Augmentation and Use of WRF-Hydro to Simulate Overland Flow- and Streamflow-Generated Debris Flow Hazards in Burn Scars

Chuxuan Li¹, Alexander L. Handwerger²³, Jiali Wang⁴, Wei Yu⁵⁶, Xiang Li⁷, Noah J. Finnegan⁸, Yingying Xie⁹¹⁰, Giuseppe Buscarnera⁷, and Daniel E. Horton¹

¹ Department of Earth and Planetary Sciences, Northwestern University, Evanston, IL, 60208, USA
² Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, CA, 90095, USA
³ Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, 91109, USA
⁴ Environmental Science Division, Argonne National Laboratory, Lemont, IL, 60439, USA
⁵ Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, CO, 80309, USA
⁶ NOAA/Global Systems Laboratory, 325 Broadway Boulder, Denver, CO, 80305-3328, USA
⁷ Department of Civil and Environmental Engineering, Northwestern University, Evanston, IL, 60208, USA
⁸ University of California Santa Cruz, Department of Earth and Planetary Sciences, Santa Cruz, CA, 95064, USA
⁹ Program in Environmental Sciences, Northwestern University, 2145 Sheridan Road, Evanston, IL, 60208, USA
¹⁰ Department of Biological Sciences, Purdue University, 915 W State St, West Lafayette, IN 47907, USA

Correspondence to: Chuxuan Li (chuxuanli2020@u.northwestern.edu)
Abstract

In steep wildfire-burned terrains, intense rainfall can produce large volumes of runoff that can trigger highly destructive debris flows. The ability to accurately characterize and forecast debris-flow hazards in burned terrains, however, remains limited. Here, we augment the Weather Research and Forecasting Hydrological modeling system (WRF-Hydro) to simulate both overland and channelized flows and assess postfire debris-flow hazards over a regional domain. We perform hindcast simulations using high-resolution weather radar-derived precipitation and reanalysis data to drive non-burned baseline and burn scar sensitivity experiments. Our simulations focus on January 2021 when an atmospheric river triggered numerous debris flows within a wildfire burn scar in Big Sur – one of which destroyed California’s famous Highway 1. Compared to the baseline, our burn scar simulation yields dramatic increases in total and peak discharge, and shorter lags between rainfall onset and peak discharge. At Rat Creek, where Highway 1 was destroyed, discharge volume increases eight-fold and peak discharge triples relative to the baseline. For all catchments within the burn scar, we find that the median catchment-area normalized discharge volume increases nine-fold after incorporating burn scar characteristics, while the 95th percentile volume increases 13-fold. Catchments with anomalously high hazard levels correspond well with post-event debris flow observations. Our results demonstrate that WRF-Hydro provides a compelling new physics-based tool to investigate and potentially forecast postfire hydrologic hazards at regional scales.

Short Summary

In January 2021 a storm triggered numerous debris flows in a wildfire burn scar in California. We use a hydrologic model to assess debris flow hazards in pre-fire and postfire scenarios. Compared to pre-fire conditions, the postfire simulation yields dramatic increases in total and peak discharge, substantially increasing debris flow hazards. Our work demonstrates the utility of 3-D hydrologic models for investigating and potentially forecasting postfire debris flow hazards at regional scales.

1 Introduction

Following intense rainfall, areas with wildfire burn scars are more prone to flash flooding (Neary et al., 2003; Bart & Hope 2010; Bart 2016) and runoff-generated debris flow hazards than unburned areas (Moody et al., 2013; Ice et al., 2004; Shakesby & Doerr, 2006). After wildfire, reduced tree canopy interception, decreased soil infiltration due to soil-sealing effects (Larsen et al., 2009), and increased soil water repellency – especially in hyper-arid environments (Dekker and Ritsema, 1994; Doerr and Thomas, 2000; MacDonald and Huffman, 2004) – increases excess surface water, and on sloped terrains leads to overland flow (Shakesby & Doerr, 2006; Stoof et al.,...
As water moves down hillslopes and erosion adds sediment to water-dominated flows, clear water floods can transition to turbulent and potentially destructive debris flows (Meyer & Wells, 1997; Cannon et al., 2001, 2003; Santi et al., 2008). In contrast to debris flows initiated by shallow landslides, this rainfall-runoff process has been identified as the major cause for postfire debris flows in the western U.S. (Cannon, 2001; Cannon et al., 2003, 2008; Kean et al., 2011; Nyman et al., 2015; Parise & Cannon, 2012), and in other regions with Mediterranean climates (Mitsopoulos & Mironidis, 2006; Bisson et al., 2005; Parise & Cannon, 2008, 2009; Rosso et al., 2007). In California, because climate change is projected to increase the intensity and frequency of wet-season precipitation (Swain et al., 2018; Polade et al., 2017), increase wildfire potential (Swain, 2021; Brown et al., 2021), and extend the wildfire season (Goss et al., 2020), occurrence and intensity of postfire debris flows are likely to increase (Cannon et al., 2009; Kean & Staley, 2021; Oakley, 2021).

To assess postfire debris flow hazards, statistical approaches including empirical models and machine-learning techniques are commonly used in both research and operational settings (Gardner et al., 2014; Cannon et al., 2010; Staley et al., 2016; Cui et al., 2019; Nikolopoulos et al., 2018; Friedel 2011a, 2011b). Statistical approaches are useful for identifying and characterizing relationships amongst contributing environmental factors and are helpful in operational settings due to low computational costs and the potential for rapid assessment. For example, the U.S. Geological Survey (USGS) currently employs a statistical approach in their Emergency Assessment of Postfire Debris-flow Hazards that consists of a logistic regression model to predict the likelihood of post-wildfire debris flows (e.g., Staley et al., 2016; Cannon et al., 2010), and a multiple linear regression model to predict debris flow volumes (Gartner et al., 2014). Machine-learning techniques have also been used to predict postfire debris flows in the western U.S. (Nikolopoulos et al., 2018; Friedel 2011a, 2011b). For example, self-organizing maps and genetic programming were used to predict postfire debris flow occurrence (Friedel 2011b) and volumes (Friedel 2011a), respectively. Compared to the current USGS predictive models, genetic programming was posited to be more useful in solving non-linear multivariate problems (Friedel 2011a), while a random forest algorithm demonstrated increased performance in predicting postfire debris flow occurrence (Nikolopoulos et al., 2018). Despite the utility and advantages of data-driven hazard prediction approaches, these techniques do not simulate the underlying physics, which limits their utility in developing a better process-based understanding of debris flow mechanics, limits their applicability in climatological and geographic settings different than their training sites, and limits their use in non-stationary conditions (e.g., under changing climatic conditions).

In contrast, physics-based models that simulate spatially-explicit hydrologic and mass wastage processes are well-suited for mechanistic sensitivity analyses in diverse settings, but applications of these models have tended to focus on landslide-induced debris flows (e.g., Iverson & George, 2014; George & Iverson, 2014), rather than runoff-generated debris flows which are more common in postfire areas (Cannon et al., 2001, 2003; Santi et al., 2008). Studies that have investigated postfire hydrologic responses using process-based models have largely focused on short-term...
responses in individual catchments at high spatiotemporal resolutions (McGuire et al., 2016, 2017; Rengers et al., 2016) or long-term runoff responses at coarse temporal resolutions (Rulli & Rosso, 2007; McMichael & Hope, 2007). For example, process-based models have employed shallow water equations to understand the triggering and transport mechanisms of postfire debris flows in single catchments (McGuire et al., 2016, 2017) and to investigate the timing of postfire debris flows in three separate catchments (Rengers et al., 2016), the latter of which also assessed the efficacy of a simplified kinematic wave approach. In addition to individual catchment applications, process-based models often adopt simplifications that can limit effective prediction and hypothesis testing to overcome computational limits. For example, the kinematic runoff and erosion model (KINEROS2) simplifies drainage basins into 1-dimensional channels and hillslope patches (Canfield & Goodrich, 2005; Goodrich et al., 2012; Sidman et al., 2015), and the Hydrologic Modeling System (HEC-HMS) uses an empirically-based curve number method to estimate saturation excess water (Cydzik et al., 2009), which cannot resolve infiltration excess overland flow, a critical process in burn scars (Chen et al., 2013).

