Cratering saturation and equilibrium: A new model looks at an old problem

James E. Richardson *

Center for Radiophysics and Space Research, Cornell University, Ithaca, NY 14853, United States

Abstract

Recent advances in computing technology and our understanding of the processes involved in crater production, ejecta production, and crater erasure have permitted me to develop a highly-detailed Cratered Terrain Evolution Model (CTEM), which can be used to investigate a variety of questions in the study of impact dominated landscapes. In this work, I focus on the manner in which crater densities on impacted surfaces attain equilibrium conditions (commonly called crater 'saturation') for a variety of impactor population size-frequency distributions: from simple, straight-line power-laws, to complex, multi-sloped distributions. This modeling shows that crater density equilibrium generally occurs near observed relative-density (R) values of 0.1–0.3 (commonly called 'empirical saturation'), but that when the impactor population has a variable power-law slope, crater density equilibrium values will also be variable, and will continue to reflect, or follow the shape of the production population long after the surface has been 'saturated.' In particular, I demonstrate that the overall level of crater density curves for heavily-cratered regions of the lunar surface are indicative of crater density equilibrium having been reached, while the shape of these curves strongly point to a Main Asteroid Belt (MAB) source for impactors in the near-Earth environment, as originally stipulated in Strom et al. [Strom, R.G., Malhotra, R., Ito, T., Yoshida, F., Kring, D.A., 2005. Science 309 (September), 1847–1850]. This modeling also validates the conclusion by Bottke et al. [Bottke, W.F., Durda, D.D., Nesvorny, D., Jedicke, R., Morbidelli, A., Vokrouhlicky, D., Levison, H., 2005. Icarus 175 (May), 111–140] that the modern-day MAB continues to reflect its ancient size-frequency distribution, even though severely depleted in mass since that time.

1. Introduction

Cratered terrain on the solid surface of a Solar-System body provides us with a valuable record of that surface’s bombardment history, material properties, weathering mechanisms and rates, and other endogenic processes. On ‘old’ surfaces with very low weathering rates (or other crater erasure mechanisms), the density of impact craters can reach equilibrium conditions, where for each new crater formed, a crater of roughly the same size is erased, and crater counts (over a given size range) level off as a function of time and further bombardment. The question as to when crater density equilibrium (also called ‘saturation equilibrium’ or just ‘saturation’) conditions occur and what such conditions look like is a long standing problem in the study of cratered surfaces, one which dates back to the intense studies conducted of the lunar surface during the build-up to the Apollo missions. Over time, the impact cratering community divided itself into roughly three viewpoints on this issue:

1. That variations observed in the shape of crater density curves for heavily-cratered regions of the Moon are indicative of a ‘production population’ (that is, a crater population that directly reflects its parent impactor population) and therefore, such regions are not in equilibrium. The fact that such variations are almost identically repeated on other inner Solar-System bodies, such as Mercury and Mars, lend credence to this viewpoint (Marcus, 1970; Woronow, 1977a,b, 1978; Chapman and McKinnon, 1986; Strom et al., 2005).

2. That the nearly identical overall levels observed in the crater density curves for heavily-cratered regions on Mercury, Mars, and the Moon (within a factor of 2 of each other) are indicative of these regions having reached an equilibrium state (saturation), and therefore, variations observed in the shape of these curves must be Poisson statistical in nature or indicative of endogenic crater erasure processes: that is, given time, such surfaces will eventually – given no additional endogenic crater erasure – show a straight-line (in log–log space) crater density curve (Hartmann, 1984, 1988, 1995; Hartmann and Gaskell, 1997).

3. That views (1) and (2) are both partially correct. On the one hand, these heavily-cratered regions do indeed represent crater density equilibrium conditions. On the other hand, even after
equilibrium has been reached, variations observed in the crater density level as a function of crater size continue to reflect variations in the distribution of impactors which produced the crater populations: similar, but not identical, to a production population (Chapman and McKinnon, 1986).

Typical examples of the crater density curves under debate are shown in Fig. 1, courtesy of Chapman and McKinnon (1986).

One method for unraveling this problem is the use of scale models (either computer-generated or physical) to investigate the crater production and erasure process as a function of increasing crater density. The pioneering work for this sort of modeling was presented by Gault (1970), in which a physical model was produced to simulate a cratered terrain using a 2 m × 2 m sandbox and a variety of small explosive and projectile devices to produce craters of various sizes on this model surface. Although of limited dynamic (crater-size) range, this model demonstrated that crater density equilibrium conditions generally occur at an overall density level of 1–10% of what Gault termed ‘geometric saturation:’ the crater density achieved when craters of the same size are placed rim-to-rim in a hexagonal close-packed arrangement. His modeling work also showed that when the impactor population has a cumulative log–log slope of <–2, equilibrium crater density conditions will be reached first by the smallest craters, then by larger craters, with the equilibrium crater population having a power-law slope of approximately –2 (not that of the steeper impactor population). Thus, he found that for a steeply-sloped impactor population (he did not explore the affects of a shallow-sloped impactor population), the relative age of the cratered terrain can be determined by comparing the position of the ‘knee’ in the crater density curve (the inflection point between the smaller craters, in equilibrium, and the larger craters, not yet in equilibrium) between different areas. Fig. 7, later in this work, shows a modeled example of this form of crater density equilibrium attainment. Gault’s work thus seemed to support the second view described above, in that once an equilibrium crater density is reached, the crater population no longer follows its production population and stabilizes at roughly the same overall level (1–10% geometric saturation), as observed in heavily bombarded, small crater (<1 km) regions on the lunar surface (Hartmann, 1988; Hartmann and Gaskell, 1997).

The first attempt to model the crater density evolution of a planetary surface via computer was performed by Woronow (1977a,b, 1978) who, rather than modeling complex topography in three-dimensions, developed a Monte-Carlo method for creating and erasing representative, circular, crater ‘rims’ in two-dimensions only; that is, a geometric model. Limited by the computer technology of that time, Woronow’s models were lacking both in geometric resolution (monitoring only selected points around each crater’s rim) and in impactor size range (having dynamic ranges of only 16 or 32 between the smallest and largest impactors). Because of these limitations, Woronow’s models displayed equilibrium crater density levels that are far above (by an order of magnitude or more) the crater density levels actually observed on heavily cratered surfaces, and hence seemed to support the first viewpoint described above: that heavily cratered surfaces in the inner Solar System have not yet reached equilibrium and therefore continue to display a production population. About a decade later, however, Chapman and McKinnon (1986) revisited Woronow’s Monte-Carlo based, geometric crater-rim modeling technique, utilizing higher rim resolutions and a much larger dynamic crater-size range. Their work demonstrated a different conclusion. When fully circular crater ‘rims’ are monitored and a sufficient dynamic impactor range is employed (a factor of 128–200 in their work), modeled crater density equilibrium levels are (a) quite close to those actually observed in heavily-cratered regions, but (b) will continue to mimic, or follow variations present in the parent impactor population. Thus, the work of Chapman and McKinnon (1986) seemed to support the third viewpoint described above.

The first attempt to computer-model the evolution of a cratered terrain in three-dimensions was performed by Gaskell (1993), Hartmann and Gaskell (1993, 1997), using a fractal-based digital elevation map (DEM) model which monitored the changing landscape (as successive craters were emplaced) on a variety of fractal scales. Rather than scaling from impactor to crater size, the model emplaced craters directly, using a variety of straight-line power-law distributions. The ejecta coverage produced from each crater was computed by estimating the volume excavated by each crater and distributing it around the crater such that the resulting blanket exhibited a ~3 power-law slope with distance from the crater rim; in effect, simulating gravity-scaled cratering (see Section 2.3). For both their medium-sloped ‘primary population’ (~1.83 cumulative power-law slope) and steeply-sloped ‘secondary population’ (~3.5 to ~4.0 cumulative power-law slope) of craters, Hartmann and Gaskell (1997) showed that crater density equilibrium occurs at roughly the same overall level as that observed on actual heavily cratered surfaces, and that the resulting crater population tends to follow a cumulative power-law slope of roughly ~1.83 even when the production population is steeper. This work thus supported the findings of Gault (1970) and seemed to support the second viewpoint described above, in that once an equilibrium crater density is reached, the crater population no longer follows a production population, and stabilizes at roughly the same overall level. However, shallow-slope impactor populations (i.e. ~1–1.83 cumulative power-law slope) were not investigated. Despite its sophistication, the primary drawback of this model was the lack of automatic (computer) crater counting, which severely limited the number of model runs and the number of time steps within each model run, for which crater counts could be manually performed – thus limiting its use to only a few specific case studies.

