Temporal changes in rainfall intensity-duration thresholds for post-wildfire flash floods ~~and sensitivity to spatiotemporal distributions of rainfall~~in Southern California

Tao Liu[1,2], Luke A. McGuire[1], Nina Oakley[3], Forest Cannon[3]
[1]Department of Geosciences, University of Arizona, Tucson, AZ 85721-0011, USA
[2]Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ 85721-0011, USA
[3]Center for Western Weather and Water Extremes, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA, USA

Correspondence to: Tao Liu (liutao@arizona.edu)

Abstract. Rainfall intensity-duration (ID) thresholds are commonly used to assess flash flood potential downstream of burned watersheds. High-intensity and/or long-duration rainfall is required to generate flash floods as landscapes recover from fire, but there is little guidance on how thresholds change as a function of time since burning. Here, we force a hydrologic model with radar-derived precipitation to estimate ID thresholds for post-fire flash floods in a 41.5 km^2 watershed in southern California, USA. Prior work in this study area constrains temporal changes in hydrologic model parameters, allowing us to estimate temporal changes in ID thresholds. Results indicate that ID thresholds increase by more than a factor of 2 from post-fire year 1 to post-fire year 5. Thresholds based on averaging rainfall intensity over durations of 15~~30~~-60 minutes perform better than those that average rainfall intensity over shorter time intervals. Moreover, thresholds based on the 75th percentile of radar-derived rainfall intensity over the watershed perform better than thresholds based on the 25th or 50th percentile of rainfall intensity. Results demonstrate how hydrologic models can be used to estimate changes in ID thresholds following disturbance and provide guidance on the rainfall metrics that are best suited for predicting post-fire flash floods.

1 Introduction

Heightened hydrologic responses are common within and downstream of recently burned areas, resulting in an increased likelihood of flash floods. Rainfall intensity-duration (ID) thresholds are commonly used to assess the potential for flash floods (Moody and Martin, 2001; Cannon et al., 2008). Many past studies aimed at defining thresholds for flash floods focus on the first 1-2 years following fire (Cannon et al., 2008; Wilson et al., 2018). Since the hydrologic impacts of fire are transient, rainfall ID thresholds associated with flash floods are likely to change as a watershed recovers (Ebel and Martin, 2017; Ebel and Moody, 2017; Moreno et al., 2019; Ebel, 2020). It may take more than a decade for hydrologic responses to return to pre-fire levels, yet there is limited guidance on how the magnitude and utility of rainfall ID thresholds change with time since burning. Given the increased frequency and size of fire in many geographic and ecological zones (e.g. Gillett et al., 2004; Westerling et al., 2006; Kitzberger et al., 2017), it is of growing importance to quantify the best metrics for assessing flash-flood potential in the immediate aftermath of fire as well as how these metrics change throughout the recovery process (e.g. Ebel, 2020).

Rainfall ID thresholds for flash floods are typically defined using historic data that relates rainfall over different intensities and durations to an observed hydrologic response, namely the presence or absence of flooding (e.g. Cannon et al., 2008). Due to the stochastic nature of rainfall over burned areas and limited observations throughout the recovery process, there is a paucity of data that can be used to derive empirical thresholds for flash flooding beyond one year of recovery. Hazards associated with flash flooding, however, may exist downstream of burned areas well beyond one year of recovery. Wildfire alters rainfall-runoff partitioning and flood routing by incinerating vegetation and reducing interception capacity (Stoof et al., 2012, Saksa et al., 2020), decreasing hydraulic roughness, and reducing soil infiltration capacity (Larsen et al., 2009, Ebel and Moody, 2013). Reductions in infiltration capacity are often attributed to fire-induced soil water repellency (Ebel and Moody, 2013), which is generally strongest immediately following a fire and then decays over time scales ranging from one year to more than five years (Dyrness, 1976; Huffman et al., 2001; Larsen et al., 2009), though surface soil sealing (Larsen et al., 2009) and hyper-dry conditions (Moody and Ebel, 2012) are also known to play important roles. Vegetation recovery, which may influence temporal changes in hydraulic roughness and canopy interception, can take five years or longer. Cannon et al. (2008) collected sufficient data over a two-year time period following fire in southern California, USA, to define separate rainfall ID thresholds for post-fire debris flows and flash floods in the first- and second-years following fire. They found that the ID thresholds for flash floods and debris flows may increase by as much as 25 mm/h after one year of recovery, a change that they attributed to a combination of vegetation growth and sediment removal as a result of rainstorms during the first post-fire year.

Rainfall ID thresholds are often defined over a range of durations, though averaging rainfall intensity over a particular duration may provide a more reliable threshold. Post-fire hydrological response in the first few years is often best related to rainfall intensity over short durations (less than 60 min) (Staley et al., 2017; Moody and Martin, 2001). In their efforts to define rainfall ID thresholds for post-fire debris flows, Staley et al. (2013) showed that averaging rainfall intensities over durations between 15 minutes and 60 minutes resulted in thresholds that performed better relative to those associated with longer durations. One potential explanation for this observation is that post-fire debris flows are often triggered by runoff in steep, low-order drainages, which both Kean et al. (2011) and Raymond et al. (2020) have found to be highly correlated with rainfall intensities averaged over similarly short time intervals (10-15 minutes). Moody and Martin (2001) have also documented a substantial increase in peak discharge following wildfire once the 30-minute rainfall intensity (I_{30}) crossed a threshold value, suggesting that I_{30} may be a consistent predictor of flash flood activity in recently burned watersheds. Moody and Martin (2001) suggest that peak I_{30} can be used to set the threshold for early-warning flood systems. The optimal duration for defining post-fire flash floods thresholds, as well as how it may change with time, remains relatively unexplored.

Rain gage records are typically used to derive rainfall ID thresholds for flash flood and post-fires debris flows (Staley et al., 2013; Staley et al., 2017). Post-fire debris flows, however, tend to initiate in small (<1 km^2), steep watersheds. In these small watersheds, the rainfall intensity responsible for initiating a debris flow can be characterized by a single rain gage installed near the initiation zone. Flash floods differ in that they tend to occur at larger spatial scales where rainfall is spatially variable and may not be adequately characterized by data from a single rain gage. Radar-derived precipitation estimates, which can provide high spatiotemporal resolution of rainfall intensity, present opportunities to develop basin-specific thresholds for post-fire flash floods. However, high spatiotemporal variability in rainfall intensity also brings new challenges when employing radar-derived precipitation in flood warning practice. In particular, what is the best way to summarize spatially and temporally variable rainfall intensity information with a single metric that can be used as a threshold? How does hydrological recovery following fire influence the generation of flash floods and the metrics that are best suited for their prediction? Data-driven approaches to answering these and related questions may be hampered by limited monitoring of post-fire hydrologic response throughout the recovery period and the stochastic occurrence of rainfall over burned areas, which limits opportunities for observations. Given a well-constrained hydrologic model that accounts for changes associated with post-fire recovery, it is possible to use numerical experiments to understand relationships between time since burning, the spatiotemporal patterns of rainfall over a watershed, and the occurrence of flash floods.

Here, we use ~~realistic~~ observed patterns of spatially and temporally varying radar-derived rainfall estimates over a 41.5 km^2 watershed in the San Gabriel Mountains of southern California, USA, to (1) determine the optimal method to define a rainfall ID threshold for flash floods, and (2) identify changes in rainfall ID thresholds for flash floods as a function of time since burning. The watershed, which we refer to as the upper Arroyo Seco, burned during the 2009 Station Fire (USDA Forest Service, 2009). Liu et al. (2021) used rain and stream gage data collected at different times following the fire to calibrate the KINEROS2 hydrologic model for this watershed, enabling them to quantify temporal changes in model parameters as a function of time since burning. Combining this calibrated model with spatially explicit, radar-derived estimates of rainfall intensity during 34 rainstorms, we explore the utility of different rainfall ID metrics as flash flood thresholds and quantify temporal changes in those thresholds through the first five years of recovery. Results provide insight into the magnitude of temporal changes in flash flood thresholds in the densely populated, fire-prone region of southern California. ~~These F~~findings also ~~inform~~ ~~the trend of~~ provide guidance for the magnitude of change expected in rainfall ID thresholds for flash floods ~~following a fire~~during the post-fire recovery period in ~~a~~ chaparral-dominated environments similar to southern California. More generally, results support the development of early warning systems for flash floods by identifying specific metrics that can be computed using spatially variable rainfall intensity estimates to assess the potential for flash flooding. The optimal rainfall ID metrics ~~as flash thresholds~~ identified in this study could be helpful when issuing flash flood warnings based on radar-derived precipitation estimates or data from several real-time rain gages within a watershed.

2 Study Area

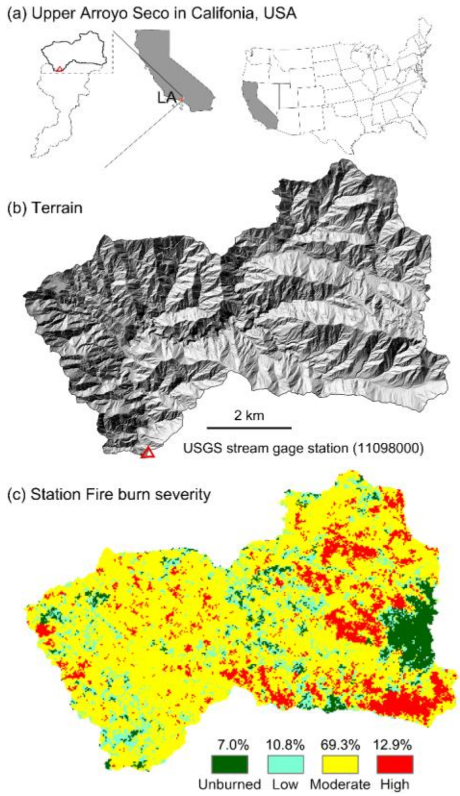


Figure 1: Modified from figure 1 in Liu et al. (2021) (a) The location of the upper Arroyo Seco watershed within California. The red triangle indicates the location of the USGS stream gage (11098000); (b) Shaded relief showing the study watershed with the USGS stream gage (red triangle; 34°13’20”, -118°10’36”); (c) Soil burn severity for the 2009 Station fire. Burn severity percentages are for planform area within each category.

