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

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

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

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

Abstract

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

1. Introduction

Over geologic timescales, landslides and their runout shape the topographic expression of mountain ranges and channel networks (e.g., Campforts et al., 2022; Korup, 2006; Larsen and Montgomery, 2012; Montgomery and Dietrich, 1988). Over more pragmatic engineering and environmental risk management timescales, landslides and their runout can inundate and destroy infrastructure (e.g., Kean et al., 2019) but also support numerous ecosystem benefits, including carbon and nutrient transport from hillslopes to channels and the creation of riparian habitat (Benda

35 et al., 2003; Bigelow et al., 2007; Goode et al., 2012). Therefore, explicit representation of landslide runout is a

36 necessary component of: (1) landslide inundation hazard assessments, with emphasis on inundation extent and flow

37 depth (e.g., Frank et al. 2015, Han et al., 2015); (2) watershed sediment yield models, with emphasis on the

38 mobilization, deposition and type of sediment carried by the landslide (e.g., Bathurst and Burton, 1998;

39 Istanbulluoglu,et al., 2005); and (3) landscape evolution models, with emphasis on topographic change prediction

40 (e.g., Tucker and Bras, 1998; Istanbulluoglu and Bras, 2005; Campforts et al., 2022);

41 Landslide runout processes can be generalized into three phases: initiation, erosion, and deposition. After a landslide

42 initiates, it may break apart and flow as a relatively dry debris slide, or it may mix with surface runoff to become a

43 debris flow. The mobility of the mass wasting material and resulting erosion/deposition pattern often varies as a

44 function of runout topography and initial relief and size of the landslide (Iverson, 1997). Mobility may also be

45 impacted by substrate liquefaction (Hungr and Evans, 2004) and landslide basal cataclasis (Shaller et al. 2020). As

46 the runout material moves downslope, flow depth varies as a function of channel width (Kean et al, 2019), which in

47 turn impacts erosion rates (Schürch et al. 2011). Theoretical, field and laboratory observations indicate that erosion

48 rates may also depend on the moisture content of the channel bed (Iverson, 2012; McCoy et al. 2012), flow grainsize

49 (Egashira et al., 2001) and granular stress within the flow (Capart et al, 2015). The slope at which deposition begins

50 is controlled by the grain to water ratio and friction angle of the slide material (Takahashi, 2014; Major and Iverson,

51 1999; Zhou et al., 2019) but the friction angle of the material may vary as a function of the grains in the flow and

52 fluidization of the flow material (Hutter et al., 1996). Lateral levees often form along the edges of the flow (Major,

53 1997; Whipple and Dunne, 1992; Shaller et al., 2020) and deposition at the distal end of the flow may occur as layered

54 accretion (Major, 1997) or as the emplacement of a single, massive deposit (Shaller et al., 2020). If the water content

55 of the runout material is high enough, as the solid fraction of the distal end of the flow compresses, the water is

56 squeezed out and may continue as an immature debris flow (sensu Takahashi, 2014) or intense bedload (sensu Capart

57 & Fraccarolo, 2011), extending the runout distance (e.g., Shaller et al. 2020).

58 Landslide inundation hazard models aim to accurately predict the runout extent and/or flow depths of a runout event

59 and may include some or most of the above processes in the model. Example models include: (1) site-specific-

60 empirical/statistical models that use simple geometric rules and an estimate of the total moblized volume (initial

61 landslide + eroded volume) or a growth factor (e.g., Reid et al. 2016); (2) detailed, continuum-based mechanistic

62 models, which conceptualize the runout process as a single-phase or multiphase flow using the depth-integrated

63 Navier-Stokes equations for an incompressible, free-surface flow (i.e., shallow water equations; Frank et al, 2015;

64 Han et al., 2015; Iverson and Denlinger, 2001) and often (though not always) require pre-knowledge of the total

65 moblized volume (e.g., Barnhart et al., 2021; Han et al. 2015); (3) reduced-complexity flow-routing models that use

66 rule-based abstractions of the key physical processes that control the flow (Clerici and Perego, 2000; Guthrie and

67 Befus, 2021; Gorr et al., 2022; Han et al., 2017, 2021; Horton et al., 2013; Liu et al, 2022) and are typically

68 implemented using just the initial landslide location and volume but often rely on heavy, site specific parameterization

69 and; (4) hybrid modelling approaches that combine mechanistic models with empirical and reduced-complexity

70 approaches (D'Ambrosio et al., 2003; Iovine et al., 2005; Lancaster et al., 2003; McDougall and Hungr 2004; Medina

71 et al., 2008).

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

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

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

Figure 1: Example landslides that are used to evaluate calibrated MWR performance: (a) Cascade Mountains, WA: a large debris avalanche over steep, broadly convergent terrain (photo credit: Stephen Slaughter). (b) Black Hills, WA: large debris flows over a broadly convergent, gently sloped valley (photo credit: Stephen Slaughter). (c) Rocky Mountains, CO: a moderate sized debris avalanche over steep, unconfined to divergent hillslope. (d) Olympic Mountains, WA: small debris flows in steep, highly convergent channels.

2.   Description of the MassWastingRunout model

2.1 Overview of the cellular-automaton Modelling approach

MWR is coded as a discrete cellular automaton (CA) model. CA models apply a set of equations or rules (deterministic or probabilistic) to individual cells of a grid to change the numerical or categorical value of a cell state (e.g., Codd, 1968). In earth sciences, CA models are widely used to model everything from vegetation dynamics (e.g., Nudurupati et al., 2023) to lava flows (e.g., Barca et al., 1993) to geomorphic transport, in which gravitationally directed erosion and depositional processes modify a digital elevation model (DEM) representation of a landscape (e.g., Chase, 1992;

113   Crave & Davy, 2001; Murray & Paola, 1994; Tucker et al., 2018). Existing CA-based landslide runout models include

114   Guthrie and Befus (2021), D'Ambrosio et al. (2003) and Han et al. (2021). In all of these models, runout behavior is

115   controlled by topographic slope and rules for erosion and deposition but conceptualization and implementation differ.

116   In MWR, mass-continuity is central to model conceptualization. Of the wide range of processes described in the

117   introduction that control observed runout, MWR explicitly represents erosion, deposition, and flow resistance due to

118   debris size and vegetation. Material exchange between the runout material and underlying terrain as well as flow

119   resistance determines runout extent and landscape evolution. Model rules are designed such that they can be

120   parameterized from field measurements. Finally, in MWR, most computations occur only at the location of moving

121   debris, in a manner analogous to the "mobile" cellular automaton implementation of Chase (1992).

122   Chase (1992) modelled precipitation-driven surface erosion by randomly placing single packets of precipitation on a

123   DEM, which then moved from higher elevation to lower elevation grid cells, eroding and transporting sediment as a

124   function of the slope between the cells. The individual packets of precipitation were referred to as precipitons. In

125   MWR, since we route the downslope progression of debris from a specified mass wasting source area, we refer to

126   these packets of debris as "debritons". The debritons represent debris flux, here defined as a volume of debris

127   transferred per model iteration per grid-cell area, [m^3/m^2/iteration] and are equivalent to the flow depth in the cell.

128   The present implementation of the MWR algorithm is coded in Python and developed as a component of the Landlab

129   earth surface modeling toolkit (Barnhart et al., 2020; Hobley et al., 2017). MWR uses the Landlab raster model grid,

130   which consists of a lattice of equally sized, rectangular cells. Topographic elevation, derived topographic attributes

131   like slope and curvature, and other spatially varying attributes such as regolith depth and grain size, are recorded at

132   nodes in the center of each cell (see Figure 5 of Hobley et al., 2017). In the subsequent sections we describe the model

133   theory. Note that all the notations of parameters and variables used in this theory are listed in Section 10.

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

135   To initiate MWR, the user provides maps of initial topography, regolith depth, and the location and depth of the mass

136   wasting source material (e.g., landslide body). Each raster model grid node in the mass wasting source material is

137   designated as a debriton (Figure 2, iteration $t = 0$) with a magnitude equal to the mass wasting source material depth

138   and basal elevation equal to the initial topography minus the mass wasting source material depth. The basal elevation

139   can be thought to represent the rupture or slip surface of the source material and the redistribution (flux) of each

140   debriton to its downslope nodes (receiver nodes) is determined as a function of the slope of the slip surface. Note that

141   if the depth of the wasting source material is spatially variable (e.g., a rotational failure), the slope of the slip surface

142   will not match the slope of the initial terrain surface. At the lowest-elevation debriton of the source material, flux to

143   its downslope nodes is determined using the surface slope of the initial DEM (see flow direction of lowest node in

144   Figure 3a). This implementation helps to ensure that the lowest-elevation debriton in the mass wasting source

145   moves downslope and movement of upslope debritons are impacted by the geometry of the mass wasting source

146   material. For example, the receiver nodes of the lowest-elevation debriton in the landslide illustrated in Figure 2

147   (iteration $t = 0$, detailed in Figure 3a) would be identified as those among the eight neighboring nodes whose initial

148   topographic elevation was less than the initial topographic elevation of the node while for the debriton at node 51, the

149 receiver nodes would be identified as those among the eight neighboring nodes whose topographic elevation is less

150 than the topographic elevation of the terrain underlying the debriton (the slip surface).

151
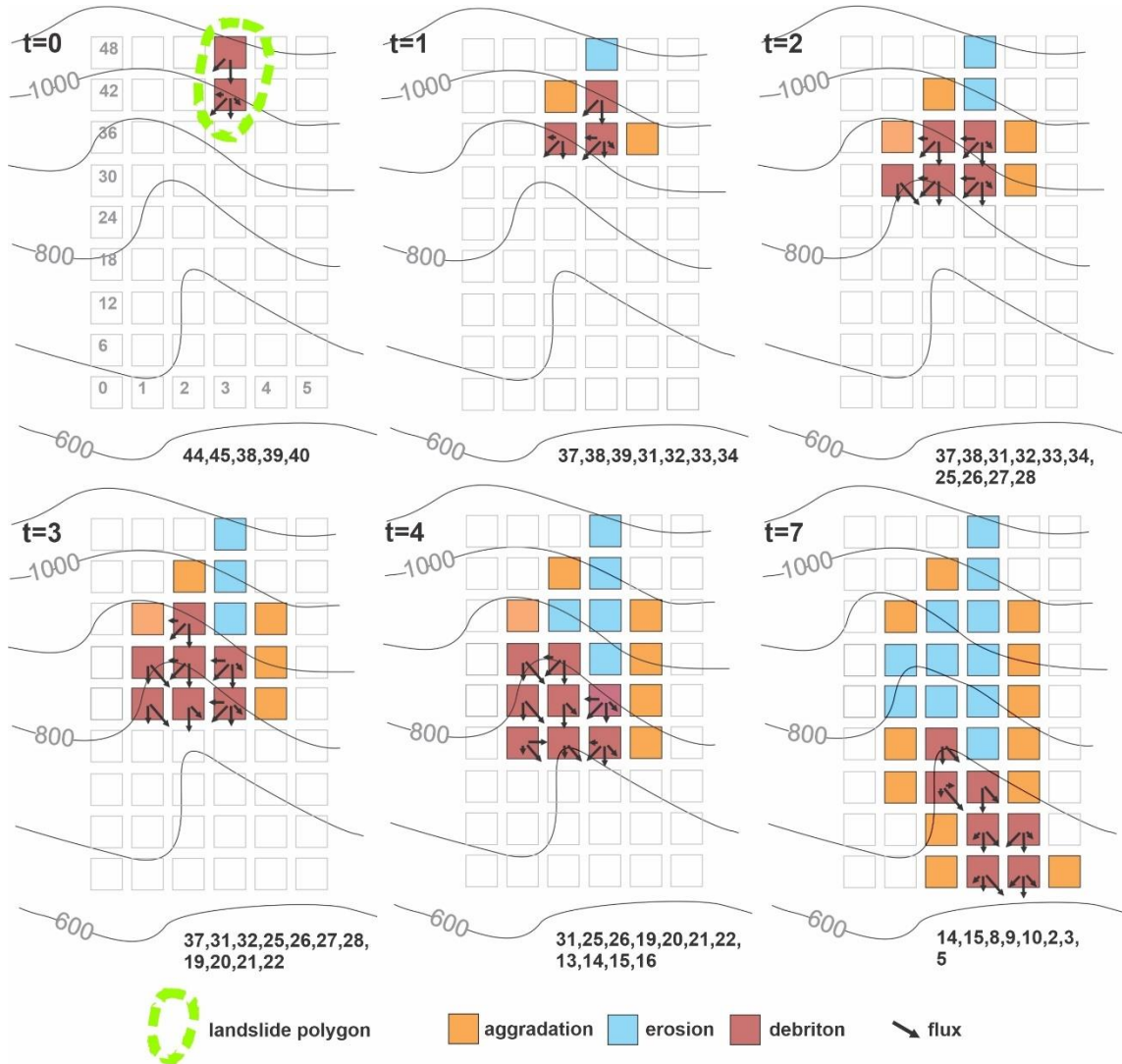
Figure 2. Illustration of initial mass wasting release (t = 0) and runout. Notice how the list of receiver nodes changes with each iteration. The flow elongates and widens as the number of receiver nodes increase and stops when the number of receiver nodes is zero. If the incoming flux (sum of all incoming debritons) to a node is less than q_c, the material stops, causing aggradation.

