https://doi.org/10.5194/egusphere-2023-1931
Preprint. Discussion started: 18 September 2023

Probabilistic assessment of postfire debris-flow inundation in response to forecast rainfall

Alexander B. Prescott[1], Luke A. McGuire[1], Kwang-Sung Jun[2], Katherine R. Barnhart[3], Nina S. Oakley[4]

[1]Department of Geosciences, The University of Arizona, Tucson, AZ, USA
[2]Department of Computer Science, The University of Arizona, Tucson, AZ, USA
[3]U.S. Geological Survey, Denver, CO, USA
[4]California Geological Survey Burned Watershed Geohazards Program, Sacramento, CA, USA

Correspondence to: Alexander B. Prescott (alexprescott@arizona.edu)

Abstract. Communities downstream from burned steeplands face increases in debris-flow hazards due to fire effects on soil and vegetation. Rapid postfire hazard assessments have traditionally focused on quantifying spatial variations in debris-flow likelihood and volume in response to design rainstorms. However, a methodology that provides estimates of debris-flow inundation downstream from burned areas based on forecast rainfall would provide decision-makers with information that directly addresses the potential for downstream impacts. We introduce a framework that integrates a 24-hour lead-time ensemble precipitation forecast with debris-flow likelihood, volume, and runout models to produce probabilistic maps of debris-flow inundation. We applied this framework to simulate debris-flow inundation associated with the 9 January 2018 debris-flow event in Montecito, California, USA. Sensitivity analyses indicate that reducing uncertainty in postfire debris-flow volume prediction will have the largest impact on reducing inundation outcome uncertainty. The study results are an initial step toward an operational hazard assessment product that includes debris-flow inundation.

1 Introduction

Debris flows threaten life and property in mountainous areas worldwide (Dowling and Santi, 2014). In the first few years following fire, debris-flow hazards are greater relative to nearby unburned areas due to reductions in soil infiltration capacity (Ebel, 2020; Larsen et al., 2009; McGuire and Youberg, 2020), decreased vegetation cover (Cerdà and Doerr, 2005; Hoch et al., 2021; Stoof et al., 2012), increases in sediment availability (Nyman et al., 2013), and reduced thresholds for sediment entrainment (Moody et al., 2005). Postfire debris flows typically initiate when short-duration, high-intensity rainfall generates runoff that rapidly entrains sediment from hillslopes and channels (DeGraff et al., 2015; Gabet and Bookter, 2008; Kean et al., 2011; McGuire et al., 2017). Fire-related debris flows pose hazards globally, including in Europe (Conedera et al., 2003; Diakakis et al., 2023; Esposito et al., 2023; García-Ruiz et al., 2013; Lourenço et al., 2012), Asia (Jin et al., 2022; Lee et al., 2022; Touge et al., 2023), western North America (e.g., Jordan, 2015; Jordan and Covert, 2009; Kean et al., 2019), and Australia (Nyman et al., 2011). Due to the potential for widespread (i.e., within hundreds of different watersheds) increases in debris-flow susceptibility following fire, it is critical to rapidly identify downstream areas most threatened by debris-flow runout and to quantify the uncertainty associated with their identification.

Empirical models that assess debris-flow likelihood and volume at the drainage basin scale have been developed for the western United States in the 1-2 years immediately following wildfire (Gartner et al., 2014; Staley et al., 2017). These models can be used to produce rapid hazard assessments in response to design rainstorms with spatially uniform intensity (e.g., Staley et al., 2016). Past studies demonstrate that, in addition to factors related to topography and soil burn severity, rainfall intensity over a 15-minute duration (I_{15}) controls the likelihood of debris-flow initiation within a basin (Staley et al., 2017) as well as debris-flow volume

https://doi.org/10.5194/egusphere-2023-1931
Preprint. Discussion started: 18 September 2023

(Gartner et al., 2014). Forecasts of debris-flow volume are highly uncertain, especially when applied in settings not represented within their training datasets (Gorr et al., 2023; Wall et al., 2023), and may exceed a factor of 10 (Gartner et al., 2014). Debris-flow likelihood and volume both increase with I_{15}, which indicates that variations in rainfall intensity, even over small spatial (i.e., a low-order basin) and temporal (15-minute) scales, will play an important role in determining the likelihood and spatial extent of debris-flow impacts. Assessing the downstream impact of debris flows, however, requires information about their runout and inundation extent.

Rapid postfire hazard assessments (e.g., Staley et al., 2016) do not currently provide information about downstream impacts, although recent debris-flow events (Kean et al., 2019) and a user needs assessment highlight the need for such a product (Barnhart et al., 2023). The probability that a downstream area will be impacted by a debris flow depends, in part, on the likelihood of a debris flow initiating in a basin upstream, the size (volume) of the debris flow, and the movement of the flow as driven by topography and flow dynamics. Several debris-flow inundation models have been applied to simulate a recent series of postfire debris flows following the 2017 Thomas Fire in southern California (Barnhart et al., 2021; Gibson et al., 2022; Gorr et al., 2022). The models used in these studies vary in their representation of debris-flow physics, from those that account for multiphase flow dynamics and pore-pressure feedbacks (George and Iverson, 2014; Iverson and George, 2014) to those that use a semi-empirical approach to route flow across the landscape (Gorr et al., 2022). All the inundation models require debris-flow volume as an input, although some models require an inflow hydrograph (Barnhart et al., 2021), whereas others require volume be specified as a single number at each initiation point representing the total volume mobilized (Gorr et al., 2022). The model proposed by Gorr et al. (2022), the Progressive Debris-Flow routing and inundation model (ProDF), performed similarly to other inundation models but required less run time for equivalent simulations. This makes it a promising tool for evaluating runout for a large number of forecast precipitation scenarios in an operational setting.

Existing debris-flow likelihood, volume, and runout models therefore provide the necessary components to create a framework for postfire hazard assessments that includes probabilistic estimates of inundation area in response to a forecast or design rainstorm, but such a framework has yet to be developed and explored. Probabilistic frameworks for predicting debris-flow runout have been explored in unburned settings, although runout models have not been directly linked with others designed to predict debris-flow likelihood and volume (Aaron et al., 2022; Sun et al., 2021). As a result, fundamental questions remain regarding the propagation of uncertainty through various model components (i.e., from rainfall to flow volume to runout) as well as the benefits and limitations of such an approach at forecast lead times (≥ 24 hours) needed for decision-making (e.g., evacuation, allocating resources).

A postfire debris-flow inundation hazard assessment should reflect uncertainty in forecast inundated area (Barnhart et al., 2023), and past work identifies at least three ways through which substantial uncertainty is likely to arise: (1) forecast peak I_{15} needed as input for debris-flow likelihood and volume models (Oakley et al., 2023); (2) simulated debris-flow volume given a peak I_{15} (Gartner et al., 2014); and (3) flow mobility parameters needed to drive a debris-flow runout model (Aaron et al., 2019). Precipitation forecasts in the weeks to hours ahead of an event include considerable uncertainty regarding short-duration and high-intensity rainfall rates (Oakley et al., 2023). Assuming a perfect rainfall forecast, debris-flow volume models have order-of-magnitude uncertainty around a given prediction (Gartner et al., 2014). Lastly, debris-flow runout model parameters that influence flow mobility require calibration, a process subject to observation biases, model assumptions, and subjective user decisions (Aaron et al., 2019). Runout model parameter uncertainty can be considerable, particularly in areas without data from prior events to calibrate against (Zeng et al., 2023), resulting in poor predictive performance of debris-flow runout (e.g., Gorr et al., 2023). Because

https://doi.org/10.5194/egusphere-2023-1931
Preprint. Discussion started: 18 September 2023

prior studies have found that debris-flow runout is sensitive to flow volume (e.g., Barnhart et al., 2021; Gorr et al., 2022), and given that uncertainty in rainfall intensity propagates forward into debris-flow volume predictions, we propose a framework for generating probabilistic debris-flow inundation maps that links atmospheric modeling with debris-flow models.

The main objectives of this work were to (1) develop an integrative atmosphere-debris-flow model framework to generate a spatially distributed forecast of inundation probability, (2) apply the proposed framework to assess debris-flow inundation downstream from burned basins using an atmospheric model ensemble designed to mimic a 24-hour lead time forecast, and (3) quantify the relative importance of key input parameters using global sensitivity analyses. A probabilistic inundation map is the final product of our model framework. In practice, this product is one that could be used to improve situational awareness for decision-makers. This study is a first step toward the development of an operational framework for probabilistic assessments of debris-flow inundation downstream from recently burned areas and provide guidance for future work aimed at further quantifying and reducing uncertainty.