The upper Arroyo Seco watershed drains the 41.5 km^2 area above USGS stream gage station (11098000) near Pasadena in the San Gabriel Mountains (Figure~~.~~ 1). The upper Arroyo Seco was burned in the August-October 2009 Station Fire, which burned more than 80% of the watershed at moderate to high soil burn severity (USDA Forest Service, 2009). Dominant shrubs and chaparral, such as chamise (*Adenostoma fasciculatum*) and manzanita (*Arctostaphylos spp.*), were completely consumed with severe soil heating in isolated patches throughout many areas burned at moderate to high severity (USDA Forest Service, 2009). Soils in this area are typically sand and silty-sand textured and thin (<1 m) with partial exposure of bedrock (Staley et al., 2014). The majority of rainfall in the study area typically occurs in the cool season, between December and March, while warm, dry conditions dominate from April to early November. The San Gabriel Mountains also experience some of the most frequent short-duration, high-intensity rainfall in the state (Oakley et al., 2018a).

Due to wildfire-induced changes in surface conditions, including canopy cover and soil-hydraulic properties, runoff generation in the first year following the fire was likely dominated by infiltration excess overland flow (Schmidt et al., 2011, Liu et al., 2021). Enhanced soil water repellency (SWR), which helps promote low infiltration capacity, and extensive dry ravel, which loads channels with fine-grained hillslope sediment, are both commonly observed after fires in the San Gabriel Mountains (e.g., Watson and Letey, 1970; Hubbert and Oriol, 2005; Lamb et al., 2011; Hubbert et al., 2012). Rengers et al. (2019) calibrated a hydrologic model using data from small watersheds (0.01-2 km^2) burned by the Station Fire and found relatively low values for saturated hydraulic conductivity (K_s), generally between 2-10 mm/h. These results are consistent with values for saturated hydraulic conductivity inferred by Liu et al. (2021) via model calibration in the upper Arroyo Seco watershed. The impact of dry ravel, which reduces grain roughness in the channel network, and reduced vegetation density led to estimates of Manning's n in the channels of the upper Arroyo Seco of approximately 0.09 s $m^{-1/3}$ in the first year following fire (Liu et al., 2021). These hydrologic changes led to widespread flooding and debris flows during multiple rainstorms in the first winter after the fire (Kean et al., 2011; Oakley et al., 2017). As hydrologic recovery began over the next several years, the watershed-scale K_s and Manning's n generally increased and likely started to mitigate the flash flood risk (Liu et al., 2021).

3 Data and Methods

3.1 Radar-derived precipitation

Weather radar coverage is adequate for estimating rainfall over the study area (NOAA 2021), and radars have been operational since the mid-1990s. This allows us to utilize observed data to capture temporal and spatial characteristics of storms impacting the study area, a region of complex terrain. We sought to identify storms in the study area that produced moderate-to-high intensity rainfall to use as inputs to a hydrologic model to simulate flood responses. Storm events were selected within the period for which observations are archived for the two operational NWS Next-Generation Weather Radar installations (NEXRAD; NOAA 1991) that cover the study area, KSOX, (Santa Ana), and KVTX (Ventura). Archives for the radars begin in 1997 and 1995, respectively.

We compiled storm events starting with those known to have produced high intensity rainfall and a debris flow response in the San Gabriel Mountains (e.g., Table 1 in Oakley et al., 2017) as well as other storms that produced high-intensity rainfall in the region (e.g., Oakley et al., 2018b, Cannon et al., 2018). We then used hourly rainfall observations from the Clear Creek (2002-present), San Rafael Hills (2005-present), and Heninger Flats (2010-present) Remote Automated Weather Stations

(RAWS, acquired from raws.dri.edu) as indicator gages for the study area. This further limited us to post-2002 events outside of the literature. All gages are <10 km from the watershed of interest; there were no long-record gages within the watershed. We used 15 mm/h as a threshold for moderate to high intensity rainfall and extracted all events from the ga~~u~~ge record meeting or exceeding this value to develop a list of events of interest. This threshold generally corresponds with a 1-year average recurrence interval storm event in the study area (NOAA Atlas 14). This value falls between the California-Nevada River Forecast Center's flash flood guidance for unburned areas in the region (~22-25 mm/h; CNRFC 2021) and regional thresholds for post-wildfire debris flows in this region at a point (12.7 mm/h, Cannon et al. 2008; Staley et al. 2013). This threshold allows us to focus on storms that have a high potential to generate floods, while keeping the number of storms to a manageable level for data processing. We reviewed the radar data for these events at which point some of the selected events could not be utilized due to radar outages or poor data quality. This exercise presented us with 34 storm events (Table S1).

Various atmospheric processes may contribute to generation of moderate-to-high rainfall intensities (e.g., Oakley et al., 2017), resulting in differing spatial and temporal precipitation patterns over a burn area. To ensure the events selected captured variability in spatial and temporal precipitation characteristics, we evaluated the spatial characteristics of the events. We found rainfall patterns could generally be categorized into four main spatial patterns at the scale of several tens of kilometers: (1) a broad pattern, a contiguous area of moderate-to-high intensity precipitation (>45 dBZ) spanning tens of kilometers; (2) a scattered pattern with numerous cells of moderate to high precipitation that are not spatially continuous; (3) an isolated pattern, with one to a few isolated cells of moderate-to-high intensity rainfall separated by non-precipitating areas several to tens of kilometers in extent; (4) a narrow cold frontal rainband (NCFR)—a north-south oriented narrow band (~3-5 km wide, tens to 100 km in length) of very high intensity rainfall (e.g., Oakley et al., 2018b; Cannon et al., 2020; Figure S1 in Supplement). At the <10 km horizontal scale (the scale of the watershed), it was harder to identify meaningful patterns and distinctions, though the larger scale signals imply varying spatial and temporal patterns of precipitation as each pass over the watershed. A table of storm events and their characteristics is available in Table S1 in the Supplement.

An approximate start and end time were determined for each event using the Clear Creek RAWS gauge as an indicator. Start time was determined by identifying the time of maximum 1h rainfall in the event and going back in time to the first of three consecutive hours of >1.5 mm/h precipitation. The end of an event was determined as the last hour where precipitation dropped below 3 mm/h for at least two consecutive hours.

Level-II base reflectivity (~~https://www.ncdc.noaa.gov/wct/)~~ https://www.ncdc.noaa.gov/wct/) between the start and end time of each event was downloaded from both the KSOX and KVTX radars. The data were used to generate spatially-distributed precipitation over the study area. Radar imagery concurrent with the gauge-based record of high intensity rainfall events was converted to a composite maximum reflectivity product at 250 m spatial and 5-minute temporal resolution. Conversion of radar reflectivity to rain rate required the application of an empirically derived reflectivity (Z) to rain rate (R) relationship (e.g.

Marshall and Palmer 1948). The Z-R relationship is conventionally represented by the equation $Z = aR^b$, which includes parameters a and b to account for variations in precipitation for a given reflectivity arising from differences in the drop size distribution. Due to the lack of previous studies investigating Z-R relationships in precipitating conditions over the region of interest, there are no standard a and b parameters to apply to the reflectivity data analyzed here. Thus, five well-known and previously published Z-R relationships were applied to the gridded reflectivity values. Supplement S3 lists the different Z-R relationships applied here and the general conditions for which they are suitable. Although the Z-R relationships used here are not based on observations from the present study's region of interest, the variation of a and b parameters yields an estimate of precipitation uncertainty. It is worth noting that a number of additional sources of radar measurement uncertainty exist that are not evaluated in depth here, including beam broadening, topographic blocking and scan elevation. However, this was not of primary concern since the goal of this study was to generate realistic spatial and temporal patterns of rainfall over the watershed with varying intensity that could be used to force the KINEROS2 hydrologic model. The goal was not to reproduce the observed hydrologic response resulting from a particular set of rainstorms.

As a range of precipitation intensities for each storm result from the application of the five different Z-R relationships (e.g., Figure S2 in Supplement), we utilize these as plausible ~~realistic~~ storms of varying precipitation intensity to increase our storm sample size, such that we apply 34 storms * 5 Z-R relations = 170 precipitation scenarios as inputs to KINEROS2. These 170 scenarios were then processed for ingestion into KINEROS2 (Figure~~.~~ 2).

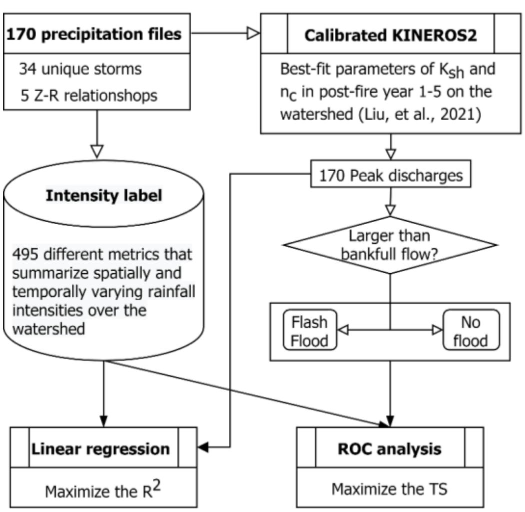


Figure 2: Delineation of rainfall intensity-duration threshold for post-fire flash flood

3.2 Summary metrics for spatially and temporally varying rainfall

In search of a spatiotemporel summary metric that may serve as a reliable flash flood threshold, we begin by describing a methodology to summarize spatially and temporally varying rainfall over a watershed. For a given rainstorm, the rainfall intensity time series at a single point, such as a single radar pixel, can be summarized by computing a moving average of intensity over a specified duration, *D*. Letting *t* denote time and *R* denote the cumulative rainfall (mm), we define the rainfall intensity over a duration *D* at any given pixel within the watershed as

$$I_D(t) = \frac{R(t) - R(t-D)}{D} \quad (1)$$

Here, we compute $I_D(t)$ for each pixel for durations of 5, 10, 15, 30, and 60 minutes. Since the intensity in each radar pixel could have a unique value, we also need a way to summarize $I_D(t)$ in space. One option would be to take the median of $I_D(t)$ to determine a typical value of I_D within the watershed at each time, *t*. However, the median may not be a good predictor of flash flooding since one could envision a scenario where it is only raining over 1/3 of the watershed, yet it is raining with sufficient intensity to generate a flash flood. We therefore compute the j^{th} percentile of $I_D(t)$ at each time, *t*, for j between 1 and 99. We denote the j^{th} percentile of $I_D(t)$ as $I_D^j(t)$. For each rainstorm, we focus our analysis on the peak value of $I_D^j(t)$ which we denote as I_D^j. As an example, I_{30}^{50} would be computed by defining I_{30} for all radar time steps within a rainstorm, determining the median value of I_{30} over the watershed at each of those time steps, and then taking the maximum of that time series of median I_{30} intensities. This analysis yields 495 different metrics (I_D^j for *j*=1,2,…,99 and *D*=5,10,15,30,60) that summarize spatially and temporally varying rainfall intensities over the watershed. In the following sections, we describe how we test the utility of each of these 495 different metrics as a flash flood threshold. A threshold defined by I_D^j would denote a threshold where 100(1-j)% of the watershed experiences rainfall of duration *D* with an intensity of *I* or greater.

