Abstract  Debris flows pose a significant hazard to communities in mountainous areas, and there is a continued need for methods to delineate hazard zones associated with debris-flow inundation. In certain situations, such as scenarios following wildfire, where there could be an abrupt increase in the likelihood and size of debris flows that necessitates a rapid hazard assessment, the computational demands of inundation models play a role in their utility. The inability to efficiently determine the downstream effects of anticipated debris-flow events remains a critical gap in our ability to understand, mitigate, and assess debris-flow hazards. To better understand the downstream effects of debris flows, we introduce a computationally efficient, reduced-complexity inundation model, which we refer to as the Progressive Debris-Flow routing and inundation model (ProDF). We calibrate ProDF against mapped inundation from five watersheds near Montecito, CA, that produced debris flows shortly after the 2017 Thomas Fire. ProDF reproduced 70% of mapped deposits across a 40 km² study area. While this study focuses on a series of post-wildfire debris flows, ProDF is not limited to simulating debris-flow inundation following wildfire and could be applied to any scenario where it is possible to estimate a debris-flow volume. However, given its ability to reproduce mapped debris-flow deposits downstream of the 2017 Thomas Fire burn scar, and the modest run time associated with a simulation over this 40 km² study area, results suggest ProDF may be particularly promising for post-wildfire hazard assessment applications.

Keywords  Debris flow · Hazard assessment · Inundation · Modeling · Wildfire

Introduction

Debris flows are a common, yet significant, hazard in alpine regions across the world (e.g., Rickenmann and Zimmerman 1993; Jakob and Hungr 2005; Webb et al. 2008; Baum and Godt 2010; Iverson 2014). The recent growth of the wildland-urban interface in mountainous regions across the world, including in the western USA (Radeloff et al. 2005, 2018), has accentuated the threats posed by debris flows. Debris flows present a direct threat to human life (Jakob and Hungr 2005; Dowling and Santi 2014), as well as to downstream infrastructure including houses, roadways, and bridges (e.g., Jakob and Hungr 2005; Kean et al. 2019). Furthermore, debris flows present challenges for water supply management by reducing the capacity of reservoirs (Rathburn et al. 2017) and degrading water quality (e.g., Smith et al. 2011). Given the numerous threats posed by debris flows, there is a continued need for methods to delineate potential hazard zones associated with debris-flow inundation. This is especially true in recently disturbed landscapes where there could be an abrupt increase in the likelihood and size of debris flows, necessitating a rapid hazard assessment.

Post-wildfire scenarios, for example, require rapid hazard assessments to limit the impacts of potential debris flows. Wildfire often promotes increases in debris-flow activity (Wells 1987; Cannon 2001; Gabet and Bookter 2008; Kean et al. 2011; Nyman et al. 2011) by decreasing vegetation cover (e.g., Parson et al. 2010; Hoch et al. 2021) and altering soil hydraulic properties (Doerr et al. 2009; Moody et al. 2015; Ebel and Martin 2017; Ebel and Moody 2017; Chen et al. 2020). Increases in fire frequency (e.g., Gillett et al. 2006; Westerling et al. 2006; Pausas and Fernandez-Munoz 2012; Nyman et al. 2019) result in increased susceptibility to post-wildfire debris flows. While post-wildfire debris-flow initiation and volume have been the focus of many past investigations (e.g., Meyer and Wells 1997; Gabet and Bookter 2008; Nyman et al. 2011; Gartner et al. 2014; Staley et al. 2017), there is a scarcity of research specifically aimed at post-wildfire debris-flow inundation (Barnhart et al. 2021). Further effort is required to develop and test methods to assess the downstream hazards associated with post-wildfire debris flows.

Current assessments of post-wildfire debris-flow hazards often involve estimates of the likelihood and volume of debris flows in response to an idealized rainstorm (Gartner et al. 2008, 2014; Nyman et al. 2011; Staley et al. 2013, 2017). The likelihood of post-wildfire debris-flow hazards are often assessed using rainfall intensity-duration (ID) thresholds (e.g., Nyman et al. 2011; Staley et al. 2017). Intensity-duration thresholds, which are based on historic observations of debris flows, define a rainfall intensity (averaged over a given duration) above which debris flows are likely to initiate. Previous studies have focused on identifying the thresholds required for post-wildfire debris-flow initiation (Cannon et al. 2010; Staley et al. 2013, 2017; Rengers et al. 2019; Tang et al. 2019; Berti et al. 2020) and on understanding the physical processes that lead to initiation (Cannon et al. 2001; Gabet and Bookter 2008; Parise and Cannon 2012; Kean et al. 2013; Langhans et al. 2017; McGuire et al. 2017; Tang et al. 2019; Palucis et al. 2021). Empirical models designed to predict post-wildfire debris-flow volume (Cannon et al. 2010; Gartner et al. 2014; Nyman et al. 2020) provide additional information on the potential for downstream hazards, since debris-flow discharge (Rickenmann 1999) and runout (Rickenmann 1999; de Haas 2015) generally increase with debris-flow volume. Existing models can provide guidance on the likelihood and magnitude of post-wildfire debris flows given information on soil burn severity, watershed morphology, and rainfall intensity (Gartner et al. 2014; Staley et al. 2017), but debris-flow runout modeling is not currently...
a standard component of post-wildfire debris-flow hazard assessments (Kean et al. 2019).

Although numerous process-based debris-flow inundation models currently exist, they are impractical for rapidly assessing debris-flow inundation scenarios where numerous debris flows are anticipated. Two-phase debris-flow models (e.g., Pudasaini 2012; Iverson and George 2014) are designed to represent the complex interactions between the solid and fluid components of a debris flow but are computationally expensive and require a number of parameters that may be poorly constrained, especially during rapid hazard assessments. Single-phase models, meanwhile, require fewer parameters and represent debris flows as idealized fluids (Cesca and D’Agostino 2008; Pudasaini 2011; Gibson et al. 2021) by assuming a constant rheology. Nevertheless, both two-phase and single-phase models are challenging to apply for the purpose of rapid hazard assessments where debris flows have the potential to initiate from numerous source watersheds. These models are more appropriate for modeling the movement of an individual debris flow (Rickenmann et al. 2006; Xu et al. 2020). Empirical models, on the other hand, may be better suited for rapidly modeling geophysical flows due to their computational efficiency (e.g., Iverson et al. 1998; Schilling 1998; Berti and Simoni 2007; Hürlimann et al. 2008).

Reduced complexity models, including empirical models that are based on observational data and do not focus on detailed, physically based processes, have proven capable of reproducing complex geomorphic processes (e.g., Nicholas 2010; Ziliani et al. 2015; Liang et al. 2015), including debris flows. Several studies have utilized established empirical inundation models like DFLOWZ, LAHARZ, and Flow-r, among others, to reproduce areas of observed debris-flow inundation (Berti and Simoni 2007; Magirl et al. 2010; Scheidl and Rickenmann 2010; Horton et al. 2013; Guthrie and Befus 2021). Despite the computational efficiency offered by empirical models to represent runout of geophysical flows, few studies have tested existing empirical models using post-wildfire debris-flow inundation data (e.g., Bessette-Kirton et al. 2019; Youberg and McGuire et al. 2018).

LAHARZ was originally developed to model lahars, and, as such, it assumes that flows are confined within a valley (Iverson et al. 1998; Berti and Simoni 2007). This assumption limits the model’s applicability in scenarios where flows avulse out of the channel. In studies where LAHARZ has been applied to nonvolcanic debris flows, it has produced unrealistically long and narrow inundation patterns compared to observed inundation (Magirl et al. 2010; Bessette-Kirton et al. 2019). Even when nonvolcanic parameters that were later developed for the LAHARZ framework are implemented (Griswold and Iverson 2008), the model underestimates the runout distance of observed inundation extent (Magirl et al. 2010; Bessette-Kirton et al. 2019). This behavior can likely be attributed to LAHARZ’s inability to simulate unconfined flow. This issue is not limited to LAHARZ, as other empirical debris-flow inundation models have struggled to replicate the lateral movement of observed debris flows (e.g., Scheidl and Rickenmann 2010).

