Received: 6 January 2022 | Revised: 24 August 2022 | Accepted: 28 August 2022

DOI: 10.1002/esp.5480

RESEARCH ARTICLE

ESPL WILEY

Predicting post-fire debris flow grain sizes and depositional volumes in the Intermountain West, United States

Sara Wall[1] | Brendan P. Murphy[2] | Patrick Belmont[1] | Larissa Yocom[3]

[1]Department of Watershed Sciences, Utah State University, Logan, UT, USA

[2]School of Environmental Science, Simon Fraser University, Burnaby, B.C., Canada

[3]Department of Wildland Resources, Utah State University, Logan, UT, USA

Correspondence
Sara Wall, Department of Watershed Sciences, Utah State University, Logan, UT 84322, USA.
Email: sara.wall019@gmail.com

Funding information
Utah State University Public Lands Initiative; Joint Fire Science Program, Grant/Award Number: 19-2-02-6; National Science Foundation, Grant/Award Number: NSF-EAR 1848667; Utah Agricultural Experiment Station

Abstract

Post-fire debris flows represent one of the most erosive consequences associated with increasing wildfire severity and investigations into their downstream impacts have been limited. Recent advances have linked existing hydrogeomorphic models to predict potential impacts of post-fire erosion at watershed scales on downstream water resources. Here we address two key limitations in current models: (1) accurate predictions of post-fire debris flow volumes in the absence of triggering storm rainfall intensities and (2) understanding controls on grain sizes produced by post-fire debris flows. We compiled and analysed a novel dataset of depositional volumes and grain size distributions (GSDs) for 59 post-fire debris flows across the Intermountain West (IMW) collected via fieldwork and from the literature. We first evaluated the utility of existing models for post-fire debris flow volume prediction, which were largely developed for Southern California. We then constructed a new post-fire debris flow volume prediction model for the IMW using a combination of Random Forest modelling and regression analysis. We found topography and burn severity to be important variables, and that the percentage of pre-fire soil organic matter was an essential predictor variable. Our model was also capable of predicting debris flow volumes without data for the triggering storm, suggesting that rainfall may be more important as a presence/absence predictor, rather than a scaling variable. We also constructed the first models that predict the median, 16th percentile, and 84th percentile grain sizes, as well as boulder size, produced by post-fire debris flows. These models demonstrate consistent landscape controls on debris flow GSDs that are related to land cover, physical and chemical weathering, and hillslope sediment transport processes. This work advances our ability to predict how post-fire sediment pulses are transported through watersheds. Our models allow for improved pre- and post-fire risk assessments across diverse ranges of watersheds in the IMW.

KEYWORDS
debris flows, grain size, post-fire erosion, sediment yields, wildfire

1 | INTRODUCTION

Wildfire activity has increased considerably in western North America over the past three decades and is expected to continue to increase in frequency, severity, and size due to increasing drought and high fuel loads (Hawbaker & Zhu, 2012; Jager et al., 2021; Murphy et al., 2018; Westerling et al., 2011). Climate change-induced drought is resulting in smaller snowpack and earlier snowmelt (Mote et al., 2005; Saley et al., 2022), which lead to longer fire seasons and drier fuels, particularly in mid-elevation forests (1680–2590 m; Abatzoglou & Williams, 2016; Westerling et al., 2006; Wilkins et al., 2021). This is concerning because mid-elevation forests host the water supply for two-thirds of the population of the western United States (Brown et al., 2008; Murphy et al., 2018).

High-severity fire can substantially reduce the infiltration capacity of soil, which causes increased runoff, surface erosion, and can generate debris flows (Cannon & Gartner, 2005; Cannon et al., 2001; Doerr et al., 2006; Mataix-Solera et al., 2011; Ren & Leslie, 2020;

Wondzell & King, 2003). However, there is a shortage of quantitative information regarding the grain size distributions of sediment eroded from post-fire landscapes, which hinders our ability to predict downstream transport and delivery of sediment and associated risk to downstream natural resources and infrastructure, such as water supply reservoirs. As fire activity increases across western North America, it is essential to understand the fundamental impacts post-fire erosion may have on downstream resources, specifically reservoirs and aquatic habitat (Jager et al., 2021).

Debris flows are one of the most erosive and potentially hazardous risks following a wildfire and have the potential to degrade water quality, water supply, and aquatic habitat (Martin, 2016; Moody & Martin, 2004; Robinne et al., 2016; Sedell et al., 2015; Smith et al., 2011). Depending on sediment composition and volume, post-fire debris flows that enter waterways may either degrade or enhance aquatic habitat (Brown et al., 2001; Burton, 2005; Gresswell, 1999; Roghair et al., 2002). Aquatic organisms, such as fish, require particular riverbed grain sizes for their survival (Kondolf, 2000), and those requirements may differ with species, life stage, and habitat purpose (e.g. Murphy et al., 2020). Sediment inputs from debris flows may significantly alter bed grain size distributions (GSDs), shifting the system closer to, or further from, optimal ecological conditions. For example, increased inputs of fine sediments into river systems as a result of post-fire erosion can result in the burying of gravels and pore spaces necessary for spawning fish (Brown et al., 2001; Gresswell, 1999; Propst & Stefferud, 1997), while deposition of coarse material can increase channel complexity and improve aquatic habitat (Bisson et al., 2003; Reeves et al., 1995; Sedell et al., 2015). Recent work has also emphasized the detrimental and costly impacts of increasing post-fire sedimentation on reservoir storage capacity (Gannon et al., 2019; Martin, 2016; McCoy et al., 2016; Moody & Martin, 2004; Murphy et al., 2018; Sankey et al., 2017). In order to accurately assess reservoir vulnerability to post-fire erosion, we must be able to predict the GSDs of that erosion as grain size exerts a first-order control on how sediment is transported through a river network (Ahammad et al., 2021; Czuba & Foufoula-Georgiou, 2014; Wilcock & Crowe, 2003).

Both post-fire hillslope and debris flow erosion pose threats to downstream water resources (e.g. Martin, 2016; Murphy et al., 2018), yet to date only the downstream impacts of hillslope erosion have typically been investigated (e.g. Gannon et al., 2019; Kampf et al., 2020; Sankey et al., 2017). One key reason for this disparity is a lack of data and knowledge regarding the size and transport dynamics of debris flow sediment. The grain sizes from hillslope erosion are controlled predominantly by soil characteristics (Pietraszek, 2006; Robichaud, 2005; Robichaud et al., 2016; Shakesby et al., 2016), whereas the controls on the grain size distribution of debris flows are not well understood (Nyman et al., 2020). Because debris flows are recognized as one of the largest contributors of erosion post-fire (Ellett et al., 2019; Moody & Martin, 2009), it is essential that debris flow characteristics such as occurrence, behaviour, and composition are better understood to predict the impacts of post-fire erosion on downstream water resources.

Several predictive models have advanced our ability to predict the occurrence and magnitude of post-fire debris flows (Cannon et al., 2010; Gartner et al., 2008, 2014; Liu & He, 2020; Sankey et al., 2017; Staley et al., 2017). In particular, Staley et al. (2017) developed an empirical model for predicting post-fire debris flow generation for the western United States. This model has been adopted by the United States Geological Survey (USGS) and the United States Forest Service (USFS) for post-fire hazards assessments and is widely used across the western United States. Gartner et al. (2008) developed a set of empirical models to predict the volumes of debris flows in burned basins in the western United States using data from Southern California, Utah, and Colorado. Gartner et al. (2014) developed another empirical model used to predict debris flow volumes for burned basins using data solely from Southern California, which has been adopted as the primary model used by the USGS and USFS for predicting debris flow volumes across the western United States. We note that the models from all three of these studies require detailed rainfall data on the triggering event storm for use in either validation or prediction. Additionally, with respect to modelling the fluvial transport of coarse sediment inputs, recent improvements to sediment routing models have allowed for improved predictions of mixed-size sediment transport through large river networks (Ahammad et al., 2021; Czuba, 2018; Czuba et al., 2016; Gilbert & Wilcox, 2020; Pfeiffer et al., 2020).

Due to the hazardous nature of debris flows, the majority of studies and models developed over the past two decades have focused on their initiation mechanisms, probability of occurrence, and potential magnitudes (e.g. Cannon & Gartner, 2005, 2010; Gartner et al., 2008, 2014; McGuire et al., 2021; Staley et al., 2017; Tang et al., 2019). However, recent studies of post-fire response have sought to link multiple predictive models to evaluate the impacts of post-fire erosion at the scale of large watersheds (>10 km^2) (e.g. Gannon et al., 2019; Langhans et al., 2016; Murphy et al., 2019; Nyman et al., 2020). Accurately modelling post-fire sediment cascades, particularly when including coarse inputs from debris flows, ultimately requires detailed information about the location, timing, volumes, and grain sizes of debris flow inputs to river networks (Murphy et al., 2019).

We identify and explore two key knowledge gaps that currently limit the development of reliable post-fire, watershed-scale models. (1) In the absence of data on triggering storm rainfall intensities, can we reasonably predict observed volumes for post-fire debris flows? (2) What controls the grain size distribution (GSD) of post-fire debris flows? Specifically, while empirical models exist to predict post-fire debris flow volumes (Gartner et al., 2008, 2014), rigorous validation of these models outside of Southern California is limited. Further, these models require high-resolution precipitation data about the triggering storm event that are often difficult to predict or constrain (Murphy et al., 2019; Nyman et al., 2015). Additionally, no models exist to predict the GSDs of post-fire debris flows. This presents a major obstacle in the development of reliable watershed-scale wildfire risk assessment models, as grain size controls the rates and modes of sediment transport through a river network (Ahammad et al., 2021; Czuba & Foufoula-Georgiou, 2014; Wilcock & Crowe, 2003).

The goal of this study is to improve predictions of post-fire debris flow characteristics that currently limit our ability to predict large watershed-scale sediment delivery, transport, and downstream impacts after wildfire. This research focuses on quantifying GSDs of post-fire debris flows, identifying upstream, landscape controls on these GSDs, evaluating existing post-fire debris flow volume models,

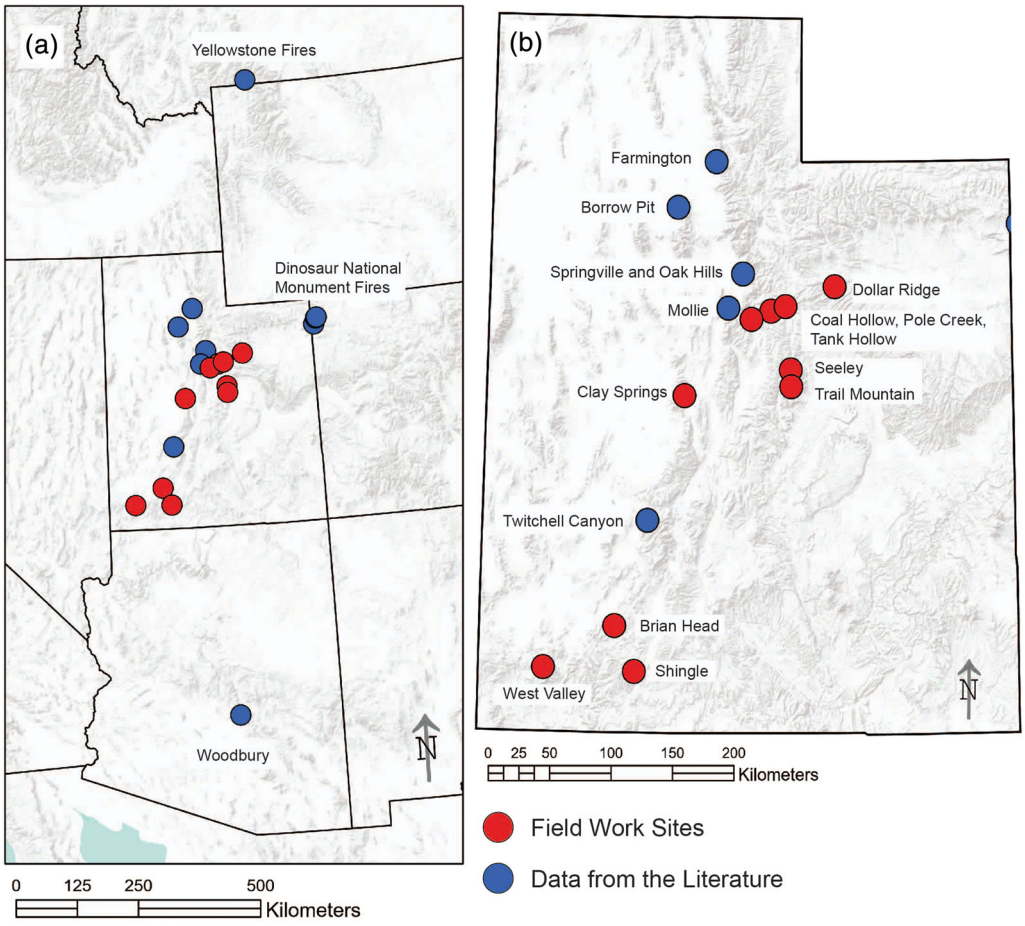

FIGURE 1 Site locations examined as part of this study. (a) Map showing the sites of wildfires across the IMW used to evaluate both post-fire debris flow grain sizes and depositional volumes. (b) Detailed map showing the sites and names of Utah wildfires examined in this study. In both panels, red circles represent sites where fieldwork was conducted for this study, and blue circles represent sites where data was previously collected and reported in published literature. [Color figure can be viewed at wileyonlinelibrary.com]

and constructing new predictive models for key GSD metrics and deposit volumes of post-fire debris flows in the IMW. This work offers the first in-depth study, of which the authors are aware, that investigates landscape controls on the GSDs of post-fire debris flows.

2 | STUDY AREA

2.1 | Intermountain West

The Intermountain West (IMW) is defined as the region of the United States between the Cascade Range and Sierra Nevada to the west and the front range of the Rocky Mountains to the east. The climate of this region is influenced by its mid-continental location, high average elevations, and complex mountain topography (WWA, 2021), which yields low humidity, large seasonal temperature changes, and steep gradients of temperature and precipitation with elevation. The region also has a strong seasonal distribution of precipitation, with the majority delivered as snow during winter months (WWA, 2021).

