8-17-2024

Estimating Post-Fire Flood Infrastructure Clogging and Overtopping Hazards

Ariane Jong-Levinger

Douglas Houston

Brett F. Sanders

Estimating Post-Fire Flood Infrastructure Clogging and Overtopping Hazards

Water Resources Research

RESEARCH ARTICLE
10.1029/2023WR036522

Estimating Post-Fire Flood Infrastructure Clogging and Overtopping Hazards

Ariane Jong-Levinger[1,2], Douglas Houston[3], and Brett F. Sanders[1,3]

[1]Department of Civil and Environmental Engineering, University of California, Irvine, Irvine, CA, USA, [2]Schmid College of Science and Technology, Chapman University, Orange, CA, USA, [3]Department of Urban Planning and Public Policy, University of California, Irvine, Irvine, CA, USA

Key Points:

- Flood risks are heightened by clogging of infrastructure with sediment, which can occur from sequences of storms especially after wildfires
- A framework for calibration and validation of a post-fire infrastructure clogging and flood hazard model is presented
- Model applications reveal whether infrastructure is adequately sized to meet design levels of protection

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. Jong-Levinger,
jonglevinger@chapman.edu

Citation:

Jong-Levinger, A., Houston, D., & Sanders, B. F. (2024). Estimating post-fire flood infrastructure clogging and overtopping hazards. *Water Resources Research*, *60*, e2023WR036522. https://doi.org/10.1029/2023WR036522

Received 30 OCT 2023
Accepted 3 AUG 2024

Abstract Cycles of wildfire and rainfall produce sediment-laden floods that pose a hazard to development and may clog or overtop protective infrastructure, including debris basins and flood channels. The compound, post-fire flood hazards associated with infrastructure overtopping and clogging are challenging to estimate due to the need to account for interactions between sequences of wildfire and storm events and their impact on flood control infrastructure over time. Here we present data sources and calibration methods to estimate infrastructure clogging and channel overtopping hazards on a catchment-by-catchment basis using the Post-Fire Flood Hazard Model (PF2HazMo), a stochastic modeling approach that utilizes continuous simulation to resolve the effects of antecedent conditions and system memory. Publicly available data sources provide parameter ranges needed for stochastic modeling, and several performance measures are considered for model calibration. With application to three catchments in southern California, we show that PF2HazMo predicts the median of the simulated distribution of peak bulked flows within the 95% confidence interval of observed flows, with an order of magnitude range in bulked flow estimates depending on the performance measure used for calibration. Using infrastructure overtopping data from a post-fire wet season, we show that PF2HazMo accurately predicts the number of flood channel exceedances. Model applications to individual watersheds reveal where infrastructure is undersized to contain present-day and future overtopping hazards based on current design standards. Model limitations and sources of uncertainty are also discussed.

Plain Language Summary Communities at the foot of the mountains face an especially dangerous type of flooding called "sediment-laden floods." Many such communities in the southwestern U.S. are protected from water floods by flood infrastructure designed to trap sediment at the mouth of mountain canyons and convey only water flows safely past developed areas to a downstream water body. Sediment-laden floods, which are more forceful and typically larger than water floods, are more likely to happen during storms over burned mountain canyons soon after a wildfire occurs. However, estimating the likelihood that sediment-laden floods fill and overtop flood infrastructure is challenging since existing sediment-laden flood models do not explicitly consider the role of flood infrastructure. Here we present the Post-Fire Flood Hazard Model (PF2HazMo), a model that can estimate the likelihood of post-fire floods on a canyon-by-canyon basis accounting for flood infrastructure. Environmental data collected following a major wildfire is used to apply PF2HazMo to three mountain canyons in southern California, and we find that it predicts the number of floods accurately relative to observed post-fire flood channel overtopping events. Further, the model is used to predict the frequency of floods due to infrastructure overtopping under both present-day and future wildfire scenarios.

1. Introduction

Cycles of wildfire and rainfall in mountainous areas produce sediment-laden flows that pose significant hazards to downstream development, including floods, mudflows, mudslides, and debris flows (e.g., Cannon & DeGraff, 2009; DiBiase & Lamb, 2020; Hamilton et al., 1954). Across the southwestern U.S., and especially in southern California, urban development encroaching upon the base of mountains has required protection from these hazards, and has utilized a combination of debris retention basins (hereafter referred to as "debris basins") to capture sediment and flood channels to convey streamflow safely past communities (Jennings et al., 1982; Johnson et al., 1991; Scott et al., 1978). In Los Angeles County alone, there are more than 150 debris basins providing protection to mountainside communities such as Pasadena, Azusa, and Arcadia (LACDPW, 2013). The use of debris basins for flood protection began following the 1933 New Year's Eve flood, during which mud and debris flowing from the San Gabriel Mountains destroyed more than 400 homes in La Cañada and La Crescenta

and caused numerous fatalities (Cobery, 2012); the practice is now widespread across the region, and debris basins or similar sediment and debris retention structures have been used for flood management in the southwestern U.S., Japan, Italy (the Alps), and Canada (Gartner et al., 2014; Gusman et al., 2011; Johnson et al., 1991). Debris basins are generally sized to contain the debris of one extreme event, or numerous smaller events, and to route flood water (minus sediment and debris) safely past development using flood channels (Johnson & McCuen, 1992; Johnson et al., 1991; Prochaska et al., 2008; Sanders & Grant, 2020; Wallerstein et al., 1997; Willardson, 2020). However, the hazard level facing communities protected by this infrastructure is difficult to quantify. Hazards include the possibility of debris basin and/or flood channel overtopping with flows of flood water, sediment, and debris made more likely by the infrastructure clogging with sediment and debris over preceding storm events. Moreover, recent exceedances have been deadly and costly. A debris flow event in Montecito, CA in 2018 resulted in 23 deaths and over 400 damaged homes (Kean et al., 2019); even 5 years later, debris flow risk was still a concern when major storms filled a debris retention basin constructed in one of the flow paths of the 2018 debris flows completely with sediment (Orozco, 2023).

Quantitative hazard estimates are a foundation of risk management, including efforts to identify, communicate, and reduce risks through infrastructure improvements. Significant progress has been made toward estimating the spatial distribution of pluvial, fluvial, and coastal flood hazards at regional (e.g., Sanders et al., 2023) and national scales (e.g., Bates et al., 2021), including improved understanding of the exposed populations and inequities (e.g., Tate et al., 2021). Progress is similarly needed to estimate sediment-laden flood hazards. Here we adopt the expression *sediment-laden flood hazards* to capture a wide range of hazardous flows across a spectrum of sediment and debris concentrations (Gartner et al., 2014). Expressions such as *hyperconcentrated flows*, *mud flows*, *debris floods*, and *debris flows* have been used to differentiate specific types of sediment-laden flow based on sediment concentration, grain size distribution, and the relative proportion of suspended load versus bedload transport (e.g., Brenna et al., 2020; Church & Jakob, 2020; Jakob et al., 2005; O'Brien, 1986). The expression *ultrahazardous flooding* is also applicable. It was introduced in the context of alluvial fans (National Research Council, 1996), a specific type of geography found in near-mountain environments, but this expression is equally relevant to a wide range of near-mountain geographies with steep slopes and fast-moving, sediment-laden flood hazards prone to unpredictable flow paths (Jong-Levinger et al., 2022; Sanders & Grant, 2020).

Considerable previous research has focused on estimating plot- to catchment-scale post-fire runoff and erosion rates, which contribute to sediment-laden flood hazards. Catchment-scale post-fire runoff models typically estimate peak discharge at the catchment outlet; model types include empirical (e.g., Hawkins & Munoz, 2011; Rowe et al., 1949; USDA, 2009), semi-empirical/semi-distributed such as HEC-HMS (USACE, 2022; Wang et al., 2020), and physically based models (Flanagan & Nearing, 1989; Van Eck et al., 2016). Studies comparing empirical, semi-distributed, and process-based models of post-fire peak discharge have not provided clear evidence for a single approach that consistently performs well across multiple study sites (L. Chen et al., 2013; Kinoshita et al., 2014). Post-fire erosion models often estimate sediment yields (sediment volume per contributing area), which are largely calculated using empirical methods derived from regression analysis (e.g., Gartner et al., 2014; Pak & Lee, 2008; Pelletier & Orem, 2014; Renard, 1997; Scott et al., 1978; Wagenbrenner & Robichaud, 2014; Wischmeier et al., 1965), though a smaller number of process-based models exist (Flanagan & Nearing, 1989; Goodrich et al., 2012; Kirkby et al., 2008). Comparison studies of empirical and process-based post-fire erosion models have found that watershed-scale model estimates span three orders of magnitude and that current models tend to consistently identify areas of low and high potential, but need better calibration for improved accuracy (Kampf et al., 2020; Rengers et al., 2021). Further, these studies present cases in which estimates from empirical models were better than those from processed-based models when compared to data from the region to which they were calibrated; this supports the conclusion that currently, no single modeling framework estimates post-fire erosion well across multiple regions. Moreover, existing frameworks do not consider how sediment-laden floods may be attenuated or exacerbated by the presence of flood control infrastructure, as only recently has attention been given to the role of infrastructure in shaping downstream hazards to development.

Consideration of event sequences is crucial to estimating sediment-laden flood hazards because debris basins and flood channels may fill with sediment over numerous rainfall events which, in turn, impact cycles of wildfire, soil and vegetation recovery, and flood control maintenance activities such as debris basin and flood channel cleaning. Indeed, sediment-laden flood hazards represent a compound hazard produced by sequences of events and the interaction of multiple hazard drivers (Sadegh et al., 2018). The importance of event sequences motivates hazard

Table 1
PF2HazMo Parameters

Parameter name	Symbol	Definition
Watershed Area	A	Drainage area of mountain canyon
Runoff Response	c_t	Ratio of the runoff coefficient to the time to peak
Time to Peak	t_p	Duration from beginning of storm to peak runoff
Fire Interval	t_f	Recurrence interval of wildfires in years
Post-fire Bulking Factor	k_1	Ratio of the combined water and sediment volumetric flux to the water volumetric flux
Recovery Timescale	t_r	Time scale for watershed recovery after fire
Debris Basin Design Capacity	V_b^{des}	Storage capacity of sediment/debris basin
Flood Channel Design Capacity	D_c^{des}	Flow capacity of flood channel downstream of debris basin
Infrastructure Cleaning Threshold	V_{thresh}	Fraction filled that triggers cleaning
Wet Season Cleaning Waiting Period	w	Number of days with no rain needed before debris basin cleaning can commence
Wet Season Debris Basin Cleaning Rate	r	Daily excavation rate for debris basin during wet season

estimation by continuous simulation, that is, where physical processes are simulated over time (e.g., hourly, daily) with a hydrologic model, the effects of antecedent conditions and system memory are explicitly resolved, and the statistics of extremes (e.g., number of exceedances per century) are deduced from long records of model outputs (Boughton & Droop, 2003; Pathiraja et al., 2012; Winter et al., 2019).

The Post-Fire Flood Hazard Model (PF2HazMo) is a continuous simulation model that was recently introduced by Jong-Levinger et al. (2022) to estimate post-fire flood infrastructure clogging and overtopping hazards, and it is a lumped model with 11 parameters (Table 1) that require estimation on a catchment-by-catchment basis. PF2HazMo solves a mass balance equation to account for the build-up of sediment within debris basins, and it builds upon the engineering practice of sediment flux modeling using the dimensionless bulking factor, k, which represents the ratio of the bulked sediment flow to the clear-water flow. Furthermore, numerous model parameters (including the bulking factor, fire occurrence, and fire severity) are treated stochastically to embrace the high level of uncertainty associated with post-fire sediment yields and possible sequences of events (Gartner et al., 2014; Kampf et al., 2020; Rengers et al., 2021), and the model is forced by a daily precipitation time series from a Monte Carlo Markov Chain (MCMC) rainfall simulator. Wildfire events are assumed to increase the bulking factor temporarily, and thus higher sediment fluxes occur from rainfall events within a few months after wildfire events compared to rainfall events that occur several years or more after a wildfire event. Finally, by comparing the magnitude of flood peaks to channel capacity on a daily basis, the frequency of channel overtopping is predicted and is taken as an estimate of the sediment-laden flood hazard. We note that while the lumped modeling approach substantially simplifies highly complex sediment generation and transport processes associated with sediment-laden flows, the adoption of a stochastic, continuous-simulation approach allows for consideration of an exhaustive set of event sequences which are arguably equally, if not more, important sources of uncertainty relative to infrastructure clogging and overtopping.

Figure 1 illustrates the functionality of PF2HazMo, namely the ability to simulate a sequence of states that contribute to infrastructure clogging and overtopping hazards. The initial state is a healthy, vegetated watershed with infrastructure at full capacity (Figure 1a). The second state (Figure 1b) captures changes to watershed properties from wildfire, which contribute to heightened flows of runoff and sediment, and the accumulation of sediment in the debris basin. The third state (Figure 1c) captures overtopping of the debris basin (with sediment) and/or the flood channel (with flood flows) from subsequent rainfall events. Finally, the fourth state (Figure 1d) captures the system following storm events or between storm events, when infrastructure capacity is restored with cleaning, and vegetation recovers over time. By using a stochastic modeling approach, thousands of different sequences are simulated over a long period of time (e.g., 100 years), and a probabilistic estimation of flood hazard levels is provided including the number of debris basin overtopping events per century and the number of channel overtopping events per century.

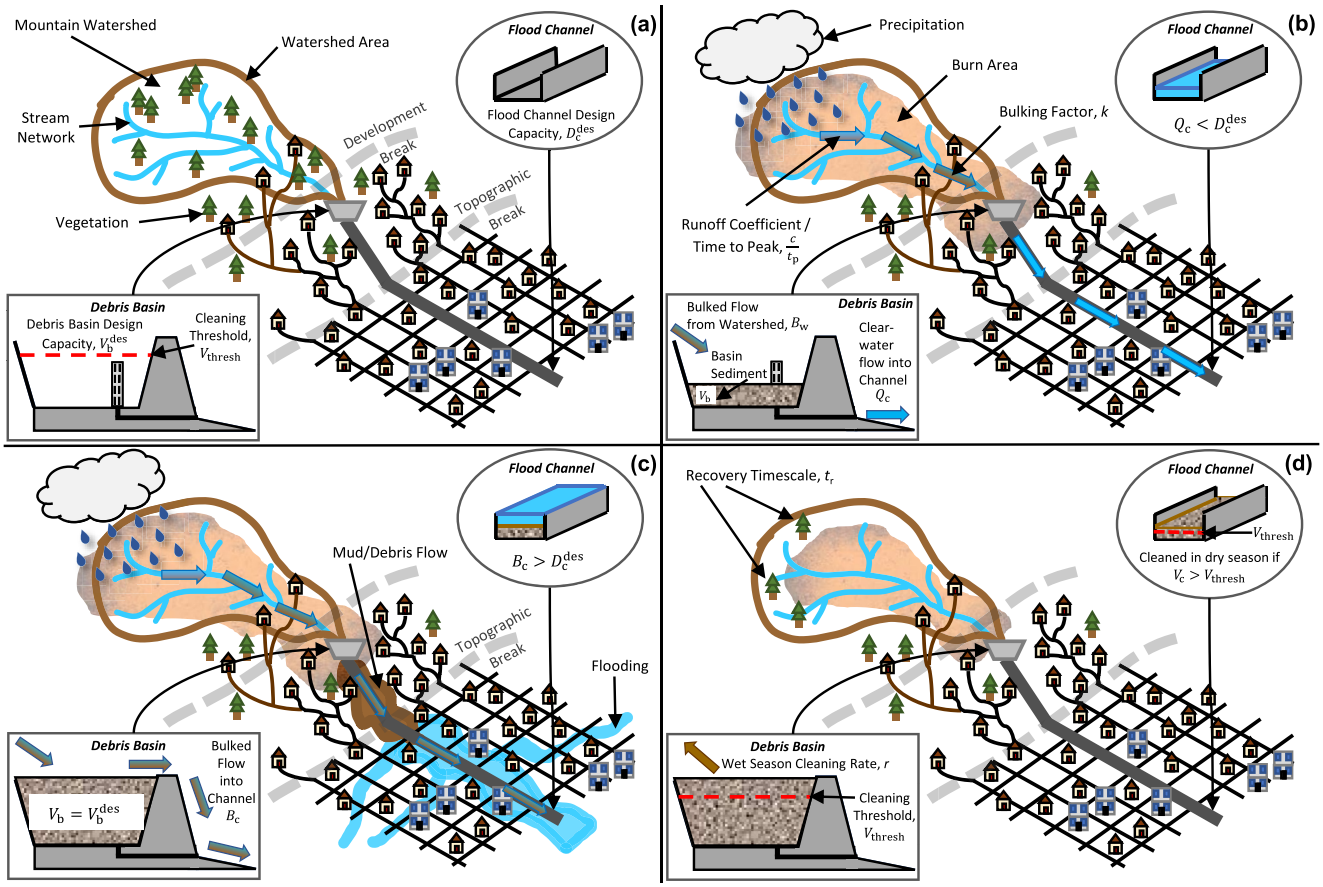

Figure 1. Progression of system states captured by Post-Fire Flood Hazard Model (PF2HazMo). (a) Pre-fire, clean infrastructure state, (b) Post-fire wet-weather state, (c) Post-fire wet-weather state with infrastructure overtopping, (d) Post-storm recovery state.