Given the current state of debris flow hazard assessment and prediction in previously burned terrains, in addition to the growing influence of anthropogenic climate change on wildfire and extreme precipitation, development of physics-based hazard assessment tools that can be used in both hindcast investigations and forecasting applications is needed. Furthermore, due to the diverse morphology of precipitation events and their interaction with geographically distributed wildfire burn scars, development of tools that can assess hazards over regional domains, particularly in operational forecasting applications, is critical. Here, to advance the field of burn scar debris flow hazard assessment, we explore the use of the physics-based and fully-distributed Weather Research and Forecasting Hydrological modeling system version 5.1.1 (WRF-Hydro). WRF-Hydro is an open-source community model developed by the National Center for Atmospheric Research (NCAR). It is the core of National Oceanic and Atmospheric Administration’s (NOAA) National Water Model forecasting system, and has been used extensively to study channelized flows (e.g., Lahmers et al., 2020; Wang et al., 2019). Here, we modify WRF-Hydro to output high temporal resolution fine-scale (100 m) debris flow-relevant overland flow, a process computed using a fully unsteady, explicit, finite difference diffusive wave formulation. Previous efforts, employing shallow water equations, diffusive, kinematic, and diffusive-kinematic wave models, have demonstrated that water-only models can provide critical insights on runoff-driven debris flow behavior (Arratano & Savage, 1994; McGuire & Youberg, 2020; Arratano & Franzi, 2010; Di Cristo et al., 2021), even in burned watersheds (Rengers et al., 2016).

To test and demonstrate the utility of WRF-Hydro in debris flow studies, we investigate the January 2021 debris flow events within the Dolan burn scar on the Big Sur coast of central California (Fig. 1a–b). We first identify multiple debris flow sites using optical and radar remote sensing data and field investigations. We then calibrate WRF-Hydro against ground-based soil moisture and streamflow observations and use it to study the effects of burn scars on debris flow hydrology and changes in hazard potential. The paper is organized as follows. Section 2 describes our debris flow identification approach and historical context. Section 3 presents a description of
Section 4 describes the simulation, calibration, and validation of WRF-Hydro. Section 5 presents the results. Section 6 discusses the results and Section 7 provides a conclusion.

**Fig. 1** | WRF-Hydro model domain and Dolan burn scar. (a) WRF-Hydro model domain depicting topography, 2020 wildfire season burn scars, and PSL soil moisture and USGS stream gage observing sites. The black rectangle outlines (b) the Dolan burn scar inset, in which debris flow locations and major streams are marked and labeled.

---

**2 Study domain and debris flow identification methodology**

The Dolan wildfire burned from August 18th till December 31st, 2020. 55% of areas within the fire perimeter were burned at moderate-to-high severity (Burned Area Emergency Response, 2020). After the fire, USGS Emergency Assessment of Postfire Debris-flow Hazards produced a debris flow hazard assessment using a design storm based statistical model (USGS, 2020). On January 27–29, 2021, an atmospheric river (AR) made landfall on the Big Sur coast, bringing more than 300 mm of rainfall to California’s Coast Ranges (Fig. 2), with a peak rainfall rate of 24 mm h\(^{-1}\). During the AR event, a section of California State Highway 1 (CA1) at Rat Creek was destroyed by a debris flow. CA1 was subsequently closed for three months and rebuilt at a cost of ~$11.5M (Los Angeles Times, 2021).
2.1 Debris flow identification from remote sensing and field work

In addition to the Rat Creek debris flow, which made national news (Los Angeles Times, 2021), we identified three other debris flows using a combination of field investigation, and open access satellite optical and synthetic aperture radar (SAR) images (Fig. 3 and Fig. B1). We examined relative differences in normalized difference vegetation index (rdNDVI) defined by (Scheip & Wegmann, 2021):

\[ \text{rdNDVI} = \frac{NDVI_{\text{post}} - NDVI_{\text{pre}}}{\sqrt{NDVI_{\text{pre}} + NDVI_{\text{post}}}} \times 100 \]  

where \( NDVI_{\text{pre}} \) and \( NDVI_{\text{post}} \) are the pre- and post-event normalized difference vegetation index (NDVI) images computed following:

\[ NDVI = \frac{NIR - R}{NIR + R} \]  

where \( NIR \) is the near-infrared response and \( R \) is the visible red response. rdNDVI was calculated from Sentinel-2 satellite data using the HazMapper v1.0 Google Earth Engine application (Scheip & Wegmann, 2021). HazMapper requires selection of an event date, pre-event window (months),
post-event window (months), max cloud cover (%) and slope threshold (°). These input
requirements filter the number of images used to calculate the rdNDVI. We set the event date to
27 January 2021 and used a 3 month pre- and post-event window with 0% max cloud cover and a
0° slope threshold to identify vegetation loss associated with the debris flows. We then created a
binary map to highlight debris flows (and other vegetation loss) pixels above a rdNDVI vegetation
loss threshold. We removed all pixels with rdNDVI > -10.

Lastly, we searched for debris flows (and other ground surface deformation) by examining SAR
backscatter change with data acquired by the Copernicus Sentinel-1 (S1) satellites (see full
description in Handwerger et al., in review). We measured the change in SAR backscatter by using
the log ratio approach, defined as

\[ I_{\text{ratio}} = 10 \times \log_{10}(\frac{\sigma_{\text{pre}}}{\sigma_{\text{post}}}) \]  

(3)

where \( \sigma_{\text{pre}} \) is a pre-event image stack (defined as the temporal median) of SAR backscatter and
\( \sigma_{\text{post}} \) is a post-event image stack. Similar to the HazMapper method, our approach requires
selection of an event date, pre-event window (months), post-event window (months) and slope
threshold (°). No cloud-cover threshold is needed since SAR penetrates clouds. We used a 3 month
pre- and post-event window and 0° slope threshold to identify ground surface changes associated
with the debris flows. We then created a binary map to highlight debris flows by removing all
pixels with \( I_{\text{ratio}} < 99\text{th percentile value.} \)

Identified debris flow source areas and deposition sites were confirmed by field investigation (N.J.
Finnegan) and named after the locations where they deposited (i.e., Big Creek, Mill Creek, and
Nacimiento). We note that there were likely more debris flows triggered during the AR event.
However, given the primary goal of this study – to demonstrate the utility of WRF-Hydro – a
comprehensive cataloging of debris flows is beyond this study’s scope.
Fig. 3| Identified debris flow sites using rdNDVI vegetation change within the Dolan burn scar.

We convert the rdNDVI data into a binary map by setting a threshold value, which yield only the likely debris flow locations. (a)–(d) Sentinel-2 rdNDVI vegetation change at (a) Rat Creek, (b) Mill Creek, (c) Big Creek, and (d) the Nacimiento River.