Historical trends in fire activity across the western United States reveal that the area burned each year prior to Euro-American settlement was much higher than the annual burn area today (Murphy et al., 2018), and overall, the west has been in a 'fire deficit' for at least six to seven decades compared to historical fire activity (Marlon et al., 2012). While the area burning in the western United States is not exceptionally high from a long-term perspective, recent trends in wildfire behaviour and the amount of infrastructure damage caused by high-severity fire across the west are unprecedented (Abatzoglou & Williams, 2016; Duane et al., 2021). As the area burned at high severity in the western United States increases, sediment yields from hillslope erosion are projected to at least double in 35% of western watersheds by 2050 (Sankey et al., 2017). Higher sediment yields as a result of increasing fire activity in the western United States pose a great threat to water storage capacity in western US reservoirs, which are essential in supporting growing populations across the region (Bladon et al., 2014; Martin, 2016; Murphy et al., 2018).

2.2 | Data from the literature

Post-fire debris flows have been extensively studied across the IMW (e.g. Cannon & Gartner, 2005, 2010; Cannon et al., 2008; Langhans et al., 2016; Santi et al., 2008; Wondzell & King, 2003), yet very little data has been reported on debris flow GSDs and volumes in this region. While post-fire debris flow GSD data are sparse, some data have been reported in the literature (Figure 1, Table 1). Here, we compile all geolocated post-fire debris flow GSD and volume data that we

TABLE 1 Wildfires examined in this study, including fire year, location (i.e. US state), data source, and the data collected or reported on from each location. For each site, the number of debris flows examined is reported in the table

Fire	Fire year	State	Source	Volume	Surface and subsurface GSD	Only subsurface GSD	Only surface GSD	Boulder D84
Yellowstone Fires	1988	WY/MT	Meyer and Wells (1997)	1	n/a	3	n/a	n/a
Dinosaur Ntl Monument	1996–2002	CO	Larsen (2003), Martin (2000)	6	n/a	n/a	9	n/a
Oak Hills	2000	UT	Giraud and McDonald (2009)	1	n/a	n/a	n/a	n/a
Farmington	2000	UT	Giraud and McDonald (2009)	2	n/a	n/a	n/a	n/a
Borrow Pit	2001	UT	Giraud and McDonald (2009)	1	n/a	n/a	n/a	n/a
Mollie Fire	2001	UT	Giraud and McDonald (2009)	5	n/a	n/a	n/a	n/a
Springville	2002	UT	Giraud and McDonald (2009)	1	n/a	n/a	n/a	n/a
Twitchell	2010	UT	Murphy et al. (2019)	n/a	2	7	n/a	n/a
Clay Springs	2012	UT	Fieldwork	2	2	n/a	n/a	2
Seeley	2012	UT	Fieldwork	1	1	n/a	n/a	1
Shingle	2012	UT	Fieldwork	1	1	n/a	n/a	1
Brianhead	2017	UT	Fieldwork	4	4	n/a	n/a	4
Tank Hollow	2017	UT	Fieldwork	2	2	n/a	n/a	2
Trail Mountain	2017	UT	Fieldwork	4	4	n/a	n/a	4
Coal Hollow	2018	UT	Fieldwork	2	2	n/a	n/a	2
Dollar Ridge	2018	UT	Fieldwork	5	5	n/a	n/a	5
Pole Creek	2018	UT	Fieldwork	5	5	n/a	n/a	5
West Valley	2018	UT	Fieldwork	3	3	n/a	n/a	3
Woodbury	2019	AZ	McGuire and Youberg (2020)	1	n/a	n/a	1	n/a
Total				**47**	**31**	**10**	**10**	**29**

could find reported in the literature from this region into one dataset for our analysis (Figure 1, Table 1). From the literature we compiled 17 post-fire debris flow volume measurements and 22 GSD measurements across fires in Utah, Arizona, Colorado, Montana, and Wyoming (Figure 1).

2.3 | Field study sites

We identified and conducted fieldwork for an additional 29 post-fire debris flow deposits from 10 fires that occurred between 2012 and 2018 across the state of Utah (Figure 1, Table 1). The contributing catchments for these sites span a wide array of lithology, climate, burn severity, and vegetation characteristics, and have drainage areas ranging from 0.1 to 10 km^2. Ten percent of the catchments have underlying unconsolidated material, 25% of the catchments have underlying igneous rock, and 64% of the catchments have underlying sedimentary rock (Hill et al., 2015). The average elevations of the study catchments range from 1600 to 3000 m and include barren land, conifer forest, deciduous forest, mixed forest, grassland, wetland, shrubland, and agricultural land (Hill et al., 2015).

3 | METHODS

We developed predictive models for post-fire debris flow characterization using a combination of field, geospatial, and statistical analysis, as detailed below.

3.1 | Measurement of variables

3.1.1 | Deposit volumes

To constrain the debris flow deposit volume at each site we measured, we used a combination of field surveying and 3D reconstruction methods. First, using a handheld GPS device, we surveyed the perimeter of the deposit to measure the planview area. We then measured the depth of the deposit at several locations within the deposit perimeter. To constrain uncertainty in debris flow volumes, we recorded detailed observations about the locations and potential magnitudes of erosion that may have occurred since initial deposition. We used the combination of mapped extent and spatially dispersed measurements of sediment depth to create 3D models of each deposit.

Specifically, we generated a triangulated irregular network (TIN) for each deposit to minimize potential overfitting between depth values that occur with other interpolation methods. Using the TIN surfaces, we calculated the volume for each debris flow deposit. Finally, we estimated the percentage error in volumes using bounds recommended for this method (i.e. −25 to +30% as per Santi, 2014).

The post-fire debris flow volume data compiled from the literature was measured using a variety of techniques. Meyer and Wells (1997) mapped debris flow deposits in Yellowstone National Park using air photos combined with compass and tape methods to map the perimeter. They then created an isopach map with approximately 50 deposit thickness measurements to estimate deposit volume. In contrast, Martin (2000), Larsen (2003), and Giraud and McDonald (2009) estimated debris flow volumes using deposit area and the average of depth measurements (these data were used in the construction of the Gartner et al., 2008 Rocky Mountain model). In total, we compiled a dataset of 47 debris flow volumes located in Utah, Wyoming, Montana, Colorado, and Arizona, which represents the largest geolocated dataset of post-fire debris flow volumes we are aware of outside of Southern California (Table 1, Figure 1).

3.1.2 | Grain size distributions

For all 29 debris flow deposits we visited, we measured both the surface and subsurface GSDs, as well as the largest boulders. At each deposit, we conducted at least one random walk Wolman pebble count with a minimum of 100 measurements (Wolman, 1954), as well as a minimum of two subsurface sieve mass measurements. Additionally, we measured the *b*-axis of the 30 largest boulders observed on each debris flow deposit to constrain the coarsest end of the GSDs, which may have unique implications for aquatic habitat. This uppermost end of GSDs is also often not well characterized by Wolman pebble counts or sieving. These three methods of measuring grain sizes provide a reasonable characterization of GSDs at each debris flow deposit.

The post-fire debris flow GSD data compiled from the literature used similar but variable methods. Meyer and Wells (1997) collected subsurface grain size data in Yellowstone National Park using sieve mass measurements. Larsen (2003) measured surface GSDs in Dinosaur National Monument using Wolman pebble counts. Murphy et al. (2019) collected both surface and subsurface GSDs using sieve mass measurements and Wolman pebble counts in the Tushar Mountains of Utah. Lastly, McGuire et al. (2021) collected surface GSDs using Wolman pebble counts in the Superstition Mountains of Arizona. From the literature, we compiled a total of 31 post-fire debris flow grain size measurements, though 10 only had subsurface data and another 10 only had surface data. No previous study reported boulder sizes (Table 1, Figure 1).

3.1.3 | Catchment characteristics

We extracted catchment topographic metrics by analysing 10 m digital elevation models from the National Elevation Dataset (NED) in ArcGIS. Variables included contributing area, average catchment gradient, catchment area with slopes ≥ 23°, and mean catchment elevation.

Burn severity data for all fires examined in this study were sourced from the USGS Monitoring Trends in Burn Severity (MTBS) project (Finco et al., 2012). MTBS provides burn severity data classified as low, medium, and high. Using the classified severity rasters, we calculated potential predictor variables for our analysis, including percentage of a catchment burned at moderate and high severity and area of a catchment burned at moderate and high severity (as per Gartner et al., 2014).

Catchment characteristics, such as lithology, vegetation cover, and climate metrics, were extracted using the US EPA StreamCatchment (StreamCat) dataset (Hill et al., 2015). From this database, we extracted each catchment's mineralogical composition and lithological composition. Additional lithologic variables extracted from the database included compressive strength and hydraulic conductivity. Soil properties (pre-fire) extracted from StreamCat included average clay, sand, and organic matter (OM) content, average soil permeability, soil depth to bedrock, and the soil erodibility factor. Additionally, we extracted the percentage vegetation type for each catchment (e.g. percentage conifer, deciduous, shrub, grasslands). Climate metrics from this database include the 30-year average precipitation, mean annual runoff, 30-year mean annual temperature, and average seasonal water table depth. Finally, we extracted the average wetness index from StreamCat, which is a metric combining the contributing catchment area, slope, and flow paths (Hill et al., 2015; Kopecky et al., 2021).

The rainfall intensities of the storms that triggered the debris flows in our field dataset are unknown and are also unreported for most of the data gathered from the literature. Therefore, using the National Oceanic and Atmosphere Administration (NOAA) Precipitation Frequency Data Server (PFDS), we extracted the 10 and 15-min duration rainfall intensities for both the 2- and 100-year storm event in each catchment (more on this approach below).

3.2 | Evaluation of existing volume models

To evaluate the existing post-fire debris flow volume prediction models, we first calculated the volumetric estimates for each debris flow using the Gartner et al. (2008) Rocky Mountain model:

$$\ln(V) = 0.72(\ln S_{30}) - 0.02(i10) + 8.54 \quad (1)$$

and the Gartner et al. (2014) Emergency Assessment model developed for the western United States:

$$\ln(V) = 4.22 + 0.39\sqrt{i15} + 0.36(\ln Bmh) + 0.13\sqrt{R} \quad (2)$$

where V = predicted debris flow sediment volume (m^3), S_{30} = catchment area with slopes greater than or equal to 30% (km^2), $i10$ = peak 10-min rainfall intensity (mm/h), $i15$ = peak 15-min rainfall intensity (mm/h), Bmh = the catchment area burned at moderate and high severity (km^2), and R = catchment relief (m), computed as the maximum minus minimum elevation.

Extremely local and high temporal resolution rainfall data for the triggering storm event are needed to accurately apply the Gartner et al. (2008, 2014) models, but these data are often not available. Even when the data are available, it can still be challenging to

impossible to accurately attribute debris flow events to the specific rainfall event that triggered them. Therefore, in the absence of such data (or alternatively, in cases where these models might be used for pre-fire risk assessments), the best available data are the 1 km^2 resolution recurrence interval (RI) average rainfall intensity data from NOAA's Precipitation Frequency Data Server (PFDS). Accordingly, we informed each Gartner et al. model using rainfall inputs representing the 2- and 100-year storm intensities of the appropriate duration (*i*10 or *i*15) from this source. Staley et al. (2020) suggested that 90% of documented post-fire debris flows in the western United States have been generated by storms with intensities of ≤5-year RI, so this range of RI intensities was chosen to reasonably constrain the possible upper and lower bounds of storm intensities that may have triggered the debris flows in our study. To evaluate the models, we then statistically compared the outputs for each storm event against the measured volumes for each deposit.

3.3 | Model construction

We developed predictive models for post-fire debris flow deposit grain sizes and volumes based on potential predictor variables characterizing catchment morphology, burn severity, climate, lithology, soil properties, and vegetation cover using a combination of machine learning and regression methods. To capture the central tendency and tails of the deposit GSDs, we constructed four separate models to predict a deposit's D16, D50, D84, and the D84 of boulders. Additionally, we constructed a volume prediction model that estimates debris flow sediment yield from burned catchments.

The initial array of predictor variables selected for inclusion in our analysis was based on previous debris flow probability and volume prediction models (Gartner et al., 2008, 2014; Staley et al., 2017), as well as additional catchment characteristics we hypothesized could influence debris flow grain sizes and volumes. We used consistent statistical methods to construct all of our GSD and volume models, which included an initial variable selection step using Random Forest modelling followed by predictive model construction using multiple linear regression (MLR).

3.3.1 | Random Forest

We first used Random Forest (RF) machine learning models to evaluate the relative importance of all included predictor variables (Breiman, 2001). RF models are a type of classification or regression tree analysis used to evaluate complex relationships between predictor and response variables by combining observations from an ensemble of trees (Cutler et al., 2007; Fisher et al., 2021; Vaughan et al., 2017). We initially removed any variables that demonstrated high covariance by using the variance inflation factor (VIF) and then examined the importance of the predictor variables using variable importance plots (Gemuer et al., 2012), which rank variables based on the mean decrease in model accuracy that would occur if they were removed from the RF model. From these outputs, we visually identified breaks in the variable importance plots (i.e. significant drops in the contributed accuracy) and selected the uppermost grouping of best predictor variables (ranging from 7 to 12 variables across our RF models). The relationships between each identified predictor and response variable were then individually evaluated based on both visual diagnostics and univariate linear regression.

3.3.2 | Multiple linear regression

We used MLR analyses to develop models for post-fire debris flow grain size metrics and deposit volumes. Based on data from the 41 post-fire debris flows with subsurface grain size data (Table 1), we developed models to predict three GSD metrics: the 16th percentile (D16), median (D50), and 84th percentile (D84). The model to predict the 84th percentile of boulder sizes (D84 boulders) was based only on data from the 29 deposits in our field study, as these were the only sites with boulder data. The volumetric prediction model was developed using a dataset that included deposit volumes from 41 post-fire debris flows (Table 1).

Requirements of linear regression analysis include: a linear relationship between the variables, normality in the residuals, and constant variance in the residuals. Therefore, we first transformed all of the GSD data into the phi scale (ϕ), or log base 2. To identify linear correlations between the predictor and response variables, we calculated the Pearson product–moment correlation coefficient (Helsel et al., 2020). We also attempted a natural log and square root transformation for each variable to investigate which produced the most linear relationship with the response variable. We moved forward in our analysis with the transformation that exhibited the highest correlation coefficient, the best linear visual diagnostics, and that met the necessary linear regression assumptions. To test for normality in the residuals we calculated a correlation between the observed residuals and the expected residuals under normality. Next, to test if there is constant variance in the residuals, we ran a Brown–Forsythe test, which has a null hypothesis that the residuals have constant variance. If the *p*-value is below $\alpha = 0.05$, then there is not constant variance in the residuals. After conducting these preliminary diagnostics on the predictor variables, we selected all the variables that met the necessary linear regression assumptions.