To address this issue, DFLOWZ was modified from LAHARZ to better simulate unconfined flows (Berti and Simoni 2007). While DFLOWZ has been shown to reproduce observed debris-flow deposits, this model requires detailed field surveys to locate areas where the flow may avulse out of the channel (Berti and Simoni 2007). This kind of survey is impractical for post-wildfire scenarios where numerous watersheds may be susceptible to debris flows. Furthermore, existing empirical models, including DFLOWZ and Flow-r, do not differentiate between the extent of deposition and the peak-flow depth of a debris flow (Berti and Simoni 2007; Horton et al. 2013). While the extent of deposition is valuable information for hazard assessments, estimates of peak-flow depth provide further insight into the potential for debris-flow impacts to cause structural damage (e.g., Kean et al. 2019).

Because few debris-flow inundation models have been tested and vetted against post-wildfire debris-flow datasets (Barnhart et al. 2021), our ability to predict and mitigate the downstream effects of anticipated debris-flow events following wildfire remains limited. Here we introduce a computationally efficient, reduced-complexity debris-flow inundation model and assess its ability to reproduce mapped inundation zones over a 40 km² area in response to debris-flow initiation in several different source watersheds.

The model is driven by a progressive variation of the multiple flow direction (MFD) routing algorithm (Freeman 1991; Pelletier 2008) that uses a series of empirical equations (Rickenmann 1999) to compute peak debris-flow depth at different points on the landscape based on debris-flow volume, topographic slope, and three parameters related to flow properties. We calibrate the model, which we refer to as the Progressive Debris-Flow routing and inundation model (ProDF), and assess its performance using mapped inundation zones and observed peak-flow depths associated with a series of post-wildfire debris flows that impacted the community of Montecito, CA, following the 2017 Thomas Fire (Kean et al. 2019). We discuss how ProDF can be used in combination with existing empirical debris-flow volume models (e.g., Gartner et al. 2014) to estimate post-wildfire inundation potential given soil burn severity, rainfall intensity, and topography. While we focus on a series of post-wildfire debris flows in this study, ProDF can be applied to any debris-flow inundation scenario where debris-flow volume can be estimated.

**Study area**

The community of Montecito is located in southern CA, USA, at the foot of the Santa Ynez Mountains (Fig. 1). The community occupies a 3 km wide piedmont plain, bordered by the mountains to the north and the Pacific Ocean to the south. The northern portion of the piedmont plain contains steep alluvial fans, while the southern portion contains more gently sloping alluvial fans. The northern and southern fans are separated by a topographic high resulting from the Mission Ridge Fault zone (Kean et al. 2019). The alluvial fans are heavily populated, containing a high density of structures and roads, including US 101, a major highway. Alluvial channels, originating in the Santa Ynez Mountains, traverse this urbanized region and terminate in the Pacific Ocean. The source areas for these drainages are underlain by Eocene sedimentary rocks (Juncal Formation, Matilija Formation, Cozy Dell Shale, and Coldwater Sandstone) (Dibblee 1966). Montecito has a Mediterranean climate that is characteristic of coastal southern California.

The Thomas Fire, which ignited in southern California on 4 December 2017, burned above Montecito, from the ridge of the Santa Ynez Mountains onto populated alluvial fans. Approximately 80% of the source areas for five major drainages, Montecito, Oak, San Ysidro, Buena Vista, and Romero Creeks, were
burned at moderate or high soil burn severity (Table 1) (Kean et al. 2019). On 9 January 2018, before the Thomas Fire was fully contained, intense rainfall above Montecito triggered debris flows in these five watersheds. Debris flows initiated in the burned catchments within the Santa Ynez Mountains before traversing the urbanized piedmont plain and emptying into the Pacific Ocean (Fig. 1). The debris flows, which ranged from 10,000 to 297,000 m$^3$ in volume (Table 1), killed 23 people, injured 180 more, and damaged over 400 homes (Kean et al. 2019). The deposits exhibited a bimodal grain-size distribution, with boulder-sized clasts originating from the Matilija Formation and Coldwater Sandstone and the finer-grained matrix consisting of the Juncal Formation and Cozy Dell Shale (Kean et al. 2019) (Fig. 2).

**Methodology**

**Model description**

ProDF employs an iterative variation of the multiple flow direction (MFD) algorithm (Freeman 1991; Pelletier 2008) to route debris flows downslope given user-specified starting points and volumes for each flow. More specifically, ProDF routes debris-flow volume downstream from the starting points and uses empirical relationships to compute flow depth based on debris-flow volume and topographic attributes (e.g., slope). Flow volume stops being routed downstream once the shear stress at the base of the flow, as determined by flow depth and local topographic slope, fails to exceed a user-defined yield strength. ProDF can therefore be used to determine the extent of debris-flow inundation and compute peak-flow depth at every point along the flow travel path, but it does not provide an estimate of sediment deposited along the travel path. The MFD algorithm routes debris-flow volume from a central cell to all downslope neighbors, weighted by slope. The iterative variation of the MFD algorithm, originally introduced by Pelletier (2008) for simulating flood flows on alluvial fans, and described in detail below, leads to a more accurate representation of lateral movement of the flow.

The basic concept behind the iterative variation of the MFD algorithm is to first route flow based on the topography, as is commonly done, to obtain a prediction for peak-flow depth across the existing topographic surface. Then, the flow is routed again over a new flow surface defined by the original topographic surface plus a small fraction of the predicted flow depth. This process is repeated for a user-specified number of iterations,

### Table 1 Characteristics of five watersheds that produced debris flows following the 2017 Thomas Fire (Kean et al. 2019)

| Drainage       | Area (km$^2$)$^a$ | Area burned at moderate/high severity (%)$^b$ | Observed volume (m$^3$)$^a$ |
|----------------|-------------------|---------------------------------------------|-----------------------------|
| Cold Spring Creek$^b$ | 8.94             | 79                                          | 179,000$^b$                |
| Hot Springs Creek$^b$      | 1.73             | 83                                          | 52,000$^b$                 |
| Oak Creek              | 0.45             | 49                                          | 10,000                     |
| San Ysidro Creek       | 7.63             | 85                                          | 297,000                    |
| Buena Vista Creek      | 1.65             | 82                                          | 41,000                     |
| Romero Creek           | 4.43             | 78                                          | 100,000                    |

$^a$Information from Kean et al. (2019)

$^b$Cold Spring and Hot Springs Creek merge to form Montecito Creek, which had a total debris-flow sediment volume of 231,000 m$^3$
allowing the flow to spread across a landscape in much the same way that a gradually rising flood wave would first fill channels and then overflow onto surrounding areas. We first describe how the model routes flow during a single iteration of this process and then describe the details of the iterative flow routing procedure.

To begin, the MFD algorithm routes debris-flow volume \( M \) [\( m^3 \)] downstream from the flow starting points, each defined by a single grid cell. When the debris-flow volume is routed from one cell to the next, the peak discharge \( Q \) [\( m^3/s \)] is computed using an empirical relationship developed by Rickenmann (1999):

\[
Q = c_1 M^{c_2}
\]

where \( c_1 = 0.135 \) [\( 1/s \)] and \( c_2 = 0.78 \) [-] are empirical constants. Discharge can then be related to flow velocity \( (V) \) [\( m/s \)], flow depth \( (h) \) [\( m \)], and flow width \( (w) \) [\( m \)] by utilizing the relationship:

\[
Q = wVh
\]

Here we set \( w \) to the resolution of the DEM (\( \Delta x \)).

We tested two equations proposed by Rickenmann (1999) that relate debris-flow velocity to flow depth and topographic slope \( (S) \) (Eqs. 10 and 11 from Rickenmann (1999)):

\[
V = \frac{\rho g h^2 S}{3\mu}
\]

\[
V = \frac{2\chi h^{1.5} S}{3}
\]

Here, \( g = 9.81 \) is gravitational acceleration [\( m/s^2 \)] and \( \rho, \mu, \) and \( \chi \) are user-defined variables that represent flow density [\( kg/m^3 \)], the dynamic viscosity of the flow [\( kg/sm \)], and the flow resistance coefficient [\( 1/(sm^{1/2}) \)] as defined by Rickenmann (1999), respectively. In the “Model calibration” section, we discuss a method for calibrating these parameters.