3.3 Hydrological modeling

We used the KINEROS2 (K2) hydrological model to simulate the rainfall partitioning, overland flow generation, and flood routing in the upper Arroyo Seco watershed. K2 is an event-scale, distributed-parameter, process-based watershed model, which has been used extensively for rainfall-runoff processes in semi-arid and arid watersheds (Smith et al., 1995; Goodrich et al., 2012). Liu et al. (2021) used rain gage data in combination with the USGS stream gage installed at the outlet of the upper Arroyo Seco watershed to calibrate K2 during different stages of the post-fire recovery process. We use the same model setup for simulations in this study. In particular, the 41.5 km^2 watershed was discretized into 1289 hillslope planes and these planes were connected by a stream network of 519 channel segments based on a one-meter LiDAR-derived digital elevation model (DEM). After accounting for a fixed interception depth of 2.97 mm based on land cover look-up table in the Automated

Geospatial Watershed Assessment toolkit (AGWA; Miller et al., 2007), infiltration of rainfall into soil is represented using the Parlange et al. (1982) approximation. Overland flow and channel flow are modeled by kinematic wave equations. Both saturated hydraulic conductivity on hillslopes (K_{sh}) and hydraulic roughness in channels (n_c) primarily determine runoff generation and the shape of hydrograph, including total runoff volume, peak discharge rate, time to peak (Canfield et al., 2005; Yatheendradas et al., 2008; Menberu et al., 2019). Other parameters, such as hydraulic roughness (n_h) and capillary drive (G_h) on hillslopes, had a relatively minor impact on modelled runoff after the Station Fire in the upper Arroyo Seco watershed (Liu et al., 2021).

Table 1. Summary of model parameters for post-fire year 1, 2, 3, and 5. The saturated hydraulic conductivity on hillslopes (K_{sh}) and hydraulic roughness in channels (n_c) are the average of values calibrated in post-fire years 1, 2, 3, and 5 (Liu et al., 2021)

Post-fire Year	Calibration Events	K_{sh} (mm/hr)	n_c (s/[m$^{1/3}$])
1	12 Dec 2009 17 Jan 2010 5 Feb 2010	7.2	0.087
2	17 Dec 2010 20 Mar 2011	13.8	0.275
3	17 Mar 2012 13 Apr 2012	18.5	0.320
5	28 Feb 2014	23.8	0.280

Liu et al. (2021) found that both K_{sh} and n_c were lowest immediately after the fire. K_{sh} increased, on average, by approximately 4 mm/h/yr during the first five years of recovery, whereas n_c increased by more than a factor of two after 1 year of recovery and then remained relatively constant. We focus here on simulating the response to rainfall in the first five years following the fire where the watershed is likely most vulnerable to extreme responses. To represent the temporal changes in K_{sh} and n documented by Liu et al. (2021) following the fire, we used different values of K_{sh} and n_c for each post-fire year (i.e. post-fire years 1, 2, 3, and 5) based on the values calibrated by Liu et al. (2021) in post-fire years 1, 2, 3, and 5 (Table. 1). Liu et al. (2021) were unable to calibrate the necessary K2 parameters in post-fire year 4 so we do not perform any simulations to constrain flash flood thresholds in that year. Initial soil moisture is set to a volumetric soil-water content of 0.1, following Liu et al. (2021). Other parameters were also given the same values as the calibrated K2 model, including saturated hydraulic conductivity of channels (1 mm/hr), net capillary drive of channels (5 mm), hydraulic roughness of hillslopes (0.1 s/(m$^{1/3}$)), net capillary drive of hillslopes (50 mm), and soil porosity of 0.4. With this model set-up, we simulate the response to each of the 170 rainstorms for post-fire years 1, 2, 3, and 5.

3.4 Rainfall intensity-duration thresholds

Each K2 simulation results in a modeled hydrograph at the watershed outlet. As a first step towards defining a flash flood threshold, it is necessary to determine, based on the modeled time series of discharge, whether or not a flash flood would have occurred. We defined the flash flood level as the discharge required to exceed bankfull flow (Sweeney, 1992), which we assumed was equal to the two-year flood (Leopold et al., 1964). To determine the discharge associated with the two-year flood, we performed a flood frequency analysis using HEC-SSP v2.2 (Bartles et al., 2019) based on annual maximum records at the USGS stream gage station (11098000). The discharge associated with the two-year flood at the stream gage station is 15.3 m^3/s, with a 95% confidence interval of 12.3-19.2 m^3/s (Figure S3). A flash flood threshold by this definition can be viewed as conservative since it may only indicate the onset of minor flooding as water begins to spill out of the channel. Based on this definition, we then used two approaches to identify the rainfall ID threshold for flash floods (Figure 2).

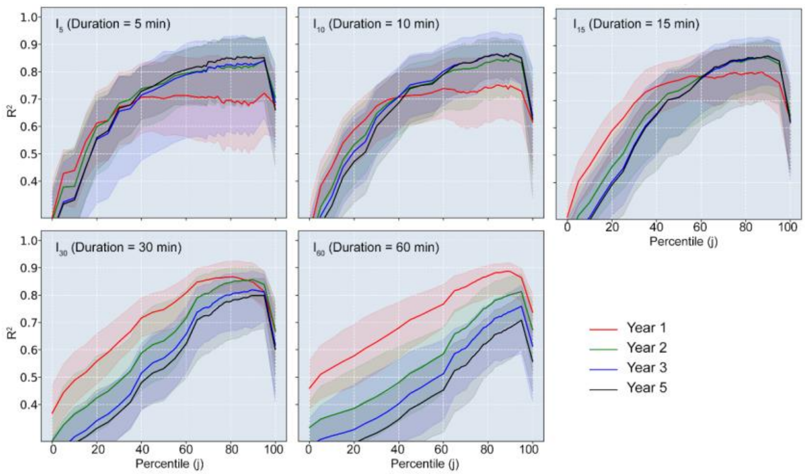
The first approach is based on a linear regression analysis that relates peak discharge with different rainfall ID metrics, namely I_D^j for different values of j and D. Using simulations of 170 rainfall-runoff events in each post-fire year, it is possible to determine a relationship for peak discharge (Q) as a function of I_D^j. Then, the rainfall ID threshold can be found by determining the rainfall intensity at which the peak discharge exceeds the bankfull capacity. The simplest quantitative relation is a linear regression:

$$Q = mI_D^j + k \tag{2}$$

where Q is the peak discharge (m^3/s) of a simulated hydrograph at the outlet, I_D^j denotes rainfall intensity (mm/hr) for the rainstorm that produced the hydrograph, and m and k denote the slope and y-intercept of the linear regression, respectively.

Considering the channel dimensions and resolution of the DEM used in the K2 model, we selected intensity-discharge (I_D^j -Q) pairs associated with Q greater than 2 m^3/s. The f~~F~~low depth associated a with~~of~~ Q less than 2 m^3/s would be~~are~~ very small and any impact from such flow would be~~potential damage is~~ negligible. The parameters in the linear equation (1) with the maximum determination coefficient (R^2_{max}) were estimated using least-squares linear regression in the SciPy Python library for the selected I_D^j-Q pairs. A total of 495 linear regressions were produced for each year because I_D^j can take on 495 different values (5 durations, 99 percentiles) for each rainstorm. For each post-fire year, we then identified the maximum R^2 value for each duration as a function of percentile from 1st to 99th (Figure 3). The rainfall ID threshold for flash flooding in each year was found, for each duration, from the linear relation associated with the largest R^2 (Figure 4).

Formatted: Font: Italic



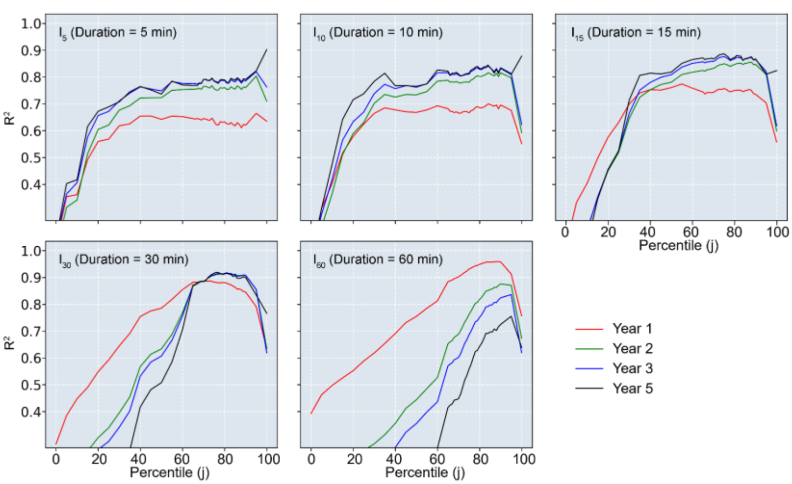
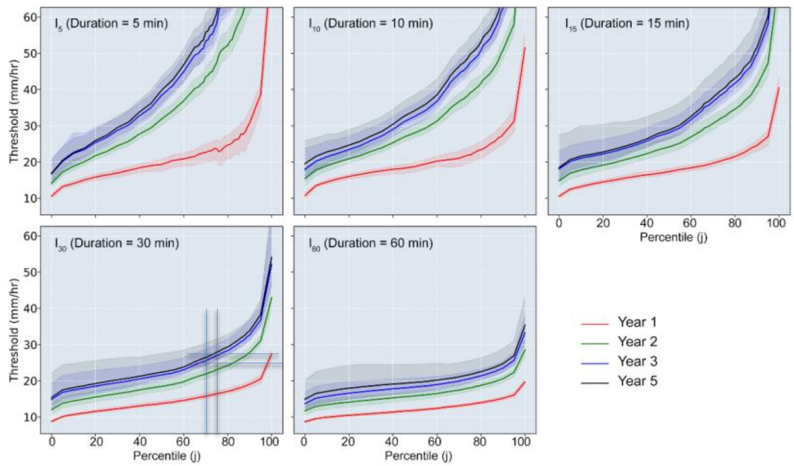
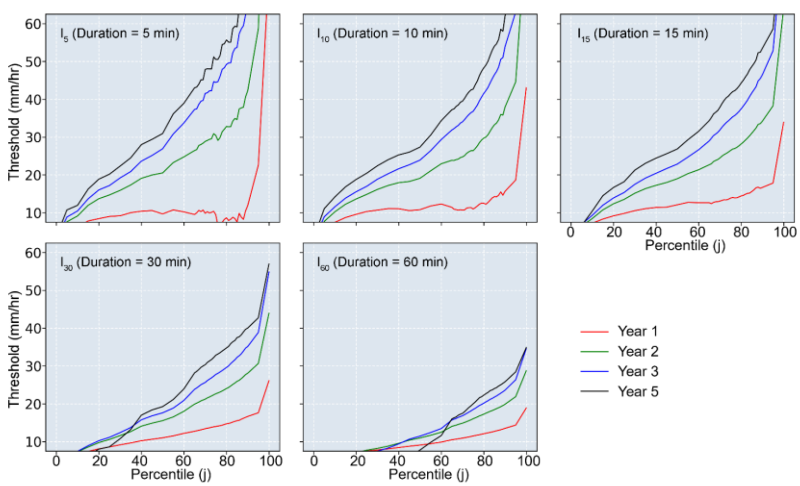


Figure 3: The determination coefficient (R^2) and 95% confidence interval associated with the linear regression between I_D^j and peak discharge in post-fire year 1, 2, 3, and 5. Data used to fit the linear relation are~~is~~ from events with peak discharge greater than 2 m^3/s.





