



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

Jeffrey Keck<sup>1,2</sup>, Erkan Istanbuluoglu<sup>1</sup>, Benjamin Campforts<sup>3</sup>, Gregory Tucker<sup>4,5</sup>, and Alexander Horner-Devine<sup>1</sup>

<sup>1</sup>University of Washington, Civil and Environmental Engineering, Seattle, WA, USA

<sup>2</sup>Washington Department of Natural Resources, Forest Resources Division, Olympia, WA, USA

<sup>3</sup>Department of Earth Sciences, Vrije University, Amsterdam, the Netherlands

<sup>4</sup>Department of Geological Sciences, University of Colorado Boulder, Boulder, CO, USA

<sup>5</sup>Cooperative Institute for Research in Environmental Sciences (CIRES),  
University of Colorado Boulder, Boulder, CO, USA

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

Received: 14 July 2023 – Discussion started: 3 August 2023

Revised: 15 July 2024 – Accepted: 29 July 2024 – Published: 8 October 2024

**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 detail typical of runout models used for landslide inundation hazard mapping. An initial digital elevation model (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 an adaptive Bayesian 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 modeling 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 in-

undate 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 an emphasis on inundation extent and flow depth (e.g., Frank et al., 2015; Han

et al., 2015); (2) watershed sediment yield models, with an emphasis on the mobilization, deposition, and type of sediment carried by the landslide (e.g., Bathurst and Burton, 1998; Istanbuluoglu, et al., 2005); and (3) landscape evolution models, with an emphasis on topographic change prediction (e.g., Tucker and Bras, 1998; Istanbuluoglu 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 the initial relief and size of the landslide (Iverson, 1997). Mobility may also be impacted by substrate liquefaction (Hungar 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 grain size (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 runout material (Takahashi, 2014; Major and Iverson, 1999; Zhou et al., 2019), but the friction angle of the runout material may vary as a function of the grains in the flow and fluidization (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 and Fracarollo, 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 and statistical models that use simple geometric rules and an estimate of the total mobilized volume (initial landslide body + 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 mobilized volume (e.g., Barnhart et al., 2021; Han et al., 2015);

(3) reduced- or appropriate-complexity flow-routing models (e.g., Murray, 2007) 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 modeling 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 collectively referred to as watershed sediment models, WSMs), the runout model must be scalable in both space and time and capable of modeling the entire runout process given an internally modeled 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 but with simplifications that can restrict model ability to accurately replicate observed inundation extent or depositional patterns. 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; Istanbuluoglu and Bras, 2005) or link of a channel network (Benda and Dunne, 1997).

We developed a new reduced-complexity landslide runout model, called MassWastingRunout (MWR), that bridges 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. 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, making it suitable for sediment yield studies. MWR can be calibrated by adjusting just two parameters ( $S_c$  and  $q_c$ , described in Sect. 2) and is augmented with a Bayesian 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, making MWR suitable for hazard assessment applications.

In this paper, we present the conceptualization and numerical implementation of the MWR model (Sect. 2), describe the calibration utility and probabilistic implementation of MWR (Sect. 3), and demonstrate basic model response to topographic convergence and slope on a series of syn-

thetic terrains (Sect. 4). Event-scale applications to replicate observed runout extent, sediment transport, and topographic change at four topographically and geologically unique field sites (see Fig. 1) are discussed (Sect. 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 (Sect. 6). We conclude with a short summary of MWR model performance and suggest how a calibrated MWR can be incorporated into WSMs.

## 2 Description of the MassWastingRunout model

### 2.1 Overview of the cellular-automaton modeling 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) and 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 and Davy, 2001; Murray and 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 landslide runout processes described in the Introduction, MWR explicitly represents erosion, deposition, and flow resistance as a function of the initial landslide body and downslope terrain. Material exchange between the runout material and underlying regolith 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) modeled 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 “debritions”. The debritions represent debris flux,

here defined as a volume of debris transferred per model iteration per grid cell area, [ $\text{m}^3 \text{m}^{-2}$  per 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; Hobbey 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 Fig. 5 of Hobbey et al., 2017). In the subsequent sections we describe the model theory. All parameters and variables used in the theory are listed in Table B1.

### 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 debrition (Fig. 2, iteration  $t = 0$ ) with a magnitude equal to the mass-wasting source material depth at the node and basal elevation equal to the initial topography minus the mass-wasting source material depth at the node. The basal elevation can be thought to represent the rupture or slip surface of the source material, and the redistribution (flux) of each debrition to its downslope nodes (receiver nodes) is determined as a function of the slope of the slip surface. At the lowest-elevation debrition of the source material, flux to its downslope nodes is determined using the surface slope of the initial DEM (see the flow direction of the lowest node in Fig. 3a). This implementation helps to ensure that the lowest-elevation debrition in the mass-wasting source material moves downslope and movement of upslope debritions is impacted by the geometry of the mass-wasting source material. For example, the receiver nodes of the lowest-elevation debrition in the initial landslide body illustrated in Fig. 2 (node 45) would be identified as those among the eight neighboring nodes whose initial topographic elevation is less than the initial topographic elevation of node 45, while for the debrition at node 51, the receiver nodes would be identified as those among the eight neighboring nodes whose topographic elevation is less than the basal elevation of the debrition (see Fig. 3a).

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

Algorithm 2 is essentially the runout model. It determines how each debrition traverses and modifies the landscape. After receiver nodes from the first model iteration are determined in Algorithm 1 (iteration  $t = 0$ ), Algorithm 2 is re-





**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 moderately sized debris avalanche over a steep, unconfined to divergent hillslope. **(d)** Olympic Mountains, WA: small debris flows in steep, highly convergent channels. Image scale varies with depth, but the approximate scale is indicated at the location of the scale bar.

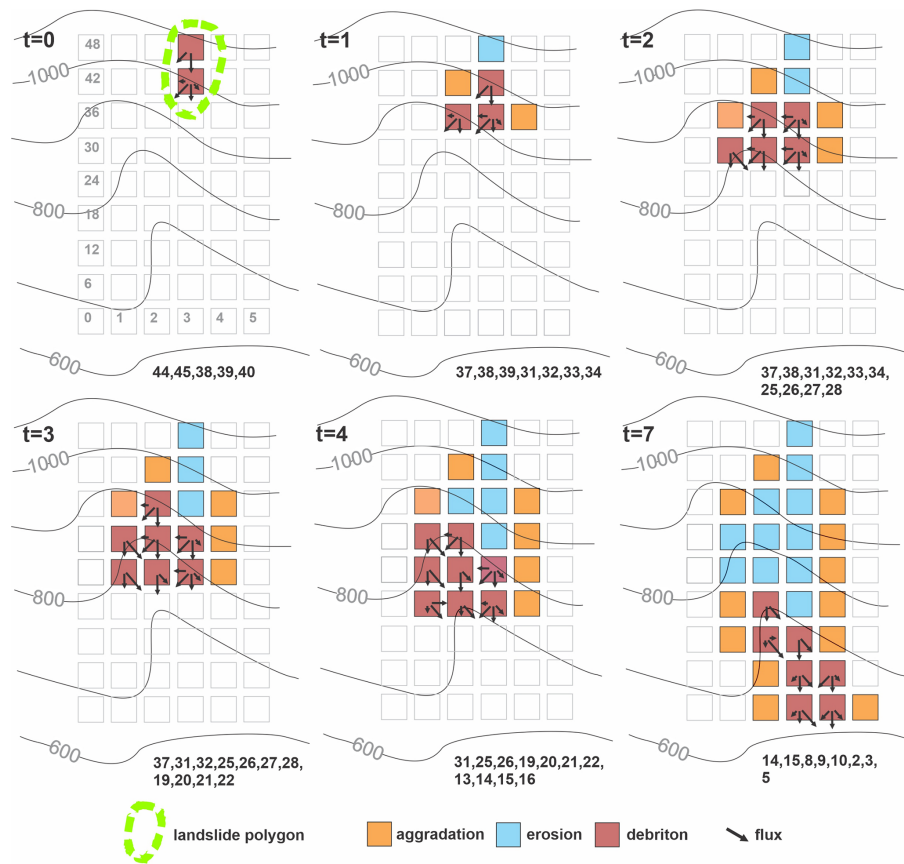
peatedly implemented until all material has deposited (i.e., there are no debritons). Each debriton moves one grid cell per model iteration; the larger the landslide size, the more iterations are necessary to evacuate the landslide slip surface. As each debriton moves, it may erode or aggrade the landscape, impacting the movement of any upslope debritons. As is common with other reduced-complexity models (e.g., Guthrie and Befus, 2021), we assume that inertial effects have a negligible impact on flow behavior (i.e., the kine-

matic flow approximation) and 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 the terrain under the debriton). We do this using the Freeman (1991) multiple-flow-direction algorithm:

$$q_{R_i} = q_0 \frac{S_i^a}{\sum_{i=1}^{N_r} S_i^a}, \quad (1)$$

where  $q_0$  is the total outgoing flux from the node and has units of depth in meters per model iteration,  $N_r$  is the number





**Figure 2.** Illustration of the mobilization and runout of an initial mass-wasting source area (landslide body) 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.

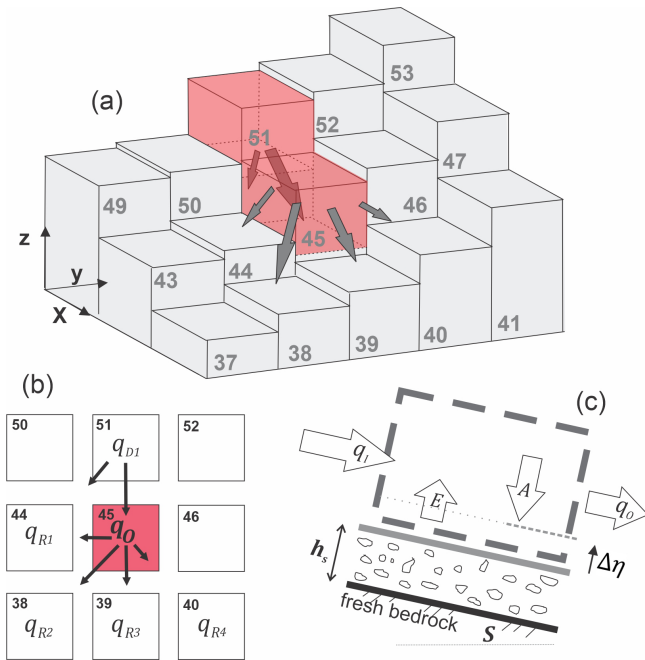
of receiving nodes,  $i$  is the index for each receiver node (e.g.,  $i = 1, 2 \dots N_r$ ), and  $S_i$  is the underlying topographic slope to the  $i$ th receiver node (Fig. 3b). The Freeman (1991) multiple-flow-direction algorithm is a commonly used approximation for two-dimensional flow, and in 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 that 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 of meters per model iteration) towards a given node ( $q_I$ ) is determined by summing the flux from each of the node's donor nodes:

$$q_I = \sum_{j=1}^{N_d} q_{D_j}, \quad (2)$$

where  $N_d$  is the number of donor nodes, and  $q_{D_j}$  is the flux from node  $D_j$  (the  $j$ th donor node,  $j = 1, 2 \dots N_d$ ; Fig. 3b).