Both Eqs. 3 and 4 have been proposed to estimate the mean velocity of debris flows, but each equation assumes a different flow rheology. Equation 3 is proposed for Newtonian laminar flows but approximates Bingham fluids with high velocities and low Bingham yield stresses (Rickenmann 1999), while Eq. 4 represents flows with a Bagnold rheology and dilatant grain shearing in the inertial regime (Rickenmann 1999). We tested both equations to determine which relationship was more applicable for the post-wildfire debris flows at our site.

By employing the relationships introduced above, it was possible to solve for flow depth:

\[
h = \left( \frac{3Q\mu}{w\rho g S} \right)^{1/3}
\]

\[
h = \left( \frac{3Q}{2w\chi S} \right)^{2/3}
\]

where Eqs. 5 and 6 are derived using Eqs. 3 and 4, respectively. Using this method, flow depth can be computed for every grid cell in the computational domain as flow volume is being routed. Flow depth can be written more explicitly as a function of debris-flow volume using Eq. 1, which yields:

\[
h = \left( \frac{3\mu c_1 M^{c_2}}{w\rho g S} \right)^{1/3}
\]

\[
h = \left( \frac{3c_1 M^{c_2}}{2w\chi S} \right)^{2/3}
\]

The proportion of volume routed through each grid cell, downstream of the starting point, is governed by the MFD algorithm, and the volume in each downstream cell changes as the flow spreads out according to the MFD algorithm. For example, if the flow starts in a single cell, 100% of the volume will be routed through that cell, regardless of slope (Fig. 3a). However, if the MFD algorithm routes the flow into two downstream cells, the flow will be routed to both downstream grid cells, weighted by the slope of each cell. In a scenario where the slope of the surface is consistent across both grid cells, 50% of the input volume is routed through each cell (Fig. 3b). Similarly, as flow moves downstream into three cells...
with a consistent slope across the width, approximately 33% of the flow volume would be routed through each grid cell (Fig. 3c). The same phenomenon occurs when a flow begins to move laterally. The proportion of the total flow volume routed through each cell will decrease as the flow spreads laterally (Fig. 3b and c). As flow volume disperses across the landscape, the discharge at each cell adjusts accordingly such that all of the volume is routed from one grid cell to a single downstream neighbor. The routing surface is updated following each iteration. By iteration 10, the routing surface has changed such that flow volume is partitioned into additional cells as it is routed downstream. This is analogous to flow spreading laterally as it spills over a channel bank. By iteration 50, the routing surface has changed sufficiently that the flow volume gets partitioned into an even greater number of cells as it moves downstream. Since the MFD algorithm will proportion the volume routed to each cell based on slope, this conceptual example assumes that slope is consistent across all downstream cells.

\[ \tau = \rho ghS \]  \hspace{1cm} (9)

ProDF only routes flow volume at a grid cell if:

\[ \tau > \tau_y \]  \hspace{1cm} (10)

If \( \tau \) is less than \( \tau_y \), which we treat as a calibrated parameter (see “Model calibration” section), at a given grid cell, then the flow volume in that cell stops being routed at that point (Fig. 3). However, the flow may continue being routed elsewhere, at points where \( \tau > \tau_y \) (Fig. 3). The flow volume has been completely routed when \( \tau < \tau_y \) at every point along the flow front (Fig. 3d).

To compare \( \tau \) and \( \tau_y \) at each grid cell, we calculate the slope of a given cell over two spatial scales. We first calculate a slope at the scale of the DEM resolution (\( \Delta x \)), which represents the slope between the current grid cell and cells immediately downslope. We also calculate a slope at the scale of three times the resolution of the DEM (3\( \Delta x \)). This value reflects the average slope of the surface over the three grid cells downslope of the original point. We use the
maximum of these two slope estimates when computing $\tau$. Computing the average slope over a distance that is greater than the resolution of the DEM helps prevent flows from stopping in cases where there may be a small-scale undulation due to a locally flat region (e.g., a narrow roadway) followed by substantially steeper slopes immediately downstream.

After initially routing the entire flow volume using the MFD algorithm to determine an estimate of flow extent and flow depth, ProDF begins the iterative routing procedure. This iterative procedure is necessary to refine estimates of peak-flow depth and inundation extent because traditional flow routing algorithms, including the MFD algorithm, route flow based solely on topographic slope. As a result, they are best at determining flow pathways assuming low flow conditions, an assumption that is likely violated in many debris-flow inundation scenarios. The traditional algorithms do not accurately route flow in situations where the depth of the flow exceeds the depth of the channel (Pelletier 2008). It would be preferable to route the flow based on the slope of the flow surface, but this is not known a priori. We address this issue by applying an iterative variation of the MFD algorithm that leads to progressive inundation in different portions of the landscape.

In the first iteration, we use the MFD algorithm to route flow volume over the original topography ($z$) and compute flow depth at each location traversed by the flow as described above. Prior to the start of each successive iteration, the model defines a new flow-routing surface ($E$), which is the original topography ($z$) plus a small portion of the flow depth ($d$) as estimated during the previous iteration (Fig. 4). Thus, the flow-routing surface represents the current model estimate for the surface of the flow rather than the topography. In particular, $E$ is defined according to:

$$E_{ij}^{k+1} = E_{ij}^k + d_{ij}^k$$  \hspace{1cm} (11)

where the subscripts $i$ and $j$ denote spatial location within the model domain, $k$ denotes the iteration number, and

$$d_{ij} = \begin{cases} \frac{h_{ij}^k}{\Delta x} & h_{ij}^k + z_{ij} > E_{ij}^k \\ 0 & \text{else} \end{cases}$$  \hspace{1cm} (12)

with $N$ denoting a prescribed number of iterations and $h_{ij}^k$ the flow depth computed during iteration $k$. At the kth iteration, ProDF routes the flow volume across the landscape using $E^k$ as the input topography and calculates new flow depths and a new inundation extent. The number of iterations ($N$) is prescribed at the start of a simulation. The model solution, a spatially variable prediction of peak-flow depth, will converge (i.e., there will be little/no change in predicted flow depth between consecutive iterations) after a certain number of iterations. To determine how the value of $N$ impacted model performance, we ran simulations using a range of iterations from 1 to 10,000. We found that 100 iterations adequately captured the lateral movement of the flow. Additional iterations led to only minor differences in flow extent and increased model run times. As such, all numerical experiments reported here use $N=100$ iterations.

At the end of the last iteration, the final estimate for flow depth is $h_{ij} = E_{ij}^N - z_{ij}$. By employing this iterative routing procedure,

![Fig. 4](image)

**Fig. 4** a Output from a model simulation at different iterations demonstrates how lateral spreading is facilitated by the iterative flow routing procedure. Model output is shown at a cross-section of San Ysidro Creek (location marked in panel b where the debris flow avulsed out of the main channel into a distributary. b A planar view of progressive inundation around San Ysidro Creek. Although the inundation extent at 100 iterations represents the final output of ProDF, model outputs at previous iterations are included to illustrate how the landscape is progressively inundated.
the model is able to represent the lateral movement of the flow that occurs when flow depths exceed bankfull capacity. Since Eq. 1 yields a prediction of peak debris-flow discharge given sediment volume, we interpret the final output of the model \(h_{f,i}\) as an estimate of peak debris-flow depth for the entire flow event at each cell in the domain (Fig. 5a). The model does not compute any estimate of deposit thickness.

The process of progressive flow routing employed here is illustrated in Fig. 4. Mapped deposits along San Ysidro Creek (Fig. 1) show that some of the debris flow avulsed out of the main channel and into a distributary channel immediately downstream of the mountain front (western channel in Fig. 4b). Figure 4a shows a cross-section of San Ysidro Creek at this location. During the first iteration, \(E\) is equivalent to the topography \(z\) shown in Fig. 4a. However, as the simulation progresses, \(E\) is updated using Eq. 12. At the end of each iteration, a fraction of the computed flow depth is added to the current iteration of \(E\) at every cell that experiences flow. The depth of the flow that is added to \(E\) is scaled based on the number of iterations prescribed at the start of the simulation. In this study, we prescribed \(N = 100\) iterations, meaning \(1/100\) of the most recent estimate of flow depth is added to \(E\) at the end of each iteration. Figure 4a shows the progression of \(E\) throughout the model simulation.