2 Study area

We focused our study on a portion of the 2017 Thomas Fire near the community of Montecito, California, USA, that was impacted by postfire debris flows in January 2018, causing 23 fatalities and substantial economic losses (Kean et al., 2019; Lancaster et al., 2021) (Fig. 1). The Thomas Fire burned more than 1100 km^2, including a series of steep basins in the Santa Ynez Mountains upstream from Montecito. The fire ignited in December 2017 and was not yet contained when a weak atmospheric river with an embedded narrow cold frontal rainband (NCFR) impacted the area on 9 January 2018 (Oakley et al., 2018). As the NCFR propagated over the burned basins upstream from Montecito, it produced rainfall with peak I_{15} between 78 mm/h and 105 mm/h (Kean et al., 2019). Runoff entrained sediment from burned hillslopes and channels (Alessio et al., 2021; Kean et al., 2019; Morell et al., 2021), producing debris flows that traveled several kilometers down the alluvial fan (Lancaster et al., 2021). We focused on six of these debris-flow-producing basins where more than 630,000 m^3 of sediment was mobilized and debris-flow inundation extent was mapped shortly following the event (Kean et al., 2019). Debris flows in the Montecito Creek Basin were sourced from two upstream burned basins, while the remaining four creeks had a single upstream source. All simulations were run using topography from a 5-meter resolution digital elevation model derived from airborne lidar collected before the event (Fig. 1).

3 Methods

3.1. Overview of model framework

We coupled rainfall output from an atmospheric model ensemble with debris-flow likelihood, volume, and runout models to generate a probabilistic forecast of postfire debris-flow inundation downstream from the six debris-flow producing basins (Fig. 2). We used a 100-member atmospheric model ensemble representing a 24-hour lead-time forecast of the 9 January 2018 precipitation event (Oakley et al., 2023). For each ensemble member, we computed basin-averaged values of peak 15-minute rainfall intensity (I_{15}) and used this as input into debris-flow likelihood (Staley et al., 2017) and volume (Gartner et al., 2014) models to predict (1) whether each basin would produce a debris flow as well as (2) the volume of sediment a debris flow, if initiated, would mobilize. We then used the ProDF debris-flow runout model (Gorr et al., 2022) to estimate downstream inundation extent and peak flow depths. In this step, we incorporated uncertainty in debris-flow volume for a given I_{15} as well as uncertainty in ProDF flow-mobility parameter values into the forecast by utilizing Monte Carlo sampling methods. Finally, we produced a map of spatially variable forecast probabilities of inundation by averaging the inundated area results from each ProDF simulation.

https://doi.org/10.5194/egusphere-2023-1931
Preprint. Discussion started: 18 September 2023

3.2. Atmospheric model ensemble design

We used the Weather Research and Forecast (WRF) atmospheric model Version 4.3 (Skamarock et al., 2021) to generate a 24-hour lead time, 100-member ensemble rainfall forecast for the 9 January 2018 event (Oakley et al., 2023). The ensemble produced a distribution of precipitation rates that reflects forecast uncertainty (Fig. 1c). Output consisted of spatially variable rainfall depths in 5-minute intervals with 1-km horizontal resolution across the study area. We spatially averaged rainfall intensities over each of the six basins (ranging from 0.45 to 8.94 km^2; Kean et al., 2019) and used a 15-minute moving window to calculate the peak I_{15} for each basin for every ensemble member. In this way, variability in the timing, location, and spatial structure of forecast precipitation translates into variability in the I_{15} subsequently used to predict debris-flow likelihood, volume, and runout.

3.3. Debris-flow likelihood and volume models

We used empirical models for postfire debris-flow likelihood (Staley et al., 2017) and volume (Gartner et al., 2014) to determine if a basin would produce a debris flow and how large it would be. The Staley et al. (2017) M1 model determines debris-flow likelihood, whereas the Gartner et al. (2014) emergency assessment volume (EAV) model predicts debris-flow volume. These models use basin-averaged metrics related to topography, soil properties, soil burn severity, and peak I_{15} as input. In this study, all input parameters for the M1 and EAV models were fixed for each basin with the exception of I_{15}.

Using the basin-averaged I_{15} from each WRF ensemble member, we computed debris-flow likelihood for each of the six basins using the M1 model. If debris-flow likelihood was less than 0.5, we assumed that a debris flow would not initiate. If likelihood was greater than 0.5, we assumed a debris flow would initiate and determined its volume using the EAV model. We defined a log-uniform distribution centered on the EAV-predicted volume with an order-of-magnitude envelope above and below the predicted volume (Fig. S1). This range of support is consistent with the prediction uncertainty of the EAV model, as well as similar models (Gartner et al., 2014). For a given ProDF simulation, we drew input volumes from these distributions by sampling from a log-uniform random variable over the range [0.1, 10.0] and multiplying the six EAV-predicted volumes (one for each of the six basins) by this scalar.

3.4 Debris-flow inundation model

We used ProDF to simulate debris-flow runout and inundation (Gorr et al., 2022). The model requires two flow mobility parameters and input debris-flow volumes at user-defined flow starting points, from which flow is iteratively routed downstream. The two flow mobility parameters, χ [s^{-1} $m^{-0.5}$] and τ_y [Pa], control the flow depth and the minimal basal shear stress that permits flow motion, respectively.

It is common to calibrate models by choosing a single set of "best" parameters based on some objective function optimization tied to an observation dataset and to then use this optimal parameter set in forward model applications (e.g., Pirulli, 2010). A limitation of this methodology is that it precludes exploration of a wider range of equifinal possible parameter sets, and different observation datasets may lead to different choices about which parameters are optimal. Instead, we used a statistical inference procedure similar to that of Aaron et al. (2019) to define a joint posterior distribution over the flow mobility parameters that uses the similarity index (Gorr et al., 2022; Heiser et al., 2017) as the objective of the maximum likelihood estimator function for nonlinear systems (Hill and Tiedeman, 2007). The similarity index varies from -1 to 1, with 1 indicating a perfect match between the simulated and observed area inundated. We generated samples from this distribution using the *emcee* Python implementation (Foreman-Mackey et al., 2013) of the Markov Chain Monte Carlo (MCMC) affine invariant ensemble sampler (Goodman and Weare, 2010). This

https://doi.org/10.5194/egusphere-2023-1931
Preprint. Discussion started: 18 September 2023

method is advantageous as we were able to sample uniformly from the MCMC output to gather flow mobility parameters for ProDF simulations. We calibrated the posterior distribution to the Oak, San Ysidro, Buena Vista, and Romero Creek Basins and reserved Montecito Creek to test the calibrated distribution. Additional details on the ProDF calibration are given in Supporting Text S1.

3.5. Inundation forecast

We generated a forecast of inundation probability by averaging together many individual ProDF simulations that were run with input parameters drawn repeatedly and independently from the calibrated distributions (Fig. 2). We ran 50 simulations for every WRF ensemble member's prediction of peak I_{15} (i.e., a total of 5,000 ProDF simulations). Debris-flow volumes for all six initiation points were drawn from the log-uniform distributions defined in Section 3.3, and the flow mobility parameters (χ, τ_y) were drawn from the calibrated joint posterior distribution defined in Section 3.4. Every simulation produced a map of peak debris-flow depth. Depth maps from all simulations were converted to binary maps of inundation presence using a threshold depth of 10 cm (e.g., Gorr et al., 2022). Averaging the binary inundation maps together with equal weights produced a map with values between 0 and 1 representing the fraction of simulations that inundated each grid cell, which we interpret as an inundation probability.

To investigate the role of input debris-flow volume on the joint distribution of forecasts and observations (described further in Section 3.6), we also generated probabilistic inundation maps for two additional scenarios, referred to as scenarios A and B. In scenario A, we use observed debris-flow volumes as input for all simulations. In this scenario, we minimize uncertainty in debris-flow volume, so we expect model performance to improve relative to the forecast. In scenario B, we assigned a debris-flow volume to each basin using the EAV prediction with the observed peak I_{15} as the input (Kean et al., 2019). This scenario also does not utilize any data from the atmospheric model ensemble. The peak I_{15} at each debris-flow initiation point was computed with inverse distance weighting of the observed rainfall rates at the KTYD and Doulton Tunnel (DT) rain gauges (78 and 105 mm/hr, respectively; Kean et al., 2019) (Fig. 1c). Evaluating model performance when debris-flow volume (scenario A) or rainfall (scenario B) are known is useful for identifying the source of any observed over or underestimation of inundated area in the forecast.

3.6 Comparing simulated and observed inundation

Debris-flow inundation model results are commonly assessed one simulation at a time by optimizing an objective function of the mapped debris-flow deposits and simulated inundation zones (e.g., Barnhart et al., 2021; Gibson et al., 2022; Gorr et al., 2022). Probabilistic forecasts cannot be directly evaluated with a similar binary classification and optimization procedure. Instead, we classified grid cells as inundated or not using a threshold probability p_t to explore the extent to which the observation was contained within the range of inundation scenarios represented by the ensemble forecast. If a given cell's forecast probability of inundation p satisfied $p \geq p_t$, it was classified as inundated, and otherwise it was not. For values of p_t between 0 and 1 (discretized every 0.01), we classified each cell in the domain and then computed the similarity index.