We next examined the narrowed response variables for multicollinearity, which can result in unreliable models with unstable and unrealistic parameter coefficients (Helsel et al., 2020). We used both visual diagnostics, such as examining correlation plots and correlation coefficients between predictors, and the VIF to diagnose multicollinearity. After identifying associated predictor variables, we kept the variable with the best diagnostics and highest correlation with the response variable and removed the other variable from the model.

Every possible combination of the predictor variables that met the above criteria were then assembled and examined as potential MLR models (combination sets included one to seven predictor variables, see the online Supporting Information). Seeking to identify the most parsimonious model for each response variable, we then selected the model that exhibited the best combination of low Akaike information criterion (AIC), which is an estimator of prediction error, and high R^2, while using the fewest number of predictor variables (Helsel et al., 2020).

3.4 | Model validation

Due to the small sample sizes of our datasets, we validated each of our selected grain size and volume models using a fivefold cross-

validation approach (Kohavi, 2001). For fivefold cross-validation, the dataset for each model is randomly shuffled and then split into five groups. The process loops through each group, holding it as the test dataset and using the remaining four groups as the training dataset. The model is fit on the training sets and evaluated on the test sets, and then the performance measures reported by each trial are averaged to report a cross-validated R^2 and root mean square error (RMSE). Ultimately, by using fivefold cross-validation we were able to evaluate whether or not overfitting of the MLR models occurred during the model construction.

4 | RESULTS AND DISCUSSION

4.1 | Grain size distribution field measurements

4.1.1 | Results

To investigate the grain size variance both within single deposits and between deposits, we examined grain size data from 51 different post-fire debris flow deposits (Table 1, Figure 2). First, comparing surface and subsurface metrics for the 31 deposits that contained both data types, we found that the surface was much coarser than the subsurface material (Figure 3). The average offset from the line representing the 1:1 relationship between the surface and subsurface distributions is 3.32 for D16, 1.15 for D50, and 0.27 for D84 (Figure 3b). This offset suggests the D16 of the surface is much coarser than the subsurface, while in contrast, the distributions for the D84 are roughly similar. Additionally, Figure 4 displays the grain size metric ranges (Figures 4a and c) and relative ranges in the phi scale (Figures 4b and d). When examining the grain size metrics for all 51 deposits in the phi scale, the D16 exhibits the greatest relative range compared to the D50 and D84, and the D84 has the least relative spread across all deposits (Figures 4b and d).

4.1.2 | Discussion

We observed inverse grading of the post-fire debris flow deposits. This type of grading is potentially the result of some combination of winnowing of fines from the deposit surface and from kinetic sieving during transport, as smaller grains pass through larger particles in

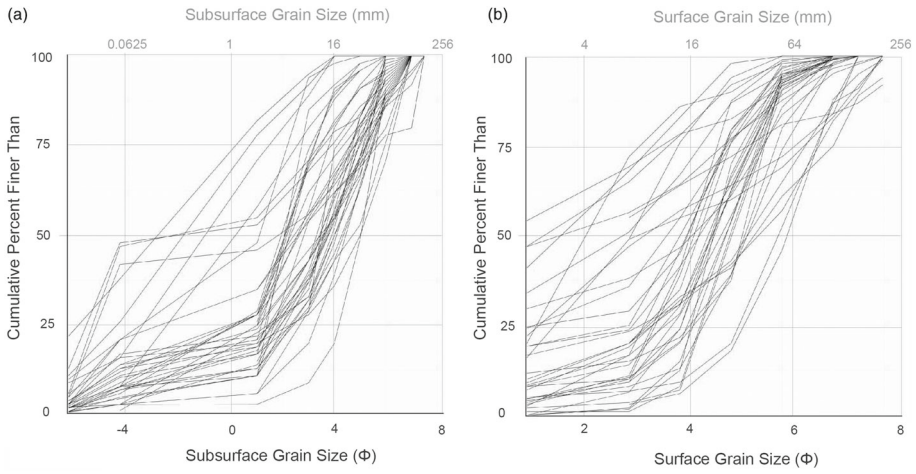

FIGURE 2 Cumulative distributions (CDFs) displaying the full GSDs for all post-fire debris flow deposits examined in this study. (a) CDF plot showing the GSDs of subsurface sediment measured by sieve mass measurements, and (b) CDF plot showing the GSDs of surface sediment measured by Wolman pebble counts.

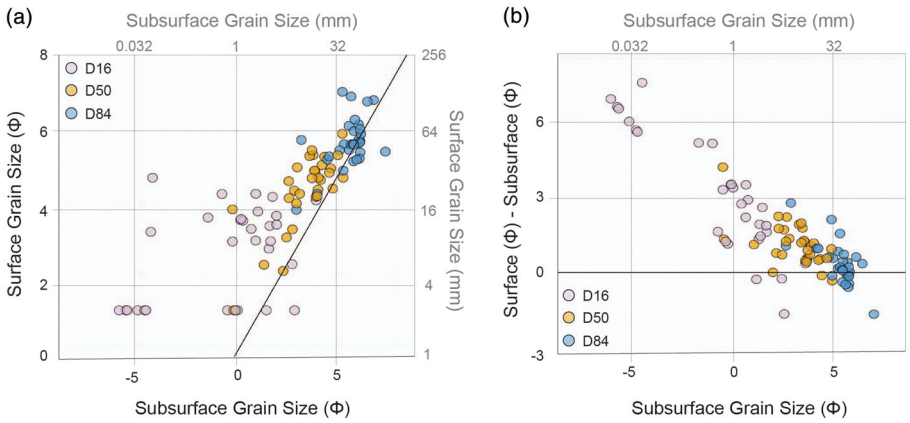

FIGURE 3 (a) The subsurface GSD metrics (D16, D50, D84) plotted against their respective surface GSD metric for the 31 debris flow deposits with both measurements. The black line represents the 1:1 relationship between the surface and subsurface distributions. (b) Plot showing the difference between the surface and subsurface values shown in panel (a). Metrics with equal values should plot along the black line (= 0 on *y*-axis). These results demonstrate that, for nearly every deposit measured, surface grain sizes are coarser than subsurface sizes in the same deposit across the full distribution of grain sizes. However, panel (b) highlights that this relationship is most pronounced in the finer tail of deposit distribution, and the difference between the surface and subsurface becomes smaller in the median and coarser tail of deposit grain sizes. [Color figure can be viewed at wileyonlinelibrary.com]

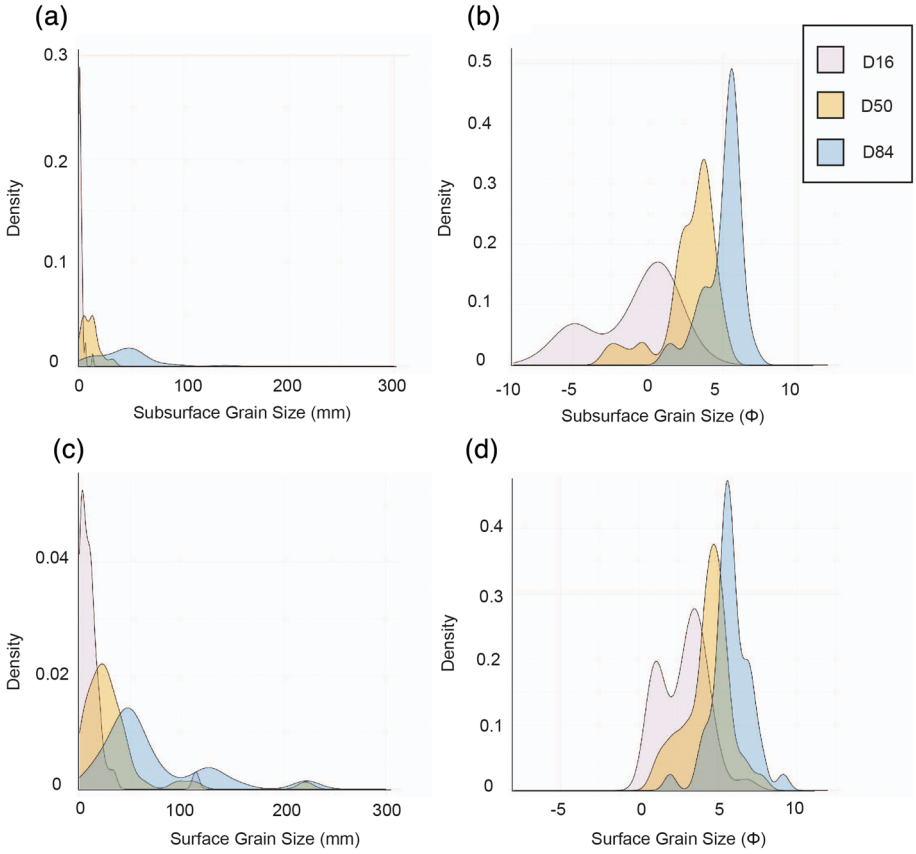

FIGURE 4 Probability density functions for the D16, D50, and D84 GSD metrics across all debris flow GSDs examined in this study. The upper row (panels a and b) shows the distribution of surface GSD metrics, and the bottom row (panels c and d) shows the distribution of subsurface GSD metrics. The left-hand column (panels a and c) shows the grain size values plotted in linear space, and the right-hand column (panels b and d) shows the grain size values plotted in the phi scale. These distributions show the overlap in the values of D16, D50, and D84 found across post-fire debris flows, highlighting that there are not distinct D16, D50, or D84 values for all debris flow deposits. [Color figure can be viewed at wileyonlinelibrary.com]

motion and displace these larger particles upwards (Betran, 2003; Naylor, 1980; Strom, 2006; Wang et al., 2012). While inverse grading has not previously been investigated or demonstrated specifically in post-fire debris flows, these findings are consistent with previous studies of debris flow rheology and sedimentology (Bridgewater, 1994; Naylor, 1980; Zou et al., 2017). It is worth noting, however, that while the identification of this phenomenon in our data is likely real, it is also likely influenced by the differing methods used to collect the data. Specifically, the Wolman pebble counts used for surface measurements are systematically biased against identifying finer sediments that would be captured by the sieving techniques.

These findings indicate that measuring the subsurface material is essential for obtaining accurate and wholistic characterizations of post-fire debris flow GSDs. Particularly in the context of understanding and predicting the impacts of post-fire sediment inputs to river systems, only measuring a deposit's surface material could result in a substantial overestimate of grain size metrics, particularly for the finer end of the GSD. These fine sediments would not only be more efficiently transported downstream if delivered to a river, but may degrade aquatic habitat and spawning grounds by infilling between river gravels (Brown et al., 2001; Gresswell, 1999; Propst & Stefferud, 1997). Thus, collecting grain size measurements of debris flow subsurface material is critical for understanding and predicting potential downstream impacts after fire.

4.2 | Predictive models

We developed four models to predict post-fire debris flow deposit GSD metrics and one volume prediction model. Due to the observed inverse grading of post-fire debris flow deposits, we developed our grain size models based on the subsurface grain size data from 41 post-fire debris flow deposits across a diversity of topographic, lithologic, and ecological conditions.

All predictor variables used in the construction of these models are publicly available and easy to obtain or derive using GIS software. General statistics were calculated for all potential predictor variables and include a correlation test for normality, the Brown–Forsythe test of constant variance, and the Pearson correlation coefficient (Table 2). Many models were developed and analysed during the MLR analysis and only the most parsimonious model is reported here. Each reported model exhibited a low AIC, high R^2, a low residual standard error, and no multicollinearity present between the predictors.

4.2.1 | Grain size models

Results

The four post-fire debris flow grain size models, including the model equation, R^2, fivefold cross-validated R^2, and sample size used to produce each model, are detailed in Table 3. Every predictor in the D16, D50, and D84 models is significant with a p-value ≤ 0.1.

The D16 model exhibits an R^2 of 0.59, a cross-validated R^2 of 0.60, and 71% of the predicted D16 values are within one residual standard error of the measured D16 value (Figure 5a). The D50 model has an R^2 of 0.71, a cross-validated R^2 of 0.69, and 76% of the predicted D50 values are within one residual standard error of the measured D50 value (Figure 5b). The D84 model has an R^2 of 0.60, a fivefold cross-validated R^2 of 0.56, and 76% of the predicted D84 values are within one residual standard error of the measured D84 value (Figure 5c). Finally, the D84 boulder model was constructed using data from 29 debris flow deposits, which were collected during

TABLE 2 Summary statistics and transformations of the predictor variables selected for the models, including the response variable related to each predictor, the correlation coefficient between predictor and response, and the *p*-value for significance of each predictor, the results from the Brown–Forsyth (BF) test for constant variance in the residuals, and results from the normality test of the residuals (normality of residuals). Refer to Section 3.3.2 for interpretations of these tests

Predictor variable	Response variable	Mean	St. dev.	Min	Max	Trans.	Corr. coeff.	Significance *p*-value	BF test	Normality of residuals
Mean annual temperature (°C)	D16	6.4	4.3	0.92	18.5	ln	0.54	<0.0001	0.59	0.94
	D50					ln	0.45	<0.0001	0.72	0.96
	D84					ln	0.38	0.025	0.98	0.97
Mean catchment elevation (m)	Boulder D84	2184	520	839	2,924	ln	0.23	0.26	0.85	0.99
% Catchment with slopes ≥ 23°	D16	52	26	0	95		−0.38	0.04	0.25	0.98
	D84					√	0.13	0.05	0.52	0.94
Catchment area with slopes ≥ 23° (km^2)	Volume	0.7	1.2	0	5.8	√	0.35	0.08	0.95	0.99
Catchment area burned at moderate and high severities (km^2)	Volume	0.7	1.7	0	9.1	√	0.3	0.02	0.14	0.99
Average catchment runoff (mm)	D84	120	112	17	463		−0.52	0.06	0.23	0.98
	Volume						0.41	0.08	0.74	0.99
Average catchment wetness index	D50	307	83	189	450	ln	0.25	0.07	0.48	0.92
	Boulder D84						−0.46	0.1	0.53	0.99
Average catchment soil permeability (cm/h)	D50	6	2.5	1	10.3		−0.46	0.0001	0.21	0.95
	D84						−0.37	0.0009	0.72	0.96
% Soil organic matter	Volume	0.8	0.3	0.2	1.4		0.49	0.0005	0.95	0.99
Soil depth to bedrock (cm)	D16	109	30.7	43	151		0.28	0.01	0.17	0.95
Average catchment rock compressive strength (mPa)	D50	79.3	26.7	30	153		−0.59	<0.0001	0.82	0.96
	D84						−0.59	0.0025	0.4	0.97
% Catchment lithological magnesium content	Boulder D84	2.9	1.4	1.1	6.1		0.42	0.2	0.97	0.99
% Catchment with conifer cover	D16	36	24.6	1.3	92		−0.49	0.03	0.05	0.97

