A landslide runout model for sediment transport, landscape evolution and hazard assessment applications

Jeffrey Keck[1,2], Erkan Istanbulluoglu[1], Benjamin Campforts[3], Gregory Tucker[4,5], Alexander Horner-Devine[1]

[1] University of Washington, Civil and Environmental Engineering, Seattle, WA, USA
[2] Washington Department of Natural Resources, Forest Resources Division, Olympia, WA, USA
[3] Department of Earth Sciences, Vrije University, Amsterdam, Netherlands
[4] Department of Geological Sciences, University of Colorado Boulder, Boulder, CO, USA
[5] Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado Boulder, Boulder, CO, USA

Correspondence to: Jeffrey Keck (keckje@gmail.com)

Abstract

We developed a new rule-based, cellular-automaton algorithm for predicting the hazard extent, sediment transport and topographic change associated with the runout of a landslide. This algorithm, which we call MassWastingRunout (MWR), is coded in Python and implemented as a component for the package Landlab. MWR combines the functionality of simple runout algorithms used in landscape evolution and watershed sediment yield models with the predictive details typical of runout models used for landslide inundation hazard mapping. An initial DEM, a regolith depth map, and the location polygon of the landslide source area are the only inputs required to run MWR to model the entire runout process. Runout relies on the principle of mass conservation and a set of topographic rules and empirical formulas that govern erosion and deposition. For the purpose of facilitating rapid calibration to a site, MWR includes a calibration utility that uses a Markov Chain Monte Carlo algorithm to automatically calibrate the model to match observed runout extent, deposition and erosion. Additionally, the calibration utility produces empirical probability density functions of each calibration parameter that can be used to inform probabilistic implementation of MWR. Here we use a series of synthetic terrains to demonstrate basic model response to topographic convergence and slope, test calibrated model performance relative to several observed landslides, and briefly demonstrate how MWR can be used to develop a probabilistic runout hazard map. A calibrated runout model may allow for region-specific and more insightful predictions of landslide impact on landscape morphology and watershed-scale sediment dynamics, and should be further investigated in future modelling studies.

1. Introduction

Over geologic timescales, landslides and their runout shape the topographic expression of mountain ranges and channel networks (e.g., Campforts et al., 2022; Korup, 2006; Larsen and Montgomery, 2012; Montgomery and Dietrich, 1988). Over more pragmatic engineering and environmental risk management timescales, landslides and their runout can inundate and destroy infrastructure (e.g., Kean et al., 2019) but also support numerous ecosystem benefits, including carbon and nutrient transport from hillslopes to channels and the creation of riparian habitat (Benda et al., 2003; Bigelow et al., 2007; Goode et al., 2012). Therefore, explicit representation of landslide runout is a necessary component of: (1) landslide inundation hazard assessments, with emphasis on inundation extent and flow depth (e.g., Frank et al. 2015; Han et al., 2015); (2) watershed sediment yield models, with emphasis on the mobilization, deposition and type of sediment carried by the landslide (e.g., Bathurst and Burton, 1998; Istanbulluoglu,et al., 2005); and (3) landscape evolution models, with emphasis on topographic change prediction (e.g., Tucker and Bras, 1998; Istanbulluoglu and Bras, 2005; Campforts et al., 2022);

Landslide runout processes can be generalized into three phases: initiation, erosion, and deposition. After a landslide initiates, it may break apart and flow as a relatively dry debris slide, or it may mix with surface runoff to become a debris flow. The mobility of the mass wasting material and resulting erosion/deposition pattern often varies as a function of runout topography and initial relief and size of the landslide (Iverson, 1997). Mobility may also be impacted by substrate liquefaction (Hungr and Evans, 2004) and landslide basal cataclasis (Shaller et al. 2020). As the runout material moves downslope, flow depth varies as a function of channel width (Kean et al, 2019), which in turn impacts erosion rates (Schürch et al. 2011). Theoretical, field and laboratory observations indicate that erosion rates may also depend on the moisture content of the channel bed (Iverson, 2012; McCoy et al. 2012), flow grainsize (Egashira et al., 2001) and granular stress within the flow (Capart et al, 2015). The slope at which deposition begins is controlled by the grain to water ratio and friction angle of the slide material (Takahashi, 2014; Major and Iverson, 1999; Zhou et al., 2019) but the friction angle of the material may vary as a function of the grains in the flow and fluidization of the flow material (Hutter et al., 1996). Lateral levees often form along the edges of the flow (Major, 1997; Whipple and Dunne, 1992; Shaller et al., 2020) and deposition at the distal end of the flow may occur as layered accretion (Major, 1997) or as the emplacement of a single, massive deposit (Shaller et al., 2020). If the water content of the runout material is high enough, as the solid fraction of the distal end of the flow compresses, the water is squeezed out and may continue as an immature debris flow (sensu Takahashi, 2014) or intense bedload (sensu Capart & Fraccarolo, 2011), extending the runout distance (e.g., Shaller et al. 2020).

Landslide inundation hazard models aim to accurately predict the runout extent and/or flow depths of a runout event and may include some or most of the above processes in the model. Example models include: (1) site-specific-empirical/statistical models that use simple geometric rules and an estimate of the total moblized volume (initial landslide + eroded volume) or a growth factor (e.g., Reid et al. 2016); (2) detailed, continuum-based mechanistic models, which conceptualize the runout process as a single-phase or multiphase flow using the depth-integrated Navier-Stokes equations for an incompressible, free-surface flow (i.e., shallow water equations; Frank et al, 2015; Han et al., 2015; Iverson and Denlinger, 2001; Medina et al., 2008) and often (though not always) require pre-knowledge of the total moblized volume (e.g., Barnhart et al., 2021; Han et al. 2015); (3) reduced complexity flow-routing models that use rule-based abstractions of the key physical processes that control the flow (Clerici and Perego, 2000; Guthrie and Befus, 2021; Gorr et al., 2022; Han et al., 2017, 2021; Horton et al., 2013; Liu et al, 2022) and are typically implemented using just the initial landslide location and volume but often rely on heavy, site specific parameterization and; (4) hybrid modelling approaches that combine mechanistic models with empirical and reduced complexity approaches (D'Ambrosio et al., 2003; Iovine et al., 2005; Lancaster et al., 2003; McDougall and Hungr 2004).

For landscape evolution and watershed sediment yield applications (herein referred to as watershed sediment models, WSMs), the runout model must be scalable in both space and time, and capable of modelling the entire runout process given an internally modelled initial landslide body (e.g. Tucker and Bras, 1998; Doten et al 2006; Campforts et al. 2022). As such, computationally efficient and parsimonious reduced complexity runout models that evolve the terrain and transfer sediment are often preferred in WSMs, however with simplifications that can restrict model ability to accurately replicate observed inundation extent or depositional patters. Such simplifications include omitting debris flow erosion and bulking in runout channels, limiting flow to only a single cell in the steepest downstream direction, and assuming debris flows only occupy the width of a single cell (e.g., Tucker and Bras, 1998; Istanbulluoglu and Bras, 2005) or link of a channel network (Benda and Dunne, 1997).

To bridge the scalable functionality of WSMs with the predictive accuracy of landslide inundation hazard models, without the computational overhead of a detailed mechanistic representation of the runout process, or difficult parameterization typical of other models, we developed a new, reduced complexity landslide runout model, called MassWastingRunout (MWR). MWR models landslide runout starting from the source area of the landslide, making it easily compatible with WSMs that internally determine the initial landslide body size and location. MWR tracks sediment transport and topographic change downstream, and evolves the attributes of the transport material. MWR can be calibrated by adjusting just two parameters (S_c and q_c, described in Section 2) and is augmented with a Markov Chain Monte Carlo (MCMC) calibration utility that automatically parameterizes model behavior to observed runout characteristics (e.g., erosion, deposition, extent). MWR also includes a built-in utility called MWR Probability, designed for running an ensemble of simulations to develop probabilistic landslide runout hazard maps.

In this paper, we present the conceptualization and numerical implementation of the MWR model (Section 2), describe the calibration utility and its probabilistic implementation (Section 3) and demonstrate basic model response to topographic convergence and slope on a series of synthetic terrains (Section 4). Event-scale applications to replicate observed runout extent, sediment transport, and topographic change at four topographically and geologically unique field sites (see Figure 1) are discussed (Section 5). We test MWR's predictive ability using the parameterization of one site to predict runout hazard at a nearby site and show a brief example of Monte Carlo model runs to determine runout probability from initial landslide source areas defined by an expert-determined potentially unstable slope or a hydrologically-driven landslide hazard model (Section 6). We conclude with a short summary of MWR model performance and discuss how a calibrated MWR can be incorporated into WSMs.

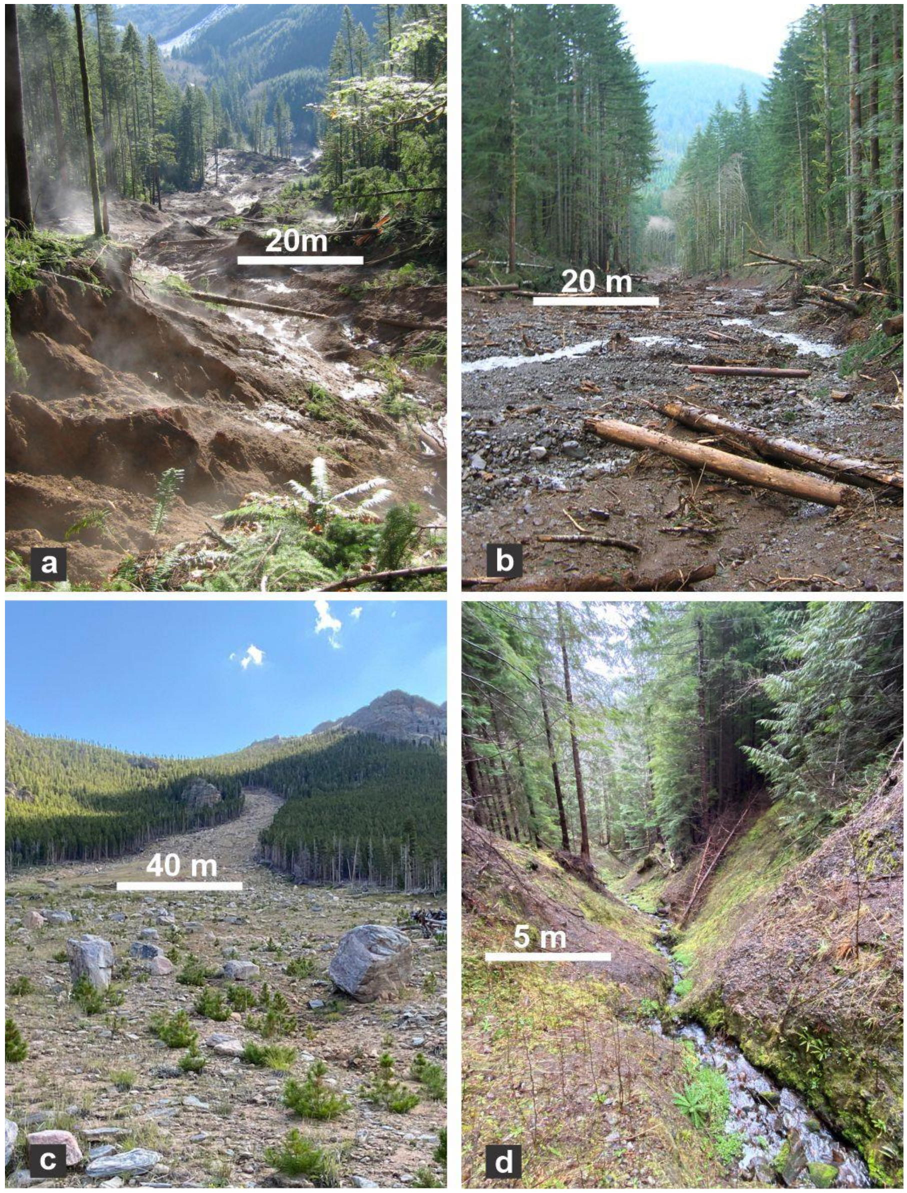

Figure 1: Example landslides used to evaluate calibrated MWR performance: (a) Cascade Mountains, WA: a large debris avalanche over steep, broadly convergent terrain (photo credit: Stephen Slaughter). (b) Black Hills, WA: large debris flows over a broadly convergent, gently sloped valley (photo credit: Stephen Slaughter). (c) Rocky Mountains, CO: a moderate sized debris avalanche over steep, unconfined to divergent hillslope. (d) Olympic Mountains, WA: small debris flows in steep, highly convergent channels. Image scale varies with depth, but approximate scale of the image is indicated at the location of the scale bar.

2. Description of the MassWastingRunout model

2.1 Overview of the cellular-automaton modelling approach

MWR is coded as a discrete cellular automaton (CA) model. CA models apply a set of equations or rules (deterministic or probabilistic) to individual cells of a grid to change the numerical or categorical value of a cell state (e.g., Codd, 1968). In earth sciences, CA models are widely used to model everything from vegetation dynamics (e.g., Nudurupati et al., 2023) to lava flows (e.g., Barca et al., 1993) to geomorphic transport, in which gravitationally directed erosion and depositional processes modify a digital elevation model (DEM) representation of a landscape (e.g., Chase, 1992; Crave & Davy, 2001; Murray & Paola, 1994; Tucker et al., 2018). Existing CA-based landslide runout models include models by Guthrie and Befus (2021), D'Ambrosio et al. (2003) and Han et al. (2021). In all of these models, runout behavior is controlled by topographic slope and rules for erosion and deposition but conceptualization and implementation differ.

In MWR, mass continuity is central to model conceptualization. Of the wide range of processes described in the introduction that control observed runout, MWR explicitly represents erosion, deposition, and flow resistance due to debris size and vegetation. Material exchange between the runout material and underlying terrain as well as flow resistance determines runout extent and landscape evolution. Model rules are designed such that they can be parameterized from field measurements. Finally, in MWR, most computations occur only at the location of moving debris, in a manner analogous to the "mobile" cellular automaton implementation of Chase (1992).

Chase (1992) modelled precipitation-driven surface erosion by randomly placing single packets of precipitation on a DEM, which then moved from higher elevation to lower elevation grid cells, eroding and transporting sediment as a function of the slope between the cells. The individual packets of precipitation were referred to as precipitons. In MWR, since we route the downslope progression of debris from a specified mass wasting source area, we refer to these packets of debris as "debritons". The debritons represent debris flux, here defined as a volume of debris transferred per model iteration per grid-cell area, [m^3/m^2/iteration] and are equivalent to the flow depth in the cell.

The present implementation of the MWR algorithm is coded in Python and developed as a component of the Landlab earth surface modeling toolkit (Barnhart et al., 2020; Hobley et al., 2017). MWR uses the Landlab raster model grid, which consists of a lattice of equally sized, rectangular cells. Topographic elevation, derived topographic properties like slope and curvature, and other spatially varying attributes such as regolith depth and grain size, are recorded at nodes in the center of each cell (see Figure 5 of Hobley et al., 2017). In the subsequent sections we describe the model theory. All parameters and variables used in the theory are listed in the Notation section.

2.2 Mobilization of the initial mass wasting source material (Algorithm 1):

To initiate MWR, the user provides maps of initial topography, regolith depth, and the location and depth of the mass wasting source material (e.g., the initial landslide body). Each raster model grid node in the mass wasting source material is designated as a debriton (Figure 2, iteration t = 0) with a magnitude equal to the mass wasting source material depth and basal elevation equal to the initial topography minus the mass wasting source material depth. The basal elevation can be thought to represent the rupture or slip surface of the source material and the redistribution (flux) of each debriton to its downslope nodes (receiver nodes) is determined as a function of the slope of the slip surface. At the lowest-elevation debriton of the source material, flux to its downslope nodes is determined using the surface slope of the initial DEM (see flow direction of lowest node in Figure 3a). This implementation helps to ensure that the lowest-elevation debriton in the mass wasting source material moves downslope and movement of upslope debritons are impacted by the geometry of the mass wasting source material. For example, the receiver nodes of the lowest-elevation debriton in the landslide illustrated in Figure 2 (iteration t = 0, detailed in Figure 3a) would be identified as those among the eight neighboring nodes whose initial topographic elevation was less than the initial topographic elevation of the node while for the debriton at node 51, the receiver nodes would be identified as those among the eight neighboring nodes whose topographic elevation is less than the topographic elevation of the terrain underlying the debriton (the slip surface).