3 WRF-Hydro

3.1 Model description

WRF-Hydro is an open-source physics-based community model that simulates land surface hydrologic processes. It includes the Noah-Multiparameterization (Noah-MP) land surface model (LSM; Niu et al., 2011), terrain routing module, channel routing module, and a conceptual baseflow bucket model. The Noah-MP LSM is a 1-dimensional column model that calculates vertical energy fluxes (i.e., sensible and latent heat, net radiation), moisture (i.e., canopy interception, infiltration, infiltration excess, deep percolation), and soil thermal and moisture states on the LSM grid (1 km in our application). The infiltration excess, ponded water depth, and soil moisture are then disaggregated using a time-step weighted method (Gochis & Chen, 2003) and sent to the terrain routing module which simulates subsurface and overland flows on a finer terrain routing grid (100 m in our application). According to the mass balance, local infiltration excess,
overland flow, and exfiltration from baseflow contribute to the surface head which flows into river channels if defined retention depth is exceeded. The channel routing module then calculates channelized flows assuming a trapezoidal channel shape (Fig. B2). Parameters related to the trapezoidal channel, such as channel bottom width ($B_w$), Manning’s roughness coefficient ($n$), and channel side slope ($z$) are functions of channel stream order (Fig. B3 and Table B1). Computed streamflow is then output on the 100 m grid. Equations used to compute infiltration excess, overland flow, and channelized flow are provided in Sect. 3.3 and 3.4.

By default, WRF-Hydro uses Moderate Resolution Imaging Spectroradiometer (MODIS) Modified International Geosphere-Biosphere Program (IGBP) 20-category land cover product as land cover (Fig. B4) and 1-km Natural Resources Conservation Service State Soil Geographic (STATSGO) database for soil type classification (Fig. B5; Miller & White, 1998). Land surface properties including canopy height (HVT), maximum carboxylation rate (VCMX25), and overland flow roughness (OV_ROUGH2D) are functions of land cover type (Table B2 & Fig. B4). Default soil hydraulic parameters in WRF-Hydro (i.e., soil porosity, grain size distribution index, and saturated hydraulic conductivity) are based on Cosby et al.’s (1984) soil analysis (Table B3) and are used to map onto the STATSGO 16 soil texture types (Fig. B5).

3.2 Meteorological forcing files
WRF-Hydro is used in standalone mode (i.e., it is not interactively coupled with the atmospheric component of WRF), but rather is forced with a combination of Phase 2 North American Land Data Assimilation System (NLDAS-2) meteorological data and Multi-Radar/Multi-Sensor System (MRMS) radar-only quantitative precipitation (Zhang et al., 2011, 2014, 2016). A description of the MRMS dataset and uncertainties therein can be found in Appendix A. NLDAS-2 provides hourly forcing data including incoming shortwave and longwave radiation, 2-m specific humidity and air temperature, surface pressure, and 10-m wind speed at 1/8-degree spatial resolution. MRMS provides hourly precipitation rate at 1-km resolution.

3.3 Overland flow routing and output
The Noah-MP LSM calculates rate of infiltration excess following Chen & Dudhia (2001):

$$
\frac{\partial h}{\partial t} = \frac{\partial P_d}{\partial t} \left( 1 - \frac{\sum_{i=1}^{n} \Delta D_i (\theta_i - \theta_0)}{P_d + \sum_{i=1}^{n} \Delta D_i (\theta_i - \theta_0)} \left[ \frac{1 - \exp \left( -k \frac{K_s}{K_{ref}} \delta_i \right)}{1 - \exp \left( -k \frac{K_s}{K_{ref}} \delta_0 \right)} \right] \right)
$$

(4)

where $h$ (m) is the surface water depth and $t$ is the time. $P_d$ (m) is the precipitation not intercepted by the canopy; $\Delta D_i$ (m) is the depth of soil layer $i$; $\theta_i$ is the soil moisture in soil layer $i$; $\theta_0$ is the
soil porosity; \( K_s \) (m s\(^{-1}\)) is the saturated hydraulic conductivity; \( K_{ref} \) is \( 2 \times 10^{-6} \) m s\(^{-1}\) which represents the saturated hydraulic conductivity of the silty–clay–loam soil texture chosen as a reference; \( \delta_t \) (s) is the model time step; and \( k \) which is equal to 3.0 is the runoff–infiltration partitioning parameter [the same as \( k d t_{ref} \) in Chen & Dudhia (2001)].

Noah-MP passes excess water to the terrain routing module, which simulates overland flow using a 2-dimensional fully-unsteady, explicit, finite-difference diffusive wave equation adapted from Julien et al. (1995) and Ogden (1997). It is considered superior to the traditionally used kinematic wave formulation in that it accounts for backwater effects and flow over adverse slopes. The diffusive wave formulation is the simplified form of the Saint Venant equations, i.e., continuity and momentum equations for a shallow water wave. The 2-dimensional continuity equation for a flood wave is:

\[
\frac{\partial h}{\partial t} + \frac{\partial q_x}{\partial x} + \frac{\partial q_y}{\partial y} = i_e
\]  

(5)

where \( h \) is the surface flow depth, \( q_x \) and \( q_y \) are the unit discharges in the x- and y-directions, respectively, and \( i_e \) is the infiltration excess. Manning’s equation which considers momentum loss is used to calculate \( q \). In the x-direction:

\[
q_x = \alpha_x h^{\beta}
\]  

(6)

where \( \beta \) is a unit dependent coefficient equal to \( \frac{5}{3} \), and

\[
\alpha_x = \frac{s_{fx}^{1/2}}{n_{ox}}
\]  

(7)

where \( n_{ox} \) is the tunable overland flow roughness coefficient. The momentum equation in the x-direction is given by:

\[
S_{fx} = S_{ox} - \frac{\partial h}{\partial x}
\]  

(8)

where \( S_{fx} \) is the friction slope, \( S_{ox} \) is the terrain slope, and \( \frac{\partial h}{\partial x} \) is the change in surface flow depth in the x-direction.

Off-the-shelf, WRF-Hydro does not output overland flow at terrain routing grids (100 m), however it is computed in the background to determine channelized streamflow. One key advance made in this work is that we modified WRF-Hydro source code to output overland flow. Overland flow depth (m) was converted to overland discharge (m\(^3\) s\(^{-1}\)) by multiplying flow depth by grid cell area (10,000 m\(^2\)) and dividing by the LSM time step (1 h).
3.4 Channel routing

If overland flow intersects grid cells identified as channel grids (2nd Strahler stream order and above; pre-defined by the hydrologically conditioned USGS National Elevation Dataset 30-m digital elevation model (DEM)), the channel routing module routes the water as channelized streamflow using a 1-dimensional, explicit, variable time-stepping diffusive wave formulation. Similarly, the continuity equation for channel routing is given as:

$$\frac{\partial A}{\partial t} + \frac{\partial q}{\partial s} = q_t$$

(9)

and the momentum equation is given as:

$$\frac{\partial q}{\partial t} + \frac{\partial \left( \frac{q^2}{A} \right)}{\partial s} + gA \frac{\partial H}{\partial s} = -gAS_f$$

(10)

where $s$ is the streamwise coordinate, $H$ is water surface elevation, $A$ is the flow cross-sectional area calculated as $(B_w + H z)H$ (Fig. B2), $q_t$ is the lateral inflow rate into the channel grid, $Q$ is the flow rate, $\gamma$ is a momentum correction factor, $g$ is acceleration due to gravity, and $S_f$ is the friction slope computed as:

$$S_f = \left( \frac{Q}{K} \right)^2$$

(11)

where $K$ is the conveyance computed from the Manning’s equation:

$$K = \frac{C_m n}{R^{2/3}}$$

(12)

where $n$ is the Manning’s roughness coefficient, $A$ is the channel cross-sectional area, $R$ is the hydraulic radius ($A/P$), $P$ is the wetted perimeter, and $C_m$ is a dimensional constant (1.486 for English units or 1.0 for SI units).

4 Model simulation, calibration, and validation

4.1 Model domain

The WRF-Hydro model domain spans regions in California including the Coast Ranges, Monterey Bay, and the Central Valley, and covers several burn scars from the 2020 wildfire season (Fig. 1a). Here we focus our analysis on the Dolan burn scar where the hazardous debris flows occurred (Fig. 1b). According to the USGS 30-m DEM, the Rat Creek debris flow site sits at the base of a 1st order catchment with a drainage area of 2.23 km$^2$. Mill Creek, Big Creek, and Nacimiento debris flows were initiated within extremely steep, intensely burned, 1st order catchments, but were deposited in 2nd, 3rd, and 3rd Strahler stream order channels, respectively.