Probabilistic forecasts can also be evaluated using a distributions-oriented approach. A distributions-oriented approach considers the entire joint distribution of forecasts and observations, $f(p, I)$, where f is the joint probability density function and I is the observed binary inundation outcome (i.e., $I = 1$ if a debris flow actually occurred in the grid cell, else $I = 0$). The joint distribution contains all the relevant information about the forecasts and observations needed for a complete verification of a forecast model (Wilks, 2019). It can be factored in two ways into conditional and marginal probabilities that are more practical for analysis, one of which is the calibration-refinement factorization (Murphy and Winkler, 1987):

https://doi.org/10.5194/egusphere-2023-1931
Preprint. Discussion started: 18 September 2023

$$f(p, I) = f(I \mid p) f(p), \quad (1)$$

This factorization allows inspection of two desirable properties of probabilistic forecasts: (i) they should be reliable in that the forecast event actually happens with a frequency close to the forecast probability, $f(I = 1 \mid p) \cong p$ (e.g., a forecast probability of 30% comes true ~30% of the time); and (ii) the distribution of forecast probabilities $f(p)$ should be dispersed toward the extreme values of 0 and 1, indicating that the model has confidence in its own predictions (Gneiting et al., 2007; Murphy, 1993). The property of (i) is referred to as calibration (i.e., reliability) and the property of (ii) is referred to as refinement (i.e., sharpness), and a general goal with probabilistic forecast models is for them to be as sharp as possible without sacrificing calibration (Gneiting et al., 2007; Wilks, 2019). We borrowed distributions-oriented methodologies developed in the weather modeling community over the last several decades (e.g., Bröcker and Smith, 2007; DeGroot and Fienberg, 1983; Gneiting et al., 2007; Murphy, 1993; Murphy and Winkler, 1987) to graphically assess the calibration-refinement factorization. Specifically, we used the reliability diagram (Bröcker and Smith, 2007; Wilks, 2019) to separately visualize the calibration and the refinement of the forecast model.

The first component of a reliability diagram is the calibration curve, a function of the conditional distribution $f(I \mid p)$ that provides a visual assessment of the fit between the distribution of forecast probabilities and the observed zones of inundation (Bröcker and Smith, 2007; Gneiting et al., 2007). Quantities on the x-axis answer the question, "what is the mean probability of inundation, $\bar{p}_k$, of all grid cells in the k^{th} bin?" On the y-axis is the frequency of observed inundation conditioned on the binned forecast probabilities, also referred to as the observed relative frequency, which provides an estimate of the calibration distribution: $y_k \approx f(I = 1 \mid p)$ (Murphy and Winkler, 1987). These answer the question, "given a forecast probability of p, how often is it correct?" With a perfect probabilistic forecast, the bin-averaged probabilities will exactly match the observed relative frequencies, $y_k = p_k$, and points will fall along the one-to-one line in the calibration curve (Bröcker and Smith, 2007). In reality, sampling variability causes deviations from the one-to-one line even for a perfectly reliable model (Wilks, 2019). Points that fall above the one-to-one line indicate that the model is under-predicting the observed extent of inundation, referred to as under-forecasting, while points below the line indicate over-forecasting. The $(\bar{p}_k, y_k)$ are computed as:

$$\bar{p}_k = \frac{1}{|\omega_k|} \sum_{\omega_k \subset \Omega} p\,, \quad (2)$$

$$y_k = \frac{1}{|\omega_k|} \sum_{\omega_k \subset \Omega} I, \quad (3)$$

where Ω is the set of all grid cells in the spatial modeling domain, ω_k is the subset of Ω that satisfies $p \in B_k$, B_k is the k^{th} bin interval, and $|\omega_k|$ is the number of model grid cells in ω_k. In all cases, we used bin widths of 10 percentage points. Figure S2 provides a visual demonstration of ω_k and the computation of y_k.

The second component of a reliability diagram shows the refinement distribution of the forecasts, answering the question, "how often is each probability of inundation predicted by the forecast model?" This plot shows the histogram of $f(p)$ using the same bins as used in constructing the calibration curve. A sharp forecast will predict probabilities close to 0 or 1 most of the time and will therefore have the highest counts near the boundaries of the histogram.

3.7 Sensitivity analyses

We performed two analyses to explore the sensitivity of inundated area to each of the three ProDF input parameters (debris-flow volume, χ, and τ_y). The goals of these sensitivity analyses were to apportion uncertainty in the model output amongst the three

https://doi.org/10.5194/egusphere-2023-1931
Preprint. Discussion started: 18 September 2023

parameters and to rank them in terms of relative importance for determining inundated area (Razavi and Gupta, 2015; Saltelli et al., 2008). We used inundated area as a summary of model output because it serves as a simple proxy for downstream impacts. We used the SALib Python package (Herman and Usher, 2017; Iwanaga et al., 2022) to implement the PAWN global sensitivity analysis method (Pianosi and Wagener, 2018). The method returns a sensitivity index between 0 and 1 for each input parameter, with higher indices indicating greater influence on model output.

First, we performed a domain-aggregated sensitivity analysis. To distinguish a significant sensitivity value from one that is due solely to approximation error of the PAWN method, we included a dummy parameter of random numbers in the analysis. Bootstrapping (n=50) was used to compute 95% confidence intervals about the median sensitivity index of each parameter. Parameters whose 95% confidence interval exceeded that of the dummy parameter were considered significant (Pianosi and Wagener, 2018).

Second, we computed spatially distributed sensitivity indices in every grid cell where simulated inundation occurred on a cell-by-cell basis. The sensitivity response variable was taken to be the local binary inundation value, I, from each ProDF simulation. This analysis allowed us to compare the relative importance of debris-flow volume, χ, and τ_y as a function of location, revealing patterns of model sensitivity over the length of the debris-flow runout paths.

4 Results

The forecast probability of debris-flow inundation is shown in Fig. 3. Areas that had a non-zero probability of inundation generally encompassed areas of observed debris-flow inundation, indicating the observation was contained within the range of outcomes represented by the ensemble. A threshold probability of 16% (i.e., $p \geq 0.16$) resulted in a binary inundation map that compared best with observations (similarity index of -0.02; Fig. S3c). For comparison, validation of the flow mobility parameter posterior distribution on Montecito Creek resulted in a similarity index of -0.047 (Fig. S4d).

The calibration component of the reliability diagram shows that the forecast average probabilities were small relative to the observed relative frequencies, indicating that the forecast tended to underestimate the observed extent of inundation (Fig. 4a). The histogram of forecast probabilities indicates a lack of forecast sharpness because the probabilities are clustered near 0 and monotonically decrease in frequency toward 1 (Fig. 4b). In contrast, analysis of the probabilistic inundation map obtained from scenario A, where debris-flow volume was set to the observed value for each basin, shows forecast probabilities that are well-calibrated to the observed frequencies and that are sharp (Fig. 4c-d). Scenario B, where EAV predictions of the observed peak I_{15} were used, resulted in a map of inundation that over-forecasted the observed inundation (Fig. 4e).

The global sensitivity analysis revealed that the runout model was significantly sensitive to all three parameters (debris-flow volume, χ, τ_y) with volume being the most influential (Fig. S5). The median PAWN sensitivity indices associated with volume, χ, τ_y, and the dummy variable were 0.38, 0.10, 0.09, and 0.05, respectively. Additionally, maps depicting spatially distributed sensitivity indices indicate that the relative importance of the input parameters varied both downstream and across-stream (Fig. 5). This implies that any of the three parameters may be most influential in determining whether an area will be inundated depending on location. The χ parameter had greatest influence at higher elevations in overbank areas removed from the primary channel, whereas the τ_y parameter had greatest influence toward the distal portion of the fan. Flow volume exerted a strong control on inundation throughout the model domain.

https://doi.org/10.5194/egusphere-2023-1931
Preprint. Discussion started: 18 September 2023

5 Discussion

Our study results indicate that reducing uncertainty in debris-flow volume predictions will have a substantial effect on reducing uncertainty associated with inundation. Even in the region for which it was developed, the prediction uncertainty associated with the EAV model is greater than an order of magnitude (Gartner et al., 2014; Barnhart et al., 2021). In addition, variations in forecast storm structure can result in highly variable precipitation intensity, duration, timing, and spatial distribution (Fig. 1c) (Oakley et al., 2023), driving further uncertainty in debris-flow volume prediction through the forecast peak-I_{15}. This uncertainty propagates from rainfall into a wide range of inundation outcomes (Fig. S6). Improved prediction of postfire debris-flow volume, whether through improved forecasts of I_{15} or improved volume models, would reduce the uncertainty associated with the inundation probabilities and result in a narrower range of inundation outcomes. However, the modular structure of the framework presented here makes it straightforward for the debris-flow likelihood and volume models to be replaced with updated or region-specific alternatives as they become available (e.g., Diakakis et al., 2023; Santi and Morandi, 2013; Staley et al., 2013; Wall et al., 2023).