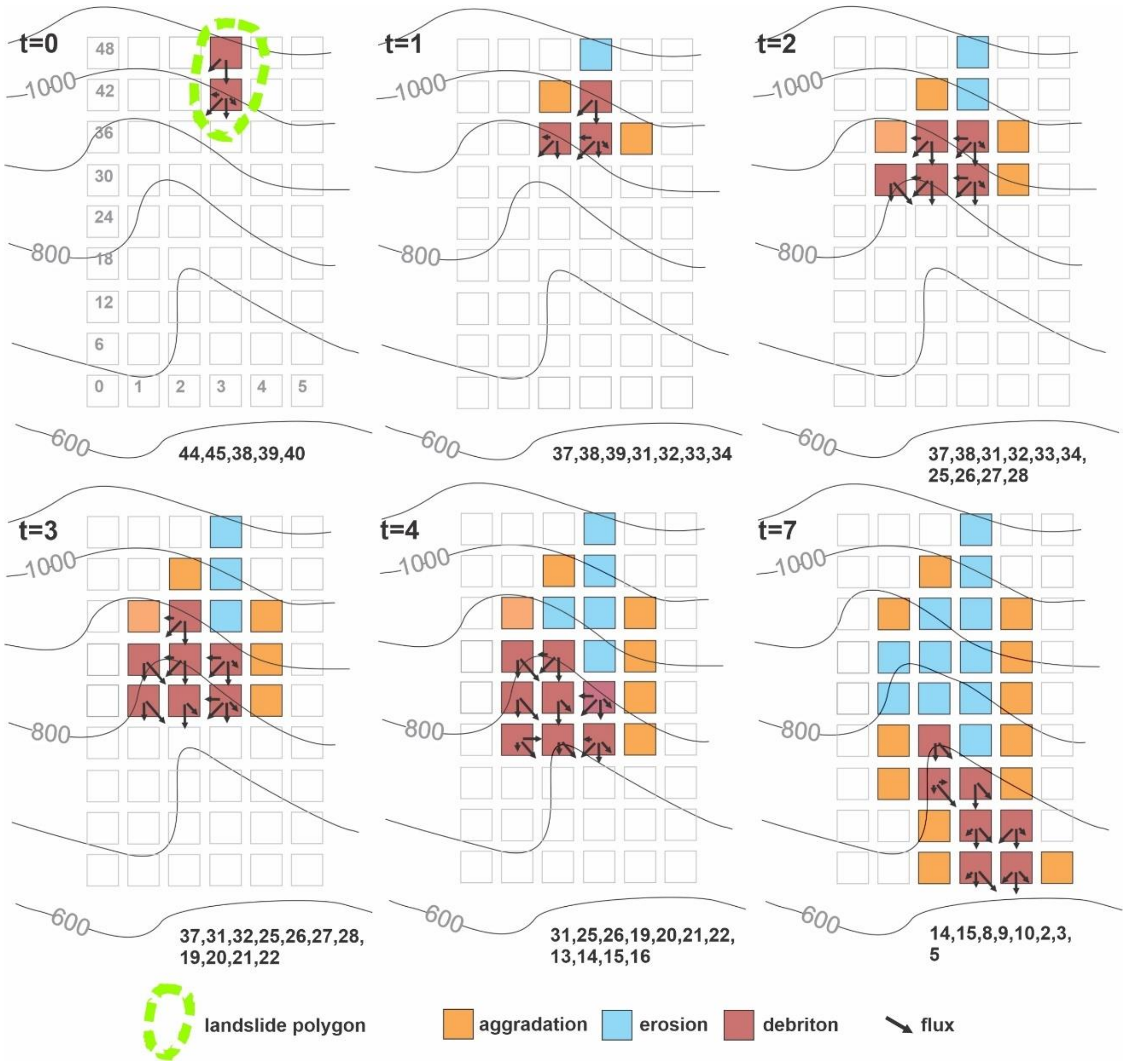

Figure 2. Illustration of initial mass wasting release and runout down a steep, convergent slope. Variable t indicates model iteration (not time). Notice how the flow elongates and widens as the model progresses and the number of receiver nodes (numbers listed at bottom of each panel) and quantity of mobilized material increase. -

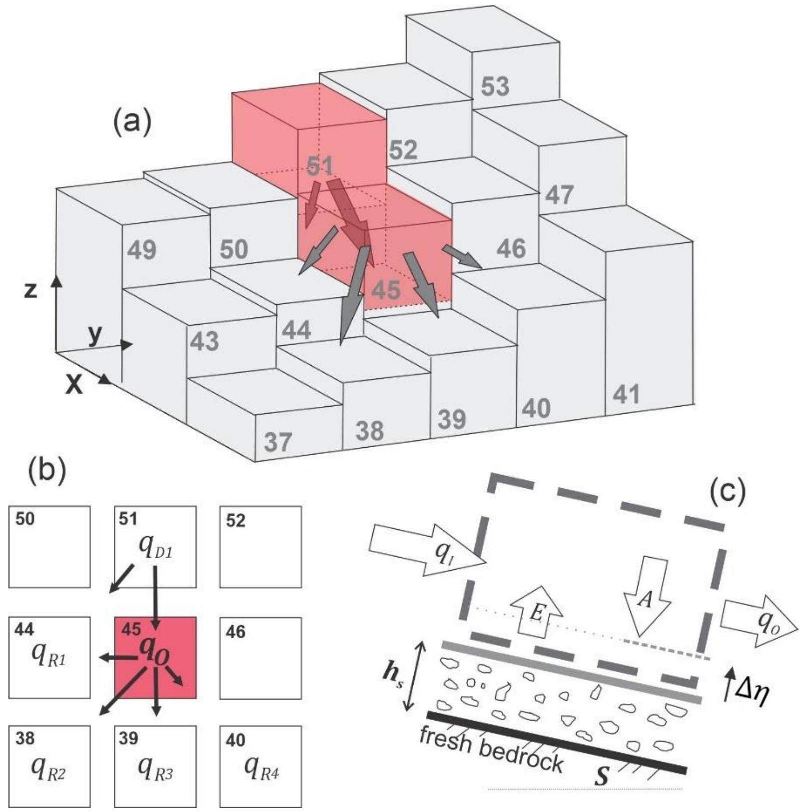

Figure 3. (a) Three-dimensional illustration of iteration $\boldsymbol{t}$ = 0 in Figure 2, showing initial source material nodes (represented by red cells) and flux towards downslope nodes. (b) Distribution of $\boldsymbol{q_O}$ to downslope nodes 38, 39, 40 and 44; (c) illustration of mass continuity applied to any node that receives a debriton.

2.3 Flow routing and rules for erosion, deposition and resistance (Algorithm 2)

Algorithm 2 is essentially the runout model. It determines how each debriton traverses and modifies the landscape. After receiver nodes from the first model iteration are determined in Algorithm 1 (iteration t=0), Algorithm 2 is repeatedly implemented until all material has deposited (i.e., there are no debritons). Each debriton moves one grid cell per model iteration, the larger the landslide size, the more iterations necessary to evacuate the landslide. As each debriton moves, it may erode or aggrade the landscape, impacting the movement of any upslope debritons. As is common with other reduced complexity models, we assume that inertial effects have negligible impact on flow behavior (i.e., the kinematic flow approximation). The downslope redistribution of a debriton or flux to each of a node's i-th receiver nodes (q_{R_i}) is determined as a function of topographic slope (slope of terrain under the debriton) using the Freeman (1991) multiple flow direction algorithm:

$$q_{R_i} = q_O \frac{S_i^a}{\sum_{i=1}^{Nr} S_i^a} \tag{1}$$

where q_O is the total out-going flux from the node and has units of depth [m] per model iteration, Nr is the number of receiving nodes, i is the index for each receiver node (e.g., $i = 1, 2 \ldots Nr$) and S_i is the underlying topographic slope to the i-th receiver node (Figure 3b). The Freeman (1991) multiple flow direction algorithm is a commonly used approximation for two-dimensional flow, and in this implementation it is handled by a pre-existing Landlab flow-routing component. The exponent a controls how material is distributed to downslope nodes, with higher values causing narrower flow (Holmgren 1994). In a braided river cellular-automaton model, Murray and Paola (1997) used an approximation for turbulent shallow water flow to justify a = 0.5 (which is the exponent on the slope factor in channel friction laws). For our application, we found MWR provided a closer fit to observed mass wasting runout if a = 1, suggesting that the material behavior is more similar to linear-viscous shear flow than to wall-bounded turbulent shear flow (e.g., as the runout debris flows downslope, it tends to spread less than shallow turbulent water). The total incoming flux (again, in units [m] per model iteration) towards a given node (q_I), is determined by summing the flux from each of the node's donor nodes:

$$q_I = \sum_{j=1}^{Nd} q_{D_j} \tag{2}$$

Where Nd is the number of donor nodes, and q_{D_j} is the flux from node D_j (the j-th donor node, j = 1,2… Nd; Figure 3b).

As noted by Tucker and Hancock (2010), the flow depths calculated from two-dimensional flow approximations like (1) can be influenced by the grid-size used to represent the terrain and depending on terrain slope and convergence, neglection of pressure and momentum forces may lead the model to underestimate or overestimate flow width in some circumstances. Rengers et al. (2016) noted that this same issue occurs when using a kinematic wave approximation of the shallow water equations because the kinematic wave approximation lacks a pressure term that would normally allow the modelled water surface to spread out. We consider flow depths determined from (2) as meaningful in the sense that they vary as a function of flux but less meaningful in the sense that they are affected by the limitations noted above. Furthermore, for the purpose of determining flow-depth-dependent erosion rates described later in this paper, and to provide a simplified representation of the effect of pressure forces, we constrain flow depths to no more than a maximum flow as:

$$h = min\ (h_{max}, q_I) \tag{3}$$

Where h_{max} is an effective upper limit to flow depth, that in practice can be approximated as the maximum observed flow depth, as inferred from field indicators or assigned based on expert judgement (See Section 5) and h is the corrected flow depth used to calculate flow shear stress. This correction allows erosion rates to vary with flux but prevents unreasonably large values. This flow depth correction does not violate the conservation of mass and runout mass balance, as h is only used to calculate flow shear stress.

To determine aggradation (A) at a node, we use a critical slope (S_c) constraint that permits computationally-rapid distribution of q_I over multiple nodes. Critical slope constraints or rules are common to many reduced complexity and landscape evolution models. Chen et al. (2023) showed that when flow inertia can be ignored, S_c can be approximated from the surface slope of observed deposits. Several landscape evolution models use a S_c-based nonlinear, nonlocal aggradation scheme (e.g., Campforts et al., 2020; Carretier et al., 2016) but when this rule is implemented with the debriton framework described above, unreasonably tall deposits result when q_I is large and slope at the node (S) $<<$

S_c. To resolve this problem, aggradation depth can be limited to $A \leq S_c \Delta x$, (where Δx grid cell length), but we found that this constraint results in long deposits that parallel the underlying slope when q_I is large. Instead, MWR computes the aggradation depth at a node assuming that the aggradation will spread over N_a nodes until all of q_I is deposited and that the surface slope of the overall deposit will be equal to S_c, as shown in Figure 4 and described as follows.

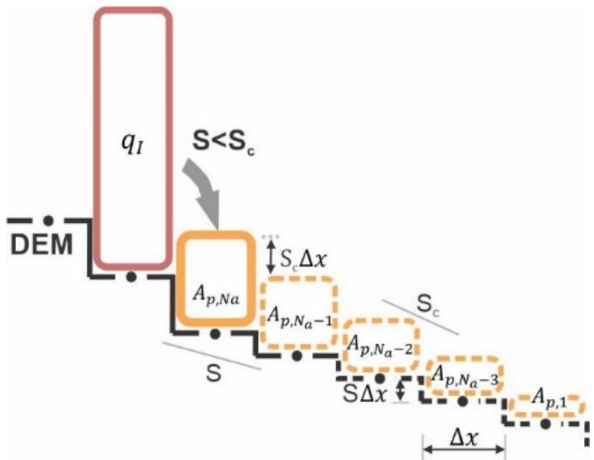

Figure 4. Illustration of aggradation rule used in MWR when q_I is assumed to spread over 5 nodes ($N_a = 5$). Solid yellow box indicates aggradation amount at a given node. Dashed yellow boxes and lines indicate the geometry of assumed the aggradation beyond the node. Dots along DEM surface are nodes.

Aggradation at a node is determined as:

$$A = \begin{cases} 0 & , \quad S \geq S_c \\ min(A_{p,N_a}, q_I) , & S < S_c \end{cases} \tag{4}$$

Where S is the steepest slope to the node's eight neighbouring nodes, A_{p,N_a} is a potential aggradation depth necessary to form a deposit that: (1) begins at the node and spreads over N_a consecutive nodes; (2) has a total volume equal to $q_I \Delta x^2$; (3) a surface slope equal the critical slope S_c and; (4) an underlying topographic slope equal to the steepest slope at the node and assumed constant over the N_a consecutive nodes of deposition. From this assumed deposit, we can analytically define A_{p,N_a} and N_a as a function of q_I, S_c and S as follows:

First, q_I, calculated from (2), can be used to calculate $A_{p,i}$ by expressing q_I as the sum of the N_a deposits that make up the overall deposit as:

$$q_I = \sum_{i=1}^{N_a} A_{p,i} \tag{5}$$

where $A_{p,i}$ is the i-th deposition amount in the deposit and $i = 1$ is the last node of deposition ($A_{p,1}$; see Figure 4). Since we assume the deposit slope and underlying topographic slope are uniform, the deposition amount at any of the N_a nodes can be determined from $A_{p,1}$ as:

$$A_{p,i} = A_{p,1} + (i-1)\Delta x(S_c - S) \tag{6}$$

From (6) we can re-write (5) as a function of $A_{p,1}$ and rearrange to define $A_{p,1}$ as a function of q_I:

$$A_{p,1} = \frac{1}{N_a} q_I - \frac{N_a - 1}{2} \Delta x (S_c - S) \quad (7)$$

Substituting (7) into (6) and solving for $i = N_a$, we get an expression for A_{p,N_a}:

$$A_{p,N_a} = \frac{1}{N_a} q_I + \frac{N_a - 1}{2} \Delta x (S_c - S) \quad (8)$$

Equation (8) can be rearranged into a quadratic equation and solved for N_a as:

$$N_a = \frac{-A_{p,1} + \frac{1}{2}\Delta x(S_c - S) \pm \sqrt{\left(A_{p,1} - \frac{1}{2}\Delta x(S_c - S) \right)^2 + 2\Delta x(S_c - S) q_I}}{\Delta x(S_c - S)} \quad (9)$$

We use (8) to solve for A_{p,N_a} and (9) to solve for N_a assuming $A_{p,1} = 1/2\, \Delta x S_c$ and rounding the positive solution to the nearest integer. When implemented using a single debriton, released on a two-dimensional hillslope as illustrated in Figure 4, the debriton deposits over N_a nodes at a uniform slope equal to S_c. When implemented on an actual three-dimensional terrain, the interaction between multiple debritons in multiple directions creates a complex deposit whose slope changes with S_c.

To determine erosion depth (E) [m/iteration], we constrain E to the lesser of a potential erosion depth, h_e, and local regolith depth, h_r:

$$E = min\,(h_r, h_e) \quad (10)$$

where h_e is computed as a function of the basal shear stress of the flow, τ [Pa], (Equations 12 and 13) and the critical shear stress (τ_c) of the regolith at the node [Pa]:

$$h_e = k(\tau - \tau_c)^f \quad (11)$$

The coefficient k is an erodibility parameter [m/Pa^f]. Stock and Dietrich (2006) showed that k encapsulates substrate properties. If h_e is used to represent erosion over geomorphic time scales, with repeated debris flow occurrences in a single model iteration, k becomes associated with debris flow length and frequency (Perron, 2017). In our application since we are modelling the erosion associated with a single runout event, as represented by the downslope movement of the debritons, the coefficient k therefore needs to scale h_e on the order of the average erosion depth caused by a single debriton. Using this logic, k can be computed using the observed average erosion depth and an estimated length of the runout material that caused the erosion. Further details on how we determine k from observed runout are included in the Appendix. The exponent f controls the non-linearity of h_e with shear stress. Many authors (Chen & Zhang, 2015; Frank et al., 2015; Shen et al., 2020) use a value of 1 for f but field measurements by Schürch et al. (2011) (see their Figure 3) suggest that f may be less than 1 if τ is assumed to vary linearly with flow depth, particularly at flow depths greater than 3 meters.

MWR includes two options for defining τ: (1) a quasi-static basal shear stress approximation or (2) a grain-size-based shear stress approximation. The quasi-static basal shear stress approximation (e.g., Takahashi, 2014) is defined as:

$$\tau = \rho g h \sin\theta \quad (12)$$

where ρ is the density of mass wasting material (grain and water mixture) [kg/m^3], g is gravity [m/s^2], h is the adjusted flow depth described in (3) and θ is the topographic slope ($\tan^{-1}(S)$) measured in degrees.

The grain-size-based shear stress approximation is defined using an empirical formula by Bagnold (1954):

$$\tau = \sigma \tan \varphi \tag{13}$$

Where σ is normal stress [Pa], φ is the collision angle between grains, measured from the vertical axis (See Bagnold, 1954), with a value of $\tan \varphi$ typically equal to 0.32. Stock and Dietrich (2006) defined σ as:

$$\sigma = \cos \theta \upsilon_s \rho_s {D_s}^2 \left(\frac{du}{dz}\right)^2 \tag{14}$$

Where υ_s is the volumetric solids concentration, ρ_s is density of the solids [kg/m^3], u is flow velocity [m/s], z is depth below the flow surface [m], du/dz is the shear strain rate [1/s] and D_s is the representative grain size [m]. Stock and Dietrich (2006) suggested that D_s corresponds to a small percentile of the coarsest fraction of the runout material (D_{88} to D_{96}) and they approximated du/dz as:

$$\frac{du}{dz} = \frac{u}{h} \tag{15}$$

Solely for the purpose of computing du/dz, we approximate velocity at a node using a grain-size dependent empirical formula for debris flow velocity by Julien and Paris (2010) as:

$$u = 5.75 u^* \log\left(\frac{h}{D_s}\right) \tag{16}$$

Where u^* is shear velocity $\left(\sqrt{gh \tan \theta}\right)$. Substituting (16), (15), (14) and (13) into (11) yields a grain-size dependent approximation for h_e that mimics the non-linear erosion response to flow depth in Schürch et al. (2011). Additionally, this form of τ is advantageous because it permits landslide-driven erosion rates to scale with landslide grain size, which can vary by lithologic region (e.g., Roda-Boluda et al. 2018). As will be shown in Section 5, we obtained reasonable model calibration at multiple sites by defining D_s from the coarser grain sizes observed in the field at existing runout-deposits, road-cuts and tree-throw pits.

