Geomorphically based predictive mapping of soil thickness in upland watersheds

Jon D. Pelletier¹ and Craig Rasmussen²

Received 29 July 2008; revised 17 June 2009; accepted 26 June 2009; published 26 September 2009.

The hydrologic response of upland watersheds is strongly controlled by soil (regolith) thickness. Despite the need to quantify soil thickness for input into hydrologic models, there is currently no widely used, geomorphically based method for doing so. In this paper we describe and illustrate a new method for predictive mapping of soil thicknesses using high-resolution topographic data, numerical modeling, and field-based calibration. The model framework works directly with input digital elevation model data to predict soil thicknesses assuming a long-term balance between soil production and erosion. Erosion rates in the model are quantified using one of three geomorphically based sediment transport models: nonlinear slope-dependent transport, nonlinear area- and slope-dependent transport, and nonlinear depth- and slope-dependent transport. The model balances soil production and erosion locally to predict a family of solutions corresponding to a range of values of two unconstrained model parameters. A small number of field-based soil thickness measurements can then be used to calibrate the local value of those unconstrained parameters, thereby constraining which solution is applicable at a particular study site. As an illustration, the model is used to predictively map soil thicknesses in two small, ~0.1 km², drainage basins in the Marshall Gulch watershed, a semiarid drainage basin in the Santa Catalina Mountains of Pima County, Arizona. Field observations and calibration data indicate that the nonlinear depth- and slope-dependent sediment transport model is the most appropriate transport model for this site. The resulting framework provides a generally applicable, geomorphically based tool for predictive mapping of soil thickness using high-resolution topographic data sets.


1. Introduction

[2] Soil thickness exerts a first-order control on the hydrologic response of upland watersheds. Relatively thin soils are more prone to saturated overland flows compared to thicker soils which have greater water storage potential. Sensitivity studies of numerical models illustrate the importance of soil thickness in controlling infiltration rates [Woolhiser et al., 2006]. Field studies show a significant inverse correlation between water residence time and terrain steepness [McGuire et al., 2005]. This inverse correlation likely reflects the increased hydrologic gradient on steeper slopes, but also the effect of generally thinner soils on steeper terrain. Hydrologic models require spatially distributed input data for soil thickness when applied to upland (i.e., soil over bedrock) landscapes. However, there is currently no widely used method for estimating soil thicknesses using readily available data. As a result, hydrologic modelers often extrapolate from simple empirical relationships between soil thickness and terrain parameters based on a limited number of field measurements [e.g., Moore et al., 1993]. In the hydrologic model developed for infiltration at Yucca Mountain, for example, soil thicknesses were assumed to be inversely proportional to the slope gradient [Civilian Radioactive Waste Management System Management and Operating Contractor, 2000]. Process-based geomorphic modeling of soil thickness [e.g., Dietrich et al., 1995], however, in addition to measured soil thicknesses in several well-documented drainage basins [Heimsath et al., 1997, 1999], indicates that soil thicknesses are most closely associated with hillslope curvature. More complex relationships between soil thicknesses and various combinations of slope, curvature, and contributing area are possible and will be discussed in this paper. However, the discrepancy between the slope-dependent assumption of the Yucca Mountain model and the curvature-dependent results of geomorphically based studies underlines the need for a better process-based understanding of how soil thickness is controlled by topography, with the ultimate goal being to construct a broadly applicable, geomorphically based method for estimating soil thicknesses that honors the complexity of geomorphic processes and the multiscale heterogeneity of real landscapes.

thickness variations in upland watersheds. Predictive mapping of soil thicknesses was the explicit goal of the Dietrich et al. [1995] study, while Roering [2008] modeled soil thicknesses within the context of a broader study aimed at understanding the feedbacks between soil thicknesses and hillslope form. As such, the approach of Roering [2008] was not explicitly proposed as a method for mapping soil thickness but his approach does predict soil thickness using modern high-resolution topographic data as input and hence should be considered as a possible geomorphically based method for this purpose. Both models calculate soil thickness as the difference between soil production and erosion using numerical landscape evolution models that begin with the modern topography (e.g., lidar-derived digital elevation models (DEMs)) and an assumed initial soil thickness. These models then simulate landscape development forward in time subject to a prescribed tectonic uplift rate. Dietrich et al. [1995], evolved the landscape forward in time for 15 ka, based on the assumption that initial soil development at that study site began in early Holocene time. Roering [2008] simulated the landscape forward in time for 0.5 Ma until a topographic steady state condition was achieved. Dietrich et al. [1995] assumed soil transport was to be proportional to the local hillslope gradient. Research in the past decade, however, has emphasized the importance of nonlinear slope-dependent transport on steep slopes [e.g., Roering et al., 1999] as well as depth-dependent transport, i.e., thick soils on hillslopes dominated by creep and bioturbation often have higher rates of sediment flux compared to thin soils [e.g., Gabet, 2000; Heimsath et al., 2005]. As such, Roering [2008] expanded the Dietrich et al. [1995] approach to include a nonlinear depth- and slope-dependent transport model. The resulting maps of soil thickness generated by these models provide a process-based prediction of how soil thicknesses vary across the landscape. Both methods have limitations, however. Dietrich et al. [1995] used the modern lidar data as the initial condition for the model at 15 ka, together with an assumed initial soil thickness at that time. The initial “age” of upland landscapes is generally not well constrained, however. Roering [2008] assumed topographic steady state with a prescribed uplift rate. This assumption is valid for some locations but does not hold generally. In this paper we describe a method that calculates soil thicknesses directly from the modern topographic data without iteration through time and without requiring a topographic steady state condition. The model contains only two “free” parameters that require site-specific calibration.