155
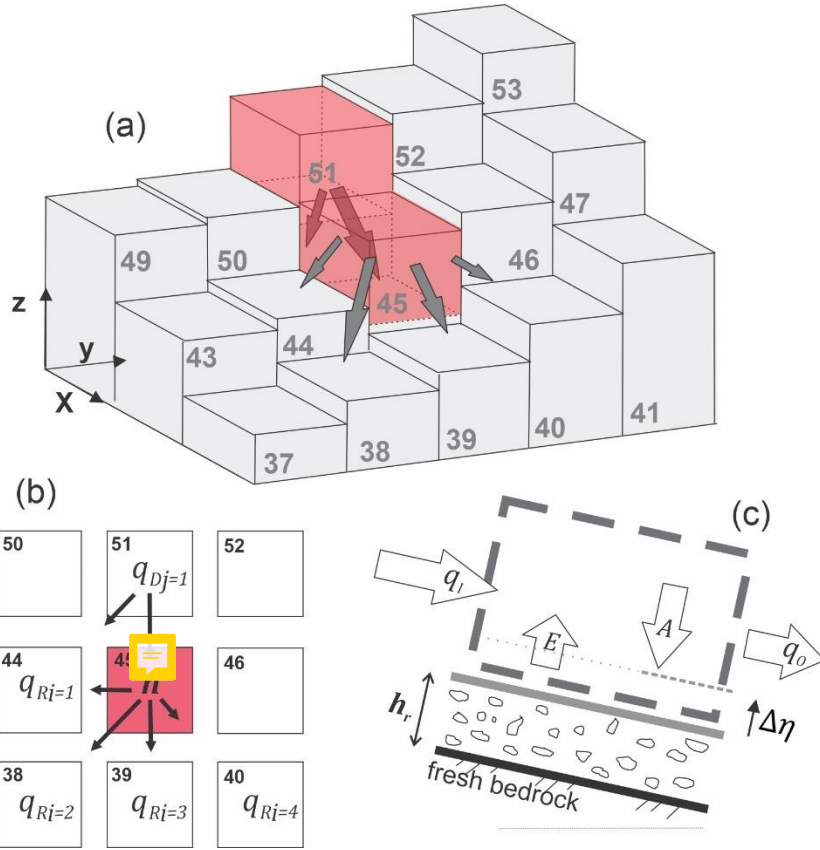
Figure 3. (a) Three-dimensional illustration of iteration t = 0 in Figure 2, showing initial source material nodes (represented by red cells) and flux towards downslope nodes. Except for the lowest elevation node in the mass wasting source material, all debritons are directed downslope based on the underlying topographic slope (compare flow directions of node 51 to node 45); (b) Distribution of q_o to downslope nodes 38, 39, 40 and 44; (c) illustration of mass continuity applied to any node that receives a debriton.

160

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

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

$$q_{R_i} = q_o \frac{S_i^a}{\sum_{i=1}^{Nr} S_i^a} \qquad (1)$$

172   where q_O is the total out-going flux from the node and has units of depth [m] per model iteration, Nr is the number

173   of receiving nodes, i is the index for each receiver node (e.g., $i = 1, 2 \dots Nr$) and S_i is the underlying topographic

174   slope to the i-th receiver node (Figure 3b). The Freeman (1991) multiflow direction algorithm is a commonly used

175   approximation for two-dimensional flow, and in this implementation it is handled by a pre-existing Landlab flow-

176   routing component. The exponent a controls how material is distributed to downslope nodes. In a braided river

177   cellular-automaton model, Murray and Paola (1997) used an approximation for turbulent shallow water flow to justify

178     $a = 0.5$ (which is the exponent on the slope factor in channel friction laws). For our application, we found MWR

179   provided a closer fit to observed mass wasting runout if $a = 1$, suggesting that the material behavior is more similar

180   to linear-viscous shear flow than to wall-bounded turbulent shear flow. The total incoming flux (again, in units [m]

181   per model iteration) towards a given node (q_I), is determined by summing the flux from each of the node's donor

182   nodes:

183     $q_I = \sum_{j=1}^{Nd} q_{D_j}$ (2)

184   Where Nd is the number of donor nodes, and q_{D_j} is the flux from node D_j (the j-th donor node, $j = 1,2 \dots Nd$; Figure

185   3b).

186   As noted by Tucker and Hancock (2010), the flow depths calculated from two-dimensional flow approximations like

187   (1) can be influenced by the grid-size used to represent the terrain and depending on the boundary conditions,

188   neglection of pressure and momentum forces may lead the model to underestimate or overestimate flow width in some

189   circumstances. Rengers et al. (2016) noted that this same issue occurs when using a kinematic wave approximation of

190   the shallow water equations because the kinematic wave approximation lacks a pressure term that would normally

191   allow the modelled water surface to spread out. For the purpose of determining flow-depth-dependent erosion rates

192   described later in this paper, and to provide a simplified representation of the effect of pressure forces, we constrain

193   flow depths to no more than a maximum flow as:

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

195   Where h_{max} is an effective upper limit to flow depth, that in practice can be approximated as the maximum observed

196   flow depth, as inferred from field indicators or assigned based on expert judgement (See Section 5) and h is the

197   corrected flow depth used to calculate flow shear stress. This correction allows erosion rates to vary with flux but

198   prevents unreasonably large values. This flow depth correction does not violate the conservation of mass and runout

199   mass balance, as h is only used to calculate flow shear stress.

200   To determine aggradation (A) at a node, we use a critical slope (S_c) constraint that permits computationally-rapid

201   distribution of q_I over multiple nodes. Critical slope constraints or rules are common to many reduced-complexity and

202   landscape evolution models. Chen et al. (2023) showed that when flow inertia can be ignored, S_c can be approximated

203   from the surface slope of observed deposits. Several landscape evolution models use a S_c-based nonlinear, nonlocal

204   aggradation scheme (e.g., Campforts et al., 2020; Carretier et al., 2016) but when this rule is implemented with the

205   debriton framework described above, unreasonably tall deposits result when q_I is large and slope at the node (S) $<<$

206   S_c. To resolve this problem, aggradation depth can be limited to $A \leq S_c \Delta x$, (where Δx grid cell length), but we found

207   that this constraint results in long deposits that parallel the underlying slope when q_I is large. Instead, MWR computes

208   the aggradation depth at a node assuming that the aggradation will spread over N_a nodes until all of q_I is deposited

209   and that the surface slope of the overall deposit will be equal to S_c, as shown in Figure 4 and described as follows.

210
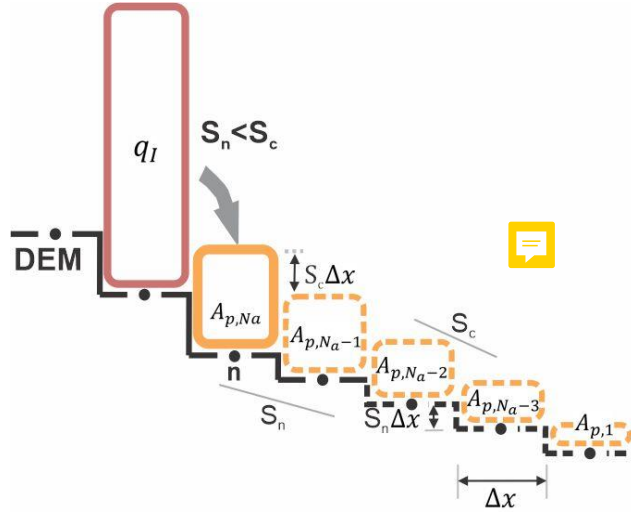
Figure 4. Illustration of aggradation rule used in MWR when q_I is assumed to spread over 5 nodes ($N_a = 5$). Solid
yellow box indicates deposition at node n. Dashed yellow boxes and lines indicate hypothetical deposition and
underlying topography. Dots along DEM surface are nodes.

214   Aggradation at a node is determined as:

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

216   Where S is the steepest slope to the node's eight neighbouring nodes, A_{p,N_a} is a potential aggradation depth (A_p)

217   necessary to form a deposit that: (1) begins at the node and spreads over N_a consecutive nodes; (2) has a total volume

218   equal to $q_I \Delta x^2$; (3) a surface slope equal the critical slope S_c and; (4) an underlying topographic slope equal to the

219   steepest slope at the node and assumed constant over the N_a consecutive nodes of deposition (S_n). From this assumed

220   deposit, we can analytically define A_{p,N_a} and N_a as a function of q_I, S_c and S_n as follows:

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

222   up the overall deposit as:

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

224   where $A_{p,i}$ is the i-th deposition amount in the deposit and $i = 1$ is the last node of deposition ($A_{p,1}$; see Figure 4).

225   Since we assume the deposit slope and underlying topographic slope are uniform, the deposition amount at any of the

226   N_a nodes can be determined from $A_{p,1}$ as:

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

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

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

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

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

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

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

234   We use (8) to solve for A_{p,N_a} and (9) to solve for N_a assuming $A_{p,1} = 1/2\, \Delta x S_c$ and rounding the positive solution to

235   the nearest integer. When implemented using a single debriton, released on a two-dimensional hillslope as illustrated

236   in Figure 4, the debriton deposits over N_a nodes at a uniform slope equal to S_c. When implemented on an actual three-

237   dimensional terrain, the interaction between multiple debritons in multiple directions creates a complex deposit whose

238   slope changes with S_c.

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

240   regolith depth, h_r:

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

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

243   shear stress (τ_c) of the regolith at the node [Pa]:

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

245   The coefficient k is an erodibility parameter [m/Pa^f]. Stock and Dietrich (2006) showed that k encapsulates substrate

246   properties. If h_e is used to represent erosion over geomorphic time scales, with repeated debris flow occurrences in a

247   single model iteration, k becomes associated with debris flow length and frequency (Perron, 2017). In our application

248   since we are modelling the erosion associated with a single runout event, as represented by the downslope movement

249   of the debritons, the coefficient k therefore needs to scale h_e on the order of the average erosion depth caused by a

250   single debriton. Using this logic, k can be computed using the observed average erosion depth and an estimated length

251   of the runout material that caused the erosion. Further details on how we determine k from observed runout are

252   included in the Supplementary Material. The exponent f controls the non-linearity of h_e. Many authors (Chen &

253   Zhang, 2015; Frank et al., 2015; Shen et al., 2020) use a value of 1 for f but field measurements by Schürch et al.

254   (2011) (see their Figure 3) suggest that f may be less than 1 if τ is assumed to vary linearly with flow depth,

255   particularly at flow depths greater than 3 meters.

256   MWR includes two options for defining τ: (1) a quasi-static basal shear stress approximation or (2) a grain-size-based

257   shear stress approximation. The quasi-static basal shear stress approximation (e.g., Takahashi, 2014) is defined as:

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

259 where ρ is the density of mass wasting material (grain and water mixture) [kg/m³], g is gravity [m/s²] and h is the

260 adjusted flow depth described in (3) and θ is the topographic slope (tan⁻¹(S)) measured in degrees.

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

262 $\tau = \sigma \tan \varphi$ $\qquad$ (13)

263 Where σ is normal stress [Pa], φ is the collision angle between grains, measured from the vertical axis (See Bagnold,

264 1954), with a value of $\tan \varphi$ typically equal to 0.32. Stock and Dietrich (2006) defined σ as:

265 $\sigma = \cos \theta v_s \rho_s D_s{}^2 \left(\frac{du}{dz}\right)^2$ $\qquad$ (14)

266 Where v_s is the volumetric solids concentration, ρ_s is density of the solids [kg/m³], u is flow velocity [m/s], z is depth

267 below the flow surface [m], du/dz is the shear strain rate [1/s] and D_s is the representative grain size [m]. Stock and

268 Dietrich (2006) suggested that D_s corresponds to a small percentile of the coarsest fraction of the runout material (D_{88}

269 to D_{96}) and they approximated du/dz as:

270 $\frac{du}{dz} = \frac{u}{h}$ $\qquad$ (15)

271 Solely for the purpose of computing du/dz, we approximate velocity at a node using a grain-size dependent empirical

272 formula for debris flow velocity by Julien and Paris (2010) as:

273 $u = 5.75 u^* log \left(\frac{h}{D_s}\right)$ $\qquad$ (16)

274 Where $u^*$is shear velocity $\left(\sqrt{gh \tan \theta}\right)$. Substituting (16), (15), (14) and (13) into (11) yields a grain-size dependent

275 approximation for h_e that mimics the non-linear erosion response to flow depth in Schürch et al. (2011). Additionally,

276 this form of τ is advantageous because it permits landslide-driven erosion rates to scale with landslide grain size,

277 which can vary by lithologic region (e.g., Roda-Boluda et al. 2018). As will be shown in Section 5, we obtained

278 reasonable model calibration at multiple sites by defining D_s from the coarser grain sizes observed in the field at

279 existing runout-deposits, road-cuts and tree-throw pits.

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

281 3c):

282 $q_O = \begin{cases} q_I - A + E, & q_I \geq q_c \\ 0 & , & q_I < q_c \end{cases}$ $\qquad$ (17)

283 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

284 conceptually represents the flow depth below which flow resistance is large enough to cease the forward momentum

285 of the flow, whether in the form of internal friction or friction due to vegetation and obstructions (e.g., large clasts or

286 logs). The density and water content of q_I, A, and E are treated as uniform and surface runoff, such as channelized

287 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

288 have been determined, change in elevation at a node ($\Delta\eta$) is calculated as:

289 $\Delta\eta = A - E$ $\qquad$ (18)

290 Attributes of the debriton and regolith are updated using a volumetric-weighted average approach. First, for each

291 regolith attribute being tracked by the model (e.g., grain size), the attribute value delivered to a node from its donor

292 nodes (ξ_D) is determined as:

293 $\xi_D = \frac{\xi_D \cdot q_D}{q_I}$ (19)

294 where $\boldsymbol{q_D}$ is a vector containing all q_{D_j} sent to the node, $\boldsymbol{\xi_D}$ is a vector containing the incoming attribute values for

295 each q_{D_j}, and q_I is the sum of incoming flux from donor nodes defined by (2).