Once A [m] and E [m] have been determined, total out-going flux per iteration, q_O [m] is determined as (see Figure 3c):

$$q_O = \begin{cases} q_I - A + E, & q_I \geq q_c \\ 0, & q_I < q_c \end{cases} \tag{17}$$

Where q_c is a threshold flux for deposition. When $q_I < q_c$, q_I deposits and q_O becomes zero. The threshold flux q_c conceptually represents the flow depth below which flow resistance is large enough to cease the forward momentum of the flow, whether in the form of internal friction or friction due to vegetation and obstructions (e.g., large clasts or logs). The density and water content of q_I, A, and E are treated as uniform and surface runoff, such as channelized stream flow or hillslope-infiltration-excess runoff, that might mix with q_I A, or E is ignored. Once q_I, A, q_O and E have been determined, change in elevation at a node ($\Delta\eta$) is calculated as:

$$\Delta\eta = A - E \tag{18}$$

Attributes (e.g., grain size, organic content or any other attribute that is transferred in the flow) of the debriton and regolith are updated using a volumetric-weighted average approach. First, for each regolith attribute being tracked by the model, the attribute value delivered to a node from its donor nodes (ξ_D) is determined as:

$$\xi_D = \frac{\boldsymbol{\xi_D} \cdot \boldsymbol{q_D}}{q_I} \tag{19}$$

where $\boldsymbol{q_D}$ is a vector containing all q_{D_j} sent to the node, $\boldsymbol{\xi_D}$ is a vector containing the incoming attribute values for each q_{D_j}, and q_I is the sum of incoming flux from donor nodes defined by (2).

Second, the attribute value sent from a node to its receiver nodes (ξ_R) is determined as:

$$\xi_R = \frac{\xi_{t-1}E + \xi_D(q_I - A)}{q_O} \tag{20}$$

where ξ_{t-1} is the attribute value at the node before any aggradation (i.e., the previous iteration attribute value). Finally, the attribute value at the node, updated to account for erosion and aggradation (ξ) is:

$$\xi = \frac{\xi_{t-1}(h_r - E) + \xi_D A}{A + h_r - E} \tag{21}$$

Regolith thickness (h_r) and topographic elevation (η) are updated at a node as:

$$\eta = \eta_{t-1} + \Delta\eta \tag{22}$$

$$h_r = h_{r\,t-1} + \Delta\eta \tag{23}$$

Where η_{t-1} and $h_{r\,t-1}$ are the topographic surface elevation and regolith thickness at the node from the previous model iteration. After regolith thickness and topographic elevation have been updated for each debriton, the multi-direction slope of the DEM, which is used for routing the debritons in the next model iteration, is recomputed from the topographic surface.

Using the above approach, debritons may become obstructed if they encounter a topographic pit or flat topography in the DEM. To allow a debriton to pass an obstruction, we rely on a simple work-around: upon encountering the obstruction, the debriton is directed to itself and some portion of the debris is deposited based on (4). At the end of the model iteration, the node elevation and slope are updated. During the next iteration, if the remaining mobile debris is no longer obstructed, it moves to its downslope node(s). If the node is still obstructed, it is again sent to itself until either all material has deposited or the elevation of the node exceeds that of its neighbour nodes, allowing the debriton to move downslope.

3. Calibration and MWR probability

3.1 Calibration utility

MWR includes an adaptive Markov Chain Monte Carlo (MCMC) calibration algorithm described by Coz et al. (2014) and Renard et al. (2006). The MCMC algorithm is implemented as a utility for MWR and identifies a single set of parameters that best match MWR output to an observed landslide runout dataset. The observed runout dataset can consist of a single or multiple landslides. Depending on user input, MWR simultaneously or sequentially models runout from each landslide source area in one model run. To use the calibration utility, the user provides an initial (prior) guess of the parameter values and their respective probability distribution functions (PDF) that calibrate the MWR to a specific site. Then, the calibration utility randomly selects a set of trial parameter values (Λ) from the prior PDFs and runs MWR using Λ. Once the model has completed the run, the algorithm evaluates the posterior likelihood of the parameter set ($L(\Lambda)$) as a lumped index of model ability to replicate observed runout (described below) and the prior likelihood of the parameter set. After the first $L(\Lambda)$ has been determined, the utility selects a new set of parameters (Λ_{t+1}) by jumping some distance (described below) from each parameter in Λ space. Depending on the value of $L(\Lambda_{t+1})$, the algorithm either stays at Λ or moves to Λ_{t+1}. This Markov process is repeated a user-specified number of times. Jump direction is random, but the algorithm is adaptive because the jump distance changes depending on if $L(\Lambda_{t+1}) > L(\Lambda)$ occurs more than a user specified threshold value. For a detailed description of the algorithm see Coz et al. (2014).

The $L(\Lambda)$ index is estimated as the product of the prior probability of the selected parameter values, $p(\Lambda)$, and three other performance metrics as:

$$L(\Lambda) = p(\Lambda) * \Omega_T * \frac{1}{\Delta\eta_E{}^2} * \frac{1}{Q_{s_E}{}^2} \tag{24}$$

where Ω_T is the Lee-Salle index (Heiser et al., 2017) for evaluating model planimetric fit and $\Delta\eta_E$ and Q_{s_E} are new dimensionless indices, proposed for this study (described below). The indice $\Delta\eta_E$ is the volumetric error of the modelled topographic change over the entire model domain normalized by the observed total mobilized volume (initial landslide body + erosion volume). The indice Q_{s_E} is the mean-cumulative sediment export error along the modelled runout path normalized by the observed mean cumulative flow. Larger values of Ω_T and smaller values of $\Delta\eta_E$ and Q_{s_E} indicate modelled runout more closely fits observed. Note that we add a value of 1 to Ω_T and use the squared-reciprocal values of $\Delta\eta_E$ and Q_{s_E} in (24) so that the magnitude of $L(\Lambda)$ is always equal to or greater than zero and increases with improved fit. The metric Ω_T is written as:

$$\Omega_T = \frac{\alpha-\beta-\gamma}{\alpha+\beta+\gamma} + 1 \tag{25}$$

where α, β and γ are the areas of matching, overestimated and underestimated runout extent, respectively.

The spatial index for volumetric error, $\Delta\eta_E$, is determined as:

$$\Delta\eta_E = \sqrt{\frac{\sum_{i=0}^{p}[(\Delta\eta_{Oi} - \Delta\eta_{Mi})\Delta x^2]^2}{V^2}}. \tag{26}$$

Where V is observed total mobilized volume and p is the number of nodes in the area made up of the matching, overestimated and underestimated areas of runout extent and $\Delta\eta_{Mi}$ and $\Delta\eta_{Oi}$ are the modelled and observed topographic change [m] at the i-th node within that extent.

To calculate Q_{s_E}, we first determine the cumulative export (flow) volume (Q_s) at each node, j along the runout profile, in a manner similar to the flow volume/mass balance curves in Fannin and Wise (2001) and Hungr and Evans (2004):

$$Q_s = -\Delta x^2 \sum_{i=1}^{u_j} \Delta\eta_{i,j} \tag{28}$$

where $\Delta\eta_{ij}$ is the topographic change [m] at the i-th node located upstream of node j, and u_j is the total number of all nodes located upstream of j. Q_s is computed for both the observed and modelled runout path (Q_{sO} and Q_{sM} respectively) and Q_{s_E} of a runout is determined as:

$$Q_{s_E} = \sqrt{\frac{\frac{1}{r}\sum_{j=1}^{r}(Q_{sO}-Q_{sM})^2}{\overline{Q_{sO}}^2}} \qquad (29)$$

Where r is the number of nodes along the center line of the runout path, and $\overline{Q_{sO}}$ is the observed mean cumulative flow.

As will be detailed in Section 5, field estimates for S_c and q_c, vary over the length of the runout path. To account for the heterogeneity of S_c and q_c, we estimate prior distributions of potential S_c and q_c values from field/remote sensing measurements. Then, from model calibration to a DEM-of-Difference (pre-runout DEM subtracted from the post-runout DEM; DoD) using the calibration utility, we find single values of S_c and q_c that allow the modelled DoD to replicate the observed DoD .

We run the calibration utility using a single Markov chain of 2000 repetitions. At most sites, the model converged relatively quickly on a solution and we therefore didn't consider burn-in or evaluate convergence (e.g., Gelman et al. 2021) and considered 2000 repetitions adequate. Future implementations of the calibration utility may include multiple chains, burn-in and a check for convergence. As a final note, many debris flow runout models are evaluated using Ω_T or variations of Ω_T alone (e.g, Gorr et al., 2022; Han et al., 2017) and the MWR calibration utility can also be run solely as a function of Ω_T. However, we found that calibration based on Ω_T (i.e., runout extent) alone results in high parameter equifinality (e.g., Beven 2006); multiple parameter sets result in an equally calibrated model as evaluated by Ω_T. As such, we recommend calibrating debris flow/ landslide runout models to an observed DoD. If repeated lidar is available, a DoD can be obtained from before and after scans of the observed runout event. Alternatively, a DoD can be created by hiking the observed runout event and mapping field-interpreted erosion and deposition depths. Additional details on how we prepared DoDs for multiple sites are included in the Supplementary Material.

3.2 Mapping landslide runout hazard

MWR includes an additional utility called MWR Probability that produces landslide runout probability maps. MWR Probability repeatedly runs MWR a user specified Np times, each repetition with a different, randomly sampled parameter set from the posterior parameter PDFs produced by the calibration utility. MWR Probability includes three options for specifying the initial mass wasting source material: (1) a user-provided landslide source area polygon(s) based on field and/or remote sensing observations; (2) a user-defined hillslope susceptible to landslides (e.g., potentially unstable slope), where landslide area and location are randomly selected within, but no larger than the hillslope; this option is useful when the extent of a potential landslide is unknown; and (3) a series of mapped landslide source areas within a watershed, as determined by an externally run Monte Carlo landslide initiation model (e.g., Hammond et al. 1992; Strauch et al., 2018) ; this option is useful for regional runout hazard applications. If using

Option 1, modelled runout probability represents uncertainty in MWR parameterization. If using Option 2 or 3, modelled runout probability reflects uncertainty in both MWR parameterization and landslide location and size.

For all three run options, each model iteration begins with the same initial topography. After Np model simulations, Np different versions of the post-runout landscape are created and, probability of runout at each node is determined as:

$$P(\Delta\eta) = \frac{\#(|\Delta\eta|>0)}{Np} \tag{30}$$

where $\#(|\Delta\eta| > 0)$ is the number of times topographic elevation at a node changes as a result of erosion or deposition from the Np model runs. Probability of erosion or aggradation can be determined by replacing the numerator in (30) with $\#(\Delta\eta < 0)$ or $\#(\Delta\eta > 0)$ respectively.

4. Basic model behavior

We evaluate basic model behavior using a series of virtual experiments. The virtual experiments consist of six synthetic terrains including: (A) a planar slope that intersects a gently sloped plane (S = 0.001), (B) a planer slope with a constriction, that intersects a gently-sloped plane, (C) a planar slope that has a bench mid-slope and then intersects a gently-sloped plane; (D) a concave up, uniform-convergence slope; (E) a concave up, variable-convergence slope that widens (convergence decreases) in the downslope direction; (F) a convex up, variable-convergence slope that widens (convergence decreases) in the downslope direction. On each terrain, a 30-meter wide, 50-meter long and 3-meter deep landslide is released from the top of the terrain. All six terrains are covered by a 1-meter thick regolith and use the same parameter values (S_c = 0.03, q_c = 0.2 m, k = 0.01, D_s = 0.2 m). Each terrain is represented using a 10-m grid. Experiment results are shown in Figure 5.

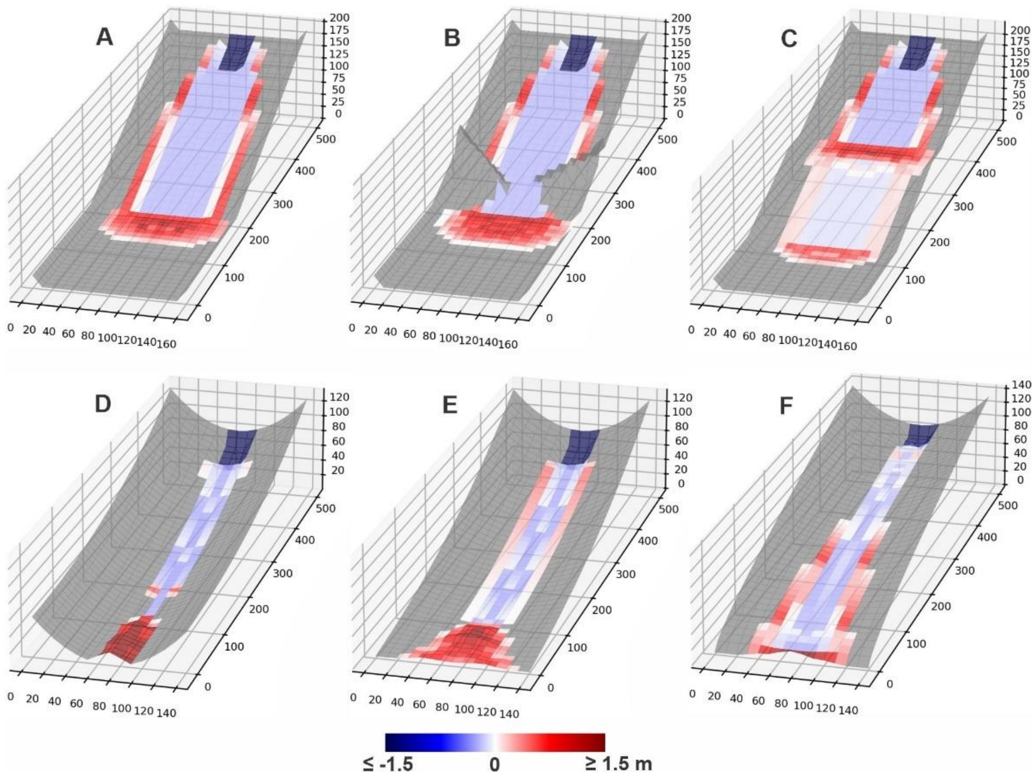

Figure 5. Shaded, 3-D visualizations of model response to six different synthetic terrains, colored according to the DoD of the final runout surface. Shading is to scale. Red indicates a positive change in the elevation of the terrain (aggradation) and blue indicates a negative change (erosion). The 3-D visualization of the DoD is exaggerated by a factor of 5 to make visible in figure. Grid size is 10 meters.

On Terrain A, the landslide spread as it moved downslope and formed levees along the edge of the runout path. The width of the spread was a function of the multiple flow direction algorithm and resistance along lateral margins of the runout as represented by q_c. At the slope break at the base of the slope, the material deposited at an angle controlled by S_c. On Terrain B, the flow initially eroded and deposited identical to the first but near the slope break, the topographic constriction forced flow depth to increase and exceed q_c, minimizing the formation of levees (because $q_O > q_c$) and resulted in a slightly larger deposit at the base of the slope. On Terrain C, landslide runout was again initially identical to the runout on Terrain A; however, upon intersecting the mid-slope bench, most of the runout material deposited. A small, thinner portion did continue past the bench but eroded at a lower rate than the initial slide upslope of the bench. Upon intersecting the flat surface at the base of the hillslope, the runout material deposited.

On Terrain D, the landslide and its runout were confined to the center of convergent terrain and only deposited once the slope was less than S_c. The slide never widened because the uniformly convergent channel shape prevented spreading and the narrower flow width maintained a higher flow depth, which prevented the formation of levees. On Terrain E, the landslide again deposited once slope was less than S_c but because topographic convergence of Terrain E decreases in the downslope direction, as the runout material moved downslope, the deposit spread more than on Terrain D, which caused thinner flow and deposition along margins of the runout path. On the final terrain, Terrain F, slope is always greater than S_c so deposition was limited to levees along the edge of the flow that formed as the runout spread in response to decreasing convergence.

MWR model behavior can be summarized as follows. The displacement and deposition of landslide material predicted by MWR responds to topography in a reasonable manner: Flow width increases as convergence decreases (e.g, Terrain F), which in turn reduces flow depth. Lower flow depths cause lower erosion rates and reduce aggradation extent. Conversely, modelled flow depth increases when convergence increases (e.g., Terrain B). Where the flow encounters broadly convergent or planer slopes, lateral levee deposits form, a common feature of landslides reported in the literature and at sites reported here (see Section 5) that detailed mechanistic models can struggle to reproduce (e.g., Barnhart et al, 2021).

We did not attempt to compare MWR modelled flow with the output of shallow-water-equation based models or observed granular flows (e.g., Medina et al, 2008; McDougall and Hungr, 2004; Iverson and Denlinger, 2001; Han et al., 2015). The cellular automaton representation in MWR does not model the time-dependent evolution of debris flow velocity and depth, and conceptually moves debris instantaneously at each iteration, as driven by changes in the evolving topographic elevation field. Because of that, only the final outcome (modelled runout extent, sediment transport and topographic change) of MWR can be compared with other models or observed runout, which we do in the next section. Also, as described in Section 2.3, behaviour of the multiple flow direction algorithm does vary with grid size. Using a coarser or finer grid, without adjusting model parametrization, could potentially change how wide the landslide spreads.