Calibration of the flow mobility parameters created a posterior distribution with a range of support over χ-τ_y space (Fig. S4a). Capturing the effect of this spectrum of possible flow behaviors was important because the debris-flow properties that we expect to influence the flow mobility parameters (e.g., grain size distribution, sediment concentration) are unknown before an event. Furthermore, it is common for debris-flow properties to change as flows move downslope (Iverson, 1997, 2003). ProDF uses constant values for χ and τ_y across the simulation domain, limiting debris-flow behavior to a single characteristic type. The forecast model, however, enabled the representation of multiple flow rheologies in the probabilistic prediction of inundation by sampling flow mobility parameters from the calibrated posterior distribution. At sites without past events to aid in calibration, we would expect greater uncertainty in any estimate of a posterior parameter distribution. Still, the greater importance of debris-flow volume in determining inundated area indicates that placing better constraints on debris-flow volume, including variations resulting from entrainment or deposition along the travel path, warrants prioritization over constraints on flow mobility parameters (Fig. 5, Fig. S5, and Fig. S6).

Analysis of the forecast inundation probabilities using the reliability curve showed that the model tended to under-forecast the observed frequencies of inundation (Fig. 4a). For example, in areas with a forecast inundation probability of 40-50%, the frequency of observed inundation was nearly 60%. In other words, the model simulated inundation less often than observed. We attribute this bias, at least in part, to the extreme nature of the 2018 precipitation event and the challenges of representing this event at a 24-hour lead time in the atmospheric model. The observed peak I_{15} values lie in the tail of the atmospheric model ensemble spread (Fig. 1c) (Kean et al., 2019; Oakley et al., 2023): thus, most values in the ensemble are lower than observed. This leads to lower ensemble predictions of debris-flow volume and likelihood with the EAV and M1 models and therefore less inundation than would be expected given the observed I_{15}. This may also explain why the best match of simulated inundation to the observed deposits occurs at a threshold probability level of 16% (Fig. 3 and Fig. S3c).

Our interpretation that the ensemble distribution of predicted I_{15} led to under-forecasting is supported by the reliability diagrams associated with forecast scenarios A and B, which were run using observed debris-flow volumes and observed rainfall intensities, respectively (Fig. 4 and Fig. S7). The calibration curves from these two scenarios indicate high sensitivity of the calibration to the input debris-flow volumes. When volumes larger than the observed were used, as was generally the case in scenario B, the model over-forecasted inundation and the calibration curve dropped below the diagonal one-to-one line (Fig. 4e). When the observed volumes were used (i.e., scenario A), the calibration curve is close to the diagonal, and the refinement distribution shows that

https://doi.org/10.5194/egusphere-2023-1931
Preprint. Discussion started: 18 September 2023

extreme values were most commonly forecast. This indicates that the calibrated forecast model is both reliable and sharp when the volumes are well-constrained (Fig. 4c-d). However, in the forecast model where debris-flow volume was a function of the peak I_{15} derived from the atmospheric model ensemble, the ultimate effects of having a lower-than-observed peak I_{15} in many atmospheric model ensemble members was likely at least partially offset by the EAV model's bias to overpredict debris-flow volumes (Barnhart et al., 2021; Kean et al., 2019).

The methods presented here take a step toward operational assessments of postfire debris-flow hazards associated with an incoming rainstorm. This information could be used to support decisions regarding evacuations, staging of equipment and emergency personnel, and debris-flow mitigation efforts. Additional assessments of the integrated modeling approach presented here in different geographic and climatic settings would help generalize findings and develop guidelines for constraining flow mobility parameters in areas where there are no historical observations that can be used to calibrate the runout model, ProDF. Computing resource constraints present a challenge for future studies and operational use because the probabilistic forecast of debris-flow inundation and the atmospheric ensemble forecast both require many core-hours of computing time. Approaches to reduce computation times include optimizing aspects of the simulation for the task at hand (e.g., number of ensemble members, horizontal grid spacing; Oakley et al., 2023), running debris-flow runout simulations massively in parallel, and limiting the spatial extent of modeling efforts.

The framework proposed here could also be applied in a pre-fire context to assess postfire hazards. In this case, the rainfall intensity input could be determined from local climatological data and soil burn severity characteristics could be simulated (e.g., Kean and Staley, 2021; Staley et al., 2018). Pre-fire assessments of postfire hazards provide valuable insight into areas of greatest concern that could assist with community planning, emergency management, and debris-flow hazard mitigation (McCoy et al., 2016; Tillery et al., 2014).

6 Conclusions

Probabilistic debris-flow inundation maps may support emergency management and hazard mitigation efforts in advance of forecasted storms over burned areas. To explore the feasibility of producing such maps, we integrated output from a 24-hour lead-time atmospheric model ensemble, empirical models for postfire debris-flow likelihood and volume, and a computationally efficient debris-flow runout model to produce a spatially distributed probabilistic forecast of debris-flow inundation. The inundation forecast reproduced the observed inundated area reasonably well at a relatively low probability threshold (16%), which likely reflects the extreme nature of the observed rainfall rates relative to those in the atmospheric model ensemble. The reliability of the forecast strongly depends on the distribution of input debris-flow volumes, and a sensitivity analysis of the forecast inundation to the model input parameters indicates that reducing the uncertainty of debris-flow volume predictions would substantially reduce the uncertainty in the forecasted inundated area. The modeling framework is generic in that each component may be replaced with improved or region-specific alternatives as they become available. This work takes initial steps toward an operational debris-flow hazard assessment framework and may help identify targeted opportunities for model improvements that would be most beneficial.

Data availability. The model code, input data, and output data used in this study are available in a Zenodo online archive under the open-access Creative Commons Attribution 4.0 International license (Prescott et al., 2023).

https://doi.org/10.5194/egusphere-2023-1931
Preprint. Discussion started: 18 September 2023

Author contributions. LAM, K-SJ, and NAO acquired funding for this project. LAM, K-SJ, and ABP conceptualized and developed the methodology. ABP developed the model code and ran simulations, with contributions from KRB. ABP and LAM prepared the manuscript draft, and all co-authors reviewed and edited the draft.

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

Disclaimer. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Acknowledgments. This work was supported by the Joint Fire Sciences Program through grant #L20AC00029 and the California Department of Water Resources Atmospheric River Program (4600013361). This material is based upon High Performance Computing (HPC) resources supported by the University of Arizona TRIF, UITS, and Research, Innovation, and Impact (RII) and maintained by the University of Arizona Research Technologies department. The authors are grateful to Matthew Simpson at Scripps Institution of Oceanography for running the WRF simulations used in this research and Brady Gales at the University of Arizona for early contributions to the experimental design.

References

Aaron, J., McDougall, S., and Nolde, N.: Two methodologies to calibrate landslide runout models, Landslides, 16, 907–920, https://doi.org/10.1007/s10346-018-1116-8, 2019.

Aaron, J., McDougall, S., Kowalski, J., Nolde, N., and Mitchell, A.: Probabilistic prediction of rock avalanche runout using a numerical model, Landslides, 19, 2853–2869, https://doi.org/10.1007/s10346-022-01939-y, 2022.

Alessio, P., Dunne, T., and Morell, K.: Post-wildfire generation of debris-flow slurry by rill erosion on colluvial hillslopes, J. Geophys. Res.-Earth, 126(11), e2021JF006108, https://doi.org/10.1029/2021JF006108, 2021.

Barnhart, K. R., Jones, R. P., George, D. L., McArdell, B. W., Rengers, F. K., Staley, D. M., and Kean, J. W.: Multi-model comparison of computed debris flow runout for the 9 January 2018 Montecito, California post-wildfire event, J. Geophys. Res.-Earth, 126(12), e2021JF006245, https://doi.org/10.1029/2021JF006245, 2021.

Barnhart, K. R., Romero, V. Y., and Clifford, K. R.: User needs assessment for postfire debris-flow inundation hazard products, U.S. Geological Survey Open-File Rep. 2023–1025, 25 pp., https://doi.org/10.3133/ofr20231025, 2023.

Bröcker, J., and Smith, L. A.: Increasing the reliability of reliability diagrams, Weather Forecast., 22(3), 651–661, https://doi.org/10.1175/WAF993.1, 2007.