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

297 $\xi_R = \frac{\xi_{t-1}E + \xi_D(q_I - A)}{q_O}$ (20)

298 where ξ_{t-1} is the attribute value at the node before any aggradation (i.e., the previous iteration attribute value). Finally,

299 the attribute value at the node, updated to account for erosion and aggradation (ξ) is:

300 $\xi = \frac{\xi_{t-1}(h_r - E) + \xi_D A}{A + h_r - E}$ (21)

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

302 $\eta = \eta_{t-1} + \Delta\eta$ (22)

303 $h_r = h_{r\ t-1} + \Delta\eta$ (23)

304 Where η_{t-1} and $h_{r\ t-1}$ are the topographic surface elevation and regolith thickness at the node from the previous

305 model iteration. After regolith thickness and topographic elevation have been updated for each debriton, the multi-

306 direction slope of the DEM, which is used for routing the debritons in the next model iteration, is recomputed from

307 the topographic surface.

308 Using the above approach, debritons may become obstructed if they encounter a topographic pit or flat topography in

309 the DEM. To allow a debriton to pass an obstruction, we rely on a simple work-around: upon encountering the

310 obstruction, the debriton is directed to itself and some portion of the debris is deposited based on (4). At the end of

311 the model iteration, the node elevation and slope are updated. During the next iteration, if the remaining mobile debris

312 is no longer obstructed, it moves to its downslope node(s). If the node is still obstructed, it is again sent to itself until

313 either all material has deposited or the elevation of the node exceeds that of its neighbour nodes, allowing the debriton

314 to move downslope.

315 **3.   Calibration and MWR probability**

316 **3.1   Calibration utility**

317 MWR includes an adaptive Markov Chain Monte Carlo (MCMC) calibration algorithm described by Coz et al. (2014)

318 and Renard et al. (2006). The user provides an initial (prior) guess of the parameter values and their respective

319 probability distribution functions (PDF) that calibrate the model to a specific site. Then, the calibration algorithm

320 randomly selects a set of parameter values (Λ) from the prior PDFs and runs MWR using Λ. Once the model has

321 completed the run, the algorithm evaluates the posterior likelihood of the parameter set ($L(\Lambda)$) as a lumped index of

322 model ability to replicate observed runout (described below) and the prior likelihood of the parameter set. After the

323 first $L(\Lambda)$ has been determined, the algorithm selects a new set of parameters (Λ_{t+1}) by jumping some distance from

324 each parameter in Λ space. Depending on the value of $L(\Lambda_{t+1})$, the algorithm either stays at Λ or moves to Λ_{t+1}. This

325 Markov process is repeated a user-specified Nc times. Jump direction is random, but the algorithm is adaptive because

326 the jump distance changes depending on how often $L(\Lambda_{t+1}) > L(\Lambda)$. For a detailed description of the algorithm see

327 Coz et al. (2014).

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

329 other performance metrics as:

330 $$L(\Lambda) = p(\Lambda) * \Omega_T * \frac{1}{\Delta\eta_E{}^2} * \frac{1}{Q_{s_E}{}^2} \tag{24}$$

331 where Ω_T is the Lee-Salle index (Heiser et al., 2017) for model planimetric fit; and $\Delta\eta_E$ and Q_{s_E} are new

332 dimensionless indices, proposed for this study. The indice $\Delta\eta_E$ is the volumetric error of the modelled topographic

333 change normalized by the observed total mobilized volume (initial landslide + erosion volume). The indice Q_{s_E} is the

334 mean-cumulative flow error along the modelled runout path normalized by the observed mean cumulative flow. Larger

335 values of Ω_T and smaller values of $\Delta\eta_E$ and Q_{s_E} indicate modelled runout more closely fits observed. Note that we

336 add a value of 1 to Ω_T and use the squared- reciprocal values of $\Delta\eta_E$ and Q_{s_E} in (24) so that the magnitude of $L(\Lambda)$ is

337 always equal to or greater than zero and increases with improved fit. The metric Ω_T is written as:

338 $$\Omega_T = \frac{\alpha - \beta - \gamma}{\alpha + \beta + \gamma} + 1 \tag{25}$$

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

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

341 $$\Delta\eta_E = \sqrt{\frac{\sum_{i=0}^{p}[(\Delta\eta_{Oi} - \Delta\eta_{Mi})\Delta x^2]^2}{V^2}} . \tag{26}$$

342 Where V is observed total mobilized volume and p is the number of nodes in the modelled runout extent, and $\Delta\eta_{Mi}$

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

344 To calculate Q_{s_E}, we first determine the cumulative debris flow volume (Q_s) at each node, j, ($Q_{s\,j}$) along the runout

345 profile, in a manner similar to the flow volume/mass balance curves in Fannin and Wise (2001) and Hungr and Evans

346 (2004):

347 $$Q_{s\,j} = -\Delta x^2 \sum_{i=1}^{u_j} \Delta\eta_{i,j} \tag{28}$$

348 where $\Delta\eta_{ij}$ is the topographic change [m] at the i-th node located upstream of node j, and u_j is the total number of all

349 nodes located upstream of j. $Q_{s\,j}$ is computed for both the observed and modelled runout ($Q_{sO\,j}$ and $Q_{sM\,j}$

350 respectively) and Q_{s_E} of a runout is determined as:

351 $$Q_{s_E} = \sqrt{\frac{\frac{1}{r}\Sigma_{j=1}^{r}\left(Q_{sO\,j}-Q_{sM\,j}\right)^2}{\overline{Q_{sO}}^2}}$$ (29)

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

353 As will be detailed in Section 5, field estimates for S_c and q_c, vary over the length of the runout path. To account for

354 the heterogeneity of S_c and q_c, we estimate prior distributions of potential S_c and q_c values from field/remote sensing

355 measurements. Then, from model calibration to a DEM-of-Difference (pre-runout DEM subtracted from the post-

356 runout DEM; DoD using the calibration utility, we find single values of S_c and q_c that allow the modelled DoD to

357 replicate the observed DoD .

358 We run the calibration utility using a single Markov chain of 2000 repetitions. At most sites, the model converged

359 relatively quickly on a solution and we therefore didn't consider burn-in or evaluate convergence (e.g., Gelman et al.

360 2021). Future implementations of the calibration utility may include multiple chains, burn-in and a check for

361 convergence. As a final note, many debris flow runout models are evaluated using Ω_T or variations of Ω_T alone (e.g,

362 Gorr et al., 2022; Han et al., 2017) and the MWR calibration utility can also be run solely as a function of Ω_T. However,

363 we found that calibration based on Ω_T alone results in high parameter equifinality (e.g., Beven 2006); multiple

364 parameter sets result in an equally calibrated model as evaluated by Ω_T. As such, we recommend calibrating debris

365 flow models to an observed DoD. If repeat lidar is available, a DoD can be obtained from before and after scans of

366 the observed runout event. Alternatively, a DoD can be created by hiking the observed runout event and mapping

367 field-interpreted erosion and deposition depths. Additional details on how we prepared DoDs for multiple sites are

368 included in the Supplementary Material.

369

370

371 **3.2   Mapping landslide runout hazard**

372 MWR includes an additional utility called MWR Probability that produces landslide runout probability maps. MWR

373 Probability repeatedly runs MWR a user specified Np times, each repetition with a different, randomly sampled

374 parameter set from the posterior parameter PDFs produced by the calibration utility. MWR Probability includes three

375 options for specifying the initial mass wasting source material: (1) a user-provided landslide source area polygon(s)

376 based on field and/or remote sensing observations; (2) a user-defined hillslope susceptible to landslides (e.g.,

377 potentially unstable slope), where landslide area and location are randomly selected within, but no larger than the

378 hillslope; this option is useful when the extent of a potential landslide is unknown; and (3) a series of mapped landslide

379 source areas within a watershed, as determined by an externally run Monte Carlo landslide initiation model (e.g.,

380 Hammond et al. 1992; Strauch et al., 2018) ; this option is useful for regional runout hazard applications. If using

381 Option 1, modelled runout probability represents uncertainty in MWR parameterization. If using Option 2 or 3,

382 modelled runout probability reflects uncertainty in both MWR parameterization and landslide location and size.

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

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

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

4.   Basic model behavior

We evaluate basic model behavior using a series of virtual experiments. The virtual experiments consist of six synthetic terrains including: (A) a planar slope that intersects a gently sloped plane (S = 0.001), (B) a planer slope with a constriction, that intersects a gently-sloped plane, (C) a planar slope that has a bench mid-slope and then intersects a gently-sloped plane; (D) a concave up, uniform-convergence slope; (E) a concave up, variable-convergence slope that widens (convergence decreases) in the downslope direction; (F) a convex up, variable-convergence slope that widens (convergence decreases) in the downslope direction. On each terrain, a 30-meter wide, 50-meter long and 3-meter deep landslide is released from the top of the terrain. All six terrains are covered by a 1-meter thick regolith and use the same parameter values (S_c = 0.03, q_c = 0.2 m, k = 0.01, D_p = 0.2 m). Experiment results are shown in Figure 5.
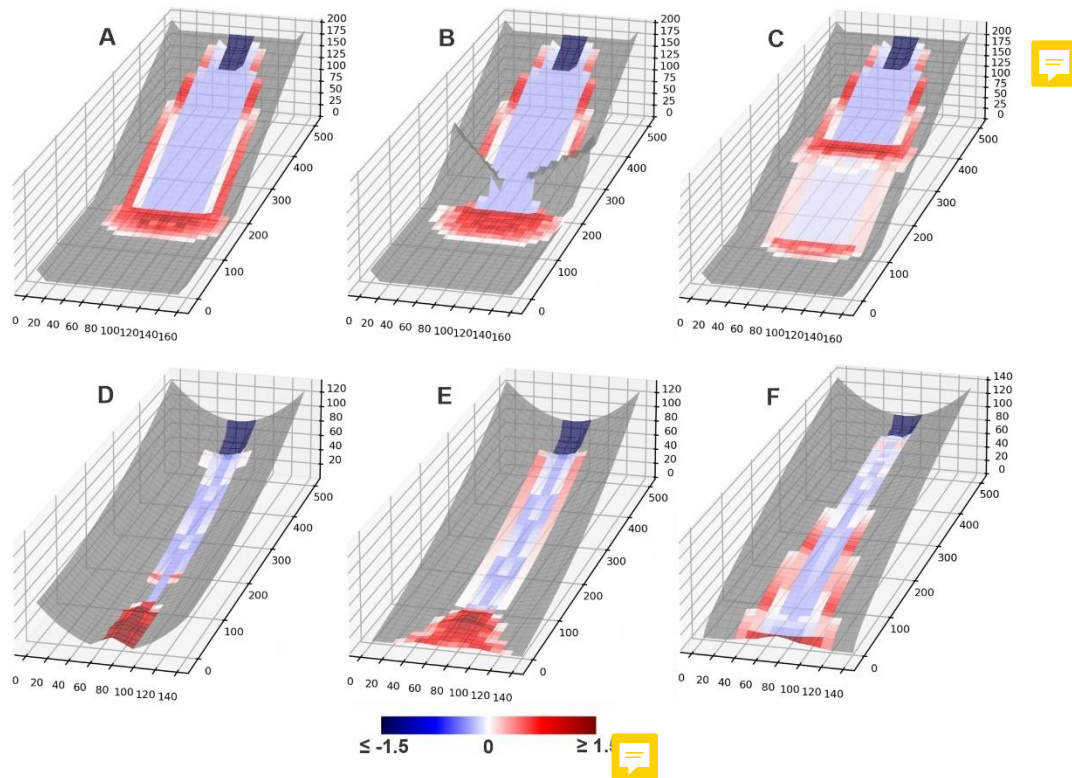
Figure 5. Shaded, 3-D visualizations of model response to six different synthetic terrains, colored according to the DoD of the final runout surface. Red indicates a positive change in the elevation of the terrain (aggradation) and blue indicates a negative change (erosion). Grid size is 10 meters. 3-D representation of DoD is exaggerated by a factor of 5 to make visible in figure.

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

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

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

We did not attempt to compare MWR modelled flow with the output of shallow-water-equation based models or observed granular flows (e.g., Medina et al, 2008; McDougall and Hungr, 2004; Iverson and Denlinger, 2001; Han et al., 2015). The cellular automaton representation in MWR does not model the time-dependent evolution of debris flow velocity and depth, and conceptually moves debris instantaneously at each iteration, as driven by changes in the evolving topographic elevation field. Because of that, only the final outcome of MWR can be compared with other models or observed runout, which we do in the next section.