5. Model Validation:

5.1 Overview

In this section, we demonstrate the ability of a calibrated MWR to replicate observed runout extent, sediment transport and topographic change at field sites located in the western USA and summarize model calibration results with an evaluation of MWR calibration relative to terrain attributes of the observed runout paths. Note that simply calibrating a model to match field data does not constitute a satisfactory test of model predictive ability (Iverson, 2003). Strategic testing, which involves calibrating the model to one site or period of time and then running the calibrated model at a separate site or period of time (Murray, 2013), is a better indicator. Two of our validation sites, the Cascade Mountain and Olympic Mountain sites, include two separate landslides and subsequent runout and we test model predictive ability at these sites in Section 6.

Calibrated model performance is demonstrated at the following field sites (see Figure 6a for locations and observed runout extent): (1) two runout events over the same hillslope in the Cascade Mountains (Washington state [WA], USA): a large debris avalanche in 2009 (Cascade Mountains, 2009) and a moderately-sized debris flow in 2022 (Cascade Mountains, 2022) that inundated and flowed within a first-to-second order channel until perpendicularly intersecting a narrow river valley several hundred meters below the landslide (Figure 1a); (2) debris flows in the Black Hills (WA) sourced from a small failure along the toe of a deep-seated landslide (Black Hills, South) and a moderately-sized debris avalanche from a large road fill (Black Hills, North) that flowed several kilometers along a relatively wide, broadly convergent channel before stopping (Figure 1b); (3) a single, moderately-sized debris avalanche in the Rocky Mountains (Rocky Mountains), the majority of which flowed several hundred meters over a broadly convergent to divergent hillslope in Colorado (Figure 1c); and (4) a 30-year chronology of small landslides and subsequent debris flows in the Olympic Mountains (WA) in steep, highly convergent channels that flowed well over a kilometer and coalesced into a single runout deposit in a dendritic, channelized watershed (Olympic Mountains; Figure 1d). All landslides initiated during heavy rainfall or rain-plus-snowmelt storm events (WRCC, 2022; NRCS, 2022; Table 1) but their runout varied in terms of erosion rate, grain size (Figure 6b), depositional behavior (Figure 6c) and the topographic convergence of the underlying terrain.

Table 1. Landslide and runout characteristics

site	Cascade Mountains, 09	Cascade Mountains, 22	Black Hills, south	Black Hills, north	Rocky Mountains	Olympic Mountains
initial landslide body length [m]	185	55	80	75	40	45
initial landslide body width [m]	80	50	15	65	35	15
initial landslide body volume [m^3]	110,000	22,000	1,500	18,500	4,600	400 - 2,200
2-day cumulative precipitation + snowmelt [mm]	120+85	140+75	205+50	205+50	193+0	100 - 220 + ?
maximum grain size [m]	0.316	0.316	0.48	0.206	0.984	0.8
slope range of positive-net deposition [%]	1 - 15	1 - 15	<1 - 10	<1 - 8	16 - 25	5 - 15
average flow depth in scour zone [m] [a]	4	2	2	3	3	3
average channel slope in scour zone [m/m]	0.25	0.25	0.15	0.15	0.4	0.3
average channel width in scour zone [m]	45	20	25	35	55	10
length of erosion, [m]	600	340	1210	1345	360	2550
erosion area, $\mathbb{A}$ [m^2]	28,400	6,600	22,800	52,400	20,800	28,900
erosion volume, $\sum E\Delta x^2$ [m^3] [b]	44,547	5,125	12,332	26,815	34,275	33,725
average erosion per unit length of runout debris, $\bar{E}/\ell$, [m/m]	0.0085	0.014	0.0068	0.0068	0.041	0.026
k	0.020	0.034	0.017	0.020	0.076	0.051
growth factor, [m^3/m]	74.2	15.1	10.2	19.9	95.2	13.2
average observed $\lvert\Delta\eta\rvert$ [m]	2.4	2.2	0.53	0.63	0.89	1.4
total erosion volume / total mobilized volume [c]	0.29	0.19	0.89	0.59	0.88	0.97

[a] rough approximation based on landslide volume, channel width and height of scour marks in erosion zone
[b] excludes landslide volume
[c] total moblized volume = initial landslide body + erosion volume

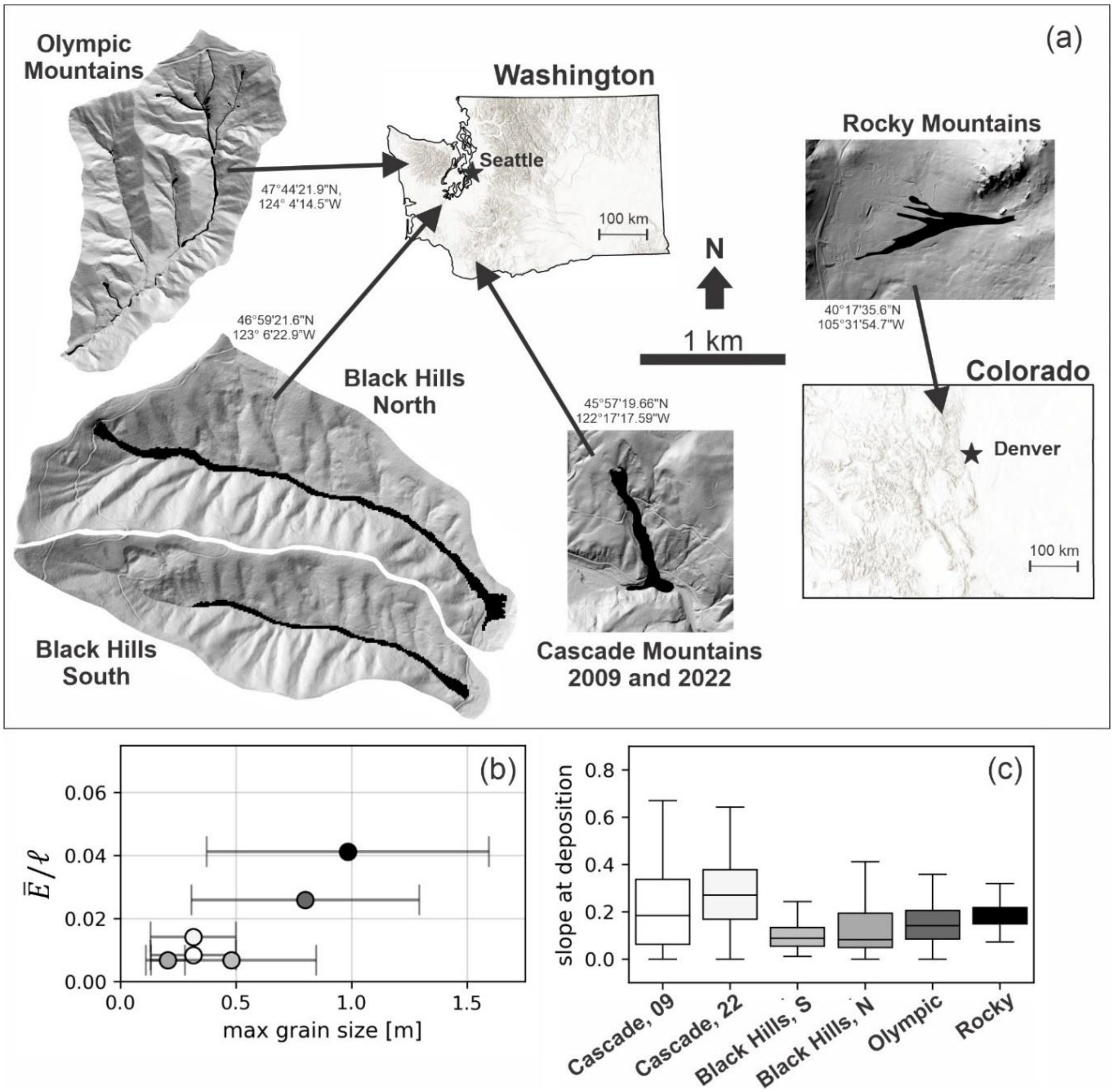

Figure 6 (**a**) Landslide locations in Washington and Colorado states. Coordinates next to each site are WGS84. Shaded DEMs of each site are shown at the same scale. (**b**) Observed average erosion rate per unit landslide length ($\bar{E}/\ell$) relative to the observed average-maximum grainsize. Error bars indicate standard deviation. (**c**) Underlying topographic slope of observed deposition.

5.2 Model setup and field parameterization

Each model was set up on a 10-meter grid representation of the pre-event DEM. The extent of the initial mass wasting source material (e.g., the initial landslide body) was interpreted from a combination of lidar, air-photo and field observations. At all locations, we use Equation (13) to approximate shear stress. We field-surveyed each site, noting the maximum flow thickness, typical deposition and erosion depths and the size of the largest grains in the runout deposits.

We estimated parameter values from these field and remote observations (See Table 1). A site-specific value for k was determined as a function of the observed average erosion depth (determined as total erosion volume divided by the erosion area, $\bar{E}$) relative to the length of the runout debris, which we approximated as the length of the initial landslide body(ℓ). Further details are described in the Appendix.

The volume of the initial landslide body ranged in volume from 400 to 110,000 m^3 across sites. At all sites, erosion and subsequent entrainment added to the total mobilized volume (initial landslide body + erosion volume), but the contribution was highly variable. The erosion volume divided by the total mobilized volume was as low as 0.19 at the Cascade Mountain, 2022 landslide to as high as 0.97 at the Olympic Mountain landslides (Table 1).

The average maximum grain size varied from 0.2 m at the Black hills sites to nearly 1 m at the Rocky Mountain Site (Figure 6b, Table 1). Values of $\bar{E}/\ell$ ranged from 0.007 to 0.041 [m/m] with the highest rate occurring at the Rocky Mountain landslide and the lowest at the Black Hills sites. Details on grain-size samples and data collected in the field are described in the Supplementary Material. In terms of growth factors (average volumetric erosion per unit length of the erosion-dominated region of the runout path, Hungr et al. 1984; Reid et al., 2016) values ranged from 10 m^3/m at the Black Hills South site to 95 m^3/m during the Rocky Mountain landslide (Table 1).

The median values of topographic slopes at which observed deposition occurred (i.e., $\Delta\eta > 0$) ranged between 0.1 and 0.3 across sites, while deposition was also observed in much steeper (>0.4) slopes, and much flatter slopes at some sites (Figure 6c) (Table 1). The slope of channel reaches where net deposition (cumulative erosion and deposition; e.g., Guthrie et al., 2010) was positive tended to be lowest at the Black Hills site (<1% to 10%) and highest at Rocky Mountain site (16% to 25%).

We defined uniform prior distributions of S_c and q_c and then used the calibration utility to find the best-fit parameter values (parameter values corresponding to the highest $L(\Lambda)$). Minimum and maximum values of S_c were initially estimated from the range of observed slope of areas of positive-net deposition (Table 1). Minimum and maximum values of q_c were set as 0.01 to 1.75, which roughly represents the range of minimum observed thickness of debris flow termini in the field at all of the validation sites. For the purpose of implementing the calibration utility, we prepared a DoD of each site. The DoD was determined either form repeated lidar or field observations as detailed in the Supplementary Material.

5.3 Calibration and model performance

Markov chains, colored according to the likelihood index, $L(\Lambda)$ are plotted in the S_c- q_c domain, along with histograms of sampled S_c and q_c values for each landslide in Figure 7. Each Markov chain includes 2000 model iterations. The runtime for 2000 model iterations depended on model domain, landslide size and number of landslides modeled but varied from roughly 1.5 for the Cascade, 2022 landslide to 6 hours for the Olympic Mountain landslides on a 2016 2.1 GHz Intel Core Xeon, 32 GB memory desktop. The chains show a wide array of sampling patterns and parameter ranges but broadly speaking, at all sites, the algorithm jumped within S_c- q_c space towards higher $L(\Lambda)$, to form bell-shaped posterior distributions for each parameter. Depending on the landslide type, the calibration algorithm converged on different S_c- q_c pairs. For example, at the Cascade Mountains site, the calibration utility converged to smaller q_c and S_c values for the 2009 event (Figure 7a), which permitted thinner flows over lower slopes and effectively made the 2009 modelled runout more mobile relative to the 2022 modelled runout (Figure 7b). At the Rocky Mountains site (Figure 7e), a relatively high q_c value helps control lateral extent of the modelled runout that in the field was controlled by standing trees (see Supplementary Material).

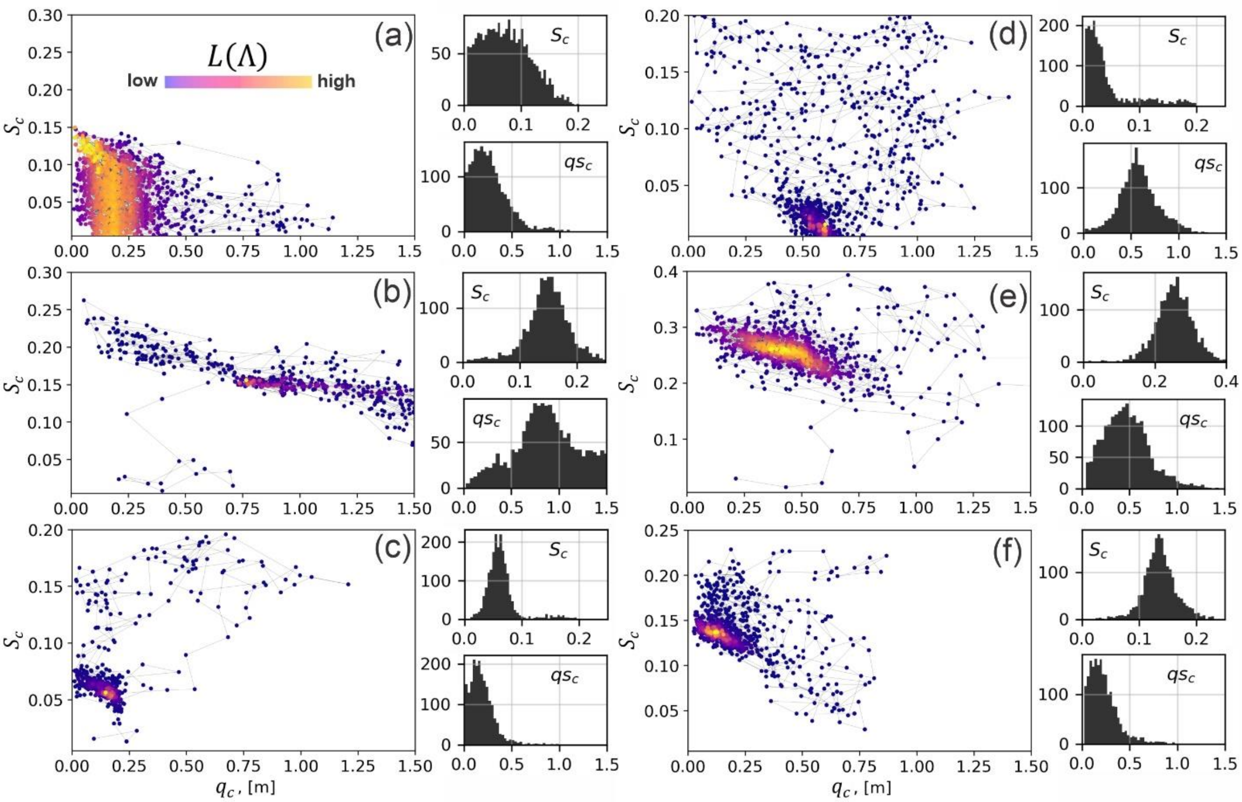

Figure 7. MWR calibration results for (a) Cascade Mountains, 2009; (b) Cascade Mountains, 2022, (c) Black Hills, South; (d) Black Hills, North; (e) Rocky Mountains and; (f) Olympic Mountains. Each result shows a scatter plot of the sampled S_c and q_c values, colored by their relative $L(\Lambda)$ value. Note y-axis scale differs between plots. To the right of each scatter plot are histograms of the iterated S_c and q_c parameters, which can be normalized to represent an empirical PDF of the possible S_c and q_c values that calibrate MWR to the site. Histogram y-axis is count and x-axis is S_c or q_c, as indicated on the histogram.

Profile plots of modelled Q_s and maps of the modelled planimetric runout extent, colored to indicate where the runout matched (α), overestimated (β) or underestimated (γ) the observed runout are shown in Figure 8. Values of Ω_T we obtained with MWR are comparable or higher than reported values of Ω_T in the literature that used a variety of models (Gorr et al., 2022; Barnhart et al., 2021; Note, to compare Ω_T values to those studies, subtract 1 from values reported in this study). Across the sites, the volumetric error of the model, $\Delta\eta_E$, ranges between 6% and 15% (median 9.1%) of the total mobilized volume from the observed DoD. An overall <10% volumetric error is reasonable considering the low number of parameters required to calibrate MWR and that empirical estimates of total mobilized volume used to run other runout models can vary by as much of an order of magnitude (e.g., Gartner et al., 2014: Barnhart et al., 2021). Model performance in predicting volume flux along the runout profile was within similar error ranges. Except for the Rocky Mountains site where MWR consistently modelled wider-than-observed flow, the cumulative flow error along the runout profile (Q_{s_E}) were limited to 5%-19% of the mean cumulative flow determined from the observed DoD.

