_016108_2305). (C) Asymmetric expansion (HiRISE image ESP_017308_2315). Note numerous pits or highly degraded small craters. (D) Expanded craters confined to high-standing unit with distinct texture. This high-standing unit approximately corresponds to the ejecta of an older, larger crater, but expanded craters also occur in mantling materials within the crater and on detached hummocks. (Anaglyph from HiRISE stereo images ESP_017875_2305 and ESP_018007_2305.) (E) Expanded craters in a crater ray composed of secondary craters. (CTX image B19_017018_2349_XN_5N144W. Credit: NASA/JPL/Malin Space Science Systems (MSSS.) North is up and light from the left in all panels.

Fig. 6. Typical expanded-crater profiles from a HiRISE DTM constructed from stereo images ESP_026446_2305 and ESP_027158_2305. The vertical axis of each profile has been shifted so that each profile begins at 0 m. Note the broad funnel-like shape steepening near the center of each crater. Digital Terrain Model credit NASA/JPL/University of Arizona.
gradient leading to diffusive loss of water vapor and retreat of the ice. An ice table deeper than equilibrium experiences the reverse process. We use a one-dimensional thermal model to estimate ice loss rates for various slopes and ice depths, and then use those loss rates to adjust the ice depth in a landscape model. Where excess ice (exceeding natural soil pore space) is present in the model, sublimation can alter the surface topography, adjusting slopes for future time steps and driving the evolution of landforms.

The thermal model solves the thermal conduction equation using a semi-implicit (Crank–Nicolson) scheme and the convergent algorithm described by Chamberlain (2006) to reach a stable solution. The surface energy balance is given by

\[ F_{\text{solar}} + F_{\text{IR}} + F_{\text{scat}} + \frac{k}{\partial T/\partial z} + L_{\text{CO}_2} + \frac{\partial m}{\partial t} = \varepsilon \sigma T^4 \]

(1) where \( F_{\text{solar}} \) is the insolation reaching the surface as a function of incidence angle and solar distance, \( F_{\text{IR}} \) is the thermal radiation from the atmosphere, \( F_{\text{scat}} \) is atmospherically scattered light, \( k \) is the thermal conductivity, \( T \) is the temperature, \( z \) is vertical distance, \( L_{\text{CO}_2} \) is the latent heat of \( \text{CO}_2 \) frost (if present), \( m \) is the mass of \( \text{CO}_2 \) frost, \( \varepsilon \) is the surface emissivity, \( \sigma \) is the Stefan–Boltzmann constant and \( T^4 \) is the surface temperature. \( \text{CO}_2 \) frost condenses if the surface temperature drops to the \( \text{CO}_2 \) frost point, and the albedo is adjusted when \( \text{CO}_2 \) frost is present. Thermophysical properties (thermal conductivity, heat capacity and density) are treated as constant with respect to temperature and atmospheric pressure. Standard model parameters are given in Table 1.

Fluxes are calculated similarly to Schorghofer and Edgett (2006):

\[ F_{\text{solar}} = \frac{S_0}{R} (1 - A)(1 - f)^{1/\max(\cos(\theta),0.04)} \sin \theta \]

(2)

\[ F_{\text{IR}} = \frac{S_0}{R^2} \cos^2 \left( \frac{\theta}{2} \right) \cos L_{\text{noon}} + F_{\text{surf}} \sin \left( \frac{\theta}{2} \right)^2 \]

(3)

\[ F_{\text{scat}} = \frac{1}{2} f_{\text{scat}} \frac{S_0}{R^2} \cos^2 \left( \frac{\theta}{2} \right) \]

(4)

where \( S_0 \) is the solar constant, \( R \) is the Mars–Sun distance, \( A \) is the surface bolometric hemispheric albedo, \( f \) is the incidence angle, \( \theta \) is the elevation of the Sun above a sloped surface and \( z \) is the surface slope. For sloped surfaces the model accounts for self-shadowing and slope effects on insolation. Incoming thermal radiation is adjusted due to the reduced view of the sky. For computational efficiency, rather than model radiative interaction with other surfaces at different temperatures, slopes receive radiation \( (F_{\text{surf}}) \) from a level surface at the temperature of flat ground that does not interact with any slopes. The coefficient \( f_{\text{surf}} \) is taken as 0.04 \( \times (P_{\text{surf}}/700 \text{ Pa}) \), where \( P_{\text{surf}} \) is the surface pressure. This is a modified version of the “2% assumption” of Kieffer et al. (1977), in which the infrared emission from the atmosphere is assumed to be 2% of the noon insolation. (We scale to a different reference pressure than Schorghofer and Edgett (2006), which is the approximate pressure of the Viking Lander 1 site where Haberle and Jakosky (1991) tested the “2% assumption.”) The coefficient \( f_{\text{scat}} \) is taken as 0.02 and \( f = f_{\text{IR}} + f_{\text{scat}} \). This model for atmospheric extinction and thermal radiation is simplistic but takes much less computer time than a more accurate model of radiative transfer. For level surfaces with the two sets of baseline parameters discussed below, annual-average temperatures from the simple model differed from a more detailed model (Dundas et al., 2014) by 3 and 1 K, respectively. The instantaneous surface temperatures often differed more significantly, but ice stability is controlled by subsurface temperature. For the purpose of understanding the general behavior of landforms, the thermal model does not require a high level of accuracy, only a reasonable approximation of the behavior of equilibrium depth with slope and aspect. Stability depths in this paper should be understood in this context. Moreover, the variation in atmospheric conditions over the hundreds of kyr timescales that we model is probably significant.

Near-surface atmospheric water vapor contents are modeled using the method of Chamberlain (2006). (Slightly different approaches have been used by Mellon et al. (2004), Schorghofer and Aharonson (2005). For the present paper we use this method since we also use column abundance predictions from Chamberlain (2006) for different orbital conditions.) A value is assumed for the annual-average atmospheric water vapor content. It is assumed that this water is mixed in an atmospheric column which is either at the saturation value for the average diurnal temperature or at some limiting constant value, whichever is less. The constant value is then tuned to match the assumed annual-average column abundance of water vapor. Note that these column abundances are simply a proxy for the annual-average near-surface water vapor number density (the parameter actually used to calculate ice stability depth), and the true relationship between column abundance and near-surface water vapor abundance may not be well-described by assumptions of uniform mixing (cf. Tamppari and Lemmon, 2014). The near-surface value is then adjusted to never exceed the saturation vapor pressure at the surface temperature, which has the effect of depleting near-surface water vapor at night. This model is simplistic, but captures the basic features of the observed annual water cycle adequately for our purposes, since we only need a reasonable approximation of the near-surface water vapor density. For the present climate we assume an