2. Background

(a) Schematic diagram of a hillslope profile from divide to channel head. (b) Models for the relationship between soil production rate and soil thickness, illustrating the exponential model of Heimsath et al. [1997] and an alternative “humped” model based upon a particular form of the function proposed by Furbish and Fagherazzi [2001].

\[
\frac{\partial h}{\partial t} = \frac{\rho_s}{\rho_w \cos \theta} P_0 e^{-h/h_0} + D \nabla^2 z
\]
\( \frac{\partial z}{\partial t} = D \nabla^2 z + U \)  

(2)

\[ z = b + h \]

(3)

where \( t \) is time, \( \rho_b \) is the bedrock density, \( \rho_s \) is the sediment density, \( P_0 \) is the maximum bedrock lowering rate on a flat surface, \( \theta \) is the slope angle, \( U \) is the rock uplift rate, \( D \) is the hillslope diffusivity, and \( h_0 \) is a characteristic soil depth [Heimsath et al., 1997, 2001]. Equations (1)–(3) state that the rate of change of soil thickness with time is the difference between a “source” term proportional to the rate of soil production associated with the bedrock surface lowering and a “sink” term equal to the curvature of the topographic profile. The \( \cos \theta \) dependence in (2) originates from the fact that soil production is an exponential function of soil thickness normal to the surface [Heimsath et al., 2001]. The curvature-based erosion model in (3) is the classic diffusion model of hillslope evolution, first proposed by Culling [1960, 1963]. Equations (1)–(3) can be solved for the steady state case in which soil thickness is independent of time:

\[ h = \frac{h_0}{\cos \theta} \ln \left( \frac{\rho_b}{\rho_s} \frac{P_0}{D \cos \theta - \nabla^2 z} \right) \]

(4)

Note that steady state in this context does not mean that the topography is in steady state, but, rather, that the soil thickness does not change through time as the landscape is denuded (i.e., a “soil thickness steady state” condition). Equation (4) suggests that given a very accurate map of topography, soil thicknesses can be estimated if the values of \( h_0 \) and \( \rho_b \) \( \rho_s \) \( D \) are known. Alternatively, calibration data for observed soil thicknesses can be used to infer the local value of \( \rho_b \) \( \rho_s \) \( D \) if \( h_0 \) is known. The analysis of in situ cosmogenic isotopes indicates that the value of \( h_0 \) (the soil thickness at which bedrock lowering falls to \( 1/e \) of its maximum value) is approximately 0.5 m for several well-studied sites around the world [e.g., Heimsath et al., 1997, 1999, 2001].

[5] Evidence suggests that the diffusion model of hillslope evolution (equation (3)) has very limited application to most upland hillslopes [e.g., Roering et al., 1999; Gabet, 2000; Heimsath et al., 2005]. In steep landscapes, sediment flux increases nonlinearly with slope gradient as the angle of stability is approached. Steep, planar hillslopes and abrupt, knife edge drainage divides are a signature of landslide dominated, nonlinear transport on hillslopes [Roering et al., 1999, 2001; Roering, 2004]. In addition to the nonlinear slope dependence of hillslope transport processes, there are also processes that require an area and soil depth dependence as well. Braun et al. [2001] proposed three end-member models of sediment transport in their study of the coevolution of process and form in evolving hillslopes: slope-dependent, contributing area- and slope-dependent, and depth- and slope-dependent transport. In this paper, we include the same three end-member models, updated to include the nonlinear slope dependence documented by Roering et al. [1999]. We consider the nonlinear slope-dependent (NSD) model given by

\[ \frac{\partial z}{\partial t} = D \nabla \cdot \left( \frac{h_n \nabla z}{1 - (|\nabla z|/S_c)^2} \right) + U \]

(5)

and the nonlinear area- and slope-dependent (NASD) model given by

\[ \frac{\partial z}{\partial t} = D \nabla \cdot \left( \frac{A^m \nabla z}{1 - (|\nabla z|/S_c)^2} \right) + U \]

(6)

where \( A \) is the contributing area (a proxy for runoff), \( m \) is an empirical exponent, and \( S_c \) is the tangent of the angle of stability. Contributing area on hillslopes is calculated using a multiple-flow direction algorithm [Freeman, 1991]. Equation (5) is generally used for erosion dominated by processes occurring on the surface that have no dependence on area, such as rain splash. The NSD model can have broader applicability, however, (i.e., at sites dominated by creep, bioturbation, and mass movements) provided that the soils are uniformly thicker than the thickness of the layer undergoing transport. In such cases, the NSD model can be applicable because the system does not operate within the range of soil thicknesses where depth dependence is significant. Area-dependent transport models such as (6) are generally used for erosion dominated by sheet flow and/or rilling. Equation (6) combines area-dependent flux with nonlinear slope-dependent transport. Although area-dependent transport and nonlinear slope-dependent transport have not been combined in the form of (6) previously, a nonlinear increase in sediment flux with slope is just as likely to occur in a steep hillslope dominated by rilling, overland flow, and mass movements as in a steep hillslope dominated by area-independent processes such as creep. As such, the combination of the two models in form of (6) is appropriate. Historically, the NSD model has been the most commonly used transport model in hillslope studies, but its application to hillslopes dominated by freeze-thaw creep and bioturbation has been questioned by a number of recent studies [Gabet, 2000; Gabet et al., 2003; Heimsath et al., 2005; Yoo et al., 2005]. These studies have clearly shown that depth-dependent transport models are more appropriate in areas dominated by those processes. The nonlinear depth- and slope-dependent (NDSD) model combines the depth dependence advocated by these studies with the nonlinear slope dependence of Roering et al. [1999, 2001] and Roering [2004] to give

\[ \frac{\partial z}{\partial t} = D \nabla \cdot \left( \frac{h_n \nabla z}{1 - (|\nabla z|/S_c)^2} \right) + U \]

(7)

where \( h_n \) is the soil thickness normal to the surface. Equation (7) is not appropriate for very thick soils (i.e., more than several meters) because the sediment flux cannot continue to increase indefinitely as soil thickness increases (unless the hillslope is near the angle of stability and hence capable of transporting the entire soil profile by mass movement). More complex depth-dependent models that include a depth-saturation effect [e.g., Roering, 2008] are
available that can be used in cases where the soil is very thick. As we will show, each of these transport models (equations (5)–(7)) predicts distinctly different relationships between soil thickness and landscape position for a given hillslope form. As such, calibration data can be used to constrain which transport relationship is most applicable for a particular study site if the topography is well constrained.