MWR generally successfully replicates observed sediment transport along the runout path via model parameterizations that are unique to each landslide. For example, the profile plots of Q_s at the Cascade Mountain site (Figure 8a and 8b) show that during the 2009 landslide, all of the runout material flowed past the first 750 meters of the runout path. During the 2022 landslide, material began to deposit just down slope of the initial landslide scar, as both observed and modelled Q_s reverse slope, indicating loss in downstream volume flux. Model comparisons in the Cascade Mountains site were limited to the upper 750 m of the hillslope because a large portion of the runout material was lost to fluvial erosion in the valley (see Supplementary Material).

MWR also successfully replicates the observed sediment transport patterns at the Olympic Mountains site (profile plot of Q_s in Figure 8f) and to a lesser degree, the Rocky Mountain site (Figure 8e). This finding is notable, because at the Olympic Mountain site, observed runout extent and sediment depositional pattern were heavily impacted by woody debris. Similarly, at the Rocky Mountains site, the width of the runout appeared to be restricted by trees. (See Supplementary Material).

Using a fixed cell size of 10-m might have impacted model performance in some areas. MWR tended to over-estimate the runout width for small landslides like the Olympic Mountains and Cascade Mountains, 2022 sites (yellow zones in Figures 8f and 8b), likely because of the 10-m grid size used to represent the terrain. A 10-m DEM is generally accepted as a good balance between model detail and computational limitations (e.g., Horton et al. 2013). However, for small landslides, the 10-m grid is close to the size of the channels that controlled observed runout (Supplementary Material) and may not have accurately represented the terrain. Modelled flow was less topographically-constrained and tended to flow over a wider area of the terrain than observed in the more confined and smaller channels within the axis of the runout valleys.

Because MWR does not have an explicit representation of flow momentum, it may show poor performance in regions of the runout path where flow momentum is the primary control on runout extent. For example, at the Cascade Mountain, 2009 slide, MWR underestimates the slope-perpendicular flow over a bench (large red zone in Figure 8a). Review of model behavior for this slide (Figure 9) shows how MWR successfully mimics diverging flow around a broad ridge upslope of the bench (iteration t=28 in Figure 9), but afterwards continues to follow topographic slope and converges too rapidly into a narrow ravine along the west edge of the bench (iteration t=40 in Figure 9; compare to runout scar in air photo and underestimated region on topographic bench in Figure 8a). At the Rocky Mountains site, in addition to standing trees, the forward momentum of the runout may have also restricted lateral spread of the observed runout. Modelled runout is consistently too wide.

Overall, calibration was best at the Cascade Mountain, 2009 landslide (values of Ω_T are highest and values of $\Delta\eta_E$ and Q_{s_E} are lowest) and poorest at the Rocky Mountain and Olympic Mountain sites (Values of Ω_T are lowest Q_{s_E} and $\Delta\eta_E$ are highest). At both the Rocky Mountain and Olympic Mountain sites, because we lacked repeat lidar, we created the DoD from a map of field estimated erosion and deposition depths and estimated the pre-event DEM. The lower calibration scores may indicate that field estimated DoDs were not as accurate as those determined via lidar differencing. Another source of uncertainty that we have not addressed in our study is regolith thickness. Using spatially accurate regolith thickness, rather than a uniform thickness, would likely improve MWR performance too. Nonetheless, although imperfect, at most sites, MWR does not appear to have a strong systematic bias in modeled output, which suggests that MWR may not have any structural weaknesses; however the consistent over-estimated width on planar to divergent topography at the Rocky Mountain site requires further investigation at similar sites to determine if this issue is due to calibration or the model.

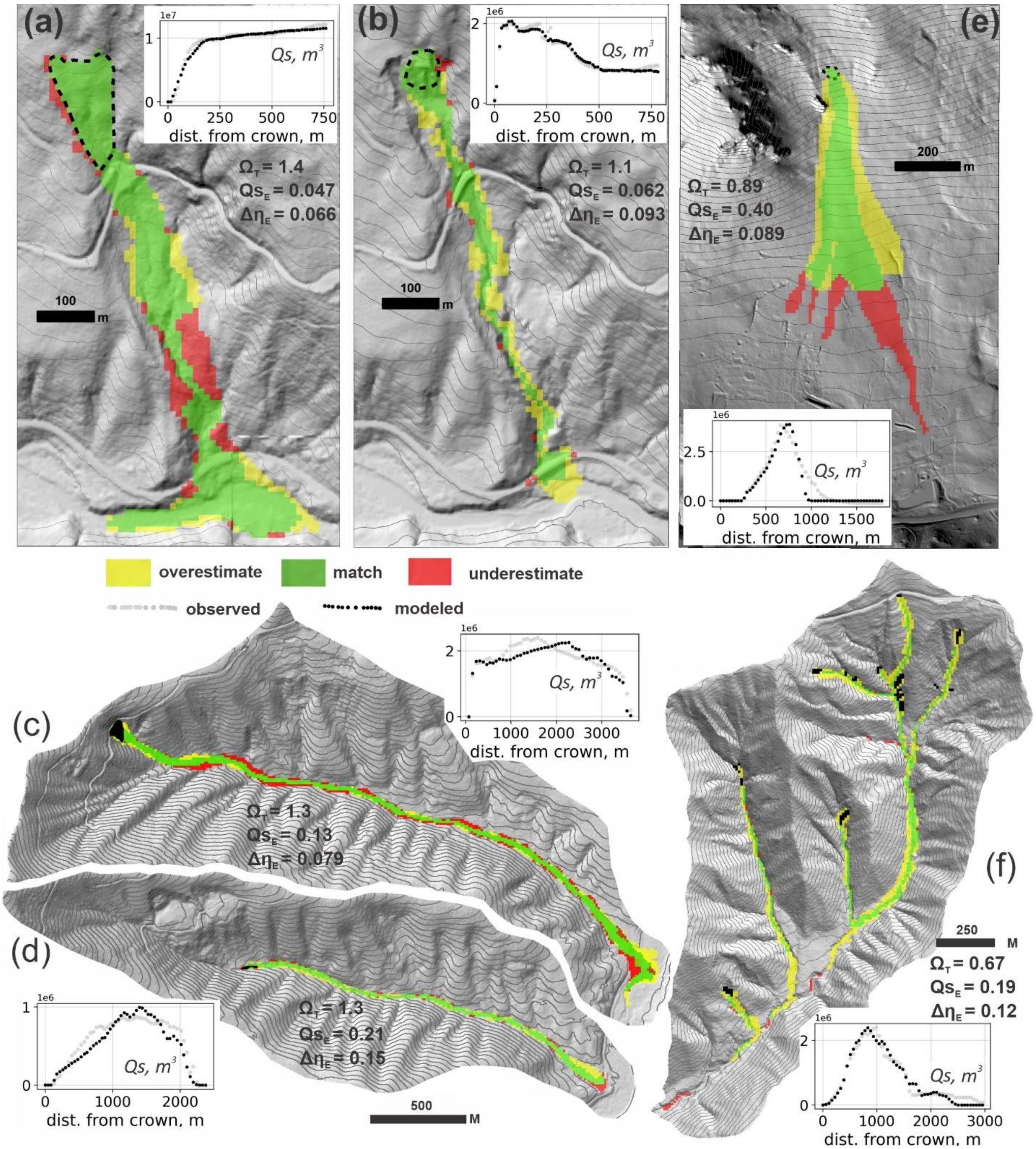

Figure 8. Calibrated model performance as indicated by maps of modeled runout extent, profile plots of observed and modeled cumulative sediment transport along the centerline of the runout path (Q_S, see equation 28) and

reported values of Ω_T, $\Delta\eta_E$ and Q_{s_E}. Y-axis label for profile plots of Q_s indicated on plot. In all maps, up is north except in (e), north is towards the left. (a) Cascade Mountains, 2009; (b) Cascade Mountains, 2022; (c) Black Hills, North; (d) Black Hills, South; (e) Rocky Mountains; (f) Olympic Mountains.

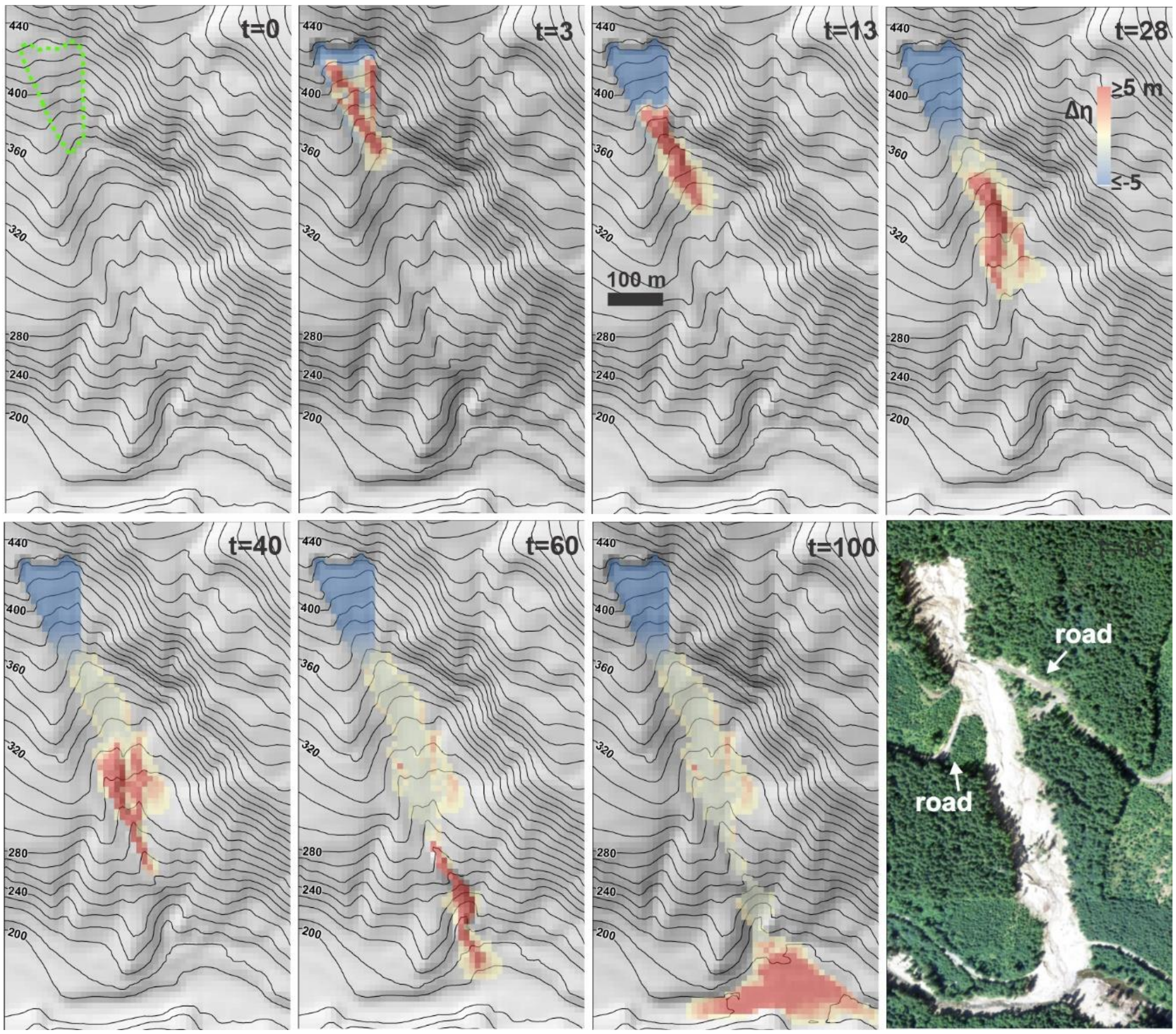

Figure 9. Illustration of modeled runout of the Cascade Mountains, 2009 landslide beginning from the initial movement of the landslide body to final deposition in the river valley that demonstrates MWR response to topography. Note how the landslide slip surface directs the initial flow. Topography lines reflect the underlying terrain, which is updated after each iteration. Air photo in last panel shows observed runout extent. Note that upper road is not part of the observed landslide runout path.

To understand whether the ability to calibrate MWR systematically varies with topography of the runout path, we compared model performance with three topographic indices described by Chen & Yu (2011). The indices are computed from the terrain in the observed runout extent and include the relief ratio (H/L), mean total curvature (κ) and the mean specific stream power index (SPI). The index H/L equals the average slope of the runout path (or relative relief), determined as the total topographic relief of the runout (measured from the center of the landslide to the end of the runout path) divided by the horizontal length of the runout and indicates the mobility of the runout. Index κ represents topographic convergence, which is the second derivative of the terrain surface, with increasingly positive values of index κ reflecting growing topographic convergence and concave-up channel profile (e.g., Istanbulluoglu et al., 2008). The index SPI is determined as the natural log of the product of the contributing area and slope. Indices κ and SPI are computed at each node in the runout extent and model performance is compared to the mean value.

Comparison of model performance with respect to the topographic indices in Figure 10 shows: slightly improved model performance over runout paths that are less convergent (lower SPI and κ values) and on steeper terrain (higher H/L) but neither trend is significant. The latter finding appears to be mostly a result of how well modelled sediment transport and topographic change (Q_{s_E} and $\Delta\eta_E$) replicated observed, as there does not appear to be a trend in Ω_T with H/L and the two best performing models (both Cascade Mountain landslides) had the lowest (best) Q_{s_E} values and low $\Delta\eta_E$ values. Both findings are likely impacted by the grid size we used to represent terrain. As noted above, at all sites we used a 10-m grid, but at some sites 10-m doesn't quite capture the relief of channelized topography that controlled observed runout, leading to modelled runout that was considerably wider than observed and causing low Ω_T value (this is especially true at the Olympic Mountains site, Figure 10a, b and c). Also, it is important to note that these indices were calculated for the extent of the observed debris flows and may not represent the topographic form that controlled the model.

In summary, using the calibration utility, we showed how MWR can be calibrated to a range of different landslide types and runout terrains. To a certain degree, though calibration, MWR can be parameterized to compensate for deficiencies in the DEM or processes not explicitly represented in the model (momentum, woody debris). A relationship between model performance and topography was not eminent. This finding is likely a result of the contributions of numerous factors other than the terrain form, such as the DEM resolution, the quality of the DoD and importance of processes not explicitly included in the model that also impact performance.

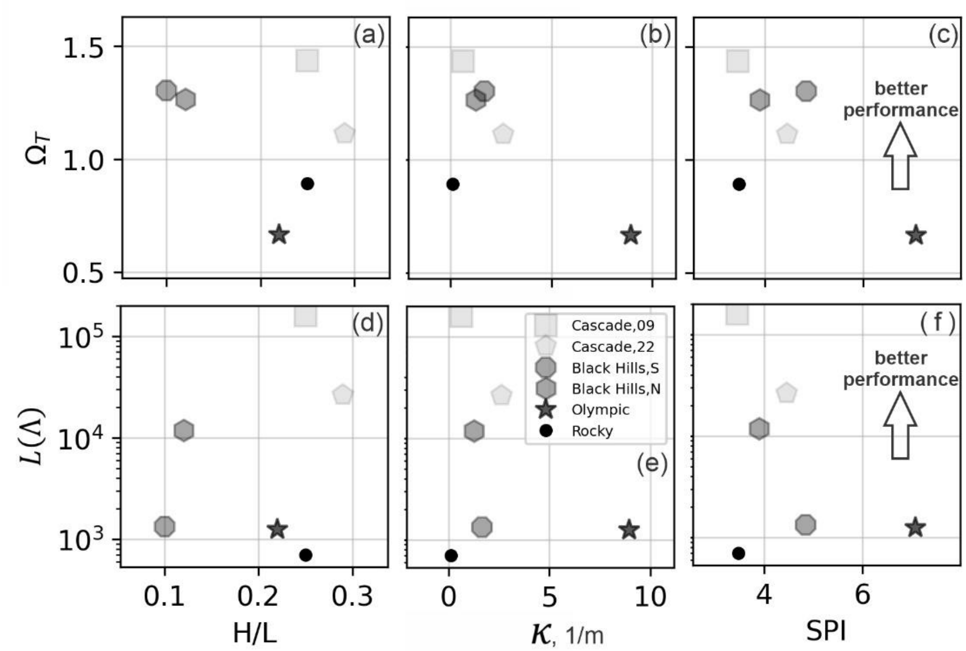

Figure 10. Illustration of model calibration, as reflected by the posterior parameter likelihood $L(\theta)$ and planimetric fit (Ω_T) relative to topographic indices. There is no strong trend between the topographic indices and calibration performance. Note, curvature values are scaled by a factor of 100.

6. Discussion

6.1 Strategic testing of MWR for hazard mapping applications

Having demonstrated basic model response to topography and that MWR can be calibrated to a variety of landslides and runout terrains, we now strategically test MWR using the Cascade Mountain and Black Hills sites. Since both of these sites include two separate landslides, we can thus test model performance by swapping best-fit model parameters at each site, rerunning the models and comparing results with the original, calibrated results. At the Cascade Mountain site, the 2009 and 2022 landslides originated on the same hillslope (Figure 8a and 8b). At Black Hills site, the two landslides occurred on different hillslopes but in adjacent east-west oriented watersheds (Figure 8c and 8d).

