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

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

5 [1] Washington Department of Natural Resources, Forest Resources Division, Olympia, WA, USA
6 [2] University of Washington, Civil and Environmental Engineering, Seattle, WA, USA
7 [3] Institute of Arctic and Alpine Research, University of Colorado Boulder, Boulder, CO, USA
8 [4] Department of Geological Sciences, University of Colorado Boulder, Boulder, CO, USA
9 [5] Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado Boulder, Boulder,
10 CO, USA
11
12 *Correspondence to*: Jeffrey Keck (keckje@gmail.com)

13 **Abstract**

14 We developed a new rule-based, cellular-automaton algorithm for predicting the hazard extent, sediment transport and
15 topographic change associated with the runout of a landslide. This algorithm, which we call MassWastingRunout
16 (MWR), is coded in Python and implemented as a component for the package Landlab. Given the location and
17 geometry of an initial landslide body (i.e., landslide polygon), MWR models the downslope progression of the runout
18 process and evolves the underlying terrain. Runout behavior is controlled by mass continuity, topography, and rules
19 for erosion and deposition, which can be informed from field observations. MWR includes a calibration utility that
20 uses a Markov Chain Monte Carlo algorithm to sample model parameter space and tune the model to match observed
21 patterns of landslide runout extent, deposition and erosion. Output from the calibration utility informs probabilistic
22 implementation of MWR. Here we demonstrate calibrated model performance relative to a range of observed runout
23 phenomena and terrain, including debris flows in channelized, low-energy-dissipation terrains and debris avalanches
24 on open-slope, moderate, to high-energy-dissipation terrains. We test model ability to predict runout probability at a
25 case study site using parameters that were determined through calibration to a different site. Finally, we show how to
26 use a calibrated MWR model to determine runout-probability from an expert-defined, potentially unstable slope and
27 a landslide hazard map.

28 **1.   Introduction**

29 Over geologic timescales, landslides (e.g., Hungr et al., 2013) and their runout shape the topographic expression of
30 mountain ranges and channel networks (Campforts et al., 2022; Korup, 2006; Larsen and Montgomery, 2012;
31 Montgomery and Deitrich, 1988). Over more pragmatic engineering and environmental risk management timescales,
32 landslides and their runout pose a primary risk to humans and infrastructure but also support numerous ecosystem
33 benefits, including carbon and nutrient transport from hillslopes to channels and the creation of ecological habitats
34 (Benda et al., 2003; Bigelow et al., 2007; Goode et al., 2012). Therefore, explicit representation of the landslide runout
35 should be an integral component of (1) landscape evolution models, with emphasis on topographic change prediction;

36   (2) landslide risk assessment tools, with emphasis on probabilistic impact and; (3) sediment budget models, with

37   emphasis on the mobilization and deposition of sediment carried by the landslide. Our proposed runout model is

38   developed as a component of the Landlab earth surface modeling toolkit (Hobley et al., 2017) and designed to serve

39   all three purposes.

40   Landslide runout processes can be generalized into three phases: initiation, erosion and deposition, briefly discussed

41   here. After a landslide initiates, it may break apart and become a granular flow such as a debris flow or dry debris

42   slide. Field observations show that erosion rates are sensitive to the moisture content of the channel bed (McCoy et

43   al. 2012) and, although highly variable, generally increase with flow depth (Schürch et al. 2011). At the same time,

44   laboratory experiments show that entrainment rates vary with granular stress (Capart et al, 2015) and increase with

45   flow grainsize (Egashira et al., 2001). Deposition occurs by layered accretion rather than emplacement of a single,

46   massive deposit (Major, 1997) and the length of the deposit is a function of resistance to the flow, the initial landslide

47   volume and the volume of material entrained during the runout (Iverson, 1997). The grain to water ratio and friction

48   angle of the slide material control the slope at which deposition begins (Takahashi, 2014; Major and Iverson, 1999;

49   Zhou et al., 2019) but the friction angle of the material varies as a function of the grains in the flow and fluidization

50   (fluctuation of interstitial fluid and the solid particles) of the flow material (Hutter et al. 1996). If the water content of

51   the runout material is high enough, as the solid fraction of the debris flow compresses, the water is squeezed out of

52   the debris flow and may continue to transport sediment as an immature debris flow (sensu Takahashi, 2014) or intense

53   bedload (sensu Capart & Fraccarolo, 2011), which also causes the deposition angle to decrease and extent of the runout

54   to increase.

55   Existing runout models attempt to represent the above processes with a variety of different numerical approaches and

56   level of detail. Most models are parameterized from measurements of frictional characteristics of the runout material

57   and model performance is validated against historic observations (e.g., measurements of an actual landslide and

58   runout). These models can broadly be grouped into four categories: (1) site-specific-empirical/statistical models

59   (Benda and Cundy, 1990; Fannin and Wise, 2001; Miller and Bennet, 2008; Reid et al. 2016); (2) detailed, continuum-

60   based mechanistic models that predict flow physics and granular dynamics (Takahashi, 1978; Iverson and Denlinger,

61   2001); (3) top-down, simplified rules-based abstractions of some of the key physical processes into reduced-

62   complexity flow-routing models (e.g., Murray, 2003; Clerici and Perego, 2000; Guthrie et al., 2008); and (4) hybrid

63   modeling approaches that combine detailed, mechanistic approaches with empirical models and top-down, rule-based

64   abstractions (Chen et al., 2023; D'Ambrosio et al., 2003; Iovine et al., 2005; Lancaster et al., 2003; McDougall and

65   Hungr 2004; Medina et al., 2008; Frank et al, 2015; Han et al., 2017, 2021).

66   While any of the above models can be implemented in a probabilistic manner for risk assessment purposes, simply by

67   repeatedly running the models for different initial landslide volumes or parameter values (e.g., Hurlimann et al., 2008),

68   these models generally either lack a built-in probabilistic implementation or are too computationally intensive for

69   Monte Carlo simulations. In addition, only a few models are designed for use in watershed sediment budget and

70   landscape evolution applications, which involves evolving both the topography and regolith during the runout event.

71   Finally, landslide hazard or risk maps typically use a hillslope stability model to delineate regional or watershed-scale

72   landslide risk at the initiation zone (e.g., Wu et al., 1995; Lee and Park, 2015; Lee et al., 2018; Strauch et al., 2018)

73   but exclude the subsequent runout hazard. To determine runout hazard, a landslide runout model is needed.

74   Of the landslide runout models used in landscape evolution applications, they generally include the removal and

75   redistribution of hillslope material via the initial landslide and subsequent runout, but ignore erosion and entrainment

76   of new material. Deposition of eroded landslide material is often calculated by redistributing it downstream either

77   based on a depositional slope threshold, or through a calculated transport distance formulation (Carretier et al., 2016;

78   Campforts et al., 2020). Fully mechanistic models can predict runout, their computational expense and heavily

79   parameterized entrainment rules can limit their application at the scale required for regional mapping and rapid

80   response assessments (Hurlimann et al., 2008; Reid et al., 2016; Gorr et al., 2022). Recently developed hybrid, top-

81   down and empirical-statistical methods provide opportunities for mapping downslope landslide hazard (Gorr et al.,

82   2022; Guthrie & Befus, 2021; Horton et al., 2013; Liu et al., 2022, Wallace et al., 2022), but the models are still

83   heavily parameterized and may require extensive regional observations to calibrate.

84   Our proposed model is a new rule-based cellular-automaton algorithm, suitable for predicting the hazard extent,

85   sediment transport and topographic change caused by the runout of a landslide and easily coupled with a landslide

86   hazard model. This algorithm is coded in Python and implemented as a set of components for the Landlab earth

87   surface modeling toolkit (Hobley et al., 2017; Barnhart et al., 2020) and is called MassWastingRunout (MWR). MWR

88   is designed to: (1) model topographic change caused by the erosion and deposition that occurs during the landslide

89   runout and; (2) map runout impact in space probabilistically as a function of uncertainties in the initial landslide

90   volume and model parameterization. Most MWR parameter values can be inferred from field or remote sensing

91   observations but two parameters (see Section 3) may rely on model calibration. For the purposes of rapidly calibrating

92   MWR to the observed runout extent and depositional patterns of a specific site, we developed a Markov Chain Monte

93   Carlo (MCMC) calibration utility. Additionally, MWR can track attributes of the regolith such as depth, organic

94   content and sediment size, represented as a numerical value on a model grid and transferred as a conserving and well-

95   mixed mass with the runout material.

96   In this paper, we present the MWR model (Section 2), describe the calibration utility and probabilistic implementation

97   of MWR (Section 3) and demonstrate calibrated model performance at four topographically and geologically unique

98   field sites (Section 4). The field sites include six different landslides and encompass the low-, moderate- and high-

99   energy-dissipation settings (Figure 1) described by Nicoletti and Sorriso-Valo (1991). The sites also span a range of

100 landslide-runout erosion rates. After demonstrating calibrated model performance, we then test the model using the

101 parameterization of one site to predict runout probability at a neighboring site. Finally, we show how to couple MWR

102 with a probabilistically modeled landslide hazard map or a field-mapped, potentially unstable slope to determine

103 runout probability (Section 5).

104

105
106
107
108

2.   Description of the MassWastingRunout model

2.1 Overview of the cellular-automaton modeling approach

MWR is coded as a discrete cellular automaton (CA) model. CA models iteratively apply a set of equations or rules
(deterministic or probabilistic) to individual cells of a grid (e.g., Codd, 1968) to change the numerical or categorical
value of a cell state or calculate fluxes of mass. In earth sciences, CA models have been used for modeling vegetation
dynamics (e.g., Nudurupati et al., 2023) and geomorphic transport, in which gravitationally directed erosion and
depositional processes modify a digital elevation model (DEM) representation of a landscape (e.g., Chase, 1992; Crave
& Davy, 2001; Murray & Paola, 1994; Tucker et al., 2018). Other cellular-automaton based landslide runout models

109
110
111
112
113
114
115
116

117   include DebrisFlow Predictor (Guthrie and Befus, 2021), SCIDDICA (D'Ambrosio et al., 2003) and a model by Han

118   et al. (2021). In the context of predicting the runout from a landslide, the CA approach often includes only the key

119   processes that control the runout at the DEM grid scale (See Section 4) and therefore may be faster than an equivalent

120   detailed, continuum-based mechanistic model that attempts to model small-scale processes such as the granular

121   physics of the runout. In MWR, computations only occur at the location of moving debris, in a manner analogous to

122   the "mobile" cellular automaton implementation of Chase (1992).

123   Chase (1992) modeled precipitation-driven fluvial erosion by randomly placing single packets of precipitation on a

124   DEM, which then moved downslope, eroding and transporting sediment as a function of slope. The individual packets

125   of precipitation were referred to as precipitons. After one iteration, both the precipiton and the landscape evolved and

126   the location of the precipiton for the next iteration was determined by the slope of the landscape and the state of the

127   precipiton. In our implementation, we route debris released from a specified landslide source area and we refer to the

128   packets of debris as "debritons". The debritons represent debris flux. Each debriton occupies a single cell. Multiple

129   debritons can traverse the landscape at once. As they move, the debritons change volume, split to form new, smaller

130   debritons or combine to form larger debritons. In this model, debris flux is defined as the volume of debris per unit

131   area of the terrain transferred downslope in one model iteration [L3/L2/iteration]. Erosion and transport capacity of

132   the debritons are a function of slope as well as properties of the debritons.

133   The present implementation of the MWR algorithm uses the Landlab raster model grid. The raster grid consists of a

134   lattice of rectangular cells. Properties of each cell, such as topographic elevation, slope, and other spatially varying

135   attributes such as regolith depth or grain size, are recorded at nodes in the center of each cell (see Figure 5 of Hobley

136   et al. 2017) and evolve as the model runs. In the next section, we describe the methodology used to determine flow

137   direction and how each debriton interacts with the terrain and regolith as it moves down slope.

2.2 Mobilization of the initial landslide (Algorithm 1):

139   Algorithm 1 determines the initial direction and flux of material over the landslide slip-surface (surface along which

140   the landslide body initially fails/moves). It operates on all nodes within the body, as represented by a user-specified

141   polygon and serves as the first iteration of MWR. All later iterations are implemented following Algorithm 2

142   (described below).

143   For each node n within the landslide polygon (Figure 2, iteration 0), Algorithm 1 determines which nodes will receive

144   debris (receiver nodes: R_n), the flux of debris sent to each of the node n's i-th receiver nodes (qs_{n,R_n}) and attributes

145   of the sent debris (ξ_n) based on topographic slope and attributes at node n (Figure 3). Movement is initiated at the

146   lowest debriton in the landslide body, where topographic slope is computed from the upper surface of the debriton. At

147   all other debritons, and for all other stages of modeled runout (Algorithm 2), topographic slope is determined from

148   the terrain underlying the debriton (See Figure 3a and 3b).

149   The depth of the slip surface is specified by the user, which can be inferred from field evidence at a specific landslide

150   (e.g., shallow failure along the fresh-bedrock to regolith contact) or determined from an external landslide model (See

151   Section 5). All nodes that send material are designated as donor nodes and appended to a list of donor nodes (D). All

152   R_n are appended to receiver node list R. Flux and attributes that will be passed from the donor nodes to the receiver

153 nodes are appended to lists qs and ξ. Lists D, R, qs and ξ are the input to Algorithm 2 and change with each iteration

154 of Algorithm 2.