In this current work, I present a new Cratered Terrain Evolution Model (CTEM) which takes advantage of modern computing tech-
ology, recent advances in the impact cratering scaling-laws, and our understanding of seismically-induced crater erosion to produce a fully three-dimensional model of crater production and erosion on a given target surface. This new model includes both downslope regolithic migration and automatic crater counting. Each simulation uses Monte-Carlo sampling of a user-supplied impactor population to bombard an initially clean digital target surface (possessing periodic boundary conditions) as a function of time, and then generates surface ‘images’ and crater counts at specified time intervals throughout the run. Although this model can be used in a variety of ways, the purpose of this particular investigation is to demonstrate how and when crater density equilibrium (crater ‘saturation’) conditions are reached, and how these conditions change with changes in the slope and shape of the parent impactor population.

2. Cratered terrain evolution model

2.1. Basic model description

As mentioned above, this Cratered Terrain Evolution Model (CTEM) uses Monte-Carlo techniques (Press et al., 1992) to populate a model surface with fresh craters as a function of time, and simultaneously allows previous craters to be eroded by superposing new craters, covered with impact ejecta, and seismically shaken and degraded by nearby impacts. The model is loosely based upon a more primitive, previous model (Richardson et al., 2005), and consists of 11 two-dimensional matrix layers (representing the surface area to be studied) to form a fully three-dimensional surface model. The primary model layer stores a digital elevation map (DEM), a secondary layer tracks regolith depth, while the remaining nine layers store crater reference and deviation information for each pixel of the DEM (see Fig. 2). Three layers are used to store crater position information, three layers store crater profile information, and three layers store crater profile deviation information (explained further in Section 2.4). The purpose in having three separate sets of ‘tracking’ information per pixel is to permit the superposition of smaller craters on top of much larger craters, while preserving information about the larger crater(s) beneath, one model ‘level’ for each log-decade of crater size permitted. The model permits a dynamic crater-size range of 1024, from smallest (1 pixel) to largest (1024 pixels) crater, but also permits the degradation of visible craters by sub-pixel craters of less than its radius. Many early works use the projectile density \( \rho \) and impact velocity \( V \), where \( \rho \) is the projectile density and \( V \) is the volume of the transient crater produced, and \( m \) is the mass of the impactor, given by \( m = \frac{4}{3} \pi \rho a^3 \) (\( \rho \) is the projectile density and \( a \) is the projectile’s spherical radius);

\[
\pi_2 = \frac{g}{\rho V^3} \left( \frac{m}{\rho} \right) \left( \frac{V}{V_t} \right) = 3.22 \left( \frac{g a}{V^3} \right). \tag{2}
\]

where \( \pi_3 \) is the dimensionless, \( \pi \) group parameters are:

\[
\pi_v = \rho V \frac{m_i}{m} \tag{1}
\]

Following each impact, the model enters an Eulerian phase, whereby unstable slopes are permitted to collapse and regolith can migrate downslope due to the seismic effects of nearby impacts (Section 2.6). At specified time intervals, crater pixel-fragments are automatically counted and used to determine the distribution of actual craters on the surface (down to the smallest crater fragment), and estimate the number of ‘observable’ craters, in addition to taking a snapshot of the DEM at that time. The algorithms used for crater creation and erasure are described in detail in the following sections.

2.2. Impact crater size scaling

In order to estimate the crater size and ejecta produced by a particular impact, we take advantage of the fact that the final crater is usually much larger than the projectile, such that projectile-specific properties (such as diameter, shape, and composition) do not affect the final outcome: a concept referred to as ‘late-stage equivalence’ (Dienes and Walsh, 1970). In this case, only a single, dimensional ‘coupling parameter’, which depends upon the projectile’s total energy and momentum, will affect the size and shape of the cratering result (Holsapple and Schmidt, 1987). When this is the case, a number of power-law scaling relationships have been observed in experimental impacts, and derived mathematically as point-source solutions, that link impacts of different sizes, velocities and gravitational accelerations. The derivation of these crater scaling relationships is based upon the Buckingham \( \pi \) theorem of dimensional analysis (Buckingham, 1914) and have undergone extensive development over the years, as described in Holsapple and Schmicht (1980, 1982, 1987), Housen et al. (1983), Schmidt and Housen (1987); and the review work, Holsapple (1993).

In impact cratering, the most commonly used set of four dimensionless, \( \pi \) group parameters are:

\[
\pi_v = \rho V \frac{m_i}{m} \tag{1}
\]

where \( \pi_v \) is called the cratering efficiency, \( \rho \) is the target surface density, \( V \) is the volume of the transient crater produced, and \( m \) is the mass of the impactor, given by \( m = \frac{4}{3} \pi \rho a^3 \) (\( \rho \) is the projectile density and \( a \) is the projectile’s spherical radius);

\[
\pi_2 = \frac{g}{\rho V^3} \left( \frac{m}{\rho} \right) \left( \frac{V}{V_t} \right) = 3.22 \left( \frac{g a}{V^3} \right). \tag{2}
\]

where \( \pi_3 \) is called the gravity-scaled size, and is a measure of the importance of gravity \( g \) in the cratering event (\( v_i \) is the impact speed). The factor of \( (4 \pi/3)^{1/3} = 3.22 \) is often neglected, or written as 1.61 if the impactor is placed in terms of its diameter \( d_i \) rather than its radius \( a_i \);

\[
\pi_3 = \frac{g}{\rho V^3} \left( \frac{m}{\rho} \right) \left( \frac{V}{V_t} \right) \tag{3}
\]

where \( \pi_4 \) is called the non-dimensional strength, and is a measure of the importance of target strength \( Y \) in the cratering event. Many early works use the projectile density \( \rho_i \) in place of target density \( \rho_t \) in the denominator, so one must carefully note which form is being used in each study. And lastly:

\[
\pi_4 = \frac{\rho_i}{\rho_t} \tag{4}
\]

where \( \pi_4 \) is the density ratio between target and impactor, and is often assumed to be \( \approx 1 \) in many applications (and is therefore negligible).

Performing dimensional analysis using these four parameters leads to a solution for the cratering efficiency, given as (see Holsapple (1993) for the full derivation):
\[\pi_V = K_1 \left( \pi_2 \pi_4^{\mu_2} + [K_2 \pi_3 \pi_4^{\mu_3}] \right)^{\frac{1}{\mu_4}},\]  

where \(\mu\) and \(\nu\) are proportionality constants. In practice, \(\nu\) can generally be taken as equal to 1/3 at all times, while \(\mu\) is variable between \(1/3 \leq \mu \leq 2/3\), depending upon whether the cratering
event is primarily governed by the impactor’s kinetic energy \((\mu = 2/3)\) or momentum \((\mu = 1/3)\) (Holtsapple and Schmidt, 1987). If we further assume that \(K_2\) is close enough to unity to permit the quantity \(K_1 Y\) to equal an ‘effective’ target strength \(Y\), then we can simplify Eq. (5) to give (Holtsapple, 1993):

\[
\pi V = K_1 \left[ \pi_2 \pi_4^{-1} + \pi_3 \pi_5 \right]^{3/2},
\]

(6)

where \(\pi_2 = (g_Y / v_Y^2)\) and \(\pi_3 = (Y / \rho_i v_Y^2)\).

Bringing everything together, we obtain the desired function for computing the transient crater volume:

\[
V = K_1 \left( \frac{m_i}{\rho_i} \right) \left( \frac{g_Y}{v_Y^2} \right)^{3/2} \left( \frac{Y}{\rho_i v_Y^2} \right)^{3/2},
\]

(7)

where \(K_1, \mu, \) and \(Y\) are experimentally derived properties of the target material. The transient crater-volume \(V\) can be related to the more easily measured transient crater diameter \(D\) or radius \(R\) by:

\[
V = \frac{1}{24} \pi D^3 = \frac{1}{3} \pi R^3,
\]

(8)

where we assume that the transient crater depth \(H\) is roughly 1/3 its diameter \(D\): in experiments this is somewhat variable, between 1/4 and 1/3 (Schmidt and Housen, 1987; Melosh, 1989).