### Table 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ice density</td>
<td>917 kg m(^{-3})</td>
<td>Non-porous</td>
</tr>
<tr>
<td>Ice specific heat</td>
<td>1577 J kg(^{-1}) K(^{-1})</td>
<td>Value for 200 K; Paterson (1994)</td>
</tr>
<tr>
<td>Ice thermal conductivity</td>
<td>3.14 W m(^{-1}) K(^{-1})</td>
<td>Value for 200 K; Paterson (1994)</td>
</tr>
<tr>
<td>Regolith specific heat</td>
<td>840 J kg(^{-1}) K(^{-1})</td>
<td>Typical silicate; consistent with Mellon et al. (2004)</td>
</tr>
<tr>
<td>Regolith bulk density and thermal conductivity</td>
<td>Variable</td>
<td>Function of thermal inertia following Mellon and Jakosky (1993)</td>
</tr>
<tr>
<td>Surface emissivity</td>
<td>1.0</td>
<td>Lower values widely used, but see Haberle et al. (2008)</td>
</tr>
<tr>
<td>( \varepsilon_{\text{CO}_2} ) frost emissivity</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>( \varepsilon_{\text{CO}_2} ) frost albedo</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>Vapor diffusion coefficient</td>
<td>( 3 \times 10^{-4} \text{ m}^2 \text{s}^{-1} )</td>
<td>In range from Hudson et al. (2007) experiments; can be somewhat reduced by various factors (Hudson and Aharonson, 2008)</td>
</tr>
<tr>
<td>Dry lag porosity(^a)</td>
<td>Variable</td>
<td>Zent et al. (2010) estimated 0.48–0.55 for the upper 15 mm at the Phoenix Lander site</td>
</tr>
<tr>
<td>Wind speed</td>
<td>2.5 m s(^{-1})</td>
<td>For forced-convection sublimation; see Dundas and Byrne (2010)</td>
</tr>
</tbody>
</table>

\(^a\) Estimated from the regolith bulk density assuming a grain density of 2600 kg m\(^{-3}\). Thermal inertias of 150–250 J m\(^{2}\) s\(^{-1}\) K\(^{-1}\) correspond to porosities of 0.48–0.38.
annual-average atmospheric water column abundance of 20 precipitable microns (pr. μm) at 0 km elevation. This is consistent with the latitude limit of ice exposed by recent impacts, which suggest that 20–25 μm is an appropriate estimate for the recent long-term value (Byrne et al., 2009; Dundas et al., 2014). Mellon et al. (2008b) suggest that a long-term average of 20 pr. μm is a good match for the global ice depth distribution. The two methods may be sensitive to somewhat different timescales for the ice to adjust; for this study we use 20 pr. μm, and note that the general nature of our results would not be affected by modest differences. We also note that the method of scaling column abundance to near-surface vapor density used by the above estimates is somewhat different from that in Chamberlain and Boynton (2007), although Chamberlain (2006) notes that they produce similar results. For high obliquities we use the values estimated by Chamberlain and Boynton (2007). For each set of conditions considered, we use this method and a level-surface thermal model to generate an estimate of the near-surface average atmospheric vapor pressure, and then run a suite of models with varying slope, aspect and ice depth to determine vapor pressures at the ice table.

The thermal model is run with a time step of 0.01 martian days to determine the mean temperature and vapor pressure at the ice table. With values for the annual-average water content at the ice table and the surface, it is possible to calculate the rate of diffusive ice loss (e.g., Hudson et al., 2007; Schorghofer, 2007, 2010):

\[ J = -D \frac{\partial \rho}{\partial z} \] (5)

where \( J \) is the mass flux, \( D \) is the diffusion coefficient, \( \rho \) is the vapor density and \( z \) is vertical distance. We then generate lookup tables of annual-average vapor pressure differences between the surface and a range of ice table depths for varying slopes and slope aspects. These are used to determine annual evaporation rates in the landscape model, with a time step of one Mars year. We assume east–west symmetry, which is observed in models of ice stability on slopes (Aharonson and Schorghofer, 2006) and halves the number of model runs needed to generate a lookup table of ice-table vapor pressures. This assumption was tested by generating mirror-image lookup tables with east- and west-facing slopes. Differences between model runs with each were very small, indicating that this approximation is sufficient. The near-surface atmospheric water content for a site is calculated using a thermal model for a level surface and the method described above. A lookup table is needed for each set of surface thermophysical properties and orbital parameters to be considered. We also calculate the rate of sublimation from an exposed ice surface by forced convection (Dundas and Byrne, 2010) using the same vapor density difference, and use this value if it is less than that resulting from Eq. (5) on the assumption that for very shallow ice, diffusion through the lag is not the rate-limiting process. (Rates from Eq. (5) go to infinity if \( \partial z \) goes to zero and \( \partial \rho \) is finite.) For simplicity, we only consider forced convection rather than the complete sublimation model of Dundas and Byrne (2010). Free convection has a more complex relationship to the vapor density difference.

The surface is described by a grid of elevation points. Given input surface topography and ice depths, we calculate ice loss rates from the lookup table using trilinear interpolation over slope, aspect and the logarithm of ice depth. Although the values within the table do not actually vary linearly, this simplification was considered acceptable in order to improve computation speed. Once the loss rate is calculated, the topography and ice depth are adjusted, depending on whether the top of the ice table consists of pore ice or excess ice. The thickness of the dry lag is also adjusted based on the retreat of pore ice or the regolith content of the excess ice, accounting for the porosity of the developing lag as in Schorghofer (2010). Sublimation retreat is assumed to be normal to the local surface.

In calculating loss rates for each point, we use the ice depth at that point and calculate the slope and aspect using a plane defined by the x and y slopes given by the surrounding four points. This loss rate calculation is somewhat analogous to solution of an advection equation (i.e., time derivative proportional to spatial derivative) via a forward-time-centered-space (FTCS) method, which is known to be numerically unstable (e.g., Press et al., 2007). However, the physical model used to calculate loss rates is complex and does not have a simple linear slope dependence, defying any simple stability analysis. The addition of diffusive mass movement, described next, serves to damp any tendency for short-wavelength perturbations to blow up and destabilize the solution. We found that model results exhibit a very weak dependence on spatial resolution. At the resolution employed (i.e., 0.25 m/post), variation with resolution is far less than the landform relief and falls rapidly as the point spacing is decreased.