As noted by Tucker and Hancock (2010), the flow depths calculated from two-dimensional flow approximations as in Eq. (1) can be influenced by the choice of grid size used to represent the terrain. Additionally, as simplified multi-directional flow models as in Eq. (1) neglect the pressure and momentum forces in the movement of flow, they can result in inaccurate flow width and depth estimates, depending on terrain slope and convergence. Rengers et al. (2016) noted these limitations when using a kinematic wave approximation of the shallow-water equations, as this approximation lacks a pressure term that facilitates the spreading of the modeled water surface. While the topographic controls on mass conservation are adequately represented by Eq. (1), our model bears such limitations when calculating flow depth and width. Moreover, in our model, flow depth is used to determine a depth-dependent erosion rate. As such, in order to avoid unrealistically high flow depths (and erosion rates), we constrain flow depth to an upper limit as

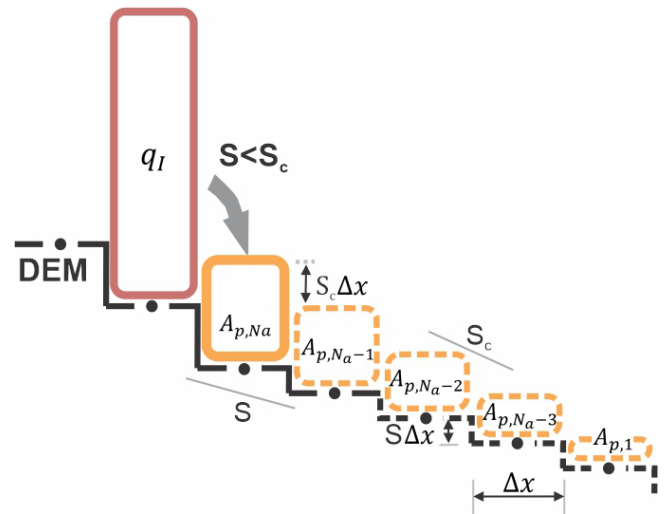
$$h = \min(h_{\max}, q_I), \quad (3)$$



**Figure 3.** (a) Three-dimensional illustration of iteration  $t = 0$  in Fig. 2, showing initial source material nodes (represented by red cells) and flux towards downslope nodes. (b) Distribution of  $q_0$  to downslope nodes 38, 39, 40, and 44; (c) illustration of mass continuity applied to any node that receives a debriton.

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 judgment (see Sect. 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 an  $S_c$ -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 the slope at the node ( $S \ll S_c$ ). To resolve this problem, aggradation depth can be limited to  $A \leq S_c \Delta x$  (where  $\Delta x$  is the 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



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

slope of the overall deposit will be equal to  $S_c$ , as shown in Fig. 4 and described as follows.

Aggradation at a node is determined as

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

where  $S$  is the steepest slope to the node's eight neighboring nodes, and  $A_{p,N_a}$  is a potential aggradation depth necessary to form the beginning of the overall 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) has a surface slope equal to the critical slope  $S_c$ , and (4) has an underlying topographic slope equal to  $S$  that is assumed to be constant over the  $N_a$  consecutive nodes of deposition. From these assumptions, 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 Eq. (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

$$q_I = \sum_{i=1}^{N_a} A_{p,i}, \quad (5)$$

where  $A_{p,i}$  is the  $i$ th deposition amount in the overall deposit and  $i = 1$  is the last node of deposition ( $A_{p,1}$ ; see Fig. 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

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

From Eq. (6) we can rewrite Eq. (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 Eq. (7) into Eq. (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 Eq. (8) to solve for  $A_{p,N_a}$  and Eq. (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 Fig. 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 per 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,  $\tau$  [Pa] (Eqs. 12 and 13), and the critical shear stress ( $\tau_c$ ) of the regolith at the node [Pa],

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

The coefficient  $k$  is an erodibility parameter [ $\text{m Pa}^{-f}$ ]. Stock and Dietrich (2006) showed that  $k$  encapsulates substrate properties. If  $h_e$  is used to represent erosion over geomorphic timescales, 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 modeling 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 Appendix A. The exponent  $f$  controls the nonlinearity of  $h_e$  with shear stress. Many authors (e.g., Chen and 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 Fig. 3) suggest that  $f$  may be less than 1 if  $\tau$  is assumed to vary linearly with flow depth, particularly at flow depths greater than 3 m.

MWR includes two options for defining  $\tau$ : (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  $\rho$  is the density of mass-wasting material (grain and water mixture) [ $\text{kg m}^{-3}$ ],  $g$  is gravity [ $\text{m s}^{-2}$ ],  $h$  is the adjusted flow depth described in Eq. (3), and  $\theta$  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):

$$\tau = \sigma \tan \varphi, \quad (13)$$

where  $\sigma$  is normal stress [Pa], and  $\varphi$  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  $\sigma$  as

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

where  $v_s$  is the volumetric solid concentration,  $\rho_s$  is density of the solids [ $\text{kg m}^{-3}$ ],  $u$  is flow velocity [ $\text{m s}^{-1}$ ],  $z$  is depth below the flow surface [m],  $du/dz$  is the shear strain rate [ $\text{s}^{-1}$ ], 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

$$\frac{du}{dz} = \frac{u}{h}. \quad (15)$$

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

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

where  $u^*$  is shear velocity ( $\sqrt{gh \tan \theta}$ ). Substituting Eqs. (16), (15), (14), and (13) into Eq. (11) yields a grain-size-dependent approximation for  $h_e$  that mimics the nonlinear erosion response to flow depth in Schürch et al. (2011). Additionally, this form of  $\tau$  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 Sect. 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 outgoing flux per iteration,  $q_O$  [m], is determined as (see Fig. 3c)

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

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

$$\Delta\eta = A - E. \quad (18)$$

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 ( $\xi_D$ ) is determined as

$$\xi_D = \frac{\xi_D * q_D}{q_I}, \quad (19)$$

where  $q_D$  is a vector containing each  $q_{D_j}$  sent to the node,  $\xi_D$  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 Eq. (2).

Second, the attribute value sent from a node to its receiver nodes ( $\xi_R$ ) is determined as

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

where  $\xi_{t-1}$  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 ( $\xi$ ), is

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

Regolith thickness ( $h_r$ ) and topographic elevation ( $\eta$ ) are updated at a node as

$$\eta = \eta_{t-1} + \Delta\eta, \quad (22)$$

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

where  $\eta_{t-1}$  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 needed for implementing Eq. (1) in the next model iteration, is recomputed.

As the DEM is updated following each model iteration, topographic pits or flat topography may form. These features have no slope or slope inwards and obstruct debriton movement. To allow a debriton to pass an obstruction, we rely on a simple workaround: upon encountering the obstruction, the debriton is directed to itself and some portion of the debris is deposited based on Eq. (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 neighbor nodes, allowing the debriton to move downslope.

### 3 Calibration and MWR Probability

#### 3.1 Calibration utility

MWR includes a calibration utility that uses an adaptive Bayesian Markov chain Monte Carlo (MCMC) algorithm described by Coz et al. (2014) and Renard et al. (2006). The calibration utility determines (1) a single set of parameters that best match MWR output to an observed landslide runout dataset and (2) posterior parameter probability distribution functions (PDFs). The observed runout dataset can consist of a single landslide 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 PDFs that calibrate the MWR to a specific site. Then, the calibration utility randomly selects a set of trial parameter values ( $\Lambda$ ) from the prior PDFs and runs MWR using  $\Lambda$ . Once the model has completed the run, the algorithm evaluates the posterior likelihood of the parameter set ( $L(\Lambda)$ ) as the product 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 ( $\Lambda_{t+1}$ ) by jumping some distance (described below) from each parameter in  $\Lambda$  space. Depending on the value of  $L(\Lambda_{t+1})$ , the algorithm either stays at  $\Lambda$  or moves to  $\Lambda_{t+1}$ . This Markov process is repeated a user-specified number of times. The jump direction is random, but the algorithm is adaptive because the jump distance changes depending on whether  $L(\Lambda_{t+1}) > L(\Lambda)$  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

$$L(\Lambda) = p(\Lambda) \cdot \Omega_T \cdot \frac{1}{\Delta\eta_E^2} \cdot \frac{1}{Q_{SE}^2}, \quad (24)$$

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

$$\Omega_T = \frac{\alpha - \beta - \gamma}{\alpha + \beta + \gamma} + 1, \quad (25)$$

where  $\alpha$ ,  $\beta$ , and  $\gamma$  are the areas of matching, overestimated, and underestimated runout extent, respectively.

The index  $\Delta\eta_E$  is determined as

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

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 modeled runout extent.  $\Delta\eta_{Mi}$  and  $\Delta\eta_{Oi}$  are the modeled and observed topographic change [m] at the  $i$ th node within that extent.

To calculate  $Q_{SE}$ , we first determine the cumulative sediment transport 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):

$$Q_s = -\Delta x^2 \sum_{i=1}^{u_j} \Delta\eta_{i,j}, \quad (27)$$

where  $\Delta\eta_{i,j}$  is the topographic change [m] at the  $i$ th node located above the elevation of node  $j$ , and  $u_j$  is the total number of nodes located above node  $j$ .  $Q_s$  is computed for both the observed and modeled runout ( $Q_{sO}$  and  $Q_{sM}$ , respectively) and  $Q_{SE}$  is determined as

$$Q_{SE} = \sqrt{\frac{\frac{1}{r} \sum_{j=1}^r (Q_{sO} - Q_{sM})^2}{Q_{sO}^2}}, \quad (28)$$

where  $r$  is the number of nodes along the runout profile, and  $Q_{sO}$  is the observed mean cumulative flow.

As will be detailed in Sect. 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 PDFs of potential  $S_c$  and  $q_c$  values from field and/or 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 produce posterior PDFs of  $S_c$  and  $q_c$  and find a single  $S_c$  and  $q_c$  pair that allows the modeled DoD to best replicate the observed DoD (i.e., the  $S_c$  and  $q_c$  pair with the highest  $L(\Lambda)$  value).

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 did not account for 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  $\Omega_T$  or variations of  $\Omega_T$  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  $\Omega_T$  (i.e., runout extent). However, we found that calibration based on  $\Omega_T$  alone results in high parameter equifinality (e.g., Beven 2006); multiple parameter sets result in an equally calibrated model as evaluated by  $\Omega_T$ . As such, we recommend calibrating debris flow and 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 Supplement.

### 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 number of times ( $N_p$ ), 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 based on field and/or remote sensing observations; (2) a user-defined hillslope that is 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, modeled runout probability represents uncertainty in MWR parameterization. If using Option 2 or 3, modeled 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  $N_p$  model simulations,  $N_p$  different versions of the post-runout landscape are created, and the probability of runout at each node is determined as

$$P(\Delta\eta) = \frac{\text{num}(|\Delta\eta| > 0)}{N_p}, \quad (29)$$

where  $\text{num}(|\Delta\eta| > 0)$  is the number of times topographic elevation at a node changes as a result of erosion or deposition from the  $N_p$  model runs. Probability of erosion or aggradation can be determined by replacing the numerator in Eq. (29) with  $\text{num}(\Delta\eta < 0)$  or  $\text{num}(\Delta\eta > 0)$ , 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 planar 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, and (f) a convex-up variable-convergence slope that widens (convergence decreases) in the downslope direction. On each terrain, a 30 m wide, 50 m long, and 3 m deep landslide is released from the top of the terrain. All six terrains are covered by a 1 m 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 Fig. 5.

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 material identical to Terrain A, but near the slope break, the topographic constriction forced flow depth to increase and exceed  $q_c$ , minimizing the formation of levees (because  $q_0 > q_c$ ), 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 upslope of the bench. Upon intersecting the flat surface at the base of the hillslope, the runout material was deposited.

On Terrain D, the landslide and its runout were confined to the center of the convergent terrain and were 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 the 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, the 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 respond 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, modeled flow depth increases when convergence increases (e.g., Terrain B). Where the flow encounters broadly convergent or planar slopes, lateral levee deposits form, a common feature of landslides reported in the literature and at sites reported here (see Sect. 5) that detailed mechanistic models can struggle to reproduce (e.g., Barnhart et al., 2021).