Application of PF2HazMo to estimate the spatial distribution of hazards requires parameter estimation (Table 1) on a catchment-by-catchment basis using available data, including calibration. In support of this need, the objective of this paper is to present a parameterization and calibration methodology that draws upon publicly available data and to estimate the uncertainty in hazard estimates through validation against observations. Hence, we demonstrate herein the estimation of parameters required for application of PF2HazMo to estimate post-fire flood infrastructure clogging and overtopping hazards and quantify uncertainties in these hazard estimates.

The remainder of the paper is organized as follows: Section 2 (Methods) describes the data collected and methods used to conduct the model sensitivity and error analysis, estimate the model parameter distributions, calibrate the model to specific watersheds, and evaluate the performance of the model. Section 3 (Results) compares simulation results after calibrating PF2HazMo to observed post-fire peak flows or to probabilistic precipitation-frequency discharge-frequency curves; simulated post-fire infrastructure exceedances are then compared to observed overtopping events. Section 4 (Discussion) compares our modeling approach and results to those of other models of post-fire hydrology and discusses how the modeling framework can be used to inform flood risk management. Section 5 presents conclusions.

2. Methods

The methodological sections that follow are organized into six subsections that cover: (a) a description of the study catchments and the data used; (b) model sensitivity and error propagation analysis; (c) estimation of parameter values for uncalibrated parameters; (d) model calibration; and (e) model validation.

Model calibration (Section 2.4) targets a runoff response parameter which is found through sensitivity analysis (Section 2.2) to be particularly important for constraining uncertainty, while model validation (Section 2.5)

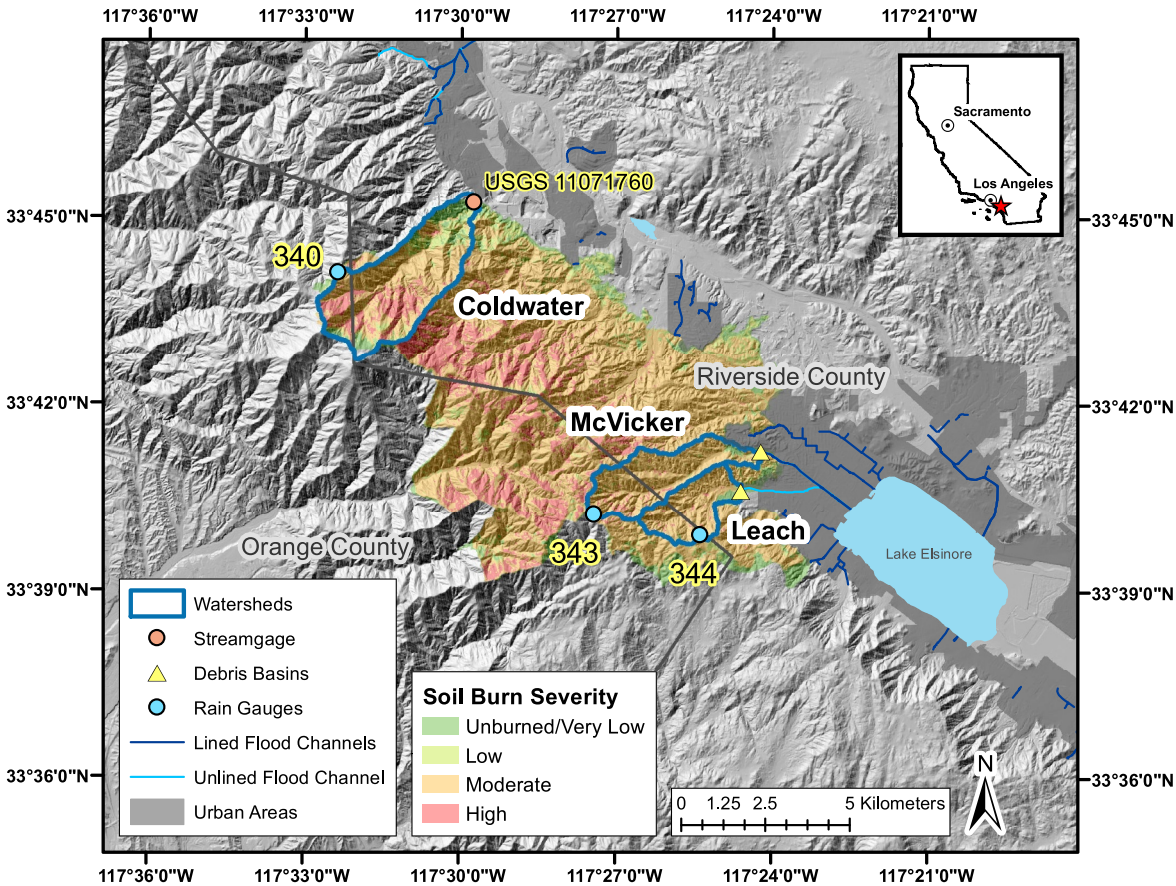

Figure 2. Map of study area showing study watersheds, burn severity for 2018 Holy Fire, and locations of rain gauges, streamgage, and flood control infrastructure.

targets the hazard variables predicted by the model, namely debris basin exceedances and flood channel exceedances.

2.1. Study Sites and Data

A set of three small mountain watersheds (<11 km^2) in Riverside County, California is utilized in this study to present methods of parameter estimation, calibration, and validation for PF2HazMo (Figure 2). These catchments were burned by the 2018 Holy Fire and have similar percentages of area burned, average slope, and aspect (Table 3). The Coldwater watershed represents a gauged watershed with a U.S. Geological Survey (USGS) streamgage at its outlet; however, there is no flood control infrastructure located downstream. McVicker and Leach watersheds are both ungauged (referring to the lack of streamflow monitoring equipment; however, these sites are both equipped with rain gauges), and each has a debris basin and flood channel located downstream. McVicker is upstream of a concrete-lined channel, while Leach is upstream of an unlined, or soft-bottom, flood channel, allowing us to compare how the model performs for different channel types. Further, all three basins are located upstream of urban areas (dark gray areas in Figure 2, based on block-level population density data from the 2020 U.S. Census), and hence estimation of post-fire flood and debris flow risks is important for public safety.

We note that gage records from the Coldwater watershed are used to calibrate the runoff response parameter (Table 1) which, in turn, is applied to validate both the McVicker and Leach watershed models where exceedance data is available. Gage data is rarely available in all of the basins where hazard information is needed, and thus the parameterization approach presented here demonstrates a pragmatic pathway for leveraging available data to apply the modeling framework to several adjacent watersheds.

2.1.1. Environmental Setting

The study watersheds are located in the Santa Ana Mountains of southern California within the Santa Ana River watershed, which drains into the Pacific Ocean. The Santa Ana Mountains are part of the Peninsular Ranges, a group of mountain ranges oriented from north to west that are characterized by Jurassic and Cretaceous plutonic, mainly granitic rock that is more resistant to erosion than the neighboring Transverse Ranges to the north (Inman & Jenkins, 1999). In particular, the underlying lithology of the study sites is Mesozoic to Precambrian metamorphic rocks intruded by granitic rocks as well as Mesozoic volcanic rocks (California Geological Survey, 2018). The climate is classified as hot-summer Mediterranean to hot semi-arid according to the Köppen climate classification (Beck et al., 2018). Average annual precipitation in the region is about 451 mm, with most of it falling as rain from October to April (PRISM Climate Group, 2023). The vegetation is Chaparral-dominated; indicative vegetation types include Diegan coastal sage scrub, mesic chaparral, and sparse coniferous vegetation (RCEPD, 2003; Wilder & Kinoshita, 2022).

The mean fire return interval, or the average time between fires under the presumed historical fire regime, for these watersheds is approximately 40 years (LANDFIRE, 2014). The 2018 Holy Fire was the first major wildfire to affect these watersheds (i.e., to burn >10% of the watershed area) since 1987 for the Coldwater watershed and since 1954–1956 for the Leach and McVicker watersheds, according to historical fire perimeter data (CAL FIRE, 2021). On 6 August 2018, the Holy Fire ignited near Trabuco Canyon in the Santa Ana Mountains and burned 9,300 ha by the time it was contained on 17 October 2018 (CAL FIRE, 2021). According to conversations with Riverside County Flood Control and Water Conservation District (hereafter referred to as "Riverside County"), mud and debris flows were observed within the headwater catchments of the study watersheds during the 2018–2019 wet season.

2.1.2. Rainfall and Runoff Data

Rainfall data from three tipping bucket rain gauges operated by Riverside County were acquired to force PF2HazMo simulations of the 2018–2019 wet season. Precipitation data was available for the Coldwater rain gauge (340, Figure 2) from 19 November 2018 to 1 April 2019, while data availability for the McVicker (343) and Leach (344) rain gauges was from 1 November 2018 to 1 March 2021. The rain gauges recorded precipitation depth at irregular time intervals in increments of 1.02 mm (0.04 inches) per tip; the raw data was aggregated by day to force model simulations at a daily timestep. Rain gauge data was also used to estimate the time to peak model parameter, t_p, as described in Section 2.2.3.

Instantaneous streamflow data at an interval of 15 min from 4 December 2018 to 31 March 2019 were retrieved for USGS streamgage 11071760, located at the outlet of the Coldwater watershed, and used to calibrate the Runoff Response parameter. Peak flows were identified based on a visual analysis of the maximum daily discharge time series and an understanding of when the major storm events of the wet season occurred from discussions with Riverside County; the total daily precipitation, maximum daily flow rate time series, and observed peak flow rates for Coldwater watershed are shown in Figure S1 in Supporting Information S1. These observed post-fire peak flows were compared to the simulated daily peak flows that occurred on the same dates as the observed flows to calculate model error metrics and calibrate the model.

2.1.3. Overtopping Data

Time-lapse cameras installed by Riverside County at the Leach and McVicker debris basins provided observations of debris basin overtopping that were used to validate the model. Time-lapse footage taken on days with major storm events (29 November 2018, 6 December 2018, 17 January 2019, 2 February 2019, and 14 February 2019) were analyzed. Three (3) overtopping events were observed at Leach debris basin on 6 December 2018, 17 January 2019, and 14 February 2019, while one (1) overtopping event was observed at McVicker debris basin on 14 February 2019. Additionally, drone survey footage of the flood channel downstream of the Leach watershed showed evidence of channel overtopping on 14 February 2019. To the authors' knowledge, no video, photographic, or anecdotal evidence exists for the overtopping of the flood channel downstream of the McVicker watershed during the 2018–2019 wet season.

Table 2
Model Parameter Ranges Used in the Sensitivity and Error Analyses

Parameter name	Symbol	Units	Minimum	Mean/Median	Maximum
Watershed area	A	km^2	0.01	3.08	32.3
Runoff response	c_t	1/s	2.33×10^{-6}	7.52×10^{-4}	1.50×10^{-3}
Time to peak	t_p	hr	0.167	4.90	11.9
Post-fire bulking factor	k_1	Unitless	1.25	2.06	2.86
Recovery timescale	t_r	days	140	637	1,134
Debris basin design capacity	V_b^{des}	m^3	1.09×10^4	1.50×10^5	2.89×10^5
Flood channel design capacity	D_c^{des}	m^3/s	1.94	484	965
Infrastructure cleaning threshold	V_{thresh}	% filled	75	85	95
Wet season cleaning waiting Period	w	days	1	7	14
Wet season debris basin cleaning rate	r	m^3/day	336	1,624	4,826

2.2. Model Sensitivity and Error Propagation

Characterizing the sensitivity of PF2HazMo to the parameters listed in Table 2 and their mutual influence on model errors is important and can be useful for guiding model parameter estimation and calibration strategies. Here, "sensitivity analysis" implies varying the value of each model parameter by a small amount, one at a time, and quantifying the resulting effect on the model output to determine which parameters have the greatest influence on model results. In contrast, "error propagation analysis" implies simultaneously varying all parameters across their respective value ranges and assessing their collective effect on changes in model outputs.

2.2.1. Sensitivity Analysis

Table 2 presents the model parameters included in the sensitivity analysis. Fire interval was not included in the sensitivity or error propagation analyses because the purpose of the tests is to determine the sensitivity of model parameters that could be calibrated to a specific site, but the fire interval is a parameter based on either the historical fire regime or is set to a value representative of future fire regimes and is not a focus of calibration efforts. Additionally, since the model of wildfire occurrence is stochastic, implementing a given fire interval entails differences in the relative timing of wildfires and precipitation between simulation trials, which makes it difficult to distinguish the effect of varying the parameter's value from the effect of varying the sequencing of wildfires and precipitation.

For each parameter except for the wet season debris basin cleaning rate (r) and the peak flow timescale (t_p), the mean parameter estimate was used as a baseline value. Observations were used to define the ranges for the r and t_p parameters, and since the distributions were skewed, we used the median instead of the mean as the baseline. For each parameter, the baseline value was increased or decreased by 10%. Model outputs used to measure sensitivity included the number of channel exceedances, number of debris basin exceedances, sum of sediment flows into the debris basin, sum of sediment flows into the channel, sum of overbank flows, and maximum overbank flows over the course of a 1-year simulation period. These six output variables were chosen to capture both hazard variables and sediment volumes that bear on management costs.

For model forcing, a 1-year daily time series of precipitation was stochastically generated using the MCMC rainfall simulator described in Jong-Levinger et al. (2022); the same precipitation time series was used in all simulation trials for consistency. The simulation starts on 1 June and continues through 31 May of the following year (365 days), proceeding at a daily time step. Since fires typically occur during July through August in California (Li & Banerjee, 2021), for all simulation trials the fire was set to occur on 1 July, 1 month after the start of the simulation, to ensure that variability in model outputs was due to changes in parameter values rather than the timing of the fire and precipitation.