In addition to sublimation, mass movement is also affected by diffusive mass movement. This process describes the cumulative effects of small movements producing surface creep. On Mars, such movement could be due to factors like seismic shaking, small meteorite impacts, seasonal frost loading and sublimation, and thermal cycling. Given the abundance of thermal contraction polygons on Mars (Mellon et al., 2008a), the latter is likely to be particularly important. This model may also reproduce the net effects of larger mass movements, but does not resolve individual events. Such slope evolution has typically been modeled by linear diffusion, but we follow recent work in incorporating nonlinear effects, causing a dramatic acceleration in mass wasting as slopes approach some critical angle (e.g., Roering et al., 1999). The flux is given by

\[ q_s = \frac{K \nabla z}{1 - (|\nabla z|/S_c)^2} \] (6)

where \( q_s \) is the sediment flux, \( K \) is the diffusivity and \( S_c \) is the critical slope. We implement this effect using an explicit numerical method detailed by Perron (2011), his Eqs. (9)–(11)). \( K \) and \( S_c \) are model parameters. Coefficients for nonlinear diffusion approximate those for linear diffusion when well below the critical slope.

The input parameters for such equations are highly variable on Earth. Roering et al. (1999) estimated a diffusivity of \( 3.2 \times 10^{-3} \text{ m}^2 \text{ s}^{-1} \) and a critical slope of 1.25 (51°) for a study site in Oregon. Hanks (2000) found that diffusivity values affecting fault scarp in the Basin and Range region of the western United States were typically \( 10^{-3} \text{ m}^2 \text{ s}^{-1} \). Martin and Church (1997) examined a range of previous studies and found a median value of \( 2 \times 10^{-3} \text{ m}^2 \text{ yr}^{-1} \) for creep, with much higher values for the net effect of landsliding.

The values for these parameters on present-day Mars are not known. Creep rates on Mars should be much less than on Earth due to lack of rainfall, lower gravity, and lower rates of other processes, but mid- and high-latitude rates on Mars are probably much higher than the planetary average due to thermal cycling of ground ice and the effects of seasonal CO2 frost (both basal sublimation and loading of the surface), as shown by the rapid disappearance of small craters (Korteniemi and Kreslavsky, 2013). This is consistent with the generally muted topography at high latitudes (e.g., Squyres and Carr, 1986), although mantling by airfall deposits likely also contributes to this appearance (e.g., Kreslavsky and Head, 2000). In the absence of measurements of this parameter for Mars, we adopt a value of \( 10^{-4} \text{ m}^2 \text{ yr}^{-1} \), which is near the low end of terrestrial values, and therefore probably an upper limit for Mars. The short high-latitude crater lifetimes inferred by Korteniemi and Kreslavsky (2013) would be consistent with higher values, but they investigated latitudes that are mostly
The model incorporates self-shadowing by a slope element, but the controlling processes are different because lunar landscape evolution is dominated by small impacts that are partially screened by the martian atmosphere. Assuming a high value is a conservative assumption for our purposes since it will tend to smooth modeled landforms and reduce the relief generated by sublimation. Martian soils are known to support steep trench walls and have some cohesion (e.g., Moore et al., 1987; Shaw et al., 2009), so we adopt a critical slope of 1.25 (51°) for this work. We apply creep modification only to the regolith cover, omitting the possibility of viscous deformation of the ice or modification by fracture infill.

The default parameters for diffusive models like this one cease to be valid for exposed bedrock or ice-cemented ground, and we use this model only to modify regolith thickness. We handle this in a simplified manner by finding nodes where the lag thickness would fall below 0.1 mm, reducing the outward flux components from those nodes by a fixed factor, and then recalculating the diffusive change to topography, iterating until there are no such nodes. This approach has no direct physical basis but suffices as a simple approximation.

The most important model parameters are the initial topography, depth to ice, excess ice thickness, and thickness of pore ice above the clean ice (if any); the regolith content of the excess ice; the diffusivity used in modeling mass wasting; and the geographic location, thermophysical properties and orbital parameters used to generate the lookup table. Standard model parameters are summarized in Table 1. We use periodic boundary conditions, which equate to an assumption that the landform being modeled is part of a field of similar landforms with spacing set by the domain size. This assumption is consistent with the occurrence of regions of abundant scallops or expanded secondary craters, although in scalloped regions it is not necessarily the case that all landforms began growth at the same time.

This model is necessarily limited in order to make it computationally tractable. The simplifications are mainly driven by the need to treat the slopes and ice table depths generically in the lookup table rather than individually modeling evolution rates for the unique conditions of each point on the surface. The thermal model incorporates self-shadowing by a slope element, but does not include mutual shadowing and has only a simplistic radiative interaction with a level slope. In generating the lookup tables we assume that the icy layer has the thermophysical properties of pure ice rather than ice-cemented ground, because we are primarily interested in cases where a significant quantity of excess ice is present near the surface. Ice and ice-cemented regolith are much more similar to each other than they are to dry regolith, so we consider this an adequate approximation. Other simplifications are used because we focus on examining the landforms produced by sublimation following disturbances to equilibrium. We did not define thermal contraction polygons or other lateral variations in the ice table, although such variations could localize sublimation (as suggested in some locations by pits along polygon troughs). Under conditions where the ice table is aggrading, we assume that the net annual flux of ice into the ground fills the pore space just above the ice table. This style of aggradation is predicted to occur when ice ages above an impermeable ice layer (Schorghofer and Forget, 2012), but in some cases pore space may be partially filled through a greater volume of the lag (e.g., Schorghofer, 2010). Since we are modeling conditions that produce thermokarst landforms, where most of the surface is stable or undergoing ice loss, this is not expected to fundamentally affect the results; cold slopes where ice is aggrading do not undergo topographic changes, because ice is assumed to fill in the pore space. In situ growth of excess ice is possible under some conditions (e.g., Fisher, 2005; Sizemore et al., 2015) but is not included. (The process modeled by Fisher (2005) is slow; the ice segregation modeled by Sizemore et al. (2015) can be fast, but primarily under warm conditions.) This simplification is likewise considered acceptable since we are primarily interested in conditions where sublimation is dominant. We also omit any viscous creep of the ice. Both viscous creep and ice lens growth may be important in some conditions on Mars, but in this paper we focus on the morphologies produced by sublimation.