Figure~~.~~ 4: The rainfall intensity-duration threshold for flash flood derived from the best linear relation for different durations and percentiles of the most intense~~ive~~ rainfall field in post-fire year 1, 2, 3, and 5.

We also estimated the 95% confidence interval (CI) of both ~~the~~ R^2 and the rainfall ID threshold by performing bootstrapping resampling on 170 rainfall-runoff events for each year. The number of replications is 50. The 95% CI was constructed with the 2.5 percentile and the 97.5 percentile of the ranked R^2 or rainfall ID threshold.

Formatted: Superscript

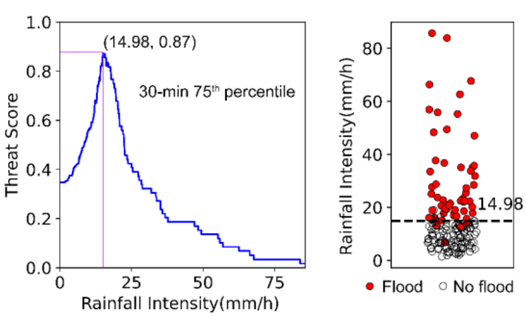
Formatted: Superscript

The second approach for determining rainfall ID thresholds is based on a receiver operating characteristic (ROC) analysis following Staley et al. (2013). We assess the utility of a potential threshold (e.g. $I_{30}^{50} = 20\ mm/hr$), by computing the threat score (TS) associated with using that threshold to define the transition between rainstorms that produce flash floods and those that do not. The TS, as one of the ROC utility functions, measures the fraction of forecast events that were correctly predicted:

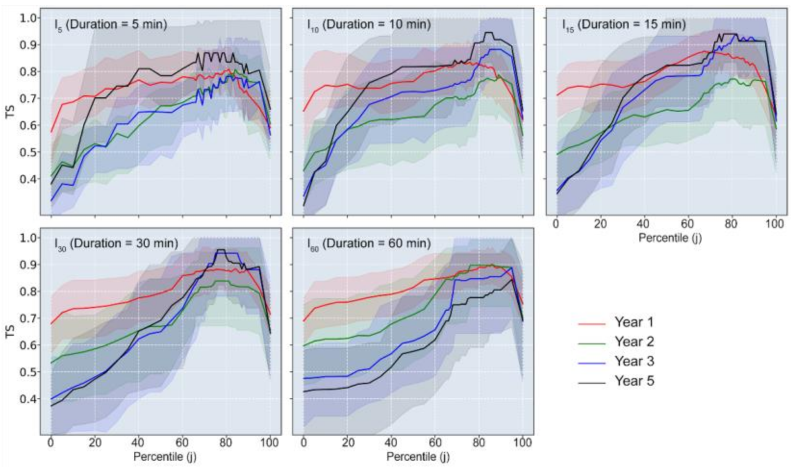
$$TS = \frac{TP}{TP + FP + FN} \tag{3}$$

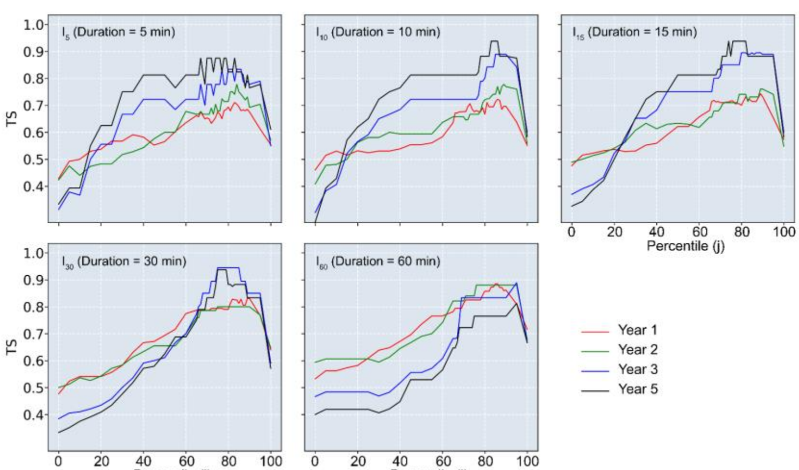
where TP, FP, and FN denote a true positive, false positive, and false negative, respectively. Flash flood occurrence (true or false) is determined by comparing the peak discharge of each simulated hydrograph with the flash flood level (15.3 m^3/s). A TP represents an event where rainfall rates exceed the threshold (e.g. $I_{30}^{50} = 20 mm/hr$), and a flash flood occurred. A FP represents an event where rainfall rates exceed the threshold, but no flash flood occurred. FN events occur when rainfall rates were below the threshold, yet a flash flood occurred. The optimal TS is 1, meaning use of the threshold resulted in no false positives or false negatives.

For a given rainfall intensity metric (e.g. the peak 75th percentile of I_{30}, I_{30}^{75}, in year 1), we calculated TS for intensities ranging from 0-100 mm/hr at 0.01 mm/hr intervals (Figure 5). We then identified the threshold associated with the maximum TS (TS_{max}). The intensity associated with TS_{max} is the optimal threshold for that rainfall metric (Figure 6). We determined the optimal threshold associated with each of the 495 rainfall metrics for each post-fire year (1,2,3, and 5) (Figure 7). We also estimated the 95% CI of TS and rainfall ID threshold for each year by performing bootstrapping resampling with 50 replications.

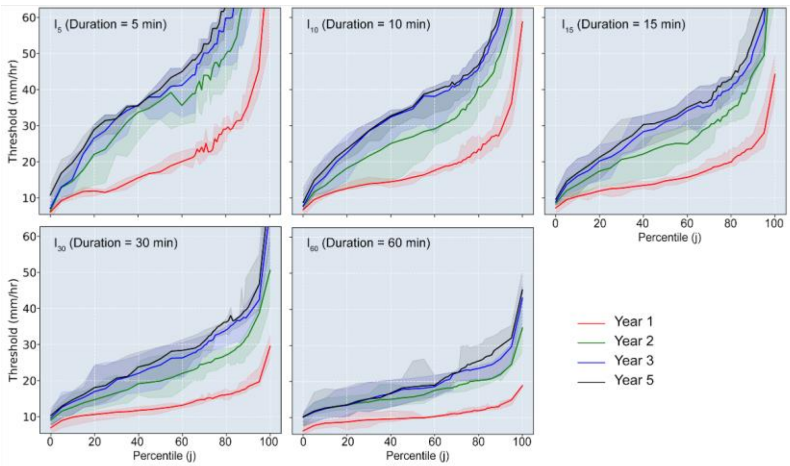


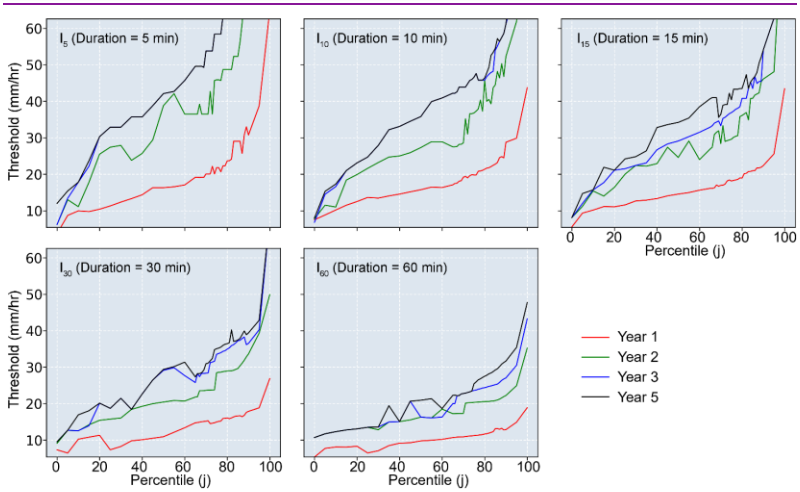
Figure~~.~~ 5: Threat score (TS) of the peak 75th percentile of I_{30} in post-fire year 1. (a) Relationship between rainfall intensity and TS; (b) Scatter plots of positive (flood, red circle) and negative (no flood, hollow circle) with the rainfall intensity associated with the maximum TS.





Figure~~.~~ 6: The threat scores (TS_{max}) associated with flood occurrence and I_D^j in post-fire years 1, 2, 3, and 5. Data used to analyze is from events with peak discharge greater than 2 m^3/s.





Figure~~.~~ 7: The rainfall intensity threshold for flash flood derived from the maximum of TS for different durations and percentiles of the most intensive rainfall field in post-fire years 1, 2, 3, and 5.