Model sensitivities were quantified using the Sensitivity Coefficient, ϕ_i, defined by the ratio of the relative change in model output to the relative change in parameter values as follows:

Table 3
Model Parameters Used to Estimate Post-Fire Flood Hazards for Each Study Watershed

Parameter name	Units	Coldwater	McVicker	Leach
Watershed area	km^2	10.9	5.86	3.86
Mean watershed slope (30-m DEM)	%	60.8	46.3	45.4
Length of longest flow path	km	6.44	6.44	3.22
Watershed relief	m	1,342	801	603
Runoff response	1/s			
Frequency		3.36×10^{-5}	2.82×10^{-5}	3.08×10^{-5}
MAE		1.47×10^{-5}	1.47×10^{-5}	1.47×10^{-5}
NSE/RMSE		1.24×10^{-5}	1.24×10^{-5}	1.24×10^{-5}
Mean % Error		2.38×10^{-6}	2.38×10^{-6}	2.38×10^{-6}
Time to peak	hr	5.58	6.02	5.73
Watershed area burned	%	98.0	99.4	99.96
Moderate severity	%	57.6	89.0	89.0
High severity	%	17.9	5.05	2.74
Wildfire hydrologic impact (Livingston et al., 2005)		Moderate	Moderate	Moderate
Post-fire bulking factor	Unitless			
Lower bound		1.001	1.03	1.04
Upper bound		1.01	1.38	1.52
Recovery timescale	days	1,103	726	737
Debris basin design capacity	m^3	–	78,912	20,642
Flood channel design capacity	m^3	–	52.5	10.0
Infrastructure cleaning threshold	% volume filled	–	85	85
Wet season cleaning waiting period	days	–	1	1
Wet season debris basin cleaning rate	m^3/day	–	1,710	2,064

Abbreviation: DEM, Digital Elevation Model.

$$\phi_{i,j} = \frac{\Delta y_i / \overline{y}}{\Delta x_j / \overline{x}} \quad (1)$$

where y_i represents model output variable ($i = 1, 2, \ldots m, m = 6$), x_j represents model parameters ($j = 1, 2, \ldots n, n = 10$) and $\overline{x}$ and $\overline{y}$ represent mean/median values (Hamby, 1994). Model parameters were subsequently ranked by sensitivity coefficient for each output variable, and a count of the number of times a parameter was ranked among the top three most sensitive parameters was taken as an overall measure of parameter importance.

2.2.2. Error Propagation Analysis

Error propagation analysis accounts for the combined influence of all parameters on model outputs when varied across their statistical distributions, unlike sensitivity analysis which isolates the effect of small changes to individual parameters (Gardner et al., 1981). A total of 500 1-year model simulations were completed with all parameters randomly sampled from a uniform distribution between the reported minimum and maximum value, except for the wet season debris basin cleaning rate (r) and the peak flow timescale (t_p). Observations were available for these parameters, so we sampled the data uniformly at random, with replacement ($n = 66$ and $n = 80$, respectively).

Once model output data were generated across the 500 parameter combinations, correlation and partial correlation coefficients were first computed between each model output variable and each parameter. Second, model parameters were ranked by highest squared correlation coefficient. Third, the number of times a parameter ranked in

the top three for each model output variable was counted to obtain an overall ranking of each parameter relative to its error propagation.

2.2.3. Model Parameter Range Estimation

The range of values for each model parameter is presented in Table 2 based on the following considerations: The range of values for the runoff response parameter, c_t, was estimated by first taking the runoff coefficient for undeveloped land between 0.1 and 0.9 Mahmoud and Alazba (2015, Table 2). Next, storm duration was computed for each site using 2018–2021 rain gauge records for Leach (344) and McVicker (343) (Figure 2), and the time to peak parameter, tp, was estimated as storm duration divided by 2 (Snider et al., 2007).

Both c_t and tp were considered in the sensitivity and error analysis. The c_t parameter is used to calculate peak discharge from the watershed outlet, but tp is used separately to calculate the volume of sediment from the watershed outlet ($V_s = Q_w t\mathrm{p}(k-1)$); thus it is important to understand the sensitivity of both parameters. Storm duration (taken as twice the time to peak) was estimated from the rain gauge records by defining storms as continuous rainfall with at least 5.08 mm (0.2 in) of depth and 2.54 mm/hr (0.1 in/hr) of intensity, identifying new storms after 6 hr with <1.27 mm (0.05 in) of rainfall. Storm durations >24 hr were excluded since the model timestep is daily.

The minimum value for the post-fire bulking factor represents the lower bound of hyperconcentrated flow, while the maximum value represents the upper range of debris flow (Gusman et al., 2011).

PF2HazMo accounts for watershed recovery from fire with an exponential decay model for the bulking factor, k, as follows,

$$k(t) = k_0 + (k_1 - k_0)\, e^{-t/t_r} \tag{2}$$

Previous studies have reported that the recovery time is between two and 10 years (Gartner et al., 2009, 2014; Lavé & Burbank, 2004), and thus the range of values for t_r was estimated by assuming $k(t) = 1.01k_0$ when t equals the observed recovery time and rearranging Equation 2 to solve for t_r.

The range of debris basin capacities was calculated according to Equation 21 and channel capacity according to Equation 20 in Jong-Levinger et al. (2022) assuming a bulked design standard. Bound estimates were based on a 50-year precipitation return level estimated from extreme value analysis of a long-term daily precipitation record (1932–2020) from the Big Tujunga Dam station in Los Angeles, California (ID: USC00040798) (NOAA, n.d.), the average area for a watershed in the Transverse Ranges of southern California as estimated by Gartner et al. (2014), and design bulking factors of 1.15, typical of normal streamflow in an alluvial fan, to 2.0, the highest design bulking factor currently in use in southern California (Gusman et al., 2011). For channel capacity estimation, we used the minimum and maximum values of t_p mentioned above. A minimum runoff coefficient of 0.1 was used as above, but the maximum was set to 0.4 since we found that the resulting number of channel exceedances was zero for most of the simulations otherwise.

The range of cleaning threshold values was based on communication with Riverside County, which typically restores infrastructure to full capacity during the dry season when 85% of the capacity is filled with sediment. The minimum and maximum values were obtained by adding or subtracting 10% of the volume filled.

The range of waiting period values is based on the assumption that maintenance crews must wait between one and 14 days after a storm occurs before excavating sediment from infrastructure based on correspondence with Riverside County.

The range of wet season debris basin cleaning rates is based on daily debris basin cleaning records for Leach and McVicker debris basins from December 2018 to April 2019 provided by Riverside County. The median was used as the measure of central tendency since the distribution was right-skewed.

2.3. Parameter Estimation Methods

PF2HazMo is applied by providing estimates for model parameters, some of which are bracketed by previous work or direct measurements, and others which are calibrated. In this section, sources of data and estimation approaches for uncalibrated model parameters are reported. Parameters selected for calibration and the calibration

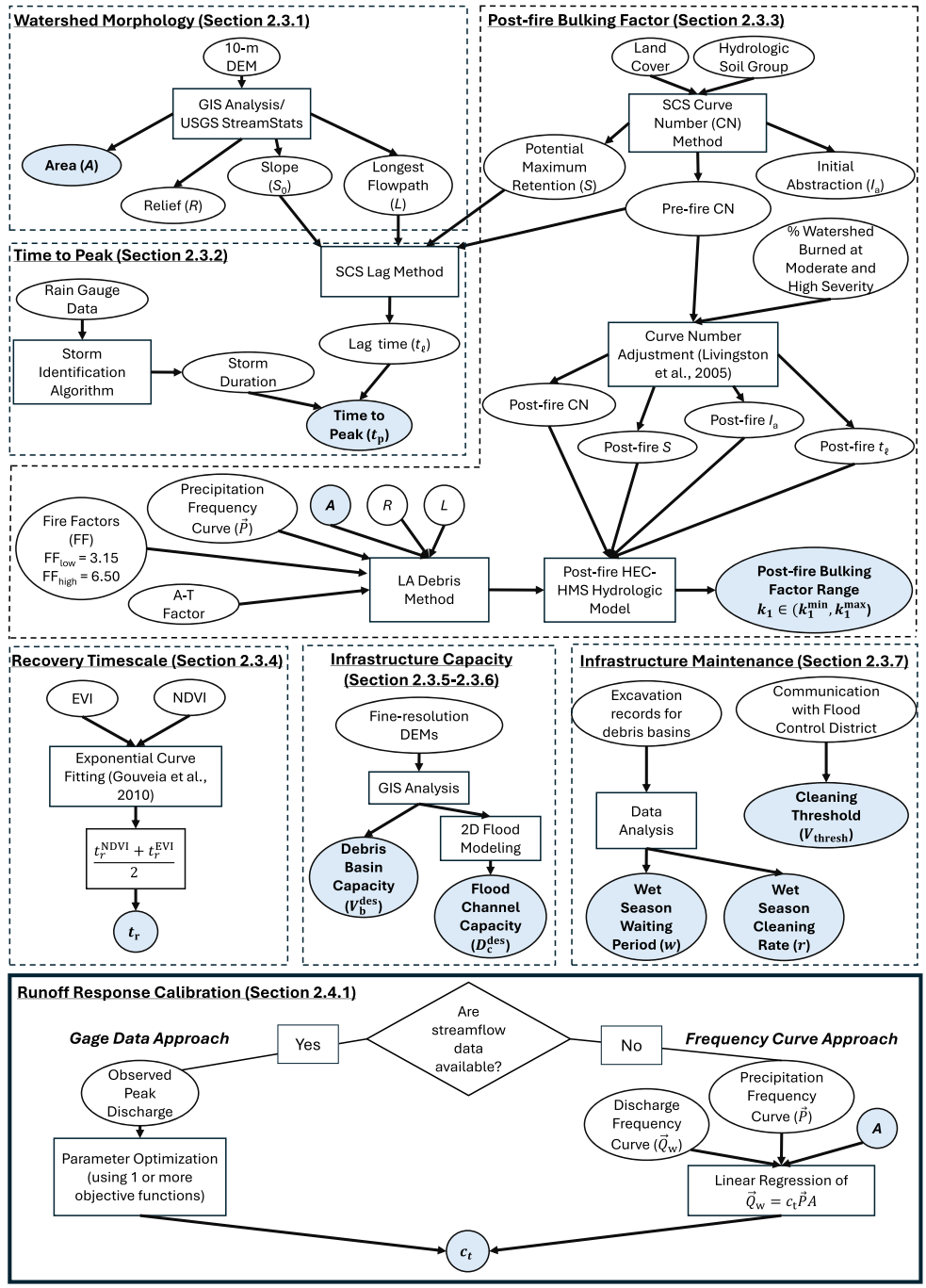

Figure 3. Methodological "roadmap" visually summarizing the inputs, processes, and outputs involved in the parameterization or calibration of each model parameter. Parameterization methods are surrounded by dashed lines while calibration methods are surrounded by solid lines.

approach are described in the following section. A flowchart visualizing the parameterization and calibration methods described in Sections 2.3 and 2.4.1 is shown in Figure 3.

2.3.1. Watershed Morphology

Watershed area was estimated using 10-m Digital Elevation Models (DEMs) from the USGS using ArcGIS (ESRI, 2021; USGS, 2019a). Additional watershed characteristics including the mean slope, length of the longest

flow path, and relief were needed to estimate the post-fire bulking factor and were calculated by the StreamStats online tool developed by the USGS (2019b). StreamStats allows the user to specify the location of a watershed outlet and delineates the watershed using a 30-m DEM; these and other watershed characteristics are provided in the Basin Characteristics report.

2.3.2. Time to Peak

While the runoff response parameter, c_t, could be used to calculate peak flows from the watershed using the Rational Method, the time to peak parameter, t_p, needed to be estimated separately to quantify the fluxes of water and sediment entering the debris basin. The t_p parameter was estimated by modeling each storm with a triangular hydrograph with equal rising and falling limbs. The time to peak parameter represents the amount of time between the start of direct runoff production and the peak of the hydrograph, as observed at the watershed outlet. The time to peak was estimated as $\Delta D/2 + t_\ell$, where ΔD is the storm duration and t_ℓ is the watershed lag (Snider et al., 2007). For each watershed, ΔD was estimated as the median storm duration for the 2018–2019 wet season using rainfall data from the rain gauge nearest each watershed. t_ℓ was estimated using the Soil Conservation Service (SCS) watershed lag method using the watershed characteristics from StreamStats (Kent et al., 2010).

2.3.3. Post-Fire Bulking Factor

The post-fire bulking factor, k_1, was estimated by developing a lumped (i.e., single-basin) model for each watershed using HEC-HMS software (USACE, 2022), calibrating the curve number parameter to pre-fire peak discharge frequency estimates, then increasing the curve number to reflect the expected post-fire increase in runoff following Livingston et al. (2005). The resulting time series of post-fire discharge and sediment load were integrated over each simulated storm to calculate a range of post-fire bulking factors.

The SCS Curve Number method was used as the loss method since there are several studies documenting how the curve number can be adjusted to reflect the post-fire condition (e.g., Carra et al., 2021; Cerrelli, 2005; Goodrich et al., 2005; Soulis, 2018). We calculated an initial estimate for the pre-fire curve number by intersecting National Land Cover Database 2019 land cover data and hydrologic soil group data from the Soil Survey Geographic Database (SSURGO) (Dewitz, 2021; Soil Survey Staff, 2023). Using a land use and soil group lookup table, an area-weighted average curve number for each watershed was estimated (USACE, 2023). This curve number was then used to calculate the maximum potential retention, S, and the initial abstraction, I_a, assuming $I_a = 0.2S$ following Mockus et al. (2004), as well as the watershed lag time following Kent et al. (2010). We used the SCS Unit Hydrograph method as the transform method, using a peak rate factor of 550 based on the steepness of the study watersheds. We forced the model with precipitation frequency estimates for a 2-year to 500-year storm using the Frequency Storm method and the same basin-averaged precipitation frequency estimates described in Section 2.4.1 to develop a precipitation-frequency curve. To distribute the rainfall throughout the storm over time, we used a precipitation intensity pattern for a 6-hr storm from the Riverside County Hydrology Manual representative of a reasonable distribution of rainfall likely to cause critical runoff conditions during a major storm event (RCFCWCD, 1978, Plate E-5.9).

The initial, uncalibrated pre-fire curve number estimate was used to simulate a peak discharge frequency curve, which was then compared to the peak discharge frequency curve retrieved from StreamStats (see Section 2.4.1). The pre-fire curve number was calibrated to the discharge frequency curve by adjusting it by a constant factor between 0 and 1 until the percent error in the area-normalized runoff volume was less than 10% and the Nash-Sutcliffe efficiency (NSE) was greater than 0.90.

Next, to obtain the post-fire curve number, the calibrated pre-fire curve number was adjusted using the following two-step procedure developed by Livingston et al. (2005): (a) the percentage of the watershed area burned at moderate and high burn severity was calculated using the Burn Area Reflectance Classification map for the Holy Fire to determine the Wildfire Hydrologic Impact classification (BAER, 2021), and (b) the relationship between pre-fire curve number and the post-fire to pre-fire curve number ratio was used to calculate the corresponding post-fire curve number. The watershed lag was updated to reflect the post-fire curve number.

The final step of calculating the post-fire bulking factor range was simulating post-fire sediment load. The range of post-fire bulking factors for each watershed was estimated by applying the Los Angeles District Debris Method (LA Debris Method) to each "n-year" return period event in the frequency curves for each watershed and setting

the lower and upper bounds to the minimum and maximum estimated bulking factor, respectively. The LA Debris Method consists of a set of five regression equations with unit sediment yield as the dependent variable (sediment volume normalized by watershed area) and watershed area, relief ratio (relief divided by longest flow path), Fire Factor, and either 1-hr peak precipitation or unit peak discharge (volumetric discharge normalized by watershed area) as the independent variables (Gatwood et al., 2000). Each of the five equations is applicable to a class of watershed areas based on a multiple linear regression analysis conducted on 536 debris yield observations from 87 watersheds in the counties of Los Angeles, Ventura, San Bernardino, and Santa Barbara (Gatwood et al., 2000). We chose the LA Debris Method for estimating post-fire sediment loads because it consists of multiple equations that can be applied to watersheds with a wide range of drainage areas (0.259–518 km^2) and accounts for the effect of time since fire with a parameter termed the "Fire Factor." The Fire Factor is a dimensionless parameter of the LA Debris Method that represents the impact of wildfire on debris yield with a value of 6.5 representing 10 years since 100% of a given watershed burned at a value of 3.15 representing 1 year since 100% of the watershed burned. Since the goal was to obtain a range of post-fire bulking factors from which to sample randomly, we ran one simulation using a Fire Factor of 6.5 (the "high sediment" scenario) and another using a Fire Factor of 3.15 (the "low sediment" scenario). The LA Debris Method requires an Adjustment-Transposition (A-T) Factor to be specified; this parameter is designed to account for differences in geomorphology between the San Gabriel Mountains of Los Angeles County where the regression equations were developed and that of the watershed of interest (Gatwood et al., 2000). For this study, the A-T factor was set to 1.0 due to a lack of sufficient debris yield data for these or surrounding watersheds to estimate it. We note that the actual debris yield of the study watersheds is likely somewhat less than predicted since the Peninsular Ranges in which they are located are known to have a lower debris potential than that of the San Gabriel Mountains (Gatwood et al., 2000; Gusman & Viducich, 2022). The equations of the LA Debris Method used in this study, according to the drainage area constraints of each equation, are as follows,

$$\log Dy = 0.65 \log P + 0.62 \log RR + 0.18 \log A + 0.12 \log FF \quad (3)$$

$$\log Dy = 0.85 \log Q + 0.53 \log RR + 0.04 \log A + 0.22 \log FF \quad (4)$$

where Dy represents unit debris yield, P represents maximum 1-hr precipitation, RR represents relief ratio, A represents watershed area, FF represents Fire Factor, and Q represents unit peak runoff. Equation 3 applies to watershed areas from 0.259 to 7.77 km^2 and thus was used for Leach and McVicker watersheds. Equation 4 applies to watershed areas from 7.77 to 25.9 km^2 and thus was used for Coldwater watershed (see Table 3 for watershed area values).