### Table 2
Model scenarios.

<table>
<thead>
<tr>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scallop case 1</td>
<td>Baseline; 50°N, thermal inertia (TI) 250 J m$^{-2}$ K$^{-1}$ s$^{-1/2}$, albedo 0.13, 0.55 m equilibrium ice depth, $\kappa = 10^{-4}$ m$^2$ yr$^{-1}$, 1 vol% regolith in thick clean ice under the equilibrium cover, circular initial disturbance with center depth 0.3 m and radius 16 m</td>
</tr>
<tr>
<td>Scallop case 2</td>
<td>As baseline, but 10 vol% regolith in clean ice</td>
</tr>
<tr>
<td>Scallop case 3</td>
<td>As baseline, but restricted thickness (1 m) of clean ice</td>
</tr>
<tr>
<td>Scallop case 4</td>
<td>As baseline, but with TI of 350 J m$^{-2}$ K$^{-1}$ s$^{-1/2}$, causing ice to be unstable under level surfaces</td>
</tr>
<tr>
<td>Scallop case 5</td>
<td>As baseline, but smaller initial disturbance (8 m radius)</td>
</tr>
<tr>
<td>Scallop case 6</td>
<td>As baseline, but shallower initial disturbance (0.15 m)</td>
</tr>
<tr>
<td>Scallop case 7</td>
<td>As baseline, but wider (elliptical) disturbance</td>
</tr>
<tr>
<td>Scallop case 8</td>
<td>As baseline, but $\kappa = 3 \times 10^{-5}$ m$^2$ yr$^{-1}$</td>
</tr>
<tr>
<td>Scallop case 9</td>
<td>As baseline, but $\kappa = 3 \times 10^{-5}$ m$^2$ yr$^{-1}$ and with 10 vol% regolith at depths between 0.95 and 1 m</td>
</tr>
<tr>
<td>Crater case 1</td>
<td>Baseline; 50°N, TI 150 J m$^{-2}$ K$^{-1}$ s$^{-1/2}$, albedo 0.25, 0.035 m equilibrium ice depth, $\kappa = 10^{-4}$ m$^2$ yr$^{-1}$, 1 vol% regolith in restricted thickness (1.5 m) of clean ice, paraboloid initial disturbance with depth 5 m and diameter 30 m</td>
</tr>
<tr>
<td>Crater case 2</td>
<td>As baseline, but at 35° obliquity (128 pr. μm water vapor), 0.011 m equilibrium ice depth</td>
</tr>
<tr>
<td>Crater case 3</td>
<td>As baseline, but at 45° obliquity (94 pr. μm water vapor), 0.019 m equilibrium ice depth</td>
</tr>
<tr>
<td>Crater case 4</td>
<td>As baseline, but at 20° obliquity (4 pr. μm water vapor), unstable ice</td>
</tr>
<tr>
<td>Crater case 5</td>
<td>As baseline, but intersecting thick clean ice (&gt; initial crater depth)</td>
</tr>
<tr>
<td>Crater case 6</td>
<td>As baseline, but with no clean ice and an initial ice depth of 1.5 m (well below equilibrium) producing a loose cover</td>
</tr>
<tr>
<td>Crater case 7</td>
<td>As baseline, but intersecting thick clean ice with 10 vol% regolith</td>
</tr>
<tr>
<td>Crater case 8</td>
<td>As baseline, but intersecting thick clean ice with an initial lens of ice-free regolith coating the crater interior (loosely simulates eolian deposition)</td>
</tr>
<tr>
<td>Crater case 9</td>
<td>As baseline, but with ice-cemented regolith extending to 1.5 m depth above thick clean ice</td>
</tr>
<tr>
<td>Crater case 10</td>
<td>As baseline, but with only pore ice</td>
</tr>
</tbody>
</table>

* Chamberlain and Boynton (2007).*

4 Model scenarios

We consider two families of scenarios, one intended to reproduce the general morphology of scalloped depressions and the other to look at crater expansion in icy targets. Scenarios are summarized in Table 2 and discussed below.
In the first group of scenarios we explore the development of scallop-like landforms from minor disturbances to a level, equilibrium landscape. An equilibrium thickness of dry regolith overlies a thick layer of excess ice. The initial disturbances in these scenarios are generated by removal of part of the protective lag layer in order to locally destabilize the surface. This method may not directly correspond to any physical process; the style and nature of realistic initial disturbances is discussed further in Section 5. We examine the effects of varying the size and thickness of this disturbance.

We focus on modeling the evolution and morphology of an archetypal scalloped depression like that in Fig. 1a. Many scallops deviate from this shape, but only the simplest case is examined here. Complex landscapes may develop from mergers and interactions of many scallop-like features (Fig. 1b), but we do not model the development of such a landscape in this paper apart from the inclusion of periodic boundary conditions.

In the second family of scenarios, we examine expansion of a crater cutting into or through an ice-rich layer and consider scenarios that might produce a funnel-like profile. The initial conditions for these scenarios comprise parabolic cavities cut into or through ice-rich layers. For simplicity, we ignore uplifted crater rims and ejecta deposited around the rim. This could thicken the lag and inhibit expansion, but much of the ejecta is likely to be ice with a low regolith content, which will be highly prone to sublimation in this scenario (Dundas and Byrne, 2010), and the non-ice component will probably be prone to eolian erosion. This assumption is questionable in cases where pore ice is dominant, but the same initial geometry is retained to simplify comparisons. Impact heating and melting are not considered; this will drive some sublimation on the crater walls but the heat will be rapidly radiated away. In both families of scenarios we assume that the initial ice depth is at or near the equilibrium depth, unless otherwise noted.

Given the occurrence of mass wasting, there is no true equilibrium or final landform shape. If the model were run indefinitely, the eventual result would be a level surface and an equilibrium ice table, although the relative thicknesses of clean and pore ice would vary. We chose a model run time of 250,000 Mars Years (MY) for all cases, which allows considerable landform development.

5. Results

Features resembling scalloped depressions develop spontaneously from initial disturbances (Figs. 7–9). The local disturbance causes sufficient collapse to generate a warm equator-facing slope that is significantly out of equilibrium in the present climate.

Ground ice beneath this slope sublimes, causing the slope to both retreat and become shallower. If the underlying ice is very clean, a significant amount of sublimation is required to restore an equilibrium cover. The slope retreats until the combination of sublimation and mass movement lower the slope and thicken the lag, restoring equilibrium. Retreat of the pole-facing scarp is much slower because it is colder, making ice more stable there, but some retreat does occur, driven by mass wasting. The resulting landforms strongly resemble simple scalloped depressions (Figs. 1–4).

Variations of initial conditions have unsurprising effects on the scallop shape (Fig. 9). For a given model timescale, a content of 10 vol% lithic regolith material rather than 1% (case 2, still consistent with very clean ice) causes an equilibrium lag to build up much more quickly and leads to a smaller, shallower depression.

![Fig. 7. Growth progression of model scallop (case 1). Panels show the initial disturbance and the developing landform after 20,000, 100,000 and 250,000 Mars Years (MY). North is up. The original surface is at 4 m (arbitrary reference).](image)

![Fig. 8. (A) South-to-north profile along the axis of the model scalloped depression from case 1 showing the initial and final (250,000 Mars years) position of the surface and the ice table. (B) Profiles of the development of the scallop from case 1, plotted every 10,000 Mars years.](image)
after the same period of evolution. A limited thickness of excess ice also leads to a shallower scallop. Smaller or shallower initial disturbances (e.g., cases 5–6 in Fig. 9) produce smaller scallops, while a wider disturbance (e.g., case 7) yields a wider and deeper landform. Low mass movement diffusivity (e.g., case 8) prevents the depression from widening significantly and leads to a deeper, narrower feature. (Note that since the assumed diffusivities are likely upper bounds, the regolith content can be traded against the mass...
movement diffusivity to generate landforms of a certain depth.) An unstable surface (e.g., case 4) produces a strange landform where initial collapse and rapid infill produce a lump of thick regolith that resists further sublimation as the surroundings deflate. The occurrence of a layer with higher regolith content (case 9) produced a small kink in the slope with little effect on the overall morphology.