TABLE 3 The best models found for predicting post-fire debris flow grain size metrics in the phi scale, including the equation, goodness-of-fit statistics, and sample size

Model	Equation	# Variables	R^2	*x*-val R^2	*n*
D16 (Φ)	$D16 = -7.21 + 2.83\ln(T) - 0.03Cp - 0.03Sp + 0.04Rd$	4	0.59	0.6	41
D50 (Φ)	$D50 = 5.52 - 0.036CS + 1.17\ln(T) - 0.35K$	3	0.71	0.69	41
D84 (Φ)	$D84 = 4.83 + 0.66\ln(T) - 0.004Ro - 0.27K + 0.16\sqrt{Sp}$	4	0.6	0.56	41
Boulder D84 (Φ)	$D84B = 3.75 - 0.002WI + 0.022\sqrt{Sp} + 0.093MgO + 0.75\ln E$	4	0.34	0.25	29

The variables are defined as T = mean annual temperature (°C), Cp = percentage area of the catchment with conifer cover, Sp = percentage area of the catchment with slopes ≥ 23°, Rd = average catchment soil depth to bedrock (cm), CS = average catchment rock compressive strength (MPa), K = average catchment soil permeability (cm/h), WI = average catchment wetness index, Ro = average catchment runoff (mm), MgO = average catchment lithological magnesium oxide content, E = mean catchment elevation.

fieldwork as part of this study. The D84 boulder model predicts the 84th percentile size of the boulders in a debris flow deposit and has an R^2 of 0.34, a cross-validated R^2 of 0.25, and 76% of the predicted values are within one residual standard error of the measured value (Figure 5d).

Discussion

Our results demonstrate that there are systematic landscape controls on post-fire debris flow GSDs that allow for the prediction of grain size metrics based on variables related to catchment land cover, physical and chemical weathering, and hillslope sediment transport processes. The four GSD models reported here are the first investigation into landscape controls on post-fire debris flow grain sizes. We examined 50 potential predictor variables, out of which we identified 10 variables as significant in the prediction of one or more key GSD metrics (Table 2). Notably, our analysis revealed that no wildfire-related metrics were identified as controls on the GSD of post-fire debris flows. This suggests debris flow GSDs are controlled by the grain sizes available for transport prior to the fire, rather than any possible hydrogeomorphic effects of the fire, which is consistent with previous observations in the literature (Kean et al., 2011, 2019).

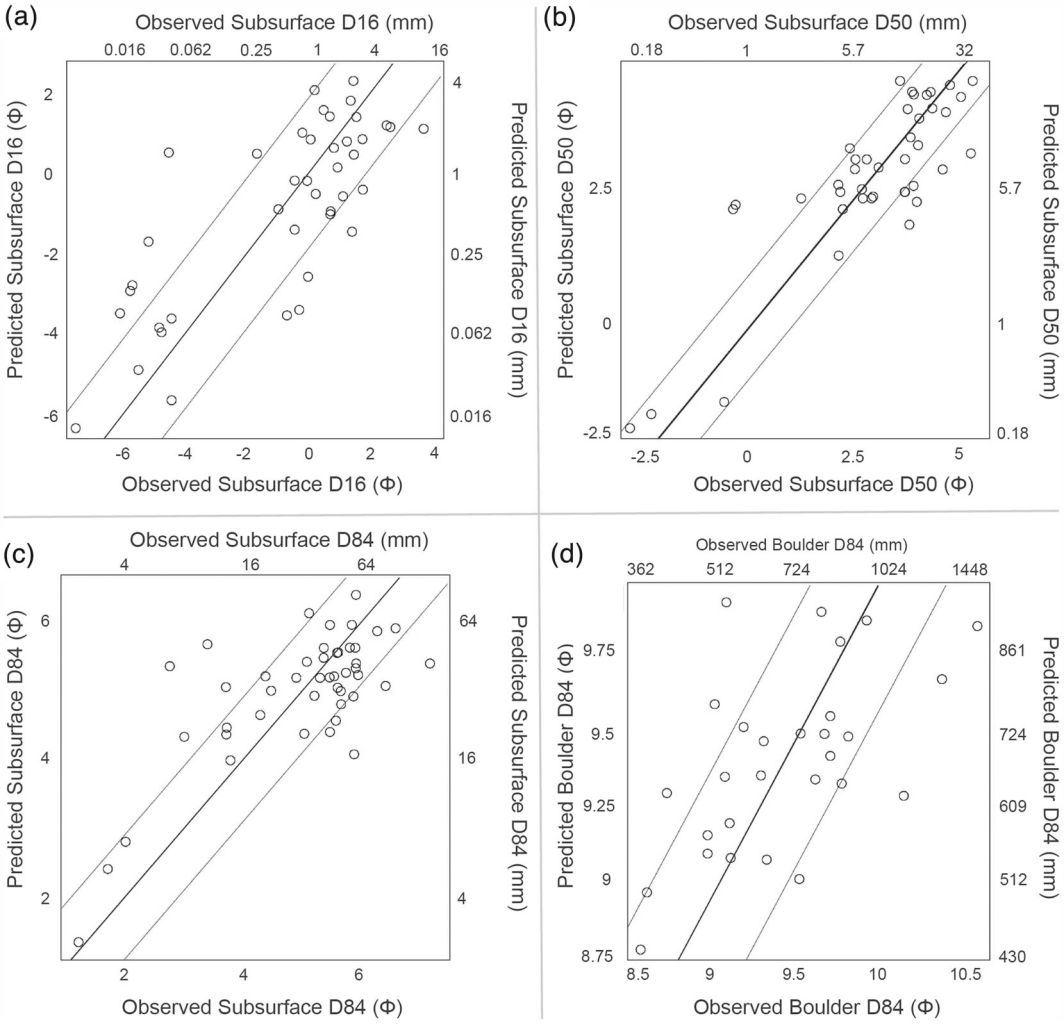

FIGURE 5 Plots showing the grain size values predicted by each model plotted against the measured subsurface grain size metrics for each debris flow deposit. Observed vs. predicted values for the D16 model (a), D50 model (b), D84 model (c), and boulder D84 model (d). The thick black line in each plot represents the 1:1 relationship, and the thin black lines above and below the 1:1 line represent the envelope for one residual standard error.

Variables that are known to influence weathering rates are present in the GSD models. For example, mean annual temperature (MAT) exhibits a significant positive relationship with D16, D50, and D84. MAT can be related to the depth and intensity of frost cracking, as it has been shown to be correlated with how much time an area spends in the frost-cracking window and the availability of water to contribute to segregation ice growth (Hales & Roering, 2007; Messenzehl et al., 2017). MAT can also drive chemical weathering rates and influence grain size; however, this effect should not be considered in isolation from local precipitation (Murphy et al., 2016; Sklar et al., 2017). Elevation is also a proxy for a location's exposure to physical and chemical weathering processes (Marshall & Sklar, 2012; Riebe et al., 2015; Sklar et al., 2016), and we found average catchment elevation to have a significant positive relationship with a deposit's boulder D84.

The percentage area of the catchment with slopes ≥23° was also found to be significant in the prediction of debris flow deposit D16 and D84, however, it exhibited a negative relationship with D16 and a positive relationship with D84. The positive relationship between D84 and slope is consistent with findings in the literature (Attal et al., 2015; Riebe et al., 2015; Whittaker et al., 2010). Attal et al. (2015) found that low-gradient slopes increase residence times, exposing particles to weathering processes for longer periods of time, resulting in the production of finer grain sizes. In contrast, the relationship between D16 and slope is negative. This could indicate that slope is not acting as a proxy for weathering processes in its control on D16, but instead reflecting debris flow transport processes for the most transportable material. Terrain steepness encourages efficiency of runoff-related erosion and sediment transport processes, which has been found to directly influence the shear stress of overland, rill, and channelized flow erosion (Cannon, 2001; Cannon et al., 2003; Prancevic & Lamb, 2015; Prancevic et al., 2014; Santi et al., 2008; Staley et al., 2017). It is possible that more overland flow and rill erosion on steeper slopes could result in the entrainment, transport and contribution of more fines to debris flows and produce a finer D16.

The catchment's average rock compressive strength has a significant negative relationship with debris flow deposit D50. We expected rock strength to be an important variable in grain size prediction, however, the negative relationship between rock strength and D50 is contrary to relationships previously documented in the literature (Allen et al., 2015; Marshall & Sklar, 2012; Roda-Boluda et al., 2018; Sklar et al., 2016). This discrepancy could result from limitations associated

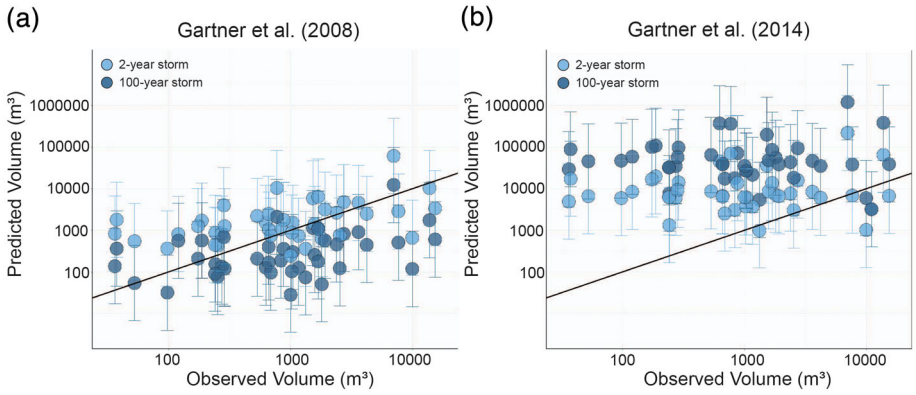

FIGURE 6 Post-fire debris flow volume predictions by the Gartner et al. (2008) model (a) and the Gartner et al. (2014) model (b) under 2-year storm (light blue) and 100-year storm (dark blue) scenarios plotted against the observed debris flow volumes. The error bars represent model error of two standard errors of the estimate. The thick black line shows the 1:1 relationship. The Gartner et al. (2008) model employs a negative relationship between rainfall intensity and debris flow volume (such that the 100-year events exhibit smaller predicted volumes), while the Gartner et al. (2014) model employs a positive relationship between rainfall intensity and debris flow volume. [Color figure can be viewed at wileyonlinelibrary.com]

with the estimation and aggregation of catchment-average rock compressive strength in StreamCat data. Examining the average compressive rock strength for each contributing catchment (i.e. average rock strength across 0.1–10 km^2) could be too coarse to capture the nuances associated with where sediment is being sourced from within the catchment and what the corresponding rock strengths are for those areas. Future work should further examine this relationship and which metrics for rock strength may best predict deposit GSDs.

These models fill a key knowledge gap in the development of reliable post-fire, watershed-scale models that incorporate debris flows (e.g. Murphy et al., 2019). Characterizing the GSD controls of post-fire debris flows is critical to predicting their impacts on downstream resources, because grain size exerts a first-order control on fluvial transport (e.g. Ahammad et al., 2021). These new models will advance our ability to predict how post-fire debris flows may impact aquatic habitat, as erosional inputs can either enhance or degrade habitat depending on the grain size composition (Brown et al., 2001; Burton, 2005; Gresswell, 1999). Additionally, since no wildfire metrics were significant in predicting post-fire debris flow grain sizes, these models could easily be implemented for pre-fire risk assessments. Specifically, by not requiring any knowledge of fire conditions, these models could be used in advance of fires to help identify which catchments might contribute large boulders, or alternatively very fine sediment, if they were to burn and produce debris flows. This predictive power could help inform pre-fire resource management and/or post-fire risk management and mitigation, since the relative grain sizes produced by the debris flows could have positive or negative implications for aquatic habitat or downstream resources and infrastructure.

4.2.2 | Evaluation of previous volume models

Results

Of the two Gartner et al. (2008, 2014) models that we evaluated, we found that the Gartner et al. (2008) Rocky Mountain model provided the best predictions of volumes for the IMW debris flows in this study. Specifically, the Rocky Mountain model with 2-year RI performed best (RMSE = 3292 m^3), followed by the Rocky Mountain model with 100-year RI (RMSE = 3698 m^3). Given our approach and assumptions using RI rainfall intensity inputs for the models, we expected the volume predictions from the two rainfall scenarios would roughly straddle the 1:1 observed vs. predicted line (i.e. the triggering storm RI was likely between the 2- and 100-year storm for most events). Using this approach with the Gartner et al. (2008) Rocky Mountain model, we found this to be the case for 48% of debris flows, with 21% overpredicted and 31% underpredicted (Figure 6A).

Eighty-five percent of debris flow volumes predicted by the 2-year RI Rocky Mountain model were within one order of magnitude of observed values. While there was a slight positive trend in the plot of residuals vs. fitted values ($m = 0.34$, $R^2 = 0.05$), we found that 100% of debris flows with observed volumes <1000 m^3 ($n = 21$) were overpredicted, and 80% of debris flows with observed volumes >4000 m^3 were underpredicted by on average a factor of 2.5 (range 0.25–15×). The majority of debris flows with observed volumes ≥1000 m^3 were underpredicted (13 of 20) but all by less than one order of magnitude. By comparison, the 100-year RI Rocky Mountain model underpredicted the volumes of 80% of debris flows, but 83% of predictions were within one order of magnitude. There was a slightly more positive trend in the residuals vs. fitted values plot ($m = 0.5$, $R^2 = 0.13$), however the majority of overpredictions (70%) were in debris flows with observed volumes <250 m^3. These overpredictions were 1.2–9.6× their observed volumes, but on average were only overestimated by 213 m^3.