During the first 50 iterations, flow is restricted to the main channel of San Ysidro Creek (Fig. 4). In the later iterations of the simulation, however, the surface described by \(E\) indicates that flow overtopped the main channel and spread laterally. This illustrates how the model responds when the depth of a debris flow exceeds the depth of the channel. The flow-routing surface at the final iteration (100) represents the peak-flow depth predicted by ProDF (Fig. 4b). Notice that the model converges towards a solution, with less change in flow depth between iterations 75–90 relative to the amount of change between iterations 10–25.

![Fig. 5](image)

**Fig. 5** a ProDF outputs a map of modeled peak-flow depth at each cell, which can be used to determine the extent of inundation. b A comparison between the mapped deposits and the output of ProDF using the globally calibrated best-fit parameters. Areas where the model overlaps with the observations are shown as true positive (TP). Areas where the model did not show flow, but sediment was observed in the field represent a model underestimation or false negative (FN). Areas where the model predicted flow, but there were no mapped deposits is shown as overestimation or false positive (FP). A large region of model underestimation caused by the presence of a US 101 bridge is highlighted by the red circle.

### Input parameters

ProDF requires several inputs, including a digital elevation model (DEM), a debris-flow sediment volume, and a user-defined debris-flow starting point. For the purpose of this study, we defined topography in the Montecito study area using a 3-m, pre-event DEM derived from airborne lidar. However, we also explored the effect of cell size with 5-m and 10-m DEMs that were resampled from the original DEM. We hydrologically conditioned the input DEM using a fill algorithm to remove all local depressions. We set the volume for each of the five modeled debris flows based on field estimates of sediment that was deposited in each of the five watersheds (Table 1) (Kean et al. 2019). The initiation points for the Montecito debris flows are unknown, but the precise initiation location does not affect downstream inundation patterns in the model. Therefore, the “starting” point for each debris flow, which is defined as a single grid cell, can be located anywhere upstream of the watershed outlets where the January 2018 debris flows lost confinement and inundated alluvial fans. We selected starting points at the upstream extent of our input DEM, near the area where the flows first exited the mountain front (see starting locations in Fig. 4). Since we define the starting point as a single cell, the discharge at this point will be unrealistically high based on Eq. 1. However, within a short distance downstream of the starting point, the discharge returns to realistic values.

In addition to the three inputs discussed above, ProDF also requires user-defined values for flow density \(\rho\), yield strength \(\tau_y\), and a flow mobility coefficient, either a viscosity \(\mu\) or a flow resistance coefficient \(\chi\) as introduced by Rickenmann (1999). We constrained the range of values for each parameter using results from previous studies. Iverson (1997) found that \(\rho\) rarely falls outside the range of 1800–2300 kg/m³ for natural debris flows. Therefore, \(\rho\) was constrained to that range. Reported values of \(\tau_y\) in debris flows vary...
considerably throughout the literature, with some studies finding $\tau_y$ to be as low as 12 Pa (Major and Pierson 1992), and others finding $\tau_y$ to be as high as 6000 Pa (Whipple and Dunne 1992). Rickenmann (1999) back calculated $\mu$ and $\chi$ as a function of peak discharge. Using a semi-theoretical relationship, Rickenmann (1999) found that values for $\mu$ and $\chi$ range between $10^{-6}$ to $10^{3}$ kg/sm and $10^{-10}$ to $10^{-6}$ m/(sm$^{1/2}$), respectively. O’Brien and Julien (1988) found that $\mu$ and $\tau_y$ for mudflows increased exponentially with sediment concentration, which may explain the wide range of values for these parameters. After constraining the range of values for $\tau_y$, $\mu$, and $\chi$ based on the above considerations, we were able to further narrow the range of plausible values at our study area based on the results of initial simulations. For this study, we therefore explored $\tau_y$ values between 10 and 500 Pa, $\mu$ values between 50 and 1200 kg/sm, and $\chi$ values between 10 and 60 m/(sm$^{1/2}$).

**Model calibration**

We calibrated ProDF input parameters by minimizing the misfit between modeled debris-flow inundation and maps of observed debris-flow inundation associated with the 2018 Montecito debris-flow event (Kean et al. 2019). In particular, we performed a series of numerical experiments to model inundation downstream of five watersheds (Montecito Creek, Oak Creek, San Ysidro Creek, Buena Vista Creek, and Romero Creek) that produced impactful debris flows (Fig. 1). The goal of the calibration was to assess the ability of the model to reproduce the mapped inundation pattern using two different formulations, one where flow depth is computed with Eq. 5 and one where flow depth is computed with Eq. 6, as well as to determine best-fit values for flow mobility parameters ($\rho$, $\chi$, $\mu$, and $\tau_y$). We assessed spatial variations in the best-fit parameters throughout our study area by determining the best-fit parameters for each of the five different debris-flow-producing watersheds, which we refer to as local calibrations. We also determined a separate set of parameters that produced the best overall fit when simulating debris-flow inundation across all five watersheds at once, which we refer to as a global calibration. Finally, after performing both local and global calibrations, we ran a series of numerical experiments to assess how well a given set of locally calibrated parameters would perform when employed to model debris-flow inundation from the remaining four watersheds.

To start the calibration process, we employed the Latin Hypercube Sampling method (McKay et al. 1979) to generate a list of unique parameter combinations. In total, we generated 6000 unique combinations: 3000 combinations of parameters for Eq. 5 ($\rho$, $\tau_y$, and $\mu$) and 3000 combinations of parameters for Eq. 6 ($\rho$, $\tau_y$, and $\chi$). For the global calibration, we performed 6000 simulations, one for each set of parameters, to model inundation downstream of all five watersheds simultaneously. Then, using the same 6000 parameter combinations, we simulated debris-flow inundation downstream of each individual watershed to determine how the best-fit parameters varied from watershed-to-watershed (i.e., local calibrations). In all cases, we saved the final inundation output from the model to compare against mapped inundation patterns. In order to limit our model observations to areas where flow depths pose a potential hazard, we only consider a grid cell to have been inundated by the flow if the modeled peak-flow depth exceeded 10 cm.

We assessed model performance by comparing modeled inundation extent with the extent of mapped inundation (Kean et al. 2019) using an evaluation metric introduced by Heiser et al. (2017) that we refer to here as the similarity index ($\Omega$). The similarity index is an evaluation concept that accounts for model overestimation (cells where there is modeled inundation, but no mapped inundation; false positive, FP), underestimation (cells where there is mapped inundation, but no modeled inundation; false negative, FN), and overlap (cells where there is both modeled and mapped inundation; true positive, TP) when compared to mapped inundation. The similarity index is defined as

$$\Omega = \alpha_T - \beta_T - \gamma_T$$

Here, $\alpha_T$ (TP/(TP + FP + FN)) represents the area of overlap between simulated and mapped inundation, while $\beta_T$ (FN/(TP + FP + FN)) and $\gamma_T$ (FP/(TP + FP + FN)) represent areas of model underestimation and overestimation, respectively. The value of $\Omega$ varies between a fixed range of 1 (total fit) and −1 (total misfit), regardless of the size of the model domain. A value of 0 indicates a scenario where the number of modeled points that match the mapped points is equal to the number of modeled points that do not match the mapped points. Therefore, scenarios that produce a similarity index greater than 0 are those in which modeled inundation better reproduces mapped inundation. We used the similarity index as an objective function to determine the parameter sets that provided the best fit between modeled and mapped inundation.

To quantify the extent of model underestimation and overestimation, we isolated the values of $\beta_T$ and $\gamma_T$ for mapped and modeled points with $>10$ cm flow depth. We assessed model underestimation using $\beta_T$. The value of $\beta_T$ can vary from 0 to 1, where 0 indicates no model underestimation and 1 indicates complete underestimation (no modeled flow). We used $\gamma_T$ to assess model overestimation. The value of $\gamma_T$ is always positive and less than 1, where 0 indicates no model overestimation and a higher value of $\gamma_T$ indicates a greater degree of overestimation (e.g., a scenario where there is no overlap between modeled and mapped deposits). These evaluation metrics provide additional insight into model behavior and are best considered in conjunction with one another.