155

Algorithm 1 Determinin the initial D, R, qs and ξ lists

sort initial landslide nodes from lowest elevation to highest
for each node n in the initial landslide nodes, beginning from the lowest node
    determine R_n, qs_{n,R_n} and ξ_n using (1)
    subtract landslide thickness from topographic elevation
    update the topographic slope using new topographic elevation
    append n, R_n, qs_{n,R_n} and ξ_n to lists D, R, qs and ξ
end for

156

157
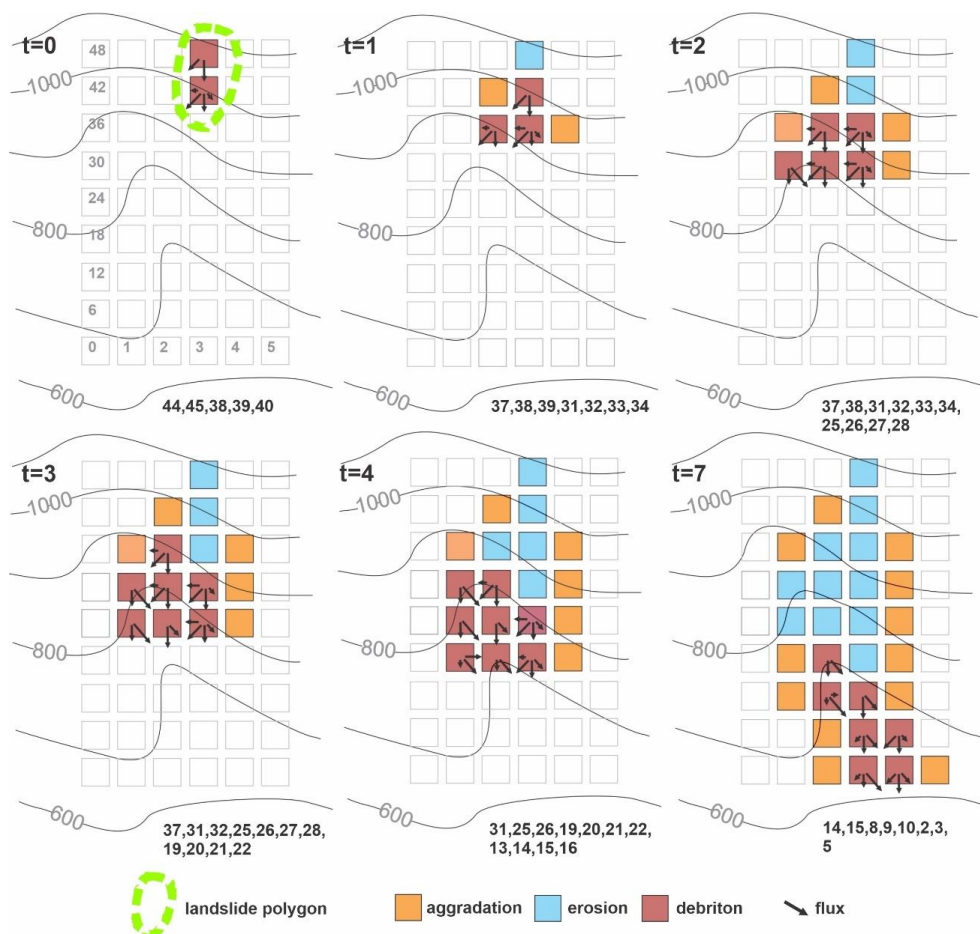
Figure 2. Illustration of initial mass wasting release (t = 0) and runout. Notice how the receiver node list $\boldsymbol{R}$, shown at the base of each iteration, changes with each iteration. The flow elongates and widens as the number of nodes in $\boldsymbol{R}$ increase and stops when the number of nodes in $\boldsymbol{R}$ is zero. If the incoming flux (sum of all incoming debritons) to a node is less than $\boldsymbol{qs_c}$, the material stops, causing aggradation.
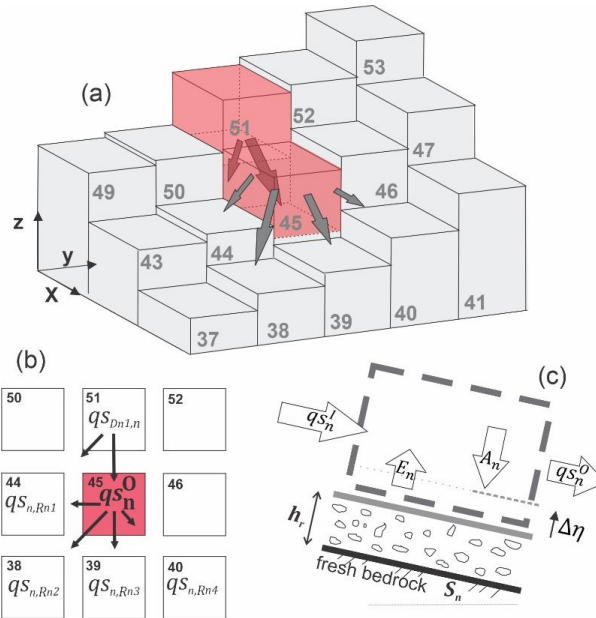
Figure 3. (a) Three-dimensional illustration of iteration 0 in Figure 1, showing initial landslide nodes (initial debritons) and flux towards downslope nodes. Except for the first released debriton (lowest elevation node in the landslide), all debritons are directed downslope based on the underlying topographic slope (compare flow directions of node 51 to node 45); (b) Distribution of qs_n^O to downslope nodes and incoming material qs_n^I (here equal to $qs_{D_{n1},n}$); (c) Aggradation (A_n) and erosion (E_n) are determined from field informed rules and qs_n^O is determined from mass continuity.

2.3 Rules for debris flow erosion and deposition (Algorithm 2)

Algorithm 2 is essentially the runout model. Algorithm 2 is repeatedly implemented while the length of R is greater than zero (while there are receiver nodes for the next model iteration). An illustration of several iterations of Algorithm 2 is shown in Figure 2 and in the Validation section (Section 4, Figure 7). Algorithm 2 determines how material traverses and interacts with the terrain, beginning with the initial debritons, described by D, R, qs and ξ lists passed from Algorithm 1. While Algorithm 2 processes each debriton, it appends n (node index of the debriton), R_n, $qs_{n,R_{ni}}$ and ξ_n to temporary lists D', R', qs' and ξ'. At the end of each implementation of Algorithm 2, D', R', qs' and ξ' become the updated D, R, qs and ξ lists.

In addition to slope, debriton movement is controlled by mass continuity and a simple threshold flux constraint. At a node n, outgoing sediment flux, qs_n^O, is determined as (Figure 3c):

$$qs_n^O = \begin{cases} qs_n^I - A_n + E_n, & qs_n^I \geq qs_c \\ 0 & , & qs_n^I < qs_c \end{cases} \tag{1}$$

Where, qs_n^I is total incoming sediment flux [L/iteration] to the node (described in Equation 2); A_n and E_n are local deposition (aggradation) and erosion [L/iteration], both of which are calculated prior to (1); qs_c is the threshold flux for deposition. When $qs_n^I < qs_c$, qs_n^I deposits and qs_n^O becomes zero. Conceptually, qs_c represents the flow depth

184   below which flow resistance is large enough to cease the forward momentum of the flow, whether in the form of

185   frictional resistance along the base of the flow or debris and vegetation in the path of the flow.

186   Summing the outgoing flux from each donor node towards node n, gives the total incoming flux, qs_n^I:

187     $$qs_n^I = \sum_{j=1}^{Nd_n} qs_{D_{nj},n} \qquad (2)$$

188   where Nd_n is the number of donors to node n, and $qs_{D_{nj},n}$ is the flux from node D_{nj} (the j-th donor to n) to node n.

189   The flux from node n to each of the node n's i-th receiver nodes (qs_{n,Rn_i}) is determined using the Freeman (1991)

190   multiflow direction algorithm:

191     $$qs_{n,R_{ni}} = qs_n^O \frac{S_i^a}{\sum_{i=1}^{Nr_n} S_i^a} \qquad (3)$$

192   where Nr_n is the number of receiving nodes of node n, and S_i is the underlying topographic slope ($\tan\theta$) to each of

193   node n's i -th receiver and qso_n.

194   The multiflow direction algorithm is common to many models, and in this implementation it is handled by a pre-

195   existing Landlab flow-routing component. The exponent a controls how material is distributed to downslope nodes

196   as a function of slope. In a braided river cellular-automaton model, Murray and Paola (1994) used an approximation

197   for turbulent shallow water flow to justify $a = 0.5$ (which is the exponent on the slope factor in channel friction laws).

198   For our application, we found MWR was easier to calibrate to observed mass wasting runout if $a = 1$.

199   Local aggradation (A_n) and erosion (E_n) in (1) are determined using a series of field-data-informed rules (Figure 3c).

200   These rules are equivalent to the "transition functions" applied in SCIDDICA (D'Ambrosio et al. 2003) and the Han

201   et al. (2021) models. If the local slope (S_n) is greater than a critical slope (S_c), then no material deposits, whereas if

202   S_n is less than S_c, then a portion or all of qs_n^I deposits:

203     $$A_n = \begin{cases} 0 & , & S_n \geq S_c \Delta x \\ min\left(A_{p,n|N_a}, qs_n^I\right), & S_n < S_c \Delta x \end{cases} \qquad (4)$$

204   Where, $A_{p,n|N_a}$ is the potential deposition depth at node n, which is the first node where the debriton begins to deposit

205   because the deposition criterion is met (described below, see Figure 4). To determine $A_{p,n|N_a}$, we initially used the

206   nonlinear, nonlocal deposition scheme described by Carretier et al. (2016) and Campforts et al. (2020); however, when

207   implemented with the flow routing method described above, unreasonably tall deposits resulted when qs_n^I was large

208   and $S_n << S_c$. Alternatively, we experimented with limiting the deposition depth to $A_n \leq S_c \Delta x$, but when qs_n^I was

209   large, this constraint resulted in long deposits that paralleled the underlying slope. Therefore, a different constraint

210   was needed. We ultimately developed a new deposition rule that determines deposition at a node by back-computing

211   the deposition amount necessary to form a deposit that spreads over multiple nodes whose surface slope is controlled

212   by S_c, as described below.

213   When the slope criteria for deposition is met at a node n, sediment that enters the node is assumed to be

214   deposited/distributed over a single row of downstream nodes following three rules: (1) the volume of the deposit will

215   be equal to $qs_n^I \Delta x^2$ and spreads over N_a nodes; (2) the surface slope of the deposit will equal the critical slope S_c and;

216   (3) the underlying topographic slope of the N_a nodes is uniform and equal to S_n. Assumption 3 is necessary because

217 the terrain evolves with each model iteration and the future downslope terrain is unknown. From this hypothetical

218 deposit, we can analytically define $A_{p,n|N_a}$ and N_a as a function of qs_n^I, S_c and S_n as follows (see Figure 4):

219 From assumption 1, qs_n^I can be expressed as the sum the N_a deposits that make up the overall deposit as:

220 $$qs_n^I = \sum_{i=1}^{N_a} A_{p,i} \tag{5}$$

221 where $A_{p,i}$ is the i-th deposition amount in the deposit. Since we assume the deposit slope and underlying topographic

222 slope are uniform, the deposition amount at any of the N_a nodes can be determined from the last deposition amount

223 ($A_{p,1}$) as:

224 $$A_{p,i} = A_{p,1} + (i-1)\Delta x (S_c - S_n) \tag{6}$$

225 From (6) we can re-write (5) as a function of $A_{p,1}$ (write each of N_a deposits in (6) in terms of $A_{p,1}$) and rearrange to

226 define $A_{p,1}$ as a function of qsi_n:

227 $$A_{p,1} = \frac{1}{N_a} qs_n^I - \frac{N_a - 1}{2} \Delta x (S_c - S_n) \tag{7}$$

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

229 $$A_{p,n|N_a} = \frac{1}{N_a} qs_n^I + \frac{N_a - 1}{2} \Delta x (S_c - S_n) \tag{8}$$

230 Equation (7) can be rearranged into a quadratic equation and solved for N_a as:

231 $$N_a = \frac{-A_{p,1} + \frac{1}{2}\Delta x(S_c - S_n) \pm \sqrt{\left(A_{p,1} - \frac{1}{2}\Delta x(S_c - S_n)\right)^2 + 2\Delta x(S_c - S_n)qs_n^I}}{\Delta x(S_c - S_n)} \tag{9}$$

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

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

234 in Figure 4 (i.e., a single row of nodes), the debriton deposits over N_a nodes at a uniform slope equal to S_c. When

235 implemented on an actual terrain, the interaction between multiple debritons in multiple directions creates a complex

236 deposit whose slope changes with S_c.

237 The third variable in Equation 1 is debriton erosion rate, E_n [L/iteration]. We constrain E_n to the lesser of a potential

238 erosion rate, h_e, and local regolith depth, h_r:

239 $$E_n = min\,(h_r, h_e) \tag{10}$$

240 where h_e is computed as a function of the basal shear stress of the granular flow (τ, described below in Equations 12

241 and 13) and the critical shear stress (τ_c) of the regolith at the node:

242 $$h_e = k(\tau - \tau_c)^m \tag{11}$$

243 Shen et al. (2020) used $\tau_c = 5$ Pa to represent a boulder-colluvium substrate. Stock and Deitrich (2006) used normal

244 rather than shear stresses and excluded τ_c ($\tau_c = 0$). The exponent m controls the non-linearity of h_e. Many authors

245 (Chen & Zhang 2015; Frank et al., 2015; Shen et al., 2020) use a value of 1 for m but field measurements by Schurch

246 et al. (2011) (see their Figure 3) suggest that m may be less than 1 if τ is assumed to vary linearly with flow depth,

247 particularly at high flow depths ($h > 3$m). We experimented with various values of m, ranging from 1 to 0.2 and

248 found that if τ is assumed to vary linearly with flow depth (see Equation 12, below), model performance improves

249 with lower values of m and if τ is assumed to vary as a function of grain-collision dependent shear stress (see Equation

250 13, below), as long as k is determined as a function of m (Section 4.3, Equation 36), the impact of the value of m on

251 model behavior is relatively minor.

252 The coefficient k in (11) is an erodibility parameter. Stock and Dietrich (2006) showed that k encapsulates rock

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

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

255 we are modeling the erosion associated with a single runout event, as represented by the debritons. The coefficient k

256 therefore needs to scale h_e so that it is on the order of the average erosion depth caused by a single debriton.

257 As noted above, to define τ, MWR includes two options: (1) a quasi-static basal shear stress approximation or (2) a

258 grain-size-based shear stress approximation. The quasi-static basal shear stress approximation (e.g., Takahashi, 2014)

259 is defined as:

260 $$\tau = \rho g h \sin \theta \qquad (12)$$

261 where ρ is the density of mass wasting material (grain and water mixture), g is gravity and h is flow depth, which in

262 MWR is equivalent to qs_n^I. The grain-size-based shear stress approximation is defined as follows:

263 Bagnold (1954) defined τ as a function of grain-collision dependent normal stress (σ).

264 $$\tau = \sigma \tan \varphi \qquad (13)$$

265 where φ is the collision angle between grains, measured from the vertical axis (See Bagnold, 1954), with a value of

266 $\tan \varphi$ typically equal to 0.32. Following Bagnold (1954) and Iverson (1997), Stock and Dietrich (2006) defined σ as:

267 $$\sigma = \cos \theta v_s \rho_s D_s^2 \left(\frac{du}{dz}\right)^2 \qquad (14)$$

268 where θ is the topographic slope measured in degrees, v_s is the volumetric solids concentration, ρ_s is density of the

269 solids, u is flow velocity, z is depth below the flow surface, du/dz is strain and D_s is the representative grain size.

270 Stock and Dietrich (2006) suggested that D_s corresponds to a small percentile of the coarsest fraction of the runout

271 material (D_{88} to D_{96}) and they approximated du/dz as:

272 $$\frac{du}{dz} = \frac{u}{h} \qquad (15)$$

273 Solely for the purpose of computing du/dz, we approximate velocity at a node following Julien and Paris (2010) as:

274 $$u = 5.75 u^* log\left(\frac{h}{D_s}\right) \qquad (16)$$

275 where $u^*$is shear velocity $\left(\sqrt{gh\tan\theta}\right)$. This grain-size dependent approximation for τ is advantageous because it

276 permits landslide-driven erosion rates to vary by lithologic region, as represented by spatial variation in D_s generated

277   by each lithologic region (e.g., Roda-Boluda et al. 2018), Furthermore, D_s can be quickly approximated in the field

278   by measuring the coarser grain sizes in existing runout-deposits, road-cuts and tree-throw pits.

279   Once A_n and E_n are determined at a node, we assume uniform material density and determine the change in elevation

280   at the node ($\Delta\eta_n$) as:

281    $\Delta\eta_n = A_n - E_n$ (17)

282   Finally, for each regolith attribute being tracked by the model (e.g., grain size), the attribute value at the node (ξ_n)

283   and the attribute value sent from the node to its receiver nodes (ξ_{Rn}) is determined using a volumetric-weighted

284   average as:

285    $\xi_n = \frac{\xi_n(h_r - E_n) + \xi_{Dn}(A_n)}{D + h_r - E_n}$ (18)

286    $\xi_{Rn} = \frac{\xi_E E_n + \xi_{Dn}(qsi_n - A_n)}{qso_n}$ (19)

287   Where ξ_{Dn} is the weighted average attribute value delivered to the node from its donor nodes:

288    $\xi_{Dn} = \frac{\xi_D \cdot qs_{D_n}}{qsi_n}$ (20)

289   and qs_{D_n} is a vector containing all $qs_{D_{nj},n}$ sent to the node and ξ_i is a vector containing the incoming attribute values

290   for each $qs_{D_{nj},n}$.