Equation (2) states that bedrock lowering is a maximum for bare bedrock slopes and decreases exponentially with increasing soil thickness. This exponential relationship has been inferred from cosmogenic isotope analyses on hillslopes [Heimsath et al., 1997, 1999]. Conceptually, this relationship represents the buffering effect that soil has on an underlying bedrock, protecting it from diurnal temperature changes and the infiltrating runoff that drive physical and chemical weathering. The exponential soil production function may not capture the full complexity of soil production, however. As soil thickness decreases below a critical value in arid and semiarid regions, the landscape may be unable to store enough water to promote weathering or support plant life. Plants act as weathering agents (e.g., root growth can fracture rock, canopy cover can decrease evaporation, etc.). As such, in some arid and semiarid environments weathering rates may increase with increasing soil thickness for thin soils, a behavior inconsistent with the exponential model. A humped or bell-shaped relationship of soil production to soil thickness (Figure 1b) has been theorized for decades [Ahnert, 1977; Cox, 1980; Dietrich et al., 1995; Anderson and Humphrey, 1989; Furbish and Fagherazzi, 2001; Minasny and McBratney, 1999, 2006]. Recent cosmogenic radionuclide data from granitic landscapes in Australia provide support for a humped production model [Heimsath, 2006]. Using a humped production model in landscape evolution models reproduces landscapes with a bimodal distribution of slopes similar to many arid region hillslopes [Anderson and Humphrey, 1989; Strudley et al., 2006]. As such, a humped production function is most likely to be valid in study sites with a bimodal distribution of slopes and a high density of exposed bedrock with limited vegetation cover. Here we use a humped production function given by the form

$$\frac{\partial h}{\partial t} = \frac{\rho_h}{\rho_s \cos \theta} \frac{P_0}{\exp \left( \frac{h - h_0}{T} \right) + D \nabla^2 z}$$

and plotted in Figure 1b for the simple case of a flat slope with $U = 0$. The method we develop does not depend on this specific form of the humped production function, however, and other forms may be used. Whether the soil production function is exponential or humped most likely depends on the study area. We will consider both possibilities in this paper.

3. Model Description

3.1. Two-Dimensional Modeling of Soil Thickness on Synthetic Hillslopes

To gain a qualitative understanding of how soil thickness varies as a function of landscape position and to understand the challenges involved in predicting soil thickness in 3-D landscapes, we first consider a 2-D version of the model applied to synthetic hillslope profiles. For simplicity, we also consider the exponential soil production function only in this section. The simplest relationship between topography and soil thickness is obtained by assuming topographic steady state and spatially uniform tectonic uplift. In this case, bedrock weathering must keep pace with a constant, uniform rate of uplift. This implies [Roering, 2008]

$$P_0 \sqrt{1 + \frac{\partial z}{\partial x}^2} \exp \left[ \frac{-h}{h_0 \sqrt{1 + \frac{\partial z}{\partial x}^2}} \right] = U$$

In (9) we used $\cos \theta = 1/\sqrt{1 + \frac{\partial z}{\partial x}^2}$ in order to express all of the topographic controls in terms of $z$ and its derivatives. The solution to (9) is [Roering, 2008]

$$h = \frac{h_0}{\sqrt{1 + \frac{\partial z}{\partial x}^2}} \ln \left( \frac{P_0}{U} \sqrt{1 + \frac{\partial z}{\partial x}^2} \right)$$

Equation (10) states that in topographic steady state, soil thicknesses are a function of slope only (not curvature) and are independent of the sediment transport process acting on the slope. Equation (10) can also be applied in three dimensions by replacing $\frac{\partial z}{\partial x}$ with $|\nabla z|$. In order to apply (10) to a particular landscape, a high-resolution DEM must be first be used to compute the slope gradient terms. Second, a range of values of the ratio $P_0/U$ must be input into (10), resulting in a family of solutions corresponding to each value of $P_0/U$. Finally, field-based calibration data for soil thickness can then be used to constrain which value of $P_0/U$ is most appropriate for a given study area. Equation (10) provides a useful starting point for modeling soil thicknesses, but the assumption of topographic steady state with spatially uniform uplift is not generally applicable.

The soil thickness steady state condition used in this paper assumes a balance between soil production and erosion at every point on the landscape. No constraint is placed on erosion rates (spatially or temporally) under this condition. Soil profiles operating under the exponential soil production function model tend to evolve toward steady state soil thickness because of the inverse relationship between soil production rate and soil thickness. A soil profile governed by this inverse relationship will tend to evolve back toward a steady state soil thickness condition if perturbed away from that state because a perturbation that causes local soil thinning will trigger an increase in soil production rates, driving the system back toward the original soil thickness in order to maintain a balance between production and erosion. If the NDSD model is applicable at a given site, this negative feedback mechanism is even stronger because a perturbation that thins the soil locally will simultaneously increase production and decrease erosion, thus promoting a return to thicker soils. Because of this negative feedback relationship, the soil thickness steady state condition is the state toward which many hillslope systems naturally evolve. The existence of this negative feedback does not guarantee that an equilibrium condition exists (e.g., the soil thickness could oscillate around, but never reach, an equilibrium value), but it makes an equilibrium condition more likely to exist. In the case of a hillslope dominated by the humped production
function, greater landscape instability can be expected [Furbish and Fagherazzi, 2001], and the conditions under which a soil thickness steady state condition occur are not well constrained. It is difficult to prove the existence of a soil thickness steady state condition precisely at a given study site. Despite this difficulty, it is common in the geomorphic literature to assume that this condition is met. For example, studies that measure erosion rates using the abundance of cosmogenic radionuclides produced in situ in bedrock assume a soil thickness steady state condition [Heimsath et al., 1997, 1999, 2001]. These studies argue that systematic deviations from soil thickness steady state would yield trends in their data that have not been observed [e.g., Heimsath et al., 2001].