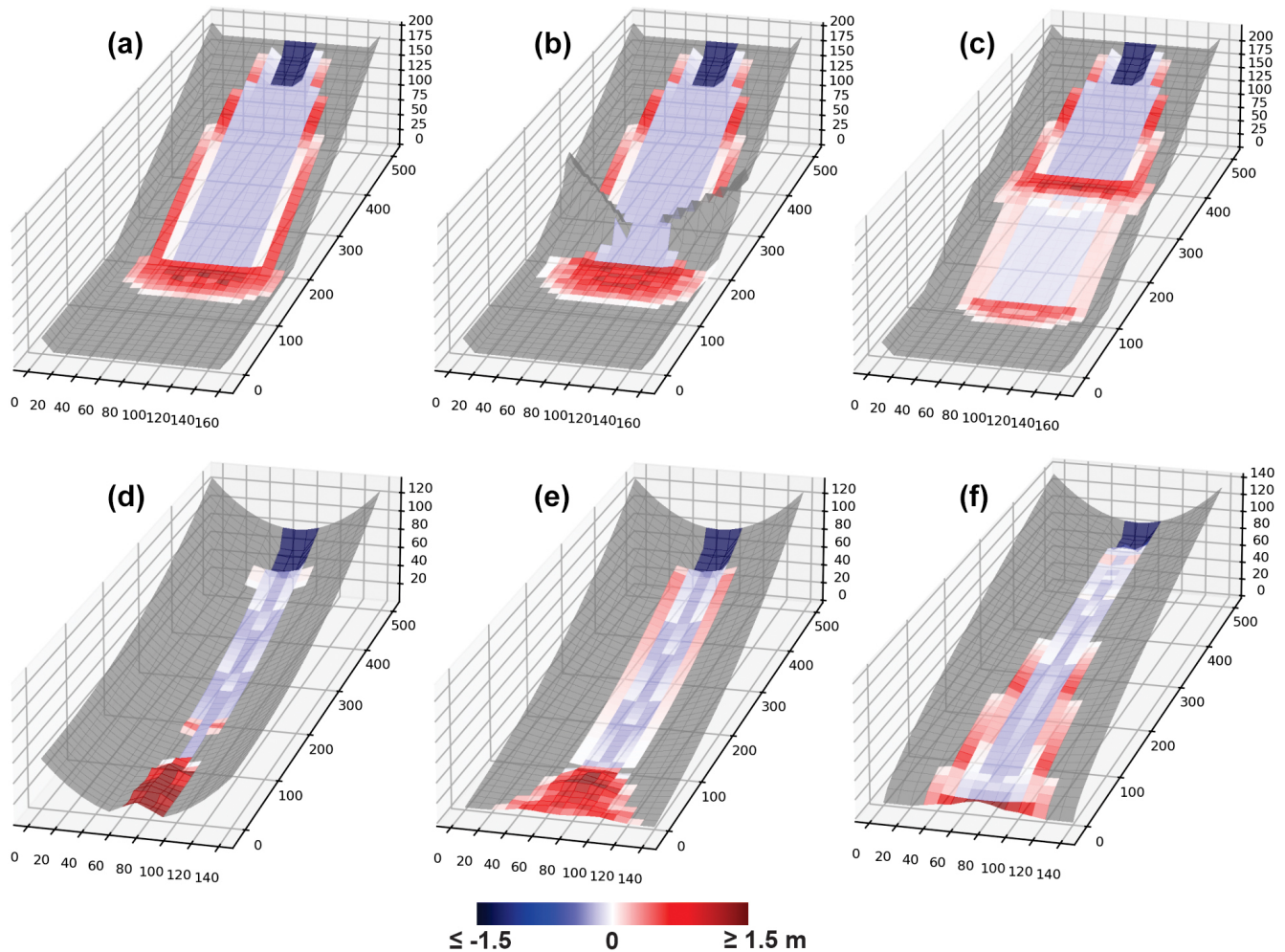
We do not attempt to compare MWR modeled flow with the output of shallow-water-equation-based models or observed granular flows (e.g., Medina et al., 2008; McDougall and Hungr, 2004; Iverson and Denlinger, 2001; Han et al., 2015). The cellular-automaton representation in MWR does not model the time-dependent evolution of debris flow velocity and depth, and it conceptually moves debris instantaneously at each iteration, as driven by changes in the evolving topographic elevation field. Because of that, only the final outcome (modeled 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 Sect. 2.3, the behavior of the multiple flow direction algorithm does vary with grid size. Using a coarser or finer grid, without adjusting model parameterization, could potentially change the runout patterns shown in Fig. 5.

## 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 satisfac-





**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 it visible in the figure. Grid size is 10 m.

tory 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 Mountains and Olympic Mountains sites, include two separate landslides and subsequent runout and we test model predictive ability at these sites in Sect. 6.

Calibrated model performance is demonstrated at the following field sites (see Fig. 6a for locations and observed runout extent): (1) two runout events over the same hill-slope 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 (Fig. 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 (Fig. 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 (Fig. 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; Fig. 1d). All landslides were initiated during heavy rainfall or rain-plus-snowmelt storm events (WRCC, 2024; NRCS, 2022; Table 1), but their runout varied in terms of erosion rate, grain size (Fig. 6b), depositional

behavior (Fig. 6c), and the topographic convergence of the underlying terrain.

## 5.2 Model setup and field parameterization

Each model was set up on a 10 m grid representation of the pre-event DEM with either a uniform or spatially varying regolith thickness (detailed for each site in the Supplement). The length ( $l$ ) and area of the initial mass-wasting source material (e.g., the initial landslide body) were interpreted from a combination of lidar DEM, air photos, and field observations. The average depth of the initial landslide body was measured in the field or from the DoD. The volume of the initial landslide body was determined as the area times the average depth. An average width was determined as the area divided by the length. At the Olympic Mountains site, where the observed runout pattern formed as a result of multiple landslides (see the Supplement), landslide depth and width values listed in Table 1 are average values, and landslide length, area, and volume values are the average cumulative value upstream of each runout path. At all locations, we use Eq. (13) to approximate shear stress. We field-surveyed each site, noting the maximum flow depth (inferred from initial landslide body volume and height of scour marks as well as width of the channel in the erosion zone), 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 from 400 to 110 000 m<sup>3</sup> 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 for the Cascade Mountains 2022 landslide and as high as 0.97 for the Olympic Mountains 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 Mountains site (Fig. 6b, Table 1). Values of  $\bar{E}/l$  ranged from 0.007 to 0.041 [m m<sup>-1</sup>], with the highest rate occurring for the Rocky Mountains landslide and the lowest at the Black Hills sites. Details on grain size samples and data collected in the field are described in the Supplement. 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<sup>3</sup> m<sup>-1</sup> at the Black Hills south site to 95 m<sup>3</sup> m<sup>-1</sup> at the Rocky Mountains site (Table 1).

The median values of topographic slopes at which observed deposition occurred (i.e.,  $\Delta\eta > 0$ , inferred from the DoD) 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 (Fig. 6c). The slope of channel reaches where net deposition (cumulative erosion and deposition; e.g., Guthrie et al., 2010, inferred from field observations) was positive tended to be lowest at the Black Hills site ( $< 1\%$  to  $10\%$ ) and highest at the Rocky Mountains 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 slopes 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 from either repeated lidar or field observations as detailed in the Supplement.

## 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 Fig. 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 h for the Cascade Mountains 2022 landslide to 6 h for the Olympic Mountains site 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 (Fig. 7a), which permitted thinner flows over lower slopes and effectively made the 2009 modeled runout more mobile relative to the 2022 modeled runout (Fig. 7b). At the Rocky Mountains site (Fig. 7e), a relatively high  $q_c$  value helps limit the lateral extent of the modeled runout that in the observed runout was controlled by standing trees (see the Supplement).

Profile plots of modeled  $Q_s$  and maps of the modeled planimetric runout extent, colored to indicate where the runout matched ( $\alpha$ ), overestimated ( $\beta$ ), or underestimated ( $\gamma$ ) the observed runout, are shown in Fig. 8. Values of  $\Omega_T$  we obtained with MWR are comparable to or higher than reported values of  $\Omega_T$  in the literature that used a variety of models (Gorr et al., 2022; Barnhart et al., 2021; note that

**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, $l$ [m] <sup>a</sup>	185	55	80	75	40	45
Initial landslide body width [m] <sup>a</sup>	86	53	13	69	25	15
Initial landslide body depth [m] <sup>a</sup>	6.9	7.5	1.4	3.6	4.6	1.5
Initial landslide body vol. [m <sup>3</sup> ] <sup>a</sup>	110 000	22 000	1500	18 500	4600	1025
2 d 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]	4	2	2	3	3	3
Average channel slope in scour zone [m m <sup>-1</sup> ]	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, $A_e$ [m <sup>2</sup> ]	28 400	6600	22 800	52 400	20 800	28 900
Erosion volume, $\sum E \Delta x^2$ [m <sup>3</sup> ] <sup>b</sup>	44 547	5125	12 332	26 815	34 275	33 725
Average erosion per unit length of runout debris, $\bar{E}/l$ [m m <sup>-1</sup> ]	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 <sup>3</sup> m <sup>-1</sup> ]	74.2	15.1	10.2	19.9	95.2	13.2
Average observed $ \Delta\eta $ [m]	2.4	2.2	0.53	0.63	0.89	1.4
Total erosion volume/total mobilized volume <sup>c</sup>	0.29	0.19	0.89	0.59	0.88	0.97

<sup>a</sup> For the Olympic Mountains site, width and depth are average values and length and volume are defined as the average cumulative value upstream of each runout path. <sup>b</sup> Excludes landslide volume. <sup>c</sup> Total mobilized volume = initial landslide body + erosion volume.

to compare  $\Omega_T$  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. 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 as an order of magnitude (e.g., Gartner et al., 2014; Barnhart et al., 2021). Model performance in predicting cumulative sediment transport along the runout profile was within similar error ranges. Except for the Rocky Mountains site where MWR consistently modeled wider-than-observed flow, the mean cumulative sediment transport error along the runout profile ( $Q_{SE}$ ) was limited to between 5 % and 19 % of the mean cumulative flow determined from the observed DoD.

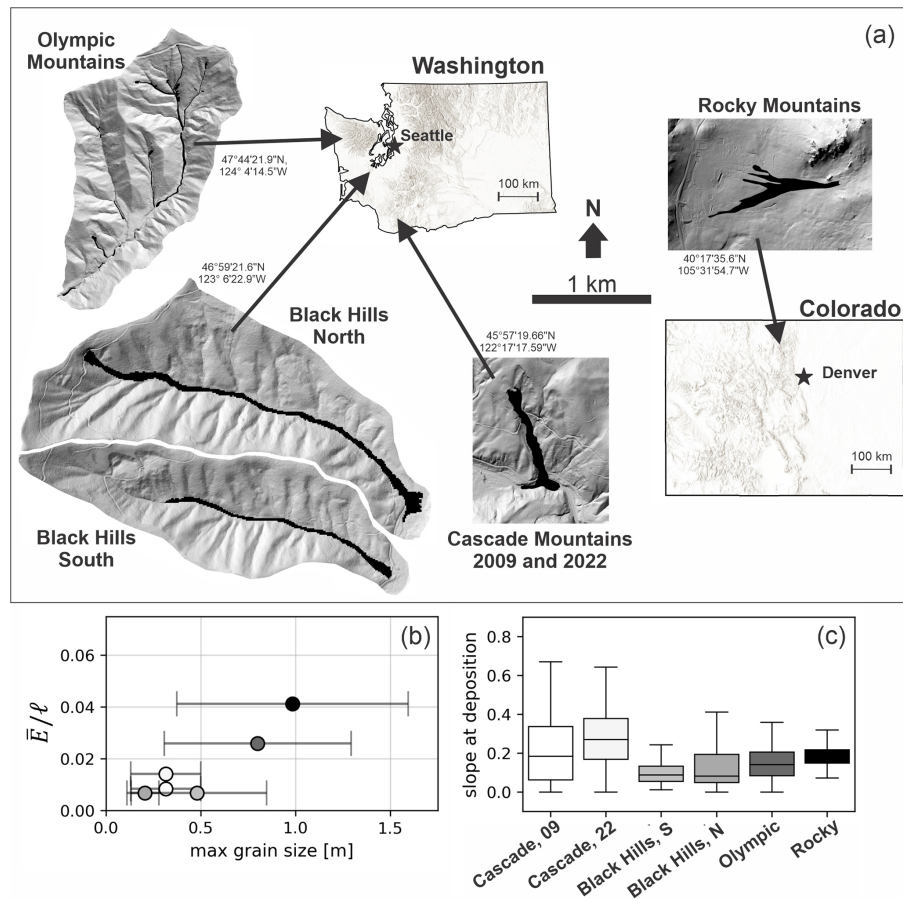
MWR generally successfully replicates observed sediment transport along the runout path via model parameterizations that are unique to each landslide. For example, the profile plots of  $Q_s$  at the Cascade Mountains site (Fig. 8a and b) show that during the 2009 landslide, all of the runout material flowed past the first 750 m of the runout path. During the 2022 landslide, material began to deposit just downslope of the initial landslide source area, as both observed  $Q_s$  and modeled  $Q_s$  reverse slope, indicating a decrease in cumulative flow. 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 the Supplement).