291   For each model iteration t, as Algorithm 2 processes each unique node n in the receiver node list (R), it appends $\Delta\eta_n$

292   to a list $\Delta\eta$ and n to D', R_n to R', $qs_{n,Rn}$ to qs' and ξ_{Rn} to ξ'. After Algorithm 2 has cycled through all receiver nodes,

293   D', R', qs' and ξ' become D, R, qs and ξ and maps of regolith thickness ($h_{r\,t}$), topographic elevation (η_t) and the

294   elevation of the runout flow surface (F_t) are updated for the entire model domain using vector operations as:

295    $\eta_t = \eta_{t-1} + \Delta\eta$ (21)

296    $h_{r\,t} = h_{r\,t-1} + \Delta\eta$ (22)

297    $F_t = \eta_t + qs$ (23)

298   After η_t, $h_{r\,t}$ and F_t are updated, the multi-direction slope of the DEM, which is used for routing the debritons the

299   next model iteration, is recomputed from η_t.

300   Note, the debritons are obstructed if they encounter a topographic pit or flat topography in the DEM. To allow a

301   debriton to pass an obstruction, we rely on a simple work-around: upon encountering the obstruction, the debriton is

302   directed to itself ($R_n = n$) and some portion of the debris is deposited based on (4). At the end of the model iteration,

303   the node elevation and slope are updated, and then during the next iteration, if the remaining mobile debris is no longer

304   obstructed, it moves to its downslope node(s). If the node is still obstructed, it is again sent to itself until either all

305   material has deposited or the elevation of the node exceeds that of its neighbor nodes, allowing the debriton to move

306   downslope.

307   Finally, as noted in Tucker and Hancock (2010), flow routing algorithms used in CA models, such as (3), can cause

308   flow-depth dependence on the grid-size used to represent the terrain. Since aggradation and erosion rates determined

309   by (8) and (11) are sensitive to flow depth, we thus constrain the flow depth used to determine A_n and E_n to no more

310   than the maximum observed flow depth as:

311     $h' = min\ (h_{max}, h)$ (24)

312   Where h' is the flow depth used in (8) and (11) and h_{max} is the maximum observed flow depth, which can be inferred

313   from field indicators (See Section 4). This constraint is not applied to mass continuity in (1).

314

Algorithm 2 Runout algorithm

while length(R) >0
   for node n in R
      determine flux and attributes of incoming material from lists qs and ξ
      compute erosion, deposition and outflow as a function of underlying topographic slope
      determine the new receiver nodes (R_n), receiving flux ($qs_{n,Rn}$), receiving properites (ξ_{Rn})
      append n to D', R_n to R', $qs_{n,Rn}$ to qs and ξ_{Rn} to ξ'.
   end for
   update the topographic elevation and slope for erosion and deposition
   lists D', R', qs' and ξ' become D, R, qs and ξ
end while

315

316
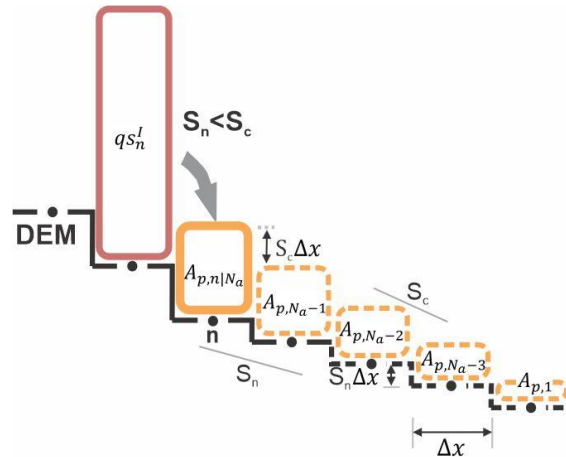
Figure 4. Conceptual model used to define deposition rule, here illustrated using $N_a = 5$. The rule relies on three
assumptions about the depositional process: (1) deposition of the debriton, which has a flux of qs_n^I at node n,
spreads over N_a nodes to form a hypothetical deposit as illustrated by yellow boxes above; (2) the underlying
topographic slope of the N_a nodes is uniform and equal to S_n and; (3) the surface slope of the hypothetical deposit
will be equal to S_c. 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.

3.   Calibration and MWR probability

3.1   Calibration utility

We developed a calibration program, written in Python as a utility for the package Landlab, to calibrate MWR. This utility uses a modified version of an adaptive Markov Chain Monte Carlo (MCMC) sampling algorithm from Coz et al. (2014), who used it to calibrate stage-discharge relationships to observed river flow in France. The MCMC algorithm automatically samples model parameter space, evaluates model performance relative to observed runout, and determines the most likely set of parameter values that calibrate the model.

For each iteration of the calibration utility, a likelihood index, $L(\theta)$, for the given set of model parameters (θ) is determined as the product of the probability of the selected parameter values, based on a prior estimate of the probability density function (PDF) of each parameter ($p(\theta)$), and three other metrics as:

$$L(\theta) = p(\theta) * \Omega_T * \frac{1}{\Delta\eta_E{}^2} * \frac{1}{Q_{s_E}{}^2} \qquad (25)$$

where Ω_T is the Lee-Salle index (Heiser et al., 2017) and indicates modeled planimetric fit to the observed debris flow extent, and $\Delta\eta_E{}^2$ and $Q_{s_E}{}^2$ are new dimensionless indices, proposed for this study, that represent modelled distributed topographic change and cumulative flow errors, respectively. Larger values of Ω_T indicate that the modeled planimetric runout extent more closely fits observed. Smaller values of $\Delta\eta_E{}^2$ and $Q_{s_E}{}^2$ indicate better fit. Note that the reciprocal values of $\Delta\eta_E{}^2$ and $Q_{s_E}{}^2$ are used in (25) so that improved fit causes $L(\theta)$ to increase. To ensure that all values of $L(\theta)$ are greater than or equal to zero, here we add a value of 1 to Ω_T so that it scales from 0 to 2 as:

$$\Omega_T = \frac{\alpha - \beta - \gamma}{\alpha + \beta + \gamma} + 1 \qquad (26)$$

where α, β and γ are the areas of matching, overestimated and underestimated runout extent, respectively. The root values of $\Delta\eta_E{}^2$ and $Q_{s_E}{}^2$ ($\Delta\eta_E$ and Q_{s_E}) reflect the volumetric error of the modeled topographic change relative to the observed total mobilized volume (observed landslide + entrained material) and the mean-modelled-cumulative flow error along the runout path relative to the observed mean cumulative flow. The index $\Delta\eta_E{}^2$ is determined by dividing the squared, cumulative, volumetric square error (SE) of modeled topographic change by the squared-observed-total-mobilized volume (V):

$$\Delta\eta_E{}^2 = \frac{SE_\alpha + SE_\beta + SE_\gamma}{V^2}. \qquad (27)$$

Here values SE_α, SE_β and SE_γ are the SE of modeled $\Delta\eta$ in areas α, β and γ respectively, individually calculated as:

$$SE = \sum_{i=0}^{p}[(\Delta\eta_{Oi} - \Delta\eta_{Mi})\Delta x^2]^2 \qquad (28)$$

where, p is the number of nodes within each of the α, β, and γ areas, and ($\Delta\eta_{Mi}$) and $\Delta\eta_{Oi}$ are the observed topographic change [L] at the i-th node (cell) in that erosion area.

Cumulative flow volume (Q_s) is determined along the channel profile, as the integrated, volumetric topographic change upslope of a point along the profile in a manner similar to Fannin and Wise (2001), who computed cumulative

13

354   debris flow volume at individual reaches along a channel profile (see their Figure 4). We compute cumulative flow at

355   each node (grid cell) j along the channel profile as:

356    $Q_s = -\Delta x^2 \sum_{i=1}^{q_j} \Delta \eta_{i,j}$ (29)

357   where $\Delta \eta_{ij}$ is the topographic change [L] at the i-th node located upstream of node j, q_j is the number of all nodes in

358   the DEM located upstream of node j, and node j is a node located along the channel profile. The negative sign converts

359   upslope erosion and deposition to a positive and negative change in cumulative flow respectively. For each node j

360   along the profile, Q_s is computed for both the observed and modeled runout ($Q_{sO\ j}$ and $Q_{sM\ j}$) and $Q_{s_E}^2$ is the mean

361   square error of the modeled-cumulative flow at all nodes along the profile, normalized by the observed-mean-squared-

362   cumulative flow along the profile:

363    $Q_{s_E}^2 = \dfrac{\frac{1}{r}\sum_{j=1}^{r}\left(Q_{sO\ j}-Q_{sM\ j}\right)^2}{\frac{1}{r}\sum_{i=0}^{r}Q_{sO\ j}^2}$ (30)

364   Where r is the number of nodes in the channel profile. As noted above, values of $Q_{s_E}^2$ and $\Delta \eta_E^2$ (or Q_{s_E} and $\Delta \eta_E$)

365   close to zero indicate a good fit and large values (>>0) indicate a poor fit and in the final likelihood index, $L(\theta)$, we

366   use the reciprocals of $\Delta \eta_E^2$ and $Q_{s_E}^2$, so that smaller values of $Q_{s_E}^2$ and $\Delta \eta_E^2$ result in larger $L(\theta)$ values.

367   A minimalistic calibration of MWR requires a shape file, representing the boundaries of a landslide and its observed

368   runout. In this application, (25) is determined as $L(\theta) = p(\theta) * \Omega_T$. A thorough calibration of MWR requires a DEM

369   of difference (DoD), which is a map of the change in topographic elevation ($\Delta \eta$) caused by the observed runout, and

370   is most easily obtained by subtracting the DEMs recorded before and after the landslide and subsequent runout but

371   can also be created from a field survey (see Section 4). In case that a minimalistic calibration is used, because the

372   likelihood index is limited to an evaluation of planimetric extent, the model tends to have high parameter equifinality

373   (e.g., Beven 2006); multiple parameter sets result in an equally calibrated model as evaluated by Ω_T. When an

374   observed DoD is used to calibrate MWR, a much narrower range of equally likely parameter values can be identified.

375   The calibration utility is presently coded to modify S_c and qs_c, k, landslide thickness, and mean soil depth; however,

376   any parameter of the MWR can potentially be calibrated using this approach.

377   As will be detailed in Section 4, most parameters in MWR can be inferred from field observations and a DoD map,

378   but parameters S_c and qs_c are best determined through calibration. A range of S_c values can be inferred from

379   deposition slopes measured at channel-filling deposits (reaches where deposition exceeds erosion and results in

380   positive net deposition; Guthrie et al., 2010) of the observed runout but given the intermittency and variability of most

381   debris flow deposits, identifying a single representative S_c value in the field can be difficult. As noted in Section 2,

382   conceptually, qs_c represents a minimum flow thickness below which resistance due to friction or debris, such as

383   vegetation or boulders in the path for the flow, causes the flow to cease. While it's possible that qs_c may be inferred

384   from the minimum observed flow thicknesses along the margins and termini of the runout, this evidence, especially

385   at the termini of debris flows, can be erased rapidly as streams erode deposits and may also vary along the runout path.

386   It is through model calibration to the DoD of the observed runout, that a single value of S_c and qs_c that best represents

387   the heterogeneity of observed depositional slope and flow thickness can be determined.

388   The MCMC algorithm is implemented as follows (see Algorithm 3): First, the user defines an initial prior PDF for

389   each parameter to be calibrated, currently only limited to normal and uniform distributions. The algorithm then

390   randomly selects a set of parameter values (θ) from the prior PDFs, evaluates the prior likelihood of the parameter set

391   ($p(\theta)$), and runs MWR using θ. Once the model has completed the run, the algorithm evaluates the posterior

392   likelihood of the parameter set ($L(\theta)$, based on how well modeled runout matches the planimetric extent, topographic

393   change and cumulative flow along the observed runout path.

394   After the first $L(\theta)$ has been determined, the algorithm selects a new set of parameters (θ_{t+1}) by jumping some

395   distance from each parameter in θ space. The jump distance is determined by randomly sampling a normal distribution

396   of possible jump distances, whose standard deviation is initially specified by the user, but changes as the algorithm

397   iteratively runs and evaluates model performance. The next iteration, $t + 1$, MWR is run using θ_{t+1} and once the

398   model has completed the run, L_{t+1} is determined. If L_{t+1} is larger than L, θ is updated to the value of θ_{t+1} and L is

399   updated to the value of L_{t+1}. If $L_{t+1} < L$, then θ and L have a probability of remaining unchanged equal to 1- L_{t+1}/L

400   where the fraction L_{i+1}/L is referred to as the acceptance ratio. This Markov process is repeated a user-specified N

401   times. The algorithm is adaptive because the jump distance changes depending on the average acceptance ratio, which

402   is evaluated every user-specified number of iterations ($Ncycles$). Two additional parameters control the jump size:

403   a_{min} and a_{max}. If the average acceptance ratio is below a_{min}, the variance of the jump-distance distribution is

404   decreased by 10 percent and if the average acceptance ratio is greater than a_{max}, the variance is increased by 10

405   percent. We initially tested the adaptive MCMC algorithm using values for $Ncycles$, a_{max}, a_{min} and the initial jump

406   size ($jump\_size$) from Coz et al. (2014) but through a series of tests, found that values listed in Algorithm 3 worked

407   best for our application.

408   An advantage of defining $L(\theta)$ as the product rather than summation of the runout extent, topographic change and

409   cumulative flow errors metrics is that a change in any one metric has an equal impact on the relative value of $L(\theta)$

410   (e.g, a 20% reduction of any of the metrics causes a 20% reduction in $L(\theta)$). Furthermore, because we use the product

411   of the reciprocals of the squared topographic and flow error (e.g., error-based weighting, Foglia et al., 2009), this

412   formulation of $L(\theta)$ is very sensitive to changes in calibration performance and results in values that vary over several

413   orders of magnitude. This sensitivity to calibration performance works well for guiding the MCMC algorithm as it

414   searches parameter space based on changes in the relative value of $L(\theta)$.

415   Additionally, one of the major advantages of using an MCMC approach for model calibration is that after an adequate

416   number of model runs have been performed, the resulting histogram of parameter indicates the posterior PDF of the

417   parameter values (Gelman et al., 2021; Renard et al., 2006), which can then be used in Monte Carlo modeling runs.

418   Typically the initial series of jumps are considered "warm up" and depending on context, may be discarded from the

419   histogram (Gelman et al., 2021). Also, the minimum number of iterations needed for convergence on best-fit

420   parameters (maximum $L(\theta)$) can be determined as a function of the ratio of between sequence variance (e.g., variance

421   between two sequences of $Ncycles$ parameters) and variance within each sequence.

422   In the Model Validation section, we run the calibration utility using a single chain of 2000 repetitions to estimate the

423   posterior PDF for each parameter. At most sites, the model converges relatively quickly on a solution and we therefore

424   don't consider burn-in, or evaluate convergence. Future implementations of the calibration utility may include multiple

425 chains, burn-in and a check for convergence. We use the parameter set with the highest $L(\theta)$ value (best-fit parameter

426 set) to compare model performance relative to observed runout. Then, in Section 5, we iteratively run MWR with

427 parameter values sampled from the posterior PDFs to illustrate the calculation of runout probability as a function of

428 parameter uncertainty.