Table 1

<table>
<thead>
<tr>
<th>Material</th>
<th>(K_1)</th>
<th>(\mu)</th>
<th>(Y) (MPa)</th>
<th>(\rho_i) (kg m(^{-3}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>2.30</td>
<td>0.55</td>
<td>0.0</td>
<td>1000</td>
</tr>
<tr>
<td>Sand</td>
<td>0.24</td>
<td>0.41</td>
<td>0.0</td>
<td>1750</td>
</tr>
<tr>
<td>Dry soil</td>
<td>0.24</td>
<td>0.41</td>
<td>0.18</td>
<td>1500</td>
</tr>
<tr>
<td>Wet soil</td>
<td>0.20</td>
<td>0.55</td>
<td>1.14</td>
<td>2000</td>
</tr>
<tr>
<td>Soft rock</td>
<td>0.20</td>
<td>0.55</td>
<td>7.6</td>
<td>2250</td>
</tr>
<tr>
<td>Hard rock</td>
<td>0.20</td>
<td>0.55</td>
<td>18.0</td>
<td>2500</td>
</tr>
</tbody>
</table>

If the force of gravity \(g\) is much greater than the effective yield strength of the target material \(Y\), that is, if it takes much more energy to loft the material out of the crater bowl than it takes to effectively break the material apart, then Eq. (7) can be simplified to:

\[
V_s = K_1 \left( \frac{m_i}{\rho_i} \right) \left( \frac{g_Y}{v_Y^2} \right)^{3/2} \left( \frac{Y}{\rho_i v_Y^2} \right)^{3/2},
\]

(9)
a condition called gravity-dominated cratering. If the force of gravity \(g\) is much smaller than the effective yield strength of the target material \(Y\), that is, if it takes much less energy to loft the material out of the crater bowl than it takes to effectively break the material apart, then Eq. (7) can be simplified to:

\[
V_s = K_1 \left( \frac{m_i}{\rho_i} \right) \left( \frac{Y}{\rho_i v_Y^2} \right)^{3/2},
\]

(10)
a condition called strength-dominated cratering. These two cratering ‘regimes’ are frequently treated as separate end-members in the development of crater scaling relationships, neglecting the effects of viscosity (Holtsapple and Schmidt, 1982). The transition strength \(Y_t\) between gravity- and strength-dominated cratering can be roughly estimated by setting the gravity and strength terms in Eq. (7) equal to each other, as suggested in Holtsapple (1993). This gives:

\[
Y_t = Y_s \left[ \frac{g_Y}{v_Y^2} \left( \frac{\rho_i}{\rho_i} \right)^{1/2} \right]^{1/2}.
\]

(11)

Table 1 lists the applicable material properties for several common target materials, as given in Holtsapple (1993), while, Fig. 3 shows an application of the above impact crater-size scaling relationships, for both the Earth’s Moon and the Asteroid 433 Eros, with each shown for a variety of effective target surface strengths \(Y\). Note the much more severe effect of increasing target strength in the low-gravity environment of the asteroid’s surface, as one would expect by examining Eqs. (7) and (11).

![Fig. 3. Examples of the impact crater scaling-relationship (from impactor to crater diameter) used by the CTEM, for the lunar surface (left) and the surface of Asteroid 433 Eros (right). The lunar plot uses an impact speed of 17.5 km/s (45° incidence) and a surface gravity of 1.63 m/s\(^2\), while the Eros plot uses an impact speed of 5.0 km/s (45° incidence) and a surface gravity of 0.00038 m/s\(^2\). Target material properties are the same in each case, with the primary variable being effective target strength \(Y\), where each line indicates: (bold solid) \(Y = 0\) Pa or pure gravity-dominated cratering, (dashed) \(Y = 5\) kPa, (one-dot) \(Y = 50\) kPa, (two-dot) \(Y = 500\) kPa, (three-dot) \(Y = 5\) MPa, and (four-dot) \(Y = 50\) MPa or pure strength-dominated cratering. The importance of target strength in the low-gravity Eros environment is well demonstrated by this general crater-size scaling-law. As one moves from smaller to larger impactors, the lower curves all eventually merge with the (bold solid) gravity-dominated cratering curve, indicating the transition from strength- to gravity-dominated cratering.](image-url)
It is important to point out that the concept of what is meant by ‘strength’ in impact cratering is presently somewhat fuzzy. Modern theories of dynamic fracture indicate that the actual failure strength of a material should be strongly rate dependent (Grady and Kipp, 1987), a factor not included in the derivation of the crater scaling relationships. Additionally, computation hydrodynamic (CHD) modeling indicates that the strength of the material surrounding an impact is often strongly degraded by shock-wave passage long before the excavation flow clears this material out of the crater interior, such that the excavation stage is governed more by a target’s ‘post-shock’ strength than its initial strength (Croft, 1981; Chapman and McKinnon, 1986; Nolan et al., 1996). Furthermore, impact energy expended in the compaction of porous target material will also manifest itself as a form of ‘strength,’ since our current scaling relationships include only a generic stress variable $Y$, equivalent to an energy expended per unit volume during crater formation, which does not specify how that energy was actually used (Housen and Holsapple, 2003; Holsapple and Housen, 2007). We must therefore be careful to refer to any assumed strength for a target’s ‘post-shock’ strength than its initial strength (Croft, 1981), Chapman and McKinnon (1986), and Melosh (1989). 

For simple craters, the transient crater is approximately 85% the transient crater diameter $D_f$; that is, the crater's momentary diameter prior to gravitational collapse. In order to compute a final crater diameter $D_f$ from the transient crater diameter $D_t$, two expressions are used, one for small, simple craters, and one for large, complex craters – see the discussion in Chapman and McKinnon (1986) for more details. For simple craters, the transient crater is approximately 85% the diameter of the final crater, and the conversion factor is linear:

$$D_f = 1.18D_t,$$  

(12) as described in Chapman and McKinnon (1986) and Melosh (1989). For complex craters the conversion follows a power-law, as described in Croft (1981), Chapman and McKinnon (1986), and Melosh (1989):

$$D_f = D_t^{0.18}D_t^{1.18},$$  

(13) where $D_t$ is a target-specific transition crater diameter, estimated by:

$$D_t = \frac{D_{ref}g_{ref}}{g},$$  

(14) which takes advantage of the $1/g$ relationship for the transition point from simple to complex cratering described in Chapman and McKinnon (1986). Adopting a reference diameter of $D_{ref} \approx 1.35$ km for the Earth ($g_{ref} = 9.81$ m/s$^2$), I use the intersection point between Eqs. (12) and (13) to demarcate the model’s transition from simple to complex craters with increasing crater size. This transition occurs at final crater diameters of $3.3$ km (Earth), $8.8$ km (Mars), and $20.2$ km (Moon), in good agreement with the plots shown in Chapman and McKinnon (1986) and Melosh (1989).

Another useful scaling relationship is the crater formation time $T_f$; that is, how much time passes between the beginning of the coupling phase to the end of the excavation stage. Schmidt and Housen (1987) give the following ‘short-form’ expression for gravity-dominates cratering:

$$T_g = K_{tg} \frac{V_s^2}{g},$$  

(15) A proportionality constant value of $K_{tg} = 0.8$ is given in Fig. 9 of Schmidt and Housen (1987), while Holsapple and Housen (2007) cite a value of $K_{tg} = 0.92$ for sand. Due to the ‘self-similarity’ of all gravity-scaled craters, this constant is applicable to the full spectrum of impact environments and target materials (from sand to hard rock), and we will take advantage of this constant feature in the next section.