**Peak-flow depths**

While we used the extent of inundation and corresponding similarity index to determine best-fit parameters, we also assessed the performance of the calibrated model by comparing modeled and observed peak-flow depths. In the days following the Montecito debris flows, Kean et al. (2019) made measurements of peak-flow depth based on the height of flow lines on buildings. A total of 534 of these measurements fall within the extent of our model domain. Of these 534 measurements, 317 were taken at points where ProDF modeled flow using the globally calibrated best-fit parameters. Because we already considered how well the model reproduced the mapped extent of inundation using the similarity index, we only compared peak-flow depth at these 317 points. We conservatively estimated the uncertainty associated with measuring peak-flow depth to be plus or minus 10 cm of the reported value from Kean et al. (2019) because field measurements varied due to uncertainty in flow splash marks. Therefore, we considered simulated peak-flow
depth to be underestimated if it underpredicted the observed peak-flow depth by more than 10 cm and overestimated if it overpredicted observed peak-flow depth by more than 10 cm. We also noted all points at which the modeled peak-flow depth was within 10 cm of the observed value.

Model sensitivity

We conducted two sensitivity analyses. The first was aimed at exploring the extent to which model performance, as measured by the similarity index, was affected by changes to the three model parameters related to flow mobility ($\rho$, $\chi$, $\tau_r$). We explored relationships between $\rho$, $\chi$, $\tau_r$, and the similarity index using the 3000 simulations associated with the global calibration using Eq. 6. Understanding sensitivity to these three input parameters will help inform future studies by providing guidance on which of these parameters, either calibrated or estimated from previous studies, are most important to constrain.

Second, we assessed model sensitivity to flow volume, $\rho$, $\chi$, and $\tau_r$, to explore how uncertainty in flow volume and mobility parameters influences forward predictions, specifically the total area inundated. Assessing sensitivity to flow volume is particularly important in this context given the uncertainty associated with estimating debris-flow volume prior to an event (e.g., Gartner et al. 2014). For this analysis, we ran numerical experiments on a single watershed, San Ysidro Creek, while varying flow volumes from 148,500 m$^3$, or 50% of the observed flow volume, to 594,000 m$^3$, or 200% of the observed flow volume. Values for the flow mobility parameters were restricted to be between the minimum and maximum values determined through the local calibrations at each of the five watersheds. We used a Latin Hypercube Sampling strategy to select 3000 unique combinations of flow volume, $\rho$, $\chi$, and $\tau_r$ within these ranges.

Comparison with existing models

To assess how ProDF performed relative to two established inundation models, LAHARZ and FLO-2D, we compared the best-fit simulations from ProDF with those presented by Bessette-Kirton et al. (2019). Bessette-Kirton et al. (2019) tested LAHARZ and FLO-2D using the mapped inundation within the San Ysidro Creek watershed. We ran simulations of ProDF on the San Ysidro model domain using the same DEM resolution (5 m) and input volume as Bessette-Kirton et al. (2019) to ensure a systemic comparison between the two studies. We compared the performance of the three models, LAHARZ, FLO-2D, and ProDF, by utilizing the similarity index ($\Omega_r$), $\beta_r$, and $\gamma_r$. We did not run any simulations with LAHARZ or FLO-2D in this study, instead relying on the results provided in Bessette-Kirton et al. (2019).

LAHARZ and FLO-2D are established models that have been used to simulate debris-flow inundation in both burned (Bessette-Kirton et al. 2019; Youberg and McGuire et al. 2018) and unburned (Rickenmann et al. 2006; Magirl et al. 2010) scenarios. LAHARZ is an empirical model that was originally developed to model lahars (Iverson et al. 1998; Schilling 1998), although subsequent work has provided model coefficients designed to represent nonvolcanic debris flows (Griswold and Iverson 2008; Magirl et al. 2010). Bessette-Kirton et al. (2019) employed empirical equations for both lahars and nonvolcanic debris flows, resulting in two separate LAHARZ simulations. FLO-2D is a process-based, two-dimensional finite difference model that was developed to predict areas of inundation for clear-water floods and debris flows (O’Brien et al. 1993).

Results

Flow resistance law, DEM resolution, and number of flow routing iterations

We assessed model performance using two different flow resistance laws (Eqs. 5 and 6), three different DEM resolutions (3 m, 5 m, 10 m), and a range of total iterations of the progressive flow routing algorithm. While assessing how the model performed using different flow resistance laws, we found that the similarity index associated with the best-fit parameters for the global calibration (using the entire five-watershed dataset) and three of the five local calibrations (Oak, San Ysidro, and Romero) was greater when using Eq. 6 relative to Eq. 5 (Online Resource 1; Table 2). However, the similarity index values for each of the five individual watersheds, as well as for the global dataset, fell within 10% of each other. Despite the similar performance between flow resistance laws, Eq. 6 outperformed Eq. 5.

We found that a DEM with a grid spacing of 5 m led to improved model performance relative to grid spacings of 3 m or 10 m. ProDF produced similar results when run on 3 m or 5 m input topography, generating similarity index values of 0.02 and 0.04, respectively, when using the best-fit parameters from the global calibration (Table 3). However, as expected, run times were shorter on the 5-m DEM relative to the 3-m DEM. Model run times were shortest when using a 10-m DEM, although model accuracy suffered from the coarser resolution (similarity index of ~0.06 for all watersheds). $\beta_r$ values decreased with coarsening resolution, while $\gamma_r$ values increased (Table 3). This indicates that ProDF is more likely to overestimate the extent of mapped inundation when using lower-resolution input topography.

Employing 100 iterations ($N$) of the progressive flow routing algorithm provided the best combination of model accuracy and efficiency. Lower values of $N$ (e.g., 5, 10, 20) resulted in widespread model underestimation, particularly in areas where the flow overtopped the main channel and spread laterally (Online Resource

| Watershed     | Equation similarity index | Equation similarity index |
|---------------|---------------------------|---------------------------|
| Montecito     | 0.22                      | 0.18                      |
| Oak           | −0.05                     | 0.01                      |
| San Ysidro    | 0.24                      | 0.25                      |
| Buena Vista   | −0.23                     | −0.23                     |
| Romero        | 0.02                      | 0.04                      |
| All watersheds| 0.00                      | 0.04                      |
Table 3 Best evaluation metric values obtained using different DEM resolutions. Runtimes were determined on a machine with an AMD Ryzen 7 4800H processor and 16 GB memory using a single thread.

| DEM resolution (m) | Similarity index | ρ_T | γ_T | Model runtime for entire study area (s) |
|-------------------|------------------|-----|-----|----------------------------------------|
| 3                 | 0.02             | 0.24| 0.25| 122                                    |
| 5                 | 0.04             | 0.22| 0.26| 45                                     |
| 10                | −0.06            | 0.20| 0.33| 14                                     |

2). The similarity index of these simulations, which was less than zero in all cases where N < 25, reflected the poor fit. As N increased, flow continued to expand laterally from areas of initial flow concentration, flowing over parts of the landscape where inundation extent was previously underestimated by the model. We found that N = 100 resulted in the best-fit for the entire five-watershed dataset with a similarity index of 0.04. Increasing N beyond 100 resulted in increased runtime, little to no change in inundation extent between subsequent iterations, and had a negligible impact on the similarity index (Online Resource 2). In the “Local and global calibration” to “Sensitivity analysis” sections, we focus on results with a model setup using Eq. 6, a 5 m input DEM, and 100 iterations of the progressive flow routing algorithm since these choices optimized model performance.

Local and global calibration

We calibrated ProDF at each of the five individual watersheds (local calibration) using 3000 combinations of density (ρ), the flow resistance coefficient (χ), and yield strength (τ_y). The locally calibrated best-fit parameters for each watershed were distinct from one another (Table 4). Calibrated values of ρ for the individual debris flows ranged from 1830 to 2277 kg/m³. Best-fit τ_y values varied between 95 and 496 Pa, while χ ranged between 16.9 and 49.6 1/(sm1/2). We also calibrated ProDF to the entire five-watershed dataset (global calibration). The set of globally calibrated best-fit parameters was unique from the sets of local parameters calibrated for each watershed. However, the global best-fit values of ρ = 2253 kg/m³, τ_y = 296 Pa, and χ = 18.5 1/(sm1/2) fell within the range of locally calibrated parameters listed in Table 4.