[5] First we consider the NSD and NASD models ((5) and (6)) within the context of the soil thickness steady state condition. The equation for the rate of change of soil thickness for these models is given by

$$\frac{\partial h}{\partial t} = \frac{\rho_h}{\rho_s} P_0 \sqrt{1 + \frac{\partial z}{\partial x}^2} \exp \left[ - \frac{h}{h_0 \sqrt{1 + |\partial z/\partial x|^2}} \right] + \frac{\partial z}{\partial t} - U$$

(11)

where \( \partial z/\partial t \) is given by

$$\frac{\partial z}{\partial t} = D_1 \frac{\partial}{\partial x} \left( \frac{\partial z/\partial x}{\sqrt{\left( 1 - \frac{\partial z/\partial x}{S_c} \right)^2}} \right) + U$$

(12)

for the NSD model and

$$\frac{\partial z}{\partial t} = D_2 \frac{\partial}{\partial x} \left( \frac{\partial^m z/\partial x}{\sqrt{\left( 1 - \frac{\partial z/\partial x}{S_c} \right)^2}} \right) + U$$

(13)

for the NASD model. Equation (13) assumes that the distance from the divide, \( x \), scales with the square root of drainage area, \( A \), in order to reduce the 3-D model to two dimensions. Generally speaking, scaling the distance from the divide to drainage area introduces an additional parameter related to the basin shape, but here we assume that this parameter is subsumed within the transport coefficient \( D_2 \). The analytic solution to (11), assuming steady state soil thickness, is given by

$$h = h_0 \sqrt{1 + \frac{\partial z}{\partial x}^2} \ln \left( \frac{\rho_h}{\rho_s} \frac{P_0}{\partial z/\partial t} \sqrt{1 + \frac{\partial z}{\partial x}^2} \right) \quad (14)$$

Equations (12)–(14) illustrate that soil thickness, even in the simplest and most idealized transport models, is a complex combination of slope, curvature, and, in cases of overland flow or rilling, landscape position. As such, empirical models that predict soil thickness as a function of slope gradients only are unlikely to capture the full complexity of soil thickness variations (unless topographic steady state applies).

[10] The equation for the NDSD model is given by

$$\frac{\partial h}{\partial t} = \frac{\rho_h}{\rho_s} P_0 \sqrt{1 + \frac{\partial z}{\partial x}^2} \exp \left[ - \frac{h}{h_0 \sqrt{1 + |\partial z/\partial x|^2}} \right] + D_3 \frac{\partial}{\partial x} \left( \frac{h \frac{\partial z}{\partial x}}{1 - \left( \frac{\partial z}{\partial x} / S_c \right)^2} \right)^2$$

(15)

Equation (15) is more challenging to solve than (14) because \( h \) cannot be isolated algebraically. Equation (15) can be solved numerically, however, by representing the hillslope as a discrete set of points. In this approach, the first step is to solve for the soil thickness at the divide, \( h_d \), by balancing soil production and erosion at that point and utilizing the fact that the slope gradient is equal to zero at divides. This gives

$$\frac{\rho_h}{\rho_s} P_0 e^{-kh_d/h_0} = D_3 h \frac{\partial z}{\partial x} \bigg|_{x=0}$$

(16)

Equation (16) is a transcendental equation that can be solved numerically using a root-finding technique. Once the soil thickness at the divide is known, the soil thickness at the next downslope point, \( h_1 \), can be computed in a similar way using a discretized version of (16):

$$\frac{\rho_h}{\rho_s} P_0 \sqrt{1 + \frac{\partial z}{\partial x}^2} \exp \left[ - \frac{h_1}{h_0 \sqrt{1 + |\partial z/\partial x|^2}} \right] = D_3 \left( \frac{h_1 \frac{\partial z}{\partial x}}{1 - \left( \frac{\partial z}{\partial x} / S_c \right)^2} \right)^2$$

(17)

where

$$\frac{\partial z}{\partial x} = \frac{z_{i+1} - z_i}{\Delta x}$$

(18)

The simplest root-finding technique is a brute force method in which all possible values of \( h_d \) and \( h_1 \) are input into (17) and (18), respectively, in order to determine the values of \( h_d \) and \( h_1 \) that minimize the difference between the left and right sides of those equations. In this paper, the values of soil thickness at each point are assumed to be between zero and some maximum prescribed depth in increments of 1 cm. This approach requires evaluating (17) and (18) several hundred times for each point on a landscape in order to find the values of \( h \) that balance production and erosion at every point. Points farther downslope from \( h_1 \) can be computed using a more general discretized version of (17), i.e.

$$\frac{\rho_h}{\rho_s} P_0 \sqrt{1 + \frac{\partial z}{\partial x}^2} \exp \left[ - \frac{h_1}{h_0 \sqrt{1 + |\partial z/\partial x|^2}} \right] = D_3 \left( \frac{h_{i-1} \frac{\partial z}{\partial x}}{1 - \left( \frac{\partial z}{\partial x} / S_c \right)^2} \right)^2$$

(19)
Figure 2. Plots of soil thickness on an idealized 2-D hillslope, using (a–c) nonlinear slope-dependent transport, (d–f) nonlinear area- and slope-dependent transport, and (g–i) nonlinear depth- and slope-dependent transport, for three different values of the angle of stability (30° (Figures 2a, 2d, and 2g), 33° (Figures 2b, 2e, and 2h), and 36° (Figures 2c, 2f, and 2i)), and a range of values of the model parameter $\rho_b P_0/\rho_s D$.

Figure 2 illustrates the model predictions for soil thickness for the three end-member transport models applied to a synthetic hillslope profile and assuming a soil thickness steady state condition. Topographic profiles in steep terrain most often have their greatest (downward) convexity near the divide and become more planar with increasing distance downslope. In order to construct a synthetic slope that honors this field observation, we derived our synthetic hillslope based on the assumption that the hillslope curvature is inversely proportional to the distance from the divide: $d^2z/dx^2 \propto 1/x$. Integrating this expression twice gives the following mathematical form for the synthetic hillslope profile:

$$z(x) = z_0 - ax(\ln x - 1)$$  \hspace{1cm} (20)

where $a$ is a model parameter that controls the overall hillslope relief for a given slope length. An alternative approach is to model the topography and soil thickness forward in time until both topographic and soil thickness steady state conditions are achieved for a prescribed uplift rate, as in the work by Roering [2008]. Here I chose to use a prescribed topographic $Pr$ however, in order to make the paper more internally consistent. The method of this paper assumes that the topography is known and that the soil thickness can be determined without having to evolve the topography forward in time to a hypothetical topographic steady state condition. By prescribing the topographic profile in (20), the approach in this section is consistent with the model framework adopted in the remainder of the paper.