2.3. Impact ejecta blanket thickness

In order to estimate the ejecta blanket coverage produced by an individual impact, the CTEM utilizes the ballistic ejecta theory developed in Richardson et al. (2007). This theory takes advantage of a basic concept described in the Maxwell-Z model of ejecta behavior (Maxwell and Seifert, 1974; Maxwell, 1977); namely, that the ejecta flow emerging from the surface at some radius r from an impact site represents a hydrodynamic streamline, which is in a steady-state condition and incompressible. If we also assume that frictional forces beyond those implicit in Eq. (18) (below) are small compared to the forces of inertia, gravity, and strength (that is, inviscid flow), we can use Bernoulli’s principle at the point of ejecta emergence to form an energy balance equation:

$$\frac{1}{2} \rho_e V_e^2 = \frac{1}{2} \rho_e V_e^2 - K_q \rho_e g r - K_T,$$  

(16) where $V_e$ is the emergence velocity (after losses due to friction), and $V_p$ is the effective ejection velocity that we desire (after losses due to gravity and strength). Beginning on the left, the first two terms describe the kinetic energy (or stagnation pressure) of the excavation flow in a single streamline, assuming that upon emergence, the hydrostatic pressure in the flow is zero. The third term describes the mean amount of gravitational potential energy needed to loft each unit volume in the flow (a function of surface radius r), and the fourth term describes the amount of energy needed to fracture or ‘break loose’ each unit volume in the flow (a function of effective target strength $T_f$). Further development yields the following equation for the final ejecta velocity $V_e$ as a function of distance $r$ from the impact site (Richardson et al., 2007):

$$v_e(r) = \left[ V_e^2 - C_{vpg} \rho_e g r - C_{vps} \frac{\rho_e}{\rho_t} \right] \frac{1}{2},$$  

(17) where $V_e$ is given by the gravity-scaled ejecta velocity equation derived by Housen et al. (1983):

$$v_e(r) = C_{vpg} \sqrt{gR_e} \left( \frac{T_f}{K_{rg}} \right)^{-1/2},$$  

(18) along with constant $C_{vpg}$ derived by Richardson et al. (2007):

$$C_{vpg} = \sqrt{\frac{\alpha}{C_{tg}}} \left( \frac{\mu}{\mu + 1} \right),$$  

(19) which, in turn, contains constant $C_{tg}$:

$$C_{tg} = K_{tg} \left( \frac{\pi}{3} \right)^{1/2},$$  

(20) for a crater in which the transient crater has a depth to diameter ratio of 1/3. In effect, $C_{tg} \approx K_{tg}$ within experimental accuracy, and is thus assigned a value of $C_{tg} = 0.85$ for all cratering conditions (see the discussion regarding Eq. (15)).

The constant $C_{vps}$ in the strength term of Eq. (17) is a little trickier to define, but was estimated in Richardson et al. (2007) as:

$$C_{vps} = C_{vpg} \left[ \frac{\rho_e g R_e}{V_f} \right]^{1/2} \left( \frac{R_e}{R_t} \right)^{1/2},$$  

(21) where the addition of $Y_t$ (Eq. (11)) to the denominator maintains equation stability below the transition from strength- to gravity-dominated cratering (low values of $Y_t$). Note that in the above expressions, the purely gravity-scaled, transient crater radius $R_t$ is
found using Eqs. (9) and (8), while the ‘true’ transient crater radius \( R_t \) is found using Eqs. (7) and (8).

Given the above method for determining the ejection velocity at some given distance \( r \) from the impact site, we next calculate the mass of ejecta expelled at that distance \( r \) by dividing the excavated portion of the transient crater into a large number of \( N \) concentric, paraboloid shells: where each thin shell approximates a Maxwell Z-model streamtube of thickness \( dr = R_s N^{-1} \), such that all the material in a given streamtube emerges from the surface at distance \( r \) and at speed \( v_e(r) \). The mass of material \( m \) contained in a particular paraboloid shell is given by:

\[
m_i = K_e \pi \rho_e \left( r_i^2 - r_{i-1}^2 \right),
\]

where \( K_e = C^2_{inl} \), and \( i \) runs from 0 to \( N - 1 \). The volume \( V_i \) of ejecta produced within each paraboloid shell is found by dividing the shell’s ejected mass \( m_i \) by \( 0.8 \rho_e \), where the ejecta blanket is assumed to be 20% more porous than that of the original target material (Melosh, 1989).

The value of \( K_e r \) here and in Eq. (16) represents a mean streamline excavation depth, and ranges from \( r/9.0 \) (at \( \mu = 0.40 \)) to \( r/5.8 \) (at \( \mu = 0.55 \)) for a \( C_e \) value of 0.85. This implies a maximum streamline excavation depths of \( r/4.5 \) to \( r/2.9 \), which is reasonable: usual values range from \( r/5 \) to \( r/4 \) (Melosh, 1989). This paraboloid shell approach represents an approximation of the more accurate method for computing the excavated mass via numerical integration between streamtubes within the Maxwell Z-model itself. Comparison between these two methods, however, has shown that the paraboloid shell approximation yields results that are within 3–4% of the numerical Z-model integration (depending upon \( Z \) value), such that the added accuracy of a full integration was deemed not worth the additional computation time.

In order to estimate the ejecta blanket thickness at some distance \( l \) from the impact site, we utilize a set of simple ballistics equations to compute landing distances for the material ejected at distances \( r \), and \( r_{i+1} \). Since the vast majority of impact ejecta will land within 1–3 crater radii of the impact site, and most craters in the model will be significantly smaller than the radius of curvature of the target body under study, we assume flat-plane geometry to obtain:

\[
v_{eh}(r_i) = v_e(r_i) \cos \left( \frac{\pi}{180} \left( 55 - \frac{r_i}{R_s} \right) \right),
\]

for the horizontal ejection velocity, and

\[
v_{ev}(r_i) = v_e(r_i) \sin \left( \frac{\pi}{180} \left( 55 - \frac{r_i}{R_s} \right) \right),
\]

for the vertical ejection velocity, where ejection angles gradually drop from 
55° to 35°, consistent with the experimental findings of Cintala et al. (1999) (see Richardson et al. (2007) for a full discussion). The above two equations are used to find the ejecta landing distances \( l_i \) and \( l_{i+1} \) for the ejecta launched at distances \( r_i \) and \( r_{i+1} \), respectively, using the simple ballistics equation:

\[
l_i = r_i + \frac{2v_{eh}(r_i) v_e(r_i)}{g},
\]

assuming no horizontal motion once the ejecta have landed. The surface area occupied by the landed ejecta can be found by:

\[
A_i = \pi \left( l_i^2 - l_{i+1}^2 \right),
\]

with the ejecta blanket thickness \( b_i \) at some distance \( l_i \) from the impact site is estimated by:

\[
b_i = \frac{V_i}{A_i}.
\]

For each impact in the model, the above equations are used to produce a variably sized look-up table of ejecta blanket thickness as a function of distance from the impact site, which is automatically scaled to the size of the resulting impact crater and the extent of ejecta coverage on the surface. This look-up table is kept to a fine enough resolution to permit linear interpolations between each

Fig. 4. Examples of the ejecta blanket thickness algorithm used by the CTEM, for a 100 m diameter impactor striking the lunar surface (left) and the surface of Asteroid 433 Eros (right). The lunar plot uses an impact speed of 17.5 km/s (45° incidence) and a surface gravity of 1.63 m/s², while the Eros plot uses an impact speed of 5.0 km/s (45° incidence) and a surface gravity of 0.069 m/s². As in Fig. 3, target material properties are the same in each case, with the primary variable being effective target strength \( T \), where each line indicates: (bold solid) \( T = 0 \) Pa or pure gravity-dominated cratering, (dashed) \( T = 5 \) kPa, (one-dot) \( T = 50 \) kPa, (two-dot) \( T = 500 \) kPa, (three-dot) \( T = 5 \) MPa, and (four-dot) \( T = 50 \) MPa or pure strength-dominated cratering. The importance of target strength in the low-gravity Eros environment is again well demonstrated. Note that far-field ejecta is less affected by target strength, such that ejecta blanket thicknesses as a function of distance from the impact eventually merges with that expected from a purely gravity-dominated crater. For Eros, the sharp cut-off in ejecta at about 17.5 km from the impact site indicates the point where ejecta velocities are greater than the escape velocity of the target body.
2.4. Crater placement and profile

Following the random selection of an impactor size and impact site, and the computation of the resulting crater dimensions and ejecta blanket coverage, the program surveys the area where the crater is to be placed, determining this area's mean elevation and slope. This survey is used to produce a 'reference plane,' upon which the new crater is superimposed. The new crater follows a standard paraboloid shape, with its depth to diameter ratio \( d/D \) designated by a user input parameter. The height of the crater's rim above the surrounding reference plane (as a fraction of crater diameter) is also designated by a user input parameter \( (h_r/D) \). Outside of the crater's rim, the amount of surface uplift outside of the crater's rim falls off as a function of distance from the impact site by a power-law factor of between \( -3 \) and \( -5 \) (user determined), with values of \(-4\) to \(-5\) yielding the most consistent results. Note that this uplifted (or upturned) ground is in addition to the material laid down in the ejecta blanket. Because the user selects the parameters determining the height and fall-off with distance of the upturned crater rim, which adds positive topography to the simulation, these parameters must be selected with care, such that the mean elevation of the cratered terrain either remains constant with time, or better, slowly deflates over time and impacts, as material is lost to space due to ejection at greater than the target's escape velocity. In some later CTEM version, these parameters may be more directly computed to avoid this potential problem.