To calibrate and validate WRF-Hydro output, we use soil moisture observations from two Physical Sciences Laboratory (PSL) monitoring stations [i.e., Lockwood (lwd) and Gilroy (gry)] (Fig. 1a).
Due to the Mediterranean climate of California, many USGS stream gages experience low or no flow during the dry season. In addition, many gages are under manual regulation to mitigate wet-season flood risks and better distribute water resources. As such, it can be challenging to obtain natural streamflow observations for model calibration. Here, three USGS stream gages [i.e., Arroyo Seco NR Greenfield, CA (ID 11151870), Arroyo Seco NR Soledad, CA (ID 11152000), and Arroyo Seco BL Reliz C NR Soledad, CA (ID 11152050)] (Fig. 1a) on streams that have measurable flows during our study period and are free of human regulation are used. These gages are located downstream of the Dolan burn scar and hence are useful in calibrating the parameters associated with burn scar effects. The PSL soil moisture observations were recorded at 2-minute intervals and USGS streamflow gage data were recorded at 15-minute intervals, but we perform all observation-model comparisons at hourly-mean resolution.

### 4.2 Baseline simulation and soil moisture calibration

WRF-Hydro was run from January 1–31 of 2021. We performed the baseline simulation by modifying WRF-Hydro default parameters (Table B3) based on a calibration using soil moisture observations from stations lwd and gry. Neither PSL station is located in a burn scar. Since the baseline simulation includes no postfire characteristics, it can also be regarded as the “pre-fire” scenario. Soil moisture at 10 cm below ground in the baseline simulation was calibrated by performing a domain-wide adjustment of soil porosity and grain size distribution index at the simulation start (Table B3). We then allowed the model to spin up from January 1–10 before using January 11–31 for validation. Using a relatively short spin-up period is justified because prior to the AR event, little rain fell on the Dolan burn scar (i.e., ~400 mm of rainfall fell from June to December 2020). As such, in the months preceding the debris flow events, soil moisture observations indicate already dry condition prior to our 10 day spin up.

After calibration, the simulated soil moisture closely mimics ground-based PSL observations (Fig. 4). Both the observed magnitude and variability are well captured, with the simulated ±1 standard deviation envelope largely encompassing PSL observations during the AR. Model performance was evaluated using four quantitative metrics, i.e., correlation coefficient, root mean square error, mean bias, and Kling-Gupta efficiency (KGE; Gupta et al., 2009; Kling et al., 2012). KGE has previously been used in soil moisture calibration applications (e.g., Lahmers et al., 2019; Vergopolan et al., 2020) and is computed as follows:

$$KGE = 1 - \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad (13)$$

where $r$ is the correlation coefficient between the observation and simulation, $\alpha$ is the ratio of the standard deviation of simulation to the standard deviation of observation, and $\beta$ is the ratio of the
mean of simulation to the mean of observation. KGEs close to 1 indicate a high-level consistency
between the simulation and observation, while negative KGEs indicate poor model performance
(Schönfelder et al., 2017; Andersson et al., 2017).

The model’s ability to simulate soil moisture substantially improves after calibration (Fig. 4; Table
1). KGE values approach 1 (0.72 at lwd and 0.88 at gry), indicating that WRF-Hydro adequately
simulates the hydrologic environment and its response to meteorological change.

Fig. 4| Precipitation, observed and simulated soil moisture at two PSL soil moisture stations.
January 11–31, 2021 MRMS precipitation (green bars) and observed (black line) and simulated
volumetric soil moisture 10 cm below ground in the baseline simulation (purple dashed line) at
PSL sites (a) Lockwood (lwd) and (b) Gilroy (gry). Envelope of purple shading depicts ±1 standard
deviation of model simulated soil moisture. KGE scores are provided at top left for each station.
Table 1 | Quantitative evaluation metrics for the simulated soil moisture and streamflow when compared against observations. The metrics include the Pearson correlation coefficient \((r)\), root mean square error (RMSE), and mean bias (Bias). In addition, the comprehensive metrics Kling-Gupta efficiency (KGE) and Nash-Sutcliffe efficiency (NSE) are used to evaluate model-simulated soil moisture and streamflow, respectively. For soil moisture, the numbers in front of “/” are calculated between the default run (i.e., uncalibrated run) and the observations, whereas the numbers following “/” are the corresponding values in the baseline simulation (the purple dashed line in Fig. 4). For streamflow, the numbers in front of “/” are computed between the baseline run (purple dashed line in Fig. 6) and the observations, while the numbers behind “/” are for burn scar simulation (red line in Fig. 6). If the model performance regarding a certain metric is enhanced in the burn scar simulation, the number after “/” is underlined.

### Soil moisture (Default / Baseline)

<table>
<thead>
<tr>
<th>Station</th>
<th>( r )</th>
<th>RMSE</th>
<th>Bias</th>
<th>KGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>lwd</td>
<td>0.97 / 0.98</td>
<td>7.06 / 4.32</td>
<td>5.21 / 4.16</td>
<td>0.10 / 0.72</td>
</tr>
<tr>
<td>gry</td>
<td>0.94 / 0.94</td>
<td>5.19 / 2.53</td>
<td>-4.79 / -1.66</td>
<td>0.80 / 0.88</td>
</tr>
</tbody>
</table>

### Streamflow (Baseline / Burn scar)

<table>
<thead>
<tr>
<th>Station</th>
<th>( r )</th>
<th>RMSE</th>
<th>Bias</th>
<th>NSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1870</td>
<td>0.28 / 0.93</td>
<td>39.29 / 14.69</td>
<td>1.65 / 3.36</td>
<td>-0.17 / 0.73</td>
</tr>
<tr>
<td>2000</td>
<td>0.26 / 0.86</td>
<td>51.22 / 24.92</td>
<td>2.47 / 4.81</td>
<td>-0.15 / 0.73</td>
</tr>
<tr>
<td>2050</td>
<td>0.25 / 0.81</td>
<td>49.96 / 27.43</td>
<td>5.70 / 8.24</td>
<td>-0.38 / 0.53</td>
</tr>
</tbody>
</table>

4.3 Burn scar simulation and streamflow calibration

To simulate effects of wildfire burn scars on hydrologic processes and debris flow hazards, we made two modifications to the baseline simulation soil moisture calibrated model configuration. First, we changed the land cover type within the burn scar perimeter to its nearest LSM analogue, i.e., “barren and sparsely vegetated”. The switch to barren land causes: (1) height of the canopy (HVT) to decrease to 0.5 m; (2) maximum rate of carboxylation at 25°C (VCMX25) to decrease to 0 \( \mu \text{mol CO}_2/(\text{m}^2 \cdot \text{s}) \); and (3) overland flow roughness coefficient (OV_ROUGH2D) to decrease to 0.035 (Fig. 5a–c) from default values (Fig. B4 and Table B2).