Several crater cases produce expansion similar to observed landforms (Figs. 10–12). The baseline case 1 (current conditions, crater cutting through an ice layer) initially expands asymmetrically, but ultimately becomes more symmetric as creep becomes more important than sublimation (Figs. 10 and 12). Ice beneath the equator-facing slope is destabilized at all depths in this example (Fig. 11a). Cases 2 and 3 (higher obliquity) are more symmetric initially. All three ultimately produce landforms broadly similar to the observed craters, with bowl-shaped central depressions and wider, shallow-sloping rims. At 20° obliquity (case 4), the equator-facing slope retreats but ice is unstable under the entire surface, so the final landform does not have the funnel-shaped form. However, in reality obliquity is not constant over these timescales.

Other perturbations have varied effects. In case 5, a crater cutting into thick ice also expands and sees the upper slopes retreat. The final landform does have a central depression and surrounding shallow apron, although the break in slope is not as well-defined as in cases 1–3. Although ice is destabilized under some slopes, the effect on the landform is counterbalanced by inward mass movement. A thicker pore ice cover (case 9) over clean ice produced changes to the crater shape and a smaller expanded lip. Conditions of pore ice with a thick or thin loose, desiccated cover (cases 6 and 10) produced little expansion. Less-clean ice (case 7) gave modest expansion but without well-defined breaks in slope. In case 8, the crater cavity was partially coated with a soil lens of varying thickness (greatest at the crater center). This loosely approximates the results of eolian dust infilling the crater, and produces a distinctly expanded crater with an inner cavity.

6. Discussion

6.1. Scalloped depressions

The thermokarst pits produced by our model reproduce the basic morphologic features of martian scalloped depressions. Disruption of an equilibrated surface with excess ice is thus a plausible mechanism for producing such depressions. The initial disturbances that we have used to generate scalloped depressions are arbitrary, but the parameters of the disturbance do have an effect on scallop morphology. These parameters could be tuned to better reproduce the detailed dimensions of real scallops, but because the initial disturbance is arbitrarily imposed this would not be meaningful. It remains to be seen whether realistic disturbances and climate variations produce any characteristic morphologies, and which properties of scallop morphometry are ultimately controlled by properties of the ground properties and climate conditions rather than the triggering instability. The precise evolutionary sequence observed in our model (Fig. 7), with the initial downward collapse followed by retreat of the poleward slope, may also be the result of the imposition of a sudden disturbance. Gradual disturbances could result in different landforms, although the factors driving growth will be similar.
More realistic triggers for an initial instability include small impacts and local darkening or erosion due to eolian processes; under some circumstances evolution of thermal contraction cracks might also trigger instabilities, since thermal contraction polygons develop small-scale topography (Sletten et al., 2003) that affects the local ice depth (Mellon et al., 2009). Although dust devil tracks are transient features and very elongated, some locations are especially conducive to their formation and could produce a persistent albedo effect. Dust devil tracks are particularly prominent in the higher mid-latitudes (Whelley and Greeley, 2008). Another possibility is differential changes in the equilibrium depth as the climate changes. In this case, a surface with ice in equilibrium with some initial minor topography might develop instabilities as the equilibrium depth changes differently for different slopes. Lefort et al. (2009) suggested that the initial stages of scallop development are hummocky terrain and subtle scarps, progressing to shallow depressions and eventually scalloped landforms; such a sequence may be consistent with this scenario. A related possibility is that a uniform increase in stability depth could cause local collapse if the upper part of the excess ice layer is heterogeneous. Thermal contraction cracks, in particular, might develop local ice concentrations if frost preferentially forms in cracks (Fisher, 2005), and ice lens formation (Sizemore et al., 2015) is highly dependent on local soil properties. Local heterogeneities in martian ground ice are observed (Mellon et al., 2009; Dundas et al., 2014), but we have not included them in our model in order to isolate the effects of disturbances. Given any disturbance that causes local destabilization and subsidence, the same basic effects will cause retreat of warm slopes with unstable ice until a new equilibrium is reached, although the scale and style of disturbance could affect the landform scale and details of the morphology.

The modeled scallop evolution largely involves sublimation and retreat of the equator-facing slope, with mass wasting causing the pole-facing scarp to retreat slightly. This style of evolution most resembles the models proposed by Morgenstern et al. (2007), Lefort et al. (2009), although the present model does not include any eolian effects as invoked by Lefort et al. apart from the simplistic imposition of a mantle in crater-expansion case 8. The equatorward-growth model proposed by Ulrich et al. (2010), Séjourné et al. (2011) has scallops forming by retreat and steepening of the pole-facing scarp at high obliquity (45°), when the north-facing slope is the warmest surface. Our model demonstrates retreat of warm slopes, but the retreating slopes shallow over time rather than steepening. The exception is the transient period when the initial disturbance is collapsing, when steep slopes are produced by the floor falling rather than the walls retreating.
Additionally, since scallop scarps are typically concave up, retreating slopes should tend to become shallower over time unless the basal slopes have greater stability depths (less stable ice) or systematically lower regolith thicknesses. The former is unlikely since the equatorward-growth hypothesis stipulates that it is the steep slope that is unstable; the latter is unlikely since mass wasting inflow and greater ice loss should make regolith thicknesses at the base of the scarp greater than near the top. Shadowing (not accounted for) will also be maximized at the base of the slope. Although we expect retreating slopes to shallow over time, they can remain significantly steeper than the floor (Fig. 8b), which is observed in many scallops.

Nevertheless, the model uncertainties and parameter space are sufficiently large that we do not rule out the possibility of maintaining a steep scarp while retreating under some conditions, particularly if eolian erosion or CO$_2$ processes strip away parts of a developing lag. Lower values of diffusivity would also favor maintaining a steep scarp during retreat since the accumulation of extra material at the base would be reduced. In comparison, steep scarps in the south polar “Swiss Cheese” terrain do remain steep while retreating (Byrne and Ingersoll, 2003). Those landforms, however, are not directly analogous, due to the lack of a lag thick enough to slow sublimation. Moreover, the model results of Byrne and Ingersoll (2003) that best reproduced observed wall slopes had albedo decreasing with depth (compensating for shadowing and increasing the retreat rate of the base of the slope), which is not the case for scallops. We note that at the latitudes of scalloped depressions, ice is less stable under current climate conditions than it is at high obliquity (e.g., Chamberlain and Boynton, 2007), which should favor thermokarst feature development in the present climate—it is possible that sublimation is ongoing under many equator-facing slopes. However, low diffusivity could permit slopes to retreat while remaining steep and some pole-facing slopes could have retreated under high-obliquity conditions. Additionally, there may be conditions where processes beyond the scope of this model (like eolian erosion or CO$_2$-driven activity) are important, and the evolution of slopes under those conditions will be different.