Recognizing that our assessment does not use the triggering storm intensity data as intended, we rearranged the Gartner et al. (2008) Rocky Mountain model, substituted measured volumes for predicted volumes [Equation 1)], and estimated the *i*10 rainfall intensity the model would require us to accurately predict the volume of each measured debris flow. This analysis allowed us to further evaluate the model, as well as our approach of using the 2- and 100-year RI rainfall intensities. Consistent with interpretations from Figure 6A, we found that approximately half (48%) of the estimated rainfall intensities were between the 2- and 100-year RI events, with 21% greater than the 100-year and 31% less than the 2-year average rainfall intensity (Figure 7A). While this lends support to our approach, it is not consistent with Staley et al. (2020), who found that 77% of post-fire debris flows are triggered by storms of ≤2-year RI intensity. Notably, this analysis also highlights that events requiring a rainfall intensity

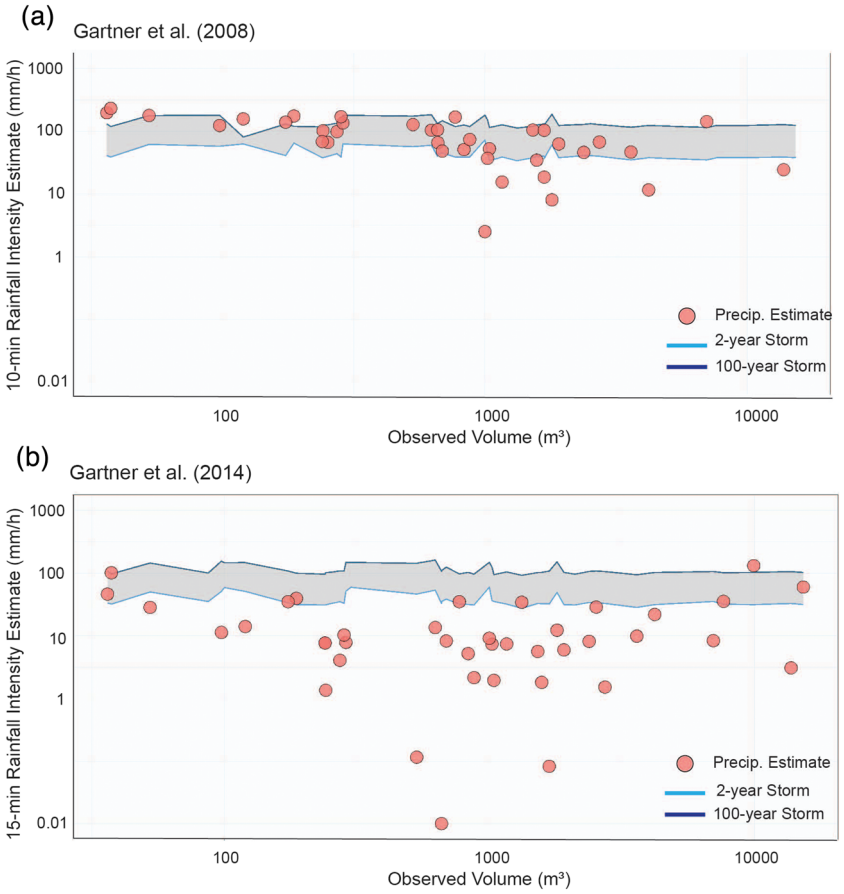

FIGURE 7 (a) The 10-min rainfall intensities of triggering storm events estimated by rearranging the Gartner et al. (2008) model for debris flow volumes in our Intermountain West dataset (red circles). (b) The 15-min rainfall intensities of triggering storm events estimated by rearranging the Gartner et al. (2014) model for debris flow volumes in our Intermountain West dataset (red circles). The grey shaded region represents the range of rainfall intensities between the 2-year storm event (light blue line) and the 100-year storm event (dark blue line) for each debris flow catchment. [Color figure can be viewed at wileyonlinelibrary.com]

RI > 100 years by the Gartner et al. (2008) model were predominantly the smaller debris flows (<1000 m^3). Although perhaps counterintuitive, this reflects the negative relationship between rainfall intensity and debris flow volume in the model [i.e. Equation 1) indicates that more intense storms produce smaller debris flows for any given slope condition]. Further, this analysis suggests that, if we assume the Gartner et al. (2008) Rocky Mountain model would perfectly predict the measured debris flow volumes had we known the triggering rainfall intensities (i.e. our back-calculated storm intensities are accurate), there is in fact no apparent relationship between post-fire debris flow volume and rainfall intensity (Figure 7A). This apparent lack of relationship is also observed with *i*15 values when the same analysis was performed with the Gartner et al. (2014) model (Figure 7B).

Using the 2- and 100-year RI 15-min rainfall intensities as model inputs, we found the Gartner et al. (2014) model performed with significantly lower accuracy (RMSE = 18 255 and 116 761 m^3, respectively) than the respective Gartner et al. (2008) model (Figure 6B). Almost 90% of debris flow volumes were overpredicted by the 2-year RI model, with 56% of predictions within one order of magnitude. For the 100-year RI model, 97% of volumes were overpredicted, with only 15% of predictions within one order of magnitude. In contrast to the Gartner et al. (2008) Rocky Mountain model, only three debris flows (or 8%) had observed volumes between the 2- and 100-year RI predictions; two of these were among the four largest observed debris flows (>7500 m^3). The Gartner et al. (2014) models also exhibited the most significant and positive trends in the residuals vs. fitted values plot (2-year RI: $m = 1.05$, $R^2 = 0.31$; 100-year RI: $m = 1.09$, $R^2 = 0.33$). This indicates an increasing degree of overprediction with the increasing size of predicted debris flows, however the largest relative overpredictions occurred in debris flows with observed volumes between 250 and 1000 m^3.

In the absence of triggering rainfall data, one possible explanation for the significant overprediction found with the Gartner et al. (2014) model could be that the events in fact all occurred with rainfall intensities much smaller than that of the 2-year RI storm, a conclusion that would be consistent with Staley et al. (2020). However, this would be contradictory to our analysis with the Gartner et al. (2008) Rocky Mountain model, which was developed specifically for this region, performed significantly better, and indicated that more than two-thirds of events would have been triggered by a storm with a >2-year RI.

Discussion

A notable limitation in our ability to definitively evaluate the accuracy of the Gartner et al. (2008, 2014) models was a lack of rainfall data for the storm events that triggered each debris flow in our dataset. However, this issue highlights an inherent challenge in validating these models in most landscapes. Dangerous and destructive post-fire debris flows are occurring in many regions and countries, yet the necessary rainfall data are often non-existent, incomplete, or too coarse in spatiotemporal resolution to identify the potential magnitude of the triggering storm. Moreover, even where adequate rainfall gauging and data are available, researchers must still be able to confidently identify the exact timing of the debris flows in order to attribute them to a measured storm intensity. Most wildfires burn in remote terrain, so unless a debris flow directly impacts lives, property, or infrastructure, the documentation and attribution to an exact hour, day, or even week can be difficult to impossible.

Therefore, we evaluated the Gartner et al. (2008, 2014) models under an assumption that most debris flows in the IMW were likely triggered by storms at or between the average 2- and 100-year RI rainfall intensities. We recognize this is not the intended approach to parameterize precipitation in the models and that the RI of average rainfall intensities may not represent those of peak storm intensities; however, in addition to our common data limitations, we highlight that in the development of the Gartner et al. (2008) Rocky Mountain model, the authors themselves did not find rainfall to be a significant predictor. Rather, the authors state that this variable was forced into their model. Presumably as a result of this decision, the magnitude of the regression coefficient for the rainfall intensity variable was not only negative but also close to zero [i.e. −0.02 in Equation 1)]. To further examine the role of precipitation in the Gartner et al. (2008) model, we analysed the dataset used in the original construction of the model (Garter, 2005) and ran a univariate regression of their data with only the slope parameter included and no rainfall metric. We found that forcing rainfall intensity into the model improved the explained variance of their debris flow volumes by 7%. Additionally, including rainfall intensity produced a model with a negative relationship between rainfall intensity and debris flow volume. This relationship is difficult to reconcile in terms of geomorphic processes and is notably inconsistent with other post-fire debris flow volume models that include rainfall metrics as a predictor (e.g. Gartner et al., 2008, western United States model; Gartner et al., 2008, Southern California model; Cannon et al., 2010; Gartner et al., 2014, Emergency Assessment model).

While other models report a positive and more significant relationship between rainfall and debris flow volumes than the Gartner et al. (2008) Rocky Mountain model, most are heavily influenced by Southern California data. Assuming one of the Gartner et al. (2008, 2014) models evaluated here would have produced accurate results had we known the triggering storm intensities, we back-calculated rainfall intensities and found neither model exhibits a significant trend with our IMW post-fire debris flow volumes. Overall, our findings suggest it is possible that the assumption of a scaling relationship between precipitation intensity and debris flow volumes may not be appropriate for post-fire debris flows in the IMW. Future work should explore whether this relationship is unique to post-fire debris flows in the Transverse Ranges of Southern California or if there are other geologic/tectonic settings that exhibit a predictable relationship between rainfall intensity and debris flow volume. Regionally specific models have frequently been suggested for the western United States, and there are many geologic, tectonic, climatic, and geographical differences that could likely influence debris flow processes.

We found that the commonly used Gartner et al. (2014) model significantly overpredicted the volumes of debris flows in our IMW dataset. While likely influenced by the previously mentioned geographic differences, we highlight that statistical discrepancies could also potentially contribute to the lack of performance with the Gartner et al. (2014) model. Specifically, there is a large disparity between the size of debris flows examined in this study and those from Southern California, analysed by Gartner et al. (2014) (Table 4). The average volume used in the construction of their model was nearly 30-fold larger than the average in our IMW dataset. While this is possibly another strong indicator that there are in fact distinct geographic differences, this statistical disparity could nonetheless contribute to our overpredictions, particularly given that the most significant residuals were observed in debris flows with the smallest volumes.

Finally, it is also possible that temporal effects could have influenced our results when applying the Gartner et al. (2014) model. Specifically, the Gartner et al. (2014) model was constructed using data from debris flows that were all known to have occurred within 2 years post-fire. In contrast, we can only confirm that 50% of the debris flows in our dataset occurred within 2 years following a fire. We cannot confirm whether this was true or not for the remaining debris flows, because we do not have documentation regarding the timing of their occurrence. It is possible that this could introduce additional variance in the predictions from the Gartner et al. (2014) model, as debris flow volumes have been found to be largest shortly after a fire and then decrease over time (Santi & Morandi, 2013).

4.2.3 | New debris flow volume prediction model

Results

Following the assessment of existing models, we developed a new model to predict sediment yield potential for debris flows produced from burned catchments in the IMW. Table 4 presents our volume prediction model compared to the Gartner et al. (2008, 2014) models, including the respective model equations and sample sizes of original datasets.

We constructed the new IMW volume model using the same approaches as with the grain size models, but with a dataset of 41 measured deposit volumes. The model predicts the natural log of the debris flow deposit volume (m^3) as a function of the catchment's pre-fire percentage soil organic matter (SOM) content, the square root of the catchment area with slopes ≥23° (km^2), the square root of the catchment burned at moderate and high severity (km^2), and the average catchment runoff (mm). Every predictor in our volume prediction model is significant using a significance threshold of ≤0.1. The model exhibited an $R^2 = 0.51$ and a cross-validated $R^2 = 0.55$. This similarity in cross-validated and general R^2 value indicates the model is not overfitted to the dataset. In comparison to the evaluated Gartner et al. (2008, 2014) models, our volume model exhibited the lowest RMSE (= 2941 m^3).

Additionally, 97% of predicted volumes were within one order of magnitude of the observed values (Figure 8), and there was no detectable trend or variance explained in the residuals vs. fitted values plot ($m = 0.0$, $R^2 = 1e\text{-}05$). While our model did overpredict 89% of debris flows <250 m^3, it was on average by just 200 m^3 (compared to 825 m^3 for the best-performing 2-year RI storm with Gartner et al., 2008). Our model underpredicted volumes for all four of the largest debris flows (those with observed volumes between 7500 and 15 000 m^3). However, this was also true of the Gartner et al. (2008) models, and our average underprediction was equivalent to that of the 2-year RI in the Gartner et al. (2008) model (~7400 m^3).

Discussion

We constructed a new post-fire debris flow volume prediction model using the largest dataset of geolocated post-fire debris flow volumes that we are aware of outside Southern California. Our model was developed and assessed with data solely from the IMW and included

TABLE 4 Equations for the post-fire debris flow volume prediction models from Gartner et al. (2008, 2014), as well as the new model developed in this study. Fit diagnostics (R^2 and cross-validated R^2) represent the results for our model applied to the Intermountain West debris flow volume dataset compiled in this study. Dataset sizes (n), along with minimum, mean, and maximum volumes of debris flows, represent those data used in the original development for each model

		Model development dataset				Applied to IMW dataset	
Model	Equation	n	Min	Mean	Max	R^2	x-val. R^2
Gartner et al. (2008)	$\ln V = 0.72(\ln S_{30}) - 0.02(i10) + 8.54$	17	n/a	n/a	n/a	n/a	n/a
Gartner et al. (2014)	$\ln V = 4.22 + 0.39\sqrt{i15} + 0.36(\ln Bmh) + 0.13\sqrt{R}$	79	29	46 200	864 300	n/a	n/a
Volume prediction model (this study)	$\ln V = 2.7 + 1.9Om + 0.175\sqrt{S_{23}} + 0.8\sqrt{Bmh} + 0.003Ro$	41	36	1600	13 750	0.5	0.55

The variables are defined as V = volume of a debris flow (m^3), S_{30} = catchment area with slopes ≥ 30% (km^2), $i10$ = peak 10-min rainfall intensity (mm/h), $i15$ = peak 15-min rainfall intensity (mm/h), Bmh = catchment area burned at moderate and high severity (km^2), R = catchment relief (m), S_{23} = catchment area with slopes ≥ 23°, Om = percentage soil organic matter, Ro = average catchment runoff (mm).