The second adjustment was to decrease soil infiltration rates within the burn scar perimeter, achieved by reducing soil saturated hydraulic conductivity (DKSAT; Fig. 5d; Scott & van Wyk, 14

Table 1  Evaluation metrics of simulated soil moisture and streamflow

<table>
<thead>
<tr>
<th>Soil moisture (Default / Baseline)</th>
<th>Station</th>
<th>( r )</th>
<th>RMSE</th>
<th>Bias</th>
<th>KGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>lwd</td>
<td>0.97 / 0.98</td>
<td>7.06 / 4.32</td>
<td>5.21 / 4.16</td>
<td>0.10 / 0.72</td>
<td></td>
</tr>
<tr>
<td>gry</td>
<td>0.94 / 0.94</td>
<td>5.19 / 2.53</td>
<td>-4.79 / -1.66</td>
<td>0.80 / 0.88</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Streamflow (Baseline / Burn scar)</th>
<th>Station</th>
<th>( r )</th>
<th>RMSE</th>
<th>Bias</th>
<th>NSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1870</td>
<td>0.28 / 0.93</td>
<td>39.29 / 14.69</td>
<td>1.65 / 3.36</td>
<td>-0.17 / 0.73</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>0.26 / 0.86</td>
<td>51.22 / 24.92</td>
<td>2.47 / 4.81</td>
<td>-0.15 / 0.73</td>
<td></td>
</tr>
<tr>
<td>2050</td>
<td>0.25 / 0.81</td>
<td>49.96 / 27.43</td>
<td>5.70 / 8.24</td>
<td>-0.38 / 0.53</td>
<td></td>
</tr>
</tbody>
</table>
1990; Cerdà, 1998; Robichaud, 2000; Martin & Moody, 2001) from default values (Table B3).

Consistent with the hydrophobicity of burned soils, we calibrate the burn scar simulation by systematically exploring a range of burn scar area saturated hydraulic conductivities (0 to 3×10^{-7} m s^{-1} with a 5×10^{-8} m s^{-1} increment), with the goal of reproducing streamflow behavior similar to USGS gage observations. We found that a value of 1.5×10^{-7} m s^{-1} gives the highest Nash-Sutcliffe efficiency (NSE; Nash & Sutcliffe, 1970) across all three USGS stream gages (Table 1). NSE and KGE are the two most widely used metrics for calibration and evaluation of hydrologic models. The NSE has previously been used in streamflow calibration applications (e.g., Xia et al., 2012; Bitew & Gebremichael, 2011), and it is calculated as follows:

\[ NSE = 1 - \frac{\sum_{t=1}^{T}(Q_{\text{sim}}(t) - Q_{\text{obs}}(t))^2}{\sum_{t=1}^{T}(Q_{\text{obs}}(t) - \bar{Q}_{\text{obs}})^2} \]

where \(T\) is the length of the time series, \(Q_{\text{sim}}(t)\) and \(Q_{\text{obs}}(t)\) are the simulated and observed discharge at time \(t\), respectively, and \(\bar{Q}_{\text{obs}}\) is the mean observed discharge. By definition, NSEs of 1 indicate perfect correspondence between the simulated and observed streamflow. Positive NSEs mean that the model streamflow has a greater explanatory power than the mean of the observations, whereas negative NSEs represent poor model performance (e.g., Moriasi et al., 2007; Schaefli & Gupta, 2007). When burn scar characteristics are included, NSEs increase from negative values in the baseline to greater than 0.5, and the NSEs at gages 1870 and 2000 reach 0.84 and 0.73, respectively. Higher NSE scores indicate the abovementioned burn scar parameter changes improve the model’s ability to simulate streamflow observations downstream of the burn scar (Table 1).
Fig. 5 | Parameter setting in the WRF-Hydro burn scar simulation. (a) The height of the canopy (HVT; m; shading), (b) maximum rate of carboxylation at 25°C (VCMX25; $\mu$mol CO$_2$/($m^2 \cdot s$); shading), (c) overland flow roughness coefficient (OV_ROUGH2D; shading), and (d) saturated hydraulic conductivity (DKSAT; m s$^{-1}$; shading) in the burn scar simulation.
5 Results

5.1 Hydrologic response due to burn scar incorporation

The pre-fire baseline simulation fails to capture the hydrologic behavior observed at the USGS gages located within the burn scar (Fig. 6). Incorporation of burn scar characteristics substantially alters the hydrologic response of the model and provides much higher fidelity streamflow.
simulations (Fig. 6). Observed hydrographs are characterized by two early streamflow peaks related to two precipitation bursts on January 27th and 28th. Our burn scar simulation captures this behavior, while the baseline simulation streamflow peaks just once, with a lower magnitude and an ~3-day lag after peak precipitation (Fig. 6). The steep rising limbs and high magnitude discharge peaks of the burn scar hydrograph are indicative of flash flooding. Compared with the pre-fire baseline scenario, the burn scar’s barren land and low infiltration rate substantially accelerate drainage rates and increase discharge volume into stream channels.

5.2 Hydrologic response at four debris flow sites

We identified locations and extent of four debris flows from remote sensing data and field work (Fig. 3 & Fig. B1). rdNDVI shows vegetation loss caused by debris flows. Mill Creek, Big Creek, and Nacimiento were relatively large debris flows with runout lengths between ~2–5 km. Rat Creek occurred in a smaller catchment and had a runout length of ~300 m. The difference in runout length and debris flow size is primarily controlled by upstream catchment size. Due to its low stream order (1st Strahler stream order), Rat Creek is the only debris flow site modeled entirely as overland flow in our WRF-Hydro simulations.

At the four debris flow sites, we use three metrics to characterize hydrologic anomalies: (1) accumulated runoff volume, (2) peak discharge, and (3) time to peak discharge. Fig. 7 depicts accumulated channelized discharge volume (blue shading) and accumulated overland discharge volume (yellow-red shading) from January 27th 00:00 to 28th 12:00 near the four debris flow sites in the burn scar simulation. Accumulation time period is chosen such that it covers the first two runoff surges in the simulated hydrographs which are likely associated with debris flows (Fig. 8) given that nearly concurrent peak rainfall intensity and peak discharge is a signature characteristic of debris flows (Kean et al., 2011). Runoff volume is on the order of $10^4$ m$^3$ at Rat Creek and $10^6$ m$^3$ at the other three sites.
Simulated overland flow and streamflow in burn scar simulation

Fig. 7| WRF-Hydro simulated overland flow and streamflow in the burn scar simulation. (a)–(d) Total volume of accumulated overland flow (yellow-red shading) and streamflow (blue shading) on $\log_{10}$ scale between January 27th 00:00 and 28th 12:00 at four debris flow sites. Black rectangles correspond to domains in Fig. 3a–d. Black circles and triangles indicate debris flow source areas and deposits, respectively.

Dramatic hydrographic changes after inclusion of burn scar characteristics are simulated at debris flow source areas (Fig. B6 and Table B4) and deposition sites (Fig. 8 and Table 2). WRF-Hydro facilitates investigation of the hydrologic response at triggering and deposition locations and along the runout path. Here, to emphasize the downstream hazards, our analysis is focused on debris flow deposits. At Rat Creek, where a section of CA1 collapsed, the magnitude of discharge substantially increases, and overland flow surges are concurrent with rainfall bursts (Fig. 8a). Total discharge accumulated during the AR event increases approximately eight-fold (791%), and peak
discharge more than triples compared to the baseline simulation (Fig. 8a and Table 2). At Mill Creek, Big Creek, and Nacimiento, baseline hydrographs are characterized by less variability, muted responses to two early precipitation bursts, and a delayed third discharge peak that does not occur until ~3 days after AR passage (Fig. 8b–d). Maximum discharge peaks in the baseline hydrographs lag those in the burn scar simulation by ~2 days (Fig. 8b–d; Table 2). In the burn scar simulation, total volume substantially increases at the three channelized sites – total volume increases ~650% at Mill Creek, ~891% at Big Creek, and ~829% at Nacimiento (Fig. 8b–d and Table 2), and the absolute increase in volume is on the order of $10^6$ m$^3$ (Table 2). Peak discharge more than triples at Mill Creek and Big Creek and more than quadruples at Nacimiento. Additionally, response times of the peak in discharge to the peak in precipitation decrease to less than an hour, highlighting the simulated flashiness of the burned catchments.