Séjourné et al. (2011) noted a systematic variation in thermal contraction polygon morphology within scallops. They noted low-centered polygons (attributed to ice table aggradation) on pole-facing scarps, transitioning to high-centered polygons (attributed to sublimation) further from the scarp. A retreating equator-facing slope is consistent with this interpretation, as the ice table may aggrade on cold pole-facing slopes while polygons degrade on the retreating slope. Séjourné et al. (2011) did not note any transition in polygon size, while our model indicates significant variation in ice table depth (Fig. 8a). On a given slope polygon dimensions should be larger for a deeper ice table (Mellon et al., 2008a), but small polygons might remain visible as the ice retreats, especially if the lag eventually becomes thick enough to inhibit cracking. Haltigin et al. (2014) reported that at the landscape scale, polygons in Utopia are systematically smaller in lower-elevation areas, and also noted dramatically higher polygon densities at the base of scallop scarps. One possible factor is that the hectometer-scale inter-scallop polygons of Utopia Planitia may not be of seasonal thermal contraction origin, but instead relate to deformation of a thin surface layer (Yoshikawa, 2003), which could equate to the ice-rich surface unit inferred by Stuurman et al. (2014).

6.2. Crater expansion

Model results also indicate that sublimation of excess ice is a plausible way to produce expanded, funnel-shaped craters. Matches to the observed morphology, with a distinct break in slope, occurred when a crater cut entirely through a layer of excess ice. Various forms of expansion occurred for other stratigraphic columns that include excess ice, even without the occurrence of distinct layering, although the breaks in slope were less distinct. The observed breaks in slope are subtle (Fig. 6), and therefore, it would be premature to conclude that they indicate the base of an excess ice layer. The alternative scenarios with thick excess ice are consistent with observations, as they do produce a broad depression around a deeper central pit. In these cases, rapid accumulation of a lag near the bottom of a crater halts expansion while the upper walls continue to retreat since mass movement keeps the ice slightly unstable. Case 8 (with a lens of deposited sediment) could explain the observed transition to a fine-grained texture in the inner depression (Fig. 5). Those cases are also consistent with the greater expansion of larger craters, whereas a finite layer thickness might yield comparable expansion regardless of crater size, depending on the importance of mass wasting.

It is clear that crater expansion is a likely consequence of impacts into targets with substantial volumes of shallow excess ice, and therefore should be expected to occur on Mars. At high latitudes Mars has few preserved small impact craters, and thermokarstic growth may contribute to crater erasure. We note that craters in a target that is undoubtedly very clean ice (the north polar cap) exhibit some signs of wall retreat by ablation (Byrne et al., 2010), but are primarily erased by ice accumulation in the interior (Banks et al., 2010). However, at lower latitudes, higher temperatures will favor sublimation rather than frost accumulation, so infilling will be primarily by atmospheric dust and may be substantially slower.

Although initial crater expansion is asymmetric in the present climate, it becomes more symmetric over time. How does this fit with the poleward growth that appears to occur in scallops? The reason is likely to be the relative importance of sublimation and mass movement. The early crater slopes are steep, so (for the chosen value of diffusivity) mass movement is relatively more important for the craters. Additionally, the lower equilibrium depth used for the crater runs means that sublimation alone will become less important early, making further sublimation dependent on the effects of mass movement on the lag. Alternatively, early asymmetry in the model was less pronounced at higher obliquity, and in fact was slightly biased equatorward at 45° obliquity. This temperature distribution is consistent with the high summer temperatures of pole-facing slopes at high obliquity noted by Kreslavsky et al. (2008). Either effect (or both) may be involved in the relative symmetry of many martian expanded craters, and some amount of eolian erosion may influence their shapes as well. Additionally, inclusion of ejecta would introduce a slight equator-facing slope to the surface adjacent to the pole-facing crater wall. This would reduce the stability of ice there, although the effect may be negated by the thickening of the lag due to ejecta deposition. We note that some scalloped depressions could initiate as impact craters under conditions where sublimation is extensive enough to completely degrade the original crater.

6.3. Implications for martian ground ice

A fundamental point (noted by previous studies of scalloped depressions, but highlighted by this model) is that generation of sublimation thermokarst pits with depths of many meters suggests that excess ice thicknesses at least this large previously existed. Furthermore, such ice is likely to remain in adjacent areas that have not undergone thermokarst collapse, or those areas should also be depressed. Consistent with this idea, SHARAD radar results suggest tens of meters of clean ice in the shallow subsurface in Arcadia Planitia, not far from the location of many expanded craters (Plaut et al., 2009; Bramson et al., 2014). Subsurface interfaces
indicating thick near-surface clean ice are also detected in Utopia Planitia in the region of scalked depressions (Stuorman et al., 2014). SHARAD can detect ice-rich layers only when thicker than ~10–20 m over significant areas (Seu et al., 2007), consistent with our interpretations of the landforms. Additionally, although Utopia Planitia and the southern highlands are not favorable for detection of fresh impacts, new craters exposing clean ice have been observed near expanded craters in and around Arcadia Planitia (Dundas et al., 2014). Loss of ice that fills the pores of a self-supporting regolith will not cause collapse. This is an assumption of the model rather than a result, but it is well-grounded in physical expectations—by definition, vapor-deposited pore ice occurs in the interstices of a regolith that previously supported itself without ice infill. The bulk ice content of the excess ice layer is uncertain and probably varies; we used a very high ice content (99 vol%, similar to that observed in a local deposit at the Phoenix landing site), and still produced shallow scallops. However, somewhat lower concentrations are plausible, because we used a diffusivity value that is likely an upper limit, thereby allowing a lag to accumulate at the bottom of the scaplo by mass movement.

The implication that significant excess ice remains today could be avoided if the thickness or regolith content of clean ice were spatially heterogeneous. It is unlikely, however, that the excess ice originally occurred in pockets that mirrored the shapes expected to be produced by sublimation. Such pockets could occur if snow infilled preexisting thermokarst depressions, but that simply moves the issue back by a generation. Alternatively, scalloped depressions could be initiated by local collapse of a thin or dirty layer of excess ice and develop further by preferential eolian erosion, even after complete removal of ice from the region. This possibility is not included in our current model and we cannot rule it out. There are few bedforms in Utopia to indicate significant eolian transport, but dust devil tracks are common. However, it is not obvious how such a scenario would produce crater expansion, and HiRISE images indicate that small craters and other depressions on Mars are usually sites of local eolian deposition rather than erosion.