As shown in Figure 11, at three of the landslides (both Cascade Mountain landslides and the Black Hills, North landslide), when the best-fit parameters from the other landslide are used to predict runout, the accuracy of modelled runout planimetric extent drops but resultant Ω_T values can still be as high or higher than values reported in other studies (compare to equivalent Ω_T values in Gorr et al., 2022 and Barnhart et al., 2021). In terms of modelled sediment transport and topographic change, swapping best-fit parameters has a more substantial effect. At the Cascade Mountain, 2009 landslide, using the 2022 best-fit parameter values causes about half of the modelled runout material to prematurely deposit on the hillslope, reducing the amount of sediment that reaches the valley floor (Q_{s_E} increases by a factor of nine; Figure 11). Using the Cascade Mountain, 2009 parameter values on the Cascade Mountain, 2022 landslide (Figure 11b) increases modelled runout extent and results in nearly four times the entrainment and transport of sediment to the valley floor, causing Q_{s_E} to increase by a factor of 20 and $\Delta\eta_E$ by 83%. At the Black Hills site, using the South basin best-fit model parameters at the North basin causes Q_{s_E} and $\Delta\eta_E$ increase by 83% and 39% respectively (Figure 11c). Unlike the other three landslides, swapping best-fit parameters at the Black Hills, South landslide results in both large sediment transport and runout extent error because the North basin best-fit parameters cause modelled landslide to entrain too little and stop only a few hundred meters from the initial source area (Figure 11d).

Although the need for calibration of MWR is a limitation for its transferability across sites, this limitation holds true for most physics-based models. Barnhart et al. (2021) compared the ability of three different detailed-mechanistic models to replicate an observed post-wildfire debris-flow runout event in California, USA. All three models used a shallow-water-equation-based approach that conserved both mass and momentum, representing the flow as either a single phase or double phase fluid. All models gave comparable results in simulating the event, suggesting that there may not be a “true” best model. Despite the high level of detail and processes explicitly included in each model, all models were sensitive to and required an estimate of the total mobilized volume, and the ability to replicate observed runout ultimately depended on calibration of the parameters used to characterize debris flow properties.

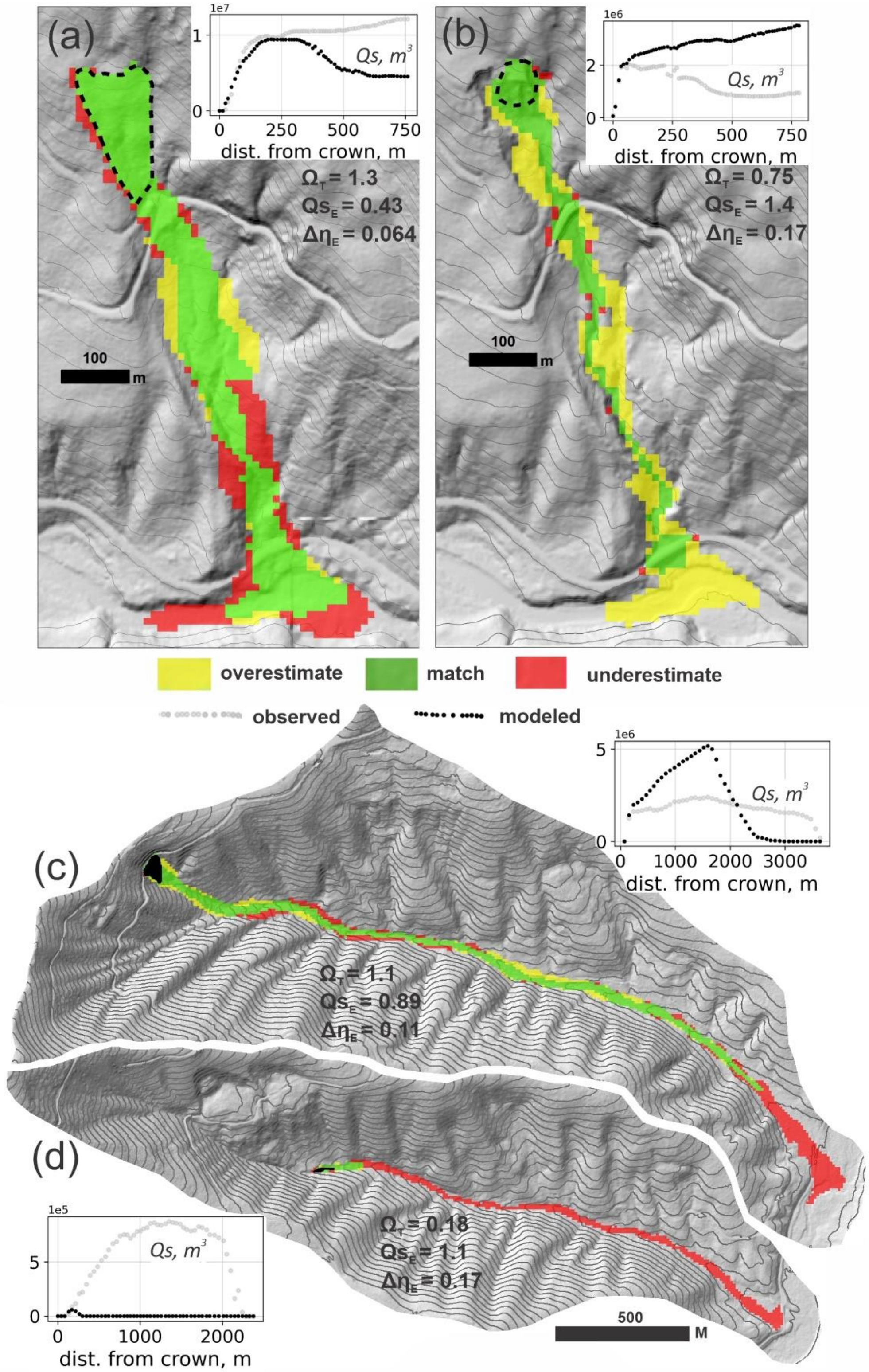

Figure 11. Model performance using the neighboring landslide parameter values, as indicated by modeled runout extent, profile plots of Q_s, reported values of Ω_T, $\Delta\eta_E$ and Q_{s_E} . Compare with Figure 8. (a) Cascade Mountain, 2009; (b) Cascade Mountain, 2022; (c) Black Hills, North; (d) Black Hills, South

As landslide hazard models often forecast hazard probabilistically, an alternative test to simply swapping the best-fit parameters is to swap parameter PDFs determined from the calibration utility and compare probability of runout at each model node (equation 30). As shown in Figure 12, similar to the first test, at three of the landslides, using the parameter distribution associated with the neighbouring landslide results in relatively minor changes in whether runout is likely to occur versus not occur (probability of runout ≥50%; Figures 12a, 12b and 12d). At the Black Hills South landslide, swapping parameter PDFs causes a large change in runout probability (Figure 12c).

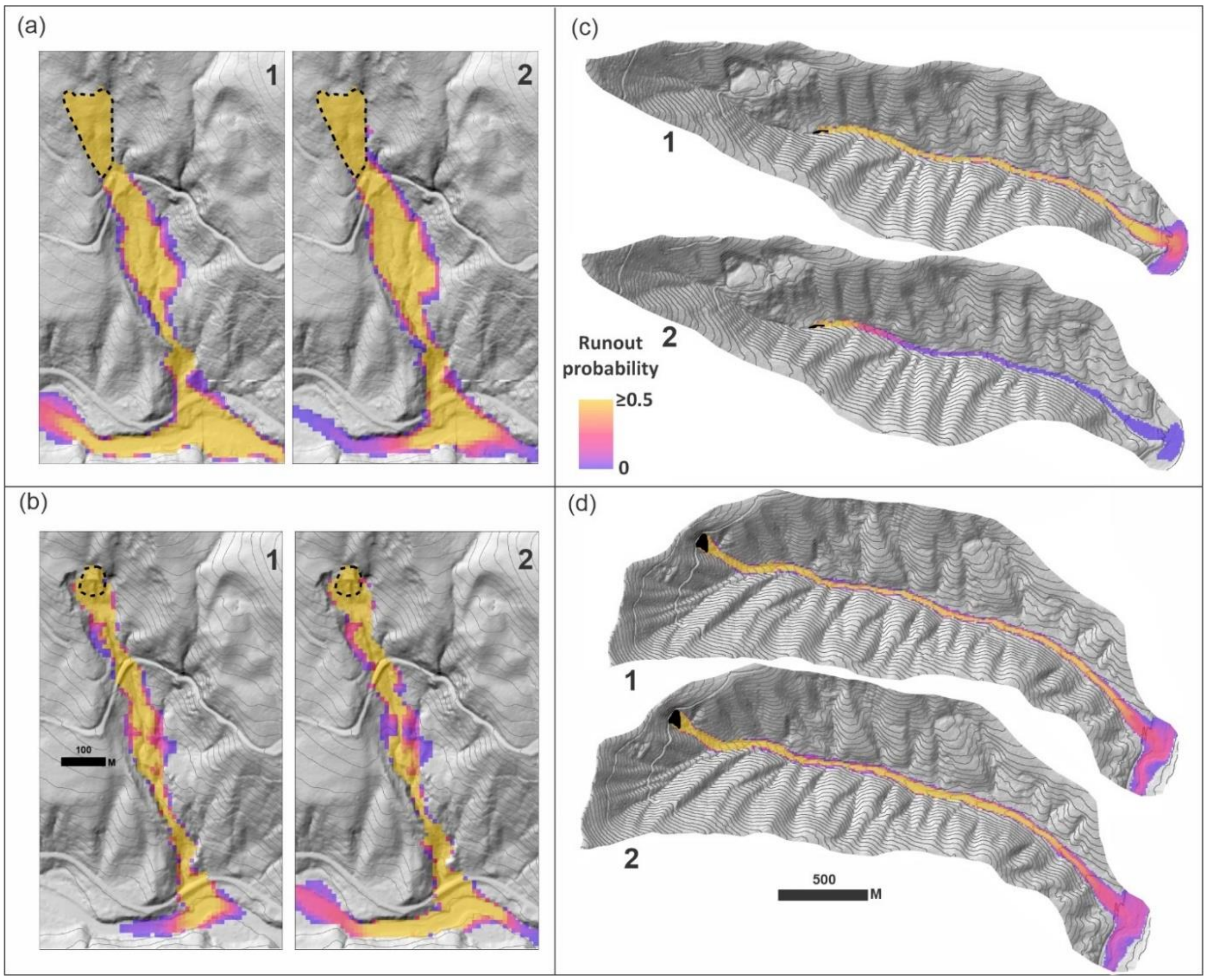

Figure 12. Model tests by swapping parameter PDFs and comparing runout probability at the (a) Cascade Mountain, 2009; (b) Cascade Mountain, 2022; (c) Black Hills, South and; (d) Black Hills, North sites. (**1**) runout using parameter distributions of the site and (**2**) runout using parameter distributions of the neighboring site.

The results of these two tests suggest that site-specific calibration may be needed when the user aims to apply MWR to sediment budget analyses; however, we suspect that this finding is a consequence of testing the model at a site with very different landslide types and runout processes. At sites like the Cascade Mountain and Black Hills sites, which consisted of a diverse range of landslide processes including small, confined debris flows to large, unconfined debris avalanches, MWR may need to be calibrated to each type of landslide and predictive applications might involve applying the appropriate parameter set based on landslide type. In regions where landslide processes are relatively uniform (like the Olympic Mountain site), calibration to one landslide might be sufficient to predict the depositional patterns and sediment transport at another. Finally, as noted in Section 3.1, we found numerous parameter combinations allowed MWR to match observed runout extent. This finding suggests that if the project aim is limited to an evaluation of runout extent, model calibration to the site may not be as critical and parameter values from calibration to nearby landslides or even globally-available repeated DEMs and airphotos that show the slope of past landslide deposits (for S_c) and how thick their frontal lobes are at the point of deposition (for q_c), might be sufficient.

6.2. MassWastingRunout probability applications

In this section we briefly demonstrate how to determine runout probability from a probabilistically determined landslide hazard map or a specific, potentially unstable slope using MWR. The first application may be appropriate for watershed- to regional-scale runout hazard assessments. The second application is an example hazard assessment for a potentially unstable hillslope. Both applications are demonstrated at the Olympic Mountain site where landslide size and type tended to be relatively uniform and parameter PDFs determined through calibration may therefore represent typical runout processes in the basin.

6.2.1. Runout probability from a landslide hazard map

To determine runout probability from a landslide hazard map, we ran MWR Probability using Option 3, reading a series of mapped landslide source areas created by an externally run Monte Carlo landslide initiation model. For the landslide initiation model, we used LandslideProbability, an existing component in Landlab that computes landslide probability by iteratively calculating Factor-of-Safety (FS: ratio of the resisting to the driving forces) at each node on the raster model grid Np times from randomly selected soil (regolith) hydrology properties (e.g., soil depth, saturated hydraulic conductivity) soil strength (friction angle, cohesion) and recharge rates (precipitation input rate minus evapotranspiration and soil storage). Landslide probability at a node is defined as the number of times FS<1 divided by Np.

We first ran LandslideProbability using a 50-year precipitation event (WRCC, 2017) to determine landslide probability (Figure 13a) over the entire Olympic Mountains model domain and create the series of Np FS maps. Details on the LandslideProbaility setup are included in the Supplementary Material. We then read the series of FS maps into MWR Probability, treating all nodes with $FS < 1$ as a landslide source, and ran MWR Np times. Each iteration, MWR read a new FS map and randomly selected a new set of parameter values from S_c- q_c parameter PDFs created by the calibration utility.

Runout probability results are illustrated in Figure 13b and show that the probability of runout is high in many of the second order channels but low at the basin outlet. As discussed in Section 3, the probability of aggradation or erosion caused by the runout can also be determined by adjusting the numerator of Eq. (30) and the probability of deposition greater than 1 meter is shown in Figure 13c. In this example, in addition to MWR parameter uncertainty, runout probability reflects uncertainty in landslide size and location caused by a 50-year precipitation event.

6.2.2 Runout probability for a specific, potentially unstable slope

When field evidence or other data indicate that a specific hillslope may be potentially unstable, but the exact area of a potential landslide on that slope is unknown, MWR can be used to generate a hazard estimate that takes into account the uncertainty in the landslide area. For this application, MWR Probability is run using Option (2), which requires a polygon representing the extent of the potentially unstable slope. We designated a 0.6 ha, convergent hillslope in the headwaters of the Olympic Mountains site as a potentially unstable slope (Figure 13d). For each model repetition, a landslide area can form anywhere within the potentially unstable slope and is at least as large as a user defined minimum size but no larger than the potentially unstable slope. This example shows that, given uncertainty in the landslide size and location, and uncertainty in MWR parameterization, if a landslide were to initiate on the potentially unstable slope, the probability of the runout reaching the basin outlet is less than 5%.

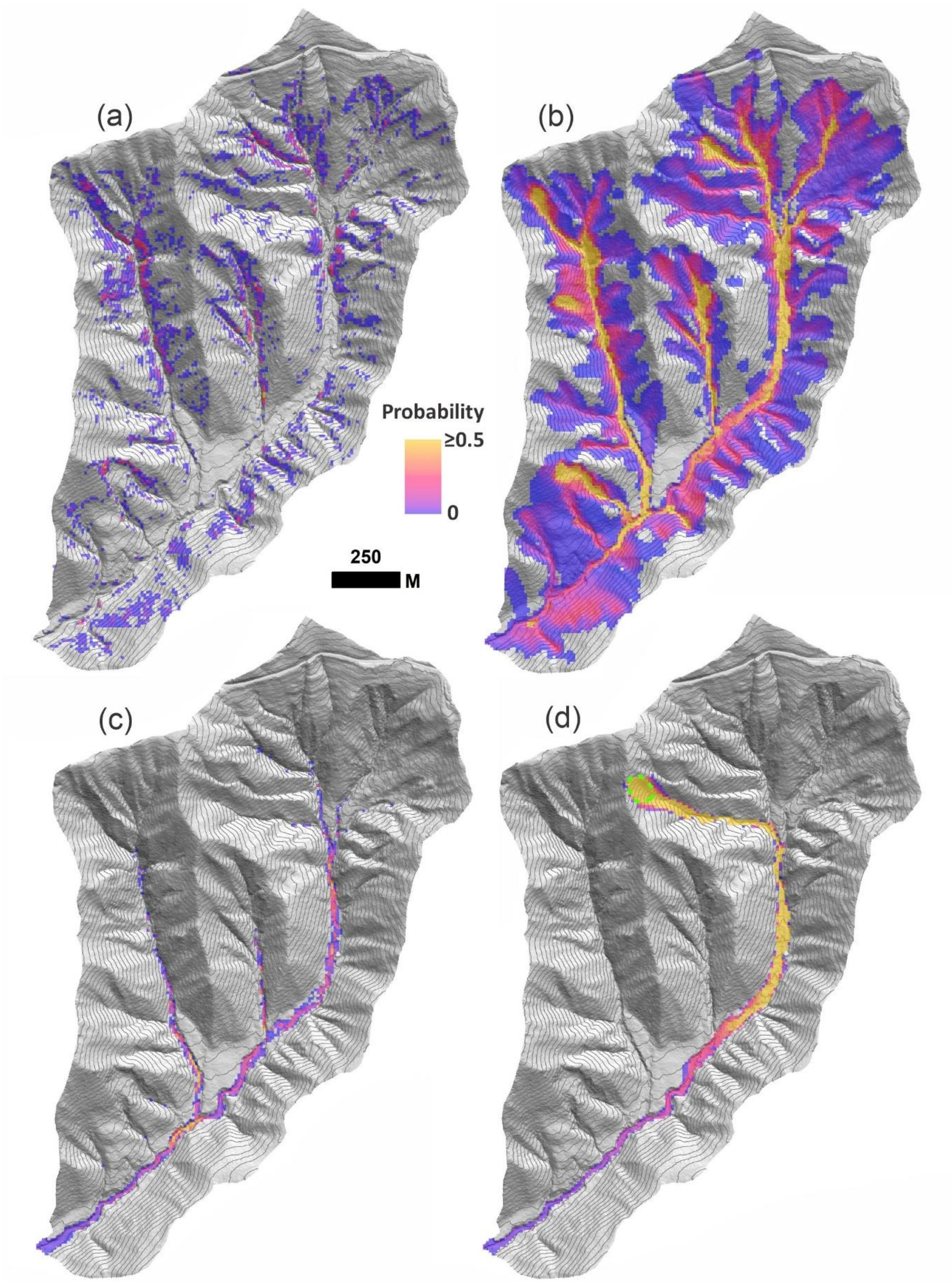

Figure 13. Olympic Mountain site: **(a)** Landslide probability, $P(FS \leq 1)$. **(b)** Corresponding runout probability, $P(\Delta\eta)$. **(c)** Probability of deposition greater than 1 m and **(d)** Runout probability for the potentially unstable slope (green-dashed polygon).