Fig. 8| WRF-Hydro simulated discharge time-series at four debris flow deposition locations. (a)–(d) MRMS precipitation (green bars) and simulated discharge time-series for January 26$^{th}$ 00:00 to 31$^{st}$ 23:00 at (a) Rat Creek, (b) Mill Creek, (c) Big Creek, and (d) Nacimiento deposition locations (black triangles in Fig. 7a–d) in baseline simulation (purple dashed line) and burn scar simulation (red line).
Table 2

The total runoff volume, peak discharge, and peak timing at debris-flow deposits

<table>
<thead>
<tr>
<th>Site name</th>
<th>Baseline simulation</th>
<th>Burn scar simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total volume (m³)</td>
<td>Peak discharge (m³ s⁻¹)</td>
</tr>
<tr>
<td>Rat Creek</td>
<td>6,897</td>
<td>0.54</td>
</tr>
<tr>
<td>Mill Creek</td>
<td>312,925</td>
<td>13.10</td>
</tr>
<tr>
<td>Big Creek</td>
<td>842,808</td>
<td>46.10</td>
</tr>
<tr>
<td>Nacimiento</td>
<td>743,531</td>
<td>33.15</td>
</tr>
</tbody>
</table>

5.3 Debris flow hazard assessment for the Dolan burn scar

Since high magnitude runoff is often the cause and precursor of runoff-generated debris flows in burned areas (Cannon et al., 2003, 2008; Rengers et al., 2016), we use simulated accumulated volume of overland flow and streamflow to assess runoff-generated debris flow hazard potential under pre-fire (i.e., baseline; Fig. 9a&d) and postfire (i.e., burn scar simulation; Fig. 9b&e) conditions. We assess changes at both stream and catchment levels and use the difference between burn scar and baseline simulations to assess added debris flow hazard potential (Fig. 9c&f). Consistent with the increasing erosive and entrainment power associated with increasing discharge, our debris flow hazard increases as the accumulated discharge volume increases. To reduce the effects of catchment size on the volume-based hazard levels, we normalize a catchment’s discharge volume by the area of the catchment (Santi et al., 2012; Fig. 9d–f). Non-normalized catchment hazard maps are also provided (Fig. B7).

In the pre-fire baseline simulation, the AR-induced precipitation produces lower debris flow hazard over most of the domain, but elevated hazards along stream channels (Fig. 9a). We note no substantial differences between areas in or out of the burn scar. In the burn scar simulation, debris
flow hazard levels increase across the Dolan burn scar and along channels outside but downstream
of the burn scar (Fig. 9b–c). The discharge volume increases by an order of magnitude near Rat
Creek, Big Creek, Mill Creek, and Nacimiento. Within the burn scar, hazards along major stream
channels, such as the Nacimiento River and San Antonio River increase. Outside the burn scar,
hazard levels along river channels downstream of the burn scar, such as the Arroyo Seco River,
also increase (Fig. 9c).

At the catchment level, debris flow hazards are assessed using accumulated discharge volumes
normalized by catchment areas (Fig. 9d–f). Accumulated discharge volumes are assessed at the
outlet of each catchment between January 27th 00:00 to 28th 12:00. In the baseline simulation, the
majority of catchments are subject to relatively low debris flow hazards compared to the burn scar
simulation with total normalized discharge volume less than 10^3 m^3 km^-2 (Fig. 9d). In the burn scar
simulation, over half of catchments within the Dolan burn scar have normalized discharge volume
greater than 10^5 m^3 km^-2, while over 1/4 of basins reach 10^6 m^3 km^-2 (Fig. 9e). The additional
debris flow hazard brought about by the inclusion of wildfire burn scar characteristics is substantial
(Fig. 9f).

To summarize changes in debris flow hazards as a result of including burn scar characteristics in
WRF-Hydro simulations, we create distributions of pre-fire baseline and burn scar catchment-area
normalized discharge from the 404 catchments located within the Dolan burn scar perimeter (Fig.
10). After incorporating burn scar characteristics, the full distribution shifts to the right, indicating
increased hazard levels – a shift considered robust by a Student’s t-test (p value: 4.6E-45). A
quantitative assessment of this shift indicates that the mean catchment area normalized discharge
volume increases by ~1300% (from ~380k to 5.5M m^3 km^-2) while the standard deviation increases
~1400% (from ~1.6M to 23.0M m^3 km^-2). We also assess shifts at a range of distribution
percentiles: 5P: 148% (~0.6k to ~1.5k m^3 km^-2), 25P: 725% (~3.7k to ~30.7k m^3 km^-2), 50P: 924%
(~13k to ~135k m^3 km^-2), 75P: 980% (~120k to ~1.3M m^3 km^-2), and 95P: 1300% (~2.1M to
~29.1M m^3 km^-2). In the burn scar simulation, more than half of catchments have normalized
volumes > 10^5 m^3 km^-2 and more than 1/4 of catchments have volumes > 10^6 m^3 km^-2 – values that
correspond to the 75P and 90P of the baseline simulation, respectively. Disproportionate shifting
of the right tail of the distribution suggests that extreme debris flow hazards increase non-linearly
under simulated burn scar conditions.

Our catchment-area normalized discharge volume-based hazard assessment also indicates that the
catchments containing Mill Creek, Big Creek, and Nacimiento had elevated hazard potential (Fig.
9d–f), consistent with our (limited) debris flow observations. Other areas with elevated hazards
include catchments containing the Arroyo Seco and San Antonio Rivers. Beyond the burn scar
perimeter, effects of fire expand to adjacent and downstream catchments, and the drainage basins
of the Arroyo Seco and Nacimiento Rivers are simulated to have potentially hazardous conditions,
i.e., normalized discharge volumes in excess of 10^6 m^3 km^-2 (Fig. 9e&f).
Fig. 9 | Discharge volume-based runoff-generated debris flow hazards. Debris flow hazards at individual stream level for the (a) baseline, (b) burn scar, and (c) difference between burn scar and baseline simulations. Hazard is estimated as total discharge volume from January 27th 00:00 to 28th 12:00. (d)–(f) Normalized debris flow hazards by catchment area at catchment level. For each catchment, the hazard is determined by total discharge volume at the catchment outlet from January 27th 00:00 to 28th 12:00 divided by catchment area.
Fig. 10 | Distributions of accumulated discharge volumes at the outlet of the 404 catchments normalized by upstream catchment areas within Dolan burn scar in the baseline simulation (purple line) and in the burn scar simulation (red line). Dashed vertical lines indicate median values.

5.4 Debris flow hazard assessment at regional scales

While the results we present above primarily focus on hazards in the Dolan burn scar, our WRF-Hydro domain includes a number of additional 2020 wildfire burn scar sites (Fig. 1a). Given the long filament-like structure of western U.S. landfalling ARs, the heterogeneous nature of landfalling trajectories, and the potential for systems to interact with diverse topographic terrains, the development of tools capable of regional hazard assessments under high-gradient precipitation events is crucial – particularly in a wildfire-prone region like California. To demonstrate the potential utility of WRF-Hydro in regional applications, we assess hazards over our full domain (Fig. 11). We find that hazard potential, from both channelized and overland flows, is greatest within the burn scar sites, with maximum hazards found in the Dolan burn scar, consistent with the location of elevated precipitation along the Coast Ranges – where more than 300 mm of rain fell over three days (Fig. 11). Other high hazard-elevated precipitation regions within our domain include the western edge of the Sierra Nevada and areas north of Monterey Bay, which collocate with the Mineral and Del Puerto burn scars, respectively. Similar to our Dolan burn scar focused analysis, areas within and downstream of these burn scar sites have elevated streamflow discharge volumes compared to the non-burned areas (Fig. 11b). Likewise, areas of heightened accumulated
overland flow are elevated in burn scar regions, but also demonstrate a strong correspondence to
the spatial distribution of precipitation (Fig. 11a & c).