Complex craters are handled in the simulation by a user parameter which specifies the maximum floor depth to which a crater profile is permitted to extend. For example, in the lunar case, complex craters generally reach maximum depths of roughly 3–5 km below the surrounding terrain (Melosh, 1989), regardless of crater diameter. This flat-floor depth parameter thus generates a highly simplified complex crater shape, without central peaks, peak-rims, or wall terraces, and this should be kept in mind when viewing the simulation synthetic images. Further complexity can be added at a later date, if deemed necessary.

When a new crater is formed, three values are stored at each pixel location occupied by that crater and its rim, in addition to the changes made to the DEM and the model layer which tracks regolith thickness. The DEM fall between look-up table distances from the impact site, for a 100 m impactor striking the lunar surface and the surface of Asteroid 433 Eros, with each shown for a variety of effective surface target strengths \( \Phi \). As also demonstrated in Fig. 3, target strength has a much more severe effect in the low-gravity environment of the asteroid, with progressively thinner ejecta blankets produced as the surface approaches the effective strength of solid rock. In the case of the zero-strength, gravity-dominated lunar impact, the double-precision diameter value is also used as its unique identifier, and this identifier is tagged onto each pixel within a radius of 1.5 final crater radii \( R_f \) of the impact site. Second, the initial crater profile, in elevation above or below its surveyed reference plane, is stored at each pixel location. And third, the current crater profile, in elevation above or below its initial profile, is stored at each pixel location: these profile 'deviation' values begin at zero when the crater is first formed.

Over time and subsequent impacts, whenever a crater pixel-element (pixel-fraction) deviates in elevation by more than 75% from its initial, reference plane profile elevation, that particular pixel-element is considered to have been eradicated and all three values associated with that pixel-element are reset to zero. Fig. 5 shows a typical, simple crater cross-sectional profile, along with the degree of profile deviation permitted in the model, by either further excavation or burial, until that part of the crater is considered to be eradicated. Permitting up to 75% deviation from the crater's initial, profile elevation allows craters to degrade to depth/diameter \( d/D \) ratios as low as 0.05, an approximate minimum relief necessary for visually counted craters under excellent (low-angle) lighting conditions.

2.5. 'Visible' crater counting

Because the craters produced by this model are handled in a per-unit-area (pixel) fashion during crater production and erasure, some rules need to be applied when they are counted for the production of a crater size-frequency distribution curve. Crater counting by human eye requires that at least enough of the rim of the crater remains, such that a radius of curvature can be estimated within the resolution limits of the image being examined. Additionally, crater counters recognize craters according to the amount of experience they have and their personal way of recognizing craters. This introduces a certain amount of subjectivity into the 'art' of manual crater counting.

Therefore, to incorporate a semblance of this feature of crater counting into this model, I use a relatively simple rule specifying that some fraction of the crater's original surface area must remain in order for it to be 'visible' or 'observable.' The remaining crater surface area is determined by summing the number of crater fragments (pixel-elements) which have not yet been eradicated via excavation or regolith coverage (see Section 2.4). This user-specified rule is currently set such that greater than 1/3 of the crater's original surface area must remain in order for it to be considered 'observable.' Note that this counting rule is heavily affected by the pixilation of each crater, especially for craters less than 5 pixels across (as is 'real' visual crater counting also). Craters less than 3 pixels across are never counted as 'observable.' This CTEM thus maintains two forms of crater counts: (1) the total (or actual) number of craters on the surface, down to the smallest, remaining cra-
ter fragment (pixel-element) and (2) the number of observable (countable) craters on the surface, based upon this user-selected observation rule.

2.6. Downslope regolith motion and seismic shaking

Perhaps the most unique feature of this CTEM is the use of a topographic modification model for the degradation and erasure of slopes and impact craters. This modeling is done using an analytical theory of erosion for transport-limited downslope regolith flow first described by Culling (1960), in which we assume that the downslope flow of regolith is controlled by the transportation rate and not by the regolith supply or production rate (weathering-limited flow) – see Nash (1980). This assumption should hold true for most of the surfaces studied, due to the ability of impact 'gardening' to produce a continuous supply of loose regolith around each new crater. In this process, classically modeled by Ross (1968), Soderblom (1970), Shoemaker et al. (1970), Gault et al. (1974) for the lunar surface, small impacts overturn a shallow portion of the regolith layer, while larger impacts overturn the regolith to greater depths. Because of their shallow excavation depths, impacts producing craters up to a few hundred meters in diameter will tend to only 'recycle' the existing regolith layers, with some small loss of material to space after each impact due to ejection at greater than the escape velocity of the target body. 'New' regolith is thus only generated by the largest impacts, which excavate a portion of the fractured bedrock below the regolith layer. It is this constant regolith recycling (with occasional replenishment) that allows us to assume a transport-limited regolith supply.

To develop this erosion theory, we begin with an expression for the conservation of mass on an infinitely small portion of a hillslope, in Cartesian coordinates (Culling, 1960):

$$\frac{\partial z}{\partial t} = -\left[ \frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} \right],$$

(28)

where $z$ is the elevation, and $f_x$ and $f_y$ are the flow rates of regolith in the $x$ and $y$ directions, respectively.

If the regolith layer is isotropic with regard to material flow, and the flow rate is linearly proportional to the slope's gradient then:

$$f_x = -K_d \frac{\partial z}{\partial x} \quad \text{and} \quad f_y = -K_d \frac{\partial z}{\partial y},$$

(29)

where $K_d$ is a downslope diffusion constant, which has units of $\text{m}^3 \text{m}^{-2} \text{s}^{-1}$ (volume flux per unit time) or $\text{m} \text{s}^{-1}$ (downslope motion per unit time).

Substituting Eq. (29) into Eq. (28) gives:

$$\frac{\partial z}{\partial t} = K_d \left[ \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} \right],$$

(30)

which I re-write in diffusion equation form:

$$\frac{\partial z}{\partial t} = K_d \nabla^2 z.$$  

(31)

Note that these equations are traditionally a function of time, applicable when the diffusion constant $K_d$ is some constant rate with respect to time (units of either $\text{m}^3 \text{m}^{-2} \text{s}^{-1}$ or $\text{m} \text{s}^{-1}$). In this case, however, we must define a specific amount of downslope

Fig. 6. Overhead view of a 250-m-diameter crater, depicted at four different times and showing its gradual degradation and erasure by impact-induced seismic shakedown, utilizing the downslope-diffusion theory developed in Section 2.6. Note the rapid initial degradation while slopes are still relatively high, followed by a more gradual degradation as slopes flatten. Compare this Eulerian, finite-difference modeling method to the analytical model developed and shown in Fig. 15 of Richardson et al. (2005).
diffusion per impact $K_i$ (units of m$^3$ m$^{-2}$ per impact, or simply m per impact) due to seismic vibrations which are themselves a function of the impactor’s size, velocity, target surface gravity, and distance from the impact site. Utilizing the numerical ‘shake-table’ test setup described in Richardson et al. (2005), Richardson (2009), a preliminary form for the function which determines the total amount of downslope diffusion experienced due to seismic ‘shakedown’ at some distance $l$ from a particular impact is given by:

$$K_i = C s \frac{v a_i D}{l^d}$$  \( (32) \)

where current (early) values for the constants are: $a = 0.5$, $b = 1.0$, $c = 0.5$, and $d = 0.5$. The seismic shaking conversion factor $C_s$ has been empirically determined to be $\approx 10^{-2}$, also utilizing seismic parameters given for the fractured lunar crust in my numerical shake-table ‘experiments.’ Further development is planned for this portion of the model.

The power-law drop off of $-1/2$ in downslope regolith diffusion (displacement) as a function of distance from the impact site is of particular interest, in that it indicates that the total amount of downslope motion is proportional to the peak velocities experienced by the slope in question, which, in turn, follows the square root of the peak accelerations experienced. Likewise, the peak accelerations experienced by the slope follow the square root of the peak seismic energy densities experienced. This implies a peak seismic energy fall-off with distance of approximately $1/l^2$ (an inverse square-law relationship); that is, most of the seismic energy produced by an impact (point-source) initially remains concentrated within an expanding hemispherical shell within the target material, even when that target material is highly fractured.