4 Results

4.1 Optimal summary metrics for defining rainfall ID thresholds

~~Linear regression analyses suggest that there is a stronger relationship between I_D^j and peak discharge (Q) as j increases, with the exception of a rapid dropoff in R^2 for j>90 and durations (D) greater than 5 minutes (Figure 3). For durations of 5-15 min, R^2 were low in the first 20-30 percentiles, then increased to 0.61-0.82 between the 30^{th}-90^{th} percentiles. Whereas the high R^2 interval for durations of 30 min and 60 min were with the largest value between 0.92-0.96 between the 60^{th}-90^{th} percentiles in year 1-5. The optimal rainfall threshold for flash floods (based on regressions of Q as a function of I_D^j) increased from 13.3 mm/hr of I_{60}^{89} (the 89^{th} percentile of 60 min peak rainfall field) in year 1 to 33.2 mm/hr of I_{30}^{76} (the 76^{th} percentile of 30 min peak rainfall field) in year 5 (Figure 4; Table 2). More generally, averaging rainfall intensity over a duration of 30 minutes and choosing a percentile, j, of approximately 75-85 leads to threat scores of approximately 0.8 or greater for all post-fire years. None of the other rainfall summary metrics performed this well across all post-fire years.~~

Linear regression analyses suggest that there is a stronger relationship between I_D^j and peak discharge (Q) as j increases, with the exception of a rapid dropoff in R^2 for j>95 (Figure 3). In post-fire year 1, the maximum R^2 increases with duration (D) from a value of 0.72 associated with I_{05}^{95}, to 0.75 associated with I_{10}^{85}, 0.80 associated with $I_{15}^{72} - I_{15}^{87}$, 0.87 associated with I_{30}^{81}, to 0.89 associated with I_{60}^{89}. In post-fire years 2-5, the ~~high~~ R^2 ~~interval for~~ values associated with durations of 5 min, 10 min, and 15 min ~~were with the large value between 079-0.86 between the~~ were maximized (0.79-0.86) within a window from the 60^{th}-95^{th} percentiles. The optimal rainfall threshold for flash floods (based on regressions of Q as a function of I_D^j) increased from 10.1 mm/hr of I_{60}^{89} (the 89^{th} percentile of 60 min peak rainfall field) in year 1 to 44.6 mm/hr of I_{15}^{90} (the 90^{th} percentile of 15 min peak rainfall field) in year 5 (Figure 4; Table 2). More generally, averaging rainfall intensity over a duration of 15 minutes and choosing a percentile, j, of approximately 75-90 produced an R^2 of approximately 0.80 or greater for all post-fire years (Figure 3). None of the other rainfall summary metrics performed this well across all post-fire years.

~~**Table. 2 The optimal metrics of rainfall ID and corresponding rainfall thresholds for flash floods in post-fire year 1-5**~~

	~~Linear regression~~				~~Receiver operating characteristic (ROC)~~		
~~Year~~	~~Rainfall metric~~	~~Equation~~	~~R^2_{max}~~	~~Intensity (mm/hr)~~	~~Rainfall metric~~	~~TS_{max}~~	~~Intensity (mm/hr)~~
~~1~~	~~I_{60}^{89}~~	~~$Q = 10.25 * I_{60}^{89} - 121.27$~~	~~0.958~~	~~13.3~~	~~$I_{60}^{85} - I_{60}^{86}$~~	~~0.89~~	~~13.1-13.2~~
~~2~~	~~I_{30}^{81}~~	~~$Q = 2.38 * I_{30}^{81} - 42.64$~~	~~0.916~~	~~24.4~~	~~$I_{60}^{76} - I_{60}^{95}$~~	~~0.88~~	~~20.4-25.0~~

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

~~3~~	~~I_{30}^{81}~~	~~$Q = 1.91 * I_{30}^{81} - 41.92$~~	~~0.917~~	~~30.0~~	~~$I_{30}^{75} - I_{30}^{85}$~~	~~0.94~~	~~33.5-37.3~~
~~5~~	~~I_{30}^{76}~~	~~$Q = 2.38 * I_{30}^{76} - 63.70$~~	~~0.919~~	~~33.2~~	~~$I_{30}^{75} - I_{30}^{79}$~~	~~0.94~~	~~35.1-36.3~~

~~Note: We denote the peak jth percentile of I_D (rainfall intensity over a duration D) as I_D^j. For example, I_{30}^{81} is the peak value of the 81st percentile of I_{30} (rainfall intensity over 30-min).~~

Table. 2 The linear regression-based optimal ~~metrics of~~ rainfall ID metrics and corresponding rainfall thresholds for flash floods in post-fire years 1-5

Year	Rainfall metric	Equation	R^2_{max} (95% CI)	Intensity (mm/hr) (95% CI)
1	I_{60}^{89}	$Q = 8.51 * I_{60}^{89} - 70.19$	0.89 (0.80, 0.92)	15.05 (14.50, 15.53)
2	I_{15}^{88}	$Q = 0.94 * I_{15}^{88} - 14.86$	0.86 (0.73, 0.92)	39.23 (36.97, 41.84)
3	I_{15}^{90}	$Q = 0.63 * I_{15}^{90} - 11.41$	0.86 (0.76, 0.93)	49.87 (36.68, 55.44)
5	I_{15}^{90}	$Q = 0.60 * I_{15}^{90} - 11.51$	0.86 (0.70, 0.92)	51.64 (48.18, 60.13)

Note: We denote the peak jth percentile of I_D (rainfall intensity over a duration D) as I_D^j. For example, I_{15}^{88} is the peak value of the 88th percentile of I_{15} (rainfall intensity over 15-min).

Table. 3 The ROC-based optimal metrics of rainfall ID and corresponding rainfall thresholds for flash floods in post-fire year 1-5

Year	Rainfall metric	TS_{max} (95% CI)	Intensity (mm/hr) (95% CI)
1	I_{60}^{86}	0.90 (0.84, 0.96)	12.91 (12.20, 13.20)
2	I_{60}^{76}	0.90 (0.74, 0.99)	19.98 (17.80, 20.40)
3	I_{30}^{75}	0.94 (0.78, 1.00)	32.60 (28.64, 33.60)
5	I_{30}^{76}	0.96 (0.82, 1.00)	34.86 (32.20, 35.40)

Note: We denote the peak jth percentile of I_D (rainfall intensity over a duration D) as I_D^j. For example, I_{60}^{86} is the peak value of the 86th percentile of I_{60} (rainfall intensity over 60-min).

Thresholds derived using the ROC method yielded broadly similar trends. The maximum threat score, TS_{max}, generally increased with *j* up to a point (approximately *j*=90) and then began to decrease regardless of the choice of duration (*D*) (Figure 6). The highest threat scores (TS), regardless of post-fire year or duration, were generally associated with the 70[th]-95[th] percentiles. For events in years 1-2, TS_{max} (0.90) occurs $I_{60}^{76} - I_{60}^{86}$ (the 76[th] -86[th] percentile of the peak I_{60} rainfall field); for events in years 3-5, the TS_{max} (0.94-0.96) occurs around I_{30}^{75} (the 75[th] percentile of the peak I_{30} rainfall field). The optimal rainfall threshold for a flash flood increased from $I_{60}^{86} = 12.9\ \mathrm{mm/hr}$ (the 86[th] percentile of 60 min peak rainfall field) in year 1 to $I_{30}^{76} = 34.9\ \mathrm{mm/hr}$ (the 76[th] percentile of 30 min peak rainfall field) in year 5 (Table 3; Figure 6). Averaging rainfall intensity over a duration of 30 minutes and choosing a percentile, *j*, of approximately 75-85 leads to threat scores of approximately 0.9 or greater for all post-fire years. Other metrics did not perform this well, on average, across all post-fire years.

4.2 Increases in rainfall intensity thresholds with time since fire

The rainfall intensity thresholds at each percentile ~~significantly~~ increased substantially from post-fire year 1 to 5 (Figures 4 and 7). However, the increase from year 1 to 2 is considerably larger than that from year 2 to 3 or from year 3 to year 5. Taking the I_{30}^{75} (the 75[th] percentile of the peak I_{30} rainfall field) as an example due to its strong performance as a threshold for all post-fire years, the thresholds based on linear regression analyses in year 1, 2, 3, and 5 are 16.8, 23.2, 26.9, and 27.6 mm/hr, respectively; the ROC-based I_{30}^{75} thresholds in year 1, 2, 3, and 5 are 16.0, 26.9, 32.6, and 34.5 mm/hr, respectively (Figure 7).

We are also able to use the model to assess the individual impacts of temporal changes in K_{sh} and n_c on temporal variations in the flash flood threshold. If K_{sh} is allowed to vary from year to year (Table 1) and n_c is held fixed at its calibrated value for year 1, then ROC analysis indicates that the optimal threshold of I_{30}^{75} still increases with time since burning (Figure 8). However, it increases slower than the case where both K_{sh} and n_c are allowed to vary with time (Figure 8). If n_c is allowed to vary from year to year (Table 1) and K_{sh} is held fixed at its calibrated value for year 1, then ROC analysis indicates that the optimal threshold associated with I_{30}^{75} increases from year 1 to year 2 but then stays roughly constant as time increases (Figure 8). Therefore, changes in K_{sh} and n_c both play important roles in determining the degree to which the flash flood threshold increases from year 1 to year 2, but that further increases in the threshold in years three and five are driven mainly by increases in K_{sh} as a function of time since burning.

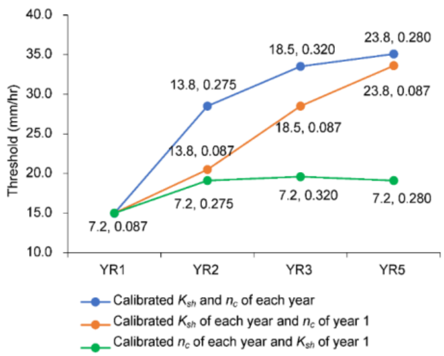


Figure 8: The ROC (receiver operating characteristic) based thresholds for I_{30}^{75} in each year with different model settings. Pairs of K_{sh} (saturated hydraulic conductivity on hillslopes) and n_c (Manning's *n* in channels) in each model are along with the data points.

5 Discussion

5.1 Implication of optimal metrics of rainfall intensity for flood warning

Rain gage records, which provide rainfall intensity data at a single point, are often used to define rainfall ID thresholds in debris-flow and flash flood studies (e.g. Moody and Martin, 2001; Cannon et al., 2008; Cannon et al., 2011; Guzzetti et al., 2008; Kean et al., 2011; Staley et al., 2013; Raymond et al., 2020; McGuire and Youberg, 2020). Using point source data to define thresholds for debris flows and flash floods is ideal when rainfall intensity does not vary substantially over the watershed, an assumption that is most appropriate for watershed areas less than several square kilometers. Radar-derived rainfall data has the advantage of providing spatially explicit information over an entire watershed at a high-temporal resolution (e.g. 5 minute). However, one challenge in using radar-derived precipitation to define thresholds is the need to condense spatially and temporally variable rainfall intensity information down to a single rainfall intensity metric. Regardless of whether the approach to determining an ID threshold involves fitting empirical relationships (e.g., Moody and Martin, 2001; Cannon et al., 2008) or using ROC analysis (e.g., Staley et al., 2013), a single metric is required to represent the rainfall intensity for each duration.

We summarized spatially variable rainfall intensity data over the watershed by computing the peak value of $I_D^j(t)$, the j[th] percentile of $I_D(t)$ for each rainstorm. We used two different techniques, one based on a linear regression analysis and one based on ROC analysis (Figure 2), to define thresholds for flash floods in post-fire years 1, 2, 3, and 5. Although the optimal metrics produced by the two approaches are not identical, they are generally similar in each post-fire year. In particular, high R^2 and TS_{max} values are associated with metrics of the peak 75th-85th percentile of rainfall intensity averaged over 15-60 minutes (I_D^j for $75 \leq j \leq 85, D = 15{,}30{,}60$). In other words, a good indicator of the potential for a flash flood is the presence of intense pulses of rainfall over durations of 15-60 minutes that cover at least 15%-25% of the watershed (Figure 9). This finding highlights the ability of rainstorms to produce flash floods even if they don't cover the majority of the watershed with intense rainfall. If rainfall over the majority of the watershed was required to produce flash floods, then we would expect that I_D^j with j<50 would be a better predictor of flash floods.