From the resulting simulated time series of post-fire flow rates and sediment loads, the post-fire bulking factor for each storm in the frequency curve was calculated as the sum of the volumes of water and sediment integrated over the duration of the storm divided by the volume of water. The minimum and maximum bulking factors estimated from both the high and low sediment scenarios were used to define the bounds of the post-fire bulking factor range.

2.3.4. Recovery Timescale

The recovery timescale parameter, t_r, represents the rate of watershed recovery after the watershed is burned by wildfire–that is, the exponential decay rate of the post-fire bulking factor, k_1, back to the pre-fire baseline bulking factor, k_0. The recovery timescale is implemented in the equation for the bulking factor (Equation 2). To estimate the recovery timescale, we assume the watershed vegetation recovers at a rate commensurate to the decay rate of the post-fire bulking factor since vegetation regrowth is the main process by which soil becomes re-stabilized following wildfires (Cerdà & Doerr, 2005; Shakesby, 2011). The recovery rate of watershed vegetation was estimated following Gouveia et al. (2010), who used post-fire Normalized Difference Vegetation Index (NDVI) time series to model vegetation recovery. NDVI is one of the most common satellite-derived vegetation indices used to study post-disturbance vegetation recovery, and several studies have used NDVI time series to evaluate post-fire vegetation recovery in the southwestern U.S. (e.g., Hope et al., 2007; Storey et al., 2021; Van Leeuwen, 2008). Essentially, since NDVI is an indicator of vegetation greenness, Gouveia et al. (2010) define the "lack of greenness," $y(t)$ as:

$$y(t) = \mathrm{NDVI}(t) - \mathrm{NDVI}^*(t) \quad (5)$$

where NDVI*(t) is an annual cycle that represents the "ideally healthy state of vegetation along the phenological year." The lack of greenness variable drops below zero when a fire burns the vegetation, and an exponential curve of the form:

$$y(t) = ae^{-bt} \tag{6}$$

where t is the time since the start of the fire and a = NDVI(t = 0)-NDVI*(t = 0), can be fit to the post-fire time series such that the rate parameter, b, may be found through linear regression. Comparing Equations 2 and 6, it becomes apparent that the recovery timescale may be estimated as $t_r = 1/b$.

This approach was applied to NDVI and Enhanced Vegetation Index (EVI) time series data from the Moderate Imaging Spectroradiometer on the Terra satellite (Didan, 2021). EVI data were included in the analysis to compare the results from the two vegetation indices since EVI is designed to minimize canopy-soil variations and have improved sensitivity where vegetation is more dense, and since it has also been used to study post-fire vegetation recovery (X. Chen et al., 2011; Kinoshita & Hogue, 2011; Wittenberg et al., 2007). NDVI and EVI images that represent 16-day Maximum Value Composites at a resolution of 250 m from 2000 to 2023 were spatially averaged by watershed area in Google Earth Engine. The resulting time series of spatially averaged NDVI and EVI were normalized according to Equation 5 by estimating NDVI*(t) and EVI*(t) as the mean monthly maximum values over the pre-fire period. Eleven missing values in years 2001 and 2002 were replaced with the long-term monthly mean NDVI/EVI values for the pre-fire period. Before curve fitting, we ensured the pre-fire NDVI/EVI time series were stable through time by conducting a Mann-Kendall test on the water-year averaged annual values and found that there was no monotonic trend with time. The fit of the exponential curves to the normalized, spatially averaged vegetation index time series, 95% confidence intervals, R^2, root mean squared error (RMSE), and bulking factor curves resulting from the estimation of t_r are shown in Figure S2 in Supporting Information S1. The recovery timescale parameter for each watershed was calculated as the average of the two recovery timescale estimates produced by curve fitting to the NDVI data and the EVI data.

2.3.5. Debris Basin Capacity

For Leach watershed, the debris basin capacity was estimated through analysis of a DEM produced from an unmanned aerial vehicle (UAV) photogrammetric survey of the debris basin conducted by Riverside County using ArcGIS (ESRI, 2021). The capacity of the debris basin was increased in November 2018 before the first post-fire wet season began in anticipation of increased sediment fluxes, and since Riverside County conducted the photogrammetric survey just after this improvement was completed, it captured the debris basin in its empty, "design capacity" state. The DEM produced from the photogrammetric data had a resolution of 15 cm (0.5 ft) and an absolute vertical accuracy of <5–7 cm (0.16–0.23 ft) at the 95% confidence level, which includes the ground control point error, according to Riverside County. Non-ground features such as vegetation were manually removed from the point cloud to produce a bare earth DEM; then the Surface Volume tool was used in ArcGIS to estimate the volume of the empty debris basin. For McVicker watershed, the debris basin design capacity was estimated by Riverside County from a UAV photogrammetric survey they conducted in September 2018 before the wet season began with similar vertical accuracy.

2.3.6. Flood Channel Capacity

Flood channel capacities for both Leach and McVicker were estimated using PRIMo, a dual-grid flood inundation model, to determine the maximum flow rate the flood channels could convey without producing overbank flows (Kahl et al., 2022; Sanders & Schubert, 2019; Sanders et al., 2023). A DEM produced from an airborne Lidar survey conducted in September 2018 before the start of the wet season with a resolution of 0.46 m (1.5 ft) and vertical accuracy of 3.02 cm (0.099 ft, RMSE) was used to estimate the design capacities of the flood channels. After hydro-conditioning the DEM in ArcGIS to remove bridges and culverts, we forced PRIMo with a constant volumetric flow rate, a no-flow boundary condition at the edge of the domain upstream of the input point source, and a dry boundary condition at the downstream end of the model domain. As shown in Figure 4, the input flow rates were increased or decreased iteratively toward a steady state until the channel overtopped, resulting in overbank flows. We used a threshold flood depth of 5 cm to signify substantial flood channel overtopping. Since the flood channel downstream of the Leach watershed is an unlined channel and that downstream of the McVicker

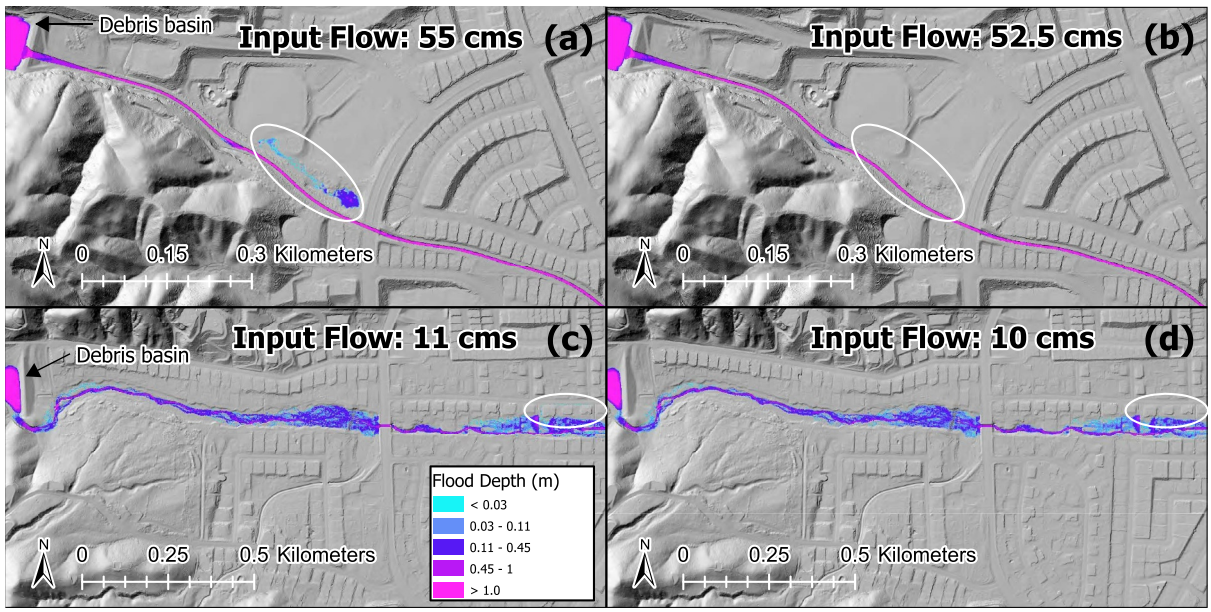

Figure 4. Illustration of channel capacity estimation for McVicker (a, b) and Leach (c, d) watersheds. McVicker channel overtopped with a constant input flow rate of 55 cm (a), but the channel contained a flow rate of 52.5 cm (b). Leach channel overtopped with a constant flow rate of 11 cm (c), resulting in flow onto the street with depths just over the 5 cm threshold (circled in white). When the input flow was decreased to 10 cm (d), no street flooding was simulated.

watershed is a concrete-lined channel, we used Manning n values of 0.030 and 0.013, respectively, based on Chow (1959). Additionally, the Leach flood channel had a brick wall along its southern bank that was not resolved by the DEM and needed to be enforced in the simulation to prevent unrealistic channel overtopping; the grid edge classification method developed by Kahl et al. (2022) was used to implement a 1.5-m tall no-flow boundary to simulate this wall.

2.3.7. Infrastructure Cleaning Standards

The infrastructure maintenance protocol simulated is Summer + After-fire + Wet Season Cleaning, based on the standard maintenance protocol employed by Riverside County. This means that the debris basin and flood channel are restored to full capacity each summer if the infrastructure cleaning threshold has been exceeded or if a wildfire occurred prior to the start of the wet season; additionally, the debris basin will be excavated during the wet season if the cleaning threshold has been exceeded and no rain has occurred for at least the number of days in the "wet season cleaning waiting period." The infrastructure maintenance model parameters include the infrastructure cleaning threshold, the wet season cleaning waiting period, and the wet season debris basin cleaning rate. The infrastructure cleaning threshold was set to when 85% of the debris basin or flood channel is filled, based on discussions with Riverside County; this threshold applies to the cleaning of both the debris basin and the flood channel. The wet season cleaning waiting period and debris basin cleaning rate were estimated based on records of debris basin excavation volumes for Leach and McVicker debris basins from Riverside County for the 2018–2019 wet season at a daily timescale. Daily excavation volumes for each debris basin were averaged over the wet season to obtain the cleaning rate, and waiting periods were roughly based on the number of days following a storm without rain that passed before sediment was excavated. Importantly, only the debris basins are cleaned during the wet season as well as the dry season according to conventional maintenance practices; channels are cleaned during the summer only, based on conversations with Riverside County.

2.4. Calibration Methods

PF2HazMo was calibrated to the first wet season following the Holy Fire, when streamflow and infrastructure overtopping data were available, to assess the performance of the modeling framework with respect to reproducing observed post-fire peak flows and infrastructure exceedances. Streamflow data was used to calibrate the runoff response parameter, c_t, based on its high sensitivity as will be revealed in the results.

For both the calibration and validation, an 8-month simulation period was used spanning a wildfire event and post-fire wet season where rainfall, streamflow, and overtopping were observed. The simulation begins on 1 August 2018 and fire occurrence is simulated on 6 August 2018, the day the Holy Fire ignited. With each simulation, a post-fire bulking factor value is chosen uniformly at random from the estimated post-fire bulking range (Section 2.3.3), and thus the sediment yield properties of each simulation exponentially decay thereafter in accordance with vegetation recovery timescales estimated for each watershed. Observed daily rainfall totals from the rain gauge nearest each watershed are used to force the simulation, with ~41 days of rainfall during the wet season depending on the rain gauge. The simulation ends on 31 March 2019, about 10 days after the last day with observed rainfall. The parameter values used to apply the model to each study watershed are shown in Table 3.

2.4.1. Rainfall-Runoff Model Calibration

PF2HazMo predicts runoff from rainfall using a type of rational method as follows:

$$Q_{\mathrm{w}} = \frac{cPA}{t_{\mathrm{p}}} \tag{7}$$

where c is the runoff coefficient, Q_{w} is the clear-water discharge at the watershed outlet, P is the daily precipitation total, A is the area of the watershed, and t_{p} is the time to peak of a triangular hydrograph with equal rising and falling limbs.

Given the availability of precipitation and streamflow data, calibration targeted the ratio of the runoff coefficient to the time to peak, $c_t = c/t_{\mathrm{p}}$ (Gage Data Approach). Additionally, a second calibration strategy was developed targeting the availability of regional precipitation-frequency and discharge-frequency curves (Frequency Curve Approach). We developed both approaches because, in an ideal scenario, streamflow observations would be available for the watershed of interest to be used to calibrate c_t directly; however, many watersheds are too steep to install gauges, or post-fire bulked flows may be so forceful that they destroy monitoring equipment, according to conversations with Riverside County. Since precipitation and discharge frequency estimates are available for the majority of the continental United States and its territories from the National Oceanic and Atmospheric Administration (NOAA) and the USGS, respectively, PF2HazMo could potentially be applied broadly given the availability of a suitable calibration method for leveraging these data.

Applying the Gage Data Approach, we used observed post-fire streamflow data for the Coldwater watershed to calibrate c_t using multiple error metrics (objective functions) commonly used in hydrologic modeling. Observed peak flows were identified based on a visual analysis of the maximum daily discharge time series and an understanding of when the major storm events of the wet season occurred from discussions with Riverside County. Through manual calibration of the model, we determined the optimal value for c_t was within the range (2×10^{-6}, 4×10^{-5}). We then conducted an automated calibration process in which we ran the model for 100 linearly spaced candidate values within this range and calculated the corresponding NSE index, the RMSE, the mean absolute error (MAE), and the mean percent error (% Error) by comparing the observed peak flows to the median of the simulated peak flow distribution on the same days of the wet season. The equations for the error metrics are as follows:

$$\mathrm{NSE} = 1 - \frac{\Sigma_{i=1}^{n} |y_i - x_i|^2}{\Sigma_{i=1}^{n} |x_i - \overline{x}|^2} \tag{8}$$

$$\mathrm{RMSE} = \left(\frac{1}{n} \Sigma_{i=1}^{n} |y_i - x_i|^2 \right)^{1/2} \tag{9}$$

$$\mathrm{MAE} = \frac{\Sigma_{i=1}^{n} |y_i - x_i|}{n} \tag{10}$$

$$\%\mathrm{Error} = \left| \frac{y_i - x_i}{x_i} \right| \times 100 \tag{11}$$

where x represents observed peak flows, y represents simulated peak flows, n is the number of measured-predicted data pairs, and $\overline{x}$ is the mean of observed peak flows. We chose this suite of error metrics rather than a single objective function because they each have advantages and disadvantages. For example, mean error metrics such as MAE and mean % error are less affected by outliers than are squared error metrics, while relative change metrics (such as % error) highlight errors in low flows more than in high flows (Jackson et al., 2019). For each error metric, we can obtain an optimal parameter estimate by minimizing (in the case of RMSE, MAE, and % error) or maximizing (in the case of NSE) the metric and then compare the outputs from the calibrated models to observations of peak flows and overtopping events. In summary, we can assess how different error metrics (objective functions) bias hazard estimation.