5.   Model Validation:

5.1 Overview

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

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

Table 1. Landslide and runout characteristics

| site	Cascade Mountains, 09	Cascade Mountains, 22	Black Hills, south	Black Hills, north	Rocky Mountains	Olympic Mountains
landslide length, ℓ [m]	185	55	80	75	40	45
landslide width [m]	80	50	15	65	35	15
landslide volume [m^3]	110,000	22,000	1,500	18,500	4,600	400 - 2,200
2-day cumulative precipitation + snowmelt [mm]	120+85	140+75	205+50	205+50	193+0	100 - 220 + ?
maximum grain size [m]	0.316	0.316	0.48	0.206	0.984	0.8
Slope range of positive-net deposition [%]	1 - 15	1 - 15	<1 - 10	<1 - 8	16 - 25	5 - 15
average flow depth in scour zone [m] [a]	4	2	2	3	3	3
average channel slope in scour zone [m/m]	0.25	0.25	0.15	0.15	0.4	0.3

average channel width in scour zone [m]	45	20	25	35	55	10
length of erosion, [m]	600	340	1210	1345	360	2550
erosion area, $\mathbb{A}$ [m^2]	28,400	6,600	22,800	52,400	20,800	28,900
erosion volume, $\sum E \Delta x^2$ [m^3] [b]	44,547	5,125	12,332	26,815	34,275	33,725
average erosion per unit length of landslide, $\bar{E}/\ell$, [m/m]	0.0085	0.014	0.0068	0.0068	0.041	0.026
k	0.020	0.034	0.017	0.020	0.076	0.051
growth factor, [m^3/m]	74.2	15.1	10.2	19.9	95.2	13.2
average observed $|\Delta\eta|$ [m]	2.4	2.2	0.53	0.63	0.89	1.4
total erosion volume / total mobilized volume [c]	0.29	0.19	0.89	0.59	0.88	0.97

461   [a] rough approximation based on landslide volume, channel width and height of scour marks in erosion zone

462   [b] excludes landslide volume

463   [c] total moblized volume = erosion volume + landslide volume

464

465
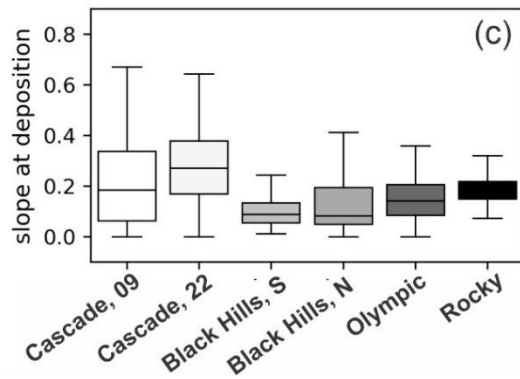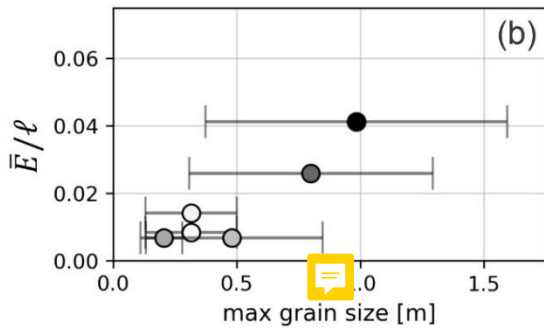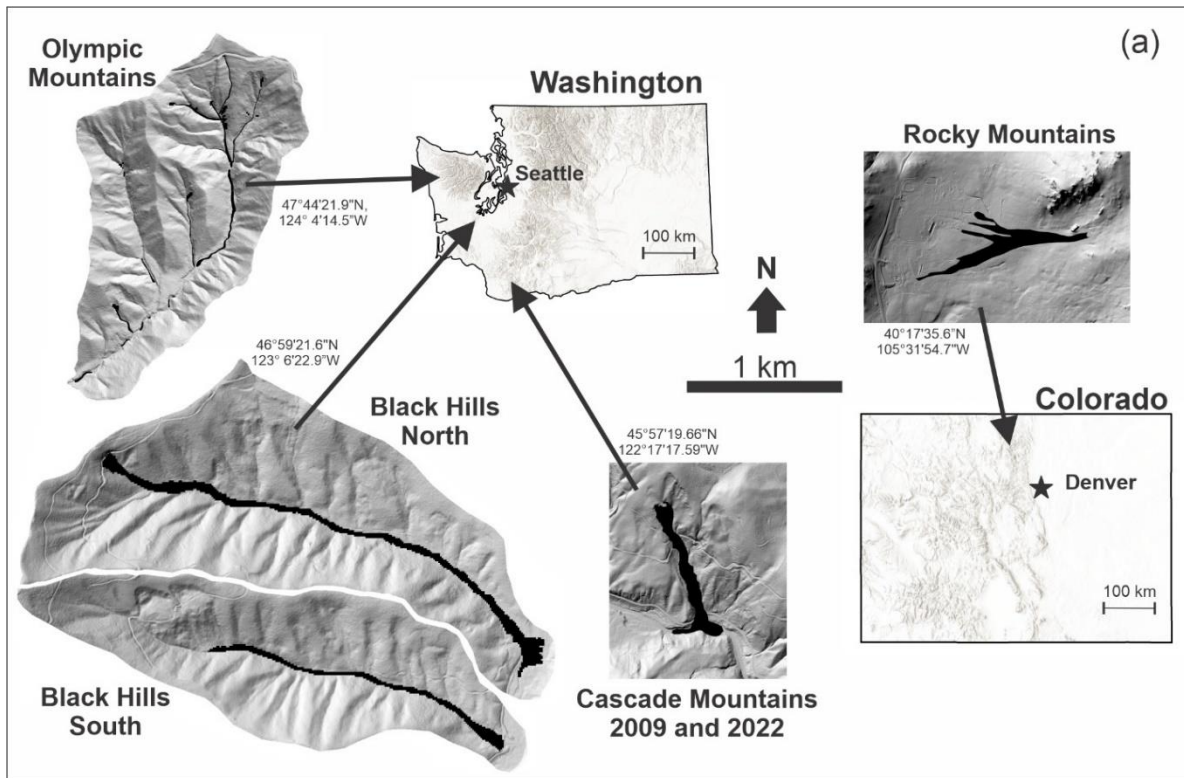
Figure 6 (**a**) Landslide locations in Washington and Colorado states. Coordinates next to each site are WGS84. Shaded DEMs of each site are shown at the same scale. (**b**) Observed average erosion rate per unit landslide length ($\bar{E}/\ell$) relative to the observed average-maximum grainsize. Error bars indicate standard deviation. (**c**) Underlying topographic slope of observed deposition locations.

5.2 Model setup and field parameterization

Each model was set up on a 10-meter grid representation of the pre-event DEM. The extent of the mass wasting source material, which in all cases was a landslide, was interpreted from a combination of lidar, air-photo and field observations. At all locations, we use (13) to approximate shear stress. We field-surveyed each site, noting the maximum flow thickness, typical deposition and erosion depths and the size of the largest grains in the runout deposits. We estimated parameter values from these field and remote observations (See Table 1). A site-specific value for k was determined as a function of the observed average erosion depth (determined as total erosion volume divided by the erosion area, $\bar{E}$) relative to the landslide length (ℓ). Further details are described in the Supplementary Material.

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

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

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

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

5.3 Calibration and model performance

Markov chains, colored according to the likelihood index, $L(\Lambda)$ are plotted in the S_c-q_c domain, along with histograms of sampled S_c and q_c values for each landslide in Figure 7. Each Markov chain includes 2000 model iterations. The runtime for 2000 model iterations depended on model domain and landslide size but varied from roughly 1.5 to 6 hours on a 2016 2.1 GHz Intel Core Xeon, 32 GB memory desktop. The chains show a wide array of sampling patterns and parameter ranges but broadly speaking, at all sites, the algorithm jumped within S_c-q_c space towards higher $L(\Lambda)$, to form bell-shaped posterior distributions for each parameter. Depending on the landslide type, the calibration algorithm converged on different S_c-q_c pairs. For example, at the Cascade Mountains site, the calibration utility converged to smaller q_c and S_c values for the 2009 event (Figure 7a), which permitted thinner flows over lower slopes and effectively made the 2009 modelled runout more mobile relative to the 2022 modelled runout (Figure 7b). At the Rocky Mountains site, a relatively high q_c value helps control lateral extent of the modelled runout that in the field was controlled by standing trees (Figure 7e).
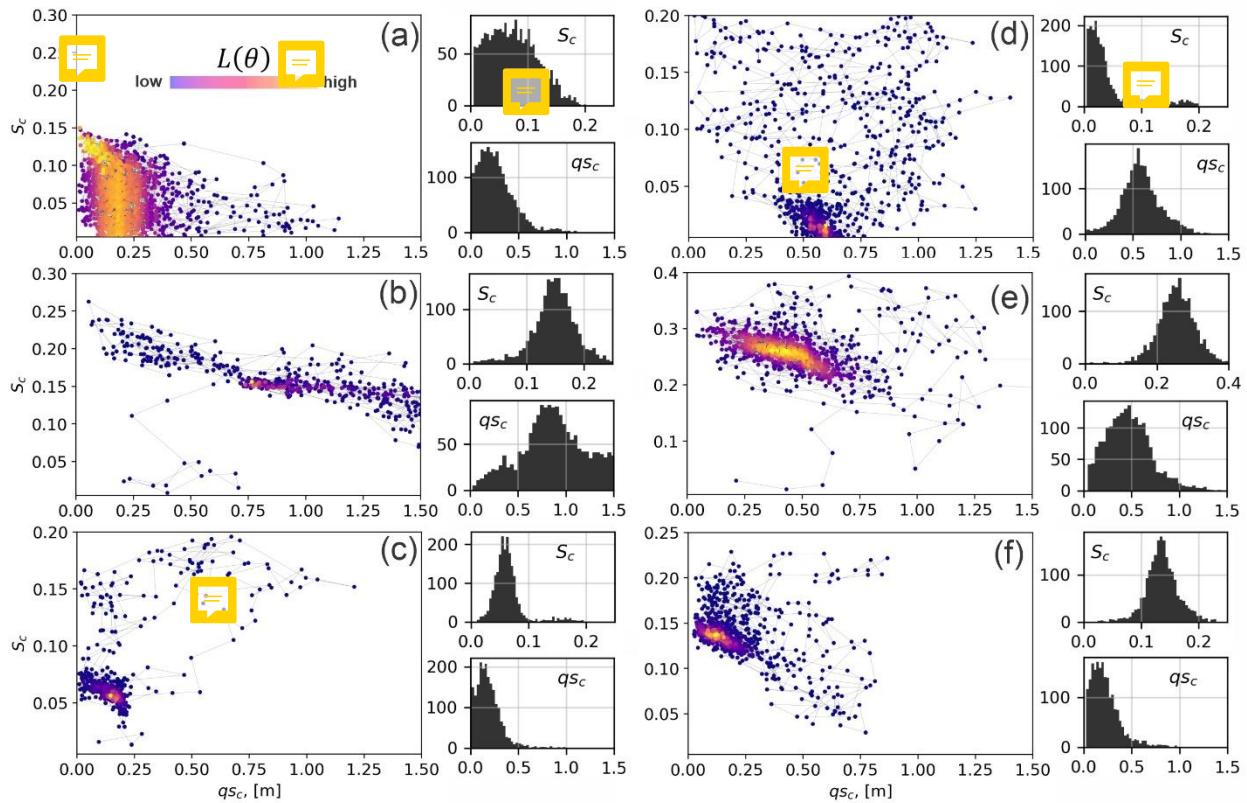
Figure 7. MWR calibration results for (a) Cascade Mountains, 2009; (b) Cascade Mountains, 2022, (c) Black Hills, South; (d) Black Hills, North; (e) Rocky Mountains and; (f) Olympic Mountains. Each result shows a scatter plot of the sampled S_c and q_c values, colored by their relative $L(\Lambda)$ value. To the right of each scatter plot are histograms of the iterated S_c and q_c parameters, which represent an empirical PDF of the possible S_c and q_c values that calibrate MWR to the site. Note y-axis scale differs between plots.

Profile plots of modelled Q_s and maps of the modelled planimetric runout extent, colored to indicate where the runout matched (α), overestimated (β) or underestimated (γ) the observed runout are shown in Figure 8. Values of Ω_T we

521   obtained with MWR are comparable or higher than reported values of Ω_T in the literature that used a variety of models

522   (Gorr et al., 2022; Barnhart et al., 2021; Note, to compare Ω_T values to those studies, subtract 1 from values reported

523   in this study). Across the sites, the volumetric error of the model, $\Delta\eta_E$, ranges between 6% and 15% (median 9.1%)

524   of the total mobilized volume from the observed DoD. An overall <10% volumetric error is reasonable considering

525   the low number of parameters required to calibrate MWR and that empirical estimates of total mobilized volume used

526   to run other runout models can vary by as much of an order of magnitude (e.g., Gartner et al., 2014: Barnhart et al.,

527   2021). Model performance in predicting volume flux along the runout profile was within similar error ranges. Except

528   for the Rocky Mountains site where MWR consistently modelled wider-than-observed flow, the cumulative flow error

529   along the runout profile (Q_{s_E}) were limited to 5%-19% of the mean cumulative flow determined from the observed

530   DoD.

531   MWR generally successfully replicates observed sediment transport along the runout path via model parameterizations

532   that are unique to each landslide. For example, the profile plots of Q_s at the Cascade Mountain site (Figure 8a and 8b)

533   show that during the 2009 landslide, all of the runout material flowed past the first 750 meters of the runout path.

534   During the 2022 landslide, material began to deposit just down slope of the initial landslide scar, as both observed and

535   modelled Q_s reverse slope, indicating loss in downstream volume flux. Model comparisons in the Cascade Mountains

536   site were limited to the upper 750 m of the hillslope because a large portion of the runout material was lost to fluvial

537   erosion in the valley (see Supplementary Material).

538   MWR also successfully replicates the observed sediment transport patterns at the Olympic Mountains site (profile plot

539   of Q_s in Figure 8f) and to a lesser degree, the Rocky Mountain site (Figure 8e). This finding is notable, because at the

540   Olympic Mountain site, observed runout extent and sediment depositional pattern were heavily impacted by woody

541   debris. Similarly, at the Rocky Mountains site, the width of the runout appeared to be restricted by trees. (See

542   Supplementary Material).

543   Using a fixed cell size of 10-m might have impacted model performance in some areas. MWR tended to over-estimate

544   the runout width for small landslides like the Olympic Mountains and Cascade Mountains, 2022 sites (yellow zones

545   in Figures 8f and 8b), likely because of the 10-m grid size used to represent the terrain. A 10-m DEM is generally

546   accepted as a good balance between model detail and computational limitations (e.g., Horton et al. 2013). However,

547   for small landslides, the 10-m grid is close to the size of the channels that controlled observed runout (Supplementary

548   Material) and may not have accurately represented the terrain. Modelled flow was less topographically-constrained

549   and tended to flow over a wider area of the terrain than observed in the more confined and smaller channels within

550   the axis of the runout valleys.

551   Because MWR does not have an explicit representation of flow momentum, it may show poor performance in regions

552   of the runout path where momentum controls runout extent. For example, at the Cascade Mountain, 2009 slide,

553   modelled extent misses a bench located along the east edge of the runout path (large red zone in Figure 8a). Review

554   of model behavior for this slide (Figure 9) shows how MWR successfully mimics diverging flow around a broad ridge

555   in the middle of the runout path (iteration t=28 in Figure 9), but afterword converges too rapidly into a narrow ravine

556   in the middle of the runout path (iteration t=40 in Figure 9). At the Rocky Mountains site, in addition to standing

557   trees, the forward momentum of the runout may have also restricted lateral spread of the observed runout. Modelled

558   runout is consistently too wide.