429

Algorithm 3 Adaptive Markov Chain Monte Carlo calibration algorithm

N : user defined number of times the while loop will run
θ_{t-1} : initial set of model parameter from which the first jump in parameter space is made
$jump\_size$: ratio of standard deviation of the jump size expressed as the ratio of the range between the
         minimum and maximum parameter values
$L(\theta_{t-1}) = 1$
$t = 0$
while t $< N$
     from θ_{t-1}, jump random distance to new parameter set, (θ_t)
     compute liklihood of parameter values using prior distribution $(p(\theta_t))$
     run MWR with with θ_t
     evaluate model calibration with $\Delta\eta_E{}^2$, $Q_{s_E}{}^2$ and Ω_T
     compute posterior likihood of parameter values ($L(\theta_t)$) using (25)
     acceptance_ratio = min(1, $L(\theta_t)$/ $L(\theta_{t-1})$)
     append acceptance_ratio to acceptance_ratio_list
     rv = random value between 0 and 1

     if rv < acceptance_ratio
         $\theta_{t-1} = \theta_t$

     if remainder of t/$Ncycles = 0$
         if mean acceptance_ratio > a_{max}
           increase variance of jump size distribution by 10%
         if mean accepatnce_ratio < a_{min}
           decrease variance of jump size distribution by 10%

     t = t+1
end while

Where $Ncycles = 100$, $a_{max} = 0.7$, $a_{min} = 0.3$, $jump\_size = 0.09$

430

431 **3.2   Mapping landslide runout risk**

432 MWR includes an additional program called MWR Probability, also written in Python as a utility for the package

433 Landlab, that relies on a Monte Carlo implementation of MWR using the parameter PDF's produced by the calibration

434 utility to produce landslide runout hazard maps. MWR Probability repeatedly runs MWR for a user specified N times,

435 each repetition with a different S_c, qs_c pair, randomly sampled from the parameter PDFs. After the N model runs, it

436 determines probability of runout impact at each model node as:

437 $$P(\Delta\eta) = \frac{n(|\Delta\eta| > 0)}{N} \tag{31}$$

438 where $n(|\Delta\eta| > 0)$ is the number of times topographic elevation at the node changed as a result of erosion or

439 deposition from the N model runs.

440 MWR Probability includes three options for specifying the initial landslide that is be used to iteratively run MWR *N*

441 times: (1) a single, fixed-size landslide, represented as a polygon with a fixed regolith depth, (2) a single, landslide

442 whose size varies each iteration within a user defined potentially unstable slope with fixed regolith depth (3) multiple

443 landslides that change location, area and depth with each iteration. If option 2 is used, for each model iteration, MWR

444 Probability randomly selects MWR parameter values from the parameter PDFs and a landslide area within the

445 potentially unstable slope polygon. The randomly selected landslide area is at least as large as a user defined minimum

446 size but no larger than the potentially unstable slope and can form anywhere within the potentially unstable slope. If

447 option (3) is used, for each model repetition, MWR reads externally determined regolith depth and landslide locations

448 from a total of *N* regolith depth and landslide maps.

449 **4.   Model Validation:**

450 **4.1 Overview**

451 In this section, we demonstrate calibrated model performance relative to observed runout at field sites located in the

452 western USA. Note that simply calibrating a model to match field data does not constitute a satisfactory test of model

453 predictive ability (Iverson, 2003). Strategic testing, which involves calibrating the model to one site or period of time

454 and then running the calibrated model at a separate site or period of time (Murray 2013), is a better indicator. Two of

455 our validation sites, the Cascade Mountain and Olympic Mountain sites, include two separate landslides and

456 subsequent runout and we test model predictive ability at these sites in Section 5.

457 Calibrated model performance is demonstrated at the following field sites (see Figure 5a for locations): (1) two runout

458 events in the Cascade Mountains (Washington state [WA], USA): a large debris avalanche in 2009 and a moderately-

459 sized debris flow in 2022 that inundated and flowed within a first-order channel until perpendicularly intersecting a

460 narrow river valley several hundred meters below the landslide (Figure 1a); (2) channelized debris flows sourced from

461 a small failure along the toe of a deep-seated landslide and a moderately-sized debris avalanche from a large road fill

462 in the Black Hills (WA) that flowed several kilometers along a flume-like channel before stopping (Figure 1b); (3) a

463 single, moderately-sized debris avalanche in the Rocky Mountains (Colorado state), the majority of which flowed

464 several hundred meters over a broadly convergent to divergent hillslope (Figure 1c); and (4) a 30-year chronology of

465 small landslides and subsequent debris flows in steep, highly convergent channels in the Olympic Mountains (WA)

466 that flowed well over a kilometer and coalesced into a single runout deposit in a dendritic, channelized watershed

467 (Figure 1d).

468 These validation sites encapsulate the three runout morphologies categorized in terms of runout-energy-dissipation by

469 Nicoletti and Sorriso-Valvo (1991): channelized runout (low-energy-dissipation; Figure 1d), unobstructed spreading

470 of the runout (moderate-energy-dissipation; Figure 1c) and runout that perpendicularly intersects a channel and

471 collides with the opposite channel wall (high-energy-dissipation; Figure 1a). The validation sites also span different

472 climates and land use. The Cascade Mountain, Black Hills and Olympic Mountain sites are located in a wet, maritime

473 climate, in landscapes actively managed as tree plantations. The Rocky Mountain site is located in a semi-arid climate,

474   in a landscape managed as a national park. Despite differences in runout morphology, land use and climate, all

475   landslides initiated during heavy rainfall or rain-plus-snowmelt storm events (WRCC, 2022; NRCS, 2022; Table 1).

476   In the next sub-sections, we summarize the remote sensing and field methods used to parameterize MWR and the

477   calibration utility. Then, for each site, we describe observed runout characteristics, modeled parameterization and

478   model performance. At all locations, we use (13) to approximate shear stress as a function of grain size. Model

479   performance is discussed in terms of the planimetric, topographic and volumetric flow metrics in $L(\theta)$ and compared

480   to topographic characteristics of the runout path.

481

482   **Table 1.** Landslide and runout characteristics

site	Cascade Mountain, 09	Cascade Mountain, 22	Black Hills, south	Black Hills, north	Rocky Mountain	Olympic Mountain
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 at positive-net deposition [%]	1 - 15	1 - 15	<1 - 10	<1 - 8	16 - 25	5 - 15
average flow depth in scour zone [m] *	4	2	2	3	3	3
average channel slope in scour zone [m/m]	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
total erosion volume, $\sum E \Delta x^2$ [m^3]**	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 $|\Delta\eta|$	2.4	2.2	0.53	0.63	0.89	1.4
total erosion volume / total mobilized volume***	0.29	0.19	0.89	0.59	0.88	0.97

483   * rough approximation based on landslide volume, channel width and height of scour marks in the scour zone of the
484   channel
485   ** excludes landslide volume
486   ***total moblized volume = total erosion volume+ landslide volume
487
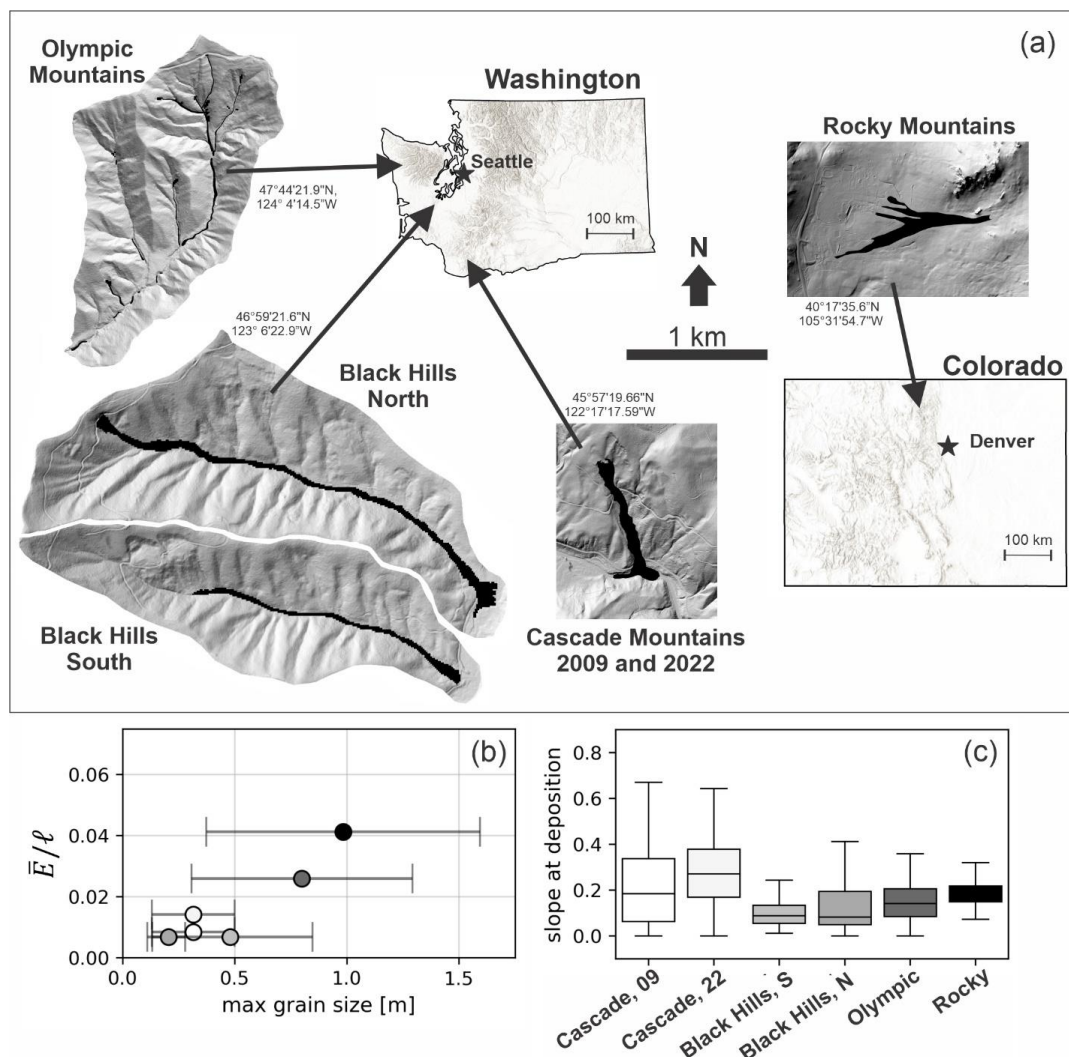
Figure 5 (**a**) Landslide locations in Washington and Colorado states. Coordinates next to each site are WGS84. Shaded DEMs of each site are sized to the same scale. (**b**) Observed average erosion rate per unit landslide length ($\bar{E}/\ell$) relative to the observed average-maximum grainsize. Error bars indicate standard deviation. (**c**) Underlying topographic slope of observed deposition locations.

4.2 Model setup, data acquisition and field reconnaissance

For each validation site, we used a combination of field and remote sensing observations to set up and parameterize MWR. Each model was set up on a 10-meter grid, with the following gridded inputs: (1) the pre-event DEM; (2) a map of regolith depth and (3) landslide boundary polygon, represented by all nodes within the polygon. If the landslide

498 slip plane depth differed from the regolith depth, a list of the slip plane depths at all nodes in the landslide polygon

499 were used to define the slip plane. For implementing the calibration utility, we prepared a DoD of each site.

500 In the field, we walked the runout path from the head of the landslide to the terminus of the runout deposit noting the

501 location and thickness of deposition and erosion, deposit composition, the slope of channel-filling deposits and

502 maximum grain size in the deposit. Geometric measurements in the runout path were measured using a combination

503 of a clinometer, steel tape and range finder. Location was measured using a handheld GPS with maximum accuracy

504 = +/- 1.8 m. Maximum grain size (which at all sites was cobble to boulder sized) was visually approximated in the

505 field or from photos that included a scale object, to the nearest 5 cm at eight to sixteen points along the runout path.

506 We estimated deposition depth from the thickness of fluvially eroded terraces and burial depth of stumps or trees. In

507 channels that were eroded to the fresh bedrock surface, we estimated erosion depth using the regolith thickness visible

508 along the edges of scoured channel walls (Olympic Mountain site). At other locations, we used vegetative indicators

509 such as the remnants of in-situ tree roots (Rocky Mountain site).

510 Remote observations included air photos (DNR, 2022; Google Earth, 2022) and lidar DEMs (DNR, 2022; Open

511 Topography, 2022) of the landslide and runout path. The lidar DEMs were used to define the model domain. The air

512 photos were used in combination with field observations and the lidar DEM to interpret the lateral extent of the

513 observed runout. Where pre- and post-mass wasting event lidar DEMs were available, we subtracted the pre-event

514 DEM from the post-event DEM to create the DoD. We found that not all DEMs or regions of the DEM were of equal

515 quality and therefore, where possible, DoD maps derived from lidar differencing were checked in the field. At sites

516 with only one post-event lidar DEM (Rocky Mountain and Olympic Mountain), we created DoD entirely from field

517 observations of deposition and erosion. Model setup details specific to each site are included in the Supplementary

518 Material.

519 **4.3 Estimating model parameter values from the observed runout**

520 We estimated all parameter values from field/remote observations of the observed runout except S_c and qs_c (See Table

521 1). For both S_c and qs_c, we defined uniform prior distributions based on field observations and then used the

522 calibration utility to find the best-fit parameter values (parameter values corresponding to the highest $L(\theta)$). Minimum

523 and maximum values of S_c were initially estimated from the range of observed slope of areas of positive-net deposition

524 (Table 1). Minimum and maximum values of qs_c were set as 0.01 to 1.75, which roughly represents the range of

525 minimum observed thickness of debris flow termini in the field at all of the validation sites. The range of S_c and qs_c

526 were later adjusted if we found the calibration utility tended to sample parameter values near the edge of the initial

527 parameter range. We approximated k from the DoD as follows:

528 The average erosion depth caused by the observed runout ($\bar{E}$) can be determined from the DoD as the total erosion

529 volume ($\sum E \Delta x^2$) divided by the erosion area ($\mathbb{A}$) in the DoD:

530 $$\bar{E} = \frac{\sum E \Delta x^2}{\mathbb{A}} \tag{32}$$

531 where $\sum E \Delta x^2$ and $\mathbb{A}$ exclude the initial landslide volume and area, areas of deposition ($+\Delta \eta$) and areas with no

532 change in elevation ($\Delta \eta = 0$). In terms of the debriton conceptualization used in MWR, $\bar{E}$ can also be written as a

533 function of the mean number of times a debriton would need to pass over a grid cell ($\bar{n}$) multiplied by an average

534 erosion depth per debriton ($\bar{h}_e$) to equal $\bar{E}$ as:

535 $$\bar{E} = \bar{n}\bar{h}_e \qquad (33)$$

536 An estimate for $\bar{n}$ can be determined from the average length of the runout material, which we approximate simply as

537 the mapped landslide length (ℓ) divided by the cell width:

538 $$\bar{n} = \frac{\ell}{\Delta x} \qquad (34)$$

539 Note that if the observed runout formed as a result of multiple landslides (as was the case at the Olympic Mountain

540 site), then ℓ was determined as the sum of the initial landslide lengths. Also, as the debritons move down slopes in

541 excess of S_c, they entrain material, split, and spread, and the runout material tends to lengthen. Using the initial

542 landslide length to represent the runout length thus represents a minimum value for $\bar{n}$ and if needed, (34) can be

543 multiplied by a coefficient to scale ℓ into a more representative runout length. Combining (33) and (34), $\bar{h}_e$ can be

544 defined as the average erosion rate per unit length of landslide ($\bar{E}/\ell$) times the cell width:

545 $$\bar{h}_e = \frac{\bar{E}\,\Delta x}{\ell} \qquad (35)$$

546 Rewriting (11) as a function of the average shear stress in the erosion-dominated reaches of the runout path ($\bar{\tau}$) and

547 assuming $\tau_c \cong 0$, debris flow erodibility parameter k can be estimated as:

548 $$k = \frac{\bar{h}_e}{\bar{\tau}^m} \qquad (36)$$

549 To solve for k, we estimated $\bar{\tau}$ from field-approximated debris flow depth and channel slope measurements in the

550 erosion-dominated reaches of the runout path. To estimate flow depth, we used the height of scour marks on the

551 channel wall or tree trunks, above the channel bed (Table 1). We used (13) to define $\bar{\tau}$. For D_s, we used the average

552 maximum grain size observed over the whole runout path. As noted in Section 2.3, if τ is defined as a function of

553 grain-collision dependent shear stress and k is determined as a function of m, as in (36), the impact of m on model

554 behavior is relatively small; however, the calibration utility did tend to more quickly converge on a set of best-fit

555 parameter values when a lower value for m was used and results herein use $m = 0.2$.