MWR also successfully replicates the observed sediment transport patterns at the Olympic Mountains site (profile plot of  $Q_s$  in Fig. 8f) and to a lesser degree the Rocky Mountains site (Fig. 8e). This finding is notable because at the Olympic Mountains and Rocky Mountains sites, observed runout extent and sediment depositional patterns were heavily impacted by woody debris or standing trees (see the Supplement).

Using a fixed grid size of 10 m might have impacted model performance in some areas like the Olympic Mountains and Cascade Mountains 2022 sites where MWR tended to overestimate the runout width (yellow zones in Fig. 8f and b). While a 10 m DEM is generally accepted as a good balance between model detail and computational limitations (e.g., Horton et al., 2013), for small landslides, the 10 m grid was close to the width of the channels that controlled observed runout (see Fig. 1d) and may not have accurately represented the terrain. Thus, modeled flow was less topographically constrained and tended to flow over a wider area than observed.

Also, 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, for the Cascade Mountains 2009 slide, MWR underestimates a short section of slope-perpendicular flow over a bench. Review of model behavior for this slide (Fig. 9) shows how MWR successfully mimics diverging flow around a broad ridge upslope of the bench (iteration  $t = 28$  in Fig. 9) but afterwards continues to follow topographic slope and converges too rapidly into





**Figure 6.** (a) Landslide locations in Washington and Colorado. 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 ( $\bar{E}/l$ ) relative to the observed average maximum grain size. Error bars indicate standard deviation. (c) Underlying topographic slope of observed deposition.

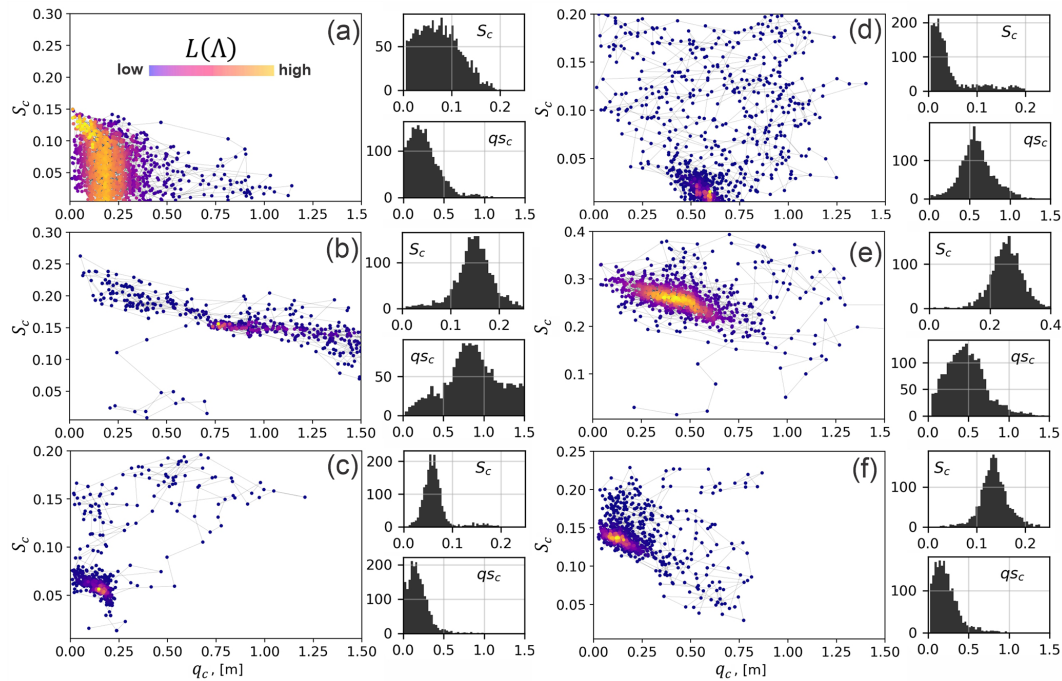
a narrow ravine along the west edge of the bench (iteration  $t = 40$  in Fig. 9; compare to runout scar in the air photo and underestimated region on the topographic bench in Fig. 8a).

Nonetheless, overall, calibration was best for the Cascade Mountains 2009 landslide (values of  $\Omega_T$  are highest and values of  $\Delta\eta_E$  and  $Q_{SE}$  are lowest) and poorest at the Rocky Mountains and Olympic Mountains sites (values of  $\Omega_T$  are lowest and  $Q_{SE}$  and  $\Delta\eta_E$  are highest). At both the Rocky Mountains and Olympic Mountains 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. At most sites, we used a uniform thickness. As regolith thickness limits maximum erosion depth (i.e., Eq. 10), using a spatially accurate regolith thickness may improve model performance. Finally, except for the Rocky Mountains site, where topography was unusually planar and the model seemed to consistently overestimate flow width, at most sites, MWR

does not appear to have a strong systematic bias in modeled output, which suggests that when applied to convergent terrain, MWR may not have any structural weaknesses. In the next section, we evaluate model performance relative to runout path topography.

#### 5.4 Runout path topography and model performance

Model behavior at the Rocky Mountains site suggests that MWR performance may systematically vary with topography (e.g., it may not perform as well on planar hillslopes). To check for systematic model variation, we compared model performance with three topographic indices described by Chen and Yu (2011). The indices are computed from the terrain in the observed runout extent and include the relief ratio ( $H/L$ ), mean total curvature ( $\kappa$ ), and 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 initial landslide to the end of the runout path) divided by the horizontal length of the runout, and indicates the mobil-



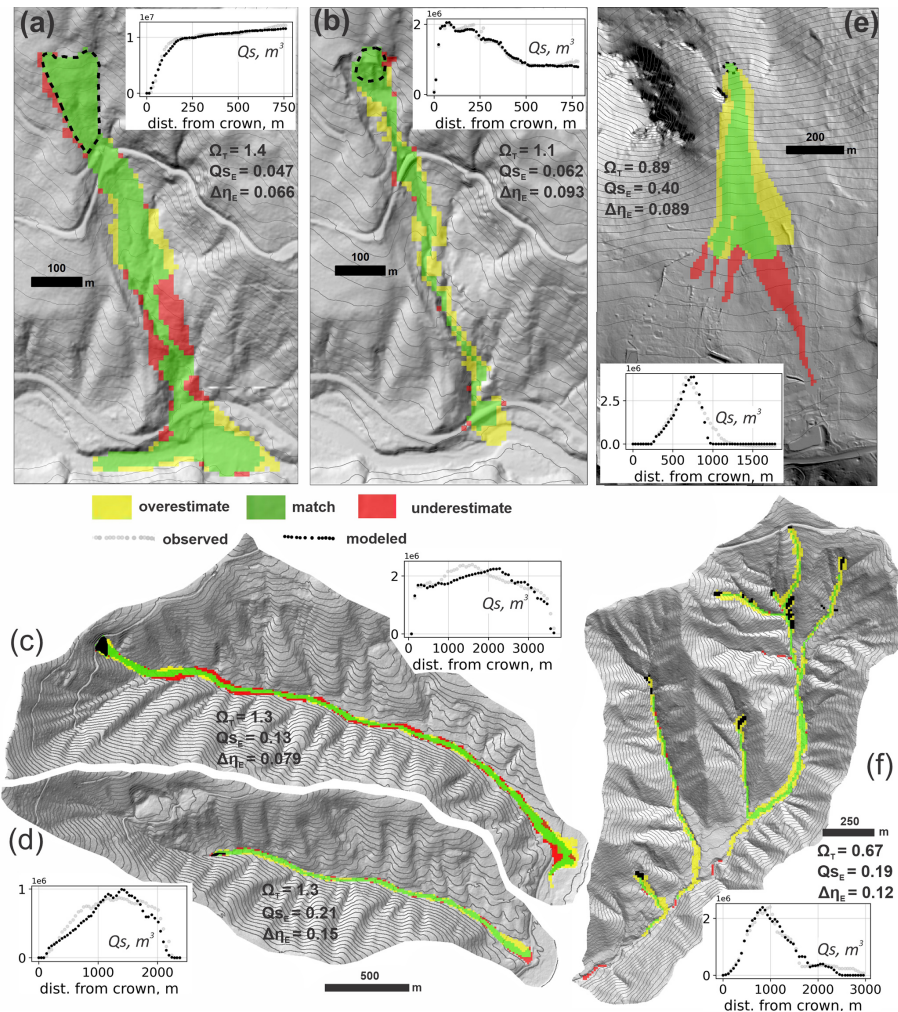
**Figure 7.** MWR calibration results for (a) Cascade Mountains 2009, (b) Cascade Mountains 2022, (c) Black Hills south, (d) Black Hills north, (e) Rocky Mountains, and (f) Olympic Mountains. Each result shows a scatter plot of the sampled  $S_c$  and  $q_c$  values, colored by their relative  $L(\Lambda)$  value. The line between points indicates the jump direction. Note that the 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, posterior PDF of the possible  $S_c$  and  $q_c$  values that calibrate MWR to the site. The histogram y axis is the count and the x axis is  $S_c$  or  $q_c$ , as indicated on the histogram.

ity of the runout. The index  $\kappa$  represents topographic convergence, which is the second derivative of the terrain surface, with increasingly positive values of the index  $\kappa$  reflecting growing topographic convergence and a concave-up channel profile (e.g., Istanbuloglu et al., 2008). The index SPI is determined as the natural log of the product of the contributing area and slope. Indices  $\kappa$  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 Fig. 10 shows slightly improved model performance over runout paths that are less convergent (lower SPI and  $\kappa$  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 modeled sediment transport and topographic change ( $Q_{SE}$  and  $\Delta\eta_E$ ) replicated observed change, as there does not appear to be a trend in  $\Omega_T$  with  $H/L$  and the two best-performing models (both Cascade Mountains landslides) had the lowest (best)  $Q_{SE}$  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 does not quite capture the relief of channelized topography that controlled observed runout, leading to modeled runout that was considerably wider than observed and causing low  $\Omega_T$

values (this is especially true at the Olympic Mountains site; Fig. 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). While model performance at the Rocky Mountains site suggests that MWR may not perform as well on planar hillslopes, 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, the accuracy of the regolith map, and the importance of processes not explicitly included in the model that also impact performance.



**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 Eq. 27), and reported values of  $\Omega_T$ ,  $\Delta\eta_E$ , and  $Q_{SE}$ . The y-axis label for profile plots of  $Q_s$  is indicated. In all maps, up is north, except in (e), where 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.