![Fig. 11](image)

**Fig. 11** MRMS accumulated precipitation and regional debris flow hazard assessment. (a) MRMS accumulated precipitation during January 27th 00:00 to 29th 23:00 over the model domain (shading; mm). Names of burn scars are labeled in black. (b) Accumulated streamflow (yellow-to-red shading; m$^3$) and (c) accumulated overland flow from 27th 00:00 to 28th 12:00 over the model domain (yellow-to-red shading; m$^3$). Wildfire perimeters of 2020 wildfire season are outlined in black in (a), and in blue in (b) and (c). The coastline of California is in grey.

### 6 Discussion

Given the historic and growing frequency of wildfires in the western U.S. (Swain 2021; Williams et al., 2019; Goss et al., 2020) and globally (Jolly et al., 2015; Flannigan et al., 2013), developing tools to investigate, better understand, and potentially predict changes in burn scar hydrology and natural hazards at regional scales is critical. Here, we demonstrate the first use of WRF-Hydro to simulate the surface hydrologic response over a burn scar during a landfalling AR. We augmented the default version of WRF-Hydro to output overland flow and to replicate burn scar behavior by adjusting vegetation type and infiltration rate parameters. WRF-Hydro simulations were validated against PSL soil moisture and USGS streamflow observations before we used simulated streamflow and overland flow volumes to characterize debris flow hazard potential.

A comparison between baseline and burn scar simulations demonstrated that changes in hydraulic properties of burned areas causes drastic changes in surface flows, including faster discharge response times, greater discharge volumes, and overall flashier hydrologic behavior in surface flows. As a result of including burn scar characteristics in WRF-Hydro simulations, median catchment-area normalized discharge volume increases nine-fold, while 95P volume increases 13-
fold. The magnitude of our simulated changes is consistent with findings from previous postfire hydrology studies (Anderson et al., 1976; Scott, 1993; Meixner & Wohlgeneth, 2003; Kinoshita & Hogue, 2015; Kean et al., 2011). At Rat Creek, where a debris flow destroyed CA1, our model simulation predicted an eight-fold increase in accumulated overland flow and a tripling in peak discharge when compared to the baseline simulation. At Mill Creek, Big Creek, and Nacimiento, the increase of runoff volume from the baseline to the burn scar simulation is on the order of 10⁶ m³. Our hazard assessments based on catchment-area normalized discharge volumes indicated that Mill Creek, Big Creek, and Nacimiento were under elevated debris flow hazards, corresponding well with identified debris flow occurrences.

Despite methodological differences, our debris flow hazard assessment for this AR event is generally consistent with the USGS’ postfire, pre-AR, design-storm-based preliminary hazard assessment (USGS, 2020). As described above, USGS preliminary hazard assessments use logistic regression models to estimate the likelihood of debris flow occurrence and multivariate linear regression models to estimate debris flow volumes. This empirical approach is trained on historical western U.S. debris flow occurrence and magnitude data and incorporates estimated burn scar soil erodibility and burn severity data (Cannon et al., 2010; Gartner et al., 2014; Staley et al., 2016). For precipitation, the USGS assessment utilizes a design storm approach that assumes 1–5 year return interval magnitude precipitation falls uniformly over a region/burn scar (USGS, 2020). For the Dolan burn scar, both assessments find that large stream channels had relatively higher hazard levels than small streams or overland areas. However, close comparison of hazard maps reveals differences in spatial distribution of high-hazard catchments. In the USGS assessment, higher hazard levels are predicted north and southeast of the burn scar, whereas in our assessment the highest hazards occur along major stream channels. We hypothesize that USGS-assessed areas of higher hazard potential are related to their use of design-storm precipitation (see Fig. 2 for MRMS precipitation footprint) and burn severity data (Burned Area Emergency Response, 2020). Comparison with the USGS assessment framework suggests room for improvement in WRF-Hydro-based assessments (i.e., inclusion of burn severity and soil erodibility data), but also highlights the potential utility of working with spatially-distributed and time-varying precipitation. However, this also means the accuracy of WRF-Hydro predictions depends on the accuracy of precipitation forcing, and in our hindcast application, MRMS precipitation data (Appendix A). Accordingly, our WRF-Hydro-based hazard assessment could benefit from precipitation products mosaiced from various sources to constrain precipitation-based uncertainties (e.g., gauge-corrected and/or Mountain Mapper MRMS), although the long processing time of these datasets inhibits timely post-event assessments.

As a water-only model, WRF-Hydro is currently restricted to simulating the hydrologic ingredients of debris flows. While water-only models have been widely used to investigate and better understand debris flow dynamics (Arattano & Savage, 1994; Arattano & Franzi, 2010; Rengers et al., 2016; McGuire & Youberg, 2020; Di Cristo et al., 2021), sediment supply, soil erodibility, and
other sedimentological factors also play important roles in determining the potential for and
severity of mass failure events (McGuire et al., 2017). Developing a debris flow model that couples
hydrologic and sediment erosion and transport processes would represent a significant advance
and be of great practical use (Banihabib et al., 2020; Shen et al., 2021). At a minimum, soil grain
size maps and domain-specific rainfall intensity-duration curves can provide guidance to define
transitions from water floods to debris flows if historical debris flow data is available in the study
domain (McGuire & Youberg, 2020; Tognacca et al., 2000; Gregoretti & Fontana, 2008; Cannon
et al., 2007).

7 Conclusion

Use of WRF-Hydro to simulate runoff-generated debris flow hazards in burn scar settings
represents a novel application. It is notable that in this application we have balanced the
computational cost of a regional domain with our choice of resolved spatial resolution for terrain
routing and overland flow calculations (100 m). However, WRF-Hydro has previously been
applied to smaller domains at higher terrain routing resolutions (~30 m). Future work could assess
the use of the model to study burn scar hydrology at finer spatial scales, should the application
warrant and should underlying data at sufficient resolution exist. Other potential applications of
our modified model framework include alpine areas and steep hillslopes with sparse vegetation
where runoff-generated debris flows dominate over landslide-initiated ones (Davies et al., 1992;
Coe et al., 2003, 2008).

Further, our burn scar parameter changes are performed to Noah-MP, which is the core land
surface component of the National Centers for Environmental Prediction Global Forecast System
(GFS) and Climate Forecast System (CFS), thus the findings presented herein, are likely to prove
useful in the broader worlds of forecast meteorology and climate science. In addition, here WRF-
Hydro is driven by historical precipitation and meteorological data, i.e., in hindcast mode. We see
no reason why this modeling framework could not also be employed to project hazards under
future climatic conditions (e.g., Huang et al., 2020), or given its relatively low computational
expense, in operational forecast mode. Indeed, modern ensemble-based meteorological forecasting
could provide high spatiotemporal forcing data with which disaster preparedness managers could
probabilistically assess debris flow hazard potential, and issue advanced life and property saving
warnings.
Appendix A

Text A1. Multi-Radar/Multi-Sensor System (MRMS) radar-only precipitation estimate and uncertainty

MRMS is a precipitation product that covers the contiguous United States (CONUS) on 1-km grids. It combines precipitation estimates from sensors and observational networks (Zhang et al., 2011, 2014, 2016), and is produced at the National Centers for Environmental Prediction (NCEP) and distributed to National Weather Service forecast offices and other agencies. Input datasets used to produce MRMS include the U.S. Weather Surveillance Radar-1988 Doppler (WSR-88D) network and Canadian radar network, Parameter-elevation Regressions on Independent Slopes Model (PRISM; Daly et al. 1994, 2017), Hydrometeorological Automated Data System (HADS) gauge data with quality control (Qi et al., 2016), and outputs from numerical weather prediction models. There are four different MRMS quantitative precipitation estimates (QPE) products incorporating different input data or combinations: radar only, gauge only, gauge-adjusted radar, and Mountain Mapper. For our study period (i.e., January 1–31, 2021), only the radar-only QPE is currently available.