The numerical ‘shake-table’ simulations described in Richardson et al. (2005), Richardson (2009) also indicate that significant downslope regolith motion occurs only when seismic accelerations on the target surface exceed $1g$, with that motion occurring in hop-slip fashion down the slope. Between $0.2g$ and $1g$, seismic accelerations are much less efficient at moving regolith downslope, which then occurs in stick-slip fashion. When the seismic accelerations drop below $0.1–0.2g$, downslope regolith displacement fails to occur at all. Since Eq. (32) was empirically determined only for conditions of a seismic acceleration ratio $a_i/g > 1$, we require a method to determine the affected area, or ‘seismic range’ where Eq. (32) can be applied for an individual impact.

At a given point some distance away from the impact site, the seismic vibration-induced acceleration $a_i$ is given by a rearrangement of Eq. (5) of Richardson et al. (2005):

$$a_i = 2\pi f \sqrt{\frac{2\epsilon_i}{\rho \ell}}$$  \( (33) \)

where $f$ is the primary seismic frequency, generally between 10 and 20 Hz, as discussed in Richardson et al. (2005), and $\epsilon_i$ is the seismic energy density within the material comprising the shaking slope. If we make the simplifying assumption that the seismic energy produced by the impact is concentrated within a thin, expanding hemispherical shell (of 1 m thickness), then the peak seismic energy density experienced at some distance $l$ from the impact site can be estimated by:

Fig. 7. Four stages of a lunar-surface CTEM run for an impactor population having a steep, cumulative, power-law slope $-2.5$. In this case, ‘sandblasting’ by small craters dominate, and the run displays classic Gault (1970) behavior: with small craters reaching equilibrium first (at about 5–10% of geometric saturation, and possessing a cumulative power-law slope of roughly $-2$), followed by successively larger craters. In this instance, the position of the ‘knee’ in the curve can be used as an indication of relative surface age. The run has a pixel-scale of 50 m, and depicts an area 100 km by 100 km in size. Crater counts are plotted in relative crater-density fashion (Arvidson et al., 1979). R-plot Legend: (bold solid) visible model crater count (‘observable’ craters), (thin solid) actual model crater count (all crater fragments), (dot-dash) production population, (horizontal dash) geometric saturation, (horizontal dotted) lines are shown at 5%, and 10% of geometric saturation. The (slanted dotted) line indicates where a single crater of the given size would plot on the graph.
\( c_s = \frac{E_i}{2\pi l^2} \) \hspace{1cm} (34)

where \( l \) is the straight-line chord distance between the impact site and the point on the CTEM surface in question, assuming that the model rests upon the surface of an assumed spherical target body of a user-defined radius. The total seismic energy \( E_s \) within this expanding hemispherical shell (including losses due to attenuation) is given by Eq. (14) of Richardson et al. (2005):

\[ E_s = \frac{1}{12}\eta \pi \rho v_i^2 l_i^2 \frac{2.377}{e \alpha \beta} \] \hspace{1cm} (35)

where \( \eta \) is the impact seismic efficiency factor, which ranges from about \( 10^{-4} \) to \( 10^{-6} \) in value (see Section 1.2 of Richardson et al. (2005)), \( K_s \) is the seismic diffusivity (in \( \text{m}^2 \text{s}^{-1} \)), and \( Q \) is the seismic quality factor (seismic attenuation parameter). The seismic diffusivity is defined as (Dainty et al., 1974; Toksoz et al., 1974):

\[ K_s = \frac{1}{3} v_i l_i \] \hspace{1cm} (36)

where \( v_i \) is the seismic P-wave velocity in competent rock, and \( l_i \) is the mean free path for the scattering of seismic waves; that is, the distance over which \( 1/e \) of the seismic energy is scattered.

The seismic velocity \( v_i \) is determined by the rock’s elastic properties, and the mean free path for scattering \( l_i \) is directly proportional to the mean fracture spacing within the target medium. To adopt reasonable assumptions for each, we use values consistent with the upper lunar crust (determined from the lunar seismic experiments): a competent rock seismic velocity of \( v_i = 3 \text{ km s}^{-1} \), and a range of mean free paths for scattering of \( l_i = 0.125 - 2.000 \text{ km} \) (Dainty et al., 1974; Toksoz et al., 1974).

The values of \( Q \) published for the upper lunar crust fall into two ranges: \( Q = 3000 - 5000 \) based upon data from the long-period (LP) instruments (Dainty et al., 1974; Toksoz et al., 1974), and \( Q = 1600 - 2300 \) based upon data from the short-period (SP) instruments (Nakamura, 1976). The difference in these values is related to different propagation distances, propagation depths, and frequency bands used in the two studies (seismic energy leakage underneath the fractured zone also plays a role). I therefore cautiously adopt values of \( Q = 1000 - 2000 \) in this model, given that a \( Q \) of 2000 represents a rough lower limit to the actual \( Q \) of the uppermost lunar crustal layers (Nakamura, 1976).

Following each impact, the area of the CTEM which is affected seismically by the impact is determined (that is, where \( a_i / g > 1 \) via Eq. (33)), and that portion of the model enters an Eulerian phase where two things occur: (1) unstable slopes above the angle of repose (a user-supplied parameter) are permitted to collapse (or diffuse down) to below the angle of stability, and (2) the applicable amount of seismically-induced downslope regolith diffusion is applied. Note that if seismic accelerations are too low to produce downslope regolith motion, the area within 1.5 crater radii is still checked for unstable slopes (and permitted to collapse). Fig. 6 shows an example of how a fresh impact crater is slowly degraded and eventually erased by the process of seismic shaking in this model: compare this to the analytical model shown in Fig. 15 of Richardson et al. (2005). Note that opposite to the process of viscous relaxation (for craters, see the discussion in Melosh (1989)), slope collapse and seismic shaking first erases the high-frequency (short-wavelength) topography, followed in succession by the low-frequency topography, with the lowest-slope portions of the crater bowl to be filled and ‘erased’ last. The explicit form of crater erosion employed in the CTEM also compares quite well with the

**Fig. 8.** Four stages of a lunar-surface CTEM run for an impactor population having a shallow, cumulative, power-law slope \(-1.5\). In this case, ‘cookie-cutting’ by large craters dominate, and the resulting crater population continues to reflect its parent impactor population even after equilibrium conditions have been established, as described in Chapman and McKinnon (1986). This model run illustrates well the downslope regolith motion produced by the collapse of unstable slopes and seismic shaking, on the walls of large craters. The run has a pixel-scale of 50 m, and depicts an area 100 km by 100 km in size. R-plot Legend: (bold solid) ‘visible’ model crater counts, (thin solid) actual model crater counts, (dot-dash) production population, (horizontal dash) geometric saturation, (horizontal dotted) lines are shown at 1%, 5%, and 10% of geometric saturation.
classic mathematical model of small-crater erosion developed by Soderblom (1970): see his Fig. 4 for comparison. Quantitative comparisons between the two models may be performed as part of a future study.

As described in this and the previous sections, the CTEM thus has two methods for covering a pixel area with regolith, and two methods for removing that regolith. In the first instance, regolith can be deposited by either impact ejecta emplacement (ballistic sedimentation), or by the downslope motion of regolith from upslope of the pixel in question. In the second instance regolith can be removed by the process of crater excavation, or by the downslope motion of regolith to regions downslope of the pixel in question. Note that upturned crater rims (uplifted rock beneath the crater’s ejecta blanket) are NOT considered to be ‘regolith’ by the model, although they can act as a source region for regolith, since this rock has been uplifted, broken, and is severely weakened compared to the surrounding country rock. Therefore, material must be transported, either ballistically or via downslope flow, to be considered as ‘regolith’ in this model.