7.0 Concluding remarks

In this study, we described, calibrated and tested MassWastingRunout (MWR), a new cellular-automaton landslide runout model that combines the functionality of simple runout algorithms used in landscape evolution and watershed sediment yield models (WSMs) with the predictive detail typical of runout models used for landslide inundation hazard mapping. MWR is implemented in Python as a component for the Landlab earth surface modelling toolkit and is designed for probabilistic landslide hazard assessments, sediment transport and landscape evolution applications. MWR includes a Markov Chain Monte Carlo calibration utility that determines the best-fit parameter values for a site as well as empirical Probability Density Functions (PDF) of the parameter values. MWR also includes a utility called MWR Probability that takes the PDF output from the calibration utility to determine runout probability.

Results indicate that despite its simple conceptualization, MWR shows skill in modeling the final runout extent, sediment transport and topographic change associated with a landslide. MWR needs only the location and geometry of an initial landslide source area to model the entire runout process. When compared to other models capable of replicating observed landslide inundation patterns, the strength of MWR lies in its use of field-inferable parameters, its ability to internally estimate the total mobilized volume (initial landslide body + erosion volume) and its relatively parsimonious model design.

MWR can be calibrated to a site using just two parameters (critical slope, S_c, and a threshold flux for deposition, q_c) and the MWR calibration utility enables the user to calibrate the model for a watershed within several hours on a standard desktop (Section 5.3). Although the predictive power of MWR hinges on calibration—a common requirement for mechanistic models—its reliance on two calibration parameters serves to constrain model uncertainty. Site-specific calibration may be needed when MWR is used for sediment budget analysis, but if the aim is limited to mapping runout extent, it may be possible to infer parameterization from nearby landslides or possibly from globally available repeated DEMs and air photos that show where past mass-wasting flows have stopped (for S_c) and how thick their frontal lobes are at the point of deposition (for q_c). Nonetheless, as a rules-based, cellular-automaton model, MWR is not designed to accurately simulate flow depth. For accurate flow depths or debris flow impact forces, a detailed-mechanistic modeling approach should be used.

MWR shows a rich set of intuitive responses to topographic curvature and slope. When calibrated to the runout of six different observed landslides, the volumetric error of MWR, $\Delta\eta_E$, ranged between 6% and 15% (median 9.1%) of the observed total mobilized volume. Except for the Rocky Mountains site where MWR consistently modelled wider-than-observed flow, the cumulative flow error along the runout profile (Q_{s_E}) were limited to 5%-19% of the mean cumulative flow determined from the observed DEM-of-Difference (DoD). These are considered acceptable levels of performance given that the total mobilized volume of many debris flow models assume an order of magnitude range of confidence. A notable finding of this paper is that MWR modeled runout did not have any strong systematic bias in predictions (toward unrealistically short or wide flows, for example), which suggests that MWR is structurally sound. However, MWR may underperform compared to mechanistic models when flow momentum is the primary driver of runout extent. (e.g., in areas of slope-perpendicular flow).

As a component of the Landlab earth surface modelling toolkit, MWR is designed to be compatible with other models. MWR can be readily coupled with a landslide initiation model (e.g., LandslideProbability) and geomorphic transport laws for hillslope diffusion and fluvial incision to investigate the role of landslides and their runout on long-term landscape evolution. In this study we showed how to couple MWR with LandslideProbability to map debris flow hazard when landslide initiation location is uncertain. Future studies will explore large-scale application in landscape evolution or sediment yield models, and characterize model parameters for different geologic and hydroclimatic conditions. The use of a calibrated runout model in WSMs might allow for region-specific and more insightful predictions of landslide impact on landscape morphology and watershed-scale sediment dynamics.

Appendix A - Determination of k

The average erosion depth caused by the observed runout ($\bar{E}$) can be determined from the DoD as the total erosion volume ($\sum E\Delta x^2$) divided by the erosion area ($\mathbb{A}$) in the DoD:

$$\bar{E} = \frac{\sum E \Delta x^2}{\mathbb{A}} \qquad \text{(A1)}$$

where $\sum E\Delta x^2$ and $\mathbb{A}$ exclude the initial landslide body volume and area, areas of deposition ($\Delta\eta > 0$) and areas with no change in elevation ($\Delta\eta = 0$). In terms of the debriton conceptualization used in MWR, $\bar{E}$ can also be written as a function of the mean number of times a debriton would need to pass over a grid cell ($\bar{n}$) multiplied by an average erosion depth per debriton ($\bar{h}_e$) to equal $\bar{E}$ as:

$$\bar{E} = \bar{n}\bar{h}_e \qquad \text{(A2)}$$

An estimate for $\bar{n}$ can be determined from the average length of the runout material, which we approximate simply as the mapped landslide length (ℓ) divided by the cell width:

$$\bar{n} = \frac{\ell}{\Delta x} \qquad \text{(A3)}$$

Note that if the observed runout formed as a result of multiple landslides (as was the case at the Olympic Mountain site, see Supplementary Material), then ℓ was determined as the sum of the initial landslide body lengths. Also, as the debritons move down slopes in excess of S_c, they entrain material, split, and spread, and the runout material tends to lengthen. Using the initial landslide length to represent the runout length thus represents a minimum value for $\bar{n}$ and if needed, (A2) can be multiplied by a coefficient to scale ℓ into a more representative runout length. Combining (A2) and (A3), $\bar{h}_e$ can be defined as the average erosion rate per unit length of runout debris ($\bar{E}/\ell$) times the cell width:

$$\bar{h}_e = \frac{\bar{E}\,\Delta x}{\ell} \qquad \text{(A4)}$$

Rewriting equation (11) as a function of the average shear stress in the erosion-dominated reaches of the runout path ($\bar{\tau}$) and assuming $\tau_c \cong 0$, debris flow erodibility parameter k can be estimated as:

$$k = \frac{\bar{h}_e}{\bar{\tau} f} \qquad \text{(A5)}$$

To solve for k, we estimated $\bar{\tau}$ from field-approximated debris flow depth and channel slope measurements in the erosion-dominated reaches of the runout path. To estimate flow depth, we used the height of scour marks on the channel wall or tree trunks, above the channel bed (Table 1). We used (13) to define $\bar{\tau}$. For D_s, we used the average maximum grain size observed over the whole runout path. If τ is defined as a function of grain-collision dependent shear stress approach (13) and k is determined as a function of f, as in (A5), the impact of f on model behavior is relatively small.

Notation

q_{R_i}	[m]	debris flux from a node to each of the node i-th receiver nodes
q_O	[m]	the total out-going debris flux
Nr		the number of receiving nodes of node n
S_i		the underlying topographic slope ($\tan\theta$) to each of the node i-th receiver nodes
a		exponent in (1) that controls how flow is distributed to downslope nodes
q_I	[m]	The total incoming flux
Nd		number of donors nodes to a node
q_{D_j}	[m]	the flux from node D_j (the j-th donor node)
h	[m]	flow depth at node, adjusted to be no more than h_{max}
h_{max}	[m]	the maximum observed flow depth
A	[m]	aggradation depth
S_c		critical slope
S		steepest slope to the node's eight neighbouring nodes
Δx	[m]	cell length
$A_{p\|N_a}$	[m]	potential aggradation depth that forms a deposit that spreads over N_a consecutive nodes
$A_{p,i}$	[m]	i-th deposition amount in the deposit illustrated in Figure 4
N_a		number of nodes qs_n^I is assumed to spreads over
E	[m]	erosion depth
h_r	[m]	regolith depth
h_e	[m]	potential erosion depth
θ	[°]	topographic slope used to determine shear stress, equal to $\tan^{-1}(S)$
τ	[Pa]	basal shear stress
τ_c	[Pa]	critical shear stress of the regolith
k		erodibility parameter in (11)
f		exponent, controls the non-linearity of h_e in (11)
ρ	[kg/m^3]	density of runout material
σ	[Pa]	normal stress at basal surface
φ		tangent of collision angle between grains, measured from the vertical axis
υ_s		volumetric solids concentration
ρ_s	[kg/m^3]	density of solids
D_s	[m]	characteristic particle diameter
u	[m/s]	depth average flow velocity
z	[m]	depth below the flow surface
u^*		shear velocity
g	[m/s]	acceleration due to gravity
$\Delta\eta$	[m]	change in elevation at node
$\boldsymbol{q_D}$		a vector containing all q_{D_j} sent to the node
$\boldsymbol{\xi_D}$		a vector containing the incoming attribute values for each q_{D_j}
ξ_D		attribute value delivered to the node
ξ_R		attribute value sent to receiver nodes

ξ		attribute value at node
η	[m]	topographic elevation
Λ		parameter set
$L(\Lambda)$		likelihood of parameter set
$p(\Lambda)$		prior probability of parameter set
Ω_T		the Lee-Salle index for evaluating model planimetric fit
α	[m^2]	modelled area of matching extent (compared to observed runout extent)
β	[m^2]	modelled area of overestimated extent
γ	[m^2]	modelled area of underestimated extent
$\Delta\eta_E$		volumetric error of the modelled topographic change relative to the observed total mobilized volume, fraction.
V	[m^3]	observed total mobilized volume
p		the number of nodes in the modelled runout extent
$\Delta\eta_{Mi}$	[m]	the modelled topographic change [m] at the i-th node within the runout extent
$\Delta\eta_{Oi}$	[m]	the observed topographic change [m] at the i-th node within the runout extent
Q_{s_E}		mean-modelled-cumulative flow error along the runout path relative to the observed mean cumulative flow, fraction.
j		index used to represent each node along a profile of the runout path.
$\Delta\eta_{ij}$	[m]	topographic change [m] at the i-th node located upstream of node j
u_j		totalnumber of all nodes located upstream of node j
r		the number of nodes along the center line of the runout path
Q_s	[m^3]	the cumulative debris flow volume at each node, j along the center line of the runout path
Q_{sO}	[m^3]	the observed cumulative debris flow volume (Q_s) at each node, j
Q_{sM}	[m^3]	the modeled cumulative debris flow volume (Q_s) at each node, j
$\Delta\eta_{ij}$	[m]	the topographic change [m] at the i-th node located upstream of node j
u_j		the total number of all nodes located upstream of j
$\overline{Q_{sO}}$	[m^3]	the observed mean cumulative flow
$P(\Delta\eta)$		
$\bar{E}/\ell$	[m/m]	average erosion per unit length of runout debris
$P(\Delta\eta)$		probability of runout, expressed as the probability that the elevation of a node changes
$\#$		number of
Np		number Monte Carlo iterations used to determine probability
$\mathbb{A}$	[m^2]	erosion area of the observed or modeled runout
$\bar{E}$	[m]	average erosion depth caused by the runout
$\sum E\Delta x^2$	[m^3]	the total erosion volume
$\bar{n}$		mean number of times a debriton would need to pass over a grid cell multiplied by an average erosion depth per debriton to equal $\bar{E}$
$\bar{h}_e$	[m]	average erosion depth per debriton
ℓ	[m]	length of runout debris, approximated as the length of the initial landslide body
H/L		the total topographic relief of the runout (measured from the center of the landslide to the end of the runout path) divided by the horizontal length of the runout
κ	[1/m]	mean total curvature
SPI		mean specific stream power index
FS		Factor-of-Safety, ratio of the resisting to the driving forces acting on a hillslope

Code availability

MassWastingRunout and several tutorial notebooks area available at: https://github.com/landlab/landlab

Acknowledgements

This research was partially supported by the following programs: National Science Foundation (NSF) PREEVENTS program, ICER-1663859; NSF OAC-2103632; and NASA Grant number 22-RRNES22-0010 and benefited from critical referee reviews. Stephen Slaughter field reviewed the Cascade Mountains, 2009 and Black Hills landslides the year they occurred and provided photos and field observations that aided author interpretation. John Jenkins helped with the 2022 field reconnaissance the Cascade Mountains landslide. Eli Schwat helped with field reconnaissance at the Olympic Mountains site. This work also benefitted from Landlab support and coding guidance from Eric Hutton and helpful feedback from many others.

Competing interests

The contact author has declared that none of the authors has any competing interests.

References

Bagnold, R. A.: Experiments on a gravity-free dispersion of large solid spheres in a Newtonian fluid under shear. Proceedings of the Royal Society of London, 225(1160), 49–63. https://doi.org/10.1098/rspa.1954.0186, 1954.

Barca, D., Crisci, G., Di Gregorio, S., and Nicoletta, F.: Cellular automata method for modelling lava flows: Simulation of the 1986–1987 eruption, Mount Etna, Sicily, in Kilburn, C., and Luongo, G., eds., Active lavas: Monitoring and modeling: London, University College of London Press, p. 291–309, 1993.

Barnhart, K. R., Hutton, E. W. H., Tucker, G. E., Gasparini, N. M., Istanbulluoglu, E., Hobley, D. E. J., Lyons, N. J., Mouchene, M., Nudurupati, S. S., Adams, J. M., & Bandaragoda, C.: Short communication: Landlab v2.0: a software package for Earth surface dynamics. Earth Surface Dynamics, 8(2), 379–397. https://doi.org/10.5194/esurf-8-379-2020, 2020.

Barnhart, K. R., Jones, R., George, D. J., McArdell, B. W., Rengers, F. K., Staley, D. M., & Kean, J. W.: Multi-Model Comparison of Computed Debris Flow Runout for the 9 January 2018 Montecito, California Post-Wildfire Event. Journal of Geophysical Research: Earth Surface, 126(12). https://doi.org/10.1029/2021jf006245, 2021.

Benda, L., & Dunne, T.: Stochastic forcing of sediment supply to channel networks from landsliding and debris flow. Water Resources Research, 33(12), 2849–2863. https://doi.org/10.1029/97wr02388, 1997.

Benda, L., Veldhuisen, C. P., & Black, J.: Debris flows as agents of morphological heterogeneity at low-order confluences, Olympic Mountains, Washington. Geological Society of America Bulletin, 115(9), 1110. https://doi.org/10.1130/b25265.1, 2003.

Beven, K.: A manifesto for the equifinality thesis. Journal of Hydrology, 320(1–2), 18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007, 2006.

Burton, A., & Bathurst, J. C.: Physically based modelling of shallow landslide sediment yield at a catchment scale. Environmental Geology, 35(2–3), 89–99. https://doi.org/10.1007/s002540050296, 1998.

Bigelow, P., Benda, L., Miller, D., & Burnett, K. M.: On Debris Flows, River Networks, and the Spatial Structure of Channel Morphology. Forest Science, 53(2), 220–238. https://doi.org/10.1093/forestscience/53.2.220, 2007

Campforts, B., Shobe, C. M., Overeem, I., & Tucker, G. E.: The Art of Landslides: How Stochastic Mass Wasting Shapes Topography and Influences Landscape Dynamics. Journal of Geophysical Research: Earth Surface, 127(8). https://doi.org/10.1029/2022jf006745, 2022

Campforts, B., Shobe, C. M., Steer, P., Vanmaercke, M., Lague, D., & Braun, J.: HyLands 1.0: a hybrid landscape evolution model to simulate the impact of landslides and landslide-derived sediment on landscape evolution. Geoscientific Model Development, 13(9), 3863–3886. https://doi.org/10.5194/gmd-13-3863-2020, 2020.

Capart, H., & Fraccarollo, L.: Transport layer structure in intense bed-load. Geophysical Research Letters, 38(20), n/a. https://doi.org/10.1029/2011gl049408, 2011.

Capart, H., Hung, C., & Stark, C. R.: Depth-integrated equations for entraining granular flows in narrow channels. Journal of Fluid Mechanics, 765. https://doi.org/10.1017/jfm.2014.713, 2015.

Carretier, S., Martinod, P., Reich, M., & Goddéris, Y.: Modelling sediment clasts transport during landscape evolution. Earth Surface Dynamics, 4(1), 237–251. https://doi.org/10.5194/esurf-4-237-2016, 2016.

Chase, C. G.: Fluvial landsculpting and the fractal dimension of topography. Geomorphology, 5(1–2), 39–57. https://doi.org/10.1016/0169-555x(92)90057-u, 1992.

Chen, C., & Yu, F.: Morphometric analysis of debris flows and their source areas using GIS. Geomorphology, 129(3–4), 387–397. https://doi.org/10.1016/j.geomorph.2011.03.002, 2011.