Previous work has also identified that 30-minute rainfall intensity works well for predicting flash floods and debris flows (Moody and Martin, 2001; Kean et al., 2011; Staley et al., 2013). The finding that I_{15}^j, I_{30}^j and I_{60}^j work best as thresholds when $75 \leq j \leq 85$ could be helpful when issuing flash flood warnings based on radar-derived precipitation estimates or data from several real-time rain gages within a watershed. Current operational forecast models such as the High Resolution Rapid Refresh model have a horizontal resolution of 3km and minimum temporal resolution of 15 minutes (Benjamin et al., 2016; NOAA 2021a), such that it is feasible to use either I_{15}^j, I_{30}^j or I_{60}^j in an operational forecast setting. Where sufficient operational NEXRAD weather radar coverage is present, radar-derived precipitation estimates such as the MRMS (Zhang et al., 2016) can provide near-real-time precipitation estimates at 1 km and as fine as 15 min temporal resolution (NOAA 2021b). In the case of poor radar coverage, gap-filling radars may be temporarily deployed or installed (e.g., Jorgensen et al., 2011; Cifelli et al., 2018) to provide information necessary for accurate precipitation estimates. While the magnitude of rainfall thresholds estimated here may only work for similar, recently burned watersheds within the San Gabriel Mountains, ~~the use of metrics such as I_{30}^{75} as a reliable predictor of post-fire flash floods may be more general~~ this work provides a general guideline of exploring a reliable predictor of post-fire flash floods for other watersheds and settings. Further testing is needed in watersheds with different watershed size, topographic characteristics, landscape, and burn severity patterns.

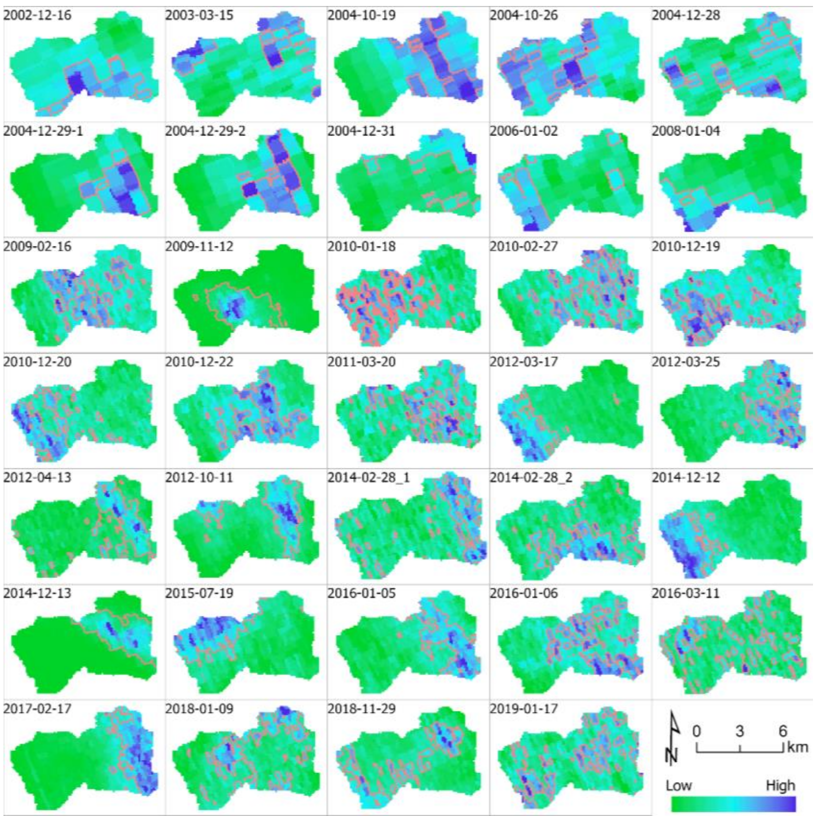
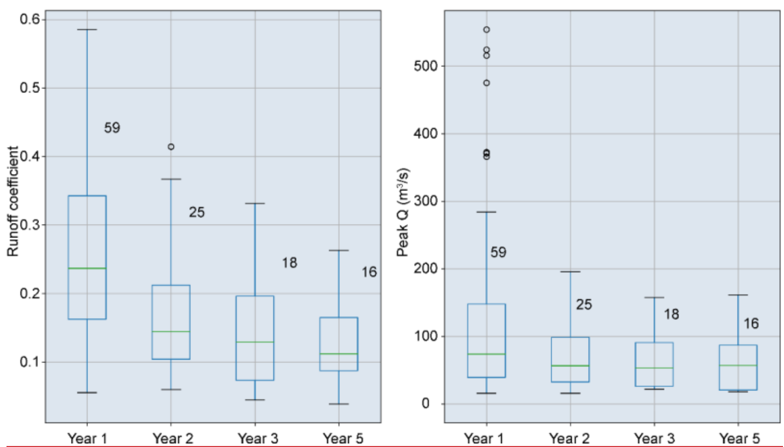


Figure 9: Snapshots of the spatial patterns of I_{30}^{75} of 34 unique storms. The peak jth percentile of I_D (rainfall intensity over a duration D) is denoted as I_D^j. I_{30}^{75} is the peak value of the 75st percentile of I_{30} (rainfall intensity over 30-min). Red contours delineate the pixels with rainfall intensities larger than I_{30}^{75} of each storm.

Several limitations are present in this work. First, we assess a small number of storm events (34) in the area as we are limited by the length of radar and gage records as well as and the number of events that impact the indicator rain gages, though applying the five Z-R relationships provides us with 170 rainfall realizations to assess. We prefer the use of observed rainfall data (radar and gauges) over simulated products, such as output from a rainfall generator (e.g., Zhao et al., 2019; Evin et al., 2018), as the radar is able to capture the spatial and temporal patterns of rainfall intensity in the study area's complex terrain. Though rainfall generators have advanced to represent some synoptic-to-mesoscale features, such as frontal and convective precipitation (e.g., Zhao et al. 2019), they are fundamentally designed to represent statistical characteristics of rainfall in places with limited observations (Wilks and Wilby 1999) and cannot be relied upon to replicate small scale storm characteristics in complex terrain (e.g., Camera et al. 2016). Future work could compare results from this hydrologic modeling experiment with observed versus simulated rainfall. ~~However, the advantage of using observed storms rather than using a rainfall generator (e.g., Zhao et al., 2019; Evin et al., 2018) is that our results represent spatial and temporal precipitation patterns that are physically realistic~~. Second, the challenges ~~of radar observations and application~~ of Z-R relationships to convert reflectivity to precipitation also presents challenges in accurately representing precipitation values. This can be addressed in future work through studies to constrain Z-R relationships for storms producing intense rainfall in this region and through the deployment or installation of high-resolution gap-filling radars (e.g., Johnson et al., 2019).

5.2 ~~Increasing rainfall intensity thresholds with time since fire~~The role of hydrological model in rainfall intensity thresholds estimation

In this study we employed the K2 model calibrated by Liu et al. (2021) to parameterize hydrologic changes affecting Hortonian overland flow within a five-year period following fire. Hillslope saturated hydraulic conductivity (K_{sh} = 7.2 mm/hr) and hydraulic roughness in channels (n_c = 0.087 s/m$^{1/3}$) were lowest immediately after fire (Table 1), resulting in high runoff coefficients and low rainfall thresholds in post-fire year 1. In later years, with K_{sh} and n_c gradually increasing (Table 1), more rainfall infiltrated into soil and there was increased attenuation of flood peaks. Simulations indicate that the number of flash-flood-producing rainstorms decreased from 59 in year 1 to 25, 18, and 16 in years 2, 3, and 5, respectively. Runoff coefficients and peak discharge of simulated hydrographs also decreased with time since fire (Figure 10). Given the same precipitation ensemble, the likelihood of flash floods significantly decreased with time. The peak discharge produced by the highest intensity rainfall event with I_{60}^{75} of 51.8 mm/hr was 554.0 m^3/s in the first year after the fire, which is three times greater than the peak discharges of 157.5 m^3/s in year 3 and 161.2 m^3/s in year 5 produced by the same rainstorm. From a flood hazard perspective, the downstream area may be exposed to a 1000-year flood under the recently burned condition (less than one year since the fire), whereas the discharge produced in years three and five would amount to roughly a 30- to 40-year flood (Figure S3).



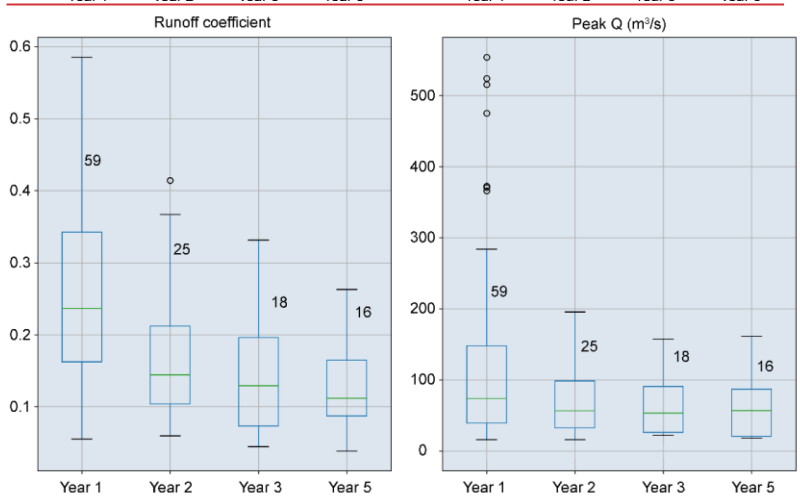


Figure 10: Box plots showing the runoff coefficient and peak discharge of flash floods in post-fire year 1, 2, 3, and 5. The numbers of flash floods in each year are displayed next to the box.

We were also able to perform numerical experiments to quantify the relative importance of temporal changes in K_{sh} and n_c on temporal variations in the flash flood threshold (Figure 8). Results suggest that changes in vegetation and grain roughness, which are likely to influence n_c, throughout the recovery process are less important for determining flash flood potential in our study area relative to changes to saturated hydraulic conductivity on hillslopes. It is worth noting that temporal changes in other model parameters (e.g., hydraulic roughness on hillslopes, capillary drive) may play more of a role in driving changes in post-fire flash flood thresholds in other settings. In this study, however, we focus on changes in K_{sh} and n_c because Liu et al. (2021) were able to detect temporal changes in n_c and K_{sh} through time and unable to detect similar temporal changes in other hydrologic parameters (e.g., hydraulic roughness on hillslopes, capillary drive) due to their relatively minor influence on runoff in the study watershed.