3. Standard example model runs

Prior to attempting to match the CTEM to a crater population on an actual imaged surface, it is illustrative to observe how the CTEM behaves when bombarded with a series of simplified, ‘standard’ impactor populations as a function of time. These standard example model runs are shown in Figs. 7–11, for a steeply-sloped, shallow-sloped, concave-up, and concave-down impactor population, respectively, as depicted on a standard relative-density (R) plot (Arvidson et al., 1979). The simplest, single-sloped populations are considered first, followed by two-slope impactor populations.
3.1. Single-sloped impactor populations

Beginning with the most well-known case-study, Fig. 7 shows the result of bombarding a lunar model surface (50 m/pixel scale) with a steep impactor population of \(-2.5\) cumulative power-law slope. In this example, the model displays classic Gault (1970) behavior, with the smallest craters reaching equilibrium conditions first at between 5% and 10% of geometric saturation (with a cumulative power-law slope of about \(-2\)), and with successively larger crater sizes reaching equilibrium over time. When this type of impactor population is involved, the position of the ‘knee’ in the curve (the inflection point between \(-2\) and \(-2.5\) power-law slopes) can be used as an indicator of relative surface age when compared to other regions exposed to the same impactor population, as Gault (1970) demonstrated and exploited. Note that in this example, crater erasure is dominated by what Woronow (1977a,b, 1978) called ‘sandblasting:’ the erasure of large craters by smaller craters over time, a process which was also analytically modeled by Soderblom (1970). Also note that using the observation rule described in Section 2.5, the ‘observed’ crater equilibrium level is consistent with the 5–10% geometric saturation level described by Gault (1970) and the ‘empirical saturation’ level described by Chapman and McKinnon (1986).

In the complementary case-study, Fig. 8 shows the result of bombarding a lunar model surface (50 m/pixel scale) with a shallow impactor population of \(-1.5\) cumulative power-law slope. In this example, the model reaches what Chapman and McKinnon (1986) described as a ‘quasi-equilibrium’ state, one in which large portions of the surface are frequently ‘reset’ by the formation of large craters, and thus the overall model continues to reflect the original production population, even after a long-period of bombardment. This process is called ‘cookie-cutting’ in Woronow (1977a,b, 1978). Note that although the resulting crater curve is sometimes slightly steeper and sometimes slightly shallower than the production population curve, once it reaches this somewhat ‘jumpy’ equilibrium state, the correspondence between the three...
curves (production population, actual crater counts, and 'visible' crater counts) remains. This particular model run also shows the downslope motion model in action quite well, particular where large craters overlap and unstable slope regions are thus produced. The above two single-slope impactor population model runs demonstrate the two primary modes by which crater density equilibrium is achieved, and they also point to why there has been such a long standing disagreement over the issue. In the case of a stee-

Fig. 13. Four stages of a large-scale, lunar-surface CTEM run using the MAB impactor populations derived by O'Brien and Greenberg (2005) (left) and Bottke et al. (2005) (right), with both compared to the crater-count data presented in Fig. 1 of Strom et al. (2005). This run has a pixel-scale of 3079.37 m, and depicts the entire lunar surface area. A good match (in both runs) first occurs at an MAB exposure age of \(10^{7.7}\) years (second row), and remains until the end of the runs at \(10^{10.8}\) years.
ply-sloped impactor population, ‘sandblasting’ dominates and equilibrium occurs in such a way as to support viewpoint (2) in Section 1. In the case of a shallow-sloped impactor population, ‘cookie-cutting’ dominates and equilibrium occurs in such a way as to support viewpoint (3) in Section 1. However, when the power-law slope of the impactor population changes as a function of impactor size, the results becomes more consistent, as demonstrated in the following sections.

3.2. Two-sloped impactor populations

In the next two model ‘experiments,’ I investigated what happens when a bombarding population’s size-frequency distribution slope changes once as a function of size. Fig. 9 shows the cumulative distribution of impactor sizes used in the following two model runs, one ‘concave-up’ and one ‘concave down,’ and each with an inflection point at an impactor diameter of 100 m.

Fig. 10 shows the result of bombarding a lunar-model surface (100 m/pixel) with a ‘concave up’ impactor population having a cumulative power-law slope of –3 for impactors <100 m diameter, and a slope of –1 for impactors >100 m diameter. Each time the steeply-sloped, small-crater population starts to level off at Chapman’s ‘empirical saturation’ point, a new large crater from the shallow-sloped, large crater population tends to ‘reset’ a significant portion of the model surface, and forces the small craters to begin anew. In this manner, ‘cookie-cutting’ dominates, and the crater counts continue to reflect the shape of the production population, even after equilibrium conditions have been reached.

Similarly, Fig. 11 shows the result of bombarding a lunar-model surface (10 m/pixel) with a ‘concave-down’ impactor population having a cumulative power-law slope of –1 for impactors <100 m diameter, and a slope of –3 for impactors >100 m diameter. In this instance, the shallow-sloped, small craters dominate the crater erasure process, via ‘sandblasting,’ and the crater counts again continue to reflect the shape of the production population, even after equilibrium conditions have been reached. A similar result is shown in Fig. 19 of Chapman and McKinnon (1986), where the ‘peaked’ crater population of Saturn’s Moon Mimas was modeled.

4. Modeling the lunar cratering record

A separate long standing debate in the cratering community has been in regard to the primary source of impactors for the inner Solar System, and the lunar surface in particular: is it either cometary or asteroidal in origin? Recently, Strom et al. (2005) showed that when mapped through the impact cratering scaling-laws (Section 2.2) to form a production population, the modern-day asteroid population yields a very good match to the cratering record on the lunar surface – thus pointing to the asteroid belt as the primary source of impactors in the inner Solar System. In the same year, O’Brien and Greenberg (2005) and Bottke et al. (2005) published the results of collisional evolution models for the Main Asteroid Belt (MAB), showing that while greatly depleted over the time since its formation, the MAB continues to reflect its Late-Heavy Bombardment (LHB) size-frequency distribution; that is, it displays a ‘fossilized’ size-frequency distribution. This is because the shape of the size-frequency distribution curve for a collisionally evolved population (such as the asteroids) is more determined by material properties and impact parameters than by the starting, or original size-frequency distribution of the population (Bottke et al., 2005).

Fig. 12 shows the size-frequency distribution of objects in the Main Asteroid Belt (MAB), as modeled by O’Brien and Greenberg (2005) and Bottke et al. (2005), and placed in terms of collisions per year per square kilometer on a target surface in the middle of the MAB; that is, exposed to a mean MAB impactor flux (Bottke and Greenberg, 1993). Note that impactors >1 km in diameter closely match the actual observed population of MAB objects, while impactors <1 km in diameter represent model extrapolations. These two model impactor populations thus give us a means to test the proposal of Strom et al. (2005) and see if the cratering record recorded in the most heavily-cratered regions of the lunar surface can be duplicated by using the MAB as the impactor source in our CTEM. It should be pointed out that surface ‘ages’ determined in the following model runs are NOT actual surface ages, but should be thought of as ‘MAB exposure ages,’ that is, the surface age obtained when the flux of impactors is constant and equal to that present within the modern-day Main Asteroid Belt. Note also that a direct application of the MAB to the lunar surface does not take into account the dynamical filtering that will occur in moving objects from the MAB to the Near-Earth Object (NEO) environment (Bottke et al., 2005). That is, I am making the simplifying assumption that the magnitude of the Late-Heavy Bombardment (LHB), which hypothetically delivered MAB material directly to the near-Earth environment without filtering (Strom et al., 2005), was such that the vast majority of craters in the heavily-cratered regions of the lunar surface will have come from this LHB “Population 1” source, and thus permit me to use the Main Asteroid Belt directly as the impactor population.

Three lunar surface scales were investigated using these two MAB populations as the impactor source: a large-scale (full lunar surface) CTEM run at 3079 m/pixel, a medium-scale run at 308 m/pixel, and a small-scale run at 30.8 m/pixel. All impacts in the lunar model runs are assumed to be from ‘stony’ objects of density $\rho_i = 2700$ kg m$^{-3}$, occur at an impact speed of $v_i = 17.5$ km/s
(Strom et al., 2005; Gallant et al., 2006) and an impact angle of 45° (Pierazzo and Melosh, 2000). Table 2 lists the model parameters and material properties selected for these CTEM runs, which are described as follows.

Fig. 13 shows the result of applying these two MAB impactor populations in a large-scale lunar simulation, which includes the full surface area of the Moon (see Table 2). Note that endogenic crater erasure processes are not included in the model, such that

![Fig. 13](image)

**Fig. 14.** Four stages of a medium-scale, lunar-surface CTEM run using the MAB impactor populations derived by O’Brien and Greenberg (2005) (left) and Bottke et al. (2005) (right), as compared to the crater-count data presented in Fig. 2 of Hartmann (1995). This run has a pixel-scale of 307.9 m, and depicts 1/100 of the lunar surface area. An acceptable match (in both runs) first occurs at an MAB exposure age of $10^{9.5}$ years (second row), and remains until the end of the runs at $10^{10}$ years.
the entire surface retains its cratering record, without Mare forma-
tion, basin filling by lava, or other processes not described previ-
ously for the CTEM. Using either asteroid impactor population,
the model-produced cratering record first shows a good match to
the observed lunar cratering record – using data from Fig. 1 of
Strom et al. (2005) – at an MAB exposure age of 5 Gyr, and retains
that good fit thereafter, out to a 40 Gyr MAB exposure age at the
end of the runs.