556 **4.4 Summary of observed runout characteristics**

557 The initial landslides (e.g., the initial landslide body) ranged in volume from 400 to 110,000 m3. At all sites, erosion

558 and subsequent entrainment added to the total mobilized volume (total erosion volume plus the initial landslide

559 volume, also referred to as the total flow volume or total debris flow volume; e.g., Barnhart et al. 2021; Reid et al;

560 2016), but the contribution was highly variable. The erosion volume divided by the total mobilized volume was as

561 low as 0.19 at the Cascade Mountain, 2022 landslide to as high as 0.97 at the Olympic Mountain landslides (Table 1).

562 The average maximum grain size varied from 0.2 m at the Black hills sites to nearly 1 meter at the Rocky Mountain

563 Site (Figure 5b, Table 1). Values of $\bar{E}/\ell$ ranged from 0.007 to 0.041 [m/m] with the highest rate occurring at the

564 Rocky Mountain landslide and the lowest at the Black Hills sites. For comparison with other reported debris flow

565 erosion rates, the observed erosion rates are also listed in terms of a yield rate or growth factor (average volumetric

566 erosion per unit length of the erosion-dominated region of the runout path, Hungr et al. 1984; Reid et al., 2016). The

567 observed growth factors ranged from 10 m^3/m at the Black Hills South site to 95 m^3/m during the Rocky Mountain

568 landslide (Table 1).

569 The topographic slope at which observed deposition occurred, measured as the slope of all cells of the 10 m DEM

570 representation of the terrain in which observed $\Delta\eta > 0$, was as high as 65% at some sites (Figure 5c); however positive

571 net-deposition tended to occur at lower slopes (Table 1). The slope of areas with positive net deposition tended to be

572 lowest at the Black Hills site (<1% to 10%) and highest at Rocky Mountain site (16% to 25%).

573 **4.5 Calibration and model performance**

574 Markov chains, colored according to the likelihood index, $L(\theta)$ are plotted in the S_c - qs_c domain, along with

575 histograms of sampled S_c and qs_c values for each landslide in Figure 6. The chains show a wide array of sampling

576 patterns and parameter ranges. Broadly speaking, at all sites, the chains converged, with growing $L(\theta)$, towards either

577 a small bullseye (b, c, d, f), or wider aerial patterns (a, e). In the former, the model shows high performance with only

578 a single or a very narrow range of few S_c- qs_c pairs while in the latter, similar model performance (nearly equivalent

579 $L(\theta)$ values) can be achieved or a wider range of sample pairs. The convergence patterns also vary.

580 The Cascade Mountain, 2009 (Figure 6a) and Cascade Mountain, 2022 (Figure 6b) landslides, which flowed over the

581 same hillslope, reveal that best-fit model parameters can vary at the same location. The calibration utility converged

582 to smaller qs_c and S_c values for the Cascade Mountain, 2009 event, which permitted thinner flow over lower slopes

583 and effectively made the 2009 modeled runout more mobile relative to the 2022 modelled runout. Even though MWR

584 model does not include an explicit representation of the rheology or momentum of the runout material, to a certain

585 degree, through calibration, it can be parameterized in relation to the initial volume and runout behavior of the slide

586 to implicitly represent those processes. On the other hand, the Cascade Mountain, 2009 and 2022 models used different

587 underlying initial DEMs and differences in parameterization could reflect the degree to which those DEMs represented

588 actual terrain (See Supplementary Material).
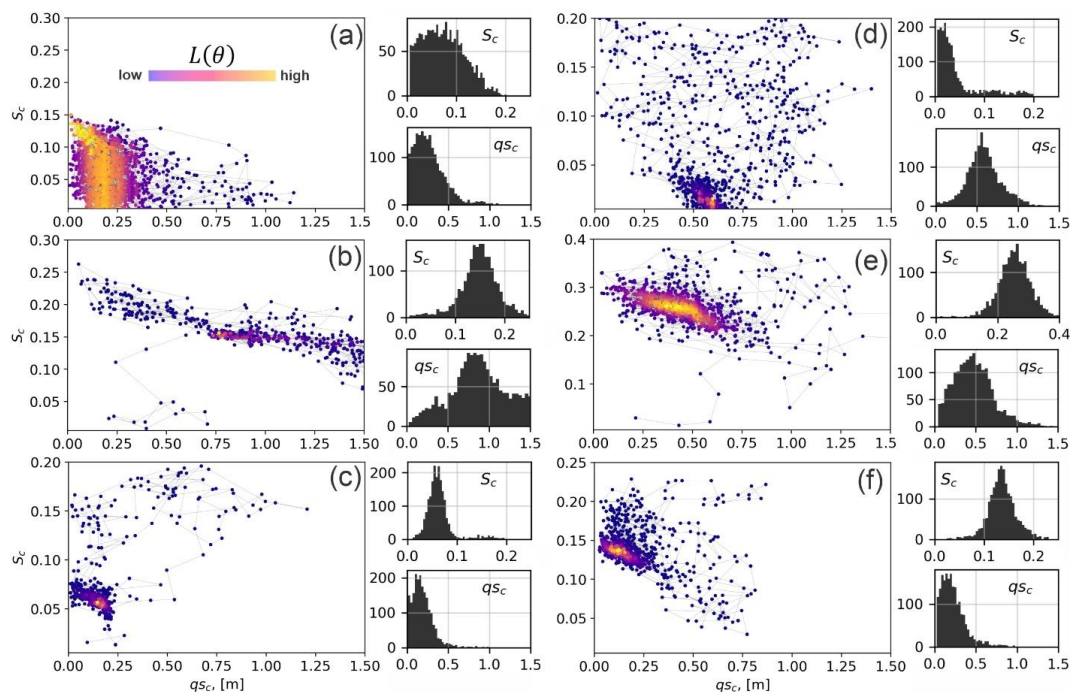
Figure 6. MWR calibration results to landslides described in Section 4 for (a) Cascade Mountain, 2009; (b) Cascade Mountain, 2022, (c) Black Hills, South; (d) Black Hills, North; (e) Rocky Mountain and; (f) Olympic Mountain. Each result shows a scatter plot of the sampled S_c and qs_c values, colored by their relative $L(\theta)$ value. To the right of each scatter plot are histograms of the iterated S_c and qs_c parameters that improved the model results, which represent an empirical PDF of the possible S_c and qs_c values that calibrate MWR to each site.

Additionally, inspection of MWR flow behavior at of the Cascade Mountain, 2009 slide (Figure 7) shows how modeled runout generally flows perpendicular to the topography lines of the underlying terrain, reflecting the impact of Equation 3 on flow direction. In Figure 8, profile plots of modeled Q_s (Equation 29) and maps of the modeled planimetric runout extent, colored to indicate where the runout matched (α), overestimated (β) or underestimated (γ) the observed runout (Equation 26) also illustrate observed and modeled flow behavior. At the Cascade Mountain, 2009 site, because a large portion of the transported sediment was lost to fluvial erosion in the valley, we limited the comparison of observed and modeled Q_s to the upper 750 m of the hillslope (see Supplementary Material).

In general, MWR successfully captures the spatial variation in Q_s and runout extent at most sites (Figure 8). The profile plots of Q_s at the Cascade Mountain site (Figure 8a and 8b) show that during the 2009 landslide, all of the runout material flowed past the first 750 meters of the runout path (Figure 8a) while during the 2022 landslide, material began to deposit just down slope of the initial landslide scar (Figure 8b). The model captures this behavior. At the Olympic Mountain site, MWR successfully replicates the observed spatially variable pattern of deposition and erosion (profile plot of Q_s in Figure 8f). This finding is notable, because at the Olympic Mountain site, the observed runout extent and sediment depositional pattern was formed by a sequence of landslides and subsequent runout events, rather than a single event, and each event altered the location of available sediment for the next. Also, at the Olympic

23

611   Mountain site, woody debris played a major role in determining the terminal extent of the runout (See Supplementary

612   Information). MWR does not include an explicit representation of woody debris impacts on runout. Furthermore, as

613   will be detailed below, topographic features of the channels that appeared to control observed runout were not well

614   represented in the 10-m DEM of the Olympic Mountain site. Nonetheless, the calibration utility found parameter

615   values that allowed the model to replicate the observed sediment transport and depositional behavior.

616   Regarding how well modelled runout replicated the planimetric extent of the observed runout, values of Ω_T are

617   comparable or higher than reported values of Ω_T for landslides in California that were modeled using a variety of

618   detailed, mechanistic and top-down models (Gorr et al., 2022; Barnhart et al., 2022; Note, to compare Ω_T values to

619   those studies, subtract 1 from values reported in this study). Overall, calibration was best at the Cascade Mountain,

620   2009 landslide (values of Ω_T are highest and values of $\Delta\eta_E$ and Q_{s_E} are lowest) and poorest at the Rocky Mountain

621   and Olympic Mountain sites (Values of Ω_T are lowest Q_{s_E} and $\Delta\eta_E$ are highest). As noted in Section 4.2, at both the

622   Rocky Mountain and Olympic Mountain sites, because we lacked repeat lidar, we created the DoD from a map of

623   field estimated erosion and deposition depths and estimated the pre-event DEM. The lower calibration scores may

624   indicate that field estimated DoDs or pre-event DEMS were not as accurate as those determined from repeat lidar

625   DEMS. Additionally, in the field at the Rocky Mountain site, standing trees and woody debris along the edge of the

626   runout path appeared to have controlled the lateral extent of the observed runout (See Supplementary Material). As a

627   noted above, MWR model does not explicitly represent of the impacts of wood on runout behavior, but by selecting

628   high qs_c values, the laterial extent of the modelled flow can be reduced. At the same time, increasingly qs_c reduces

629   the longitudinal extent of the modelled flow. The calibration utility identified S_c- qs_c pairs that allowed MWR to best

630   mimic the combined impact of woody debris and topography on the observed runout.

631   Finally, MWR also tended to underestimate the runout width of large landslides where the apparent momentum of the

632   observed the runout allowed it to flow upslope or slope-parallel. For example, after the modeled Cascade Mountain,

633   2009 landslide initially failed in the uplands, the debritons split and then rejoined to flow around a broad ridge in the

634   middle of the runout path and then were funneled into a narrow ravine (iteration t=40 in Figure 7), rather than flowing

635   over and inundating the broad ridge as was observed (Large red zone in Figure 8a; see Supplementary Material). On

636   the contrary, for small to moderately sized landslides that flowed in narrow, highly convergent channels, (Olympic

637   Mountain and Cascade Mountain, 2022 sites; Figures 8f and 8b). Overestimation of width is particularly pronounced

638   across most locations of the channel network at the Olympic Mountain site.

639

640
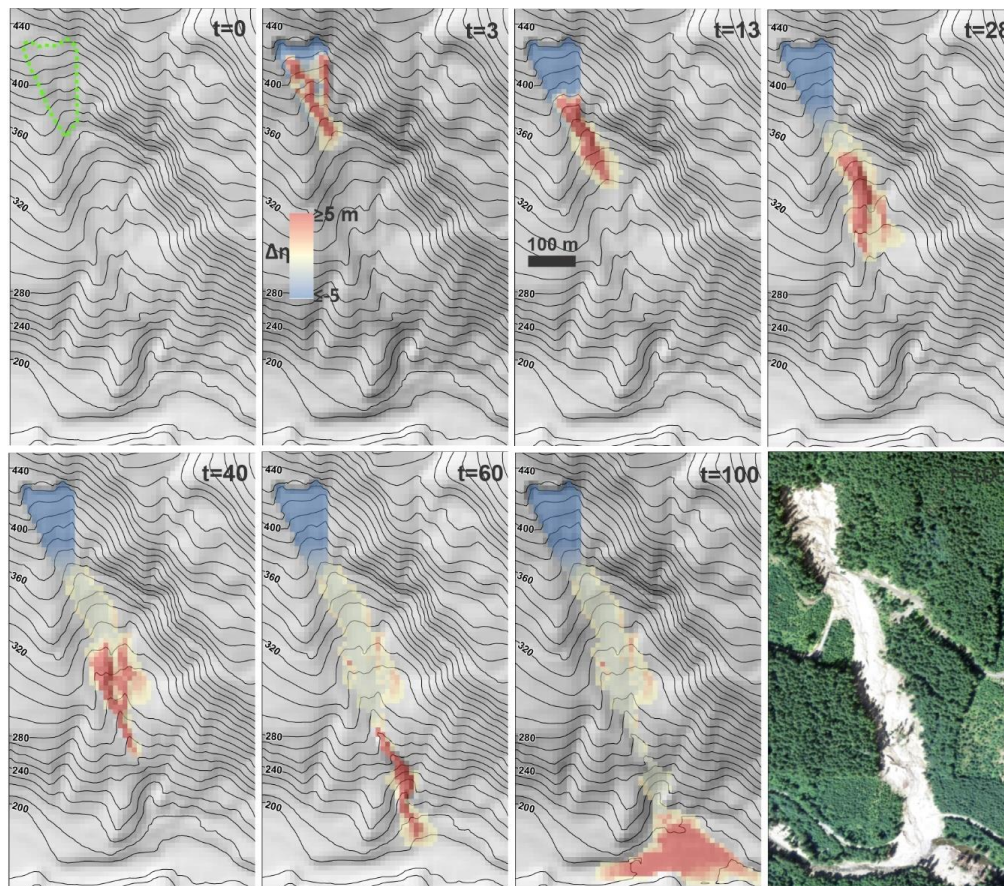
Figure 7. Illustration of modeled runout at the Cascade Mountain, 2009 landslide. At iteration $t = 0$, Algorithm 1 determines the direction and flux of the initial debritons (all nodes located in the landslide green-dashed polygon). In later iterations, Algorithm 2 routes the debritons down slope, entrains material, forms new debritons and updates 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.
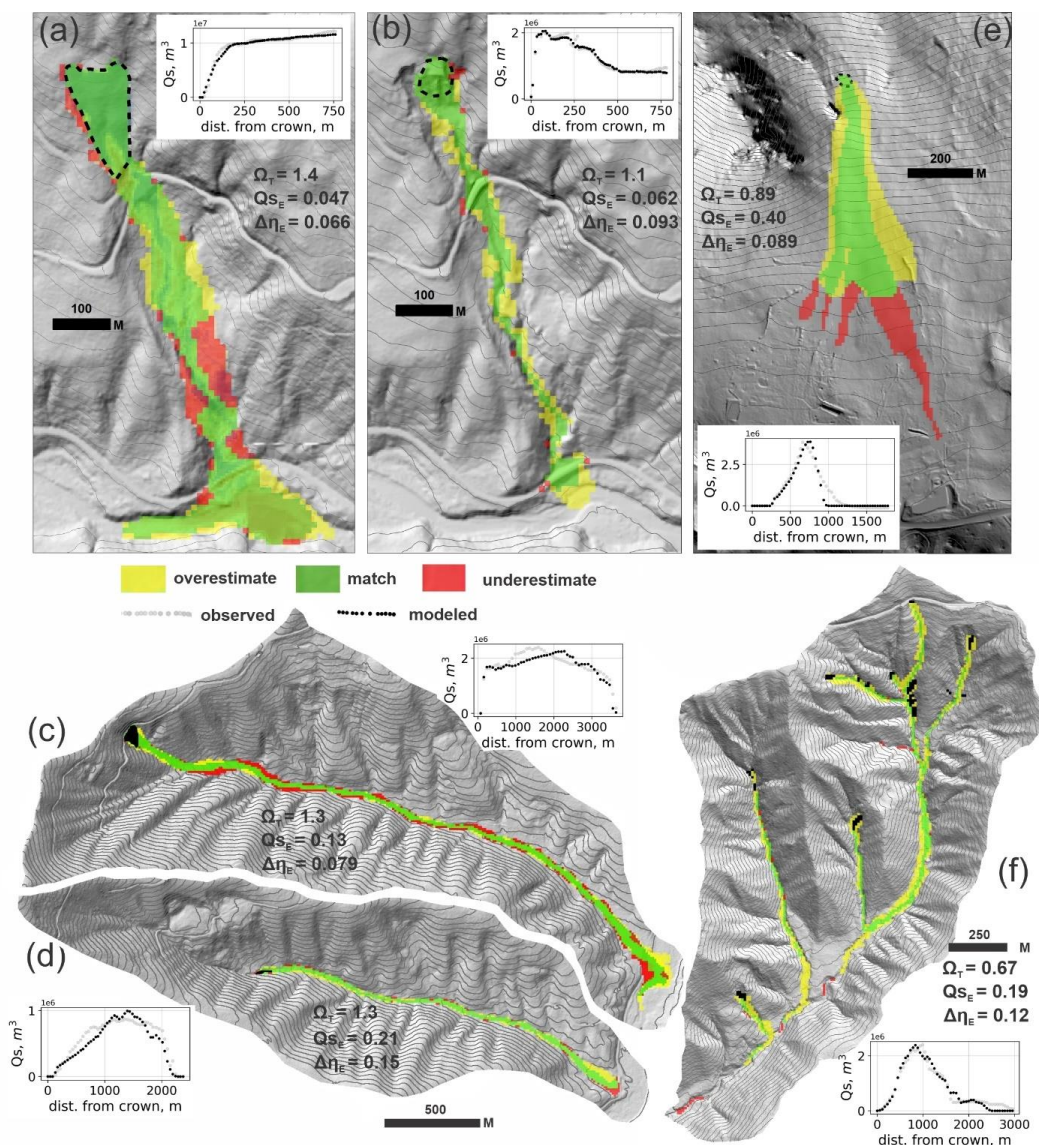
Figure 8. Calibrated model performance as indicated by modeled runout extent, profile plots of Q_s, and reported values of Ω_T, $\Delta\eta_E$ and Q_{s_E}. In all maps, up is north except in (e), north is towards the left. (a) Cascade Mountain, 2009; (b) Cascade Mountais, 2022; (c) Black Hills, South; (d) Black Hills, North; e) Rocky Mountain; (f) Olympic Mountain.