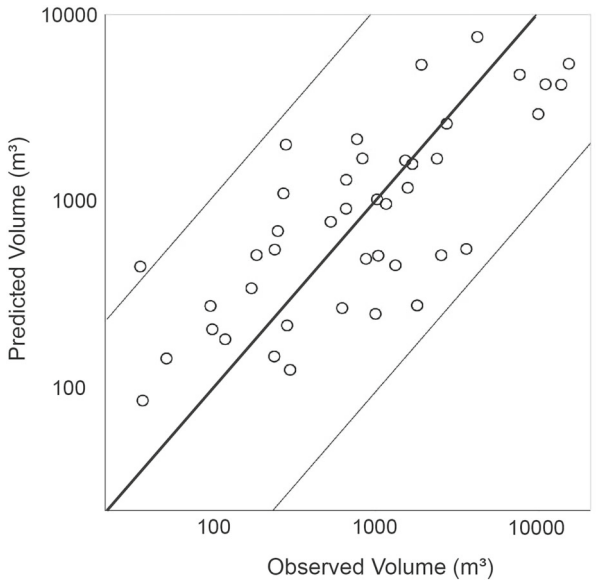

FIGURE 8 Post-fire debris flow volumes predicted using the new volume prediction model plotted against measured volumes (Table 4). The thick black line represents the 1:1 relationship. The thin lines represent the envelope of one residual standard error. The model exhibits an $R^2 = 0.51$ and a cross-validated $R^2 = 0.55$.

many of the same sites used to construct the Gartner et al. (2008) Rocky Mountain model, plus many more. Notably, our model's predictor variables were all identified and selected for inclusion by machine learning algorithms that assessed a large set of potential metrics; and in contrast with the Gartner et al. (2008, 2014) models, our model does not include any precipitation information related to the triggering storm event (as it was not known). Further, across all assessed statistical metrics, our rainfall-less, post-fire debris flow volume model performed as well or better than either of the Gartner et al. (2008, 2014) models when they were run and evaluated using a realistic range of RI storm intensities, rather than triggering storm intensities.

The variables identified and used to predict volumes in our model reflect upslope characteristics that are all previously understood to influence post-fire debris flow process (e.g. Cannon et al., 2010; Gartner et al., 2008, 2014; Pietraszek, 2006; Prancevic & Lamb, 2015; Prancevic et al., 2014; Staley et al., 2017). They include: the catchment area with slopes ≥23°, the catchment area burned at moderate and high severity, average catchment runoff, and soil organic matter content. While average catchment runoff has not specifically been included in previous volume models, it is well understood that post-fire debris flows are generated by runoff, and it stands to reason that for any given burn severity, post-fire runoff may reasonably scale with the average pre-fire runoff conditions. Our model did not improve with the inclusion of local precipitation metrics, though we only evaluated average rainfall metrics and none specific to the triggering storm event. While some previous research has found that short-duration intensities are more important for post-fire debris flow generation and sediment yield (Kean et al., 2011; Staley et al., 2013), Cannon et al. (2010) found average storm intensity to be significant for five post-fire debris flow volume prediction models.

Pre-fire OM content was found to be a particularly significant predictor for debris flow volumes, exhibiting a positive relationship. Cannon et al. (2010) also found a significant relationship between OM and post-fire debris flow volumes in two of their models, though it is not a variable in either of the Gartner et al. (2008, 2014) models. Additionally, Rupert et al. (2008) reported OM to be the strongest predictor in their post-fire debris flow generation models based on data from Southern California. OM is an important constituent for maintaining soil structure, and the quantity of OM has been observed to directly control soil aggregation (Mataix-Solera et al., 2011). When soils are burned, particularly at high severity, the severity of soil water repellency (or hydrophobicity) is dependent on the pre-fire OM content, along with soil texture and soil water content (DeBano, 1981, 2000; DeBano et al., 1979; Doerr et al., 2009). Additionally, the degree of change in soil structure, hydrophobicity, and soil infiltration capacity after wildfire has been directly linked to the quantity of OM incinerated by the fire (DeBano, 2000; DeBano et al., 1979). It is thus likely that the observed relationships between pre-fire OM and post-fire debris flow volume reflects the magnitude of post-fire hydrologic response due to incineration of OM. While the magnitude of change in soil conditions and resulting surface runoff may depend on the soil burn severity for any given OM content, burn severity is also included as a predictor in our model.

We constructed a new model for IMW post-fire debris flow volumes that exhibits good predictive capability and can be used in the absence of triggering storm event rainfall data. It is possible that the inclusion of triggering storm rainfall metrics would further improve our model's accuracy, but this data is notoriously difficult to obtain.

Despite this reality, most commonly used post-fire debris flow volume prediction models rely on metrics that require varying information about the storm that triggered the debris flow. In four of the models presented by Gartner et al. (2008), they found that volume had a significant positive relationship with the total storm rainfall. In their Rocky Mountain model, no precipitation metrics were found to be significant, however, they forced the peak 10-min rainfall intensity into the model, resulting in a weak and negative relationship. They also found no precipitation metrics to be significant in their Sedimentary Rock model. Cannon et al. (2010) reported the average storm intensity from the triggering storm event to have a significant positive relationship with post-fire debris flow volume for five of their volume prediction models. Gartner et al. (2014) found volume to have a significant positive relationship with the triggering storm peak 15-min rainfall intensity, which is consistent with the rainfall metric found to influence debris flow generation (Staley et al., 2017). Nyman et al. (2015) evaluated the Gartner et al. (2008) western United States model, which uses total storm rainfall as a predictor, for a burned area in southeast Australia. Although they reported this model worked well for predicting their observed volumes, they also found that it only explained 3% more of the variance in their volumes than an extremely parsimonious, relief-only model that included no precipitation metrics (Nyman et al., 2015).

Across all these studies and models, there is no consensus or clearly demonstrated and consistent relationship between post-fire debris flow volumes and the characteristics of the storm that generate them. Rather, despite the often-presumed positive relationship between rainfall and post-fire debris flow volume, previous work in this area has presented wildly contradictory relationships, proposed an array of potential storm metrics, and sometimes predicted volumes without the need of any precipitation metric. Although we know rainfall intensity is a critical predictor for debris flow generation (Staley et al., 2017), there is no similarly compelling evidence to demonstrate if or how post-fire debris flow volumes scale with rainfall. We posit that the role of rainfall may instead serve as more of an occurrence/non-occurrence predictor linked to debris flow initiation, rather than a scaling variable for volume. In essence, rainfall is required in that it controls the ability of a burned catchment to generate a debris flow, but the debris flow volumes produced from catchments across the IMW may be insensitive to rainfall intensity because they are often limited by the sediment availability along the debris flow path (i.e. in the channel). In line with this proposal, debris flow channels in the IMW are often observed to erode down to the underlying bedrock (e.g. DeLong et al., 2018; Murphy et al., 2019).

Moreover, rainfall thresholds have frequently been identified to control the generation of debris flows (Cannon et al., 2011; McGuire & Youberg, 2020; Staley et al., 2013, 2017) and have largely been found to occur with low-RI storm events (Staley et al., 2020). Across the western United States, there is a relatively small range of storm intensities responsible for producing post-fire debris flows that have volumes spanning many orders of magnitude (~80% triggering $i15$ are <55 mm/h; Staley et al., 2020). While higher-intensity storm events may be capable of producing larger debris flows, the reality may be that burned landscapes are so hydrologically responsive that lower-RI events often trigger debris flows and evacuate most available sediment, at least in the IMW, before the lower-probability but higher-intensity storm events occur. Thus, despite the intuitive hydrogeomorphic link, these low-rainfall thresholds may limit our ability to document and interpret this relationship from empirical datasets. Additionally, the low range of triggering rainfall intensities relative to that of debris flow volumes presents potential statistical issues with empirical observation. Even small levels of uncertainty or error in rainfall intensity values, which are often from gauges many kilometres away from the debris flows (e.g. Staley et al., 2017), could degrade and propagate large uncertainty into empirically derived relationships. These issues may contribute to the inconsistent and contradictory relationships previously observed, and ultimately, more work is required on the relationships between post-fire debris flow volume, the characteristics of the triggering storm events, and sediment availability in the contributing catchments.

This study provides a new approach and insights into the potential controls on and predictions of post-fire debris flow volumes in the IMW. To construct our volume model, we used a strictly statistical approach rather than process-based methods. Nonetheless, the predictor variables and relationships in our volume model all have robust, process-based explanations. Specifically, we find that greater volumes of debris flow sediment are predicted when larger areas of a catchment burn at moderate to high severity, where there are larger areas of the catchment with slopes ≥23°, where there is more organic material in the soil before the fire, and where there is more predicted catchment runoff based on pre-existing soil conditions. Using machine-learning statistical approaches, as we have here, can offer new insights that might otherwise be overlooked in models constructed based on previous understanding and assumptions.

5 | CONCLUSIONS

In this study, we compiled a novel dataset of deposit volumes and grain sizes for post-fire debris flows across 19 fires that were located across the IMW. Utilizing this dataset, we developed and presented four new predictive grain size models to capture the fine, median, and coarse ends of a deposit GSD. These models use topographic, climate, and lithological metrics to predict debris flow grain sizes, and as no wildfire metrics were found to be significant, they can easily be used to inform risk assessments (such as to stream habitat) conducted either before or after a wildfire occurs (i.e. no data or predictions of wildfire severity are required). We have also developed and presented a new post-fire debris flow volume prediction model, which uses predictor variables from publicly available datasets across the IMW region and that are all grounded in previously understood post-fire debris flow controls and process. This model can be applied after a fire using available burn severity maps to predict debris flow sediment yields from burned catchments without requiring information about the triggering storm. While we recognize that not using any precipitation metrics in our volume model departs from some previous approaches (Cannon et al., 2010; Gartner et al., 2008, 2014), the existing models are challenging to validate or apply without detailed precipitation data. Additionally, the current understanding of rainfall controls on post-fire debris flow volumes is limited and inconsistent, particularly outside of Southern California. Overall, the suite of models presented here advance our geomorphic understanding of burned landscapes, improve upon common approaches in post-fire risk assessment, and better inform watershed management in response to post-fire debris flows.

As the area burned at high severity is projected to increase across the western United States (Liu et al., 2010; Westerling et al., 2011), watersheds will become more vulnerable to post-fire erosion and debris flows (Cannon & Gartner, 2005; Moody & Martin, 2009; Ren & Leslie, 2020). Simultaneously, yearly decreases in snowpack and increasing evapotranspiration are resulting in increased water scarcity across the western United States (U.S. Environmental Protection Agency [EPA], 2016), making water storage in reservoirs even more essential (Abatzoglou & Williams, 2016; Cayan, 1996; Westerling et al., 2006). As advancements in modelling seek to predict how increasing post-fire erosion will impact downstream water resources and reservoir storage capacities, it is essential to have better and more detailed information about the location, timing, volumes, and grain sizes of debris flow inputs to river networks after wildfire. The models presented here offer insights into post-fire debris flows in the IMW and provide a critical step towards predicting post-fire debris flow grain sizes, which can help inform watershed-scale modelling frameworks (e.g. Murphy et al., 2019). To improve these models further and our ability to understand downstream impacts of post-fire debris flows, it is paramount that the geomorphic and wildfire communities expand efforts to collect more post-fire rainfall and debris flow data, especially in regions beyond Southern California.

ACKNOWLEDGEMENTS

This research was supported by funding from the National Science Foundation (NSF-EAR 1848667 to B. P. Murphy and P. Belmont), the Joint Fire Science Program (Award ID 19-2-02-6 to L. Yocom, B. P. Murphy, and P. Belmont), the Utah State University Public Lands Initiative (P. Belmont and B. P. Murphy), and the Utah Agricultural Experiment Station. The authors thank Alec Arditti and Morgan Reid for their contributions to the fieldwork, Scott David and Jon Czuba for their critical insights, Richard Giraud, Jen Pierce, and Luke McGuire for providing detailed debris flow data from their publications, and the anonymous reviewers who provided comments that helped improve this work.

AUTHOR CONTRIBUTIONS

a. Sara Wall
b. Brendan Murphy, Patrick Belmont, and Larissa Yocom
c. Sara Wall, Brendan Murphy, and Patrick Belmont
d. Sara Wall
e. Sara Wall
f. Sara Wall
g. Brendan Murphy and Patrick Belmont
h. Sara Wall
i. Brendan Murphy, Patrick Belmont, and Larissa Yocom

DATA AVAILABILITY STATEMENT

The primary data used for analysis and included in the figures can be found in the online Supporting Information.

ORCID

Sara Wall https://orcid.org/0000-0002-3673-1126
Brendan P. Murphy https://orcid.org/0000-0001-8025-1253
Patrick Belmont https://orcid.org/0000-0002-8244-3854
Larissa Yocom https://orcid.org/0000-0003-2459-0765

REFERENCES

Abatzoglou, J.T. & Williams, A.P. (2016) Impact of anthropogenic climate change on wildfire across western US forests. *Proceedings of the National Academy of Sciences*, 113(42), 11770–11775. Available from: https://doi.org/10.1073/pnas.1607171113

Ahammad, M., Czuba, J.A., Pfeiffer, A., Murphy, B.P. & Belmont, P. (2021) Simulated sediment-pulse dynamics in a gravel-bedded river, Nisqually River, Washington, USA. *Journal of Geophysical Research: Earth Surface*, 126(10), e2021JF006194. Available from: https://doi.org/10.1029/2021JF006194

Allen, P., Armitage, J., Whittaker, A., Michael, N., Roda-Boluda, D. & D'Arcy, M. (2015) Fragmentation model of the grain-size mix of sediment supplied to basins. *Journal of Geology*, 123(5), 405–427. Available from: https://doi.org/10.1086/683113

Attal, M., Mudd, S., Hurst, M., Weinman, B., Yoo, K. & Naylor, M. (2015) Impact of change in erosion rate and landscape steepness on hillslope and fluvial sediments grain size in the Feather River basin (Sierra Nevada, California). *Earth Surface Dynamics and Landforms*, 2(1), 1047–1092. Available from: https://doi.org/10.5194/esurf-3-201-2015

Betran, P. (2003) The rock-avalanche of February 1995 at Claix (French Alps). *Geomorphology*, 54(3–4), 339–346. Available from: https://doi.org/10.1016/S0169-555X(03)00041-2

Bisson, P., Rieman, B., Luce, C., Hessburg, P., Lee, D., Kersner, J., Reeves, G. & Gresswell, R. (2003) Fire and aquatic ecosystems of the western USA: Current knowledge and key questions. *Forest Ecology and Management*, 178(1–2), 213–229. Available from: https://doi.org/10.1016/S0378-1127(03)00063-X

Bladon, K., Emelkos, M., Silins, U. & Stone, M. (2014) Wildfire and the future of water supply. *Environmental Science and Technology*, 48(16), 8936–8943. Available from: https://doi.org/10.1021/es500130g

Breiman, L. (2001) Random forests. *Machine Learning*, 45, 5–32. Available from: https://doi.org/10.1023/A:1010933404324

Bridgewater, J. (1994) Mixing and segregation mechanisms in particle flow. In: Mehta, A. (Ed.) *Granular Matter*. New York: Springer, pp. 161–193.