Cerdà, A., and Doerr, S. H.: Influence of vegetation recovery on soil hydrology and erodibility following fire: An 11-year investigation, Int. J. Wildland Fire, 14(4), 423437, https://doi.org/10.1071/WF05044, 2005.

https://doi.org/10.5194/egusphere-2023-1931
Preprint. Discussion started: 18 September 2023

Conedera, M., Peter, L., Marxer, P., Forster, F, Rickenmann, D., and Re, L.: Consequences of forest fires on the hydrogeological response of mountain catchments: A case study of the Riale Buffaga, Ticino, Switzerland, Earth Surf. Proc. Land., 28(2), 117–129, https://doi.org/10.1002/esp.425, 2003.

DeGraff, J. V., Cannon, S. H., and Gartner, J.E.: The timing of susceptibility to post-fire debris flows in the western United States, Environ. Eng. Geosci., 21(4), 277–292, https://doi.org/10.2113/gseegeosci.21.4.277, 2015.

DeGroot, M. H., and Fienberg, S. E.: The comparison and evaluation of forecasters, J. Roy. Stat. Soc. D-Sta., 32(1–2), 12–22, https://doi.org/10.2307/2987588, 1983.

Diakakis, M., Mavroulis, S., Vassilakis, E., and Chalvatzi, V.: Exploring the application of a debris flow likelihood regression model in Mediterranean post-fire environments, using field observations-based validation, Land, 12(3), 555, https://doi.org/10.3390/land12030555, 2023.

Dowling, C. A., and Santi, P. M.: Debris flows and their toll on human life: A global analysis of debris-flow fatalities from 1950 to 2011, Nat. Hazards, 71, 203–227, https://doi.org/10.1007/s11069-013-0907-4, 2014.

Ebel, B. A.: Temporal evolution of measured and simulated infiltration following wildfire in the Colorado Front Range, USA: Shifting thresholds of runoff generation and hydrologic hazards, J. Hydrol., 585, 124765, https://doi.org/10.1016/j.jhydrol.2020.124765, 2020.

Esposito, G., Gariano, S. L., Masi, R., Alfano, S., and Giannatiempo, G.: Rainfall conditions leading to runoff-initiated post-fire debris flows in Campania, Southern Italy, Geomorphology, 423, 108557, https://doi.org/10.1016/j.geomorph.2022.108557, 2023.

Foreman-Mackey, D.: Corner.py: Scatterplot matrices in Python, J. Open Source Softw., 1(2), 24, https://doi.org/10.21105/joss.00024, 2016.

Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J.: Emcee: The MCMC Hammer, Publ. Astron. Soc. Pac., 125, 306–312, https://doi.org/10.1086/670067, 2013.

Gabet, E. J., and Bookter, A.: A morphometric analysis of gullies scoured by post-fire progressively bulked debris flows in southwest Montana, USA, Geomorphology, 96(3-4), 298–309, https://doi.org/10.1016/j.geomorph.2007.03.016, 2008.

García-Ruiz, J. M., Arnáez, J., Gómez-Villar, A., Ortigosa, L., and Lana-Renault, N.: Fire-related debris flows in the Iberian Range, Spain, Geomorphology, 196, 221–230, https://doi.org/10.1016/j.geomorph.2012.03.032, 2013.

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

George, D. L., and Iverson, R. M.: A depth-averaged debris-flow model that includes the effects of evolving dilatancy. II. Numerical predictions and experimental tests, P. Roy. Soc. A-Math. Phy., 470(2170), 20130820, https://doi.org/10.1098/rspa.2013.0820, 2014.

Gibson, S., Moura, L. Z., Ackerman, C., Ortman, N., Amorim, R., Floyd, I., Eom, M., Creech, C., and Sánchez, A.: Prototype scale evaluation of non-Newtonian algorithms in HEC-RAS: Mud and debris flow case studies of Santa Barbara and Brumadinho, Geosciences, 12(3), 134, https://doi.org/10.3390/geosciences12030134, 2022.

Gneiting, T., Balabdaoui, F., and Raftery, A. E.: Probabilistic forecasts, calibration, and sharpness, J. R. Stat. Soc., 69(2), 243–268, https://doi.org/10.1111/j.1467-9868.2007.00587.x, 2007.

Goodman, J., and Weare, J.: Ensemble samplers with affine invariance, Comm. App. Math. Com. Sc., 5(1), 65–80, http://doi.org/10.2140/camcos.2010.5.65, 2010.

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

Gorr, A. N., McGuire, L. A., Beers, R., and Hoch, O. J.: Triggering conditions, runout, and downstream impacts of debris flows following the 2021 Flag Fire, Arizona, USA, Nat. Hazards, 117, 2473–2504, https://doi.org/10.1007/s11069-023-05952-9, 2023.

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

Herman, J., and Usher, W.: SALib: An open-source Python library for sensitivity analysis, J. Open Source Softw., 2(9), 97, https://doi.org/10.21105/joss.00097, 2017.

Hill, M. C., and Tiedeman, C. R.: Effective Groundwater Model Calibration: With Analysis of Data, Sensitivities, Predictions, and Uncertainty, John Wiley & Sons, Hoboken, New Jersey, 480 pp., ISBN 9780471776369, 2007.

Hoch, O. J., McGuire, L. A., Youberg, A. M., and Rengers, F. K.: Hydrogeomorphic recovery and temporal changes in rainfall thresholds for debris flows following wildfire, J. Geophys. Res.-Earth, 126(12), e2021JF006374, https://doi.org/10.1029/2021JF006374, 2021.

Iverson, R. M.: The physics of debris flows, Rev. Geophys., 35(3), 245–296, https://doi.org/10.1029/97RG00426, 1997.

Iverson, R. M.: The debris-flow rheology myth, in: Debris-flow Hazards Mitigation: Mechanics, Prediction, and Assessment, edited by: Rickenmann, D., and Chen, C. L., Millpress, Rotterdam, Netherlands, 303–314, ISBN 907701778X, 2003.

Iverson, R. M., and George, D. L.: A depth-averaged debris-flow model that includes the effects of evolving dilatancy. I. Physical basis, P. Roy. Soc. A-Math. Phy., 470(2170), 20130819, https://doi.org/10.1098/rspa.2013.0819, 2014.

Iwanaga, T., Usher, W., and Herman, J.: Toward SALib 2.0: Advancing the accessibility and interpretability of global sensitivity analyses, Socio-Environmental Syst. Modell., 4, 18155, https://doi.org/10.18174/sesmo.18155, 2022.

Jin, T., Hu, X., Liu, B., Xi, C., He, K., Cao, X., Luo, G., Han, M., Ma, G., Yang, Y., and Wang, Y.: Susceptibility prediction of post-fire debris flows in Xichang, China, using a logistic regression model from a spatiotemporal perspective, Remote Sens.-Basel, 14(6), 1306, https://doi.org/10.3390/rs14061306, 2022.

Jordan, P.: Post-wildfire debris flows in southern British Columbia, Canada, Int. J. Wildland Fire, 25(3), 322–336, https://doi.org/10.1071/WF14070, 2015.

Jordan, P., and Covert, S. A.: Debris flows and floods following the 2003 wildfires in southern British Columbia, Environ. Eng. Geosci., 15(4), 217–234, https://doi.org/10.2113/gseegeosci.15.4.217, 2009.

Kean, J. W., and Staley, D. M.: Forecasting the frequency and magnitude of postfire debris flows across southern California, Earth's Future, 9(3), e2020EF001735, https://doi.org/10.1029/2020EF001735, 2021.

Kean, J. W., Staley, D. M., and Cannon, S. H.: In situ measurements of post-fire debris flows in southern California: Comparisons of the timing and magnitude of 24 debris-flow events with rainfall and soil moisture conditions, J. Geophys. Res.-Earth, 116, F04019, https://doi.org/10.1029/2011JF002005, 2011.

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

Lancaster, J. T., Swanson, B. J., Lukashov, S. G., Oakley, N. S., Lee, J. B., Spangler, E. R., Hernandez, J. L., Olson, B. P. E., DeFrisco, M. J., Lindsay, D. N., Schwartz, Y. J., McCrea, S. E., Roffers, P. D., and Tran, C. M.: Observations and analyses of the 9 January 2018 debris-flow disaster, Santa Barbara County, California, Environ. Eng. Geosci., 27(1), 3–27, https://doi.org/10.2113/EEG-D-20-00015, 2021.

Larsen, I. J., MacDonald, L. H., Brown, E., Rough, D., Welsh, M. J., Pietraszek, Libohova, Z., Benavides-Solorio, J. D., and Schaffrath, K.: Causes of post-fire runoff and erosion: Water repellency, cover, or soil sealing?, Soil Sci. Soc. Am. J., 73(4), 1393–1407, https://doi.org/10.2136/sssaj2007.0432, 2009.

Lee, K. H., Uchida, T., and Seo, J. P.: Monitoring of post-fire bedload transport using hydrophone in a small burnt catchment, South Korea, Forests, 13(11), 1774, https://doi.org/10.3390/f13111774, 2022.