Applying the Frequency Curve Approach, which targets the estimation of c_t in an ungauged watershed, we obtained precipitation frequency estimate grids from NOAA Atlas 14, which represent spatially interpolated point precipitation frequency estimates from a national network of precipitation-reporting stations at a resolution of 30 arc-seconds (Perica et al., 2011). We retrieved annual maximum series for a 6-hr storm for California for return periods of 2, 5, 10, 25, 50, 100, 200, and 500 years. We chose a 6-hr duration because it most closely matches the time to peak of the studied watersheds (Table 3); additionally, a 6-hr storm duration is more typical of the relatively short-duration, high-intensity storms that commonly trigger hyperconcentrated and debris flows in the region than a 24-hr storm, which is quite uncommon in southern California. We then calculated the watershed area-averaged precipitation depth for each return period for the three study watersheds using ArcGIS and used these spatial averages as our precipitation frequency estimates. Discharge frequency estimates for the same return periods were obtained by applying the StreamStats online tool to each watershed individually (USGS, 2019b). These discharge estimates are calculated using regional regression equations developed by fitting a log-Pearson Type III distribution to annual peak flow data through water year 2006 from a network of USGS streamgages in California with 10 or more years of data (Gotvald et al., 2012). The three study watersheds are located in the South Coast Region (Region 5) of the USGS regional regression analysis, which includes non-desert southern California. Once the precipitation and discharge frequency estimates for each return period were obtained, the c_t parameter was estimated through linear regression of $P \times A$ and Q_w, where c_t is the slope of the line with intercept forced to zero. In the following results sections, we refer to this parameter estimation method as the "Frequency" method.

2.5. Validation Methods

PF2HazMo was validated with applications to the Leach and McVicker catchments in Riverside County. Here, detailed monitoring was carried out by Riverside County Flood Control and Water Conservation District from November 2018 to April 2019, a period that immediately followed the Holy Fire. Overtopping events are documented from timelapse videography, and sediment volumes removed from the debris basin and channels are documented by drone surveys and sediment disposal contracts which include volumes of material transported away from the site. Infrastructure overtopping data are described in Section 2.1.3. Finally, as a check to ensure the model conserved sediment mass/volume, we compared the sediment influx, outflux, and storage change of the debris basins and found that the residual was comparable to numerical precision 10^{-11} (Figure S4 in Supporting Information S1).

3. Results

3.1. Model Sensitivity and Error Analyses

3.1.1. Sensitivity Analysis

Figure 5 compares the sensitivity coefficients between model parameters for each model output variable. For some output variables, one model parameter is clearly more sensitive than the others, such as the cleaning waiting period, w, for the number of basin exceedances, or the post-fire bulking factor, k_1, for the number of channel exceedances. However, there are a number of parameters for which the sensitivity coefficient is greater than or equal to one for multiple model outputs, such as the runoff response, c_t.

To better understand which model parameters were consistently sensitive across output parameters, we ranked the parameters by sensitivity coefficient separately for each output variable and then counted the number of times each parameter ranked in the top three most sensitive model parameters. The result of this overall sensitivity

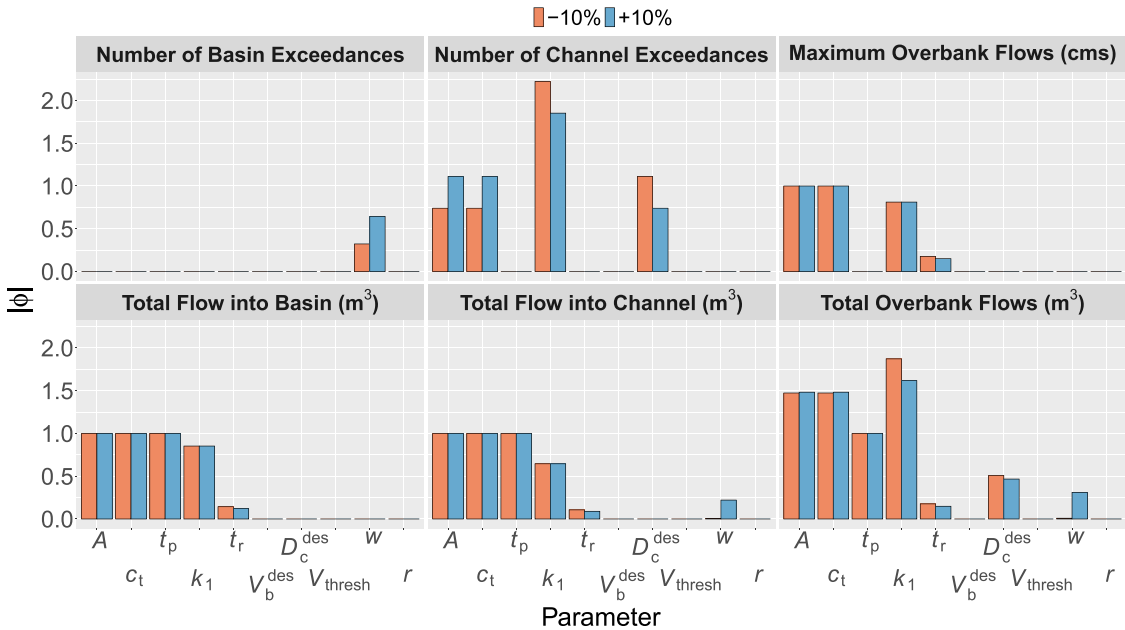

Figure 5. Absolute value of sensitivity coefficients, ϕ, for each model parameter and output variable resulting from a 10% increase or decrease in the parameter value from the mean/median.

ranking is as follows: the number of times a parameter was ranked in the top three was five for watershed area, four for runoff response, three for post-fire bulking factor, and two for time to peak. Flood channel design capacity, debris basin design capacity, and wet season cleaning waiting period each ranked in the top three once. Hence, the most sensitive model parameters are those that contribute to the peak (bulked) flow that drives the hazards.

3.1.2. Error Propagation Analysis

Correlation coefficients linking changes in model outputs to changes in model inputs were also used to rank the importance of input parameters. Ranking by greatest squared correlation coefficient and squared partial correlation coefficient produced the same results: the number of times a parameter was ranked in the top three was six for watershed area, five for runoff response, and four for time to peak. The post-fire bulking factor, flood channel design capacity, and wet season cleaning waiting period were also each ranked in the top three for one model output variable. These results are similar to the sensitivity analysis, in that watershed area, runoff response, and time to peak parameters all ranked highly.

Based on the results from the sensitivity and error analyses, model calibration targeted the runoff response parameter c_t, the post-fire bulking factor was treated as a random variable (bounded by post-fire erosion estimates from HEC-HMS modeling), and all other parameters were estimated using the methods presented in Section 2.3.

3.2. Calibration to Observed Post-Fire Peak Flows

Figure 6 compares observed peak bulked flows to the distribution of simulated peak bulked flow in the Coldwater catchment when c_t is calibrated in five different ways: the first four options rely on in-situ gage data and one of four different objective functions (Gage Data Approach), and the final option designed for ungauged basins relies on regional hazard data (Frequency Curve Approach). First, results show that calibrating PF2HazMo to regional hazard data leads to significantly higher peak bulked flows than suggested by local streamgage measurements, irrespective of the objective function used. The median of the peak flow distribution based on regional hazard data ("Frequency" in Figure 6) is almost four times greater than that of the observed peak flows for the 2018–2019 wet season. The other simulated peak flow distributions were produced by calibrating the c_t parameter to the observed peak flows, each using a different error metric as the objective function for the parameter optimization. Using NSE and RMSE as the objective function resulted in the same peak flow distribution since both are squared error

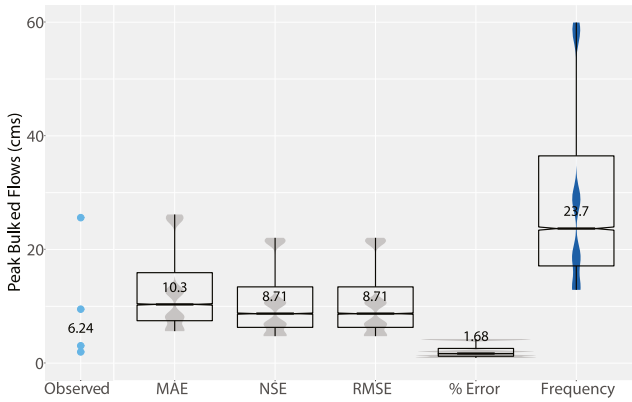

Figure 6. Observed (light blue points), gage-data calibrated (gray), and frequency-curve calibrated (dark blue) peak bulked flow distributions for Coldwater watershed during 2018–2019 wet season. Simulation results for different calibration methods are compared, where mean absolute error (MAE)/Nash-Sutcliffe efficiency (NSE)/root mean squared error (RMSE)/% Error refer to the error metric used to calibrate c_t to observed peak flows and "Frequency" refers to estimating c_t using regional precipitation and discharge frequency estimates. Numbers indicate median values. Sample size for observed peak flows is $n = 4$; sample size for each simulation trial is $n = 12{,}000$.

metrics; hereafter we refer to the simulation results produced by using NSE or RMSE as the objective function as "RMSE" for simplicity. The medians of the peak flow distributions using MAE, NSE, and RMSE (as an objective function) all fall within the notch of the boxplot of the observed peak flows; since the width of the notch roughly represents a 95% confidence interval and the notches of the simulated and observed distributions overlap, this suggests that the medians are not significantly different. However, it is important to interpret the notch of the observed distribution with caution due to the small sample size of observed flows ($n = 4$). The median of the peak flow distribution based on RMSE calibration is closest to that of the observed distribution in terms of percent error relative to the median of the observed distribution. Finally, the peak flow distribution when calibrated by % Error has the lowest median of all the simulated distributions, which is expected since % Error is a relative change metric and this type of error metric tends to emphasize errors in low flows. The MAE, NSE, RMSE, and % Error values corresponding to calibration with each of the five calibration methods are reported in Table S1 in Supporting Information S1.

3.3. Validation With Observed Overtopping Events

Validation of the model is approached with applications to the Leach and McVicker catchments that adopt site-specific parameters for measurable properties (e.g., watershed area, debris basin capacity, flood channel capacity, infrastructure cleaning attributes), the calibrated value of c_t taken from the nearby Coldwater Creek catchment (Gage Data Approach) or calculated using regional hazard data (Frequency Curve Approach), and parameter ranges for other variables as described previously. Validation is pursued with the four different calibrated values of c_t described in the previous section, and a comparison between the simulated distribution of exceedances and the observed number of exceedances is shown in Figure 7 for both channels and debris basins. Figure 7a shows that for Leach watershed, calibration using RMSE as the objective function produces the best fit between the median simulated and observed number of channel exceedances, though the simulated number of basin exceedances was overpredicted (median = 17, observed = 3). For McVicker watershed, all calibration methods correctly and precisely predicted zero channel exceedances, while the number of basin exceedances was overpredicted by roughly an order of magnitude for all but the % Error calibration method, which predicted zero basin exceedances (observed = 1, Figure 7b). Across watersheds and infrastructure types, the Frequency calibration method resulted in a higher number of exceedances than observed (by a factor of ~8–28), while calibrating the model to observed post-fire flows produced similar results for the MAE, NSE, and RMSE objective functions, with the % Error objective function tending to underpredict either the number of channel or basin exceedances.

3.4. Long-Term Compound Hazard Predictions With Calibrated Model

Following calibration and validation, PF2HazMo offers a platform to study future changes in risks from factors such as changes in wildfire frequency and intensity, changes in precipitation, and changes in infrastructure design and maintenance. As an illustration, we used PF2HazMo to simulate the number of channel exceedances over 100 years for Leach and McVicker watersheds under both a "present fire regime" scenario and a "future fire regime" scenario in which the fire frequency is increased. We used the model parameters in Table 3, selecting the NSE/RMSE estimate for the c_t parameter since it resulted in the best overall fit between simulated and observed peak bulked flows, number of infrastructure exceedances, and timing of overtopping events (Sections 3.2 and 3.3). The present fire regime scenario was defined by a mean fire recurrence interval of 40 years, which represents the average period between wildfires under the presumed historical fire regime, spatially averaged across the area of the study watersheds (LANDFIRE, 2014). The future fire regime, characterized by a fire recurrence interval of 20 years, was based on a projection from California's Fourth Climate Change Assessment that under a high emissions scenario (Representative Concentration Pathway 8.5) by 2100, the frequency of large fires would increase by nearly 50% (Bedsworth et al., 2018). The 100-year precipitation time series used to force the model was stochastically generated using the MCMC rainfall simulator described in Jong-Levinger et al. (2022).

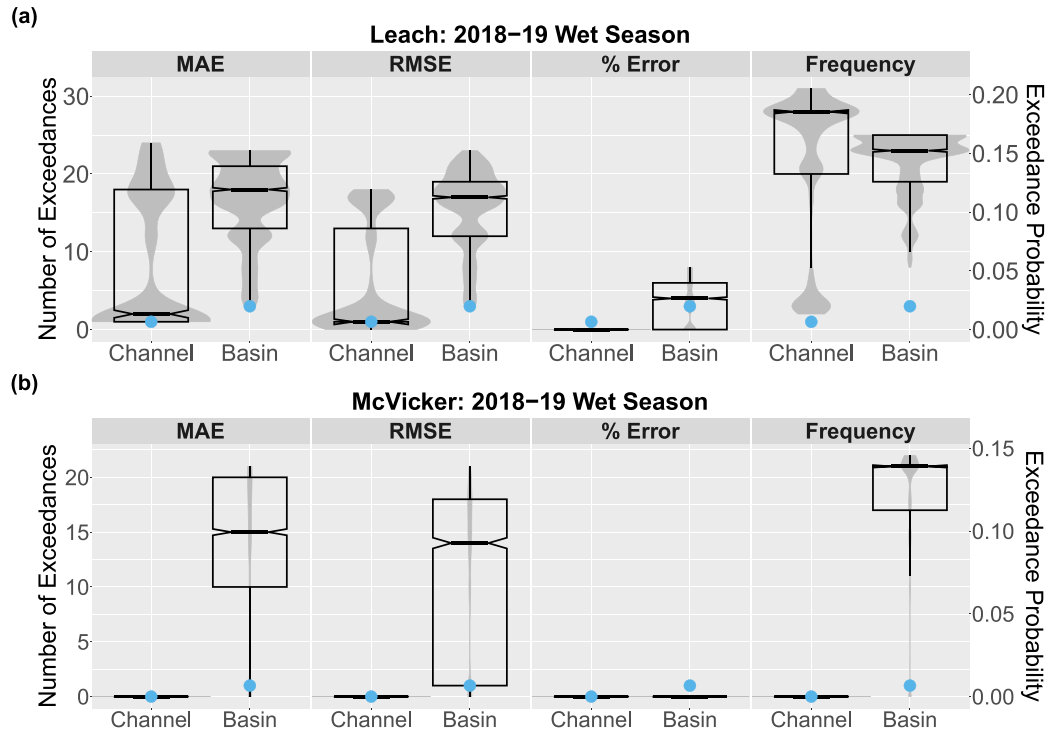

Figure 7. Simulated number of flood channel and debris basin exceedances for each calibration method (gray distributions) with observed number of exceedances shown as blue points for (a) Leach watershed and (b) McVicker watershed. Exceedance probability refers to the daily infrastructure exceedance probability during a given wet season, defined from 1 November to 31 March.