In this study, the optimal flash flood thresholds increased from I_{30}^{75} =16.0-16.8 mm/hr in post-fire year 1, to 23.2-26.9 mm/hr in year 2, and 27.6-34.5 mm/hr in post-fire year 5 (Figure 4 and 7; Table 2-3). In the San Gabriel Mountains and nearby San Bernardino and San Jacinto Mountains, Cannon et al. (2008) estimated rainfall thresholds of I_{30}=9.5 mm/hr for flash floods and debris flows in the first winter rainy season following fire. They found that the thresholds for flash floods and debris flows increased to I_{30}=19.8 mm/hr in post-fire year 2. The thresholds that we infer from hydrological modeling are greater than those reported by Cannon et al. (2008), which may be partly due to differences in (1) data and methods used and (2) the size of the studied watersheds. Our results are driven by a hydrologic model, forced with a radar precipitation ensemble that consists of 170 rainstorms that contain a variety of storm types that impact southern California. The occurrence of a flash flood is based on exceedance of the maximum channel capacity and we summarize temporal changes in the rainfall ID threshold using I_{30}^{75} since we find this to be a reliable metric for all post-fire years included in this study. In contrast, Cannon et al. (2008) established rainfall ID relations by using observations of rainstorms and hydrological response in the two years following fire in 87 small watersheds (0.2-4.6 km^2). They base their thresholds on rainfall characteristics that produced either flash floods or debris flows whereas we focus solely on flash floods. In their dataset, flash floods and debris flows were identified by investigating flood and debris flow deposits at the outlet of those small watersheds in the field. Despite differences in the magnitude of the thresholds, the increase in the threshold from post-fire year 1 to year 2 in both studies are quite close. This agreement provides support for the use of simulation-based approaches to inform temporal shifts in rainfall ID thresholds.

During the recovery process, increasing thresholds for flash floods and debris flows have also been identified in other areas at different scales by either observation- or simulation-based studies, such as hillslopes in the Colorado Front Range (Ebel, 2020) and small watersheds in Australia (Noske et al., 2016). The consistent increase in rainfall ID thresholds with time since fire in different geographic and ecological zones implies that hydraulic and hydrologic models may be useful tools for exploring how transient effects of fire translate into changes in water-related hazards. Particularly when historic data is limited and traditional empirical methods are impractical for defining thresholds, the role of hydraulic and hydrological models becomes more important.

6 Conclusions

We used 250 m, 5-minute radar-derived precipitation estimates over a 41.5 km^2 watershed in combination with a calibrated hydrological model to estimate the rainfall intensity thresholds for post-fire flash floods as a function of time since burning. The main ~~findings~~outcomes of this study are 1) identification of~~ the~~ optimal ~~metrics of~~ radar-derived rainfall metrics for~~in~~ post-fire flash flood prediction in southern California, 2) ~~demonstration of~~estimates of temporal changes in~~creasing~~ rainfall ID thresholds for flash floods following disturbance in a chapparal-dominated ecosystem, and 3) ~~proposal of~~ a methodology ~~in warning~~for using~~ of post-fire hydrologic hazards using~~ a hydrological model to assess changes in post-fire flash flood thresholds.

The optimal threshold for predicting the occurrence of a flash flood in our study areas is the 75th-85th percentile of peak rainfall intensity averaged over 15-60 minutes, i.e., I_{30}^{75}-I_{30}^{85}. In other words, a flash flood tends to be produced when rainfall intensity over 15%-25% of the watershed area exceeds a critical value. A threshold based on I_{30}^{75} performs consistently well for post-fire years 1, 2, 3, and 5, although the magnitude of the threshold increases with time since burning.

For the watershed studied, the I_{30}^{75} threshold increases from 16.0-16.8 mm/hr for year 1 to 23.2-26.9 mm/hr, 26.9-32.6 mm/hr, and 27.6-34.5 mm/hr, for years 2, 3, and 5 respectively. Increases in the threshold value of I_{30}^{75} can be primarily attributed to increases in K_{sh} rather than n_c during the hydrological recovery process. The increase in the magnitude of the threshold from year 1 to year 2 is consistent with previous observations from nearby areas in southern California.

Results provide a methodology for using radar-derived precipitation estimates and hydrological modeling to estimate flash flood thresholds for improved warning and mitigation of post-fire hydrologic hazards. Thresholds developed through these methods can then be built into operational tools that use incoming radar data to evaluate flash flood hazard in near-real time or precipitation forecasts to evaluate potential for flash flood hazard in burned watersheds.

Author contributions

TL and LM conceived the study. TL, LM, NO and FC contributed to the development and design of the methodology. TL ~~analysed~~analyzed and prepared the manuscript with review and analysis contributions from LM, NO and FC.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments:

Haiyan Wei, Carl L. Unkrich, and David C. Goodrich, who are from the KINEROS2 development group in the USDA ARS Southwest Watershed Research Center in Tucson, helped with the setting up of the KINEROS2 model and ingestion of the RADAR precipitation data into the model. We are thankful for their great help.

Financial support

This work was supported by the National Oceanic and Atmospheric Administration (NOAA) Collaborative Science, Technology, and Applied Research (CSTAR) Program under grant NA19NWS4680004 and by the National Integrated Drought Information System (NIDIS) through Task Order 1332KP20FNRMT0012.

References

Bartles, M., Brunner, G., Fleming, M., Faber, B., Karlovits, G., and Slaughter, J.: HEC-SSP Statistical Software Package Version 2.2, June 2019.

Benjamin, S. G., Weygandt, S. S., Brown, J. M., Hu, M., Alexander, C. R., Smirnova, T. G., Olson, J. B., James, E. P., Dowell, D. C., Grell, G. A., Lin, H., Peckham, S. E., Smith, T. L., Moninger, W. R., Kenyon, J. S., & Manikin, G. S. (2016). A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh, *Monthly Weather Review*, *144*(4), 1669-1694. https://doi.org/10.1175/MWR-D-15-0242.1

Camera, C., Bruggeman, A., Hadjinicolaou, P., Michaelides S., & Lange, M. A. (2017). Evaluation of a spatial rainfall generator for generating high resolution precipitation projections over orographically complex terrain. *Stoch Environ Res Risk Assess* 31, 757–773. https://doi.org/10.1007/s00477-016-1239-1

Canfield, H. E., Goodrich, D. C., & Burns, I. S. (2005). Selection of parameters values to model post-fire runoff and sediment transport at the watershed scale in southwestern forests. *Proceedings of the 2005 Watershed Management Conference - Managing Watersheds for Human and Natural Impacts: Engineering, Ecological, and Economic Challenges*, *40763*(June 2015), 561–572. https://doi.org/10.1061/40763(178)48

Cannon, F., and Coauthors, 2020: Observations and Predictability of a High-Impact Narrow Cold-Frontal Rainband over Southern California on 2 February 2019. *Wea. Forecasting*, 35, 2083–2097, https://doi.org/10.1175/WAF-D-20-0012.1.

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

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

Carsel, R. F., & Parrish, R. S. (1988). Developing joint probability distributions of soil water retention characteristics. *Water Resources Research*, *24*(5), 755–769. https://doi.org/10.1029/WR024i005p00755

Cifelli, R., Chandrasekar, V., Chen, H., & Johnson, L. E. (2018). High resolution radar quantitative precipitation estimation in the San Francisco Bay Area: Rainfall monitoring for the urban environment. *Journal of the Meteorological Society of Japan. Ser. II*, *96*, 141-155.

Dyrness, C. (1976), Effect of Wildfire on Soil Wettability in the High Cascades of Oregon, U.S. Dept. of Agriculture, Forest Service, Pacific Northwest forest and Range Experiment Station, Portland, Oreg.

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

Ebel, B. A., & Martin, D. A. (2017). Meta - analysis of field - saturated hydraulic conductivity recovery following wildland fire : Applications for hydrologic model parameterization and resilience assessment. Hydrological Processes. 2017; 31: 3682–3696. https://doi.org/10.1002/hyp.11288

Ebel, B. A., & Moody, J. A. (2013). Rethinking infiltration in wildfire-affected soils. Hydrological Processes, 27(10), 1510–1514. https://doi.org/10.1002/hyp.9696

Ebel, B. A., & Moody, J. A. (2017). Synthesis of soil-hydraulic properties and infiltration timescales in wildfire-affected soils. *Hydrological Processes*, *31*(2), 324–340. https://doi.org/10.1002/hyp.10998

Evin, G., Favre, A.-C., & Hingray, B. (2018). Stochastic generation of multi-site daily precipitation focusing on extreme events, Hydrol. Earth Syst. Sci., 22, 655–672, https://doi.org/10.5194/hess-22-655-2018.

Gillett NP, Weaver AJ, Zwiers FW, Flannigan MD. (2004). Detecting the effect of climate change on Canadian forest fires. *Geophysical Research Letters* 31. https://doi-org.ezproxy2.library.arizona.edu/10.1029/2004

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

Guzzetti, F., Peruccacci, S., Rossi, M., & Stark, C. P. (2008). The rainfall intensity-duration control of shallow landslides and debris flows: An update. *Landslides*, *5*(1), 3–17. https://doi.org/10.1007/s10346-007-0112-1

Hubbert, K. R., & Oriol, V. (2005). Temporal fluctuations in soil water repellency following wildfire in chaparral steeplands, southern California. *International Journal of Wildland Fire*, *14*(4), 439–447. https://doi.org/10.1071/WF05036

Hubbert, K. R., Wohlgemuth, P. M., & Beyers, J. L. (2012). Effects of hydromulch on post-fire erosion and plant recovery in chaparral shrublands of southern California. *International Journal of Wildland Fire*, *21*(2), 155–167. https://doi.org/10.1071/WF10050

Huffman, E. L., MacDonald, L. H., & Stednick, J. D. (2001). Strength and persistence of fire-induced soil hydrophobicity under ponderosa and lodgepole pine, Colorado Front Range. *Hydrological Processes*, *15*(15), 2877–2892. https://doi.org/10.1002/hyp.379

Johnson, LE, Cifelli, R, White, A. Benefits of an advanced quantitative precipitation information system. *J Flood Risk Management*. 2020; 13 (Suppl. 1):e12573. https://doi.org/10.1111/jfr3.12573

Jorgensen, D. P., Hanshaw, M. N., Schmidt, K. M., Laber, J. L., Staley, D. M., Kean, J. W., & Restrepo, P. J. (2011). Value of a dual-polarized gap-filling radar in support of southern California post-fire debris-flow warnings. *Journal of Hydrometeorology*, *12*(6), 1581-1595.