Lourenço, L., Nunes, A. N., Bento-Gonçalves, A., and Vieira, A.: Soil erosion after wildfires in Portugal: What happens when heavy rainfall events occur?, in: Research on Soil Erosion, edited by: Godone, D., and Stanchi, S., InTechOpen, London, United Kingdom, 65–88, https://doi.org/10.5772/50447, 2012.

MacKay, D. J. C.: Information Theory, Inference, and Learning Algorithms, Cambridge University Press, Cambridge, England, 640 pp., ISBN 9780521642989, 2003.

McCoy, K., Krasko, V., Santi, P., Kaffine, D., and Rebennack, S.: Minimizing economic impacts from post-fire debris flows in the western United States, Nat. Hazards, 83(1), 149–176, https://doi.org/10.1007/s11069-016-2306-0, 2016.

McGuire, L. A., and Youberg, A. M.: What drives spatial variability in rainfall intensity-duration thresholds for post-wildfire debris flows? Insights from the 2018 Buzzard Fire, NM, USA, Landslides, 17, 2385–2399, https://doi.org/10.1007/s10346-020-01470-y, 2020.

McGuire, L. A., Rengers, F. K., Kean, J. W., and Staley, D. M.: Debris flow initiation by runoff in a recently burned basin: Is grain-by-grain sediment bulking or en masse failure to blame?, Geophys. Res. Lett., 44(14), 7310–7319, https://doi.org/10.1002/2017GL074243, 2017.

Moody, J. A., Smith, J. D., and Ragan, B. W.: Critical shear stress for erosion of cohesive soils subjected to temperatures typical of wildfires, J. Geophys. Res.-Earth, 110(F1), F01004, https://doi.org/10.1029/2004JF000141, 2005.

Morell, K. D., Alessio, P., Dunne, T., and Keller, E.: Sediment recruitment and redistribution in mountain channel networks by post-wildfire debris flows, Geophys. Res. Lett., 48(24), e2021GL095549, https://doi.org/10.1029/2021GL095549, 2021.

Murphy, A. H.: What is a good forecast? An essay on the nature of goodness in weather forecasting, Weather Forecast., 8(2), 281–293, https://doi.org/10.1175/1520-0434(1993)008%3C0281:WIAGFA%3E2.0.CO;2, 1993.

Murphy, A. H., and Winkler, R. L.: A general framework for forecast verification, Mon. Weather Rev., 115(7), 1330–1338, https://doi.org/10.1175/1520-0493(1987)115%3C1330:AGFFFV%3E2.0.CO;2, 1987.

Nyman, P., Sheridan, G. J., Smith, H. G., and Lane, P. N. J.: Evidence of debris flow occurrence after wildfire in upland catchments of south-east Australia, Geomorphology, 125(3), 383–401, https://doi.org/10.1016/j.geomorph.2010.10.016, 2011.

Nyman, P., Sheridan, G. J., Moody, J. A., Smith, H. G., Noske, P. J. and Lane, P. N.: Sediment availability on burned hillslopes, J. Geophys. Res.-Earth, 118(4), 2451–2467, https://doi.org/10.1002/jgrf.20152, 2013.

Oakley, N. S., Cannon, F., Munroe, R., Lancaster, J. T., Gomberg, D., and Ralph, F. M.: Brief communication: Meteorological and climatological conditions associated with the 9 January 2018 post-fire debris flows in Montecito and Carpinteria, California, USA, Nat. Hazards Earth Syst. Sci., 18, 3037–3043, https://doi.org/10.5194/nhess-18-3037-2018, 2018.

Oakley, N. S., Liu, T., McGuire, L. A., Simpson, M., Hatchett, B. J., Tardy, A., Kean, J. W., Castellano, C., Laber, J. L., and Steinhoff, D.: Toward probabilistic post-fire debris-flow hazard decision support, B. Am. Meteorol. Soc., Advance online publication, https://doi.org/10.1175/BAMS-D-22-0188.1, 2023.

Pianosi, F., and Wagener, T.: Distribution-based sensitivity analysis from a generic input-output sample, Environ. Modell. Softw., 108, 197–207, https://doi.org/10.1016/j.envsoft.2018.07.019, 2018.

Pirulli, M.: On the use of the calibration-based approach for debris-flow forward-analyses, Nat. Hazards Earth Syst. Sci., 10, 1009–1019, https://doi.org/10.5194/nhess-10-1009-2010, 2010.

Prescott, A. B., McGuire, L. A., Jun, K.-S., Barnhart, K. R., and Oakley, N. S.: Online repository for code and data used in: Probabilistic assessment of postfire debris-flow inundation in response to forecast rainfall, Zenodo [dataset][code], https://doi.org/10.5281/zenodo.7838914, 2023.

Razavi, S., and Gupta, H. V.: What do we mean by sensitivity analysis? The need for comprehensive characterization of "global" sensitivity in Earth and Environmental systems models, Water Resour. Res., 51(5), 3070–3092, https://doi.org/10.1002/2014WR016527, 2015.

https://doi.org/10.5194/egusphere-2023-1931
Preprint. Discussion started: 18 September 2023

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S.: Global Sensitivity Analysis: The Primer, John Wiley & Sons, West Sussex, England, 304 pp. https://doi.org/10.1002/9780470725184, 2008.

Santi, P. M., and Morandi, L.: Comparison of debris-flow volumes from burned and unburned areas, Landslides, 10, 757–769, https://doi.org/10.1007/s10346-012-0354-4, 2013.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D., and Huang, X.-Y.: A description of the advanced research WRF model version 4.3, National Center for Atmospheric Research, Boulder, Colorado, Technical Notes NCAR/TN-556+STR, 162 pp., https://doi.org/10.5065/1dfh-6p97, 2021.

Staley, D. M., Kean, J. W., Cannon, S. H., Schmidt, K. M., and Laber, J. L.: Objective definition of rainfall intensity–duration thresholds for the initiation of post-fire debris flows in southern California, Landslides, 10, 547–562, https://doi.org/10.1007/s10346-012-0341-9, 2013.

Staley, D. M., Negri, J. A., Kean, J. W., Laber, J. M., Tillery, A. C., and Youberg, A. M.: Updated logistic regression equations for the calculation of post-fire debris-flow likelihood in the western United States, U.S. Geological Survey Open-File Rep. 2016–1106, 13 pp., https://doi.org/10.3133/ofr20161106, 2016.

Staley, D. M., Negri, J. A., Kean, J. W., Laber, J. L., Tillery, A. C., and Youberg, A. M.: Prediction of spatially explicit rainfall intensity–duration thresholds for post-fire debris-flow generation in the western United States, Geomorphology, 278, 149–162, https://doi.org/10.1016/j.geomorph.2016.10.019, 2017.

Staley, D. M., Tillery, A. C., Kean, J. W., McGuire, L. A., Pauling, H. E., Rengers, F. K., and Smith, J. B.: Estimating post-fire debris-flow hazards prior to wildfire using a statistical analysis of historical distributions of fire severity from remote sensing data, Int. J. Wildland Fire, 27(9), 595–608, https://doi.org/10.1071/WF17122, 2018.

Stoof, C. R., Vervoort, R. W., Iwema, J., Van Den Elsen, E., Ferreira, A. J. D., and Ritsema, C. J.: Hydrological response of a small catchment burned by experimental fire, Hydrol. Earth Syst. Sc., 16(2), 267–285, https://doi.org/10.5194/hess-16-267-2012, 2012.

Sun, X., Zeng, P., Li, T., Wang, S., Jimenez, R., Feng, X., and Xu, Q.: From probabilistic back analyses to probabilistic run-out predictions of landslides: A case study of Heifangtai terrace, Gansu Province, China, Eng. Geol., 280, 1105950, https://doi.org/10.1016/j.enggeo.2020.105950, 2021.

Tillery, A. C., Haas, J. R., Miller, L. W., Scott, J. H., and Thompson, M.P.: Potential postwildfire debris-flow hazards: a prewildfire evaluation for the Sandia and Manzano Mountains and surrounding areas, central New Mexico, U.S. Geological Survey Sci. Invest. Rep. 2014–5161, 24 pp., http://doi.org/10.3133/sir20145161, 2014.

Touge, Y., Hasegawa, M., Minegishi, M., Kawagoe, S., and Kazama, S.: Multitemporal UAV surveys of geomorphological changes caused by postfire heavy rain in Kamaishi city, northeast Japan, Catena, 220A, 106702, https://doi.org/10.1016/j.catena.2022.106702, 2023.

U.S. Forest Service: Thomas Fire burned area report, U.S. Department of Agriculture, Technical Report FSH 2509.13, 56 pp., https://www.fs.usda.gov/detail/lpnf/home/?cid=fseprd570093 (last access: 22 August 2023), 2020.

https://doi.org/10.5194/egusphere-2023-1931
Preprint. Discussion started: 18 September 2023

U.S. Geological Survey: National Elevation Dataset, https://www.usgs.gov/the-national-map-data-delivery (last access: 22 August 2023), 2020.