Excess ice is geographically widespread at middle and high latitudes (e.g., Boynton et al., 2002; Feldman et al., 2004; Mouginot et al., 2010; Dundas et al., 2014). However, the ground ice concentration and thickness probably varies at regional and local scales, in addition to latitudinally. While ice is stable at high latitudes under most orbital conditions and centimeter-scale thin lags can prevent sublimation there, it is often unstable or barely stable in the mid-latitudes (Mellon and Jakosky, 1995; Chamberlain and Boynton, 2007). Hence, crater expansion and other thermokarstic effects should be more substantial in the mid-latitudes, given appropriate triggers. Therefore, we should expect thermokarst landforms to be largest and most abundant in the mid-latitudes, but only where there are sufficient amounts of excess ice. The optimal locations will be those where ice is unstable often enough to form thermokarst features but not so unstable that it is removed completely. Regions of scallops and abundant expanded craters may highlight local concentrations of thick clean ice, consistent with radar detections of subsurface interfaces. Those concentrations (e.g., Bramson et al., 2014; Stuorman et al., 2014) are consistent with the mid-latitudes making a significant contribution to the global water inventory (c.f. Carr and Head, 2015). Sublimation landforms might also occur on debris-covered glaciers (e.g., Holt et al., 2008), but could be suppressed if the debris lag is thick enough. If ice becomes unstable under most of the surface, discrete landforms are unlikely to form.

The origin of thick excess ice on Mars remains incompletely understood. Vapor diffusion is expected to allow ice to come and go in the pore space of regolith (e.g., Mellon and Jakosky, 1993, 1995; Hudson et al., 2009). However, this is expected to produce mainly pore ice, and only in the relatively shallow subsurface. On Earth, most forms of excess ice result from migration and freezing of liquid water (e.g., French, 2007), and similar processes have been proposed for Mars (e.g., Soare et al., 2015). However, Mars is a cold and dry planet, where liquid water is uncommon, and a variety of processes involving little or no liquid have been proposed. Production of ice in excess of the natural pore space could result from thermal cycling coupled with vapor diffusion (Fishier, 2005). Alternatively, ice lens growth could result from the occurrence of thin liquid films at sub-freezing temperatures (Mellon et al., 2009; Sizemore et al., 2015). A final possibility is the deposition of snow or ice-dust mantles under past climate conditions, preserved to the present by thin lag deposits (e.g., Mustard et al., 2001).

The settings of martian thermokarst landforms provide some insights into the origins of local excess ice concentrations. In some cases expanded craters are confined to high-standing material like the ejecta of older craters (Viola et al., 2015). This confined occurrence could be explained by local preservation of an ice-rich layer beneath ejecta (as suggested for pedestal craters (e.g., Kadish et al., 2009), but expanded craters also occur within material filling the older craters and are absent in some low-lying parts of the ejecta. Therefore, at least some of the excess ice postdates the crater, so ancient frozen floodwaters or ice sheets preserved under the ejecta, if any, cannot be the only source. More likely candidates are ice lensing or deposition of ice-rich mantling deposits. The enhanced vapor-diffusion model suggested by Fisher (2005) is unlikely to be the main source of thick excess ice, since >10 Myr was required to generate even 85 vol% ice in the upper two meters of regolith. Recent modeling indicates that ice lensing can effectively generate excess ice on Mars and is strongly dependent on soil properties (Sizemore et al., 2015). This is consistent with confinement to particular surface units, although it is not clear whether crater ejecta would be particularly favorable. Substantial volumes of shallow excess ice could accumulate by this process in as little as 10^5–10^6 yr, but the amount may be highly variable, depending on regolith properties and the rate of supply from atmospheric water vapor (Sizemore et al., 2015). The latter is likely to vary significantly under different orbital conditions. Hence, it is likely that this process operates on Mars, but not certain whether it can produce the local thicknesses suggested by thermokarst features.

An origin from an ice-rich mantling deposit of variable thickness is consistent with confinement of expanded craters to high-standing areas (Fig. 5d), and the remains of a mantle might comprise the bulk mass of isolated knobs where expansion is observed. Such a deposit has also been invoked by some workers for the host material of scalked terrain (e.g., Morgenstern et al., 2007; Lefort et al., 2009, 2010; Zanetti et al., 2010), which appears to be a surface-draping unit (e.g., Fig. 8a of Zanetti et al., 2010). Modeling of ice deposition at high obliquity (e.g., Madeleine et al., 2009) is consistent with the location of radar reflectors suggesting thick ice (Bramson et al., 2014; Stuorman et al., 2014). However, it is not clear why a mantling deposit would show correspondence with the ejecta of older craters, as suggested by some expanded craters. A second issue is the age of a mantle deposit. Estimates of the age of the most recent mantling include ~0.1 Ma (Mustard et al., 2001), 2.1–0.4 Ma (Head et al., 2003), and either ~2.1 or ~5 Ma (Schon et al., 2012). Very young mantles would likely postdate the fields of secondary craters that have undergone expansion (Fig. 5e; Viola et al., 2015) and are inconsistent with widespread boulders on the surface. However, it is likely that recent mantling is part of a long and varied history of deposition and degradation. Kadish and Head (2014) suggest recurring deposition of mid-latitude ice over the last 200 Ma and likely longer, and several workers (Skinner et al., 2012; Soare et al., 2012) propose Mid-Amazonian deposition of widespread ice-rich
loess. Viola et al. (2015) suggest that expanded secondary craters in Arcadia Planitia are implying long-term persistence of ice. The current extent of a postulated remnant icy mantle depends on the assumed emplacement time (Schorghofer and Forget, 2012), while the time required for impacts to scatter boulders across the surface is not known and should depend on the mantle thickness. If the excess ice producing thermokarst features is the surface is not known and should depend on the mantle age of at least tens of Myr, coupled with slow and incomplete degradation, is favored at those locations. Thorough regional studies and morphological analyses as well as landed science will be needed to fully understand the implications of thermokarst landforms for the origins and history of ice.

6.4. Sublimation thermokarst on Mars

The modeling above predicts the morphologies produced by sublimation and mass movement processes on Mars. The results demonstrate that such a model reproduces the major characteristics of observed landforms on Mars. Although it is not possible to rule out proposed past conditions of widespread melting, melting does not need to be invoked.