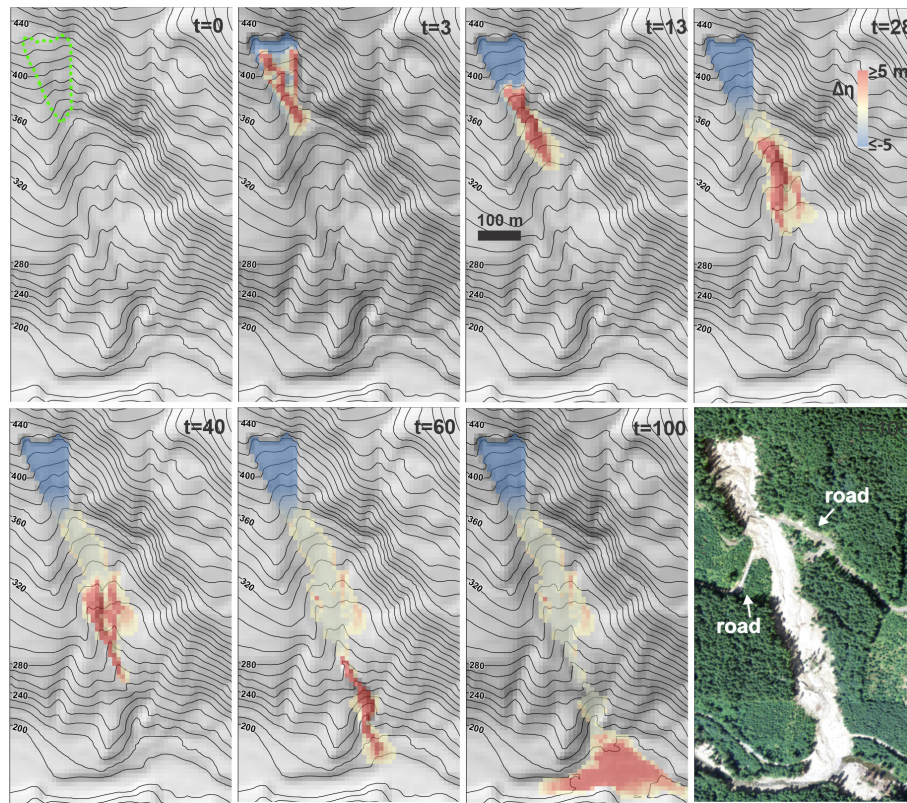
## 6 Discussion

### 6.1 Strategic testing of MWR for hazard mapping applications

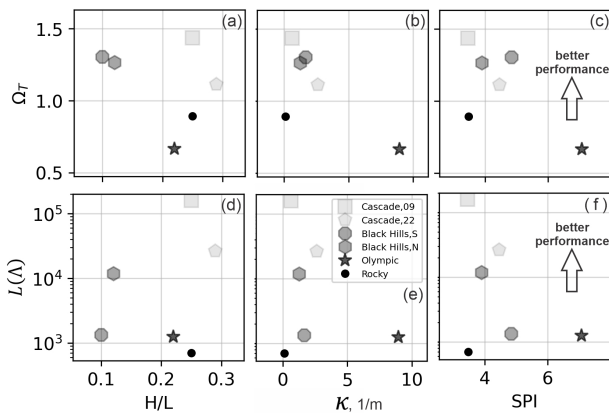
Having demonstrated calibrated model performance for a variety of landslides and runout terrains, we now strategically test MWR using the Cascade Mountains and Black Hills sites. Since both of these sites include two separate landslides, we can thus test model performance by swapping best-fit model parameters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

As shown in Fig. 11, at three of the landslides (both the Cascade Mountains landslides and the Black Hills north landslide), when the best-fit parameters from the other landslide are used to predict runout, the accuracy of modeled runout planimetric extent drops but resultant  $\Omega_T$  values can still be as high as or higher than values reported in other studies (compare to equivalent  $\Omega_T$  values in Gorr et al., 2022, and Barnhart et al., 2021). In terms of modeled sediment transport and topographic change, swapping best-fit parameters has a more substantial effect. For the Cascade Mountains 2009 landslide, using the 2022 best-fit parameter values causes about half of the modeled runout material to be prematurely deposited on the hillslope, reducing the amount of sediment that reaches the valley floor ( $Q_{SE}$  increases by a factor of 9; Fig. 11). Using the Cascade Mountains 2009 parameter values on the Cascade Mountains 2022 landslide





**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. The air photo in the last panel shows observed runout extent. Note that the upper road is not part of the observed landslide runout path.



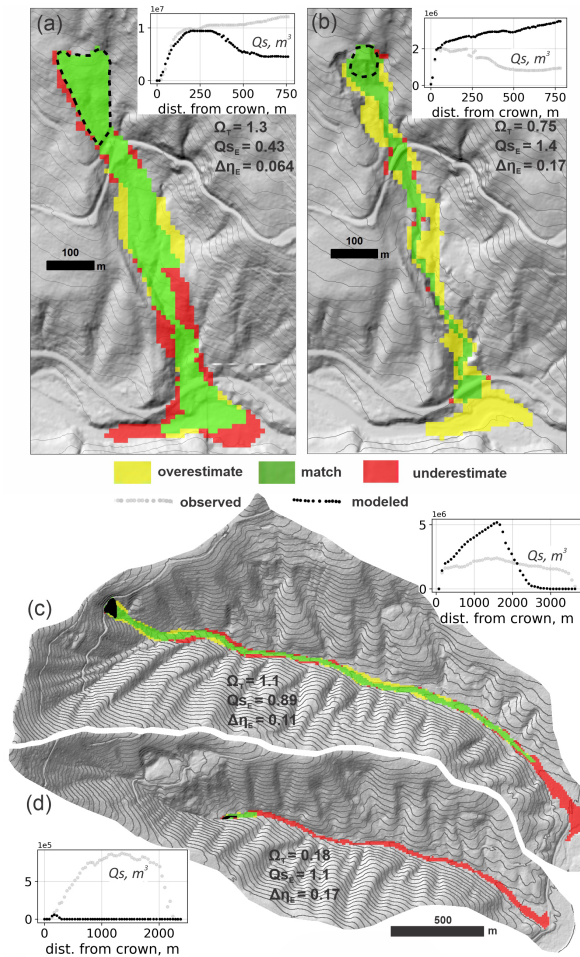
**Figure 10.** Illustration of model calibration, as reflected by the posterior parameter likelihood  $L(\Lambda)$  and planimetric fit ( $\Omega_T$ ) relative to topographic indices. There is no strong trend between the topographic indices and calibration performance. Note that the curvature values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

(Fig. 11b) increases modeled runout extent and results in nearly 4 times the entrainment and transport of sediment to the valley floor, causing  $Q_{SE}$  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  $Q_{SE}$  and  $\Delta\eta_E$  to increase by 83 % and 39 %, respectively (Fig. 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 landslide best-fit parameters applied to the south landslide causes the model to entrain too little and stop prematurely (Fig. 11d).

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 the probability of runout at each model node (Eq. 29). As shown in Fig. 12, similar to the first test, for three of the landslides, using the parameter distribution associated with the neighboring landslide results in relatively minor changes in model output, as indicated by where runout is more likely to occur versus not occur (probability of runout  $\geq 50\%$ ; Fig. 12a, b, and d), but at the Black Hills south landslide, swapping parameter PDFs causes a large change in runout probability (Fig. 12c).

The results of these two tests suggest that site-specific or even landslide-type-specific calibration may be needed to



**Figure 11.** Model performance using the neighboring landslide parameter values, as indicated by maps of modeled runout extent, profile plots of  $Q_S$ , and values of  $\Omega_T$ ,  $\Delta\eta_E$ , and  $Q_{SE}$ . Compare with Fig. 8. (a) Cascade Mountains, 2009; (b) Cascade Mountains, 2022; (c) Black Hills, north; (d) Black Hills, south.

accurately predict runout behavior using MWR, especially when the user aims to apply MWR to sediment yield analyses and the study site consists of numerous landslide types, like the Black Hills and Cascade Mountain sites. While the need for calibration of MWR may limit model 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.

We suspect that in regions where landslide processes are relatively uniform (like the Olympic Mountains site), parameterization determined through calibration to one landslide might be transferable across sites. Additionally, as noted in Sect. 3.1, we found that 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 air photos that show the slope of past landslide deposits (for  $S_c$ ) and deposit thickness (for  $q_c$ ) might be sufficient.

## 6.2 MassWastingRunout probability applications

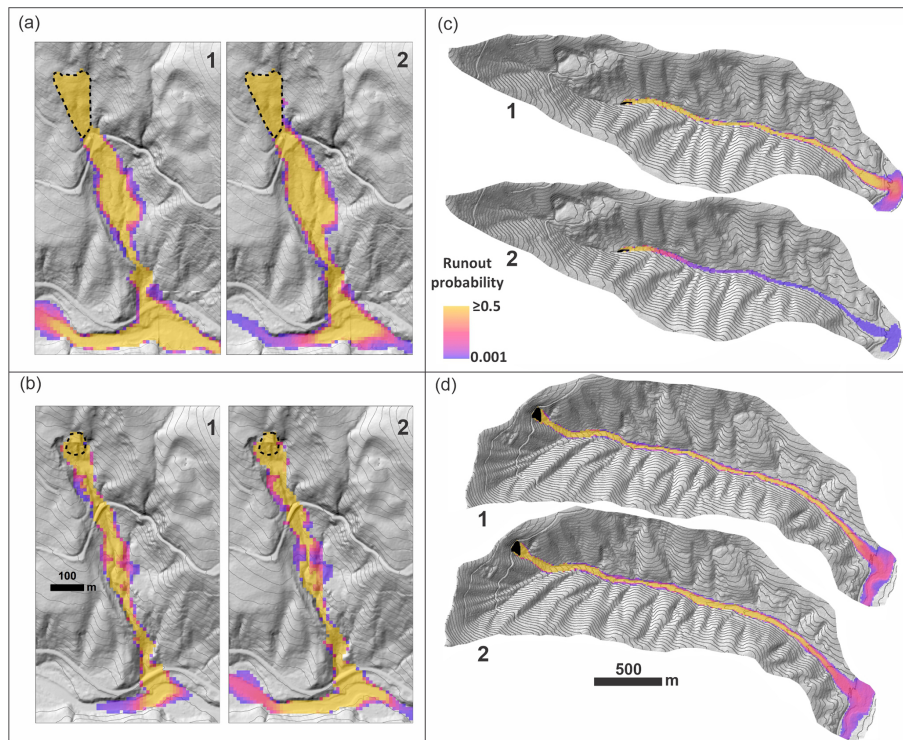
In this section we briefly demonstrate how to determine runout probability from a probabilistically determined landslide hazard map or a specific, potentially unstable slope using MWR. The first application may be appropriate for watershed- to regional-scale runout hazard assessments. The second application is an example hazard assessment for a potentially unstable hillslope. Both applications are demonstrated at the Olympic Mountains 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 (Strauch et al., 2018), an existing component in Landlab that computes landslide probability by iteratively calculating the factor of safety (FS: ratio of the resisting to the driving forces) at each node on the raster model grid  $N_p$  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  $N_p$ .

We first ran LandslideProbability using a 50-year precipitation event (WRCC, 2024) to determine landslide probability (Fig. 13a) over the entire Olympic Mountains model domain and create the series of  $N_p$  FS maps. Details on the LandslideProbability setup are included in the Supplement. We then read the series of  $N_p$  FS maps into MWR Probability, treating all nodes with  $FS < 1$  as a landslide source, and ran MWR  $N_p$  times. Each iteration, MWR read a new





**Figure 12.** Model tests by swapping parameter PDFs and comparing runout probability at the (a) Cascade Mountains 2009, (b) Cascade Mountains 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.