We assessed model performance by comparing the extent of modeled inundation against mapped inundation in each of the five individual watersheds. Using the locally calibrated values of ρ, χ, and τ_y (Table 4), ProDF yielded similarity index values greater than zero (i.e., these simulations reproduced the mapped deposit extent more often than not) for Montecito, Oak, San Ysidro, and Romero Creeks (Fig. 6). Buena Vista Creek was the only watershed that had a best-fit similarity index less than zero (−0.23). The γ_T value for Buena Vista was the largest of any of the individual watersheds, indicating a high degree of model overestimation compared to the mapped extent of inundation (Fig. 6). Despite the lower similarity index, the model still reproduced 59% of the inundation extent in Buena Vista. However, this was also the lowest percentage among the five watersheds, as the best-fit simulations resulted in overlaps in the other four watersheds of 66% (Montecito), 68% (Oak), 69% (Romero), and 81% (San Ysidro).

We also ran numerical experiments with a model domain consisting of the entire five-watershed dataset. The modeled extent of inundation overlapped 70% of the mapped extent of inundation when using the globally calibrated flow mobility parameters (Fig. 5b). The best-fit similarity index associated with this global calibration was 0.04, indicating that the modeled inundation matched the majority of observed inundation.

To assess the transferability of model parameters calibrated for one watershed to other nearby watersheds, we ran a series of simulations where local best-fit parameters from one watershed were applied to simulate inundation at the remaining four watersheds. For example, we applied the local best-fit parameters calibrated for Montecito Creek to the remaining four watersheds (Oak, San Ysidro, Buena Vista, and Romero) (Online Resource 3), which resulted in a similarity index of −0.09 for the four-watershed dataset. Depending on which watershed was chosen as the calibration watershed, these simulations yielded similarity index values between −0.05 and −0.14 (Online Resource 4). These similarity index values are lower than that associated with the globally calibrated parameters (similarity index of 0.04). However, depending on which watershed was chosen as the calibration watershed and which watersheds were chosen as the test watersheds in these simulations, ProDF reproduced between 47 and 77% of the mapped debris-flow deposits in the test watersheds.

Peak-flow depths

Using the globally calibrated best-fit parameters, the correlation coefficient between the observed and modeled peak-flow depths was 0.35 (Fig. 7a). The difference between the modeled and observed peak-flow depth was less than 1 m at 256 points (81%), less than 0.5 m at 172 points (54%), and was within 0.2 m at 96 points (30%). Qualitatively, ProDF generally reproduced patterns of peak-flow depth observed in the field (Fig. 7b). The midpoint of Montecito Creek and the upper reaches of San Ysidro Creek, which were the locations of the deepest observed peak-flow depths (> 2 m), also had some of the highest modeled peak-flow depths (> 2 m). Furthermore, watersheds with lower observed peak-flow depths (< 0.3 m), such as Buena Vista and Oak Creeks, had some of the lowest modeled peak-flow depths (< 0.3 m) (Fig. 7b).
Sensitivity analysis

Impact of flow mobility parameters on similarity index

There is no apparent relationship between similarity index and \( \rho \) (Online Resource 5a). However, there is a relationship between the similarity index and the parameters \( \chi \) and \( \gamma_i \) (Online Resource 5b and c). This indicated that ProDF was sensitive to both of these parameters and prompted further investigation into the relationship between \( \chi, \gamma_i \), and model behavior. Results suggest a tradeoff between \( \chi \) and \( \gamma_i \), where the effect of increasing \( \chi \) on the similarity index can be partially offset in some cases by decreasing \( \gamma_i \). This relationship was observed in each of the five locally calibrated datasets (Fig. 8a–e), as well as in the globally calibrated dataset (Fig. 8f).

Impact of volume and flow mobility parameters on inundated area

To better assess model applicability in predictive scenarios, we explored how uncertainty in debris-flow sediment volume and mobility parameters influenced the area inundated by ProDF. We observed a clear positive relationship between debris-flow volume and area inundated (Fig. 9a). However, there was a substantial amount of scatter, suggesting that flow mobility parameters also affected area inundated. Flow density, \( \rho \), had minimal influence on area inundated compared to flow volume (Fig. 9b). However, \( \chi \) and \( \gamma_i \) both influenced area inundated in more meaningful ways. All else being equal, lower values of \( \chi \) (Fig. 9c) and lower values of \( \gamma_i \) led to greater area inundated (Fig. 9d).

**Fig. 6** Comparisons between the locally calibrated, best-fit modeled deposits and the mapped extent of inundation for **a** Montecito, **b** Oak, **c** San Ysidro, **d** Buena Vista, and **e** Romero Creeks. The inset tables show our objective function results. \( \Omega_T \) represents the similarity index, \( \beta_T \) represents areas of underestimation, and \( \gamma_T \) represents areas of overestimation. Runtimes were determined on a machine with an AMD Ryzen 7 4800H processor and 16 GB memory using a single thread.
An efficient method for delineating the potential hazard zones associated with debris-flow inundation is necessary in recently disturbed landscapes where there could be an abrupt increase in debris-flow likelihood and magnitude. Results here show that ProDF is capable of reproducing mapped inundation over a 40 km² study area with modest run times, which increases its utility in cases where a number of watersheds may be susceptible to debris flows. Since ProDF also ingests debris-flow volume, which is an output of existing models used for post-wildfire hazard assessments (e.g., Gartner et al., 2014), it has the potential to augment existing post-wildfire hazard assessment frameworks by providing estimates of debris-flow inundation.

Model performance

While ProDF performed well in a complex inundation scenario that included roadways, houses, culverts, bridges, and other infrastructure, model performance varied from watershed to watershed. Differences in model performance among the five watersheds can be partially attributed to the model’s assumption that debris flows have a rheology that is fixed in space and time. Debris-flow grain size, sediment concentration, and pore-fluid pressure vary in space and time in natural settings (e.g., Iverson 2003). We expect these differences to be broadly reflected by differences among the calibrated best-fit parameters for each of the five watersheds. Using the best-fit parameters derived from the global calibration, ProDF performed well at Montecito and San

Fig. 7 We compared the modeled peak-flow depth using the best-fit parameters from the global calibration to the observed peak-flow depth at 317 points. The correlation coefficient between modeled and observed peak-flow depth was 0.35 a. The pattern of observed peak-flow depth generally followed the pattern of modeled peak-flow depth. The areas of deepest observed peak-flow depth overlay the areas of deepest modeled peak-flow depth b. The midpoint of Montecito Creek and the upstream reaches of San Ysidro Creek had the deepest peak-flow depths, while Buena Vista and Oak Creeks experienced some of the lowest peak-flow depths b.
Ysidro Creeks, while it generally overestimated inundation in Buena Vista Creek and underestimated inundation in Oak and Romero Creeks (Fig. 5b).

Past work indicates that the effective viscosity ($\mu$) and yield strength ($\tau_y$) of debris flows vary substantially with sediment concentration (O’Brien et al. 1993), suggesting that sediment concentration is directly related to flow mobility. The sediment concentrations of the Montecito debris flows varied significantly from one watershed to another (Kean et al. 2019). Kean et al. (2019) estimated the sediment concentration of each flow using the ratio of mean sediment thickness to flow depth and interpreted higher ratios of sediment thickness to flow depth as being associated with a higher sediment concentration. The ratio of sediment thickness to flow depth was greatest at San Ysidro Creek (0.38), lowest at Buena Vista Creek (0.11), and between 0.14 and 0.23 at the three other watersheds (Kean et al. 2019). These differences suggest potentially substantial variations in typical sediment concentrations among the five watersheds, which makes it challenging for a model with spatially constant parameters to reproduce inundation across the entire study area. Differences in flow behavior among the different watersheds due to variations in sediment concentration and grain-size distribution could be represented by changes in $\chi$ and $\tau_y$, but it is also possible that differences among the five watersheds were sufficient to alter the relationship between flow discharge and volume. Rickenmann (1999), for example, derived different empirical formulae to estimate how peak discharge was related to flow volume in granular debris flows and muddy debris flows.