As noted above, our model makes several simplifications that must be understood when considering the implications for martian landforms. The real history of thermokarst morphologic development on Mars is undoubtedly more complex than modeled in this paper, which assumes a simple ice table structure and steady-state climate. Heterogeneities in the ice will affect the morphology. We find that substantial sublimation landforms can develop on timescales on the order of $10^4$–$10^5$ yr. Since our landscape diffusivity is likely an upper bound, and our triggering disturbances are abruptly imposed, real features probably develop more slowly. The tens-of-Ma age of expanded secondary craters suggested by Viola et al. (2015) is consistent with this idea, since mass movement may play a role in thinning lag deposits and advancing sublimation. Additionally, climate variations and mantling by surface ice deposits (e.g., Madeleine et al., 2009) could suspend surface evolution by sublimation. Better information about landscape diffusivity and/or the age of martian features is needed to improve quantitative interpretations of landform morphology. The vertical relief produced under conditions where sublimation is dominant will likely be greater than we have modeled for lower diffusivities, since mass wasting will contribute less to lag development and infill (e.g., Fig. 8b; this effect is seen in comparing Fig. 9a and h). This is consistent with the substantial depths of some scalloped depressions (e.g., Lefort et al., 2009; Zanetti et al., 2010). The diffusion coefficient for water vapor through the regolith may also be somewhat different than our assumption, which would affect the rate of landform growth.

After a substantial lag has developed, landforms might survive with little modification for much longer periods, depending on the rate of mass movement. Apart from the effects of mass movement, sublimation landforms will equilibrate with the deepest ice and least-stable conditions that they experience for a significant length of time. In the mid-latitudes ice may become completely unstable at low obliquity, but obliquities below 20° have occurred only for short periods over the last 5 Ma. It is interesting to note that, although high obliquities cause stable ice to become more geographically widespread, at high latitudes the stability depth can actually become greater (Chamberlain and Boynton, 2007). It is not necessarily the case, however, that landforms reflect only the least stable conditions that they have experienced, because growth is history-dependent. It remains to be seen what specific aspects of thermokarst morphology can be linked to details of their history.

Scallop-like features without the classic poleward orientation are seen elsewhere on Mars (Fig. 13). One possibility is that this reflects primarily eolian erosion of the regolith on the steep slope, as suggested by Lefort et al. (2009) for pole-oriented scallops. Halligin et al. (2014) suggested that wind could enhance sublimation on pole-facing scarps, but for thick lags the rate of sublimation will be controlled by diffusion. For very shallow ice, however, free and forced convection of water vapor should control sublimation rates. Wind-driven forced convection in particular should be affected by the orientation of the surface relative to the wind, but this effect will only be important while the ice is very shallow. Initial experiments (Dundas, 2015) suggest that this effect is weak. These depressions north of Alba Mons (Fig. 13) also differ from those in Utopia Planitia and the Amphitrites/Peneus region south of Hellas in that they commonly have central knobs or pits, which might indicate a different formative mechanism. However, they have not yet been studied in sufficient detail to assess their origin.

It is intriguing to consider the ultimate fate of thermokarst landforms upon removal of all excess ice. It might be supposed that no trace of the landform would remain, but inward mass movement and trapping of eolian material while the surface was depressed could result in remnant hummocks serving as markers for vanished thermokarst features. This effect is somewhat analogous to the formation of thermokarst mounds from the residual material during degradation of terrestrial ice-wedge networks (e.g., French, 2007), but could develop from thermokarst landforms in a uniform surface. Fig. 9d shows an evolutionary step in this direction where a mound has developed on the floor of a sublimation pit. This morphology has an intriguing similarity to mound-in-pit features formed in ice-rich targets in the mid-latitudes (Mangold, 2003; Head and Weiss, 2014), and could also be involved in the formation of mid-latitude circular mounds (c.f. Dundas et al., 2008; Dundas and McEwen, 2010) and isolated hummocks on the northern plains. More work is needed to better understand this evolution, particularly since we have used a likely overestimate for the mass-motion diffusivity, but gradual eolian infill could produce similar effects.

The basic similarity of model results and observed landforms provides a major step toward removing the uncertainties in the sublimation model noted by Soare et al. (2011). First, although we do not address the lack of a terrestrial analog landscape, sublimation-dominated environments are rare on Earth and it is
necessary to theoretically consider the likely nature of such a surface. However, such conditions are expected on current and recent Mars. We demonstrate that scallop-like landforms are an expected result of disturbances in this environment. Although the Antarctic Dry Valleys are in many ways an excellent analog, the high latitude reduces the directionality of insolation, and the relative importance of sublimation and mass-movement processes could be different. Soare et al. (2011) also noted that the grain size of sands in Utopia may be very different from that in the Antarctic Dry Valleys. Our model makes no assumption about grain sizes except to the extent that they affect the diffusion coefficient, which is based on Mars regolith analogs. Reasonable differences would only affect the rate of evolution, not the general morphology. We also find that the slow evolution of scalloped landforms means that the surface is always covered by a lag, and the soil-over-ice stratigraphy is never exposed in the walls, an objection raised by Soare et al. (2011). The origin of ridges within the scallops remains uncertain, but plausible explanations consistent with sublimation include layering in the substrate and climate-driven variations in the rate of slope retreat. In agreement with Soare et al. (2011), the ice thicknesses required to produce the largest scallops are significant and unlikely to be due to vapor diffusion alone, particularly since excess ice is required.

7. Conclusions

A model of landscape evolution including sublimation and creep of surface material reproduces the basic morphology of scalloped depressions on Mars, starting from initial disturbances over a very ice-rich subsurface. In the model, scallops grow by collapse and retreat of the warm slope, which also shallows over time. The model also demonstrates crater expansion by sublimation in a manner broadly similar to features observed on the northern plains. This indicates that both types of features are consistent with formation by sublimation thermokarst processes. The similarity of modeled and observed landforms supports sublimation models and suggests that significant amounts of excess ice (a prerequisite for sublimation thermokarst) remains in the subsurface near those landforms. Moreover, development of such sublimation landforms is an expected result of the current environment on Mars if excess ice is present and disturbed from equilibrium. Landscapes and processes involving substantial melting have been proposed as analogs and cannot be ruled out, but there is no need to invoke such conditions, since sublimation features are an expected consequence of the recent climate. Given the considerable thicknesses implied, the origin of regional concentrations of ice may be atmospheric deposition, but it remains to be seen whether survival of ice in such a scenario is quantitatively consistent with timescales required to garden boulders to the surface. An alternative possibility is growth of ice lenses in certain materials, and in some cases older ice sheets may be preserved beneath crater ejecta. Detailed regional studies will be needed to understand the origins of excess ice.

Acknowledgments

This work was funded by the Mars Reconnaissance Orbiter HiRISE project and Mars Fundamental Research Program grant NNH13AV59I. We thank Ken Herkenhoff and Justin Hagerty for helpful comments on an early draft. Two reviewers provided useful comments.

References