Brown, D.K., Echelle, A.A., Propst, D.L., Brooks, J.E. & Fisher, W.L. (2001) Catastrophic wildfire and number of populations as factors influencing risk of extinction for Gila trout (*Oncorhynchus gilae*). *Western North American Naturalist*, 61, 139–148.

Brown, T.C., Hobbins, M.T. & Ramirez, J.A. (2008) Spatial distribution of water supply in the coterminous United States. *Journal of the American Water Resources Association*, 44(6), 1474–1487. Available from: https://doi.org/10.1111/j.1752-1688.2008.00252.x

Burton, T.A. (2005) Fish and stream habitat risks from uncharacteristic wildfire: Observations from 17 years of fire-related disturbances on the Boise National Forest, Idaho. *Forest Ecology and Management*, 211(1–2), 140–149. Available from: https://doi.org/10.1016/j.foreco.2005.02.063

Cannon, S., Bigio, E. & Mine, E. (2001) A process for fire-related debris flow initiation, Cerro Grande fire, New Mexico. *Hydrological Processes*, 15, 3011–3023. Available from: https://doi.org/10.1002/hyp.388

Cannon, S., Boldt, E., Laber, J., Kean, J. & Staley, D. (2011) Rainfall intensity-duration thresholds for postfire debris-flow emergency response planning. *Natural Hazards*, 59(1), 209–236. Available from: https://doi.org/10.1007/s11069-011-9747-2

Cannon, S. & Gartner, J. (2005) Wildfire-related debris flow from a hazards perspective. In: Jakob, M. & Hungr, O. (Eds.) *Debris-Flow Hazards and Related Phenomena*. Berlin: Springer, pp. 363–385. Available from: https://doi.org/10.1007/3-540-27129-5_15

Cannon, S., Gartner, J., Rupert, M., Michael, J., Rea, A. & Parrett, C. (2010) Predicting the probability and volume of postwildfire debris flows in the IMWern United States. *GSA Bulletin*, 122(1), 127–144. Available from: https://doi.org/10.1130/B26459.1

Cannon, S., Gartner, J., Wilson, R., Bowers, J. & Laber, J. (2008) Storm rainfall conditions for floods and debris flows from recently burned areas in southwestern Colorado and southern California. *Geomorphology*, 96(3–4), 250–269. Available from: https://doi.org/10.1016/j.geomorph.2007.03.019

Cannon, S.H. (2001) Debris-flow generation from recently burned watersheds. *Environmental and Engineering Geoscience*, 7(4), 321–341. Available from: https://doi.org/10.2113/gseegeosci.7.4.321

Cannon, S.H., Gartner, J.E., Parrett, C. & Parise, M. (2003) Wildfire-related debris-flow generation through episodic progressive sediment-bulking processes, western USA. In: Rickenmann, D. & Chen, C.L. (Eds.) *Debris-Flow Hazards Mitigation – Mechanics, Prediction, and Assessment. Proceedings of the Third International Conference on Debris-Flow Hazards Mitigation, Davos, Switzerland, 10–12 September 2003*. Rotterdam: A.A. Balkema, pp. 71–82.

Cayan, D.R. (1996) Interannual climate variability and snowpack in the western United States. *Journal of Climate*, 9(5), 928–948. Available from: https://doi.org/10.1175/1520-0442(1996)009<0928:ICVASI>2.0.CO;2

Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J. & Lawler, J.J. (2007) Random forests for classification in ecology. *Ecology*, 88(11), 2783–2792. Available from: https://doi.org/10.1890/07-0539.1

Czuba, J. & Foufoula-Georgiou, E. (2014) A network-based framework for identifying potential synchronizations and amplifications of sediment delivery in river basins. *Water Resources Research*, 50(5), 3826–3851. Available from: https://doi.org/10.1002/2013WR014227

Czuba, J., Foufoula-Georgiou, E., Gran, K., Belmont, P. & Wilcock, P. (2016) Interplay between spatially explicit sediment sourcing, hierarchical river-network structure, and in-channel bed material sediment transport and storage dynamics. *Journal of Geophysical Research: Earth Surface*, 122(5), 1090–1120. Available from: https://doi.org/10.1002/2016JF003965

Czuba, J.A. (2018) A Lagrangian framework for exploring complexities of mixed-size sediment transport in gravel-bedded river networks. *Geomorphology*, 321, 146–152. Available from: https://doi.org/10.1016/j.geomorph.2018.08.031

DeBano, L.F. (1981) *Water repellent soils: A state of the art*. USDA Forest Service Research Paper PSW-46.

DeBano, L.F. (2000) The role of fire and soil heating on water repellency in wildland environments: A review. *Journal of Hydrology*, 231–232, 195–206. Available from: https://doi.org/10.1016/S0022-1694(00)00194-3

DeBano, L.F., Rice, R.M. & Conrad, C.E. (1979) *Soil heating in chaparral fires: Effects on soil properties, plant nutrients, erosion, and runoff*. USDA Forest Service Research Paper PSW-145.

DeLong, S., Youberg, A., DeLong, W. & Murphy, B. (2018) Post-wildfire landscape change and erosional processes from repeat terrestrial lidar in a steep headwater catchment, Chiricahua Mountains, Arizona, USA. *Geomorphology*, 300, 13–30.

Doerr, S., Shakesby, R. & MacDonald, L. (2009) Soil water repellency – a key factor in post-fire erosion? In: Cerda, A. (Ed.) *Fire Effects on Soils and Restoration Strategies*. Boca Raton, FL: CRC Press chapter 7.

Doerr, S.H., Shakesby, R.A., Blake, W.H., Chafer, C.J., Humphreys, G.S. & Wallbrink, P.J. (2006) Effects of differing wildfire severities on soil wettability and implications for hydrological response. *Hydrology*, 319(1–4), 295–311. Available from: https://doi.org/10.1016/j.jhydrol.2005.06.038

Duane, A., Castellnou, M. & Brotons, L. (2021) Towards a comprehensive look at global drivers of novel extreme wildfire events. *Climatic Change*, 165(3), 1–21. Available from: https://doi.org/10.1007/s10584-021-03066-4

Ellett, N.G., Pierce, J.L. & Glenn, N.F. (2019) Partitioned by process: Measuring post-fire debris-flow and rill erosion with structure from motion photogrammetry. *Earth Surface Processes and Landforms*, 44(15), 3128–3146. Available from: https://doi.org/10.1002/esp.4728

Finco, M., Quayle, B., Zhang, Y., Lecker, J., Megown, K. & Brewer, K. (2012) Monitoring trends and burn severity (MTBS): Monitoring wildfire activity for the past quarter century using Landsat data. *Moving from Status to Trends: Forest Inventory and Analysis Symposium*. US Forest Service, Northern Research Station.

Fisher, A., Belmont, P., Murphy, B.P., MacDonald, L., Ferrier, K.L. & Hu, K. (2021) Natural and anthropogenic controls on sediment rating curves in northern California coastal watersheds. *Earth Surface Processes and Landforms*, 46(8), 1610–1628. Available from: https://doi.org/10.1002/esp.5137

Gannon, B., Wei, Y., MacDonald, L., Kampf, S., Jones, K., Cannon, J., Wolk, B., Cheng, A., Addington, R. & Thompson, M. (2019) Prioritizing fuels reduction for water supply protection. *International Journal of Wildland Fire*, 30, 733–744.

Garter, J. (2005) *Relations between wildfire related debris-flow volumes and basin morphology, burn severity, material, properties, and triggering storm rainfall*. Graduate thesis, University of Colorado, USA.

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, 45–46. Available from: https://doi.org/10.1016/j.enggeo.2014.04.008

Gartner, J.E., Cannon, S.H., Santi, P.M. & deWolfe, V.G. (2008) Empirical models to predict the volumes of debris flows generated by recently burned basins in the western U.S. *Geomorphology*, 96(3–4), 339–354. Available from: https://doi.org/10.1016/j.geomorph.2007.02.033

Gemuer, R., Poggi, J. & Tuleau-Malot, C. (2012) Variable selection using Random Forest. *Pattern Recognition Letters*, 31(14), 2225–2236. Available from: https://doi.org/10.1016/j.patrec.2010.03.014

Gilbert, J. & Wilcox, A. (2020) Sediment routing and floodplain exchange (SeRFE): A spatially explicit model of sediment balance and connectivity through river networks. *Journal of Advances in Modeling Earth Systems*, 12(9), 9. Available from: https://doi.org/10.1029/2020MS002048

Giraud, R. & McDonald, G. (2009) *The 2000–2004 Fire-Related Debris Flows in Northern Utah*. Salt Lake, UT: Utah Geological Survey.

Gresswell, R.E. (1999) Fire and aquatic ecosystems in forested biomes of North America. *Transactions of the American Fisheries Society*, 128(2), 193–221. Available from: https://doi.org/10.1577/15488659

Hales, T. & Roering, J. (2007) Climatic controls on frost-cracking and implications for the evolution of bedrock landscapes. *Journal of Geophysical Research*, 112(F2), 2033. Available from: https://doi.org/10.1029/2006JF000616

Hawbaker, T.J. & Zhu, Z. (2012) Projected future wildland fires and emissions for the Western United States. In: Zhiliang, Z. & Reed, B.C. (Eds.) *Baseline and Projected Future Carbon Storage and Greenhouse-Gas Fluxes in Ecosystems of the Western United States*. Reston, VA: US Geological Survey.

Helsel, D., Hirsch, R., Ryberg, K., Archfield, S. & Gilroy, E. (2020) *Statistical Methods in Water Resources*. Reston, VA: US Geological Survey.

Hill, R., Weber, M., Leibowitz, S., Olsen, A. & Thornbrugh, D. (2015) The stream-catchment (StreamCat) dataset: A database of watershed metrics for the conterminous United States. *JAWRA*, 52, 120–128.

Jager, H.I., Long, J.W., Malison, R.L., Murphy, B.P., Rust, A., Silva, L.G., Sollman, R., Steel, Z., Bowen, M., Dunham, J., Ebersole, J. & Flitcroft, R.L. (2021) Resilience of terrestrial and aquatic fauna to historical and future wildfire regimes in western North America. *Ecology and Evolution*, 11(18), 12259–12284. Available from: https://doi.org/10.1002/ece3.8026

Kampf, S., Gannon, B., Wilson, C., Saavedra, F., Miller, M., Heldmyer, A., Livneh, B., Nelson, P. & MacDonald, L. (2020) PEMIP: Post-fire erosion model inter-comparison project. *Journal of Environmental Management*, 268, 110704. Available from: https://doi.org/10.1016/j.jenvman.2020.110704

Kean, J., Staley, D. & Cannon, S. (2011) In situ measurements of post-fire debris flows in southern California: Comparisons of the timing and magnitude of 24 debris-flow events with rainfall and soil moisture conditions. *Journal of Geophysical Research*, 116(F4), 4019. Available from: https://doi.org/10.1029/2011JF002005

Kean, J., Staley, D., Lancaster, J., Rengers, F., Swanson, B., Coe, J., Hernandez, J., Sigman, A. & Lindsay, D. (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*, 4(4), 1140–1163. Available from: https://doi.org/10.1130/GES02048.1

Kohavi, R. (2001) A study of cross-validation and bootstrap for accuracy estimation and model. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI)*, vol. 2, pp. 1137–1143.

Kondolf, G.M. (2000) Assessing salmonid spawning gravel quality. *Transactions of the American Fisheries Society*, 129(1), 262–281. Available from: https://doi.org/10.1109/LCOMM.2018.2864265

Kopecky, M., Macek, M. & Wild, J. (2021) Topographic wetness index calculation guidelines based on measured soil moisture and plant species composition. *Science of the Total Environment*, 757, 143785. Available from: https://doi.org/10.1016/j.scitotenv.2020.143785

Langhans, C., Smith, H., Chong, D. & Nyman, P. (2016) A model for assessing water quality risk in catchments prone to wildfire. *Journal of Hydrology*, 534, 407–426. Available from: https://doi.org/10.1016/j.jhydrol.2015.12.048

Larsen, I. (2003) *From rim to river: The geomorphology of debris flows in the Green River Canyons of Dinosaur National Monument, Colorado and Utah*. Graduate thesis, Utah State University, USA.

Liu, Y., Stanturf, J. & Goodrick, S. (2010) Trends in global wildfire potential in a changing climate. *Forest Ecology and Management*, 259(4), 685–697. Available from: https://doi.org/10.1016/j.foreco.2009.09.002

Liu, W. & He, S. (2020) Comprehensive modelling of runoff-generated debris flow from formation to propagation in a catchment. *Landslides*, 17, 1529–1544.

Marlon, J.R., Bartlein, P.J., Gavin, D.G., Long, C.J., Anderson, R.S., Briles, C. E., Brown, K.J., Colombaroli, D., Hallett, D.J., Power, M.J. & Scharf, E. A. (2012) Long-term perspective on wildfires in the western USA. *Proceedings of the National Academy of Sciences*, 109(9), E535–E543. Available from: https://doi.org/10.1073/pnas.1112839109

Marshall, J. & Sklar, L. (2012) Mining soil databases for landscape-scale patterns in the abundance and size distribution of hillslope rock fragments. *Earth Surface Processes and Landforms*, 37(3), 287–300. Available from: https://doi.org/10.1002/esp.2241

Martin, D.A. (2016) At the nexus of fire, water and society. *Philosophical Transactions of the Royal Society B*, 371(1696), 20150172. Available from: https://doi.org/10.1098/rstb.2015.0172

Martin, J. (2000) *Debris-flow activity in canyon of Lodore, Colorado: Implications for debris-fan formation and evolution*. Graduate thesis, Utah State University, USA.