Wall, S., Murphy, B. P., Belmont, P., and Yocom, L.: Predicting post-fire debris flow grain sizes and depositional volumes in the Intermountain West, United States, Earth Surf. Proc. Land., 48(1), 179–197, https://doi.org/10.1002/esp.5480, 2023.

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, Fourth Edition, Elsevier, Amsterdam, Netherlands, 818 pp., https://doi.org/10.1016/C2017-0-03921-6, 2019.

Zeng, P., Zhang, L., Li, T., Sun, X., Zhao, L, Dong, X., and Xu, Q: Constructing a region-specific rheological parameter database for probabilistic run-out analyses of loess flowslides, Landslides, 20, 1167–1185, https://doi.org/10.1007/s10346-023-02037-3, 2023.

https://doi.org/10.5194/egusphere-2023-1931
Preprint. Discussion started: 18 September 2023

Figures

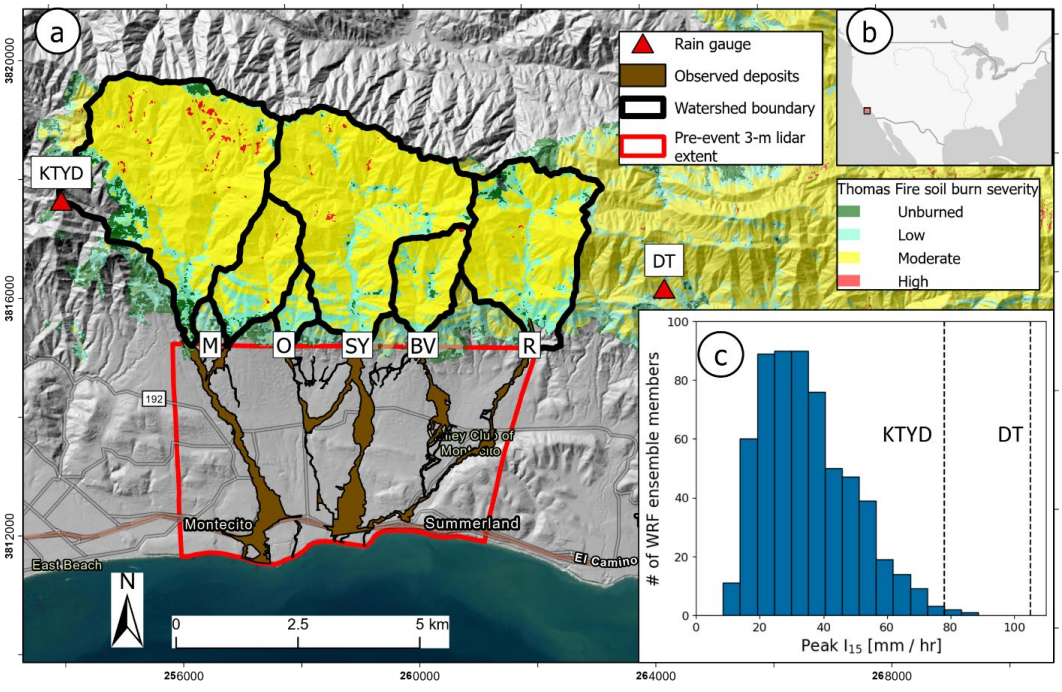


Figure 1: (a) Site map of the 9 January 2018 Montecito debris-flow event. Burned basins drain into (M) Montecito, (O) Oak, (SY) San Ysidro, (BV) Buena Vista, and (R) Romero Creeks (note that 2 burned basins drain into Montecito Creek); (b) Site location in context of the USA; (c) histogram of peak I_{15} extracted from the atmospheric model ensemble. Dashed lines show observations at the KTYD and Doulton Tunnel (DT) rain gauges. Ticks along the boundaries of (a) give coordinates in NAD 1983 UTM zone 11N. In this figure and in all following maps, the base map was sourced from ESRI and the U.S. Department of Agriculture Farm Services Agency, the hillshade layer was generated from the 10-m resolution National Elevation Dataset (U.S. Geological Survey, 2020), and the soil burn severity layer was sourced from U.S. Forest Service (2020).

https://doi.org/10.5194/egusphere-2023-1931
Preprint. Discussion started: 18 September 2023

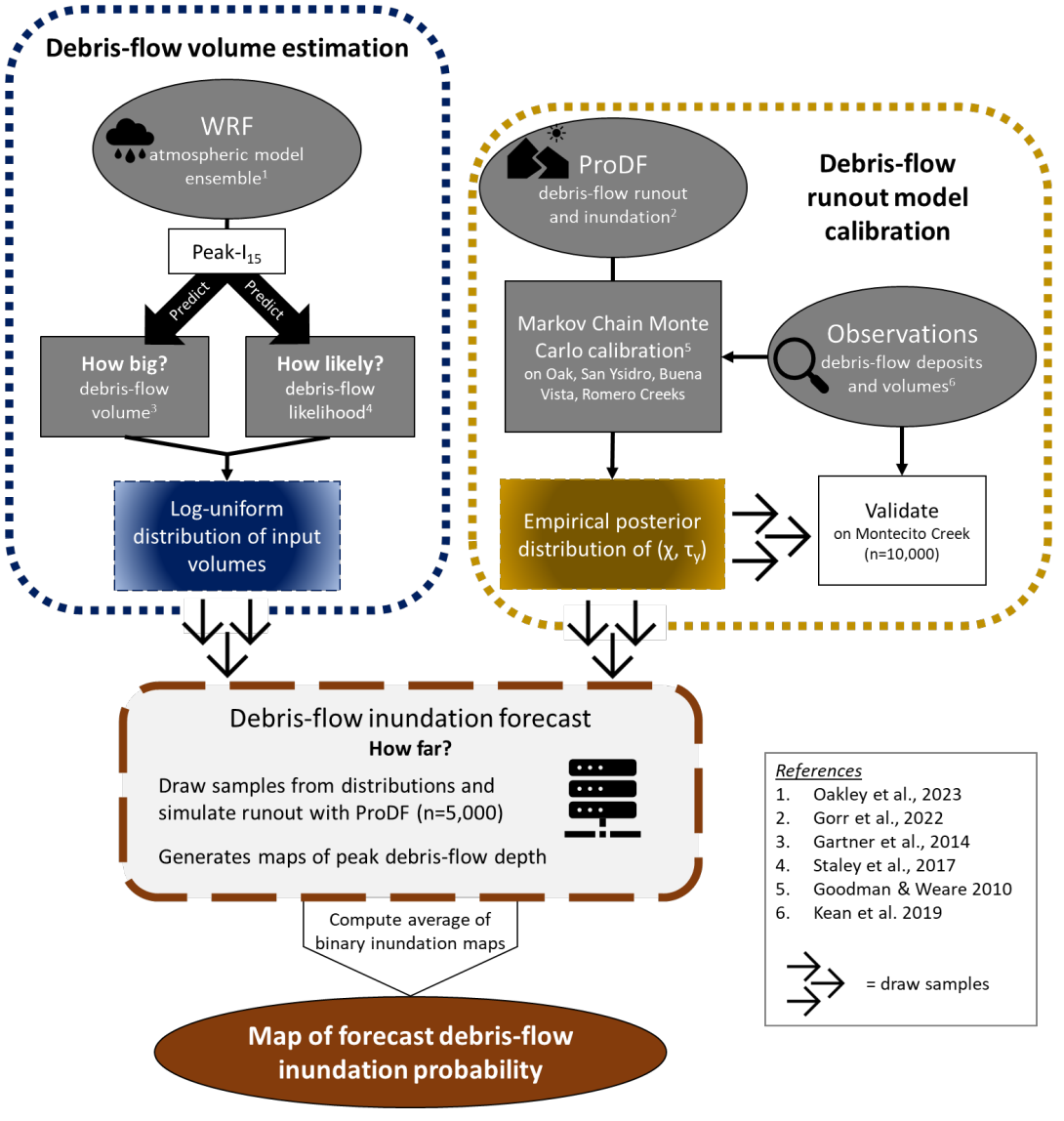
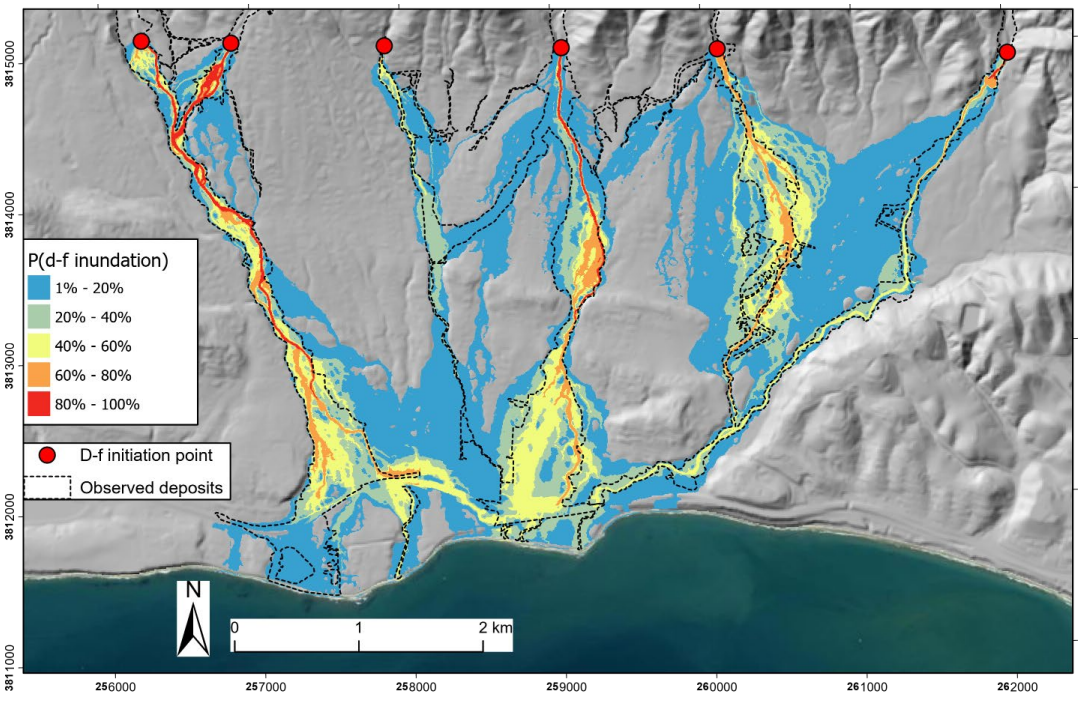