To understand whether the ability to calibrate MWR systematically varies with topography of the runout path, we compared model performance with three topographic indices computed from the terrain in the observed runout extent. The indices 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

657   of the runout (measured from the top of the initial landslide to the end of the runout path) divided by the horizontal

658   length of the runout and indicates the internal and external resistance to the observed mass wasting runout (Iverson,

659   1997). Index κ represents topographic curvature, which is the second derivative of the terrain surface, with

660   increasingly positive values of index κ reflecting increasing topographic concave-up and convergent topography (e.g.,

661   Istanbulluoglu et al., 2008). The index SPI is determined as the natural log of the product of the contributing area and

662   slope. We computed indices κ and SPI at each node in the observed runout extent and the mean was determined from

663   all nodes in observed runout extent. For a more detailed description of the topographic indices, see , Chen and Yu,

664   (2011).

665   Comparison of model performance with respect to the topographic indices in Figure 9 shows: slightly improved model

666   performance over runout-paths that are less convergent (SPI and κ values of the observed runout path are lower) and

667   slightly improved performance on steeper terrain (higher H/L). The former finding may have to do with how well

668   topographic features that controlled the observed runout were represented by the model DEM. In all models, we used

669   a 10-meter DEM (see Section 4.2), which may have lacked sufficient topographic detail to represent the narrow first

670   and second order channels that controlled observed runout of the Olympic Mountain and the Cascade Mountain, 2022

671   sites (Supplementary Material). Without a sufficiently accurate topographic representation of the topographic features

672   that controlled observed flow, modeled flow tended to flow over a wider area of the hillslope than observed (see Figure

673   8b and 8g). The latter finding appears to be mostly a result of how well modeled sediment transport and topographic

674   change (Q_{s_E} and $\Delta\eta_E$) replicated observed, as there does not appear to be a trend in Ω_T with H/L and the two best

675   performing models (Cascade Mountain landslides) both had the lowest Q_{s_E} values and low $\Delta\eta_E$ values.

676   In summary, using a variety of methods to prepare the DoD and model terrain, the calibration utility was able to

677   identify clear best-fit S_c and qs_c values at a range of different landslide types. Calibration likelihood ($L(\theta)$) tended to

678   vary with topographic indices of the runout path (H/L, κ and SPI). Model performance was impacted by how well the

679   DEM represented topographic features that controlled observed runout, the importance of runout processes not

680   explicitly represented in MWR on observed runout behavior (such standing trees and woody debris along the edge of

681   the Rocky Mountain landslide) and the accuracy of the DoD dataset (field mapped vs lidar DEM determined). To a

682   certain degree, though calibration, MWR can be parameterized to compensate for deficiencies in the DEM or processes

683   not explicitly represented in the model.
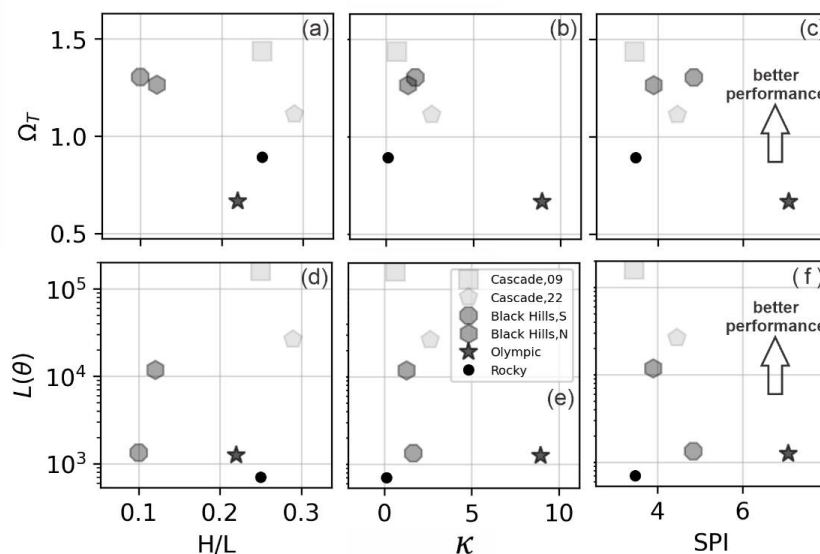
Figure 9. Illustration of model calibration, as reflected by the posterior parameter likelihood $L(\theta)$ and planimetric fit (Ω_T) relative to topographic indices.

5. Discussion

5.1 Strategic testing of MWR for hazard mapping applications

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

As shown in Figure 10, at three of the landslides (both Cascade Mountain landslides and the Black Hills, North landslide), when the best-fit parameters from the other landslide are used to predict runout, the accuracy of modeled runout planimetric extent drops but resultant Ω_T values can still be as high or higher than values reported in other studies (compare to equivalent Ω_T values in Gorr et al., 2022 and Barnhart et al., 2022). In terms of modeled sediment transport and topographic change, swapping best-fit parameters has a larger effect. At the Cascade Mountain, 2009 landslide, using the 2022 best-fit parameter values causes about half of the modeled runout material to prematurely deposit on the hillslope, reducing the amount of sediment that reaches the valley floor (Q_{s_E} increases by a factor of nine; Figure 10). Using the Cascade Mountain, 2009 parameter values for Cascade Mountain, 2022 landslide (Figure 10b) increases modeled runout extent and results in nearly four times the entrainment and transport of sediment to the valley floor, causing Q_{s_E} to increase by a factor of 20 and $\Delta\eta_E$ by 83%. At the Black Hills site, using the South basin best-fit model parameters at the North basin causes Q_{s_E} and $\Delta\eta_E$ increase by 83% and 39% respectively (Figure 10c).

706 Unlike the other three landslides, swapping best-fit parameters at the Black Hills, South landslide results in both large

707 sediment transport and runout extent error because the North basin best-fit parameters cause modeled landslide to

708 entrain too little and stop only a few hundred meters from the initial source area (Figure 10d).
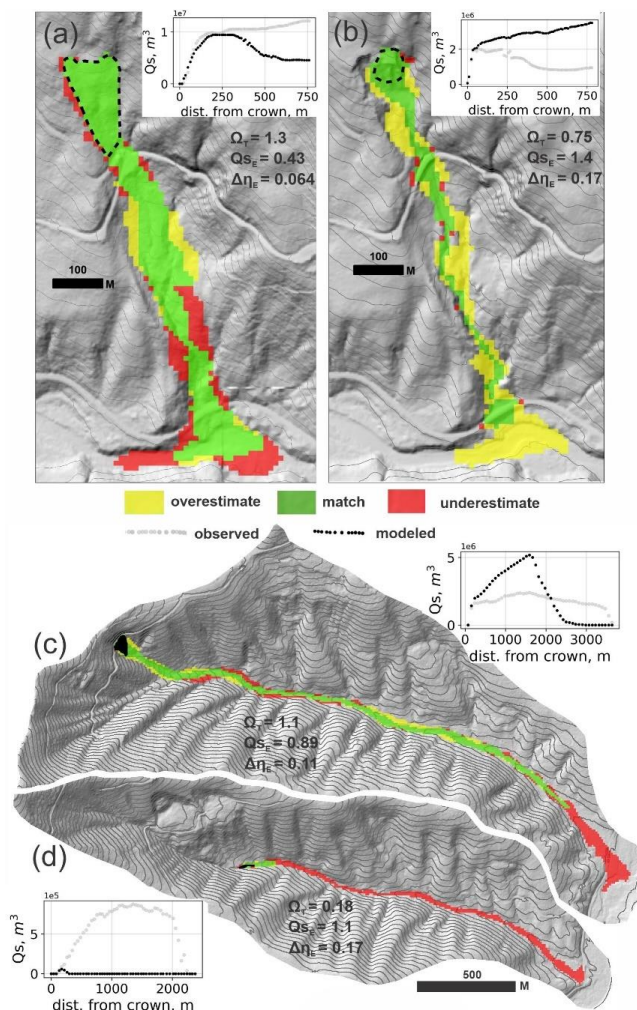
Figure 10. Model performance using the neighboring landslide parameter values, as indicated by modeled runout
extent, profile plots of Q_s, reported values of Ω_T, $\Delta\eta_E$ and Q_{s_E} and profile plots of Q_s. Compare with Figure 8. (a)
Cascade Mountain, 2009; (b) Cascade Mountain, 2022; (c) Black Hills, North; (d) Black Hills, South

714 An alternative test to simply swapping the best-fit parameters is to swap parameter PDFs determined from the

715 calibration utility and compare probabilistic runout. For this test, we run MWR Probability with option (1), using the

716 field mapped landslide polygons. As shown in Figure 11, similar to the first test, at three of the landslides, using the

717 parameter distribution associated with the neighboring landslide results in relatively minor changes in the probability

718 of runout (probability of ≥50%; Figures 11a, 11b and 11d) and mostly changes the terminal extent of the possible

719   landslide runout area. At the Black Hills South landslide, swapping parameter PDFs causes a large change in runout

720   probability (Figure 11c).

721
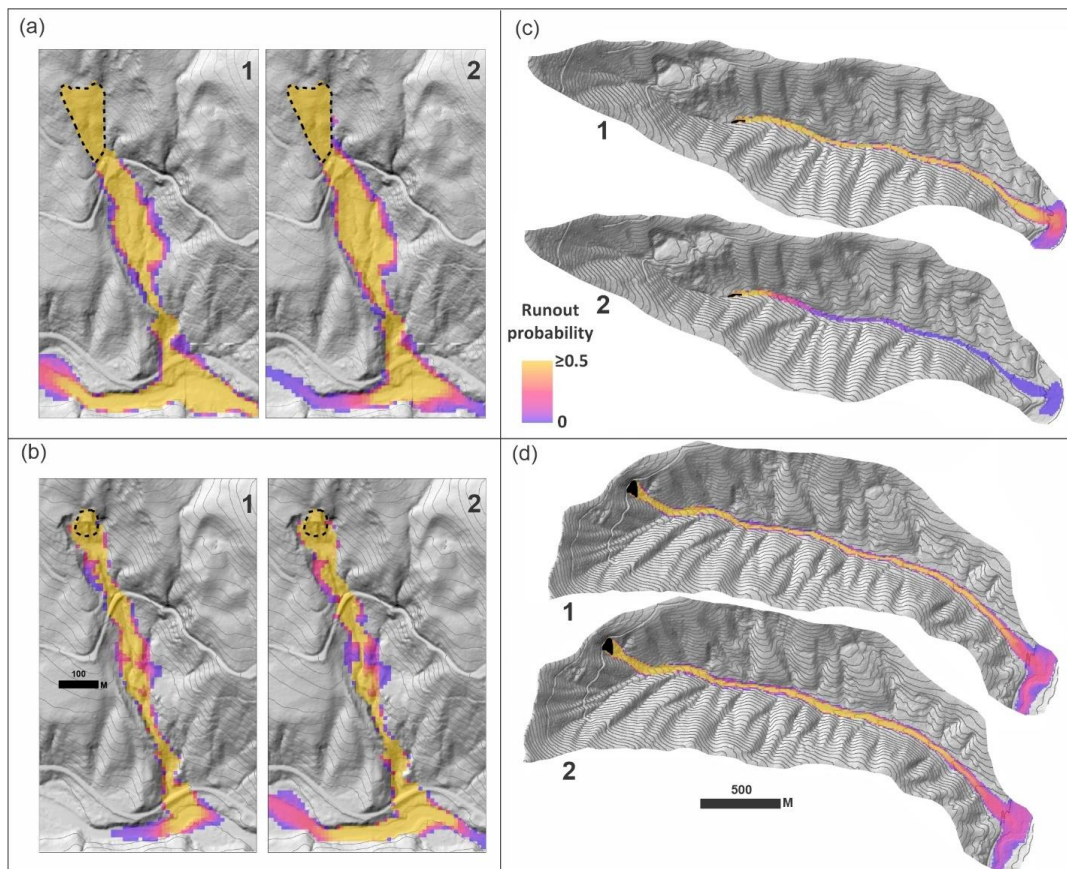
722
723   **Figure 11.** Model tests by swapping parameter PDFs and comparing runout probability at the (a) Cascade Mountain,
724   2009; (b) Cascade Mountain, 2022; (c) Black Hills, South and; (d) Black Hills, North sites. (**1**) runout using
725   parameter distributions of the site and (**2**) runout using parameter distributions of the neighboring site.
726

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

728   established for a landslide, those parameter/PDF values can then be used to assess runout extent at nearby sites.