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, which reflects MWR parameter uncertainty (as illustrated in Fig. 7f) and uncertainty in the initial landslide size and location caused by a 50-year precipitation event, is illustrated in Fig. 13b and shows that the probability of runout is high in many of the second-order channels but low at the basin outlet. As discussed in Sect. 3, the probability of aggradation or erosion caused by the runout can be determined by adjusting the numerator of Eq. (29). As an example, the probability of deposition greater than 1 m is shown in Fig. 13c.

### 6.2.2 Runout probability for a specific, potentially unstable slope

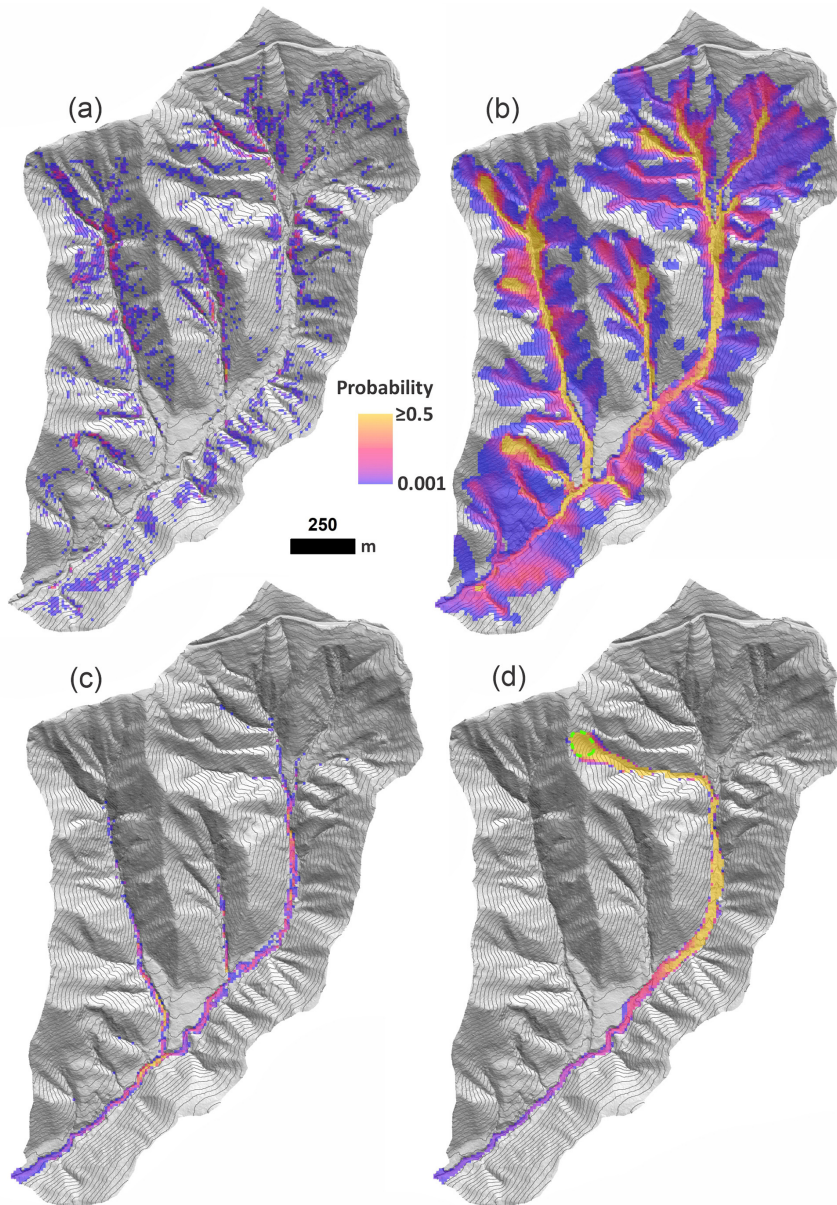
When field evidence or other data indicate that a specific hill-slope may be potentially unstable but the exact area of a potential landslide on that slope is unknown, MWR can be used to generate a hazard estimate that takes into account the uncertainty in the landslide area. For this application, MWR 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 (Fig. 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 %.

## 7 Concluding remarks

In this study, we described, calibrated, and tested Mass-WastingRunout (MWR), a new cellular-automaton landslide runout model that combines the functionality of simple runout algorithms used in WSMs (landscape evolution and watershed sediment yield models) 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 modeling toolkit and is designed for WSM and probabilistic landslide hazard assessment 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 (PDFs) 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.





**Figure 13.** Olympic Mountains 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 (dashed green polygon).

Given a DEM and map of approximate regolith depth, MWR needs only the location and geometry of an initial landslide source area to model the entire runout process. 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. 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 (Sect. 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 yield 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 rule-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 (Sect. 4). 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 modeled wider-than-observed flow, the cumulative flow error along the runout profile ( $Q_{SE}$ ) was 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 assumes an order-of-magnitude range of confidence. A notable finding of this paper is that at most sites, 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).

Finally, we briefly showed how to couple MWR with the landslide initiation model `LandslideProbability` to map debris flow hazard when the initial landslide location is uncertain. As a component of the `Landlab` earth surface modeling toolkit, MWR is designed to be compatible with other models and thus relatively easy to integrate into a WSM. An example WSM that incorporates MWR might include models for landslide initiation, hillslope diffusion, and fluvial incision to investigate the role of landslides and their runout in long-term landscape evolution. Future studies will explore large-scale application in landscape evolution or sediment yield models and characterize model parameters for different landslide types 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 ( $A_e$ ):

$$\bar{E} = \frac{\sum E \Delta x^2}{A_e}, \tag{A1}$$

where  $\sum E \Delta x^2$  and  $A_e$  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

$$\bar{E} = \bar{n} \bar{h}_e. \tag{A2}$$

An estimate for  $\bar{n}$  can be determined from the average length of the runout material divided by the cell width:

$$\bar{n} = \frac{l}{\Delta x}. \tag{A3}$$

At most sites, we approximate the average length of the runout material simply as the mapped landslide length ( $l$ ). 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, the numerator of Eq. (A3) can be multiplied by a coefficient to better match runout length. Combining Eqs. (A2) and (A3),  $\bar{h}_e$  can be defined as the average erosion rate per unit length of runout debris ( $\bar{E}/l$ ) times the cell width:

$$\bar{h}_e = \frac{\bar{E} \Delta x}{l}. \tag{A4}$$

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

$$k = \frac{\bar{h}_e}{\bar{\tau} f}. \tag{A5}$$

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 (Table 1). We used Eq. (13) to define  $\bar{\tau}$ . For  $D_s$ , we used the average maximum grain size observed over the whole runout path. If  $\tau$  is defined as a function of the grain-collision-dependent shear stress approach in Eq. (13) and  $k$  is determined as a function of  $f$ , as in Eq. (A5), the impact of  $f$  on model behavior is relatively small.

## Appendix B

Table B1. Notation.

Parameter	Unit	Description
$q_{R_i}$	[m]	debris flux from a node to each of the node $i$ th receiver nodes
$i$		variable used to represent count or index in Eqs. (1), (5), and (27)
$q_O$	[m]	the total outgoing debris flux
$N_r$		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 Eq. (1) that controls how flow is distributed to downslope nodes
$q_I$	[m]	the total incoming flux
$N_d$		number of donor nodes to a node
$q_{D_j}$	[m]	the flux from node $D_j$ (the $j$ th donor node)
$j$		variable used to represent count or index in Eqs. (2) and (28)
$h$	[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 neighboring nodes
$\Delta x$	[m]	cell length
$A_{p, N_a}$	[m]	potential aggradation depth that forms a deposit spread over $N_a$ consecutive nodes
$A_{p, i}$	[m]	$i$ th deposition amount in the deposit illustrated in Fig. 4
$N_a$		number of nodes $q_s^i$ is assumed to spread over
$E$	[m]	erosion depth
$h_r$	[m]	regolith depth
$h_e$	[m]	potential erosion depth
$\theta$	[°]	topographic slope used to determine shear stress, equal to $\tan^{-1}(S)$
$\tau$	[Pa]	basal shear stress
$\tau_c$	[Pa]	critical shear stress of the regolith
$k$		erodibility parameter in Eq. (11)
$f$		exponent, controls the nonlinearity of $h_e$ in Eq. (11)
$\rho$	[kg m <sup>-3</sup> ]	density of runout material
$\sigma$	[Pa]	normal stress at basal surface
$\varphi$		tangent of collision angle between grains, measured from the vertical axis
$u_s$		volumetric solids concentration
$\rho_s$	[kg m <sup>-3</sup> ]	density of solids
$D_s$	[m]	characteristic particle diameter
$u$	[m s <sup>-1</sup> ]	depth average flow velocity
$z$	[m]	depth below the flow surface
$u^*$		shear velocity
$g$	[m s <sup>-1</sup> ]	acceleration due to gravity
$\Delta\eta$	[m]	change in elevation at node
$\mathbf{q}_D$		a vector containing each $q_{D_j}$ sent to the node
$\xi_D$		a vector containing the incoming attribute values for each $q_{D_j}$
$\xi_D$		attribute value delivered to the node
$\xi_R$		attribute value sent to receiver nodes
$\xi$		attribute value at node
$\eta$	[m]	topographic elevation
$\Lambda$		parameter set
$L(\Lambda)$		likelihood of parameter set
$p(\Lambda)$		prior probability of parameter set
$\Omega_T$		an index for evaluating model planimetric fit
$\alpha$ [m <sup>2</sup> ]		modeled area of matching extent (compared to observed runout extent)
$\beta$ [m <sup>2</sup> ]		modeled area of overestimated extent



$\gamma$ [m <sup>2</sup> ]		modeled area of underestimated extent
$\Delta\eta_E$		volumetric error of the modeled topographic change relative to the observed total mobilized volume, fraction
$V$	[m <sup>3</sup> ]	observed total mobilized volume
$p$		the number of nodes in the modeled runout extent
$\Delta\eta_{Mi}$	[m]	the modeled 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
$Q_{SE}$		mean modeled cumulative flow error along the runout path relative to the observed mean cumulative flow, fraction
$\Delta\eta_{ij}$	[m]	topographic change [m] at the $i$ th node located upstream of node $j$
$u_j$		total number of all nodes located upstream of node $j$
$r$		the number of nodes along the centerline of the runout path
$Q_s$	[m <sup>3</sup> ]	the cumulative debris flow volume at each node, $j$ along the centerline of the runout path
$Q_{sO}$	[m <sup>3</sup> ]	the observed cumulative debris flow volume ( $Q_s$ ) at each node, $j$
$Q_{sM}$	[m <sup>3</sup> ]	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 $j$
$\frac{u_j}{Q_{sO}}$	[m <sup>3</sup> ]	the observed mean cumulative flow
$\frac{P(\Delta\eta)}{\bar{E}/l}$	[m m <sup>-1</sup> ]	average erosion per unit length of runout debris
$\frac{P(\Delta\eta)}{\text{num}() \cdot N_p}$		probability of runout, expressed as the probability that the elevation of a node changes number of number of Monte Carlo iterations used to determine probability
$A_e$	[m <sup>2</sup> ]	erosion area of the observed or modeled runout
$\bar{E}$	[m]	average erosion depth caused by the runout
$\sum E \Delta x^2$	[m <sup>3</sup> ]	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
$\kappa$	[m <sup>-1</sup> ]	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> (Landlab, 2024).

**Data availability.** Validation field site data, including the landslide source area and DEM of difference, are available at <http://www.hydroshare.org/resource/55813a5e01764546b76641a7385c2236> (Keck, 2024).