Equation (20) is plotted in Figure 2 with $z_0 = 25$ m and $a = 0.15$. Figure 2 illustrates that for each of the three sediment transport models, soil thicknesses increase systematically with increasing values of $\rho_b P_0/\rho_s D$. In the 2-D and 3-D versions of the model of this paper, soil thicknesses depend only on the ratio $\rho_b P_0/\rho_s D$ rather than on the absolute values of each parameter. It is for this reason that the model predicts soil thicknesses with only two “free” parameters, i.e., the ratio $\rho_b P_0/\rho_s D$ and the tangent of the angle of stability, $S_c$. The model also depends on the value of $h_0$, but, as noted in the introduction, cosmogenic radiocarbon studies that infer the value of $h_0$ indicate that this value varies little from 0.5 m everywhere it has been measured. As such, it is appropriate to consider $h_0$ to be a constant for the purposes of this paper.
[12] The soil thickness profiles predicted by the NASD model are the most distinctive of the three models. This transport model (shown here for \( m = 1/2 \)) predicts soils that are thickest at the divide, in contrast to the other two models where soil thicknesses are relatively thin on divides and reach a maximum value part way down the slope. A comparison of the soil thickness profiles for the NSD and NDSD models reveals that soil thicknesses are more spatially uniform in the depth-dependent model. In the extreme case of the depth-dependent model with slope gradients that are substantially below the stability threshold (i.e., Figure 2i), the predicted soil thicknesses are almost completely uniform except for the region of high curvature close to the divide. Within each transport model, the effects of an increase in the angle of stability are illustrated Figure 2. As the angle of stability increases from 30° (Figures 2a, 2d, and 2g) to 36° (Figures 2c, 2f, and 2i), the effects of nonlinearity become less pronounced, resulting in thicker soils at the base of the profile compared to models with greater nonlinearity. On the basis of this result, steeper slopes in a given watershed can be expected to have greater spatial variability in soil thicknesses between the top and bottom of the slope compared to gentler slopes where nonlinear transport is less of a factor. In the topographic profile I chose, the gradient near the base of the slope reaches the angle of stability, the flux gets very large, and hence erosion dominates over production for a wide range of model parameters. For this reason, bare slopes are predicted form under a wide range of values of \( P_0 \) and \( D \). The patterns illustrated in Figure 2 apply only to this particular hillslope form. Nevertheless, to the extent that (20) is a useful representation of a typical hillslope form in steep upland terrain, the patterns illustrated in Figure 2 are generally applicable to slopes of this form.

[13] Equations (17)–(19) provide a template for solving the NDSD model in three dimensions using either the exponential or humped production functions. Because the production and erosion terms are evaluated explicitly at each point, the humped production function can be evaluated just as easily as the exponential function. In moving from two to three dimensions, however, pixels in the model must be evaluated in a specific order. In two dimensions, the NDSD model is evaluated starting at the divide and then increasing with distance from the divide. This order is necessary because the value of \( h_{i-1} \) must be known before the value of \( h_i \) can be evaluated in (19). In three dimensions, however, there is no single direction in which the pixels can be evaluated from divides to channels. The solution to this problem is to evaluate the pixels systematically from highest elevation to lowest, as described in section 2.2.

3.2. Three-Dimensional Modeling of Soil Thickness for Real Landscapes

[14] An optional first step in estimating soil thicknesses in three dimensions is to preprocess the high-resolution DEM prior to input. DEM preprocessing is necessary in some cases because high-resolution DEMs often contain significant small-scale roughness (e.g., tree throw, animal burrows, etc. (Figure 3)). These are assumed to be short-term, local variations in soil thickness around a time-averaged thickness at a given point. Because the goal is to model that long-term average thickness, it is necessary to smooth the local topography. In addition, small-scale variability can represent a problem for modeling soil thicknesses because the model predictions depend sensitively on the topographic slope and curvature. Whenever a noisy data set such as a lidar DEM is differentiated, the small-scale variability in the data is amplified relative to the large-scale “signal” that represents the overall hillslope form. Without smoothing, some of the models of this paper predict wildly varying soil thicknesses, as we will show. However, the necessity of smoothing input DEMs prior to running the application depends on the sediment transport model. For the NSD and NASD models, it is essential to smooth the DEM in order to obtain reasonable results. For the NDSD model, accurate results can be obtained without preprocessing of the input DEM.

[15] In the model, we preprocess the DEM by applying the diffusion equation to hillslope pixels in the DEM for a prescribed length of time. How long the diffusion equation should be applied depends on the smoothing length scale desired. Diffusion should be applied long enough to eliminate most of the topographic variability at scales less than a few meters, so a smoothing length scale of between 1 and 3 m is reasonable. The smoothing length scale \( l \) is related to time and the diffusivity \( \kappa \) by the relationship \( l = \sqrt{\kappa t} \). Therefore, to achieve a prescribed value of \( l \), the diffusivity \( \kappa \) and time \( t \) should be chosen so that the square root of their product gives the desired value of \( l \). Other methods of smoothing...
(i.e., a Gaussian filter of a prescribed width) could also be used for this purpose. Prior to applying the diffusion equation, a flow routing algorithm was applied to automatically identify channels in the DEM (based on a threshold contributing area). The threshold contributing area was chosen by trial and error, minimizing the mismatch between the channel head locations predicted by the model and the location of the actual channel heads in the study site. The diffusion equation was then applied only to the nonchannel (i.e., hillslope) pixels. Once the DEM was smoothed, the soil thicknesses corresponding to the NSD and NASD models with the exponential humped production function can be easily computed using the 3-D versions of (17)–(19). The 3-D version of the NDSD model can be written as

$$\frac{\rho_b}{\rho_s} \frac{\partial}{\partial t} \left[ 1 + \left( \frac{\partial z}{\partial x} \right)^2 \right] \exp \left[ - \frac{h_y}{\left( \frac{h_0}{1 + \left( \frac{\partial z}{\partial x} \right)^2} \right)} \right]$$

$$= \frac{D_3}{\Delta x} \left( \frac{\partial^2 z}{\partial x^2} \frac{h_{jup}}{h_{jup} + h_{jup}} - \frac{\partial^2 z}{\partial y^2} \frac{h_{jup}}{h_{jup} + h_{jup}} \right) + \frac{\partial^2 z}{\partial x^2} \frac{h_{jup}}{h_{jup} + h_{jup}} \left( 1 - \left( \frac{\partial z}{\partial x} \right)^2 \right) \left( 1 - \left( \frac{\partial z}{\partial y} \right)^2 \right)$$