559   Overall, calibration was best at the Cascade Mountain, 2009 landslide (values of Ω_T are highest and values of $\Delta\eta_E$

560   and Q_{s_E} are lowest) and poorest at the Rocky Mountain and Olympic Mountain sites (Values of Ω_T are lowest Q_{s_E}

561   and $\Delta\eta_E$ are highest). At both the Rocky Mountain and Olympic Mountain sites, because we lacked repeat lidar, we

562   created the DoD from a map of field estimated erosion and deposition depths and estimated the pre-event DEM. The

563   lower calibration scores may indicate that field estimated DoDs were not as accurate as those determined via lidar

564   differencing. Another source of uncertainty that we have not addressed in our study is regolith thickness. Using

565   spatially accurate regolith thickness, rather than a uniform thickness, would likely improve MWR performance too.

566   Nonetheless, although imperfect, at most sites, MWR does not appear to have a strong systematic bias in modeled

567   output, which suggests that MWR may not have any structural weaknesses; however the consistent over-estimated

568   width on planar to divergent topography at the Rocky Mountain site requires further investigation at similar sites to

569   determine if this issue is due to calibration or the model.

570
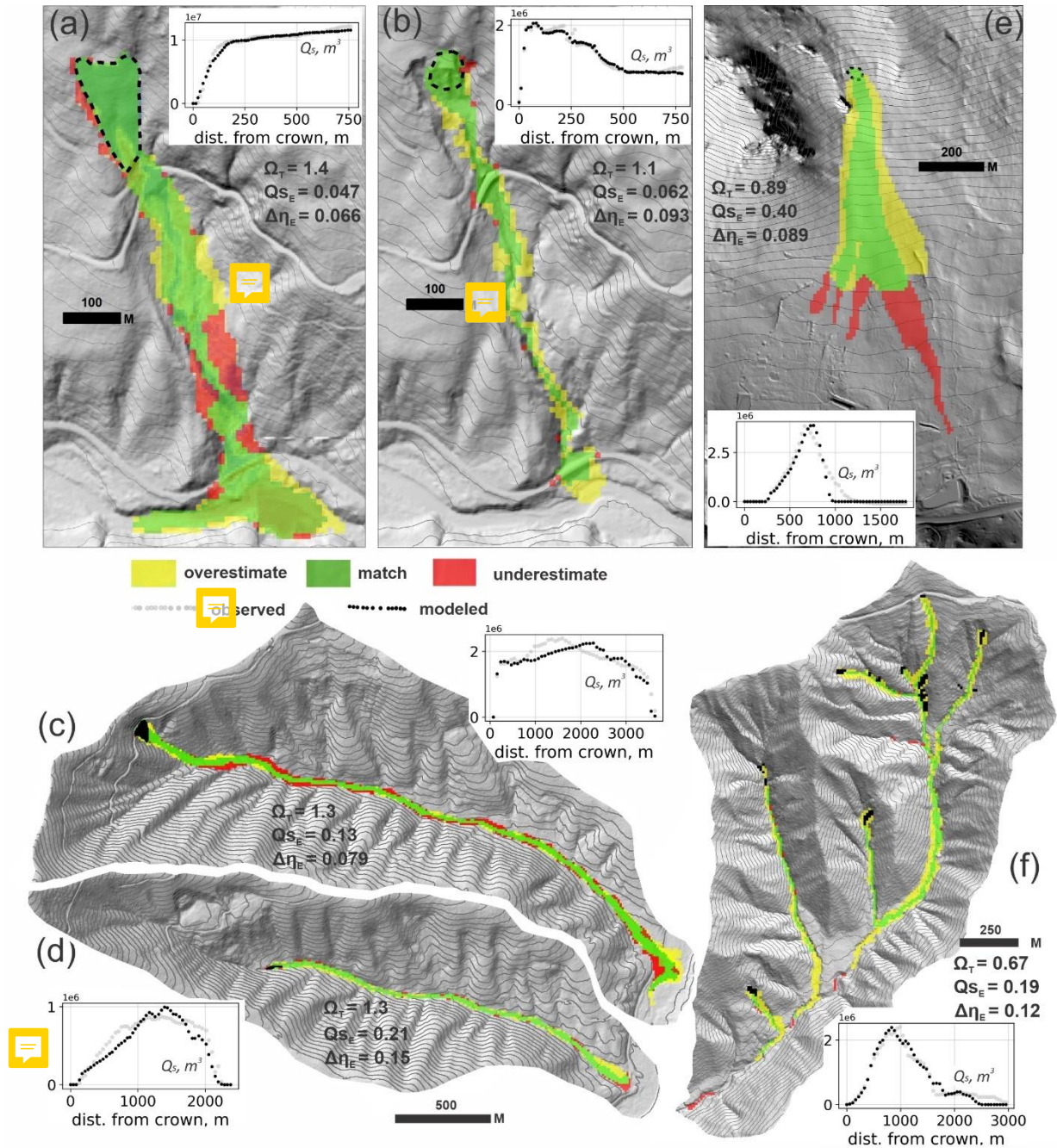
Figure 8. Calibrated model performance as indicated by modeled runout extent, profile plots of Q_s, and reported values of Ω_T, $\Delta\eta_E$ and Q_{s_E}. In all maps, up is north except in (e), north is towards the left. (a) Cascade Mountains, 2009; (b) Cascade Mountains, 2022; (c) Black Hills, North; (d) Black Hills, South; (e) Rocky Mountains; (f) Olympic Mountains.
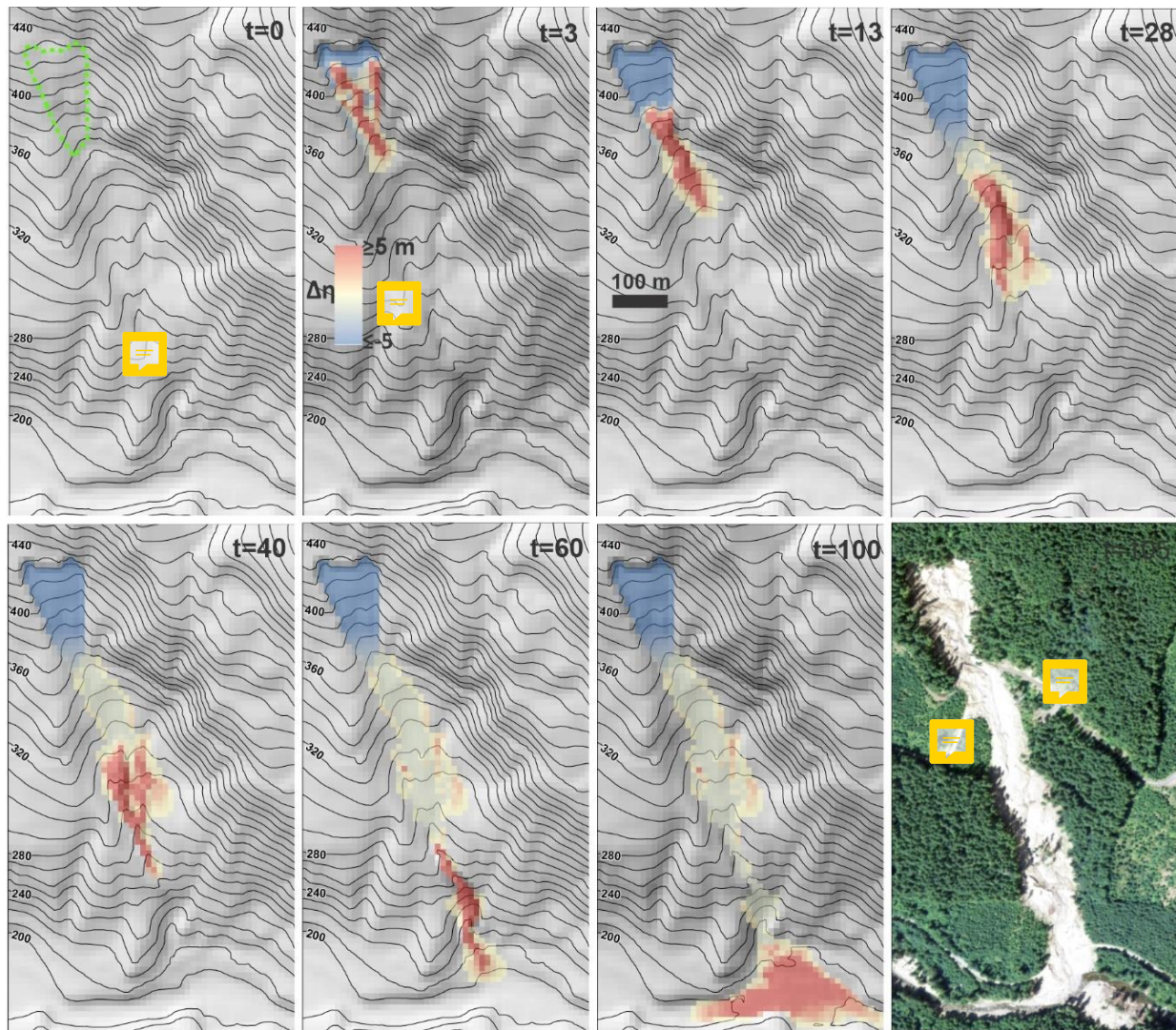
Figure 9. Illustration of modeled runout at the Cascade Mountains, 2009 landslide. At iteration $t = 0$, Algorithm 1 determines the direction and flux of the initial debritons over the slip surface of the landslide (all nodes located in the landslide green-dashed polygon). Note how the landslide slip surface directs the initial flow. In later iterations, Algorithm 2 routes the debritons down slope, updating the debritons and the terrain. By the end of the modeled runout, a colluvial fan forms at the base of the slope. Topography lines reflect the underlying terrain, which is updated after each iteration. MWR successfully replicates diverging flow at iteration $t = 28$ but misses a region of the observed runout path at iteration $t = 40$ where momentum likely controlled flow direction (compare to runout scar in air photo and underestimated region on topographic bench in Figure 8a)

To understand whether the ability to calibrate MWR systematically varies with topography of the runout path, we compared model performance with three topographic indices described by Chen & Yu (2011). The indices are computed from the terrain in the observed runout extent and include the relief ratio (H/L), mean total curvature (κ) and the mean specific stream power index (SPI). The index H/L equals the average slope of the runout path (or relative relief), determined as the total topographic relief of the runout (measured from the center of the landslide to the end of the runout path) divided by the horizontal length of the runout and indicates the mobility of the runout. Index κ represents topographic convergence, which is the second derivative of the terrain surface, with increasingly positive

594 values of index κ reflecting growing topographic convergence and concave-up channel profile (e.g., Istanbulluoglu et
595 al., 2008). The index *SPI* is determined as the natural log of the product of the contributing area and slope. Indices κ
596 and *SPI* are computed at each node in the runout extent and the mean values are computed from all nodes in the extent.
597 Comparison of model performance with respect to the topographic indices in Figure 10 shows: slightly improved
598 model performance over runout-paths that are less convergent (*SPI* and κ values of the observed runout path are
599 lower) and on steeper terrain (higher H/L) but neither trend is significant. The latter finding appears to be mostly a
600 result of how well modelled sediment transport and topographic change (Q_{s_E} and $\Delta\eta_E$) replicated observed, as there
601 does not appear to be a trend in Ω_T with H/L and the two best performing models (both Cascade Mountain landslides)
602 had the lowest (best) Q_{s_E} values and low $\Delta\eta_E$ values. Both findings are likely impacted by the grid size we used to
603 represent terrain. As noted above, at all sites we used a 10-m grid, but at some sites 10-m doesn't quite capture the
604 relief of channelized topography that controlled observed runout, leading to modelled runout that was considerably
605 wider than observed and causing low Ω_T value (this is especially true at the Olympic Mountains site, Figure 10a, b
606 and c).
607 In summary, using the calibration utility, we showed how the MWR can be calibrated to a range of different landslide
608 types and runout terrains. To a certain degree, though calibration, MWR can be parameterized to compensate for
609 deficiencies in the DEM or processes not explicitly represented in the model (momentum, woody debris). We were
610 unable to establish a clear pattern between calibration performance and topographic indices. This finding is likely
611 because numerous factors other than the terrain form, such as the DEM resolution, the quality of the DoD and
612 importance of processes not explicitly included in the model also impact performance.

613
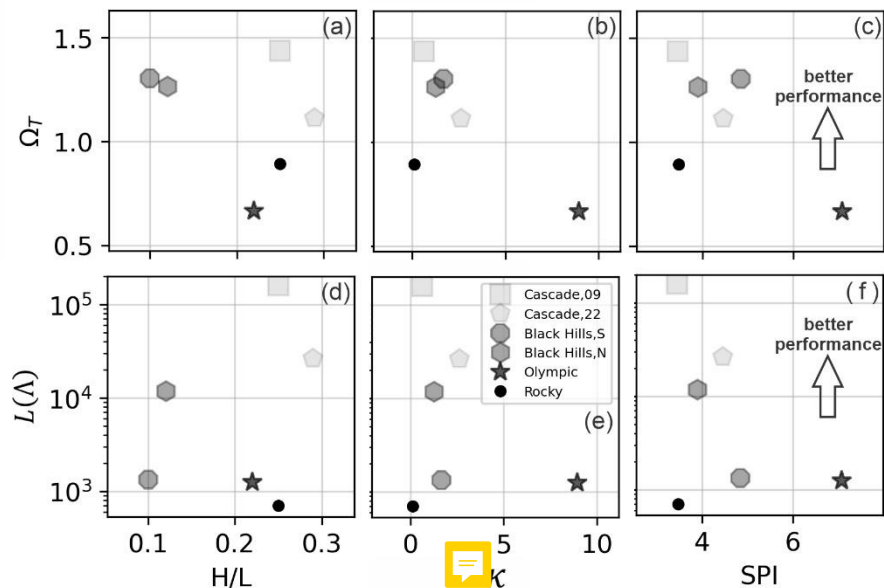
614
615 **Figure 10.** Illustration of model calibration, as reflected by the posterior parameter likelihood $L(\theta)$ and planimetric
616 fit (Ω_T) relative to topographic indices. There is no strong trend between the topographic indices and calibration
617 performance.