729   Unfortunately, it appears the same cannot be said for sediment transport and topographic change prediction, both of

730   which appear to be more sensitive to model parameterization. However, we suspect that this finding is a consequence

731   of comparing very different landslide types to conduct the tests. Had landslide processes at the Cascade Mountain and

732   Black Hills sites been more uniform, test results would have been better. For example, at the Olympic Mountain site,

733   a single parameterization seemed to capture both the runout extent and depositional patterns of all slides. There, most

734   landslides tended to be stream adjacent and relatively uniform in size. In contrast, at the Cascade Mountain site, the

735   2009 landslide was five times the volume of the 2022 landslide and both the volume and apparent momentum of the

736   2009 slide allowed it to flow topographically unconstrained while the much smaller 2022 landslide was

737 topographically confined to a narrow channel ravine. The calibration utility found that relatively low S_c and qs_c values

738 best replicated the Cascade Mountian, 2009 flow behavior while higher S_c and qs_c values best replicated the Cascade

739 Mountain, 2022 behavior. Similarly, at the Black Hills site, the north watershed landslide was over ten times the

740 volume of the southern watershed landslide, and initiated 150 hundred meters above the channel, on a steep hillside

741 (detailed in Supplementary Material) while the much smaller south watershed landslide initiated at the elevation of

742 the channel bed. Again, differences in landslide behavior equated to different parameterizations of MWR. Had the

743 Cascade Mountain, 2009 and 2022 landslides or the Black Hills North and South landslides been similar in size and

744 runout behavior, best-fit S_c and qs_c values and PDFs may have been similar and test results may have been better.

745 In summary, the Cascade Mountain and Black Hills sites each contained two landslides that we could use to

746 independently calibrate MWR and thereby evaluate model predictive ability by swapping parameter values and

747 comparing model results with calibrated model results. At three of the landslides, the tests revealed a modest decrease

748 in the accuracy of modeled runout extent (though accuracy of the reduced model extent was still comparable to values

749 reported in the literature) but a substantial decrease in the accuracy of modelled sediment transport and modeled

750 topographic change. We suspect that the poor sediment transport and topographic change test results are a consequence

751 of comparing very different landslide types and runout processes. In regions where landslide processes are relatively

752 uniform (like the Olympic Mountain site), calibration to one landslide might be sufficient to predict the depositional

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

754 landslide processes, MWR may need to be calibrated to each type of landslide (e.g., small-low-slope vs. large-steep-

755 slope) and predictive application of might involve applying the appropriate parameter set based on landslide type.

756 **5.2 MassWastingRunout probability applications**

757 In this section we demonstrate how to determine runout probability from a probabilistically determined landslide

758 hazard map or a specific, potentially unstable slope using MWR. The first application may be appropriate for

759 watershed- to regional-scale runout hazard assessments. The second application is an example site-scale hazard

760 assessment. Both applications are demonstrated at the Olympic Mountain site where landslide size and type tended to

761 be relatively uniform and parameter PDFs determined through calibration may therefore represent typical runout

762 processes in the basin.

763 **5.2.1 Runout probability from a landslide hazard map**

764 To determine runout probability from a landslide hazard map, we ran MWR Probability using option (3) at the

765 Olympic Mountain site. We use LandslideProbability (Strauch et al., 2018), an existing component in Landlab that

766 computes landslide probability by repeatedly determining the hillslope Factor-of-Safety (FS: ratio of the resisting to

767 the driving forces) at each node on the raster model grid from stochastically selected soil (regolith) hydrology

768 properties, (including soil depth, saturated hydraulic conductivity) soil strength (friction angle, cohesion) and recharge

769 rates (precipitation input rate minus evapotranspiration and soil storage). We setup LandslideProbability using soil

770 strength and hydrologic parameters based on the soil class (SSURGO, 2020) and vegetation type following Strauch et

771 al. (2018). We defined soil depth as a function of contributing area, slope gradient and minimum and maximum depths

772   following Westrick (1999). Minimum and maximum soil depths were estimated from field observations at channel

773   banks, tree-throw pits and road cuts. For recharge, we assumed evapotranspiration and soil storage losses were small

774   and used a uniform daily precipitation depth equivalent to the 50 year event (determined from rainfall records, WRCC,

775   2017).

776   We ran $N = 1000$ model runs. For each model run, FS and soil depth maps were saved and after all N runs were

777   completed, a single map of landslide probability was produced. We then ran MWR Probability using each of the N

778   FS and soil depth maps and treated all nodes with $FS < 1$ as a landslide source. Each repetition began with the same

779   initial topography but different landslide locations and regolith thickness. Landslide probability and landslide runout

780   probability are shown in Figures 12a and 12b. If changes in the elevation of the channel associated with landslide

781   runout were a concern, results of MWR Probability can also be used determine the probability of scour or deposition.

782   For example, the probability of deposition greater than 1 meter is shown in Figure 12c.

783   This example application of MWR Probability shows that while landslide probability is relatively low over most of

784   the Olympic Mountain basin, runout probability is high in many of the second order channels but then low again at

785   the basin outlet. In a field study of the spatial and temporal patterns of debris-flows, May and Gresswell (2004)

786   observed that debris flow return interval tends to be low (debris flow probability tends to be high) in channels which

787   are fed by a greater number of upslope landslide source areas; however, our model results illustrate how debris flow

788   (landslide runout) probability is both a function of the number of upslope landslide source areas as well as the typical

789   runout distance of those source areas. For example, despite having the largest number of potential landslide source

790   areas, runout probability at the basin outlet is low because it is located beyond the typical runout distance of landslides

791   at the Olympic Mountain site.

5.2.2 Runout probability for a specific, potentially unstable slope

793   When field evidence or other data indicate that a specific slope may be potentially unstable, but the exact area of a

794   potential slide on that slope is unknown, MWR can be used to determine potential runout hazard of the slope. For this

795   application, MWR Probability is run using option (2) which requires a polygon representing the maximum extent of

796   the potentially unstable slope. As noted in Section 3.2, for each model repetition, a landslide area is randomly selected

797   within the potentially unstable slope, it can form anywhere within the potentially unstable slope and is least as large

798   as a user defined minimum size but no larger than the potentially unstable slope. After a user defined N repetitions,

799   runout probability of the potentially unstable slope is determined at each downslope node from (31).

800   As an example application of using MWR Probability option (2), we designated a 0.6 ha, zero-order basin in the

801   headwaters of the Olympic Mountain site as a potentially unstable slope and modeled runout probability using $N =$

802   1000 (Figure 12d). This example shows that if a landslide were to initiate at the potentially unstable slope, the

803   probability of that material running out to the basin outlet as a mass wasting runout process is less than 5%. This result

804   is again a function of the typical runout distance of landslides at the Olympic Mountain site.
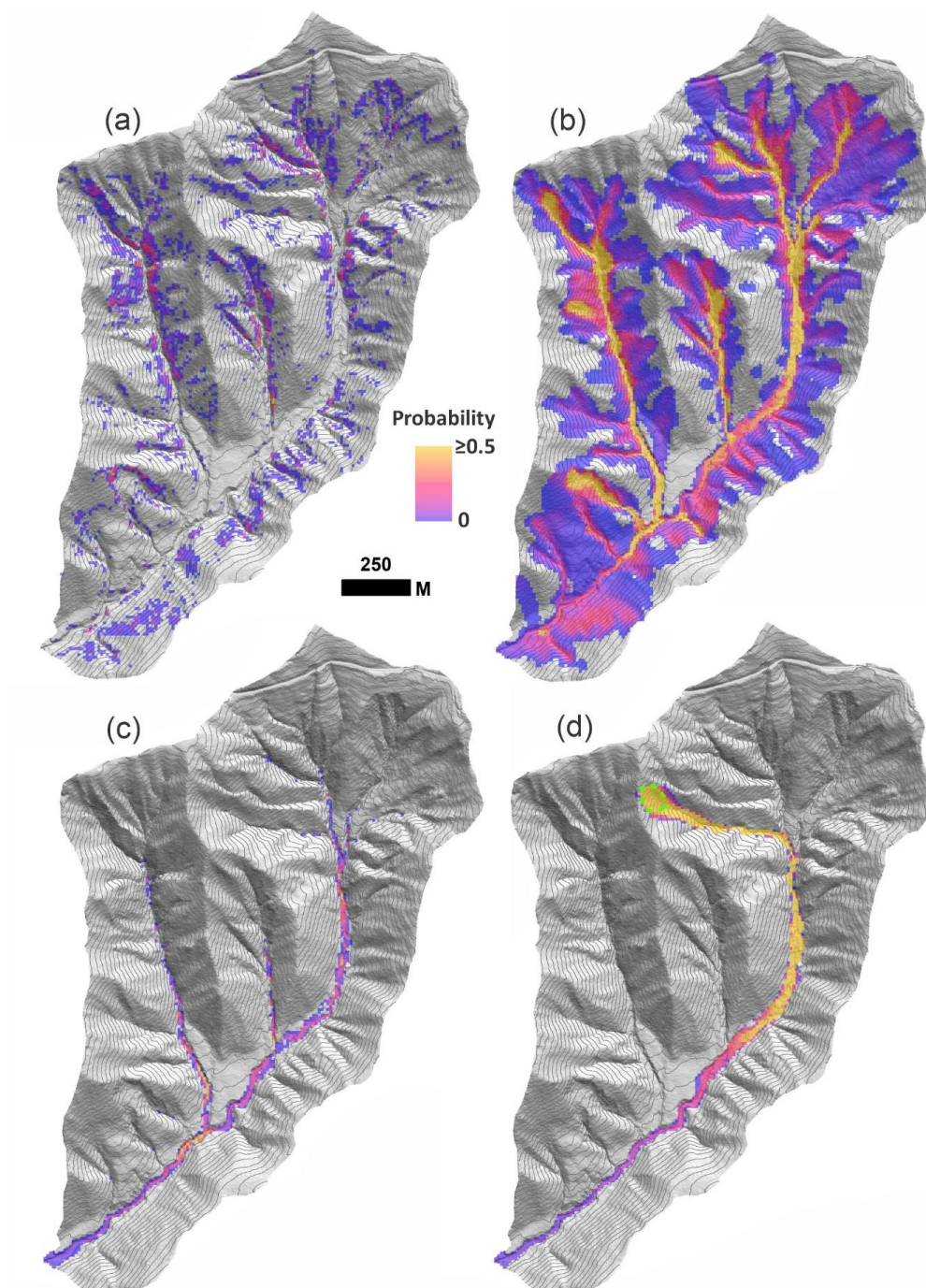
Figure 12. Olympic Mountain site: **(a)** Landslide probability, $P(FS \leq 1)$. **(b)** Corresponding runout probability, $P(\Delta \eta)$. **(c)** Probability of deposition greater than 1 m and **(d)** Runout probability for the potentially unstable slope (green-dashed polygon).

5.3 Model limitations and strengths

809

810 In the field, we observed a number of runout phenomenon that are not explicitly represented by MWR (detailed in

811 Supplementary Material), such as the apparent affect of momentum, debris fluidization, sorted deposition and the

812 impact of standing vegetation and woody debris on the runout process. For example, observed runout at the Olympic

813 Mountain site appeared to be restricted by log jams which in turn lead to large, valley-filling sediment deposits. At

814 the Rocky Mountain site, the lateral extent of the runout path appeared to have been restricted by standing trees. Also,

815 the momentum or perhaps fluidization of the initial landslide appears to have impacted small areas of the runout extent

816 at Black Hills North landslide and the Cascade Mountain, 2009 landslide.

817 Additionally, grid-size appears to have affected the lateral extent of modeled runout at some of the sites. At all sites,

818 we used a 10-m grid size and the topographic detail of the channels that constrained observed flow to a narrow width

819 at the Olympic Mountain and Cascade Mountain, 2022 sites may not have been well represented. On the other hand,

820 for landslides like the Cascade Mountain, 2009 slide and the Rocky Mountain slide, the 10-m grid size more than

821 adequately captured the key topographic features that controlled runout width and a lower resolution (larger grid size)

822 DEM may have resulted in equivalent model performance. This last point is important because the number of debritons

823 and computations needed to model the initial landslide scales with the initial landslide area divided by the square of

824 the grid size (e.g., a 100 m^2 landslide on a 10 m grid is represented by 1 debriton but on a 1-m grid is represented by

825 100 debritons). Using an excessively small grid size may result in an unnecessary number of computations. We used

826 a 10-m grid size because we assumed it was the coarsest grid size that also captured the topographic features that

827 controlled observed runout. In future work, we hope to provide a more comprehensive assessment of grid size impact

828 on model performance.

829 Nonetheless, as noted in Section 4.5, to a certain degree, through calibration, MRW can be parameterized to

830 compensate for low DEM resolution or the absence of explicit representation of all processes that controlled the

831 observed flow: at the Rocky Mountain site, the qs_c parameter could be used to mimic the resistance to lateral spreading

832 caused by the trees; at the Olympic Mountain site, even though modeled flows tended to be wider than observed and

833 woody debris affected deposition, the calibration utility identified qs_c and S_c values that allowed MWR to replicate

834 the longitudinal sediment transport behavior of the observed runout; and at the Cascade, 2009 site, the calibration

835 utility identified parameter values that allowed modelled flow to better mimic the high momentum behavior of the

836 observed runout.

837 Finally, through strategic testing, we showed that once MWR is calibrated to a landslide it may reasonably evaluate

838 the extent of runout or runout probability at neighboring slopes. We suspect that this may be particularly true in areas

839 that have relatively uniform landslide processes (Olympic Mountain site). At sites like the Black Hills and Cascade

840 Mountain sites, where landslide processes are highly variable, MWR may have limited predictive ability for sediment

841 transport and topographic change modeling and in some cases even runout extent unless the model is calibrated to

842 each type of landslide.

843 The utility of a geomorphic model is ultimately judged by its ability to explain or predict natural phenomena, even if

844 the interactions within the model greatly simplify the physical processes controlling the phenomena (Murray 2007;

845 Murray 2013). Furthermore, a model should be rejected if it fails to react to changes in forcing in a reasonable way or

846   cannot replicate the intended phenomena. From this standard, MWR appears to pass as an acceptable model, because

847   it was able to replicate a range of landslide behavior, including the Olympic Mountain site, where the final depositional

848   pattern was the result of a sequence of landslides and because it responds in an expected way to changes in terrain or

849   parameters like S_c, qs_c and grain size. MWR is a parameterized model and while the predictive application of MWR

850   may not be as restricted as implied by Iverson (2003), we suspect that calibration may be required for most

851   applications. For the purpose of facilitating rapid calibration to different sites and developing parameter PDFs, we

852   made MWR parameterization as field-relatable as possible and developed a calibration utility.

853   **6.0 Conclusion**

854   In this study, we described, calibrated and tested MassWastingRunout (MWR), a new cellular-automata landslide

855   runout model designed for probabilistic landslide hazard assessments, sediment transport and topographic change

856   applications. MWR is implemented in Python as a component for the Landlab earth surface modeling toolkit. MWR

857   includes a Markov Chain Monte Carlo calibration utility that determines the best-fit parameter values for a site as well

858   as empirical Probability Density Functions (PDF) of the parameter values. MWR also includes a utility called MWR

859   Probability that takes the PDF output from the calibration utility to model runout probability. It has three options for

860   defining the initial landslide: (1) a single, fixed-size landslide, (2) a single, variable-size landslide within a user defined

861   potentially unstable slope or (3) a single or multiple landslides that change in location and size each model iteration.