A comparison of the number of channel exceedances per century under present and future fire regimes for Leach watershed is shown in Figure 8. The Leach flood channel has a design capacity capable of conveying a roughly 50-year flood based on the discharge-frequency curve for the watershed; this means that the median number of floods per century under present conditions estimated by PF2HazMo, five, represents three more exceedances than would be expected by ignoring post-fire sediment fluxes and their impact on infrastructure clogging. Under

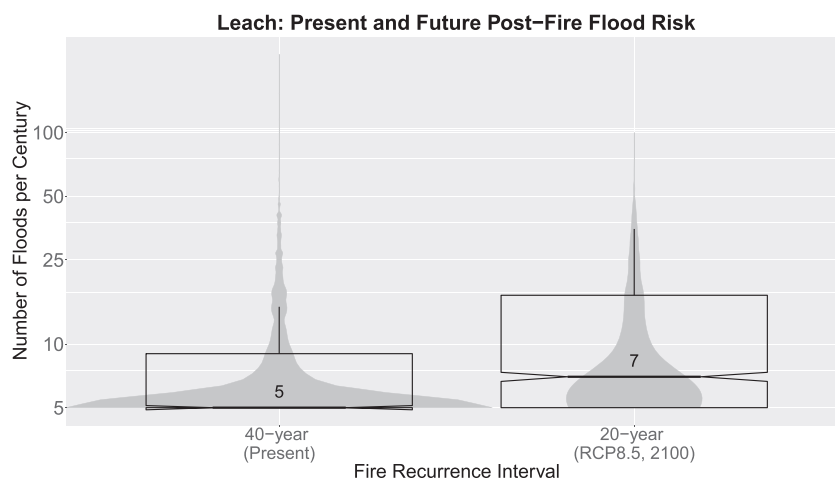

Figure 8. Estimated number of flood channel overtopping events per century under historical (40-year) and projected future (20-year) fire recurrence intervals for Leach watershed. Numbers indicate median values. Simulation results for McVicker watershed estimated zero flood channel overtopping events for both scenarios. Y-axis was log transformed.

future conditions, PFHazMo estimates a median of seven channel exceedances per century, or a roughly 14-year flood return period. To put these flood frequencies into perspective, we define a "hazard amplification factor,"

$$A_F = T_p/T_c = p_c/p_p \tag{12}$$

which represents a ratio of the return period of channel exceedances for the marginal hazard (precipitation in the absence of wildfire) to that of the compound hazard (considering both wildfire and precipitation as hazard drivers). The amplification factor for Leach watershed under the present fire regime is 2.5, while that under the future fire regime is 3.6, highlighting the importance of considering the probability of post-fire bulked flows when designing flood control infrastructure. Lastly, turning attention to the shape of the distributions in Figure 8, the tails are quite long and narrow, especially for the present hazards distribution, indicating it is possible that much greater flood frequencies than the median could be observed, though the probability is small. Notably, the density of the present distribution is more concentrated about its minimum value of five than is the future distribution, indicating a future shift toward a greater likelihood of increased flood frequencies.

In contrast to the hazard estimates for Leach watershed, for McVicker watershed, zero floods per century were predicted for both the present and future fire regime scenarios. The design flood channel capacity for McVicker watershed that we estimated using 2D flood inundation modeling is 52.5 cm (1,854 cfs). Based on the discharge-frequency curve for this watershed (USGS, 2019b), the capacity of the flood channel is greater than even the magnitude of a 500-year flood (30.9 cm or 1,090 cfs). Thus, in this case the concrete-lined flood channel provides a level of protection against floods that can handle the projected compound flood hazards under the future fire regime.

4. Discussion

4.1. Model Applications

The preceding results indicate that different calibration methods are suited to different project goals. For example, the Frequency calibration method appears to work better for simulating higher magnitude peak flow events with longer return periods. To this point, comparing the discharge-frequency curves produced from HEC-HMS simulations to the discharge frequency estimates from StreamStats used to calibrate them reveals a better fit for events with a 25-year return period or longer (Figure S3 in Supporting Information S1). Thus, the Frequency method may be best suited for applications for which more conservative peak flow estimates over longer time horizons are needed, such as the design of debris basins or flood channels downstream of watersheds that burn frequently or are known to produce high sediment yields. This calibration method may also be the best option available if the watershed of interest is ungauged and no nearby streamflow records exist.

On the other hand, if sufficient streamflow records are available, we have demonstrated that PF2HazMo can be calibrated based on these observations and an appropriate range for the post-fire bulking factor to accurately simulate channel exceedances. During the critical period just after a wildfire yet before the start of the wet season, the simulation framework can be used to forecast the number of infrastructure exceedances over the course of the wet season to direct resources to infrastructure maintenance and capacity enhancement efforts or identify infrastructure systems with low levels of protection for emergency planning purposes. By calibrating a HEC-HMS model to pre-fire streamflow using the curve number parameter and subsequently updating the curve number and lag time with satellite-derived burn severity data following Livingston et al. (2005), a range of bulking factors that reflects the burned watershed can be estimated. Since the lumped model and curve number approaches are relatively simple to implement and the required data is publicly available (from land cover, soil property, and topographic databases managed by government agencies), the bulking factor range can be estimated for a given watershed within a matter of days. This post-fire bulking range is then randomly sampled from to quantify the uncertainty regarding post-fire infrastructure exceedance probabilities and provide flood risk managers with a range of potential infrastructure exceedances and bulked flows to plan around. Based on the results presented in Section 3.2, we recommend using NSE/RMSE or MAE as the objective function for calibrating the c_t parameter, as % error produces estimates that underpredict exceedances.

4.2. Comparison to Other Studies of Post-Fire Streamflow and Erosion

By conducting sensitivity and error analyses we identified that the model parameters that have the greatest influence on model response are the runoff response, c_t, the post-fire bulking factor, k_1, the time to peak, t_p, and the flood channel design capacity, D_c^{des}. The runoff coefficient and time to peak parameters represent the influence of land cover and precipitation intensity/storm duration on peak flows, while the post-fire bulking factor and its decay over time implicitly are estimated using data on watershed morphology (e.g., slope, relief, length of longest flow path), the time since the most recent fire, watershed area burned at moderate to high severity, and soil properties–factors that have all been identified as important drivers of post-fire hydrologic response (Gartner et al., 2014; Gatwood et al., 2000; Kampf et al., 2020; Kinoshita et al., 2014; Pelletier & Orem, 2014; Saxe et al., 2018; Wagenbrenner & Robichaud, 2014; Wilder et al., 2021). Our analysis both underscores the importance of these variables and calls attention to the influence of the capacity of the flood channel downstream of the burned watershed on post-fire flood hazards.

Calibrating the model to observed post-fire peak flows produced a range of simulated peak flows that ranged from ∼1–60 cm with median estimates that differed based on the calibration method and objective function used (Figure 6). The observed peak flows (median = 6.24 cm) for Coldwater watershed had a return period of less than 25 years based on its discharge-frequency curve (Figure S3 in Supporting Information S1), but estimates from the California Geological Survey (CGS) for the peak flow during the 6 December 2018 storm were 34 (±17) cm, which is roughly a 50-year event. These estimates represent the clear-water equivalent flows needed to generate peak flow conditions observed in the field and were developed through either direct measurement of flow velocity and cross-sectional area or slope-area back-calculation using the Manning Equation and the broad-crested weir formula. The CGS estimates are much better aligned with the simulated distribution of peak flows produced using the Frequency method than with the streamgage data for Coldwater watershed (Figure 6); they were provided to the authors in a memorandum from CGS that has been included in the data repository that accompanies this study (Swanson & Lindsay, 2024). Additionally, Wilder et al. (2021) simulated post-fire streamflow for watersheds burned by the Holy Fire in the Santa Ana Mountains, including Coldwater, using a five-parameter Random Forest model; their estimates for the 13 watersheds in that region were on the same order of magnitude as our estimates using the Frequency method and those of the CGS. While it is possible that the streamgage used in this study had a negative bias, after calibrating PF2HazMo to the observed streamflow the number of flood channel overtopping events was accurately predicted, with most simulations overpredicting the number of basin overtopping events. This discrepancy between measured and estimated post-fire peak flows highlights the need for more post-fire streamgage data.

4.3. Model Assumptions and Limitations

PF2HazMo is designed to be applied at the watershed scale to capture cycles of wildfire and precipitation over multiple decades to quantify post-fire flood risk due to infrastructure overtopping; thus, certain assumptions were made when formulating the model. First, site-specific variability in sediment yields is captured by randomly sampling post-fire bulking factor values within a range constrained by post-fire erosion simulations. That is, each time a fire is simulated to occur, a post-fire bulking factor is selected randomly from a uniform distribution of values between the minimum and maximum post-fire bulking factor, estimated for a given site using an erosion model. This Monte Carlo sampling method entails that the model does not enforce an absolute limit on sediment supply, though post-fire sediment fluxes from the watershed are constrained by a lower limit determined by factors such as site-specific vegetation cover and watershed slope. Studies have found that sediment supply can limit the occurrence of post-fire debris flows in steep, bedrock-dominated landscapes, especially when multiple runoff events occur during the same wet season (e.g., DiBiase & Lamb, 2020; Guilinger et al., 2020; Palucis et al., 2021); thus, when applying the model to watersheds with these geomorphic characteristics, one can expect that sediment yields and the likelihood of infrastructure overtopping would be somewhat overestimated. Since the model is intended to estimate flood risk over multiple decades rather than over a single wet season, a timescale at which sediment supply changes are more noticeable and relevant, the authors feel that the lack of an absolute limit on sediment supply is an appropriate assumption. If flood risk estimates over shorter simulation durations are desired, future model versions could constrain sediment availability using site-specific data on long-term erosion rates, geology (e.g., bedrock-dominated or soil-mantled landscapes), and dominant erosion types.

Second, while the bulking factor parameter is used to represent different types of sediment-laden flows to the extent that they transport different sediment loads, PF2HazMo does not use process-based models to resolve the physics of sediment-laden flows. Since the goal of the model is to quantify the filling of protective infrastructure with sediment during post-fire storms over time, mass balance equations are used to track the flow of sediment from the watershed outlet to the debris basin to the flood channel. We assume that different types of sediment-laden flow can occur in a given watershed based on precipitation rates and the specific range of bulking factors estimated with erosion modeling for each site. For example, we consider hyperconcentrated flows to have sediment concentrations by volume of 20%–40% and debris flows to have sediment concentrations of 50%–65%, which are represented by bulking factor ranges of 1.25–1.67 and 2.00–2.86, respectively. The approach that PF2HazMo employs to estimate post-fire peak flow is to take the peak clear-water flow and increase it based on the stochastically sampled bulking factor rather than calculating this increase by accounting for the complex, fundamental changes in flow dynamics involved in post-fire debris flows, such as channel erosion. Substantial work has been done to develop physically based, single-phase (i.e., rheologically based) or two-phase models of sediment-laden flows (George & Iverson, 2014; Iverson, 1997; Liu & Huang, 2006; Mikoš & Bezak, 2021; O'Brien et al., 1993), which we consider to be complementary to the perspective of post-fire flood hazards provided by PF2HazMo.

Third, the application of PF2HazMo in this study employs two hydrologic models: the Rational Method, which is used to calculate peak clear-water runoff at the watershed outlet (before considering the proportion of the flow that is sediment), and HEC-HMS, which is used to estimate the range of post-fire bulking factors through a combination of hydrologic and erosion modeling (USACE, 2022). The Rational Method assumes peak flows scale with watershed area and that rainfall intensity is constant during a given storm; thus, simulated sediment yields and peak discharges scale with area, which may lead to inaccuracies when applied to large watersheds, and the effects of time-varying rainfall intensity are not captured. Uncertainties in post-fire bulking factor estimates stem from the assumptions of the hydrologic and erosion models selected for implementation in HEC-HMS, which are discussed in detail in Section 2.3.3. Lastly, it is important to note that even while the watershed is in the "post-fire" condition, that is, while the watershed hydrology is still affected by a wildfire, the model does not increase the peak clear-water discharge relative to the "pre-fire" condition. In this version of PF2HazMo, any increase in peak flow from the pre-fire to post-fire state is quantified by the bulking factor.

The methods for estimating model parameters and the stochastic hazard framework described herein will not perform well in all applications. The post-fire bulking factor range estimated for Coldwater watershed was 1.001–1.01, which is essentially no different than the baseline bulking factor value (Table 3). While the bulking factors estimated for McVicker and Leach included those representative of hyperconcentrated flow, those estimated for Coldwater did not; one potential reason for this is that the drainage area for Coldwater was large enough to fall into a different area class than that of McVicker and Leach, resulting in the use of LA Debris Method Equation 2 within HEC-HMS to estimate sediment loads (Equation 4 in this paper). This empirical model weights drainage area much less than does LA Debris Method Equation 1 (Equation 3 in this paper), used to estimate sediment loads for McVicker and Leach, and since the drainage area for Coldwater was on the lower end of the range of areas for which the regression equation was developed (11 km^2 vs. 7.8–26 km^2), this likely led to an underestimate in the post-fire bulking factor range. Previous studies of post-fire sediment yields find that unit-area sediment yields for watersheds with drainage areas similar to those of the study watersheds decrease as drainage area increases, so this result could be due to the larger watershed area of Coldwater than Leach or McVicker (Gatwood et al., 2000; Pelletier & Orem, 2014; Wagenbrenner & Robichaud, 2014).

Other erosion models implemented in HEC-HMS were tested, including the USGS Emergency Assessment and Long-Term Debris Models, but these also resulted in bulking factors close to 1. Another possibility was that Coldwater watershed simply did not experience much erosion during the first wet season following the Holy Fire: this idea is supported by an estimate from the CGS that sediment concentrations in runoff from Coldwater were roughly 5%–10% based on an analysis of runoff images from a storm event on 29 November 2018 and field observations following the storm (Swanson & Lindsay, 2024). The lower threshold for hyperconcentrated flow is often considered to be a sediment concentration of 20%, so while the CGS estimates are greater than the estimates we produced using the LA Debris Method Equations in HEC-HMS, they are still typical of normal streamflow. Post-fire erosion estimates produced from the differencing of DEMs derived from airborne Lidar show that among 10 other watersheds burned by the Holy Fire, Coldwater produced relatively low estimates; however, these same estimates indicate that McVicker produced similarly low erosion levels and yet we estimated substantially

higher bulking factors for McVicker watershed (Guilinger et al., 2023). Without sediment concentration data it is difficult to know whether the low bulking factor range estimated for Coldwater was realistic, but modelers interested in post-fire erosion modeling with HEC-HMS should note that lower bulking factor estimates for larger watersheds are possible.

In general, it is important to acknowledge the uncertainty associated with the post-fire bulking factor estimates in this study: any inaccuracies inherent in the land cover, soil properties, burn severity, precipitation frequency, or watershed morphology data used to estimate the end-members of the post-fire bulking factor range will carry over into the sediment yield estimates produced by the model. An additional source of uncertainty is the choice to use a uniform distribution when random sampling from the post-fire bulking factor range—a so-called "uniform prior"—based on the lack of evidence for a more accurate parameter distribution. That said, the model is designed to quantify the uncertainty in the post-fire bulking factor parameter by defining a range from which to sample stochastically. This means there is a systematic way to improve model accuracy over time, in that as additional observations and/or better models of post-fire sediment yields become available, uncertainties in the parameter estimates and the shape of the parameter distribution will decrease. Since our sensitivity analysis showed the model is indeed sensitive to the post-fire bulking factor (Figure 5), we note that to obtain the best model results, it is important to constrain uncertainties in the post-fire bulking factor estimates as much as possible. A comparison of box-and-whisker plots of the post-fire bulking factor estimates for each of the study watersheds is shown in Figure S5 in Supporting Information S1.

The present study was limited in scope to the first wet season following fire due to the lack of a longer data set. The infrastructure overtopping observations used to validate the model were only available for two burned watersheds for a single post-fire wet season, which limited both the temporal and spatial scope of the study. While the range of post-fire bulking factors in this study was based on data from the first wet season following the fire, actual sediment concentrations will vary based on the sediment supply, the intensity of subsequent precipitation events, and the recovery of vegetation, which is dependent on climatic factors such as drought (e.g., Hope et al., 2007; Wilder & Kinoshita, 2022). Since the hazard estimates are based on the wet season immediately following the fire and an A-T factor of 1 for the LA Debris Method (representative of the San Gabriel Mountains, which tend to produce higher sediment yields than do the Santa Ana Mountains), they can be considered a "worst case scenario" estimate of post-fire flood hazards. Evaluation of model predictions against longer post-fire streamflow and infrastructure overtopping time series is needed to better assess how the model performs in the years following the first post-fire wet season.