(21)

where the indices $iup$ and $jup$ refer to the adjacent pixels that are upslope in the $x$ and $y$ directions, respectively, from pixel $i$, $j$. In the model, the left and right sides of (21) are evaluated for all possible values of $h_i$ in order to find the value that minimizes the difference between the two sides of the equation. In order to implement (21), it is necessary to evaluate the pixels from highest to lowest elevation. To do this, the model first creates an index table using the Numerical Recipes routine “indexx” [Press et al., 2007]. The model then begins by calculating the soil thickness at the highest elevation pixel in the DEM. This pixel, by definition, is a divide with no up slope neighbors. As such, the $iup$ and $jup$ terms are zero in this case. As the model moves from pixels with high elevations to pixels with progressively lower elevations, the highest-to-lowest sequence guarantees that for each pixel $i$, $j$, the model will have already evaluated the values of $h_{iup}$ and $h_{jup}$ pixel prior to that step. This approach yields well defined solutions for each pixel in the DEM.


4.1. Study Site

[16] In this section we illustrate the application of the model to two small (~0.1 km²) semiarid watersheds located in the Marshall Gulch drainage basin, a high-elevation site within the Santa Catalina Mountains north of Tucson at 31°16´41”W and 33°25´46”N (Figure 4). Located within an elevation range of approximately 2300–2600 m above sea level, this site has an annual precipitation of 65 and 92 cm. Basin A is composed of predominantly schist, with some granite and quartzite at higher elevations. Basin B is composed entirely of granite. These basins were chosen because of their relatively similar size, shape, and slope aspect (both drain north shall Gulch). Freeze-thaw-driven and bioturbation (floral and faunal) are the dominant transport processes acting on the north facing hillslopes of Marshall Gulch. A dense layer of litter protects the hillslope from rain splash and overland flow, and no rills are present. Given the predominance of these processes, and the presence of soils ranging in thickness from 0 to 1.2 m, it is reasonable to expect that the NDSD model is the most applicable transport model for this site.

[17] A total of 35 soil pits were dug in basin A and 20 pits in basin B in order to characterize the variations in soil thickness along transects from ridge to channel. Soil pit locations were surveyed with a total station to coregister the measured values with model predictions. Soil profiles throughout the study area were characterized by a distinctive upper layer of colluvial deposition dominated by angular, cobble-sized clasts and varying in thickness from 30 to 50 cm. Below this colluvial layer was a layer of saprolite that ranged in thickness from 10 to 90 cm. We use the term saprolite to refer to “residual” soil derived from bedrock locally as well as the saprolite transported from upslope by freeze-thaw-driven creep and bioturbation (but not from mass movements, the material from which is present in the colluvial layer). Below the saprolite is a gradual transition to unweathered bedrock.

[18] Figure 5 illustrates the relationships between elevation, slope, and soil thickness on a typical hillslope profile located along the dashed line in Figure 5a. Soils are thin (i.e., 0–40 cm) on the ridgetops and generally thicker (50–120 cm) on the side slopes throughout basin A. Spatial variations in soil thickness on side slopes are controlled largely by whether the slope is locally converging or diverging, with converging slopes having thicker soils. Several small bedrock outcrops occur in the study area. These outcrops are associated with steep (>45°) slopes and are referred to in Figure 5 as “cliff-forming” outcrops. Three such outcrops occur in the first 70 m of the transect illustrated in Figure 5b. These outcrops provide potential additional data points for soil thickness because they are, by definition, devoid of soil cover. We chose not to include these outcrops in our database of soil thicknesses, however, because the lidar data do not sufficiently resolve the topography at the small scales over which these outcrops form. The main reason why the lidar data does not sufficiently resolve these features is that the Marshall Gulch basin is has extensive tree canopy vegetation cover (e.g., Figure 3). This cover limits the resolution obtainable in the lidar data to approximately 1 m. The small size of the outcrops, the limited resolution of the bare ground DEM, and the steep slopes and abrupt changes in slope in the vicinity of these outcrops make it difficult for the model to...
predict soil covers in these areas. In addition, any smoothing performed prior to input into the model exacerbates the already limited ability of the DEM to resolve the topography in the immediate vicinity of these small outcrops.

4.2. Comparison Between Model Results and Measured Thicknesses

[19] Figure 6 illustrates the results of applying the smoothing procedure to basin A with no smoothing (Figure 6a) and for smoothing length scales of 1 and 2 m (Figures 6b and 6c). Figure 7 illustrates the best fit results of the NSD sediment transport model with the exponential production function. As Figure 7 illustrates, the results of the NSD are strongly dependent on the smoothing length scale and the predicted values vary wildly from zero to values in excess of 2 m (white in Figure 7 represents a maximum “cutoff” value of 2 m). In areas that are concave up, the model breaks down entirely because, as in (4), there is no finite value for the soil thickness that can balance erosion if the landscape is depositing instead of eroding. In these areas, the cutoff value of 2 m is assumed. The results of the NASD model with different smoothing values (not shown) show a very similar pattern to those of the NSD model, i.e., values vary wildly from zero to greater than 2 m. As a result, both the NSD and NASD models predict soil thicknesses that are very different from those measured in the field at this study site.

Figure 4. Shaded relief images of (a) Marshall Gulch drainage basin and its location within Pima County, Arizona, (b) basin A and (c) basin B. Open circles show the locations of soil pits. The measured soil thickness (in cm) corresponding to each pit is listed next to the circle.
Figure 5. Illustration of relationships between elevation, slope, and soil thickness along a transect at the study site. (a) Gray scale map of slope gradient (black is gentle slope, white is steep slope). Cliff-forming outcrops are noted in the image. (b) Plot of elevation, slope, and soil thickness along the transect identified in Figure 5a.