862   We used the calibration utility to parameterize the model to six different landslides. The following results were

863   obtained:

864   •   MWR can be calibrated to a range of landslide types. The ability to calibrate MWR is slightly better on a

865   steep (high H/L) and broadly convergent (lower κ and SPI) runout path. But this finding may partly be a

866   consequence of the 10-m grid-size used to model runout at each of the six landslides, and needs further

867   research. The six landslides varied from small to large and the ability to calibrate MWR appears to be better

868   at the sites in which the terrain features that controlled runout extent were larger than the grid size.

869   •   The two model parameters, critical slope (S_c), and a threshold flux for deposition (qs_c) are found to

870   effectively control model performance over a range of landslide and landscape conditions and, to a certain

871   degree, through calibration, MWR can be parameterized to mimic the effect of processes or terrain not

872   explicitly represented in the model.

873   •   Once MWR is calibrated to an observed-runout dataset, it may work as a predictive tool for assessing the

874   runout extent of neighboring potential landslides. To predict depositional and topographic change, depending

875   on the diversity of landslide types in the region, a more rigorous calibration may be required such as

876   calibrating the model to different landslide types.

877   •   We demonstrated how to use MWR to determine mass wasting runout probability from an expert-defined

878   potentially unstable slope polygon and a landslide hazard map produced by the model LandslideProbability.

Competing interests

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

References

ArcGIS, https://desktop.arcgis.com/en/arcmap/latest/tools/spatial-analyst-toolbox/filter.htm, accessed 2022

Bagnold, R. A. (1954). 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

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. (2020). 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

Barnhart, K. R., Jones, R., George, D. J., McArdell, B. W., Rengers, F. K., Staley, D. M., & Kean, J. W. (2021). 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

Benda, L., & Cundy, T. W. (1990). Predicting deposition of debris flows in mountain channels. Canadian Geotechnical Journal, 27(4), 409–417. https://doi.org/10.1139/t90-057

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

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

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

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

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

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

909   Capart, H., Hung, C., & Stark, C. R. (2015). Depth-integrated equations for entraining granular flows in narrow

910   channels. Journal of Fluid Mechanics, 765. https://doi.org/10.1017/jfm.2014.713

911   Carretier, S., Martinod, P., Reich, M., & Goddéris, Y. (2016). Modelling sediment clasts transport during landscape

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

913   Chase, C. G. (1992). Fluvial landsculpting and the fractal dimension of topography. Geomorphology, 5(1–2), 39–57.

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

915   Chen, C., & Yu, F. (2011). Morphometric analysis of debris flows and their source areas using GIS. Geomorphology,

916   129(3–4), 387–397. https://doi.org/10.1016/j.geomorph.2011.03.002

917   Chen, H., & Zhang, L. (2015). EDDA 1.0: integrated simulation of debris flow erosion, deposition and property

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

919   Chen, T.-Y. K., Wu, Y.-C., Hung, C.-Y., Capart, H., and Voller, V. R. (2023). A control volume finite-element model

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

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

922   Clerici, A., & Perego, S. (2000). Simulation of the Parma River blockage by the Corniglio landslide (Northern Italy).

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

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

925   Coz, J. L., Renard, B., Bonnifait, L., Branger, F., & Boursicaud, R. L. (2014). Combining hydraulic knowledge and

926   uncertain gaugings in the estimation of hydrometric rating curves: A Bayesian approach. Journal of Hydrology, 509,

927   573–587. https://doi.org/10.1016/j.jhydrol.2013.11.016

928   Crave, A., & Davy, P. (2001). A stochastic "precipiton" model for simulating erosion/sedimentation dynamics.

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

930   D'Ambrosio, D., Di Gregorio, S., Iovine, G., Lupiano, V., Rongo, R., & Spataro, W. (2003). First simulations of the

931   Sarno   debris   flows   through   Cellular   Automata   modelling.   Geomorphology,   54(1–2),   91–117.

932   https://doi.org/10.1016/s0169-555x(03)00058-8

933   Department   of   Natural   Resources,   https://www.dnr.wa.gov/programs-and-services/buy-maps-aerial-photos-or-

934   survey-data

935   Egashira, S., Honda, N., & Itoh, T. (2001). Experimental study on the entrainment of bed material into debris flow.

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

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

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

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

940   Fannin, R. J., & Wise, M. P. (2001). An empirical-statistical model for debris flow travel distance. Canadian

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

942 Frank, F., McArdell, B. W., Huggel, C., & Vieli, A. (2015). The importance of entrainment and bulking on debris
943 flow runout modeling: examples from the Swiss Alps. Natural Hazards and Earth System Sciences, 15(11), 2569–
944 2583. https://doi.org/10.5194/nhess-15-2569-2015

945 Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2021). Bayesian Data Analysis
946 (3rd ed.). Electronic Edition.

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

950 Google Earth, https://earth.google.com/, accessed 2022

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

953 Guthrie, R. H., Deadman, P., Cabrera, A. R., & Evans, S. D. (2008). Exploring the magnitude–frequency distribution:
954 a cellular automata model for landslides. Landslides, 5(1), 151–159. https://doi.org/10.1007/s10346-007-0104-1

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

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

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

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

967 Horton, P., Jaboyedoff, M., Rudaz, B., & Zimmermann, M. N. (2013). Flow-R, a model for susceptibility mapping of
968 debris flows and other gravitational hazards at a regional scale. Natural Hazards and Earth System Sciences, 13(4),
969 869–885. https://doi.org/10.5194/nhess-13-869-2013

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

972 Hungr, O., Leroueil, S., & Picarelli, L. (2013). The Varnes classification of landslide types, an update. Landslides,
973 11(2), 167–194. https://doi.org/10.1007/s10346-013-0436-y

974   Hürlimann, M., Rickenmann, D., Medina, V., & Bateman, A. (2008). Evaluation of approaches to calculate debris-
975   flow parameters for hazard assessment. Engineering Geology, 102(3–4), 152–163.
976   https://doi.org/10.1016/j.enggeo.2008.03.012

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

979   Iovine, G., D'Ambrosio, D., & Di Gregorio, S. (2005). Applying genetic algorithms for calibrating a hexagonal
980   cellular automata model for the simulation of debris flows characterised by strong inertial effects. Geomorphology,
981   66(1–4), 287–303. https://doi.org/10.1016/j.geomorph.2004.09.017

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

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

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

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

992   Julien, P. Y., & Paris, A. (2010). Mean Velocity of Mudflows and Debris Flows. Journal of Hydraulic Engineering,
993   136(9), 676–679. https://doi.org/10.1061/(asce)hy.1943-7900.0000224

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

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

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

1000   Lee, C., Huang, W., Chang, Y., Chi, S., & Liao, W. (2018). Regional landslide susceptibility assessment using multi-
1001   stage remote sensing data along the coastal range highway in northeastern Taiwan. Geomorphology, 300, 113–127.
1002   https://doi.org/10.1016/j.geomorph.2017.10.019

1003   Lee, J. P., & Park, H. (2015). Assessment of shallow landslide susceptibility using the transient infiltration flow model
1004   and GIS-based probabilistic approach. Landslides, 13(5), 885–903. https://doi.org/10.1007/s10346-015-0646-6

1005   Liu, J., Wu, Y., Gao, X., & Zhang, X. (2022). A Simple Method of Mapping Landslides Runout Zones Considering
1006   Kinematic Uncertainties. Remote Sensing, 14(3), 668. https://doi.org/10.3390/rs14030668

1007 Major, J. J. (1997). Depositional Processes in Large-Scale Debris-Flow Experiments. The Journal of Geology, 105(3),

1008 345–366. https://doi.org/10.1086/515930

1009 Major, J. J., & Iverson, R. M. (1999). Debris-flow deposition: Effects of pore-fluid pressure and friction concentrated

1010 at flow margins. Geological Society of America Bulletin, 111(10), 1424–1434. https://doi.org/10.1130/0016-

1011 7606(1999)111

1012 May, C. N., & Gresswell, R. E. (2004). Spatial and temporal patterns of debris-flow deposition in the Oregon Coast

1013 Range, USA. Geomorphology, 57(3–4), 135–149. https://doi.org/10.1016/s0169-555x(03)00086-2

1014 McCoy, S. W., Kean, J. W., Coe, J. A., Tucker, G. S., Staley, D. M., & Wasklewicz, T. A. (2012). Sediment

1015 entrainment by debris flows: In situ measurements from the headwaters of a steep catchment. Journal of Geophysical

1016 Research, 117(F3), n/a. https://doi.org/10.1029/2011jf002278

1017 McDougall, S., & Hungr, O. (2004). A model for the analysis of rapid landslide motion across three-dimensional

1018 terrain. Canadian Geotechnical Journal, 41(6), 1084–1097. https://doi.org/10.1139/t04-052

1019 Medina, V., Hürlimann, M., & Bateman, A. (2008). Application of FLATModel, a 2D finite volume code, to debris

1020 flows in the northeastern part of the Iberian Peninsula. Landslides, 5(1), 127–142. https://doi.org/10.1007/s10346-

1021 007-0102-3

1022 Miller, D., & Burnett, K. M. (2008). A probabilistic model of debris-flow delivery to stream channels, demonstrated

1023 for the Coast Range of Oregon, USA. Geomorphology, 94(1–2), 184–205.

1024 https://doi.org/10.1016/j.geomorph.2007.05.009

1025 Montgomery, D. R., & Dietrich, W. E. (1988). Where do channels begin? Nature, 336(6196), 232–234.

1026 https://doi.org/10.1038/336232a0

1027 Murray, A. S., & Paola, C. (1994). A cellular model of braided rivers. Nature, 371(6492), 54–57.

1028 https://doi.org/10.1038/371054a0

1029 Murray, B.A. (2003) Contrasting the goals, strategies and predictions associated with simplified numerical models

1030 and detailed simulations. In Wilcock, P., & Iverson, R. (Eds.), Prediction in Geomorphology. American Geophysical

1031 Union.

1032 Murray, A. B. (2007). Reducing model complexity for explanation and prediction. Geomorphology, 90(3–4), 178–

1033 191. https://doi.org/10.1016/j.geomorph.2006.10.020

1034 Murray A.B. (2013) Which Models Are Good (Enough), and When?. In: John F. Shroder (ed.) Treatise on

1035 Geomorphology, Volume 2, pp. 50-58. San Diego: Academic Press.

1036 Natural Resources Conservation Service | Snow and Water Interactive Map (n.d.). Natural Resources Conservation

1037 Service. Retrieved April 1, 2022, from

1038 Nicoletti, P. G., & Sorriso-Valvo, M. (1991). Geomorphic controls of the shape and mobility of rock avalanches.

1039 Geological Society of America Bulletin, 103(10), 1365–1373. https://doi.org/10.1130/0016-7606(1991)103

1040   Nudurupati, S. S., Istanbulluoglu, E., Tucker, G. E., Gasparini, N. M., Hobley, D. E. J., Hutton, E. W. H., Barnhart,

1041   K. R., & Adams, J. M. (2023). On transient semi-arid ecosystem dynamics using Landlab: vegetation shifts,

1042   topographic refugia, and response to climate. Water Resources Research, 59(4).

1043   https://doi.org/10.1029/2021wr031179

1044   Open topography, https://opentopography.org/, accessed 2022

1045   Perron, J. T. (2017). Climate and the Pace of Erosional Landscape Evolution. Annual Review of Earth and Planetary

1046   Sciences, 45(1), 561–591. https://doi.org/10.1146/annurev-earth-060614-105405

1047   Reid, M. J., Coe, J. A., & Brien, D. L. (2016). Forecasting inundation from debris flows that grow volumetrically

1048   during travel, with application to the Oregon Coast Range, USA. Geomorphology, 273, 396–411.

1049   https://doi.org/10.1016/j.geomorph.2016.07.039

1050   Renard, B., Garreta, V., & Lang, M. J. (2006). An application of Bayesian analysis and Markov chain Monte Carlo

1051   methods to the estimation of a regional trend in annual maxima. Water Resources Research, 42(12).

1052   https://doi.org/10.1029/2005wr004591

1053   Roda-Boluda, D. C., D'Arcy, M., McDonald, J., & Whittaker, A. C. (2018). Lithological controls on hillslope

1054   sediment supply: insights from landslide activity and grain size distributions. Earth Surface Processes and Landforms,

1055   5), 956–977. https://doi.org/10.1002/esp.4281

1056   Schürch, P., Densmore, A. L., Rosser, N., & McArdell, B. W. (2011). Dynamic controls on erosion and deposition on

1057   debris-flow fans. Geology, 39(9), 827–830. https://doi.org/10.1130/g32103.1

1058   Shen, P., Zhang, L. M., Wong, H., Peng, D., Zhou, S., Zhang, S., & Chen, C. (2020). Debris flow enlargement from

1059   entrainment: A case study for comparison of three entrainment models. Engineering Geology, 270, 105581.

1060   https://doi.org/10.1016/j.enggeo.2020.105581

1061   https://www.nrcs.usda.gov/resources/data-and-reports/snow-and-water-interactive-map

1062   Soil Survey Geographic Database (SSURGO) Ag Data Commons. (n.d.). https://data.nal.usda.gov/dataset/soil-

1063   survey-geographic-database-ssurgo, accessed 2020

1064   Stock, J. P. J., & Dietrich, W. E. (2006). Erosion of steepland valleys by debris flows. Geological Society of America

1065   Bulletin, 118(9–10), 1125–1148. https://doi.org/10.1130/b25902.1

1066   Strauch, R. L., Istanbulluoglu, E., Nudurupati, S. S., Bandaragoda, C., Gasparini, N. M., & Tucker, G. E. (2018). A

1067   hydroclimatological approach to predicting regional landslide probability using Landlab. Earth Surface Dynamics,

1068   6(1), 49–75. https://doi.org/10.5194/esurf-6-49-2018

1069   Takahashi, T. (1978). Mechanical Characteristics of Debris Flow. Journal of the Hydraulics Division, 104(8), 1153–

1070   1169. https://doi.org/10.1061/jyceaj.0005046

1071   Takahashi, T. (2014). Debris Flow (2nd ed.). CRC Press, Taylor & Francis Group.

1072   Tucker, G. E., Hancock, G. J. (2010). Modelling landscape evolution. Earth Surface Processes and Landforms, 35(1),

1073   28–50. https://doi.org/10.1002/esp.1952

1074   Tucker, G. E., McCoy, S., & Hobley, D. E. J. (2018). A lattice grain model of hillslope evolution. Earth Surface

1075   Dynamics, 6(3), 563–582. https://doi.org/10.5194/esurf-6-563-2018

1076   Wallace, C. A., Santi, P. M., & Walton, G. (2022). Scoring system to predict landslide runout in the Pacific Northwest,

1077   USA. Landslides, 19(6), 1449–1461. https://doi.org/10.1007/s10346-021-01839-7

1078   Western Regional Climate Center. (n.d.). Retrieved 2017, from https://wrcc.dri.edu/

1079   Westrick, K. (1999). Soil depth calculation script implemented in python. Retrieved from

1080   https://github.com/pnnl/DHSVM-PNNL

1081   Wu, W., & Sidle, R. C. (1995). A Distributed Slope Stability Model for Steep Forested Basins. Water Resources

1082   Research, 31(8), 2097–2110. https://doi.org/10.1029/95wr01136

1083   Zhou, G. G. D., Li, S., Song, D., Choi, C. E., & Chen, X. (2019). Depositional mechanisms and morphology of debris

1084   flow: physical modelling. Landslides, 16(2), 315–332. https://doi.org/10.1007/s10346-018-1095-9

1085

1086