Finally, we note that the impact of short-duration, high-intensity rainfall on peak bulked flows is not resolved well by the model due to the use of a daily timestep. Short-duration, high-intensity rainfall has been shown to be important for generating post-fire debris flows and hyperconcentrated flows as well as generating watershed-scale erosion (e.g., Kampf et al., 2016, 2020; Staley et al., 2013, 2017; Wagenbrenner & Robichaud, 2014). We aimed to partly address this by using a 6-hr storm when selecting the precipitation frequency estimates used to calibrate the c_t parameter (Frequency calibration method), but we acknowledge this is an indirect way of incorporating the impact of high-intensity precipitation. Additionally, we note that "future" projections of post-fire floods were forced with a daily rainfall time series that has the statistical properties of the long-term precipitation gauge record with which it was parameterized and does not reflect non-stationary trends such as an increased frequency of high-intensity events (see Jong-Levinger et al. (2022)). Increased temporal resolution to resolve the effects of high-intensity storm events as well as a modified precipitation probability distribution that reflects the projected effects of climate change on the regional precipitation regime are features of the stochastic modeling framework that are being considered for future work.

5. Conclusion

PF2HazMo is a stochastic hazard estimation framework that can be calibrated and parameterized with publicly available data to estimate post-fire bulked flows and flood control infrastructure overtopping events over a range of timescales, from a single wet season to multiple decades. The present study delineates methods for calibrating the model, whether or not site-specific streamflow data is available, and compares model performance across a range of objective functions used for calibration. An appropriate calibration method should be chosen based on the objective of the modeling effort; then reasonable estimates of bulked flows and the likelihood of infrastructure overtopping can be used to inform infrastructure design and maintenance approaches, emergency planning

strategies, and risk communication efforts. Following model calibration, we show that the expected number of overtopping events per century downstream of the Leach watershed in Riverside County increases by 40% as a result of a 50% increase in the frequency of wildfire events. Using the parameter estimation and calibration methods developed by this work, the model may be applied to multiple watersheds across a region to quickly obtain a first-order estimate of the spatial distribution of post-fire flood risk and inform flood management strategies through stakeholder engagement. The modeling framework can also be used to identify areas that would benefit from the construction of new flood control infrastructure or from the expansion of the capacity of existing infrastructure.

Data Availability Statement

The data and model scripts associated with this manuscript have been made available in the open repository DRYAD (Jong-Levinger, 2024). The topographic data used to estimate watershed morphology are publicly available via the USGS National Map Data Viewer and Downloader (USGS, 2019a). The land cover data used to estimate the curve number are publicly available at the Multi-Resolution Land Characteristics (MRLC) Consortium (Dewitz, 2021). The soil properties data used to estimate the curve number are publicly available at the Web Soil Survey (Soil Survey Staff, 2023). The spatially distributed precipitation frequency estimates used to estimate the runoff response parameter and post-fire bulking factors are publicly available at the NOAA Atlas 14 Precipitation Frequency Data Server (Perica et al., 2011). The watershed-scale discharge frequency estimates used to estimate the runoff response parameter and post-fire bulking factors are publicly available at the USGS StreamStats website (USGS, 2019b). The vegetation indices used to estimate the recovery timescale are publicly available via Google Earth Engine (Didan, 2021). The precipitation record used to develop the stochastic rainfall model is publicly available from the Climate Data Online database (NOAA, n.d.).

Software Availability Statement: MATLAB was used to develop the model and to conduct the majority of the statistical analysis of the observed data and the simulation results (MATLAB, 2022). HEC-HMS version 4.10 was used to estimate post-fire bulking factors (USACE, 2022). R was used to generate boxplot visualizations and for minor data analysis (R Core Team, 2022).

Acknowledgments
This work was supported by the National Science Foundation (NSF) under Grant HDBE-2031535, and AJL acknowledges support from the Chapman University Grand Challenges Initiative. We gratefully acknowledge Riverside County Flood Control and Water Conservation District for providing access to data and knowledge about flood management that greatly contributed to this research. In particular, the authors would like to thank Jason Uhley for providing feedback and context regarding site selection and Joshua Tremba for sharing data and insight that continues to move this research forward. We gratefully acknowledge Donald Lindsay and Brian Swanson with the California Geological Survey for sharing their estimates of post-fire discharge as well as their field and modeling expertise in post-fire hydrology. The authors also thank James J. Guilinger for sharing estimates of post-fire erosion as well as for insightful discussions of post-fire erosion and sediment transport processes. Finally, we would like to thank Jochen Schubert and Daniel T. Kahl for assistance with infrastructure capacity estimation, Samuel J. Weber for assistance with Google Earth Engine, and Ved N. Bhoot for discussions that assisted with recovery timescale estimation.

References

BAER. (2021). Burned area reflectance classification (BARC) products [Dataset]. *Burned Area Emergency Response Teams*. Retrieved from https://burnseverity.cr.usgs.gov/baer/baer-imagery-support-data-download

Bates, P. D., Quinn, N., Sampson, C., Smith, A., Wing, O., Sosa, J., et al. (2021). Combined modeling of us fluvial, pluvial, and coastal flood hazard under current and future climates. *Water Resources Research*, *57*(2), e2020WR028673. https://doi.org/10.1029/2020wr028673

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*(1), 1–12. https://doi.org/10.1038/sdata.2018.214

Bedsworth, L., Cayan, D., Franco, G., Fisher, L., & Ziaja, S. (2018). *California's fourth climate change assessment: Statewide summary report* (Tech. Rep. No. SUM-CCCA4-2018-013). California Energy Commission.

Boughton, W., & Droop, O. (2003). Continuous simulation for design flood estimation—A review. *Environmental Modelling & Software*, *18*(4), 309–318. https://doi.org/10.1016/s1364-8152(03)00004-5

Brenna, A., Surian, N., Ghinassi, M., & Marchi, L. (2020). Sediment–water flows in mountain streams: Recognition and classification based on field evidence. *Geomorphology*, *371*, 107413. https://doi.org/10.1016/j.geomorph.2020.107413

CAL FIRE. (2021). Historic fire perimeter data [Dataset]. Retrieved from https://www.fire.ca.gov/what-we-do/fire-resource-assessment-program/fire-perimeters

California Geological Survey. (2018). Geology of California. Retrieved from https://www.conservation.ca.gov/cgs/Documents/Publications/Map-Sheets/MS_057-2018-Geology-of-California.pdf

Cannon, S. H., & DeGraff, J. (2009). The increasing wildfire and post-fire debris-flow threat in western USA, and implications for consequences of climate change. In *Landslides–Disaster risk reduction* (pp. 177–190). Springer.

Carra, B. G., Bombino, G., Lucas-Borja, M. E., Denisi, P., Plaza-Álvarez, P. A., & Zema, D. A. (2021). Modelling the event-based hydrological response of Mediterranean forests to prescribed fire and soil mulching with fern using the curve number, Horton and USLE-family (universal soil loss equation) models. *Land*, *10*(11), 1166. https://doi.org/10.3390/land10111166

Cerdà, A., & Doerr, S. H. (2005). Influence of vegetation recovery on soil hydrology and erodibility following fire: An 11-year investigation. *International Journal of Wildland Fire*, *14*(4), 423–437. https://doi.org/10.1071/wf05044

Cerrelli, G. (2005). FIRE HYDRO: A simplified method for predicting peak discharges to assist in the design of flood protection measures for western wildfires. In *Managing watersheds for human and natural impacts: Engineering, ecological, and economic challenges* (pp. 1–7).

Chen, L., Berli, M., & Chief, K. (2013). Examining modeling approaches for the rainfall-runoff process in wildfire-affected watersheds: Using San Dimas experimental forest. *JAWRA Journal of the American Water Resources Association*, *49*(4), 851–866. https://doi.org/10.1111/jawr.12043

Chen, X., Vogelmann, J. E., Rollins, M., Ohlen, D., Key, C. H., Yang, L., et al. (2011). Detecting post-fire burn severity and vegetation recovery using multitemporal remote sensing spectral indices and field-collected composite burn index data in a ponderosa pine forest. *International Journal of Remote Sensing*, *32*(23), 7905–7927. https://doi.org/10.1080/01431161.2010.524678

Chow, V. T. (1959). *Open-channel hydraulics*. McGraw-Hill.

Church, M., & Jakob, M. (2020). What is a debris flood? *Water Resources Research*, *56*(8), e2020WR027144. https://doi.org/10.1029/2020wr027144

Cobery, A. (2012). *The Great Crescenta Valley flood: New Year's Day 1934*. Arcadia Publishing.

Dewitz, J. (2021). National Land Cover Database (NLCD) 2019 products [Dataset]. *U.S. Geological Survey*. https://doi.org/10.5066/P9KZCM54

DiBiase, R. A., & Lamb, M. P. (2020). Dry sediment loading of headwater channels fuels post-wildfire debris flows in bedrock landscapes. *Geology*, *48*(2), 189–193. https://doi.org/10.1130/g46847.1

Didan, K. (2021). MODIS/Terra Vegetation Indices 16-day L3 global 250m SIN grid V061 [Dataset]. *NASA EOSDIS Land Processes Distributed Active Archive Center*. Retrieved from https://developers.google.com/earth-engine/datasets/catalog/MODIS_061_MOD13Q1

ESRI. (2021). *ArcMap 10.8.2* [Software]. Environmental Systems Research Institute.

Flanagan, D., & Nearing, M. (1989). USDA-water erosion prediction project: Hillslope profile and watershed model documentation (pp. 1–123). NSERL Report No., 2, 10.

Gardner, R., O'neill, R., Mankin, J., & Carney, J. (1981). A comparison of sensitivity analysis and error analysis based on a stream ecosystem model. *Ecological Modelling*, *12*(3), 173–190. https://doi.org/10.1016/0304-3800(81)90056-9

Gartner, J. E., Cannon, S. H., Helsel, D. R., & Bandurraga, M. (2009). *Multivariate statistical models for predicting sediment yields from Southern California watersheds* (Tech. Rep.). Citeseer.

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

Gatwood, E., Pedersen, J., & Casey, K. (2000). *Los Angeles District method for prediction of debris yield* (Tech. Rep.). USACE. Retrieved from https://www.hec.usace.army.mil/confluence/hmsdocs/hmsguides/files/86908188/89621118/1/1646347766266/Gatwood_2000_Los+Angeles+District+Debris+Method.pdf

George, D. L., & Iverson, R. M. (2014). A depth-averaged debris-flow model that includes the effects of evolving dilatancy. II. Numerical predictions and experimental tests. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *470*(2170), 20130820. https://doi.org/10.1098/rspa.2013.0820

Goodrich, D., Burns, I., Unkrich, C., Semmens, D. J., Guertin, D., Hernandez, M., et al. (2012). KINEROS2/AGWA: Model use, calibration, and validation. *Transactions of the ASABE*, *55*(4), 1561–1574. https://doi.org/10.13031/2013.42264

Goodrich, D., Canfield, H. E., Burns, I. S., Semmens, D., Miller, S., Hernandez, M., et al. (2005). Rapid post-fire hydrologic watershed assessment using the AGWA GIS-based hydrologic modeling tool. In *Managing watersheds for human and natural impacts: Engineering, ecological, and economic challenges* (pp. 1–12).

Gotvald, A. J., Barth, N. A., Veilleux, A. G., & Parrett, C. (2012). Methods for determining magnitude and frequency of floods in California, based on data through water year 2006. US Geological Survey Scientific Investigations Report, 5113 (p. 38).

Gouveia, C., DaCamara, C., & Trigo, R. (2010). Post-fire vegetation recovery in Portugal based on spot/vegetation data. *Natural Hazards and Earth System Sciences*, *10*(4), 673–684. https://doi.org/10.5194/nhess-10-673-2010

Guilinger, J. J., Foufoula-Georgiou, E., Gray, A. B., Randerson, J. T., Smyth, P., Barth, N. C., & Goulden, M. L. (2023). Predicting postfire sediment yields of small steep catchments using airborne lidar differencing. *Geophysical Research Letters*, *50*(16), e2023GL104626. https://doi.org/10.1029/2023gl104626

Guilinger, J. J., Gray, A. B., Barth, N. C., & Fong, B. T. (2020). The evolution of sediment sources over a sequence of postfire sediment-laden flows revealed through repeat high-resolution change detection. *Journal of Geophysical Research: Earth Surface*, *125*(10), e2020JF005527. https://doi.org/10.1029/2020jf005527

Gusman, A. J., Teal, M., Todesco, D., & Bandurraga, M. (2011). *Sediment/debris bulking factors and post-fire hydrology for Ventura County*. Final Report. WEST Consultants Inc.

Gusman, A. J., & Viducich, J. (2022). *San Diego County hydrology manual technical studies: Bulking factor study* (Tech. Rep.). River Focus, Inc.

Hamby, D. M. (1994). A review of techniques for parameter sensitivity analysis of environmental models. *Environmental Monitoring and Assessment*, *32*(2), 135–154. https://doi.org/10.1007/bf00547132

Hamilton, E., Horton, J., Rowe, P., & Reimann, L. (1954). Fire-flood sequences on the San Dimas experimental forest Tech. Paper(6).

Hawkins, R., & Munoz, A. (2011). *Wildcat 5 for Windows (W5W): Documentation and manual*. University of Arizona, School of Natural Resources and Environment.

Hope, A., Tague, C., & Clark, R. (2007). Characterizing post-fire vegetation recovery of California chaparral using TM/ETM+ time-series data. *International Journal of Remote Sensing*, *28*(6), 1339–1354. https://doi.org/10.1080/01431160600908924

Inman, D. L., & Jenkins, S. A. (1999). Climate change and the episodicity of sediment flux of small California Rivers. *Journal of Geology*, *107*(3), 251–270. https://doi.org/10.1086/314346

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

Jackson, E. K., Roberts, W., Nelsen, B., Williams, G. P., Nelson, E. J., & Ames, D. P. (2019). Introductory overview: Error metrics for hydrologic modelling–a review of common practices and an open source library to facilitate use and adoption. *Environmental Modelling & Software*, *119*, 32–48. https://doi.org/10.1016/j.envsoft.2019.05.001

Jakob, M., Hungr, O., & Jakob, D. M. (2005). *Debris-flow hazards and related phenomena* (Vol. 739). Springer.

Jennings, P. C., & Brooks, N. H. (1982). Storms, floods, and debris flows in Southern California and Arizona, 1978 and 1980. In *Proceedings of a symposium, September 17–18, 1980*.

Johnson, P. A., & McCuen, R. H. (1992). Effect of debris flows on debris basin design. *Critical Reviews in Environmental Science and Technology*, *22*(1–2), 137–149. https://doi.org/10.1080/10643389209388431

Johnson, P. A., McCuen, R. H., & Hromadka, T. V. (1991). Debris basin policy and design. *Journal of Hydrology*, *123*(1–2), 83–95. https://doi.org/10.1016/0022-1694(91)90070-x

Jong-Levinger, A. (2024). Post-fire flood hazard model (PF2HazMo) version 1.0.0: Model scripts and parameterization and validation data [Dataset]. *Dryad*. https://doi.org/10.5061/dryad.8w9ghx3t2

Jong-Levinger, A., Banerjee, T., Houston, D., & Sanders, B. F. (2022). Compound post-fire flood hazards considering infrastructure sedimentation. *Earth's Future*, *10*(8), e2022EF002670. https://doi.org/10.1029/2022ef002670

Kahl, D. T., Schubert, J. E., Jong-Levinger, A., & Sanders, B. F. (2022). Grid edge classification method to enhance levee resolution in dual-grid flood inundation models. *Advances in Water Resources*, *168*, 104287. https://doi.org/10.1016/j.advwatres.2022.104287

Kampf, S. K., Brogan, D. J., Schmeer, S., MacDonald, L. H., & Nelson, P. A. (2016). How do geomorphic effects of rainfall vary with storm type and spatial scale in a post-fire landscape? *Geomorphology*, *273*, 39–51. https://doi.org/10.1016/j.geomorph.2016.08.001

Kampf, S. K., Gannon, B. M., Wilson, C., Saavedra, F., Miller, M. E., Heldmyer, A., et al. (2020). PEMIP: Post-fire erosion model intercomparison project. *Journal of Environmental Management*, *268*, 110704. https://doi.org/10.1016/j.jenvman.2020.110704

Kean, J. W., Staley, D. M., Lancaster, J. T., Rengers, F. K., Swanson, B. J., Coe, J. A., et al. (2019). 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

Kent, K. M., Woodward, D. E., Hoeft, C. C., Humpal, A., & Cerrelli, G. (2010). Chapter 15: Time of concentration. In *National engineering handbook* (pp. 15-i–15B-3). U.S. Department of Agriculture.