Mataix-Solera, J., Cerdà, A., Arcenegui, V., Jordán, A. & Zavala, L. (2011) Fire effects on soil aggregation: A review. *Earth-Science Reviews*, 109(1–2), 44–60. Available from: https://doi.org/10.1016/j.earscirev.2011.08.002

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

McGuire, L. & Youberg, A. (2020) What drives spatial variability in rainfall intensity-duration thresholds for post-wildfire debris flows? Insights from the 2018 buzzard fire, NM, USA. *Landslides*, 17(10), 2385–2399. Available from: https://doi.org/10.1007/s10346-020-01470-y

McGuire, L.A., Rengers, F.K., Oakley, N., Kean, J.W., Staley, D.M., Tang, H., de Orla-Barile, M. & Youberg, A.M. (2021) Time since burning and rainfall characteristics impact post-fire debris-flow initiation and magnitude. *Environmental & Engineering Geoscience*, 27(1), 43–56. Available from: https://doi.org/10.2113/EEG-D-20-00029

Messenzehl, K., Viles, H., Otto, J., Ewald, A. & Dikau, R. (2017) Linking weathering, rockwall instability and rockfall supply on talus slopes in glaciated hanging valleys (Swiss Alps). *Permafrost and Periglacial Processes*, 29(3), 135–151. Available from: https://doi.org/10.1002/ppp.1976

Meyer, G. & Wells, S. (1997) Fire-related sedimentation events on alluvial fans, Yellowstone National Park, USA. *Journal of Sediment Research*, 67, 776–791.

Moody, J. & Martin, D. (2004) Wildfire impacts on reservoir sedimentation in the western United States. In *Proceedings of the Ninth International Symposium on River Sedimentation*.

Moody, J.A. & Martin, D.A. (2009) Synthesis of sediment yields after wildland fire in different rainfall regimes in the western United States. *International Journal of Wildland Fire*, 18(1), 96–115. Available from: https://doi.org/10.1071/WF07162

Mote, P.W., Hamlet, A.F., Clark, P.M. & Lettenmaier, D.P. (2005) Declining mountain snowpack in Western North America. *Bulletin of the American Meteorological Society*, 86(1), 39–49. Available from: https://doi.org/10.1175/BAMS-86-1-39

Murphy, B., Yocom, L. & Belmont, P. (2018) Beyond the 1984 perspective: Narrow focus on modern wildfire trends underestimate future risks to water security. *AGU Earth's Future*, 6(11), 1496–1497. Available from: https://doi.org/10.1029/2018EF001006

Murphy, B.P., Czuba, J.A. & Belmont, P. (2019) Post-wildfire sediment cascades: A modeling framework linking debris flow generation and network-scale sediment routing. *Earth Surfaces Processes and Landforms*, 44(11), 2095–2315. Available from: https://doi.org/10.1002/esp.4635

Murphy, B.P., Johnson, J.P., Gasparini, N.M. & Sklar, L.S. (2016) Chemical weathering as a mechanism for the climatic control of bedrock river incision. *Nature*, 532(7598), 223–227. Available from: https://doi.org/10.1038/nature17449

Murphy, B.P., Walsworth, T.E., Belmont, P., Conner, M.M. & Budy, P. (2020) Dynamic habitat disturbance and ecological resilience (DyHDER): Modeling population responses to habitat condition. *Ecosphere*, 11(1), e03023. Available from: https://doi.org/10.1002/ecs2.3023

Naylor, M. (1980) The origin of inverse grading in muddy debris flow deposits—a review. *Journal of Sedimentary Research*, 50, 1111–1116.

Nyman, P., Box, W., Stout, J., Sheridan, G., Keesstra, S., Lane, P. & Langhans, C. (2020) Debris-flow-dominated sediment transport through a channel network after wildfire. *Earth Surface Processes and Landforms*, 45(5), 1155–1167. Available from: https://doi.org/10.1002/esp.4785

Nyman, P., Smith, H., Sherwin, C., Langhans, C., Lane, P. & Sheridan, G. (2015) Predicting sediment delivery from debris flows after wildfire. *Geomorphology*, 250, 173–186. Available from: https://doi.org/10.1016/j.geomorph.2015.08.023

Pfeiffer, A.M., Barnhart, K.R. & Czuba, J.A. (2020) NetworkSedimentTransporter: A Landlab component for bed material transport through river networks. *Journal of Open Source Software*, 5(53), 2341. Available from: https://doi.org/10.21105/joss.02341

Pietraszek, J. (2006) *Controls on post-fire debris flow erosion at the hillslope scale, Colorado Front Range*. MS thesis, Colorado State University, USA.

Prancevic, J. & Lamb, M. (2015) Particle friction angles in steep mountain channels. *JGR Earth Surface*, 120(2), 242–259. Available from: https://doi.org/10.1002/2014JF003286

Prancevic, J., Lamb, P. & Fuller, B. (2014) Incipient sediment motion across the river to debris-flow transition. *Geology*, 42(3), 191–194. Available from: https://doi.org/10.1130/G34927.1

Propst, D.L. & Stefferud, J.A. (1997) Population dynamics of Gila trout in the Gila River drainage of the South-Western United States. *Journal of Fish Biology*, 51(6), 1137–1154. Available from: https://doi.org/10.1006/jfbi.1997.0512

Reeves, G.H., Benda, L.E., Burnett, K.M., Bisson, P. & Sedell, J.R. (1995) A disturbance-based ecosystem approach to maintaining and restoring freshwater habitats of evolutionarily significant units of anadromous salmonids in the Pacific Northwest. *American Fisheries Society Symposium*, 17, 334–349.

Ren, D. & Leslie, L. (2020) Climate warming enhancement of catastrophic southern Californian debris flows. *Scientific Reports*, 10(1), 10507. Available from: https://doi.org/10.1038/s41598-020-67511-7

Riebe, C., Sklar, L., Lukens, C. & Shuster, D. (2015) Climate and topography control the size and flux of sediment produced on steep mountain slopes. *PNAS*, 112(51), 15574–15579. Available from: https://doi.org/10.1073/pnas.1503567112

Robichaud, P. (2005) Measurement of post-fire hillslope erosion to evaluate and model rehabilitation treatment effectiveness and recovery. *International Journal of Wildland Fire*, 14(4), 475–485. Available from: https://doi.org/10.1071/WF05031

Robichaud, P., Elliot, W., Lewis, S. & Miller, E. (2016) Validation of a probabilistic post-fire erosion model. *International Journal of Wildland Fire*, 25(3), 337–350. Available from: https://doi.org/10.1071/WF14171

Robinne, F.N., Miller, C., Parisien, M.A., Emelko, M.B., Bladon, K.D., Silins, U. & Flannigan, M. (2016) A global index for mapping the exposure of water resources to wildfire. *Forests*, 7(1), 22. Available from: https://doi.org/10.3390/f7010022

Roda-Boluda, D., D'Arcy, M., McDonald, J. & Whittaker, A. (2018) Lithological controls on hillslope sediment supply: Insights from landslide activity and grain size distributions. *Earth Surface Processes and Landforms*, 43(5), 956–977. Available from: https://doi.org/10.1002/esp.4281

Roghair, C.N., Dolloff, C.A. & Underwood, M.K. (2002) Response of a brook trout population and instream habitat to a catastrophic flood and debris flow. *Transactions of the American Fisheries Society*, 131(4), 718–730. Available from: https://doi.org/10.1577/1548-8659(2002)131<0718:ROABTP>2.0.CO;2

Rupert, M.G., Cannon, S.H., Gartner, J.E., Michael, J.A. & Helsel, D.R. (2008) *Using Logistic Regression to Predict the Probability of Debris Flows in Areas Burned by Wildfires, Southern California, 2003–2006*. Open-File Report 2008–1370. US Geological Survey: Reston, VA.

Saley, T., Akbar, H., Hager, R., Wilkins, E.J., Elkin, C., Belmont, P. & Flint, C. G. (2022) Climate change at Utah ski resorts: Impacts, perceptions, and adaptation strategies. *Mountain Research & Development*. PMID: In review.

Sankey, J.B., Kreitler, J., Hawbaker, T.J., McVay, J.L., Miller, M.E., Mueller, E.R., Vaillant, N.M., Lowe, S.E. & Sankey, T.T. (2017) Climate, wildfire, and erosion ensemble foretells more sediment in western USA watersheds. *Geophysical Research Letters*, 44(17), 8884–8892. Available from: https://doi.org/10.1002/2017GL073979

Santi, P. (2014) Precision and accuracy in debris-flow volume measurements. *Environmental Engineering and Geoscience*, 20(4), 349–359. Available from: https://doi.org/10.2113/gseegeosci.20.4.349

Santi, P., deWolfe, V., Higgins, J., Cannon, S. & Gartner, J. (2008) Sources of debris flow material in burned areas. *Geomorphology*, 96(3–4), 310–321. Available from: https://doi.org/10.1016/j.geomorph.2007.02.022

Santi, P.M. & Morandi, L. (2013) Comparison of debris-flow volumes from burned and unburned areas. *Landslides*, 10(6), 757–769. Available from: https://doi.org/10.1007/s10346-012-0354-4

Sedell, E.R., Gresswell, R.E. & McMahon, T.E. (2015) Predicting spatial distribution of post-fire debris flows and potential consequences for native trout in headwater streams. *Freshwater Science*, 34(4), 1558–1570. Available from: https://doi.org/10.1086/684094

Shakesby, R., Moody, J., Martin, D. & Robichaud, P. (2016) Synthesizing empirical results to improve predictions of post-wildfire runoff and erosion response. *International Journal of Wildland Fire*, 25(3), 257–261. Available from: https://doi.org/10.1071/WF16021

Sklar, L., Riebe, C., Marshall, J., Genetti, J., Leclere, S., Lukens, C. & Merces, V. (2016) The problem of predicting the size distribution of sediment supplied by hillslopes to rivers. *Geomorphology*, 277, 31–49. Available from: https://doi.org/10.1016/j.geomorph.2016.05.005

Sklar, L.S., Riebe, C.S., Marshall, J.A., Genetti, J., Leclere, S., Lukens, C.L. & Merces, V. (2017) The problem of predicting the size distribution of sediment supplied by hillslopes to rivers. *Geomorphology*, 277, 31–49. Available from: https://doi.org/10.1016/j.geomorph.2016.05.005

Smith, H.G., Sheridan, G.J., Lane, P.N., Nyman, P. & Haydon, S. (2011) Wildfire effects on water quality in forest catchments: A review with implications for water supply. *Journal of Hydrology*, 396(1–2), 170–192. Available from: https://doi.org/10.1016/j.jhydrol.2010.10.043

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

Staley, D., Kean, J. & Rengers, F. (2020) The RI of post-fire debris-flow generating rainfall in the southwestern United States. *Geomorphology*, 370, 107392. Available from: https://doi.org/10.1016/j.geomorph.2020.107392

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. Available from: https://doi.org/10.1016/j.geomorph.2016.10.019

Strom, A. (2006) Morphology and internal structure of rockslides and rock avalanches: Grounds and constraints for their modeling. In: Evans, S. G., Mugnozza, G.S., Strom, A. & Hermanns, R.L. (Eds.) *Landslides from Massive Rock Slope Failure*. Dordrecht: Springer, pp. 305–326.

Tang, H., McGuire, L.A., Rengers, F.K., Kean, J.W., Staley, D.M. & Smith, J. B. (2019) Evolution of debris-flow initiation mechanisms and sediment sources during a sequence of postwildfire rainstorms. *Journal of Geophysical Research: Earth Surface*, 124(6), 1572–1595. Available from: https://doi.org/10.1029/2018JF004837

U.S. Environmental Protection Agency (EPA). (2016) Climate Change Indicators in the United States. https://www.epa.gov/climate-indicators/downloads-indicators-report

Vaughan, A.A., Belmont, P., Hawkins, C.P. & Wilcock, P. (2017) Near-channel versus watershed controls on sediment rating curves. *Journal of Geophysical Research: Earth Surface*, 122(10), 1901–1923. Available from: https://doi.org/10.1002/2016JF004180

Wang, Y., Cheng, Q. & Zhu, Q. (2012) Inverse grading analysis of deposit from rock avalanches triggered by Wenchuan earthquake. *Chinese Journal of Rock Mechanics and Engineering*, 31, 1089–1106.

Westerling, A.L., Hidalgo, H.G., Cayan, D.R. & Swetnam, T.W. (2006) Warming and earlier spring increase western US forest wildfire activity. *Science*, 313(5789), 940–943. Available from: https://doi.org/10.1126/science.1128834

Westerling, A.L., Turner, M.G., Smithwick, E.A., Romme, W.H. & Ryan, M. G. (2011) Continued warming could transform greater Yellowstone fire regimes by mid-21st century. *Proceedings of the National Academy of Sciences*, 108(32), 13165–13170. Available from: https://doi.org/10.1073/pnas.1110199108

Western Water Assessment. (2021) *IMW Climate*. https://wwa.colorado.edu/climate/index.html

Whittaker, A., Attal, M. & Allen, P. (2010) Characterizing the origin, nature and fate of sediment exported from catchments perturbed by active tectonics. *Basin Research*, 22, 809–828.

Wilcock, P.R. & Crowe, J.C. (2003) Surface-based transport model for mixed-size sediment. *Journal of Hydraulic Engineering*, 129(2), 120–128.

Wilkins, E.J., Akbar, H., Saley, T.C., Hager, R., Elkin, C.M., Belmont, P., Flint, C.G. & Smith, J.W. (2021) Climate change and Utah ski resorts: Impacts, perceptions, and adaptation strategies. *Mountain Research and Development*, 41(3), R12. Available from: https://doi.org/10.1659/MRD-JOURNAL-D-20-00065.1

Wolman, M.G. (1954) A method of sampling coarse river-bed material. *Eos, Transactions of the American Geophysical Union*, 35(6), 951–956. Available from: https://doi.org/10.1029/TR035i006p00951

Wondzell, S. & King, J. (2003) Postfire erosional processes in the Pacific Northwest and Rocky Mountain regions. *Forest Ecology and Management*, 178(1–2), 75–87. Available from: https://doi.org/10.1016/S0378-1127(03)00054-9

Zou, Z., Tang, H., Xiong, C., Su, A. & Criss, R. (2017) Kinetic characteristics of debris flows as exemplified by field investigations and discrete element simulation of the catastrophic Jiweishan rockslide, China. *Geomorphology*, 295, 1–15. Available from: https://doi.org/10.1016/j.geomorph.2017.06.012

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Wall, S., Murphy, B.P., Belmont, P. & Yocom, L. (2023) Predicting post-fire debris flow grain sizes and depositional volumes in the Intermountain West, United States. *Earth Surface Processes and Landforms*, 48(1), 179–197. Available from: https://doi.org/10.1002/esp.5480