**Interactive computing environment.** Jupyter Notebook tutorials that allow the user to recreate and modify the plots in Fig. 5 or model runout at the Cascade Mountains 2022 site can be run online or downloaded to a personal computer. The online tutorial and directions for downloading to a personal computer are available at the Landlab website: <https://landlab.github.io/#/> (year of publication: 2024, Landlab, 2017).

**Supplement.** The supplement related to this article is available online at: <https://doi.org/10.5194/esurf-12-1165-2024-supplement>.

**Author contributions.** JK conceptualized and coded MWR and the calibration and probability utilities, hiked and surveyed each validation site, prepared the validation site model input data, calibrated and assessed model performance, wrote the manuscript, and prepared the figures. EI served as an advisor and helped with model conceptualization, writing, figure formatting, and funding. BC helped with field surveys, model conceptualization, and writing. GT helped with model conceptualization and writing. AHD helped with figure formatting and model conceptualization.

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

**Disclaimer.** Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

**Acknowledgements.** This research was partially supported by the following programs: National Science Foundation (NSF) PREEVENTS program (ICER-1663859), NSF OAC-2103632, and NASA Disasters Program (grant number 80NSSC23K1103). Critical and helpful referee reviews as well as multiple detailed reviews by associate editor Wolfgang Schwanghart significantly improved the manuscript. Early conceptualization of the model as well as development of the  $Q_{SE}$  metric greatly benefited from discussion with Hervé Capart. Discussions with Tzu-Yin Kasha Chen and Chi-Yao Hung led to model improvements. Stephen Slaughter field-reviewed the Cascade Mountains 2009 landslide and the Black Hills landslides the year they occurred and provided photos and field observations that aided author interpretation. John Jenkins helped with field reconnaissance, and Miles Micheletti prepared the structure-from-motion DEM at the Cascade Mountains 2022 landslide. Eli Schwat helped with field reconnaissance at the Olympic Mountains site. This work also benefitted from coding guidance from Eric Hutton and support from the staff and researchers at CSDMS.

**Financial support.** This research has been supported by the National Science Foundation (NSF) PREEVENTS program (grant nos. ICER-1663859 and OAC-2103632) and the NASA Disasters Program (grant no. 80NSSC23K1103).

**Review statement.** This paper was edited by Wolfgang Schwanghart and reviewed by Saskia de Vilder and two anonymous referees.

## References

- Bagnold, R. A.: Experiments on a gravity-free dispersion of large solid spheres in a Newtonian fluid under shear, *P. R. Soc. Lond.*, 225, 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 *Active lavas: Monitoring and modeling*, edited by: Kilburn, C., and Luongo, G., University College of London Press, London, 291–309, <https://doi.org/10.4324/9781003325178-15>, 1993.
- Barnhart, K. R., Hutton, E. W. H., Tucker, G. E., Gasparini, N. M., Istanbuloglu, E., Hobbey, D. E. J., Lyons, N. J., Mouchene, M., Nudurupati, S. S., Adams, J. M., and Bandaragoda, C.: Short communication: Landlab v2.0: a software package for Earth surface dynamics, *Earth Surf. Dynam.*, 8, 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., and Kean, J. W.: Multi-Model Comparison of Computed Debris Flow Runout for the 9 January 2018 Montecito, California Post-Wildfire Event, *J. Geophys. Res.-Earth*, 126, e2021JF006245, <https://doi.org/10.1029/2021jf006245>, 2021.
- Benda, L. and Dunne, T.: Stochastic forcing of sediment supply to channel networks from landsliding and debris flow, *Water Resour. Res.*, 33, 2849–2863, <https://doi.org/10.1029/97wr02388>, 1997.
- Benda, L., Veldhuisen, C. P., and Black, J.: Debris flows as agents of morphological heterogeneity at low-order confluences, Olympic Mountains, Washington, *Geol. Soc. Am. Bull.*, 115, 1110, <https://doi.org/10.1130/b25265.1>, 2003.
- Beven, K.: A manifesto for the equifinality thesis, *J. Hydrol.*, 320, 18–36, <https://doi.org/10.1016/j.jhydrol.2005.07.007>, 2006.
- Burton, A. and Bathurst, J. C.: Physically based modelling of shallow landslide sediment yield at a catchment scale, *Environ. Geol.*, 35, 89–99, <https://doi.org/10.1007/s002540050296>, 1998.
- Bigelow, P., Benda, L., Miller, D., and Burnett, K. M.: On Debris Flows, River Networks, and the Spatial Structure of Channel Morphology, *Forest Sci.*, 53, 220–238, <https://doi.org/10.1093/forestscience/53.2.220>, 2007.
- Campforts, B., Shobe, C. M., Overeem, I., and Tucker, G. E.: The Art of Landslides: How Stochastic Mass Wasting Shapes Topography and Influences Landscape Dynamics, *J. Geophys. Res.-Earth*, 127, e2022JF006745, <https://doi.org/10.1029/2022jf006745>, 2022.
- Campforts, B., Shobe, C. M., Steer, P., Vanmaercke, M., Lague, D., and Braun, J.: HyLands 1.0: a hybrid landscape evolution model to simulate the impact of landslides and landslide-derived sediment on landscape evolution, *Geosci. Model Dev.*, 13, 3863–3886, <https://doi.org/10.5194/gmd-13-3863-2020>, 2020.
- Capart, H. and Fraccarollo, L.: Transport layer structure in intense bed-load, *Geophys. Res. Lett.*, 38, L20402, <https://doi.org/10.1029/2011gl049408>, 2011.
- Capart, H., Hung, C., and Stark, C. R.: Depth-integrated equations for entraining granular flows in narrow channels, *J. Fluid Mech.*, 765, R4, <https://doi.org/10.1017/jfm.2014.713>, 2015.
- Carretier, S., Martinod, P., Reich, M., and Godderis, Y.: Modelling sediment clasts transport during landscape evolution, *Earth Surf. Dynam.*, 4, 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, 39–57, [https://doi.org/10.1016/0169-555x\(92\)90057-u](https://doi.org/10.1016/0169-555x(92)90057-u), 1992.
- Chen, C. and Yu, F.: Morphometric analysis of debris flows and their source areas using GIS, *Geomorphology*, 129, 387–397, <https://doi.org/10.1016/j.geomorph.2011.03.002>, 2011.
- Chen, H. X. and Zhang, L. M.: EDDA 1.0: integrated simulation of debris flow erosion, deposition and property changes, *Geosci. Model Dev.*, 8, 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 Surf. Dynam.*, 11, 325–342, <https://doi.org/10.5194/esurf-11-325-2023>, 2023.
- Clerici, A. and Perego, S.: Simulation of the Parma River blockage by the Corniglio landslide (Northern Italy), *Geomorphology*, 33, 1–23, [https://doi.org/10.1016/s0169-555x\(99\)00095-1](https://doi.org/10.1016/s0169-555x(99)00095-1), 2000.
- Codd, E. F.: *Cellular Automata* (1st ed.), Academic Press, New York, 122 pp., ISBN 0121788504, 1968.