Kinoshita, A. M., & Hogue, T. S. (2011). Spatial and temporal controls on post-fire hydrologic recovery in Southern California watersheds. *Catena*, *87*(2), 240–252. https://doi.org/10.1016/j.catena.2011.06.005

Kinoshita, A. M., Hogue, T. S., & Napper, C. (2014). Evaluating pre-and post-fire peak discharge predictions across western us watersheds. *JAWRA Journal of the American Water Resources Association*, *50*(6), 1540–1557. https://doi.org/10.1111/jawr.12226

Kirkby, M., Irvine, B., Jones, R. J., Govers, G., & Team, P. (2008). The PESERA coarse scale erosion model for Europe. I.–Model rationale and implementation. *European Journal of Soil Science*, *59*(6), 1293–1306. https://doi.org/10.1111/j.1365-2389.2008.01072.x

LACDPW. (2013). *Sediment management strategic plan: 2012–2032* (Tech. Rep.). Los Angeles County Department of Public Works. Retrieved from https://dpw.lacounty.gov/lacfcd/sediment/stplan.aspx

LANDFIRE. (2014). Mean fire return interval, LANDFIRE 1.4.0 [Dataset]. *Wildland Fire Science, Earth Resources Observation and Science Center, U.S. Geological Survey*. Retrieved from https://landfire.gov/viewer/

Lavé, J., & Burbank, D. (2004). Denudation processes and rates in the transverse ranges, southern California: Erosional response of a transitional landscape to external and anthropogenic forcing. *Journal of Geophysical Research*, *109*(F1), F01006. https://doi.org/10.1029/2003jf000023

Li, S., & Banerjee, T. (2021). Spatial and temporal pattern of wildfires in California from 2000 to 2019. *Scientific Reports*, *11*(1), 1–17. https://doi.org/10.1038/s41598-021-88131-9

Liu, K. F., & Huang, M. C. (2006). Numerical simulation of debris flow with application on hazard area mapping. *Computational Geosciences*, *10*(2), 221–240. https://doi.org/10.1007/s10596-005-9020-4

Livingston, R. K., Earles, T. A., & Wright, K. R. (2005). Los Alamos post-fire watershed recovery: A curve-number-based evaluation. In *Managing watersheds for human and natural impacts: Engineering, ecological, and economic challenges* (pp. 1–11).

Mahmoud, S. H., & Alazba, A. (2015). Hydrological response to land cover changes and human activities in arid regions using a geographic information system and remote sensing. *PLoS One*, *10*(4), e0125805. https://doi.org/10.1371/journal.pone.0125805

MATLAB. (2022). Version 9.13.0.2049777 (r2022b) [Software]. *The MathWorks Inc*. Retrieved from https://www.mathworks.com/products/matlab.html

Mikoš, M., & Bezak, N. (2021). Debris flow modelling using RAMMS model in the alpine environment with focus on the model parameters and main characteristics. *Frontiers in Earth Science*, *8*, 605061. https://doi.org/10.3389/feart.2020.605061

Mockus, V., Hjelmfelt, A. T., & Moody, H. F. (2004). Chapter 10: Estimation of direct runoff from storm rainfall. In *National engineering handbook* (pp. 10-i–10A-51). U.S. Department of Agriculture.

National Research Council. (1996). *Alluvial fan flooding*. National Academies Press.

NOAA. (n.d.). Climate data online search [Dataset]. Retrieved from https://www.ncdc.noaa.gov/cdo-web/datasets/GHCND/stations/GHCND:USC00040798/detail

O'Brien, J. S. (1986). Physical processes, rheology and modeling of mud flows (hyperconcentration, sediment flow) (Vol. *8625109*). ProQuest Dissertations Theses Global. (303469384). Retrieved from https://chapman.idm.oclc.org/login?url=https://www.proquest.com/dissertations-theses/physical-processes-rheology-modeling-mud-flows/docview/303469384/se-2

O'Brien, J. S., Julien, P. Y., & Fullerton, W. (1993). Two-dimensional water flood and mudflow simulation. *Journal of Hydraulic Engineering*, *119*(2), 244–261. https://doi.org/10.1061/(asce)0733-9429(1993)119:2(244)

Orozco, L. (2023). Racing nature: California National Guard clears huge Montecito debris basin filled by massive storm. Retrieved from https://www.kclu.org/local-news/2023-01-24/beating-nature-california-national-guard-clears-huge-montecito-debris-basin-filled-by-massive-storm. (Section: Local News).

Pak, J. H., & Lee, J.-J. (2008). A statistical sediment yield prediction model incorporating the effect of fires and subsequent storm events 1. *JAWRA Journal of the American Water Resources Association*, *44*(3), 689–699. https://doi.org/10.1111/j.1752-1688.2008.00199.x

Palucis, M. C., Ulizio, T. P., & Lamb, M. P. (2021). Debris flow initiation from ravel-filled channel bed failure following wildfire in a bedrock landscape with limited sediment supply. *Bulletin*, *133*(9–10), 2079–2096. https://doi.org/10.1130/b35822.1

Pathiraja, S., Westra, S., & Sharma, A. (2012). Why continuous simulation? The role of antecedent moisture in design flood estimation. *Water Resources Research*, *48*(6), W06534. https://doi.org/10.1029/2011wr010997

Pelletier, J. D., & Orem, C. A. (2014). How do sediment yields from post-wildfire debris-laden flows depend on terrain slope, soil burn severity class, and drainage basin area? Insights from airborne-lidar change detection. *Earth Surface Processes and Landforms*, *39*(13), 1822–1832. https://doi.org/10.1002/esp.3570

Perica, S., Dietz, S., Heim, S., Hiner, L., Maitaria, K., Martin, D., et al. (2011). Precipitation-frequency atlas of the United States, California: NOAA atlas 14 volume 6 version 2.3 [Dataset]. *NOAA, National Weather Service*. Retrieved from https://hdsc.nws.noaa.gov/pfds/pfds_gis.html

PRISM Climate Group. (2023). Oregon State University. Retrieved from https://prism.oregonstate.edu

Prochaska, A. B., Santi, P. M., & Higgins, J. D. (2008). Debris basin and deflection berm design for fire-related debris-flow mitigation. *Environmental and Engineering Geoscience*, *14*(4), 297–313. https://doi.org/10.2113/gseegeosci.14.4.297

RCEPD. (2003). Santa Ana Mountains bioregion: Volume 2 - Section 2.4. In *Western Riverside County multiple species habitat conservation plan (MSHCP)*. Riverside County Environmental Programs Division.

RCFCWCD. (1978). *Hydrology manual* (Tech. Rep.). Riverside County Flood Control and Water Conservation District. Retrieved from https://www.cityofdhs.org/documents/riverside-county-hydrology-manual/?filter_categories[]=375

R Core Team. (2022). R: A language and environment for statistical computing [Software]. Retrieved from https://www.R-project.org/

Renard, K. G. (1997). *Predicting soil erosion by water: A guide to conservation planning with the revised universal soil loss equation (RUSLE)*. United States Government Printing.

Rengers, F. K., McGuire, L. A., Kean, J. W., Staley, D. M., Dobre, M., Robichaud, P. R., & Swetnam, T. (2021). Movement of sediment through a burned landscape: Sediment volume observations and model comparisons in the San Gabriel Mountains, California, USA. *Journal of Geophysical Research: Earth Surface*, *126*(7), e2020JF006053. https://doi.org/10.1029/2020jf006053

Rowe, P. B., Countryman, O., & Storey, H. (1949). *Probable peak discharges and erosion rates from southern California watersheds as influenced by fire*. US Department of Agriculture, Forest Service.

Sadegh, M., Moftakhari, H., Gupta, H. V., Ragno, E., Mazdiyasni, O., Sanders, B., et al. (2018). Multihazard scenarios for analysis of compound extreme events. *Geophysical Research Letters*, *45*(11), 5470–5480. https://doi.org/10.1029/2018gl077317

Sanders, B. F., & Grant, S. B. (2020). Re-envisioning stormwater infrastructure for ultrahazardous flooding. *Wiley Interdisciplinary Reviews: Water*, *7*(2), e1414. https://doi.org/10.1002/wat2.1414

Sanders, B. F., & Schubert, J. E. (2019). PRIMo: Parallel raster inundation model. *Advances in Water Resources*, *126*, 79–95. https://doi.org/10.1016/j.advwatres.2019.02.007

Sanders, B. F., Schubert, J. E., Kahl, D. T., Mach, K. J., Brady, D., AghaKouchak, A., et al. (2023). Large and inequitable flood risks in Los Angeles, California. *Nature Sustainability*, *6*(1), 47–57. https://doi.org/10.1038/s41893-022-00977-7

Saxe, S., Hogue, T. S., & Hay, L. (2018). Characterization and evaluation of controls on post-fire streamflow response across western us watersheds. *Hydrology and Earth System Sciences*, *22*(2), 1221–1237. https://doi.org/10.5194/hess-22-1221-2018

Scott, K. M., & Williams, R. P. (1978). *Erosion and sediment yields in the transverse ranges, Southern California* (Vol. 1030). US Government Printing Office.

Shakesby, R. (2011). Post-wildfire soil erosion in the Mediterranean: Review and future research directions. *Earth-Science Reviews*, *105*(3–4), 71–100. https://doi.org/10.1016/j.earscirev.2011.01.001

Snider, D., Woodward, D. E., Hoeft, C. C., Merkel, W. H., Chaison, K. E., & Moody, H. F. (2007). Chapter 16: Hydrographs. In *National engineering handbook* (pp. 116-i–16B-13). U.S. Department of Agriculture.

Soil Survey Staff. (2023). Web soil survey [Dataset]. *Natural Resources Conservation Service, United States Department of Agriculture*. Retrieved from https://websoilsurvey.nrcs.usda.gov/

Soulis, K. X. (2018). Estimation of SCS curve number variation following forest fires. *Hydrological Sciences Journal*, *63*(9), 1332–1346. https://doi.org/10.1080/02626667.2018.1501482

Staley, D. M., Kean, J. W., Cannon, S. H., Schmidt, K. M., & Laber, J. L. (2013). Objective definition of rainfall intensity–duration thresholds for the initiation of post-fire debris flows in southern California. *Landslides*, *10*(5), 547–562. https://doi.org/10.1007/s10346-012-0341-9

Staley, D. M., Negri, J. A., Kean, J. W., Laber, J. L., Tillery, A. C., & Youberg, A. M. (2017). 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

Storey, E. A., Stow, D. A., O'Leary, J. F., Davis, F. W., & Roberts, D. A. (2021). Does short-interval fire inhibit postfire recovery of chaparral across southern California? *Science of the Total Environment*, *751*, 142271. https://doi.org/10.1016/j.scitotenv.2020.142271

Swanson, B., & Lindsay, D. (2024). Discharge data estimated for flows generated from the 2018 Holy Fire by storms of November 29, 2018, and December 6, 2018 [Memorandum]. https://doi.org/10.5061/dryad.8w9ghx3t2

Tate, E., Rahman, M. A., Emrich, C. T., & Sampson, C. C. (2021). Flood exposure and social vulnerability in the United States. *Natural Hazards*, *106*(1), 435–457. https://doi.org/10.1007/s11069-020-04470-2

USACE. (2022). Hydrologic modeling system HEC-HMS user's manual version 4.10 [Software]. *Institute for Water Resources Hydrologic Engineering Center*. Retrieved from https://www.hec.usace.army.mil/software/hec-hms/downloads.aspx

USACE. (2023). *Creating a curve number grid and computing subbasin average curve number values* [Tutorial]. U.S. Army Corps of Engineers Hydrologic Engineering Center. Retrieved from https://www.hec.usace.army.mil/confluence/hmsdocs/hmsguides/gis-tutorials-and-guides/creating-a-curve-number-grid-and-computing-subbasin-average-curve-number-values

USDA. (2009). *Small watershed hydrology: WinTR-55 user guide*. Conservation Engineering Division.

USGS. (2019a). 3D elevation program 10-meter resolution digital elevation model [Dataset]. Retrieved from https://apps.nationalmap.gov/downloader/

USGS. (2019b). The StreamStats program [WebApplication]. Retrieved from https://streamstats.usgs.gov/ss/

Van Eck, C. M., Nunes, J. P., Vieira, D. C., Keesstra, S., & Keizer, J. J. (2016). Physically-based modelling of the post-fire runoff response of a forest catchment in central Portugal: Using field versus remote sensing based estimates of vegetation recovery. *Land Degradation & Development*, *27*(5), 1535–1544. https://doi.org/10.1002/ldr.2507

Van Leeuwen, W. J. (2008). Monitoring the effects of forest restoration treatments on post-fire vegetation recovery with MODIS multitemporal data. *Sensors*, *8*(3), 2017–2042. https://doi.org/10.3390/s8032017

Wagenbrenner, J. W., & Robichaud, P. R. (2014). Post-fire bedload sediment delivery across spatial scales in the interior western United States. *Earth Surface Processes and Landforms*, *39*(7), 865–876. https://doi.org/10.1002/esp.3488

Wallerstein, N., Thorne, C. R., & Abt, S. R. (1997). Debris control at hydraulic structures in selected areas of the United States and Europe.

Wang, J., Stern, M. A., King, V. M., Alpers, C. N., Quinn, N. W., Flint, A. L., & Flint, L. E. (2020). PFHydro: A new watershed-scale model for post-fire runoff simulation. *Environmental Modelling & Software*, *123*, 104555. https://doi.org/10.1016/j.envsoft.2019.104555

Wilder, B. A., & Kinoshita, A. M. (2022). Incorporating ECOSTRESS evapotranspiration in a paired catchment water balance analysis after the 2018 Holy Fire in California. *Catena*, *215*, 106300. https://doi.org/10.1016/j.catena.2022.106300

Wilder, B. A., Lancaster, J. T., Cafferata, P. H., Coe, D. B., Swanson, B. J., Lindsay, D. N., et al. (2021). An analytical solution for rapidly predicting post-fire peak streamflow for small watersheds in southern California. *Hydrological Processes*, *35*(1), e13976. https://doi.org/10.1002/hyp.13976

Willardson, B. (2020). Sustainable debris basins for post-fire protection. In *2020 intermountain engineering, technology and computing (IETC)* (pp. 1–6).

Winter, B., Schneeberger, K., Dung, N., Huttenlau, M., Achleitner, S., Stötter, J., et al. (2019). A continuous modelling approach for design flood estimation on sub-daily time scale. *Hydrological Sciences Journal*, *64*(5), 539–554. https://doi.org/10.1080/02626667.2019.1593419

Wischmeier, W. H., & Smith, D. D. (1965). Predicting rainfall-erosion losses from cropland east of the Rocky Mountains.

Wittenberg, L., Malkinson, D., Beeri, O., Halutzy, A., & Tesler, N. (2007). Spatial and temporal patterns of vegetation recovery following sequences of forest fires in a Mediterranean landscape, Mt. Carmel Israel. *Catena*, *71*(1), 76–83. https://doi.org/10.1016/j.catena.2006.10.007