618   **6.   Discussion**

619   **6.1 Strategic testing of MWR for hazard mapping applications**

620   Having demonstrated basic model response to topography and that MWR can be calibrated to a variety of landslides

621   and runout terrains, we now strategically test MWR using the Cascade Mountain and Black Hills sites. Since both of

622   these sites include two separate landslides, we can thus test model performance by swapping best-fit model parameters

623   at each site, rerunning the models and comparing results with the original, calibrated results. At the Cascade Mountain

624   site, the 2009 and 2022 landslides originated on the same hillslope (Figure 8a and 8b). At Black Hills site, the two

625   landslides occurred on different hillslopes but in adjacent east-west oriented watersheds (Figure 8c and 8d).

626   As shown in Figure 11, at three of the landslides (both Cascade Mountain landslides and the Black Hills, North

627   landslide), when the best-fit parameters from the other landslide are used to predict runout, the accuracy of modelled

628   runout planimetric extent drops but resultant Ω_T values can still be as high or higher than values reported in other

629   studies (compare to equivalent Ω_T values in Gorr et al., 2022 and Barnhart et al., 2021). In terms of modelled sediment

630   transport and topographic change, swapping best-fit parameters has a more substantial effect. At the Cascade

631   Mountain, 2009 landslide, using the 2022 best-fit parameter values causes about half of the modelled runout material

632   to prematurely deposit on the hillslope, reducing the amount of sediment that reaches the valley floor (Q_{s_E} increases

633   by a factor of nine; Figure 11). Using the Cascade Mountain, 2009 parameter values on the Cascade Mountain, 2022

634   landslide (Figure 11b) increases modelled runout extent and results in nearly four times the entrainment and transport

635   of sediment to the valley floor, causing Q_{s_E} to increase by a factor of 20 and $\Delta\eta_E$ by 83%. At the Black Hills site,

636   using the South basin best-fit model parameters at the North basin causes Q_{s_E} and $\Delta\eta_E$ increase by 83% and 39%

637   respectively (Figure 11c). Unlike the other three landslides, swapping best-fit parameters at the Black Hills, South

638   landslide results in both large sediment transport and runout extent error because the North basin best-fit parameters

639   cause modelled landslide to entrain too little and stop only a few hundred meters from the initial source area (Figure

640   11d).

641   Although the need for calibration of MWR is a limitation for its transferability across sites, this limitation holds true

642   for most physics-based models. Barnhart et al. (2021) compared the ability of three different detailed-mechanistic

643   models to replicate an observed post-wildfire debris-flow runout event in California, USA. All three models used a

644   shallow-water-equation-based approach that conserved both mass and momentum, representing the flow as either a

645   single phase or double phase fluid. All models gave comparable results in simulating the event, suggesting that there

646   may not be a "true" best model. Despite the high level of detail and processes explicitly included in each model, all

647   models were sensitive to and required an estimate of the total mobilized volume, and the ability to replicate observed

648   runout ultimately depended on the selection of the parameters used to characterize debris flow properties.

649

650
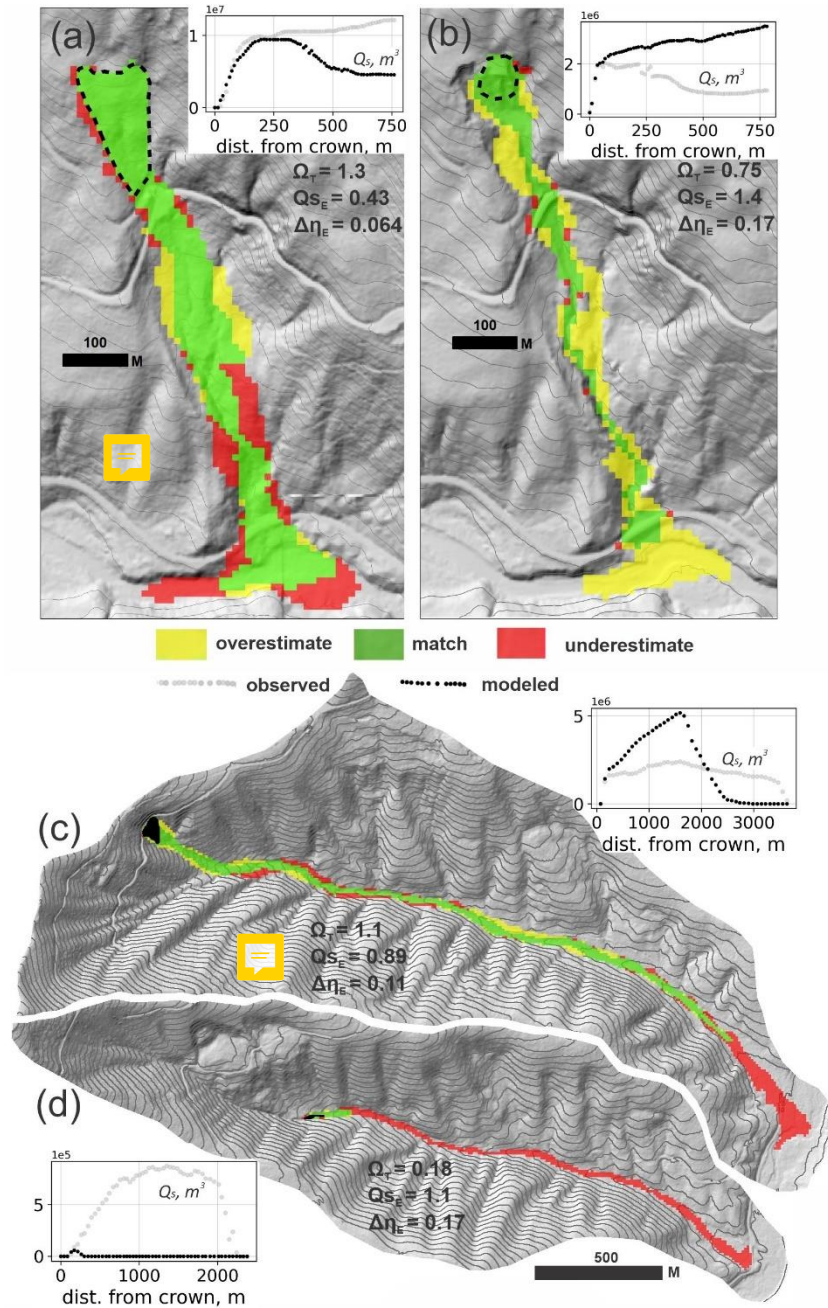
Figure 11. Model performance using the neighboring landslide parameter values, as indicated by modeled runout extent, profile plots of Q_s, reported values of Ω_T, $\Delta\eta_E$ and Q_{s_E} . Compare with Figure 8. (a) Cascade Mountain, 2009; (b) Cascade Mountain, 2022; (c) Black Hills, North; (d) Black Hills, South

As landslide hazard models often forecast hazard probabilistically, an alternative test to simply swapping the best-fit parameters is to swap parameter PDFs determined from the calibration utility and compare probability of runout at each model node (equation 30). As shown in Figure 12, similar to the first test, at three of the landslides, using the parameter distribution associated with the neighbouring landslide results in relatively minor changes in whether runout

660 is likely to occur versus not occur (probability of runout ≥50%; Figures 12a, 12b and 12d). At the Black Hills South

661 landslide, swapping parameter PDFs causes a large change in runout probability (Figure 12c).

662
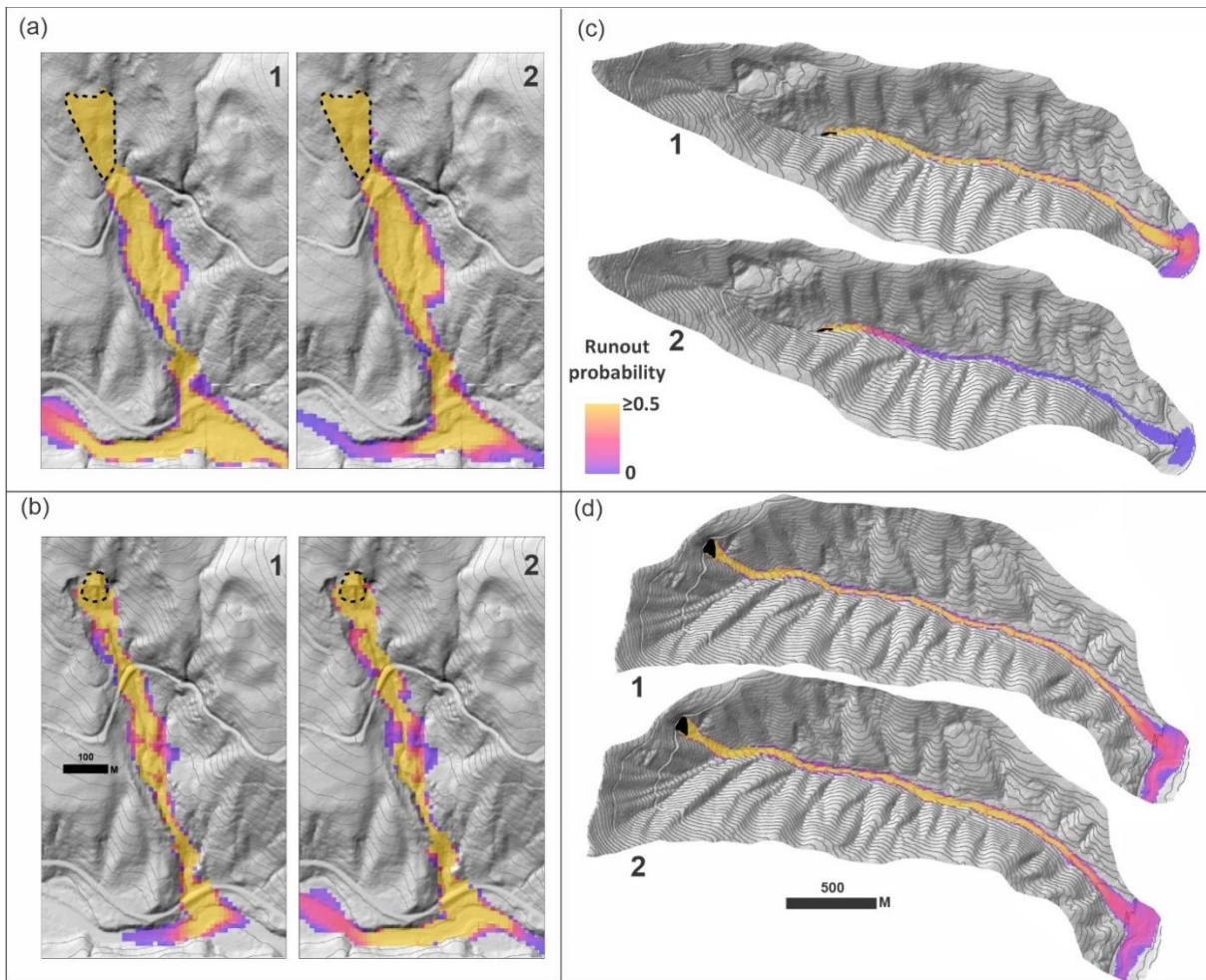
663
664 **Figure 12.** Model tests by swapping parameter PDFs and comparing runout probability at the (a) Cascade Mountain,
665 2009; (b) Cascade Mountain, 2022; (c) Black Hills, South and; (d) Black Hills, North sites. (**1**) runout using
666 parameter distributions of the site and (**2**) runout using parameter distributions of the neighboring site.

667

668 The results of these two tests suggest that in most cases, once best-fit parameters or parameter PDFs have been

669 established for a landslide, those parameter/PDF values may be useful for assessing runout extent but not useful for

670 sediment transport and topographic change prediction at nearby sites. However, we suspect that these results are a

671 consequence of comparing very different landslide types and runout processes. In regions where landslide processes

672 are relatively uniform (like the Olympic Mountain site), calibration to one landslide might be sufficient to predict the

673 depositional patterns of another. At sites like the Cascade Mountain and Black Hills sites, which consisted of a diverse

674 range of landslide processes including small, confined debris flows to large, unconfined debris avalanches, MWR may

675 need to be calibrated to each type of landslide and predictive applications might involve applying the appropriate

676 parameter set based on landslide type.

677 **6.2. MassWastingRunout probability applications**

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

684 **6.2.1. Runout probability from a landslide hazard map**

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

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

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

704 **6.2.2 Runout probability for a specific, potentially unstable slope**

705 When field evidence or other data indicate that a specific hillslope may be potentially unstable, but the exact area of
706 a potential landslide on that slope is unknown, MWR can be used to generate a hazard estimate that takes into account
707 the uncertainty in the landslide area. For this application, MWR Probability is run using option (2), which requires a
708 polygon representing the extent of the potentially unstable slope. For each model repetition, a landslide area can form
709 anywhere within the potentially unstable slope and is at least as large as a user defined minimum size but no larger
710 than the potentially unstable slope.