Figure 6. Gray scale maps of (a–c) shaded relief DEM resulting from diffusing the topography over a length scale of 0 m (no smoothing) (Figure 6a), 1 m (Figure 6b), and 2 m (Figure 6c), with (d–f) the corresponding color maps of best fit soil thicknesses predicted by the NDSD model with exponential soil production functi
The strong spatial variability in soil thickness predicted by the NSD and NASD models is a result of the fact that soil thicknesses in these two models are very sensitive to the local topographic curvature. It is likely that the failure of both the NSD and NASD models to predict reasonable soil thicknesses is fundamental at this study site and not simply related to the presence of small-scale roughening of the topography by tree throw, animal burrowing, and/or other roughening processes. In the diffusion model of hill-slope evolution (a.k.a. linear slope-dependent model), the soil thickness in steady state is inversely related to the local topographic curvature as shown in (4). The NSD and NASD models include a nonlinear slope dependence not present in the diffusion model, but the same concept applies, i.e., soil thickness is inversely related to the divergence of sediment flux, which is related primarily to changes in slope along distance. With no smoothing, the soil thicknesses predicted by the NSD and NASD models oscillate wildly as shown in Figure 7 for the NSD model. Part of this problem is associated with the fact that there are small-scale roughening processes present in this landscape that are not described by any of these transport models. As Figure 7 illustrates, however, smoothing the data to extract the larger-scale hillslope shape does not solve this problem. The reason is that hillslides at Marshall Gulch have variations in curvature at a range of scales from small scales all the way up to the hillslope scale. As a result, smoothing the data to remove the small-scale variability does not eliminate the problem of soil thicknesses oscillating unrealistically between low and high values. When the NSD model is smoothed at scales of 4 m, the predicted soil thicknesses still oscillate between very thin and very thick soils (as they did with no smoothing), just over a larger spatial scale. Smoothing the data at still larger scales (8 m, 16 m, etc.) does not eliminate this strong “oscillation” pattern, at scales of approximately 10 m and greater, smoothing begins to alter the shape of the large-scale hillslope profile (i.e., divides become artificially rounded). Nevertheless, the possibility that the assumption of steady state does not apply to some of the locations in the study area cannot be ruled out as a reason for the discrepancy between the predictions of the NSD and NASD models and observed soil thicknesses.

Figure 7. Gray scale maps of soil thickness predicted by the NSD model using the input DEM diffusively smoothed over a length scale of (a) 0 m (no smoothing), (b) 2 m, and (c) 4 m. The results of the NASD model (not shown) are very similar to those of the NSD model. Both models are characterized by “oscillations” in soil thickness between relatively thin and relatively thick values at this site.
NASD models break down because there is no soil thickness that can provide the (negative) soil production rate necessary to balance a hillslope undergoing deposition. In contrast, erosion in areas that are concave up can still be achieved in the NDSD model provided that the soil thickens sufficiently with distance downslope. To see this, consider the NDSD model equation (written without uplift, for simplicity):

\[
\frac{\partial z}{\partial t} = D_3 \left( \nabla^2 h_0 \cdot \nabla z + h_0 \nabla \cdot \left( \frac{\nabla z}{1 - (|\nabla z|/S_c)^2} \right) \right) \tag{22}
\]

In (22), the first term on the right side of (7) is expanded to emphasize that erosion in the NDSD model is controlled by both the soil thickness and the gradient of soil thickness locally. Therefore, the second term on the right side of (22) can be positive (i.e., concave up) and yet the surface can still erode as long as the gradient of soil thickness is sufficiently large that the magnitude of the first term on the right side of (22) is larger than that of the second term, i.e., if the soil thickness thickens sufficiently downslope. In summary, the NDSD model is far less sensitive to local curvature and results in finite soil thicknesses everywhere on the landscape because the erosion term is sensitive to both the soil thickness and the gradient of soil thickness. In contrast, erosion in the NSD and NASD models is not sensitive to soil thickness, making soil thickness steady state conditions possible under a smaller range of conditions.

Figures 8 and 9 plot soil thicknesses predicted by the NDSD model for the exponential and humped production functions for a characteristic thickness of \(h_0 = 0.5\) m for the exponential production function and \(h_0 = 0.3\) m for the humped production function, with best fit values of \(S_c = 1.0\) and \(\rho_0 P_0/\rho D_3 = 0.04\). The best fit values of \(S_c\) and \(\rho_0 P_0/\rho D_3\) were determined by a brute force search through the parameter space to find the parameter set that minimized the root-mean-square difference between the predicted and measured data for all of the soil pits. The values of \(h_0 = 0.5\) m for the exponential production function and \(h_0 = 0.3\) m for the humped production function were assumed to be fixed based upon the limited variation in the values of \(h_0\) measured by cosmogenic radionuclide studies. Model predictions yield good agreement between the model and measured soil thicknesses for both basins despite the difference in lithology between the two sites (e.g., dominantly schist (basin A) versus granite (basin B)). Qualitatively, the model predicts thin soils on ridgetops (0–50 cm) and thicker soils on side slopes (50–150 cm), a pattern that is consistent with the measured soil thicknesses in Figure 4. The model results also predict that soils are thicker in areas of convergence compared to soils in areas of divergence that have comparable slope gradients. Both of these patterns are consistent with spatial variations in soil thickness documented by Heimsath et al. [1997, 1999, 2001] at a several well-studied areas worldwide.

The root-mean-square differences between the model predictions and measured soil thicknesses for all of the soil pits are as follows: 0.82 m for the NSD model with exponential soil production function, 0.87 m for the NASD with exponential soil production function, 0.16 m for the NDSD model with exponential soil production function, and 0.14 m for the NDSD model with humped production function. All of the models use 2-m smoothing of the DEM prior to input into the model, but the root-mean-square difference is relatively insensitive to the smoothing length scale employed. On the basis of these data, the NDSD model is
clearly superior to the NSD and NASD models at this study site. This result is not surprising given the dominant role of depth-dependent transport processes (i.e., creep, bioturbation, and mass movements) at this location.