- Coz, J. L., Renard, B., Bonnifait, L., Branger, F., and Bour-sicaud, R. L.: Combining hydraulic knowledge and uncertain gaugings in the estimation of hydrometric rating curves: A Bayesian approach, *J. Hydrol.*, 509, 573–587, <https://doi.org/10.1016/j.jhydrol.2013.11.016>, 2014.
- Crave, A. and Davy, P.: A stochastic “precipiton” model for simulating erosion/sedimentation dynamics, *Comput. Geosci.*, 27, 815–827, [https://doi.org/10.1016/s0098-3004\(00\)00167-9](https://doi.org/10.1016/s0098-3004(00)00167-9), 2001.
- D’Ambrosio, D., Di Gregorio, S., Iovine, G., Lupiano, V., Rongo, R., and Spataro, W.: First simulations of the Sarno debris flows through Cellular Automata modelling, *Geomorphology*, 54, 91–117, [https://doi.org/10.1016/s0169-555x\(03\)00058-8](https://doi.org/10.1016/s0169-555x(03)00058-8), 2003.
- Doten, C. O., Bowling, L. C., Lanini, J. S., Maurer, E. P., and Lettenmaier, D. P.: A spatially distributed model for the dynamic prediction of sediment erosion and transport in mountainous forested watersheds, *Water Resour. Res.*, 42, W04417, <https://doi.org/10.1029/2004wr003829>, 2006.
- Egashira, S., Honda, N., and Itoh, T.: Experimental study on the entrainment of bed material into debris flow, *Phys. Chem. Earth, Parts A/B/C*, 26, 645–650, [https://doi.org/10.1016/s1464-1917\(01\)00062-9](https://doi.org/10.1016/s1464-1917(01)00062-9), 2001.
- Fannin, R. J. and Wise, M. P.: An empirical-statistical model for debris flow travel distance, *Can. Geotech. J.*, 38, 982–994, <https://doi.org/10.1139/t01-030>, 2001.
- Frank, F., McArdeell, B. W., Huggel, C., and Vieli, A.: The importance of entrainment and bulking on debris flow runout modeling: examples from the Swiss Alps, *Nat. Hazards Earth Syst. Sci.*, 15, 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, *Comput. Geosci.*, 17, 413–422, [https://doi.org/10.1016/0098-3004\(91\)90048-i](https://doi.org/10.1016/0098-3004(91)90048-i), 1991.
- Gartner, J. E., Cannon, S. H., and 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, *Eng. Geol.*, 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., and Rubin, D. B.: *Bayesian Data Analysis* (3rd ed.), Electronic Edition, <http://www.stat.columbia.edu/~gelman/book/BDA3.pdf> (last access: 17 April 2023), 2021.
- Goode, J. R., Luce, C. H., and 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., and Rengers, F. K.: A progressive flow-routing model for rapid assessment of debris-flow inundation, *Landslides*, 19, 2055–2073, <https://doi.org/10.1007/s10346-022-01890-y>, 2022.
- Guthrie, R., Hockin, A., Colquhoun, L., Nagy, T., Evans, S. G., and Ayles, C. P.: An examination of controls on debris flow mobility: Evidence from coastal British Columbia, *Geomorphology*, 114, 601–613, <https://doi.org/10.1016/j.geomorph.2009.09.021>, 2010.
- Guthrie, R. and Befus, A.: DebrisFlow Predictor: an agent-based runout program for shallow landslides, *Nat. Hazards Earth Syst. Sci.*, 21, 1029–1049, <https://doi.org/10.5194/nhess-21-1029-2021>, 2021.
- Hammond, C. J., Prellwitz, R. W., and Miller, S. M.: Landslides hazard assessment using Monte Carlo simulation. in: *Proceedings of 6th international symposium on landslides*, edited by: Bell, D. H., Christchurch, New Zealand, Balkema, 2, 251–294, ISBN 9789054100324, 1992.
- Han, Z., Chen, G., Li, Y., Tang, C., Xu, L., He, Y., Huang, X., and Wang, W.: Numerical simulation of debris-flow behavior incorporating a dynamic method for estimating the entrainment, *Eng. Geol.*, 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., and Shang, Y.: Numerical simulation for run-out extent of debris flows using an improved cellular automaton model, *B. Eng. Geol. Environ.*, 76, 961–974, <https://doi.org/10.1007/s10064-016-0902-6>, 2017.
- Han, Z., Ma, Y., Li, Y., Zhang, H., Chen, N., Hu, G., and Chen, G.: Hydrodynamic and topography based cellular automaton model for simulating debris flow run-out extent and entrainment behavior, *Water Res.*, 193, 116872, <https://doi.org/10.1016/j.watres.2021.116872>, 2021.
- Heiser, M., Scheidl, C., and Kaitna, R.: Evaluation concepts to compare observed and simulated deposition areas of mass movements, *Comput. Geosci.*, 21, 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., Istanbuluoglu, E., and 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 Surf. Dynam.*, 5, 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, *Hydrol. Process.*, 8, 327–334, <https://doi.org/10.1002/hyp.3360080405>, 1994.
- Horton, P., Jaboyedoff, M., Rudaz, B., and Zimmermann, M.: Flow-R, a model for susceptibility mapping of debris flows and other gravitational hazards at a regional scale, *Nat. Hazards Earth Syst. Sci.*, 13, 869–885, <https://doi.org/10.5194/nhess-13-869-2013>, 2013.
- Hungr, O., Morgan, G. J., and Kellerhals, R.: Quantitative analysis of debris torrent hazards for design of remedial measures, *Can. Geotech. J.*, 21, 663–677, <https://doi.org/10.1139/t84-073>, 1984.
- Hungr, O. and Evans, S. G.: Entrainment of debris in rock avalanches: An analysis of a long run-out mechanism, *Geol. Soc. Am. Bull.*, 116, 1240, <https://doi.org/10.1130/b25362.1>, 2004.
- Hutter, K., Svendsen, B., and Rickenmann, D.: Debris flow modeling: A review, *Continuum Mech. Therm.*, 8, 1–35, <https://doi.org/10.1007/bf01175749>, 1996.
- Iovine, G., D’Ambrosio, D., and 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, 287–303, <https://doi.org/10.1016/j.geomorph.2004.09.017>, 2005.
- Istanbuluoglu, E. and Bras R. L.: Vegetation-modulated landscape evolution: Effects of vegetation on landscape processes, drainage density, and topography, *J. Geophys. Res.*, 110, F02012, <https://doi.org/10.1029/2004jf000249>, 2005.
- Istanbuluoglu, 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. Geophys. Res.*, 110, F01014, <https://doi.org/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, *Geophys. Res. Lett.*, 35, L14403, <https://doi.org/10.1029/2008GL034477>, 2008.
- Iverson, R. M.: The physics of debris flows, *Rev. Geophys.*, 35, 245–296, <https://doi.org/10.1029/97rg00426>, 1997.
- Iverson, R. M.: How should mathematical models of geomorphic processes be judged?, in: *Prediction in Geomorphology*, edited by: Wilcock, P. and Iverson, R., American Geophysical Union, ISBN 0875909930, 2003.
- Iverson, R. M.: Elementary theory of bed-sediment entrainment by debris flows and avalanches, *J. Geophys. Res.*, 117, F03006, <https://doi.org/10.1029/2011jf002189>, 2012.
- Iverson, R. M. and Denlinger, R. P.: Flow of variably fluidized granular masses across three-dimensional terrain: 1. Coulomb mixture theory, *J. Geophys. Res.*, 106, 537–552, <https://doi.org/10.1029/2000jb900329>, 2001.
- Julien, P. Y. and Paris, A.: Mean Velocity of Mudflows and Debris Flows, *J. Hydraul. Eng.*, 136, 676–679, [https://doi.org/10.1061/\(asce\)hy.1943-7900.0000224](https://doi.org/10.1061/(asce)hy.1943-7900.0000224), 2010.
- Kean, J. W., Staley, D. M., Lancaster, J., Rengers, F., Swanson, B., and Coe, J.: 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, 1140–1163, <https://doi.org/10.1130/GES02048.1>, 2019.
- Keck, J.: MassWastingRunout field site data, HydroShare [data set], <http://www.hydroshare.org/resource/55813a5e01764546b76641a7385c2236> (last access: 17 September 2024), 2024.
- Korup, O.: Effects of large deep-seated landslides on hillslope morphology, western Southern Alps, New Zealand, *J. Geophys. Res.*, 111, F01018, <https://doi.org/10.1029/2004jf000242>, 2006.
- Lancaster, S. T., Hayes, S. K., and Grant, G. E.: Effects of wood on debris flow runout in small mountain watersheds, *Water Resour. Res.*, 39, 1168, <https://doi.org/10.1029/2001wr001227>, 2003.
- landlab: landlab, GitHub [code], <https://github.com/landlab/landlab>, 15 August 2024.
- Landlab: Landlab, <https://landlab.github.io/#/> (last access: 17 September 2024), 2017.
- Larsen, I. J. and Montgomery, D. R.: Landslide erosion coupled to tectonics and river incision, *Nat. Geosci.*, 5, 468–473, <https://doi.org/10.1038/ngeo1479>, 2012.
- Liu, J., Wu, Y., Gao, X., and Zhang, X.: A Simple Method of Mapping Landslides Runout Zones Considering Kinematic Uncertainties, *Remote Sens.-Basel*, 14, 668, <https://doi.org/10.3390/rs14030668>, 2022.
- Major, J. J.: Depositional Processes in Large-Scale Debris-Flow Experiments, *J. Geol.*, 105, 345–366, <https://doi.org/10.1086/515930>, 1997.
- Major, J. J. and Iverson, R. M.: Debris-flow deposition: Effects of pore-fluid pressure and friction concentrated at flow margins, *Geol. Soc. Am. Bull.*, 111, 1424–1434, [https://doi.org/10.1130/0016-7606\(1999\)111<1424:DFDEOP>2.3.CO;2](https://doi.org/10.1130/0016-7606(1999)111<1424:DFDEOP>2.3.CO;2), 1999.
- McCoy, S. W., Kean, J. W., Coe, J. A., Tucker, G. S., Staley, D. M., and Wasklewicz, T. A.: Sediment entrainment by debris flows: In situ measurements from the headwaters of a steep catchment, *J. Geophys. Res.*, 117, F03016, <https://doi.org/10.1029/2011jf002278>, 2012.
- McDougall, S. and Hungr, O.: A model for the analysis of rapid landslide motion across three-dimensional terrain, *Can. Geotech. J.*, 41, 1084–1097, <https://doi.org/10.1139/t04-052>, 2004.
- Medina, V., Hürlimann, M., and Bateman, A.: Application of FLAT-Model, a 2D finite volume code, to debris flows in the northeastern part of the Iberian Peninsula, *Landslides*, 5, 127–142, <https://doi.org/10.1007/s10346-007-0102-3>, 2008.
- Montgomery, D. R. and Dietrich, W. E.: Where do channels begin?, *Nature*, 336, 232–234, <https://doi.org/10.1038/336232a0>, 1988.
- Murray, B. A. and Paola, C.: A cellular model of braided rivers, *Nature*, 371, 54–57, <https://doi.org/10.1038/371054a0>, 1994.
- Murray, A. B. and Paola, C.: Properties of a cellular braided-stream model, *Earth Surf. Proc. Land.*, 22, 1001–1025, [https://doi.org/10.1002/\(SICI\)1096-9837\(199711\)22:11<1001::AID-ESP798>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1096-9837(199711)22:11<1001::AID-ESP798>3.0.CO;2-O), 1997.
- Murray, A. B.: Reducing model complexity for explanation and prediction, *Geomorphology*, 90, 178–191, <https://doi.org/10.1016/j.geomorph.2006.10.020>, 2007.
- Murray A. B.: Which Models Are Good (Enough), and When?. in: *Treatise on Geomorphology*, edited by: John F. Shroder, Volume 2, 50–58, San Diego, Academic Press, ISBN 9780123983534, 2013.
- Natural Resources Conservation Service (NRCS): Snow and Water Interactive Map (n.d.), Natural Resources Conservation Service [data set], <https://www.nrcs.usda.gov/resources/data-and-reports/snow-and-water-interactive-map> (last access: 15 April 2022), 2022.
- Nudurupati, S. S., Istanbuluoglu, E., Tucker, G. E., Gasparini, N. M., Hobbey, D. E. J., Hutton, E. W. H., Barnhart, K. R., and Adams, J. M.: On transient semi-arid ecosystem dynamics using Landlab: vegetation shifts, topographic refugia, and response to climate, *Water Resour. Res.*, 59, e2021WR031179, <https://doi.org/10.1029/2021wr031179>, 2023.
- Perron, J. T.: Climate and the Pace of Erosional Landscape Evolution, *Annu. Rev. Earth Pl. Sc.*, 45, 561–591, <https://doi.org/10.1146/annurev-earth-060614-105405>, 2017.
- Reid, M. J., Coe, J. A., and 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., and 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 Resour. Res.*, 42, W12422, <https://doi.org/10.1029/2005wr004591>, 2006.
- Rengers, F. K., McGuire, L. A., Kean, J. W., Staley, D. M., and Hobbey, D. E. J.: Model simulations of flood and debris flow timing in steep catchments after wildfire, *Water Resour. Res.*, 52, 6041–6061, <https://doi.org/10.1002/2015WR018176>, 2016.
- Roda-Boluda, D. C., D'Arcy, M., McDonald, J., and Whittaker, A. C.: Lithological controls on hillslope sediment supply: insights from landslide activity and grain size distributions, *Earth Surf. Proc. Land.*, 43, 956–977, <https://doi.org/10.1002/esp.4281>, 2018.

- Shaller, P. J., Doroudian, M., and Hart, M. W.: The Eureka Valley Landslide: Evidence of a dual failure mechanism for a Long-Runout Landslide, *Lithosphere*, 2020, 8860819, <https://doi.org/10.2113/2020/8860819>, 2020.
- Schürch, P., Densmore, A. L., Rosser, N., and McArdell, B. W.: Dynamic controls on erosion and deposition on debris-flow fans, *Geology*, 39, 827–830, <https://doi.org/10.1130/g32103.1>, 2011.
- Shen, P., Zhang, L. M., Wong, H., Peng, D., Zhou, S., Zhang, S., and Chen, C.: Debris flow enlargement from entrainment: A case study for comparison of three entrainment models, *Eng. Geol.*, 270, 105581, <https://doi.org/10.1016/j.enggeo.2020.105581>, 2020.
- Stock, J. P. J. and Dietrich, W. E.: Erosion of steepland valleys by debris flows, *Geol. Soc. Am. Bull.*, 118, 1125–1148, <https://doi.org/10.1130/b25902.1>, 2006.
- Strauch, R., Istanbuluoglu, E., Nudurupati, S. S., Bandaragoda, C., Gasparini, N. M., and Tucker, G. E.: A hydroclimatological approach to predicting regional landslide probability using Landlab, *Earth Surf. Dynam.*, 6, 49–75, <https://doi.org/10.5194/esurf-6-49-2018>, 2018.
- Takahashi, T.: *Debris Flow* (2nd ed.), CRC Press, Taylor and Francis Group, ISBN 978-0-203-57674-8, 2014.
- Tucker, G. E. and Bras, R. L.: Hillslope processes, drainage density, and landscape morphology, *Water Resour. Res.*, 34, 2751–2764, <https://doi.org/10.1029/98wr01474>, 1998.
- Tucker, G. E. and Hancock, G. J.: Modelling landscape evolution, *Earth Surf. Proc. Land.*, 35, 28–50, <https://doi.org/10.1002/esp.1952>, 2010.
- Tucker, G. E., McCoy, S. W., and Hobbey, D. E. J.: A lattice grain model of hillslope evolution, *Earth Surf. Dynam.*, 6, 563–582, <https://doi.org/10.5194/esurf-6-563-2018>, 2018.
- Western Regional Climate Center (WRCC): RAWs USA Climate Archive, <https://wrcc.dri.edu/>, last access: 17 September 2024.
- Whipple, K. X. and Dunne, T.: The influence of debris-flow rheology on fan morphology, Owens Valley, California, *Geol. Soc. Am. Bull.*, 104, 887–900, [https://doi.org/10.1130/0016-7606\(1992\)104<0887:TIODFR>2.3.CO;2](https://doi.org/10.1130/0016-7606(1992)104<0887:TIODFR>2.3.CO;2), 1992.
- Zhou, G. G. D., Li, S., Song, D., Choi, C. E., and Chen, X.: Depositional mechanisms and morphology of debris flow: physical modelling, *Landslides*, 16, 315–332, <https://doi.org/10.1007/s10346-018-1095-9>, 2019.