Fig. 14 shows the result of applying these two MAB impactor
populations in a medium-scale lunar Highlands simulation, which
includes 1/100 the surface area of the Moon (see Table 2). Here, the
model-produced cratering records first show an acceptable match
to the observed lunar cratering record – using data from Fig. 2 of
Hartmann (1995) – at an MAB exposure age of 2.5 Gyr, and retains
that good fit thereafter, out to a 16 Gyr MAB exposure age at the
end of the runs. In general, the smaller the scale of the CTEM
run, the shorter the amount of time it takes for crater density equi-
librium to be reached using either MAB impactor population.
Nonetheless, the long, 2.5–5 Gyr time periods of direct, modern-
day MAB exposure needed to match the existing lunar cratering re-
cord are indicative of there having been a much higher impactor
flux in the inner Solar System at some time in the past. That is,
the current NEO population (lunar environment impactor flux)
cannot have produced the observed lunar cratering record within
the lifetime of the Solar System.

The excellent fits between the previous four CTEM runs and the
observed lunar cratering record verify the finding by Strom et al.
(2005) that the primary source of impactors for the lunar surface
has been the Main Asteroid Belt (MAB). Additionally, given that
most large-scale lunar impacts were produced during early periods
of heavy bombardment (Strom et al., 2005), these runs also support
the finding by Bottke et al. (2005) that the modern-day MAB rep-
resents a ‘fossilized’ size-frequency distribution dating back to
the Late-Heavy Bombardment (LHB). Note that these four runs also
support viewpoint (3) of Section 1, in that once a surface has
reached crater density equilibrium, its age can no longer be defi-
nitely determined, but instead represents only a lower limit. Thus,
crater density equilibrium techniques, which assume viewpoint
(1) of Section 1, need to be revised to take the attainment
of crater density equilibrium conditions into account.

Shifting downward in size still further, Fig. 15 shows the result of
applying the O’Brien and Greenberg (2005) MAB impactor pop-
ulations in a small-scale lunar Maria simulation, which includes
1/10,000 the surface area of the Moon (see Table 2). This particular
model run is interesting in that it centers on the small-crater pop-
ulation most often attributed to secondary cratering, where an-
other long-standing debate has existed in regard to what fraction
of the craters in this size range are primaries and what fraction
are secondaries. Gault (1970) concluded that secondaries, while
present, represented only a minority of craters observed. Hartmann
(1988, 1995), Hartmann and Gaskell (1997) concluded that
secondary craters make up a majority of the craters seen on this
scale, while Strom et al. (2005) does not include small craters in
their plots at all, due to possibility of secondary crater ‘contamina-
tion.’ For this model run, I utilized the more steeply-sloped O’Brien
and Greenberg (2005) asteroid population, and as this run shows,
the small-crater record can be crudely, but not well duplicated
using only primary impactors: indicating the necessity for a signif-
icant contribution by secondary impacts on this size-scale. In this
case, the model-produced cratering record shows a very rough
match to the observed lunar cratering record – using data from
Figs. 1 and 2 of Hartmann and Gaskell (1997) – at an MAB exposure
age of 1.3 Gyr, but thereafter deviates from the observed crater
population as crater density equilibrium is approached, and thus

Fig. 15. Four stages of a small-scale, lunar-surface CTEM run using the MAB impactor populations derived by O’Brien and Greenberg (2005), as compared to the crater-count
data presented in Figs. 1 and 2 of Hartmann and Gaskell (1997). This run has a pixel-scale of 30.8 m, and depicts 1/10,000 of the lunar surface area. The ‘elbow’ of the
production population merges with the Hartmann and Gaskell (1997) counts at an MAB exposure age of 10^{11} years (upper right), with Hartmann’s crater counts displaying a
slightly steeper power-law distribution (to the left of the ‘elbow’) than that produced by the MAB impactor population, likely due to the contribution of secondary craters in
Hartmann’s counts. Also compare this plot to the lunar Maria and lunar Post Orientale curves depicted in Fig. 1.
indicates that in the case of the Mare Cognitum and Mare Tranquil-litatis areas under study, these surfaces are young enough not to have reached equilibrium conditions yet.

5. Conclusion

Recent advances in computing technology, and our understanding of the processes involved in crater production, ejecta production, and crater erasure have permitted me to develop a highly-detailed Cratered Terrain Evolution Model (CTEM), which can be used to investigate a variety of questions in the study of impact bombarded landscapes. In this study, I have looked in particular at the manner in which crater densities attain equilibrium (commonly called crater saturation conditions). The model runs described in Sections 3 and 4 have demonstrated the following general points:

- Crater density equilibrium generally occurs near observed crater densities of about 0.1–0.3 on a relative-density ($R$) plot (Arvidson et al., 1979), or about 2–10% geometric saturation (Gault, 1970). However, this may vary significantly, depending upon the parent impactor population’s size distribution.
- If the impactor/production population has a cumulative power-law slope of $< 2$, then small-crater ‘sandblasting’ dominates the crater erasure process and crater density equilibrium values tend to follow classic Gault (1970) behavior – with crater densities leveling off at roughly 5–10% geometric saturation, with a cumulative power-law slope of about $< 2$.
- If the impactor/production population has a cumulative power-law slope of $> 2$, then large-crater ‘cookie-cutting’ dominates the crater erasure process and crater density equilibrium values will continue to reflect, or follow the shape of the production population, as large regions are continuously ‘reset’ and then repopulated with small craters again.
- If the impactor/production population has a variable cumulative power-law slope, then a mixture of ‘sandblasting’ and ‘cookie-cutting’ will occur, depending upon the size range of craters involved, and crater density equilibrium values will continue to reflect, or follow the shape of the production population (viewpoint (3) in Section 1). This behavior thus allows the shape of an impactor population to be determined for a particular surface, even after that surface has reached crater density equilibrium.
- The ages determined, using crater density chronology techniques, for surfaces which have likely reached a crater density equilibrium condition represent only a lower limit to the age of the surface studied.

More specific to the lunar cratering record, I have found that:

- The heavily-cratered regions of the lunar surface (such as the lunar Highlands) represent a crater population which is in crater density equilibrium, but which also continues to reflect, or follow the impactor/production population which produced it.
- The shape of the crater density curve for heavily-cratered regions of the lunar surface (on all scales) is consistent with a Main Belt asteroid impactor population. That is, the size-frequency distribution of the impactor population which best reproduces the lunar cratering record using the CTEM is nearly identical to that of the current Main Asteroid Belt, as stated by Strom et al. (2005).
- The extreme age for the heavily-cratered regions of the lunar surface, dating back to the heavy-bombardment period of lunar history, along with the finding that this cratering record can be matched using the modern-day MAB, is consistent with the finding of Bottke et al. (2005): that the modern-day MAB represents what they term a ‘fossilized’ size-frequency distribution. That is, although the MAB is severely depleted in mass as compared to its ancient counterpart, it continues to maintain the shape of its Late-Heavy Bombardment (LHB) size-frequency distribution.

In addition to the above three points, preliminary modeling of the small-scale lunar surface (see Fig. 15) suggests that NEO primary impactors (originally from the MAB) dominate even the small-crater population, with secondary impactors (ejecta from very large craters) in the minority. However, because both the small-scale NEO population (Bottke et al., 2005) and the secondary impactor population (Hartmann, 1995; Hartmann and Gaskell, 1997) exhibit steep power-law distributions (because both are produced by impacts) disentangling these two sources may be more problematic than this first treatment seems to indicate. Further investigation will be necessary before more definite conclusions can be reached.

In general, this new Cratered Terrain Evolution Model has shown itself to be a very effective tool for investigating the evolution of an airless, cratered landscape over time, when impacts are the primary geologic process in play. It has been particularly gratifying to finally determine an answer to the old question of when, and how crater density equilibrium conditions occur.

Acknowledgments

The author wishes to thank Clark Chapman of the Southwest Research Institute (SWRI) for his review and insightful commentary on the history of craters studies in the planetary science community, as well as Robert Strom (University of Arizona) for his kind review comments. This research was funded by NASA Grant No. NNX07AG04G.

References