711   As an example application of using MWR Probability option (2), we designated a 0.6 ha, convergent hillslope in the

712   headwaters of the Olympic Mountains site as a potentially unstable slope and modelled runout probability, again using

713    $Np = 1000$ (Figure 13d). This example shows that, given uncertainty in the landslide size and location, and

714   uncertainty in calibrated parameterization of MWR, if a landslide were to initiate on the potentially unstable slope,

715   the probability of the runout reaching the basin outlet is less than 5%.
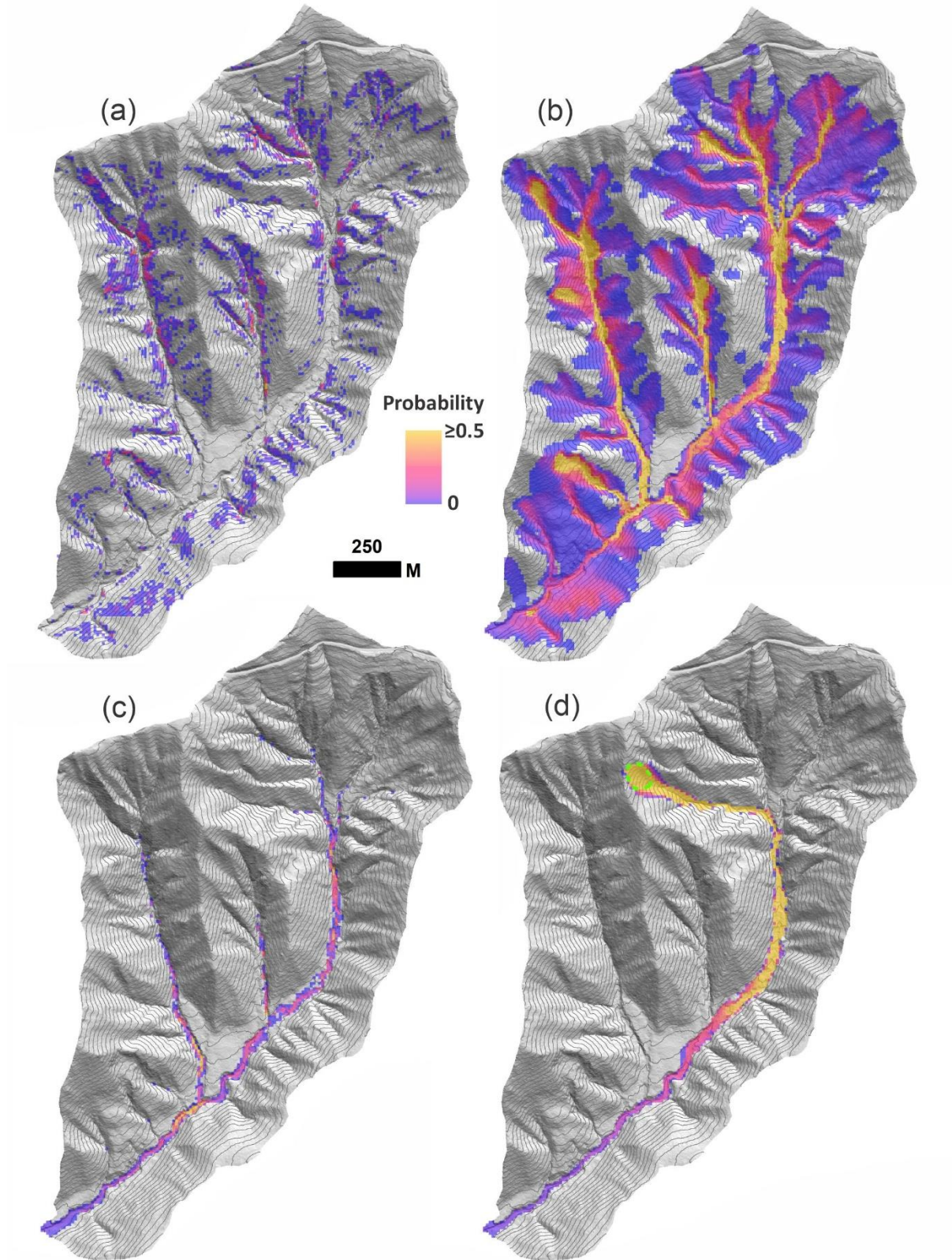
716
Figure 13. Olympic Mountain site: **(a)** Landslide probability, $P(FS \leq 1)$. **(b)** Corresponding runout probability,
$P(\Delta\eta)$. **(c)** Probability of deposition greater than 1 m and **(d)** Runout probability for the potentially unstable slope
(green-dashed polygon).

7.0 Concluding remarks

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

Results show that despite its simple conceptualization, MWR can replicate observed erosion, deposition and sediment transport patterns. A notable finding of this paper is that MWR modeled runout did not have any strong systematic bias in predictions (toward unrealistically short or wide flows, for example), which suggests that MWR may not have any structural weaknesses. When compared to other models capable of replicating inundation patterns of observed runout events, the strength of MWR lies in its potential computation efficiency, use of field-inferable parameters, limited reliance on calibration parameters (only two, critical slope, S_c, and a threshold flux for deposition, q_c) and its ability to internally estimate the total mobilized volume. MWR needs only the location and geometry of an initial source area to model the entire runout process.

MWR shows a rich set of intuitive responses to topographic curvature and slope and model performance over a range of landslide and landscape conditions across the four sites we used for this study was sufficiently controlled with the two calibration parameters. When calibrated to each individual site, the volumetric error of MWR, $\Delta\eta_E$, ranged between 6% and 15% (median 9.1%) of the observed total mobilized volume. Except for the Rocky Mountains site where MWR consistently modelled wider-than-observed flow, the cumulative flow error along the runout profile (Q_{s_E}) were limited to 5%-19% of the mean cumulative flow determined from the observed DoD. These are considered acceptable levels of performance given that the total mobilized volume of many debris flow models assume an order of magnitude range of confidence.

Once MWR is calibrated to runout observations, it can be linked to other landslide hazard models and may be useful as a regional runout hazard mapping tool in areas with relatively uniform landslide processes. In this study we showed how to use MWR to map debris flow hazard for an expert-defined potentially unstable slope and for a landslide hazard map produced from an externally run Monte Carlo landslide initiation model (Figure 13).

As a component of the Landlab earth surface modelling toolkit, MWR is designed to be compatible with other models. MWR can be readily coupled with a landslide initiation model (e.g., LandslideProbability) and geomorphic transport laws for hillslope diffusion and fluvial incision to investigate the role of landslides and their runout on long-term landscape evolution. We did not explore the use of MWR in landscape evolution or sediment yield models in this study, however its ability to replicate observed topographic change and sediment transport at multiple sites shows promise for this application. 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.

757 **8.0 Notation**

758	q_{R_i}	[m]	debris flux from a node to each of the node i-th receiver nodes
759	q_O	[m]	the total out-going debris flux
760	Nr		the number of receiving nodes of node n
761	S_i		the underlying topographic slope ($\tan\theta$) to each of the node i-th receiver nodes
762	a		exponent in (1) that controls how flow is distributed to downslope nodes
763	q_I	[m]	The total incoming flux
764	Nd		number of donors nodes to a node
765	q_{D_j}	[m]	the flux from node D_j (the j-th donor node)
766	h	[m]	flow depth at node, adjusted to be no more than h_{max}
767	h_{max}	[m]	the maximum observed flow depth
768	A	[m]	aggradation depth
769	S_c		critical slope
770	S		steepest slope to the node's eight neighbouring nodes
771	Δx	[m]	cell length
772	$A_{p|N_a}$	[m]	potential aggradation depth that forms a deposit that spreads over N_a consecutive nodes
773	$A_{p,i}$	[m]	i-th deposition amount in the deposit illustrated in Figure 4
774	N_a		number of nodes qs_n^I is assumed to spreads over
775	E	[m]	erosion depth
776	h_r	[m]	regolith depth
777	h_e	[m]	potential erosion depth
778	θ	[°]	topographic slope used to determine shear stress, equal to $\tan^{-1}(S)$
779	τ	[Pa]	basal shear stress
780	τ_c	[Pa]	critical shear stress of the regolith
781	k		erodibility parameter in (11)
782	f		exponent, controls the non-linearity of h_e in (11)
783	ρ	[kg/m^3]	density of runout material
784	σ	[Pa]	normal stress at basal surface
785	φ		tangent of collision angle between grains, measured from the vertical axis
786	v_s		volumetric solids concentration
787	ρ_s	[kg/m^3]	density of solids
788	D_s	[m]	characteristic particle diameter
789	u	[m/s]	depth average flow velocity
790	z	[m]	depth below the flow surface
791	u^*		shear velocity
792	g	[m/s]	acceleration due to gravity
793	$\Delta\eta$	[m]	change in elevation at node
794	ξ_D		attribute value delivered to the node
795	ξ_R		attribute value sent to receiver nodes
796	ξ		attribute value at node
797	Λ		parameter set
798	$L(\Lambda)$		likelihood of parameter set
799	$p(\Lambda)$		prior probability of parameter set
800	Ω_T		omega metric, nondimensional
801	α	[m^2]	modelled area of matching extent (compared to observed runout extent)
802	β	[m^2]	modelled area of overestimated extent
803	γ	[m^2]	modelled area of underestimated extent
804	$\Delta\eta_E$		volumetric error of the modelled topographic change relative to the observed total
805			mobilized volume, fraction.
806	V	[m^3]	observed total mobilized volume
807	p		the number of nodes in the modelled runout extent
808	$\Delta\eta_{Mi}$	[m]	the modelled topographic change [m] at the i-th node within the runout extent
809	$\Delta\eta_{Oi}$	[m]	the observed topographic change [m] at the i-th node within the runout extent

810			
811	Q_{s_E}		mean-modelled-cumulative flow error along the runout path relative to the observed
812			mean cumulative flow, fraction.
813	$Q_{s\,j}$	[m³]	the cumulative debris flow volume (Q_s) at each node, j,
814	$\Delta\eta_{ij}$	[m]	the topographic change [m] at the i-th node located upstream of node j
815	u_j		the total number of all nodes located upstream of j
816	$\overline{Q_{sO}}$	[m³]	the observed mean cumulative flow
817	$P(\Delta\eta)$		probability of runout at a model node
818	Np		number Monte Carlo iterations used to determine probability

9.0 Acknowledgements

10.0 Competing interests

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

11.0 References

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

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

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

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

844   Benda, L., & Dunne, T.: Stochastic forcing of sediment supply to channel networks from landsliding and debris flow.

845   Water Resources Research, 33(12), 2849–2863. https://doi.org/10.1029/97wr02388, 1997.

846   Benda, L., Veldhuisen, C. P., & Black, J.: Debris flows as agents of morphological heterogeneity at low-order

847   confluences, Olympic Mountains, Washington. Geological Society of America Bulletin, 115(9), 1110.

848   https://doi.org/10.1130/b25265.1, 2003.

849   Beven, K.: A manifesto for the equifinality thesis. Journal of Hydrology, 320(1–2), 18–36.

850   https://doi.org/10.1016/j.jhydrol.2005.07.007, 2006.

851   Burton, A., & Bathurst, J. C.: Physically based modelling of shallow landslide sediment yield at a catchment scale.

852   Environmental Geology, 35(2–3), 89–99. https://doi.org/10.1007/s002540050296, 1998.

853   Bigelow, P., Benda, L., Miller, D., & Burnett, K. M.: On Debris Flows, River Networks, and the Spatial Structure of

854   Channel Morphology. Forest Science, 53(2), 220–238. https://doi.org/10.1093/forestscience/53.2.220, 2007

855   Campforts, B., Shobe, C. M., Overeem, I., & Tucker, G. E.: The Art of Landslides: How Stochastic Mass Wasting

856   Shapes Topography and Influences Landscape Dynamics. Journal of Geophysical Research: Earth Surface, 127(8).

857   https://doi.org/10.1029/2022jf006745, 2022

858   Campforts, B., Shobe, C. M., Steer, P., Vanmaercke, M., Lague, D., & Braun, J.: HyLands 1.0: a hybrid landscape

859   evolution model to simulate the impact of landslides and landslide-derived sediment on landscape evolution.

860   Geoscientific Model Development, 13(9), 3863–3886. https://doi.org/10.5194/gmd-13-3863-2020, 2020.

861   Capart, H., & Fraccarollo, L.: Transport layer structure in intense bed-load. Geophysical Research Letters, 38(20),

862   n/a. https://doi.org/10.1029/2011gl049408, 2011.

863   Capart, H., Hung, C., & Stark, C. R.: Depth-integrated equations for entraining granular flows in narrow channels.

864   Journal of Fluid Mechanics, 765. https://doi.org/10.1017/jfm.2014.713, 2015.

865   Carretier, S., Martinod, P., Reich, M., & Goddéris, Y.: Modelling sediment clasts transport during landscape evolution.

866   Earth Surface Dynamics, 4(1), 237–251. https://doi.org/10.5194/esurf-4-237-2016, 2016.

867   Chase, C. G.: Fluvial landsculpting and the fractal dimension of topography. Geomorphology, 5(1–2), 39–57.

868   https://doi.org/10.1016/0169-555x(92)90057-u, 1992.

869   Chen, C., & Yu, F.: Morphometric analysis of debris flows and their source areas using GIS. Geomorphology, 129(3–

870   4), 387–397. https://doi.org/10.1016/j.geomorph.2011.03.002, 2011.

871   Chen, H., & Zhang, L.: EDDA 1.0: integrated simulation of debris flow erosion, deposition and property changes.

872   Geoscientific Model Development, 8(3), 829–844. https://doi.org/10.5194/gmd-8-829-2015, 2015.

873   Chen, T.-Y. K., Wu, Y.-C., Hung, C.-Y., Capart, H., and Voller, V. R.: A control volume finite-element model for

874   predicting the morphology of cohesive-frictional debris flow deposits, Earth Surface Dynamics, 11, 325–342,

875   https://doi.org/10.5194/esurf-11-325-2023, 2023.

876   Clerici, A., & Perego, S.: Simulation of the Parma River blockage by the Corniglio landslide (Northern Italy).

877   Geomorphology, 33(1–2), 1–23. https://doi.org/10.1016/s0169-555x(99)00095-1, 2000.

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

879   Coz, J. L., Renard, B., Bonnifait, L., Branger, F., & Boursicaud, R. L.: Combining hydraulic knowledge and uncertain

880   gaugings in the estimation of hydrometric rating curves: A Bayesian approach. Journal of Hydrology, 509, 573–587.