Kean, J. W., Staley, D. M., & Cannon, S. H. (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: Earth Surface, 116(4), 1–21. https://doi.org/10.1029/2011JF002005

Kean, J. W., Staley, D. M., & Cannon, S. H. (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: Earth Surface*, *116*(4), 1–21. https://doi.org/10.1029/2011JF002005

Kitzberger, T., Falk, D. A., Westerling, A. L., & Swetnam, T. W. (2017). Direct and indirect climate controls predict heterogeneous early-mid 21st century wildfire burned area across western and boreal North America. *PLoS ONE*, *12*(12), 1–24. https://doi.org/10.1371/journal.pone.0188486

Lamb, M. P., Scheingross, J. S., Amidon, W. H., Swanson, E., & Limaye, A. (2011). A model for fire-induced sediment yield by dry ravel in steep landscapes. *Journal of Geophysical Research: Earth Surface*, *116*(3), 1–13. https://doi.org/10.1029/2010JF001878

Larsen, I. J., MacDonald, L. H., Brown, E., Rough, D., Welsh, M. J., Pietraszek, J. H., et al. (2009). Causes of post-fire runoff and erosion: Water repellency, cover, or soil sealing? *Soil Science Society of America Journal*, *73*(4), 1393–1407. https://doi.org/10.2136/sssaj2007.0432

Leopold, L.B., Wolman, M.G., and Miller, J.P. 1964. Fluvial Processes in Geomorphology. Dover, New York.

Liu, T., McGurire, L. A., Wei, H. Y., Renger,s F. K., Gupta, H., Ji, L., Goodrich, D. C. (2021). The timing and magnitude of changes to Hortonian overland flow at the watershed scale during the post-fire recovery process. *Hydrological Processes.* 35(5). https://doi.org/10.1002/hyp.14208

Marshall, J. S., and W. M. K. Palmer, 1948: THE DISTRIBUTION OF RAINDROPS WITH SIZE. *J. Meteor.*, **5**, 165–166, https://doi.org/10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2.

McGuire, L. A., & Youberg, A. M. (2020). What drives spatial variability in rainfall intensity-duration thresholds for post-wildfire debris flows? Insights from the 2018 Buzzard Fire, NM, USA. *Landslides*, *December 2019*. https://doi.org/10.1007/s10346-020-01470-y

Meles, M .B., Goodrich, D. C., Gupta, H. V., Burns, S. I, Unkrich, C. L., Razavi, S., Guertin, D. P. (2019). Uncertainty and parameter sensitivity of the KINEROS2 physically-based distributed sediment and runoff model. In *Federal Interagency Sedimentation and Hydrologic Modeling Conference*.

Miller, S. N., Semmens, D. J., Goodrich, D. C., Hernandez, M., Miller, R. C., Kepner, W. G., & Guertin, D. P. (2007). The automated geospatial watershed assessment tool. 851 *Environmental Modelling and Software*. *22*(3), 365-377. 852 https://doi.org/10.1016/j.envsoft.2005.12.004

Moody, J. A., & Ebel, B. A. (2012). Hyper-dry conditions provide new insights into the cause of extreme floods after wildfire. *Catena*, *93*, 58–63. https://doi.org/10.1016/J.CATENA.2012.01.006

Moody, J. A., & Martin, D. A. (2001). Post-fire, rainfall intensity – peak discharge relations for three mountainous watersheds in the western USA. 2993(July), 2981–2993. https://doi.org/10.1002/hyp.386

Moreno, H. A., Gourley, J. J., Pham, T. G., & Spade, D. M. (2020). Utility of satellite-derived burn severity to study short- and long-term effects of wildfire on streamflow at the basin scale. *Journal of Hydrology*, *580*(October 2019), 124244. https://doi.org/10.1016/j.jhydrol.2019.124244

NOAA National Weather Service (NWS) Radar Operations Center (1991): NOAA Next Generation Radar (NEXRAD) Level 2 Base Data. NOAA National Centers for Environmental Information. doi:10.7289/V5W9574V

NOAA. (2021a). High-Resolution Rapid Refresh (HRRR) Model [Dataset]. https://registry.opendata.aws/noaa-hrrr-pds/

NOAA. (2021b) Multi-Radar/Multi-Sensor System (MRMS). [Dataset]. https://www.nssl.noaa.gov/projects/mrms/

Noske, P. J., Nyman, P., Lane, P. N. J., & Sheridan, G. J. (2016). Effects of aridity in controlling the magnitude of runoff and erosion after wildfire. *Water Resources Research*, *52*(6), 4338–4357. https://doi.org/10.1002/2015WR017611

Oakley, N. S., Lancaster, J. T., Kaplan, M. L., & Ralph, F. M. (2017). Synoptic conditions associated with cool season post-fire debris flows in the Transverse Ranges of southern California. *Natural Hazards*, *88*(1), 327–354. https://doi.org/10.1007/s11069-017-2867-6

Oakley, N. S., J. T. Lancaster, B. J. Hatchett, J. Stock, F. M. Ralph, S. Roj, & S. Lukashov. (2018a). A 22-Year climatology of cool season hourly precipitation thresholds conducive to shallow landslides in California. *Earth Interact.*, 22, 1–35, https://doi.org/10.1175/EI-D-17-0029.1.

Oakley, N. S., Cannon, F., Munroe, R., Lancaster, J. T., Gomberg, D., & Ralph, F. M. (2018b). Brief communication: Meteorological and climatological conditions associated with the 9 January 2018 post-fire debris flows in Montecito and Carpinteria, California, USA. *Natural Hazards and Earth System Sciences*, *18*(11), 3037-3043. https://doi.org/10.5194/nhess-18-3037-2018

Parlange, J. Y., Lisle, I., Braddock, R. D., & Smith, R. E. (1982). The three-parameter infiltration equation. *Soil Science*, *133*(6), 337-341, https://doi.org/10.1097/00010694-198206000-00001

Raymond, C. A., McGuire, L. A., Youberg, A. M., Staley, D. M., & Kean, J. W. (2020). Thresholds for post-wildfire debris flows: Insights from the Pinal Fire, Arizona, USA. Earth Surface Processes and Landforms, 45(6), 1349–1360. https://doi.org/10.1002/esp.4805

Rengers, F. K., McGuire, L. A., Kean, J. W., Staley, D. M., & Youberg, A. M. (2019). Progress in simplifying hydrologic model parameterization for broad applications to post-wildfire flooding and debris-flow hazards. In *Earth Surface Processes and Landforms* (Vol. 44, Issue 15, pp. 3078–3092). https://doi.org/10.1002/esp.4697

Saksa, P. C., Bales, R. C., Tobin, B. W., Conklin, M. H., Tague, C. L., & Battles, J. J. (2020). Fuels treatment and wildfire effects on runoff from Sierra Nevada mixed-conifer forests. Ecohydrology, 13, 1–16. https://doi.org/10.1002/eco.2151

Schmidt, K. M., M. N. Hanshaw, J. F. Howle, J. W. Kean, D. M. Staley, J. D. Stock, and G. W. Bawden (2011), Hydrologic conditions and terrestrial laser scanning of post-fire debris flows in the San Gabriel Mountains, CA, U.S.A., in Debris-Flow Hazards Mitigation, Mechanics, Prediction, and Assessment, edited by R. Genevois, D. L. Hamilton, and A. Prestininzi, pp. 583–593, Casa Editrice Univ. La Sapienza, Rome.

Smith, R. E., Goodrich, D. C., Woolhiser, D. A., & Unkrich, C. L. (1995). KINEROS—A kinematic runoff and erosion model. In VijayP. Singh (Ed.), *Computer models of watershed hydrology* (p. 1130). Fort Collins, CO: Water Resources Publications.

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

Staley, D.M., Wasklewicz, T.A. and Kean, J.W., 2014, Characterizing the primary material sources and dominant erosional processes for post-fire debris-flow initiation in a headwater basin using multi-temporal terrestrial laser scanning data: *Geomorphology*, v. 214, 324-338.

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

Stoof, C. R., Vervoort, R. W., Iwema, J., Ferreira, A. J. D., Ritsema, C. J., Group, D., Resources, N., Water, Q., Group, M., & Green, A. (2012). Hydrological response of a small catchment burned by experimental fire. 267–285. https://doi.org/10.5194/hess-16-267-2012

Sweeney, T. L. (1992). Modernized areal flash flood guidance. *NOAA Technical Memorandum, NWS HYDRO*, 44.

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. https://doi.org/10.1029/2018JF004837

USDA Forest Service (2009). Station fire burned area emergency response (BAER). unpublished Report. Available: http://www.fs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb5245056.pdf

Watson, C. L., & Letey, J. (1970). Indices for Characterizing Soil-Water Repellency Based upon Contact Angle-Surface Tension Relationships. *Soil Science Society of America Journal*, *34*(6), 841–844. https://doi.org/10.2136/sssaj1970.03615995003400060011x

Wilks, D. S., & Wilby, R. L. (1999). The weather generation game: a review of stochastic weather models. *Progress in Physical Geography: Earth and Environment*, 23(3), 329–357. https://doi.org/10.1177/030913339902300302

Westerling AL, Hidalgo HG, Cayan DR, Swetnam TW. 2006. Warming and earlier spring increase western US forest wildfire activity. *Science* 313: 940–943.

Wilson, C., Kampf, S. K., Wagenbrenner, J. W., & Macdonald, L. H. (2018). Forest Ecology and Management Rainfall thresholds for post- fi re runoff and sediment delivery from plot to watershed scales. *Forest Ecology and Management*, 430, 346–356. https://doi.org/10.1016/j.foreco.2018.08.025

Yatheendradas, S., Wagener, T., Gupta, H., Unkrich, C., Goodrich, D., Schaffner, M., & Stewart, A. (2008). Understanding uncertainty in distributed flash flood forecasting for semiarid regions. *Water Resources Research*, *44*(5), 1–17. https://doi.org/10.1029/2007WR005940

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., Grams, H., Wang, Y., Cocks, S., Martinaitis, S., Arthur, A., Cooper, K., Brogden, J., & Kitzmiller, D. (2016). Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimation: Initial Operating Capabilities, *Bulletin of the American Meteorological Society*, *97*(4), 621-638. Retrieved May 23, 2021~~, from~~. https://journals.ametsoc.org/view/journals/bams/97/4/bams-d-14-00174.1.xml

Zhao, Y., Nearing, M.A., & Guertin, D.P., 2019. A daily spatially explicit stochastic rainfall generator for a semi-arid climate. J. Hydrol. 574, 181–192.