In addition to disparities in model performance between watersheds, we observed zones of underestimation and overestimation within individual watersheds that may be explained by temporal
variations in flow properties (e.g., sediment concentration and flow resistance) along the runout path. Field observations and large-scale flume experiments indicate that debris-flow rheology evolves in time and space during the course of runout (Iverson 2003). Temporal changes in pore-fluid pressure, for example, can lead to substantial variations in effective basal normal stress and flow resistance (Savage and Iverson 2003; George and Iverson 2014). Systematic over- and underestimations by the model within the San Ysidro Creek watershed may be explained by variations in effective flow resistance along the runout path. Deposits near the fan apex had a high concentration of boulders, which transitioned to finer deposits downstream, sometimes over short distances (Kean et al. 2019) (Fig. 2). The change in observed sediment size (Fig. 2) suggests that the physical properties of the flow varied along the length of the flow path. ProDF underestimated the extent of inundation near the mountain front along San Ysidro Creek, where boulder concentration was high, and overestimated inundation downstream, where boulder concentration was low (Fig. 10). Similar patterns of model behavior can be seen on Montecito and Oak Creeks (Fig. 10). One possible explanation for this model behavior is that flow resistance decreased, due to reductions in boulder content, as the flow traversed the fan.

In some cases, areas of model underestimation or overestimation can be attributed to the infrastructure located within the zone of inundation. For example, US Highway 101 runs across the southern extent of Montecito Creek, perpendicular to the direction of flow. During the debris flow on Montecito Creek, this portion of US 101 acted as a sink and captured a large amount of water and sediment. However, a bridge that crosses US 101 parallel to the flow direction acted as a conduit that transported mobile material over the highway, allowing sediment to deposit further downstream. ProDF simulations do not reproduce this flow path, as the flow stops being routed when it hits US 101 (Fig. 5b). US 101 is roughly 70 m wide at the point where the flow encounters the highway, and the bridge that acted as a flow path is nearly 100 m long. The modeled flow gets trapped behind the highway and is either stopped or rerouted elsewhere. The result is a large area of model underestimation at the southwestern extent of Montecito Creek (red circle in Fig. 5b). Therefore, some differences in model performance among the five watersheds may be attributed to the degree to which roadways, bridges, and culverts affected the extent of mapped and modeled deposits.

**Sensitivity analysis**

ProDF was sensitive to $\chi$ and $\tau_y$, but not to $\rho$ (Online Resource 5). This suggests that a constant value of $\rho$ can be employed in future modeling scenarios without substantial sacrifices to model performance. However, because the values of $\chi$ and $\tau_y$ more strongly influenced the similarity index (Online Resource 5b and c), these parameters should be calibrated or estimated from values calibrated at other sites. Results further illustrated a tradeoff between $\chi$ and $\tau_y$, indicating that there is a range of $\chi$ and $\tau_y$ values that perform similarly well, although that range

![Fig. 9](image_url) The area inundated by ProDF was sensitive to the input debris-flow volume *a*. However, the substantial scatter indicates that the flow mobility parameters also influence the area inundated. Density, $\rho$, had minimal impact on the area inundated when compared to volume *b*, while $\chi$ and $\tau_y$ both meaningfully influenced the area inundated. Lower values of $\chi$ *c* and $\tau_y$ *d* led to larger areas of inundation. Areas in red indicate parameter combinations that result greater area inundated, while areas in blue indicate parameter combinations that produce smaller area inundated.
varies from watershed-to-watershed (Fig. 8). In order to be able to estimate flow resistance parameters in advance of an event, it would be beneficial to calibrate the model to past events in a wide range of settings to determine the factors that influence the best-fit ranges of $\chi$ and $\tau_y$.

The area inundated by ProDF was sensitive to the input volume (Fig. 9a), consistent with another recent study of the Montecito debris flows (Barnhart et al. 2021). Results also revealed that the inundated area was sensitive to $\chi$ (Fig. 9c) and $\tau_y$ (Fig. 9d), which highlights the importance of having inundation scenarios for model calibration and testing where volume is well-constrained. Volumes were well-constrained for our post-event analyses of the Montecito debris flows, but that is rarely the case. For hazard assessments and other predictive scenarios, debris-flow volume must be estimated based on watershed and/or rainfall characteristics. Current post-wildfire hazard assessment frameworks in the western USA, for example, utilize the volume model introduced by Gartner et al. (2014). This model was developed and tested in a limited geographic area, the Transverse Ranges of southern California, and even there, the uncertainty associated with modeled volumes is substantial.

Because there is a clear relationship between flow volume and the area inundated by ProDF, volume uncertainty will impact the performance of ProDF in predictive frameworks. This issue underscores the importance of reducing uncertainty in existing volume models, as well as the need to continue to develop models that can be used to constrain debris-flow volume in different settings (e.g., burned versus unburned) and geographic areas (Santi and Morandi 2013; Gartner et al. 2014; Nyman et al. 2020).

**Comparison to existing models**

Previous studies have utilized the extent of inundation mapped by Kean et al. (2019) to test debris-flow inundation models. Bessette-Kirton et al. (2019) ran simulations of LAHRARZ and FLO-2D on San Ysidro Creek and compared results with mapped inundation. Previous studies have utilized the extent of inundation mapped by Kean et al. (2019) to test debris-flow inundation models. Bessette-Kirton et al. (2019) ran simulations of LAHRARZ and FLO-2D on San Ysidro Creek and compared results with mapped inundation.
Fig. 11 Best-fit model simulations for a FLO-2D, b LAHARZ debris flow, c LAHARZ lahar, and d ProDF compared with mapped deposits from the San Ysidro Creek watershed. Areas where the model overlaps with the observations are shown as true positive (TP). The inset tables show our objective function results. Ω represents the similarity index, β represents areas of underestimation, and γ represents areas of overestimation. FLO-2D and LAHARZ results are from Bessette-Kirton et al. (2019)
Funding

This work was supported in part by the U.S. Geological Survey Landslide Hazards Program and the Joint Fire Sciences Program through grant #L20AAC0029. This material is based upon work supported by the National Science Foundation under Grant No. 1951274. This work was also supported by the Bert Butler Scholarship through the University of Arizona Department of Geosciences.

Availability of data and material

Topographic data used for model simulations are located at https://www.hydroshare.org/resource/c45c7c602b9746da9e8f99bd94b0676/.

Code availability

ProDF was coded in C and utilized MATLAB scripts to write input files and read output files. Code for the numerical model (ProDF) is located at https://www.hydroshare.org/resource/c45c7c602b9746da9e8f99bd94b0676/.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Barnhart KR, Jones RP, George DL et al (2021) Multi-model comparison of computed debris flow runout for the 9 January 2018 Montecito, California post-wildfire event. J Geophys Res Earth Surf. https://doi.org/10.1029/2021JF006245

Baum RL, Godt JW (2010) Early warning of rainfall-induced shallow landslides and debris flows in the USA. Landslides 7:259–272. https://doi.org/10.1007/s10346-009-0177-0

Berti M, Bernard M, Gregoretti C, Simon A (2020) Physical interpretation of rainfall thresholds for runoff-generated debris flows. J Geophys Res Earth Surf. https://doi.org/10.1029/2019JF005513

Berti M, Simon A (2007) Prediction of debris flow inundation areas using empirical mobility relationships. Geomorphology 90:144–161. https://doi.org/10.1016/j.geomorph.2007.01.014

Bessest-Kiton E, Kean JW, Coe JA et al (2019) An evaluation of debris-flow runout model accuracy and complexity in Montecito, CA: towards a framework for regional inundation-hazard forecasting. Contained in: Proceedings of the seventh international conference on debris-flow hazards mitigation, Golden, Colorado, USA, 257–264

Cannon S, Gartner J, Rupert M et al (2010) Predicting the probability and volume of postwildfire debris flows in the intermountain western United States. Geol Soc Am Bull 122:127–144. https://doi.org/10.1130/B26459.1