[24] The NDSD model with the humped production function has a slightly lower root-mean-square difference relative to that of the NDSD model with the exponential production function. This is also evident in Figures 8 and 9. As Figures 8 and 9 illustrate, the NDSD model with exponential production function predicts that soils are everywhere thicker than 20 cm, a result inconsistent with the presence of several bare outcrops in the study area. It is difficult to use these results to argue unambiguously that the humped production function is operating at this site, however, because it is difficult to disprove the alternative hypothesis that these outcrops form because the bedrock is locally more resistant. Heimsath et al. [1997, 1999, 2001], for example, has measured lower soil production rates on tors at all of their study sites. These authors attributed these lower rates to more resistant bedrock (i.e., “core stones”) rather than the operation of a humped production function, as proposed by Strudley et al. [2006]. Despite this caveat, the humped production function is a slightly better predictor of the occurrence of thin soils in the vicinity of the steep cliff-forming outcrops in the two basins (see, especially, the southernmost profile of basin A (outcrop locations in Figure 5)). As such, the results of this study provide at least preliminary support for a humped production function at this study site.

4.3. Further Discussion

[25] Although it is difficult to prove the existence of a soil thickness steady state condition at the Marshall Gulch study site, it is possible to quantify the likelihood that a soil thickness steady state condition is operable by estimating a soil residence time and relating that value to the time since the last since major soil-altering climatic change at that location. Pelletier and Rasmussen [2009], for example, estimated the soil residence time on a semiarid granitic hillslope with a soil thickness of 0.5–1.0 m (e.g., the north facing slopes of Marshall Gulch) to be approximately 2–5 ka (see Pelletier and Rasmussen’s Figure 5c). To arrive at this result, Pelletier and Rasmussen [2009] used a climatically controlled model to estimate \( P_0 \) in granitic terrain and then used the soil production function to estimate production rates as a function of soil depth, balancing production and erosion to infer a residence time. Since this residence time is within the late Holocene time frame and the Pleistocene-Holocene transition was the last major soil-altering climatic event in this region (triggering the onset of the modern monsoon-dominated precipitation regime), a steady state soil thickness condition is likely to be achieved at this site.

[26] The NSD and NASD models do not predict realistic soil thickness values for the Marshall Gulch study site using the soil thickness steady state framework of this paper. It is reasonable to ask under what conditions the NSD and NASD models could be successfully used at other sites within this framework given that both of those models are very sensitive to the presence of hummocky topography at small and intermediate scales. First, in order to apply the NSD and NASD models, the dominant processes must be limited to processes occurring at the surface or within a near-surface layer that is everywhere smaller than the soil thickness on the slope. Such landscapes may have less hummocky topography than that of Marshall Gulch, given that rain splash-dominated hillslopes (i.e., “badland” hillslopes) often have classic convex topographic profiles with “roughening” processes limited to small scales. As a result, the small and intermediate-scale hummocky topography that presents such a difficulty for the NSD and NASD
models at the Marshall Gulch study site may not be present at study sites where the NSD or NASD models are applicable based on the processes active at those sites. It is also important to note that the assumption of a soil thickness steady state condition may not apply to some study sites or at all locations within a given study site. In such cases, the best approach may be to model the topography forward in time to a topographic steady state condition, as in the work by Roering [2008]. This approach removes the hummocky topography but comes at the cost of assuming a topographic steady state condition for a landscape that might not be in topographic steady state. In addition, the Roering [2008] approach introduces an additional free parameter into the application, i.e., the rock uplift rate U. Finally, in many cases the resistance of the bedrock to weathering may be sufficiently heterogeneous that the assumption of uniform P0 values may not be valid. Spatial variations in P0 due to lithologic heterogeneities pose a challenge for all soil production and landscape evolution models, but the sensitivity of the NSD and NASD models to the presence of hummocky topography may make lithologic heterogeneities especially problematic for predicting soil thicknesses using those transport models within the soil thickness steady state model framework of this paper.

5. Conclusions

[27] The model framework developed in this paper provides a flexible, generally applicable approach to predicting soil thicknesses based on a balance between soil production and erosion. Application of the model to example basins within the Marshall Gulch watershed highlights the value of the nonlinear depth- and slope-dependent transport model for modeling soil thicknesses in landscapes dominated by creep, bioturbation, and mass movements. Although the example applications in this paper are limited to small drainage basins, there is no limitation on the size of the watershed that can be analyzed using this model framework. It should be noted, however, that the model as it is applied in the example of this paper assumes that the maximum weathering rate, P0, is uniform throughout the area of calibration. Weathering rates vary with lithology and, often, with slope aspect. In the semiarid regions of the southwestern United States, for example, it is common for north facing slopes to have thicker regolith, and hence higher effective weathering rates, compared to south facing slopes of comparable slope gradients. If the model is to be applied to a study site in which P0 varies (due to varying lithology, aspect, or some other factor), the model subregions with different values of P0 must be separately calibrated to field-based measurements.

[28] Acknowledgments. We wish to thank Amy Rice, Matt Jorgensen, and Maite Guardiola-Claramonte for assistance with field work. We also thank Evan Canfield and the Pima County Flood Control District for generously sharing their lidar data set for portions of the Santa Catalina Mountains. Associate Editor Andrea Rinaldo and two anonymous reviewers provided helpful comments that greatly improved the manuscript. This study was partially funded by the Water Sustainability Program of the State of Arizona.

References


Pelletier, J. D., and C. Rasmussen (2009), Quantifying the climatic and tectonic controls on hillslope steepness and erosion rate, Lithosphere, 7, 73 – 80, doi:10.1130/1.3


Roering, J. J. (2004), Soil creep and convex-upward velocity profiles: Theoretical and experimental investigation of disturbance-driven sedi-


---

J. D. Pelletier, Department of Geosciences, University of Arizona, 1040 East Fourth Street, Tucson, AZ 85721-0077, USA. (jdpellet@email.arizona.edu)

C. Rasmussen, Department of Soil, Water, and Environmental Science, University of Arizona, 1177 East Fourth Street, Tucson, AZ 85721-0038, USA.