Chen, H., & Zhang, L.: EDDA 1.0: integrated simulation of debris flow erosion, deposition and property changes. Geoscientific Model Development, 8(3), 829–844. https://doi.org/10.5194/gmd-8-829-2015, 2015.

Chen, T.-Y. K., Wu, Y.-C., Hung, C.-Y., Capart, H., and Voller, V. R.: A control volume finite-element model for predicting the morphology of cohesive-frictional debris flow deposits, Earth Surface Dynamics, 11, 325–342, https://doi.org/10.5194/esurf-11-325-2023, 2023.

Clerici, A., & Perego, S.: Simulation of the Parma River blockage by the Corniglio landslide (Northern Italy). Geomorphology, 33(1–2), 1–23. https://doi.org/10.1016/s0169-555x(99)00095-1, 2000.

Codd, E. F.: Cellular Automata (1st ed.). New York, Academic Press, 1968.

Coz, J. L., Renard, B., Bonnifait, L., Branger, F., & Boursicaud, R. L.: Combining hydraulic knowledge and uncertain gaugings in the estimation of hydrometric rating curves: A Bayesian approach. Journal of Hydrology, 509, 573–587. https://doi.org/10.1016/j.jhydrol.2013.11.016, 2014.

Crave, A., & Davy, P.: A stochastic “precipiton” model for simulating erosion/sedimentation dynamics. Computers & Geosciences, 27(7), 815–827. https://doi.org/10.1016/s0098-3004(00)00167-9, 2001.

D’Ambrosio, D., Di Gregorio, S., Iovine, G., Lupiano, V., Rongo, R., & Spataro, W.: First simulations of the Sarno debris flows through Cellular Automata modelling. Geomorphology, 54(1–2), 91–117. https://doi.org/10.1016/s0169-555x(03)00058-8, 2003.

Egashira, S., Honda, N., & Itoh, T.: Experimental study on the entrainment of bed material into debris flow. Physics and Chemistry of the Earth, Parts a/B/C, 26(9), 645–650. https://doi.org/10.1016/s1464-1917(01)00062-9, 2001.

Foglia, L., Hill, M. C., Mehl, S. W., and Burlando, P. (2009), Sensitivity analysis, calibration, and testing of a distributed hydrological model using error-based weighting and one objective function, Water Resources Research, 45, W06427, doi:10.1029/2008WR007255.

Fannin, R. J., & Wise, M. P.: An empirical-statistical model for debris flow travel distance. Canadian Geotechnical Journal, 38(5), 982–994. https://doi.org/10.1139/t01-030, 2001.

Frank, F., McArdell, B. W., Huggel, C., & Vieli, A.: The importance of entrainment and bulking on debris flow runout modeling: examples from the Swiss Alps. Natural Hazards and Earth System Sciences, 15(11), 2569–2583. https://doi.org/10.5194/nhess-15-2569-2015, 2015.

Freeman, T. G.: Calculating catchment area with divergent flow based on a regular grid. Computers & Geosciences, 17(3), 413–422. https://doi.org/10.1016/0098-3004(91)90048-i, 1991.

Gartner, J. E., Cannon, S. H., & Santi, P. M.: Empirical models for predicting volumes of sediment deposited by debris flows and sediment-laden floods in the transverse ranges of southern California. Engineering Geology, 176, 45–56. https://doi.org/10.1016/j.enggeo.2014.04.008, 2014.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B.: Bayesian Data Analysis (3rd ed.). Electronic Edition, 2021.

Goode, J. R., Luce, C. H., & Buffington, J. M.: Enhanced sediment delivery in a changing climate in semi-arid mountain basins: Implications for water resource management and aquatic habitat in the northern Rocky Mountains. Geomorphology, 139–140, 1–15. https://doi.org/10.1016/j.geomorph.2011.06.021, 2012.

Gorr, A., McGuire, L. A., Youberg, A., & Rengers, F. K.: A progressive flow-routing model for rapid assessment of debris-flow inundation. Landslides, 19(9), 2055–2073. https://doi.org/10.1007/s10346-022-01890-y, 2022

Guthrie, R., Hockin, A., Colquhoun, L., Nagy, T., Evans, S. G., & Ayles, C. P.: An examination of controls on debris flow mobility: Evidence from coastal British Columbia. Geomorphology, 114(4), 601–613. https://doi.org/10.1016/j.geomorph.2009.09.021, 2010.

Guthrie, R. H., & Befus, A. D.: DebrisFlow Predictor: an agent-based runout program for shallow landslides. Natural Hazards and Earth System Sciences, 21(3), 1029–1049. https://doi.org/10.5194/nhess-21-1029-2021, 2021.

Hammond C.J., Prellwitz R.W., Miller S.M.: Landslides hazard assessment using Monte Carlo simulation. In: Bell DH (ed) Proceedings of 6th international symposium on landslides, Christchurch, New Zealand, Balkema, vol 2. pp 251–294, 1992.

Han, Z., Chen, G., Li, Y., Tang, C., Xu, L., He, Y., Huang, X., & Wang, W.: Numerical simulation of debris-flow behavior incorporating a dynamic method for estimating the entrainment. Engineering Geology, 190, 52–64. https://doi.org/10.1016/j.enggeo.2015.02.009, 2015.

Han, Z., Li, Y., Huang, J., Chen, G., Xu, L., Tang, C. Y., Zhang, H., & Shang, Y.: Numerical simulation for run-out extent of debris flows using an improved cellular automaton model. Bulletin of Engineering Geology and the Environment, 76(3), 961–974. https://doi.org/10.1007/s10064-016-0902-6, 2017.

Han, Z., Ma, Y., Li, Y., Zhang, H., Chen, N., Hu, G., & Chen, G.: Hydrodynamic and topography based cellular automaton model for simulating debris flow run-out extent and entrainment behavior. Water Research, 193, 116872. https://doi.org/10.1016/j.watres.2021.116872, 2021.

Heiser, M., Scheidl, C., & Kaitna, R.: Evaluation concepts to compare observed and simulated deposition areas of mass movements. Computational Geosciences, 21(3), 335–343. https://doi.org/10.1007/s10596-016-9609-9, 2017.

Hobley, D. E. J., Adams, J. M., Nudurupati, S. S., Hutton, E. W. H., Gasparini, N. M., Istanbulluoglu, E., & Tucker, G. E.: Creative computing with Landlab: an open-source toolkit for building, coupling, and exploring two-dimensional numerical models of Earth-surface dynamics. Earth Surface Dynamics, 5(1), 21–46. https://doi.org/10.5194/esurf-5-21-2017, 2017.

Holmgren, P.: Multiple flow direction algorithms for runoff modelling in grid based elevation models: An empirical evaluation. Hydrological Processes, 8(4), 327–334. https://doi.org/10.1002/hyp.3360080405, 1994

Horton, P., Jaboyedoff, M., Rudaz, B., & Zimmermann, M. N.: Flow-R, a model for susceptibility mapping of debris flows and other gravitational hazards at a regional scale. Natural Hazards and Earth System Sciences, 13(4), 869–885. https://doi.org/10.5194/nhess-13-869-2013, 2013.

Hungr, O., Morgan, G. J., & Kellerhals, R.: Quantitative analysis of debris torrent hazards for design of remedial measures. Canadian Geotechnical Journal, 21(4), 663–677. https://doi.org/10.1139/t84-073, 1984.

Hungr, O., & Evans, S. G.: Entrainment of debris in rock avalanches: An analysis of a long run-out mechanism. Geological Society of America Bulletin, 116(9), 1240. https://doi.org/10.1130/b25362.1, 2004.

Hutter, K., Svendsen, B., & Rickenmann, D.: Debris flow modeling: A review. Continuum Mechanics and Thermodynamics, 8(1), 1–35. https://doi.org/10.1007/bf01175749, 1996.

Iovine, G., D'Ambrosio, D., & Di Gregorio, S.: Applying genetic algorithms for calibrating a hexagonal cellular automata model for the simulation of debris flows characterised by strong inertial effects. Geomorphology, 66(1–4), 287–303. https://doi.org/10.1016/j.geomorph.2004.09.017, 2005.

Istanbulluoglu, E. Bras R. L.: Vegetation-modulated landscape evolution: Effects of vegetation on landscape processes, drainage density, and topography. Journal of Geophysical Research, 110(F2). https://doi.org/10.1029/2004jf000249, 2005.

Istanbulluoglu, E., Bras R. L., Flores-Cervantes, H., and Tucker, G. E..: Implications of bank failures and fluvial erosion for gully development: Field observations and modeling, J. Geophysical Research, 110, F01014, doi:10.1029/2004JF000145, 2005.

Istanbulluoglu, E., O. Yetemen, E. R. Vivoni, H. A. Gutie´rrez-Jurado, and R. L. Bras, Eco-geomorphic implications of hillslope aspect: Inferences from analysis of landscape morphology in central New Mexico, Geophysical. Research Letters, 35, L14403, 10.1029/ 2008GL034477, 2008.

Iverson, R. M.: The physics of debris flows. Reviews of Geophysics, 35(3), 245–296. https://doi.org/10.1029/97rg00426, 1997.

Iverson, R. M., & Denlinger, R. P.: Flow of variably fluidized granular masses across three-dimensional terrain: 1. Coulomb mixture theory. Journal of Geophysical Research, 106(B1), 537–552. https://doi.org/10.1029/2000jb900329, 2001.

Iverson, R.M., How should mathematical models of geomorphic processes be judged?. In Wilcock, P., & Iverson, R. (Eds.), Prediction in Geomorphology. American Geophysical Union, 2003.

Julien, P. Y., & Paris, A.: Mean Velocity of Mudflows and Debris Flows. Journal of Hydraulic Engineering, 136(9), 676–679. https://doi.org/10.1061/(asce)hy.1943-7900.0000224, 2010.

Kean, J. W., Staley, D. M., Lancaster, J., Rengers, F., Swanson, B., Coe, J., et al.: Inundation, flow dynamics, and damage in the 9 January 2018 Montecito debris-flow event, California, USA: Opportunities and challenges for post-wildfire risk assessment. Geosphere, 15(4), 1140–1163. https://doi.org/10.1130/GES02048.1, 2019

Korup, O.: Effects of large deep-seated landslides on hillslope morphology, western Southern Alps, New Zealand. Journal of Geophysical Research, 111(F1). https://doi.org/10.1029/2004jf000242, 2006

Lancaster, S. T., Hayes, S. K., & Grant, G. E.: Effects of wood on debris flow runout in small mountain watersheds. Water Resources Research, 39(6). https://doi.org/10.1029/2001wr001227, 2003.

Larsen, I. J., & Montgomery, D. R.: Landslide erosion coupled to tectonics and river incision. Nature Geoscience, 5(7), 468–473. https://doi.org/10.1038/ngeo1479, 2012.

Liu, J., Wu, Y., Gao, X., & Zhang, X.: A Simple Method of Mapping Landslides Runout Zones Considering Kinematic Uncertainties. Remote Sensing, 14(3), 668. https://doi.org/10.3390/rs14030668, 2022.

Major, J. J.: Depositional Processes in Large-Scale Debris-Flow Experiments. The Journal of Geology, 105(3), 345–366. https://doi.org/10.1086/515930, 1997.

Major, J. J., & Iverson, R. M.: Debris-flow deposition: Effects of pore-fluid pressure and friction concentrated at flow margins. Geological Society of America Bulletin, 111(10), 1424–1434. https://doi.org/10.1130/0016-7606(1999)111, 1999.

McCoy, S. W., Kean, J. W., Coe, J. A., Tucker, G. S., Staley, D. M., & Wasklewicz, T. A.: Sediment entrainment by debris flows: In situ measurements from the headwaters of a steep catchment. Journal of Geophysical Research, 117(F3), n/a. https://doi.org/10.1029/2011jf002278, 2012.

McDougall, S., & Hungr, O.: A model for the analysis of rapid landslide motion across three-dimensional terrain. Canadian Geotechnical Journal, 41(6), 1084–1097. https://doi.org/10.1139/t04-052, 2004.

Medina, V., Hürlimann, M., & Bateman, A.: Application of FLATModel, a 2D finite volume code, to debris flows in the northeastern part of the Iberian Peninsula. Landslides, 5(1), 127–142. https://doi.org/10.1007/s10346-007-0102-3, 2008.

Montgomery, D. R., & Dietrich, W. E.: Where do channels begin? Nature, 336(6196), 232–234. https://doi.org/10.1038/336232a0, 1988.

Murray, B.A., & Paola, C.: A cellular model of braided rivers. Nature, 371(6492), 54–57. https://doi.org/10.1038/371054a0, 1994.

Murray, A. B., & Paola, C.: Properties of a cellular braided-stream model. Earth Surface Processes and Landforms, 22(11), 1001–1025. https://doi.org/10.1002/(sici)1096-9837(199711)22:11, 1997.

Murray A.B.: Which Models Are Good (Enough), and When?. In: John F. Shroder (ed.) Treatise on Geomorphology, Volume 2, pp. 50-58. San Diego: Academic Press, 2013.

Natural Resources Conservation Service | Snow and Water Interactive Map (n.d.). Natural Resources Conservation Service. https://www.nrcs.usda.gov/resources/data-and-reports/snow-and-water-interactive-map, Accessed April, 2022

Nudurupati, S. S., Istanbulluoglu, E., Tucker, G. E., Gasparini, N. M., Hobley, D. E. J., Hutton, E. W. H., Barnhart, K. R., & Adams, J. M.: On transient semi-arid ecosystem dynamics using Landlab: vegetation shifts, topographic refugia, and response to climate. Water Resources Research, 59(4). https://doi.org/10.1029/2021wr031179, 2023.

Perron, J. T.: Climate and the Pace of Erosional Landscape Evolution. Annual Review of Earth and Planetary Sciences, 45(1), 561–591. https://doi.org/10.1146/annurev-earth-060614-105405, 2017.

Reid, M. J., Coe, J. A., & Brien, D. L.: Forecasting inundation from debris flows that grow volumetrically during travel, with application to the Oregon Coast Range, USA. Geomorphology, 273, 396–411. https://doi.org/10.1016/j.geomorph.2016.07.039, 2016.

Renard, B., Garreta, V., & Lang, M. J.: An application of Bayesian analysis and Markov chain Monte Carlo methods to the estimation of a regional trend in annual maxima. Water Resources Research, 42(12). https://doi.org/10.1029/2005wr004591, 2006.

Rengers, F. K., McGuire, L. A., Kean, J. W., Staley, D. M., and Hobley, D. E. J.: Model simulations of flood and debris flow timing in steep catchments after wildfire, Water Resources Research, 52, 6041–6061, doi:10.1002/2015WR018176, 2016.

Roda-Boluda, D. C., D'Arcy, M., McDonald, J., & Whittaker, A. C.: Lithological controls on hillslope sediment supply: insights from landslide activity and grain size distributions. Earth Surface Processes and Landforms, 5), 956–977. https://doi.org/10.1002/esp.4281, 2018.

Shaller, P. J., Doroudian, M., & Hart, M. W.: The Eureka Valley Landslide: Evidence of a dual failure mechanism for a Long-Runout Landslide. Lithosphere, 2020(1). https://doi.org/10.2113/2020/8860819, 2020.

Schürch, P., Densmore, A. L., Rosser, N., & McArdell, B. W.: Dynamic controls on erosion and deposition on debris-flow fans. Geology, 39(9), 827–830. https://doi.org/10.1130/g32103.1, 2011.

Shen, P., Zhang, L. M., Wong, H., Peng, D., Zhou, S., Zhang, S., & Chen, C.: Debris flow enlargement from entrainment: A case study for comparison of three entrainment models. Engineering Geology, 270, 105581. https://doi.org/10.1016/j.enggeo.2020.105581, 2020.

Stock, J. P. J., & Dietrich, W. E.: Erosion of steepland valleys by debris flows. Geological Society of America Bulletin, 118(9–10), 1125–1148. https://doi.org/10.1130/b25902.1, 2006.

Strauch, R. L., Istanbulluoglu, E., Nudurupati, S. S., Bandaragoda, C., Gasparini, N. M., & Tucker, G. E.: A hydroclimatological approach to predicting regional landslide probability using Landlab. Earth Surface Dynamics, 6(1), 49–75. https://doi.org/10.5194/esurf-6-49-2018, 2018.

Takahashi, T.: Debris Flow (2nd ed.). CRC Press, Taylor & Francis Group, 2014.

Tucker, G. E., & Bras, R. L.: Hillslope processes, drainage density, and landscape morphology. Water Resources Research, 34(10), 2751–2764. https://doi.org/10.1029/98wr01474, 1998.

Tucker, G. E., Hancock, G. J.: Modelling landscape evolution. Earth Surface Processes and Landforms, 35(1), 28–50. https://doi.org/10.1002/esp.1952, 2010.

Tucker, G. E., McCoy, S., & Hobley, D. E. J.: A lattice grain model of hillslope evolution. Earth Surface Dynamics, 6(3), 563–582. https://doi.org/10.5194/esurf-6-563-2018, 2018.

Western Regional Climate Center. (n.d.), from https://wrcc.dri.edu/, accessed 2017 and 2022

Whipple, K. X., & Dunne, T.: The influence of debris-flow rheology on fan morphology, Owens Valley, California. Geological Society of America Bulletin, 104(7), 887–900. https://doi.org/10.1130/0016-7606(1992)104, 1992.

Zhou, G. G. D., Li, S., Song, D., Choi, C. E., & Chen, X.: Depositional mechanisms and morphology of debris flow: physical modelling. Landslides, 16(2), 315–332. https://doi.org/10.1007/s10346-018-1095-9, 2019.