Figure 2: Schematic of the probabilistic debris-flow inundation forecast model.

https://doi.org/10.5194/egusphere-2023-1931
Preprint. Discussion started: 18 September 2023



545 **Figure 3: Map of forecast debris-flow (d-f) inundation probabilities (P) ≥1% with dashed lines showing the extent of observed debris-flow deposits. Ticks along the boundaries of map give coordinates in NAD 1983 UTM zone 11N.**

https://doi.org/10.5194/egusphere-2023-1931
Preprint. Discussion started: 18 September 2023

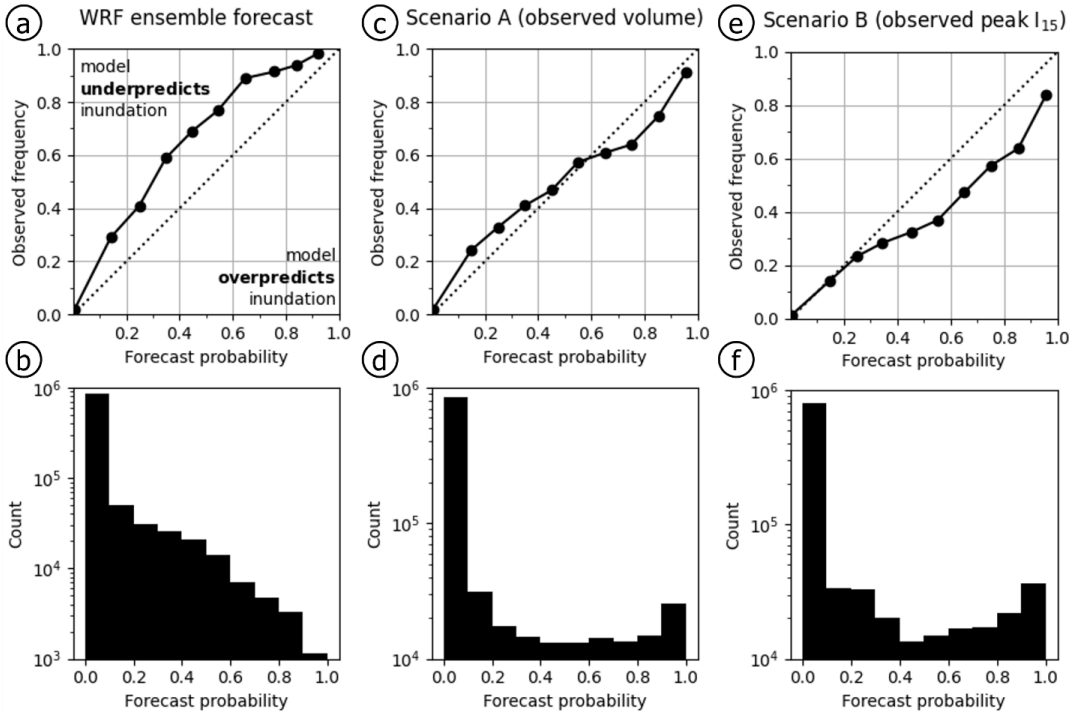
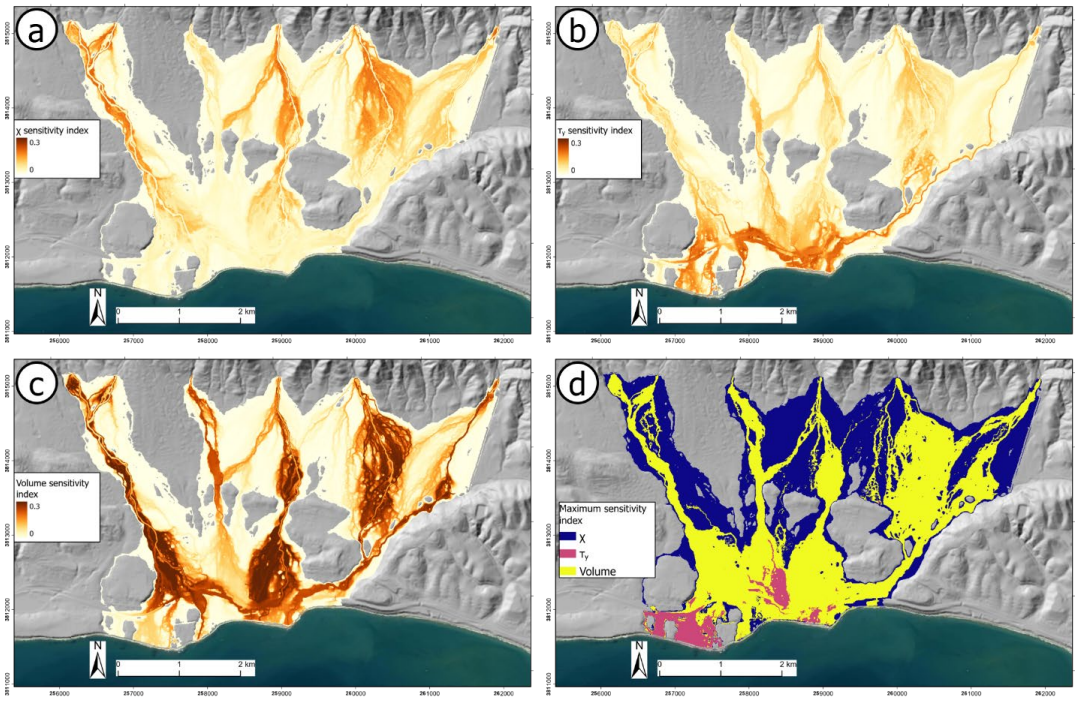


Figure 4: Two-part reliability diagrams showing the calibration-refinement factorization of the joint distribution of forecasts and observations (Section 3.6) from three separate inundation simulations: (a and b) the WRF ensemble forecast, producing the forecast inundation map in Fig. 3; (c and d) scenario A, simulated using the constant observed volumes (Kean et al., 2019), (e and f) scenario B, simulated using the constant EAV predicted volumes computed using observed peak I_{15} (Kean et al., 2019). The top row shows the calibration curve of the conditional distribution. A perfectly calibrated probabilistic model will lie along the 1:1 line, and points above (below) the diagonal indicate that the model is under-forecasting (over-forecasting) the observed frequency of inundation. The histograms in the bottom row show the refinement distribution of the forecast probabilities. A forecast that is sharp will have the highest density of probabilities toward the extreme values of 0 or 1.

https://doi.org/10.5194/egusphere-2023-1931
Preprint. Discussion started: 18 September 2023



560 **Figure 5: Spatially distributed PAWN sensitivity indices for forecast inundation against the model parameters (a) χ; (b) τ_y; and (c) debris-flow volume; and (d) the parameter that resulted in the maximum sensitivity index in each grid cell. Depending on location, each of the input model parameters may be the most influential. Ticks along the boundaries of each map give coordinates in NAD 1983 UTM zone 11N.**