Cannon SH (2001) Debris-flow generation from recently burned watersheds. Environ Eng Geosci 7:321–341. https://doi.org/10.2113/ gsegeosci.7.4.321

Cannon SH, Bigio ER, Mine E (2001) A process for fire-related debris flow initiation, Cerro Grande fire, New Mexico. Hydropl Process 15:3011–3023. https://doi.org/10.1002/hyp.388

Cesca M, D’Agostino V (2008) Comparison between FLO-2D and RAMMS in debris-flow modelling: a case study in the Dolomites. Monitoring, simulation, prevention and remediation of dense debris flows II. WIT Press, The New Forest, UK, pp 197–206

Chen J, Pangle LA, Gannon JP, Stewart RD (2020) Soil water repellency after wildfires in the Blue Ridge Mountains, United States. Int J Wildl Fire 29:1009–1020. https://doi.org/10.1071/WF20055

de Haas T, Braat L, Leuven JRFW et al (2015) Effects of debris flow composition on runout, depositional mechanisms, and deposit morphology in laboratory experiments. J Geophys Res Earth Surf 120:1949–1972. https://doi.org/10.1002/2015JF003525

Dibblee TW (1966) Geology of the central Santa Ynez Mountains, Santa Barbara County, California. California Division of Mines and Geology, San Francisco

Doerr SH, Shakesby R, MacDonald L (2009) Soil water repellency: a key factor in post-fire erosion. In: Cerda A, Robichaud PR (eds) Fire effects on soils and restoration strategies, 1st edn. CRC Press, Boca Raton, pp 213–240

Dowling CA, Santi PM (2014) Debris flows and their toll on human life: a global analysis of debris-flow fatalities from 1950 to 2011. Natural Hazards 71:203–227. https://doi.org/10.1007/s11069-013-0907-4

Ebel BA, Martin DA (2017) Meta-analysis of field-saturated hydraulic conductivity recovery following wildland fire: applications for hydrologic model parameterization and resilience assessment. Hydropl Process 31:3682–3696. https://doi.org/10.1002/hyp.11288

Ebel BA, Moody JA (2017) Synthesis of soil-hydraulic properties and infiltration timescales in wildfire-affected soils. Hydropl Process 31:324–340. https://doi.org/10.1002/hyp.10998

Freeman TG (1991) Calculating catchment area with divergent flow based on a regular grid. Comput Geosci 17:413–422. https://doi.org/10.1016/0098-3004(91)90048-1

Gabet E, Bookter A (2008) A morphometric analysis of gullies scoured by post-fire progressively bulked debris flows in southwest Montana, USA. Geomorphology 96:298–309. https://doi.org/10.1016/j.geomorph.2007.03.016

Gartner JE, Cannon SH, Santi PM (2014) Empirical models for predicting volumes of sediment deposited by debris flows and sediment-laden floods in the transverse ranges of southern California. Eng Geol 176:45–56. https://doi.org/10.1016/j.enggeo.2014.04.008

Gartner JE, Cannon SH, Santi PM, Dewolfe VG (2008) Empirical models to predict the volumes of debris flows generated by recently burned basins in the western U.S. Geomorphology 96:339–354. https://doi.org/10.1016/j.geomorph.2007.02.033

George DL, Iverson RM (2014) A depth-averaged debris-flow-model that includes the effects of evolving dilatancy. II. Numerical predictions and experimental tests. Proc Royal Soc A: Math Phys Eng Sci 470. https://doi.org/10.1098/rspa.2013.0820

Gibson S, Floyd I, Sánchez A, Heath R. (2021) Comparing single-phase, non-Newtonian approaches with experimental results: validating flume-scale mud and debris flow in HEC-RAS. Earth Surface Process Landforms 46:540–553. https://doi.org/10.1002/esp.5044

Gillett NP, Weaver AJ, Zwiers FW, Flannigan MD (2004) Detecting the effect of climate change on Canadian forest fires. Geophys Res Lett. https://doi.org/10.1029/2004GL020876

Griswold JP, Iverson RM (2008) Mobility statistics and automated hazard mapping for debris flows and rock avalanches. USGS Sci Investigations Rep 2007–5276

Landslides 19 • (2022) | 2071
Guthrie R, Befus A (2021) DebrisFlow Predictor: an agent-based runout model for shallow landslides. Nat Hazard 21:1029–1049. https://doi.org/10.1007/s11071-021-03715-4

Hoch OJ, McGuire LA, Youberg AM, Rengers FK (2021) Hydrogeomorphological recovery and temporal changes in rainfall thresholds for debris flows following wildfire. J Geophys Res (Earth Surface). https://doi.org/10.1029/2021JF006374

Heiser M, Scheidl C, Kaitna R (2017) Evaluation concepts to compare observed and simulated deposition areas of mass movements. Comput Geosci 21:335–343. https://doi.org/10.1016/j.cageo.2016.09.022

Horton P, Jaboyedoff M, Rudaz B, Zimmermann MN (2013) Flow-R, a model for susceptibility mapping of debris flows and other gravitational hazards at a regional scale. Nat Hazard 13:869–885. https://doi.org/10.1007/s11071-013-0860-8

Hürlimann M, Rickenmann D, Medina V, Bateman A (2008) Evaluation of approaches to calculate debris-flow parameters for hazard assessment. Eng Geol 102:152–163. https://doi.org/10.1016/j.enggeo.2008.03.012

Iverson RM (1992) Debris flow rheology: experimental analysis of bulking or en masse failure to blame? Geophys Res Lett 44:7310–7319. https://doi.org/10.1002/2017GL074243

Jaboyedoff M, Conover WL (1979) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21:239–245. https://doi.org/10.2307/1268522

Meyer GA, Wells SG (1997) Fire-related sedimentation events on alluvial fans, Yellowstone National Park, U.S.A. J Sediment Res 67:776–791. https://doi.org/10.1306/D426863A-2B26-11D7-864B0001020C1865D

Moody JA, Ebel BA, Nyman P et al (2015) Relations between soil hydraulic properties and burn severity. Int J Wildland Fire 25:279–293. https://doi.org/10.1071/WF14062

Nicholas AP (2010) Reduced-complexity modeling of free bar morphodynamics in alluvial channels. J Geophys Res Earth Surf. https://doi.org/10.1029/2010JF001774

Nyman P, Box WAC, Stout JC, Llobera M et al (2020) Debris-flow-dominated sediment transport through a channel network after wildfire. Earth Surf Proc Land 45:1155–1167. https://doi.org/10.1002/esp.4785

Nyman P, Rutherford ID, Lane PNJ, Sheridan GJ (2019) Debris flows in southeast Australia linked to drought, wildfire, and the El Niño–Southern Oscillation. Geology 47:491–494. https://doi.org/10.1130/G45939.1

Palaciu MC, Ulzio TP, Lamb MP (2021) Debris flow initiation from ravel-filled channel bed failure following wildfire in a bedrock landscape with limited sediment supply. GSA Bull. https://doi.org/10.1130/B35822.1

Parise M, Cannon SH (2012) Wildfire impacts on the processes that generate debris flows in burned watersheds. Nat Hazards 61:217–227. https://doi.org/10.1007/s11069-011-9769-9

Pelletier J (2008) Quantitative modeling of earth surface processes. Cambridge University Press. https://doi.org/10.1017/CBO9780511813849

Pudasaini SP (2012) A general two-phase debris flow model. J Geophys Res Earth Surf. https://doi.org/10.1002/jgrf.20148

Pudasaini SP (2011) Some exact solutions for debris and avalanche flows. Earth Surf Dynam 3:67–86. https://doi.org/10.5194/esurf-3-67-2015

Radeloff VC, Hammer RB, Stewart SI et al (2005) The wildland–urban interface in the United States. Ecol Appl 15:799–805. https://doi.org/10.1890/1051-0761(2005)015[0799:TWUIIT]2.0.CO;2

Rickenmann D (1999) Empirical relationships for debris flows. Natural Hazards: Journal of the International Society for the Prevention and Mitigation of Natural Hazards 19:47–77