881   https://doi.org/10.1016/j.jhydrol.2013.11.016, 2014.

882   Crave, A., & Davy, P.: A stochastic "precipiton" model for simulating erosion/sedimentation dynamics. Computers

883   & Geosciences, 27(7), 815–827. https://doi.org/10.1016/s0098-3004(00)00167-9, 2001.

884   D'Ambrosio, D., Di Gregorio, S., Iovine, G., Lupiano, V., Rongo, R., & Spataro, W.: First simulations of the Sarno

885   debris flows through Cellular Automata modelling. Geomorphology, 54(1–2), 91–117. https://doi.org/10.1016/s0169-

886   555x(03)00058-8, 2003.

887   Egashira, S., Honda, N., & Itoh, T.: Experimental study on the entrainment of bed material into debris flow. Physics

888   and Chemistry of the Earth, Parts a/B/C, 26(9), 645–650. https://doi.org/10.1016/s1464-1917(01)00062-9, 2001.

889   Foglia, L., Hill, M. C., Mehl, S. W., and Burlando, P. (2009), Sensitivity analysis, calibration, and testing of a

890   distributed hydrological model using error-based weighting and one objective function, Water Resources

891   Research, 45, W06427, doi:10.1029/2008WR007255.

892   Fannin, R. J., & Wise, M. P.: An empirical-statistical model for debris flow travel distance. Canadian Geotechnical

893   Journal, 38(5), 982–994. https://doi.org/10.1139/t01-030, 2001.

894   Frank, F., McArdell, B. W., Huggel, C., & Vieli, A.: The importance of entrainment and bulking on debris flow runout

895   modeling: examples from the Swiss Alps. Natural Hazards and Earth System Sciences, 15(11), 2569–2583.

896   https://doi.org/10.5194/nhess-15-2569-2015, 2015.

897   Freeman, T. G.: Calculating catchment area with divergent flow based on a regular grid. Computers &

898   Geosciences, 17(3), 413–422. https://doi.org/10.1016/0098-3004(91)90048-i, 1991.

899   Gartner, J. E., Cannon, S. H., & Santi, P. M.: Empirical models for predicting volumes of sediment deposited by debris

900   flows and sediment-laden floods in the transverse ranges of southern California. Engineering Geology, 176, 45–56.

901   https://doi.org/10.1016/j.enggeo.2014.04.008, 2014.

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

903   Electronic Edition, 2021.

904   Goode, J. R., Luce, C. H., & Buffington, J. M.: Enhanced sediment delivery in a changing climate in semi-arid

905   mountain basins: Implications for water resource management and aquatic habitat in the northern Rocky Mountains.

906   Geomorphology, 139–140, 1–15. https://doi.org/10.1016/j.geomorph.2011.06.021, 2012.

907   Gorr, A., McGuire, L. A., Youberg, A., & Rengers, F. K.: A progressive flow-routing model for rapid assessment of

908   debris-flow inundation. Landslides, 19(9), 2055–2073. https://doi.org/10.1007/s10346-022-01890-y, 2022

909   Guthrie, R., Hockin, A., Colquhoun, L., Nagy, T., Evans, S. G., & Ayles, C. P.: An examination of controls on debris

910   flow mobility: Evidence from coastal British Columbia. Geomorphology, 114(4), 601–613.

911   https://doi.org/10.1016/j.geomorph.2009.09.021, 2010.

912   Guthrie, R. H., & Befus, A. D.: DebrisFlow Predictor: an agent-based runout program for shallow landslides. Natural

913   Hazards and Earth System Sciences, 21(3), 1029–1049. https://doi.org/10.5194/nhess-21-1029-2021, 2021.

914   Hammond C.J., Prellwitz R.W., Miller S.M.: Landslides hazard assessment using Monte Carlo simulation. In: Bell

915   DH (ed) Proceedings of 6th international symposium on landslides, Christchurch, New Zealand, Balkema, vol 2. pp

916   251–294, 1992.

917   Han, Z., Chen, G., Li, Y., Tang, C., Xu, L., He, Y., Huang, X., & Wang, W.: Numerical simulation of debris-flow

918   behavior incorporating a dynamic method for estimating the entrainment. Engineering Geology, 190, 52–64.

919   https://doi.org/10.1016/j.enggeo.2015.02.009, 2015.

920   Han, Z., Li, Y., Huang, J., Chen, G., Xu, L., Tang, C. Y., Zhang, H., & Shang, Y.: Numerical simulation for run-out

921   extent of debris flows using an improved cellular automaton model. Bulletin of Engineering Geology and the

922   Environment, 76(3), 961–974. https://doi.org/10.1007/s10064-016-0902-6, 2017.

923   Han, Z., Ma, Y., Li, Y., Zhang, H., Chen, N., Hu, G., & Chen, G.: Hydrodynamic and topography based cellular

924   automaton model for simulating debris flow run-out extent and entrainment behavior. Water Research, 193, 116872.

925   https://doi.org/10.1016/j.watres.2021.116872, 2021.

926   Heiser, M., Scheidl, C., & Kaitna, R.: Evaluation concepts to compare observed and simulated deposition areas of

927   mass movements. Computational Geosciences, 21(3), 335–343. https://doi.org/10.1007/s10596-016-9609-9, 2017.

928   Hobley, D. E. J., Adams, J. M., Nudurupati, S. S., Hutton, E. W. H., Gasparini, N. M., Istanbulluoglu, E., & Tucker,

929   G. E.: Creative computing with Landlab: an open-source toolkit for building, coupling, and exploring two-dimensional

930   numerical models of Earth-surface dynamics. Earth Surface Dynamics, 5(1), 21–46. https://doi.org/10.5194/esurf-5-

931   21-2017, 2017.

932   Horton, P., Jaboyedoff, M., Rudaz, B., & Zimmermann, M. N.: Flow-R, a model for susceptibility mapping of debris

933   flows and other gravitational hazards at a regional scale. Natural Hazards and Earth System Sciences, 13(4), 869–885.

934   https://doi.org/10.5194/nhess-13-869-2013, 2013.

935   Hungr, O., Morgan, G. J., & Kellerhals, R.: Quantitative analysis of debris torrent hazards for design of remedial

936   measures. Canadian Geotechnical Journal, 21(4), 663–677. https://doi.org/10.1139/t84-073, 1984.

937   Hungr, O., & Evans, S. G.: Entrainment of debris in rock avalanches: An analysis of a long run-out mechanism.

938   Geological Society of America Bulletin, 116(9), 1240. https://doi.org/10.1130/b25362.1, 2004.

939   Hutter, K., Svendsen, B., & Rickenmann, D.: Debris flow modeling: A review. Continuum Mechanics and

940   Thermodynamics, 8(1), 1–35. https://doi.org/10.1007/bf01175749, 1996.

941 Iovine, G., D'Ambrosio, D., & Di Gregorio, S.: Applying genetic algorithms for calibrating a hexagonal cellular
942 automata model for the simulation of debris flows characterised by strong inertial effects. Geomorphology, 66(1–4),
943 287–303. https://doi.org/10.1016/j.geomorph.2004.09.017, 2005.

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

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

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

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

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

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

960 Julien, P. Y., & Paris, A.: Mean Velocity of Mudflows and Debris Flows. Journal of Hydraulic Engineering, 136(9),
961 676–679. https://doi.org/10.1061/(asce)hy.1943-7900.0000224, 2010.

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

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

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

969 Larsen, I. J., & Montgomery, D. R.: Landslide erosion coupled to tectonics and river incision. Nature Geoscience,
970 5(7), 468–473. https://doi.org/10.1038/ngeo1479, 2012.

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

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

975 Major, J. J., & Iverson, R. M.: Debris-flow deposition: Effects of pore-fluid pressure and friction concentrated at flow
976 margins. Geological Society of America Bulletin, 111(10), 1424–1434. https://doi.org/10.1130/0016-7606(1999)111,
977 1999.

978 McCoy, S. W., Kean, J. W., Coe, J. A., Tucker, G. S., Staley, D. M., & Wasklewicz, T. A.: Sediment entrainment by
979 debris flows: In situ measurements from the headwaters of a steep catchment. Journal of Geophysical Research,
980 117(F3), n/a. https://doi.org/10.1029/2011jf002278, 2012.

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

983 Medina, V., Hürlimann, M., & Bateman, A.: Application of FLATModel, a 2D finite volume code, to debris flows in
984 the northeastern part of the Iberian Peninsula. Landslides, 5(1), 127–142. https://doi.org/10.1007/s10346-007-0102-
985 3, 2008.

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

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

990 Murray, A. B., & Paola, C.: Properties of a cellular braided-stream model. Earth Surface Processes and Landforms,
991 22(11), 1001–1025. https://doi.org/10.1002/(sici)1096-9837(199711)22:11, 1997.

992

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

995 Natural Resources Conservation Service | Snow and Water Interactive Map (n.d.). Natural Resources Conservation
996 Service. https://www.nrcs.usda.gov/resources/data-and-reports/snow-and-water-interactive-map, Accessed April,
997 2022

998 Nudurupati, S. S., Istanbulluoglu, E., Tucker, G. E., Gasparini, N. M., Hobley, D. E. J., Hutton, E. W. H., Barnhart,
999 K. R., & Adams, J. M.: On transient semi-arid ecosystem dynamics using Landlab: vegetation shifts, topographic
1000 refugia, and response to climate. Water Resources Research, 59(4). https://doi.org/10.1029/2021wr031179, 2023.

1001 Perron, J. T.: Climate and the Pace of Erosional Landscape Evolution. Annual Review of Earth and Planetary Sciences,
1002 45(1), 561–591. https://doi.org/10.1146/annurev-earth-060614-105405, 2017.

1003 Reid, M. J., Coe, J. A., & Brien, D. L.: Forecasting inundation from debris flows that grow volumetrically during
1004 travel, with application to the Oregon Coast Range, USA. Geomorphology, 273, 396–411.
1005 https://doi.org/10.1016/j.geomorph.2016.07.039, 2016.

1006 Renard, B., Garreta, V., & Lang, M. J.: An application of Bayesian analysis and Markov chain Monte Carlo methods
1007 to the estimation of a regional trend in annual maxima. Water Resources Research, 42(12).
1008 https://doi.org/10.1029/2005wr004591, 2006.

1009 Rengers, F. K., McGuire, L. A., Kean, J. W., Staley, D. M., and Hobley, D. E. J.: Model simulations of flood and
1010 debris flow timing in steep catchments after wildfire, Water Resources Research, 52, 6041–6061,
1011 doi:10.1002/2015WR018176, 2016.

1012 Roda-Boluda, D. C., D'Arcy, M., McDonald, J., & Whittaker, A. C.: Lithological controls on hillslope sediment
1013 supply: insights from landslide activity and grain size distributions. Earth Surface Processes and Landforms, 5), 956–
1014 977. https://doi.org/10.1002/esp.4281, 2018.

1015 Shaller, P. J., Doroudian, M., & Hart, M. W.: The Eureka Valley Landslide: Evidence of a dual failure mechanism for
1016 a Long-Runout Landslide. Lithosphere, 2020(1). https://doi.org/10.2113/2020/8860819, 2020.

1017 Schürch, P., Densmore, A. L., Rosser, N., & McArdell, B. W.: Dynamic controls on erosion and deposition on debris-
1018 flow fans. Geology, 39(9), 827–830. https://doi.org/10.1130/g32103.1, 2011.

1019 Shen, P., Zhang, L. M., Wong, H., Peng, D., Zhou, S., Zhang, S., & Chen, C.: Debris flow enlargement from
1020 entrainment: A case study for comparison of three entrainment models. Engineering Geology, 270, 105581.
1021 https://doi.org/10.1016/j.enggeo.2020.105581, 2020.

1022 Stock, J. P. J., & Dietrich, W. E.: Erosion of steepland valleys by debris flows. Geological Society of America Bulletin,
1023 118(9–10), 1125–1148. https://doi.org/10.1130/b25902.1, 2006.

1024 Strauch, R. L., Istanbulluoglu, E., Nudurupati, S. S., Bandaragoda, C., Gasparini, N. M., & Tucker, G. E.: A
1025 hydroclimatological approach to predicting regional landslide probability using Landlab. Earth Surface Dynamics,
1026 6(1), 49–75. https://doi.org/10.5194/esurf-6-49-2018, 2018.

1027 Takahashi, T.: Debris Flow (2nd ed.). CRC Press, Taylor & Francis Group, 2014.

1028 Tucker, G. E., & Bras, R. L.: Hillslope processes, drainage density, and landscape morphology. Water Resources
1029 Research, 34(10), 2751–2764. https://doi.org/10.1029/98wr01474, 1998.

1030 Tucker, G. E., Hancock, G. J.: Modelling landscape evolution. Earth Surface Processes and Landforms, 35(1), 28–50.
1031 https://doi.org/10.1002/esp.1952, 2010.

1032 Tucker, G. E., McCoy, S., & Hobley, D. E. J.: A lattice grain model of hillslope evolution. Earth Surface Dynamics,
1033 6(3), 563–582. https://doi.org/10.5194/esurf-6-563-2018, 2018.

1034 Western Regional Climate Center. (n.d.), from https://wrcc.dri.edu/, accessed 2017 and 2022

1035 Whipple, K. X., & Dunne, T.: The influence of debris-flow rheology on fan morphology, Owens Valley, California.
1036 Geological Society of America Bulletin, 104(7), 887–900. https://doi.org/10.1130/0016-7606(1992)104, 1992.

1037   Zhou, G. G. D., Li, S., Song, D., Choi, C. E., & Chen, X.: Depositional mechanisms and morphology of debris flow:

1038   physical modelling. Landslides, 16(2), 315–332. https://doi.org/10.1007/s10346-018-1095-9, 2019.

1039

1040