We acknowledge that precipitation data has uncertainties. Use of different precipitation products may produce different results. A study comparing different gridded precipitation datasets including satellite-based precipitation data, gauge dataset, and multi-sensor products revealed large uncertainties in precipitation intensity (Bytheway et al., 2020). However, comparing different precipitation datasets to characterize uncertainties is beyond the scope of this study. MRMS provides gridded precipitation at high temporal (hourly) and spatial (1-km) resolutions, making it a useful tool to demonstrate the utility of WRF-Hydro in post-wildfire debris flow hazard assessments.
Appendix B

Fig. B1 Optical- and SAR-based remote sensing data of four debris flows. Optical data from Sentinel-2 show pre- and post-debris flow imagery in real color. rdNDVI calculated from the Sentinel-2 data show a decrease in vegetation corresponding to debris flow locations. Sentinel-1 backscatter change shows the change in ground surface properties determined by calculating the log ratio of pre- and post-event SAR images. The pre-event, post-event satellite images, Sentinel-1 Backscatter, and Sentinel-2 rdNDVI change at (a) Rat Creek, (b) Mill Creek, (c) Big Creek, and (d) Nacimiento.
Fig. B2 Schematic trapezoidal shape and related parameters of channels in WRF-Hydro. $B_w$ is the channel bottom width (m), $z$ is the channel side slope (m), and $H$ is water elevation (m). The cross-sectional area of flow is calculated as $(B_w + H z)H$.  

https://doi.org/10.5194/nhess-2021-345
Preprint. Discussion started: 3 December 2021
© Author(s) 2021. CC BY 4.0 License.
Table B1 Parameters of trapezoidal channels in WRF-Hydro.

<table>
<thead>
<tr>
<th>Stream order</th>
<th>Channel bottom width $B_w$ (m)</th>
<th>Channel side slope $z$ (m)</th>
<th>Manning’s roughness coefficient $n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.5</td>
<td>3</td>
<td>0.33</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1</td>
<td>0.21</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>0.5</td>
<td>0.09</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>0.18</td>
<td>0.06</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>0.05</td>
<td>0.04</td>
</tr>
<tr>
<td>6</td>
<td>40</td>
<td>0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>7</td>
<td>60</td>
<td>0.05</td>
<td>0.02</td>
</tr>
<tr>
<td>8</td>
<td>70</td>
<td>0.05</td>
<td>0.02</td>
</tr>
<tr>
<td>9</td>
<td>80</td>
<td>0.05</td>
<td>0.01</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>0.05</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Table B1 Parameters of the trapezoidal channels in WRF-Hydro including channel bottom width $B_w$ (m), channel side slope $z$ (m), and Manning’s roughness coefficient $n$. 
Fig. B3 (a) Stream order defined by the USGS 30-m DEM in our WRF-Hydro model domain and (b) the channel bottom width (m) which is a function of stream order (Table B1).
Table B2 MODIS IGBP 20-category land cover type and properties in Noah-MP LSM

<table>
<thead>
<tr>
<th>Land cover code</th>
<th>Land cover type</th>
<th>Canopy height (m)</th>
<th>Max carboxylation rate at 25°C ($\mu$mol CO$_2$/($m^2 \cdot s$))</th>
<th>Overland flow roughness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Evergreen Needleleaf Forest</td>
<td>20</td>
<td>50</td>
<td>0.2</td>
</tr>
<tr>
<td>2</td>
<td>Evergreen Broadleaf Forest</td>
<td>20</td>
<td>60</td>
<td>0.2</td>
</tr>
<tr>
<td>3</td>
<td>Deciduous Needleleaf Forest</td>
<td>18</td>
<td>60</td>
<td>0.2</td>
</tr>
<tr>
<td>4</td>
<td>Deciduous Broadleaf Forest</td>
<td>16</td>
<td>60</td>
<td>0.2</td>
</tr>
<tr>
<td>5</td>
<td>Mixed Forests</td>
<td>16</td>
<td>55</td>
<td>0.2</td>
</tr>
<tr>
<td>6</td>
<td>Closed Shrublands</td>
<td>1.1</td>
<td>40</td>
<td>0.055</td>
</tr>
<tr>
<td>7</td>
<td>Open Shrublands</td>
<td>1.1</td>
<td>40</td>
<td>0.055</td>
</tr>
<tr>
<td>8</td>
<td>Woody Savannas</td>
<td>13</td>
<td>40</td>
<td>0.055</td>
</tr>
<tr>
<td>9</td>
<td>Savannas</td>
<td>10</td>
<td>40</td>
<td>0.055</td>
</tr>
<tr>
<td>10</td>
<td>Grasslands</td>
<td>1</td>
<td>40</td>
<td>0.055</td>
</tr>
<tr>
<td>11</td>
<td>Permanent wetlands</td>
<td>5</td>
<td>50</td>
<td>0.07</td>
</tr>
<tr>
<td>12</td>
<td>Croplands</td>
<td>2</td>
<td>80</td>
<td>0.035</td>
</tr>
<tr>
<td>13</td>
<td>Urban and Built-Up</td>
<td>15</td>
<td>0</td>
<td>0.025</td>
</tr>
<tr>
<td>14</td>
<td>Cropland/natural vegetation mosaic</td>
<td>1.5</td>
<td>60</td>
<td>0.035</td>
</tr>
<tr>
<td>15</td>
<td>Snow and Ice</td>
<td>0</td>
<td>0</td>
<td>0.01</td>
</tr>
<tr>
<td>16</td>
<td>Barren or Sparsely Vegetated</td>
<td>0</td>
<td>0</td>
<td>0.035</td>
</tr>
<tr>
<td>17</td>
<td>Water</td>
<td>0</td>
<td>0</td>
<td>0.005</td>
</tr>
<tr>
<td>18</td>
<td>Wooded Tundra</td>
<td>4</td>
<td>50</td>
<td>0.055</td>
</tr>
<tr>
<td>19</td>
<td>Mixed Tundra</td>
<td>2</td>
<td>50</td>
<td>0.055</td>
</tr>
<tr>
<td>20</td>
<td>Barren Tundra</td>
<td>0.5</td>
<td>50</td>
<td>0.055</td>
</tr>
</tbody>
</table>
Fig. B4 MODIS IGBP 20-category land cover type in the model domain. Red polylines are 2020 wildfire burn scar perimeters.
Fig. B5 1-km STATSGO data with 16 soil texture types. Red polylines are 2020 wildfire burn scar perimeters.
Table B3
Default and calibrated soil parameters in WRF-Hydro

<table>
<thead>
<tr>
<th>Soil type</th>
<th>Default grain size distribution index</th>
<th>Default Porosity</th>
<th>Default Saturated hydraulic conductivity (m s⁻¹)</th>
<th>After calibration grain size distribution index</th>
<th>After calibration Porosity</th>
<th>After calibration Saturated hydraulic conductivity (m s⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sand</td>
<td>2.79</td>
<td>0.339</td>
<td>4.66E-5</td>
<td>2.51</td>
<td>0.315</td>
<td>1.5 x 10⁻⁷ m s⁻¹ for all the burn scars, and original values elsewhere.</td>
</tr>
<tr>
<td>Loamy sand</td>
<td>4.26</td>
<td>0.421</td>
<td>1.41E-5</td>
<td>3.83</td>
<td>0.392</td>
<td></td>
</tr>
<tr>
<td>Sandy loam</td>
<td>4.74</td>
<td>0.434</td>
<td>5.23E-6</td>
<td>4.27</td>
<td>0.404</td>
<td></td>
</tr>
<tr>
<td>Silt loam</td>
<td>5.33</td>
<td>0.476</td>
<td>2.81E-6</td>
<td>4.80</td>
<td>0.442</td>
<td></td>
</tr>
<tr>
<td>Silt</td>
<td>3.86</td>
<td>0.484</td>
<td>2.18E-6</td>
<td>3.47</td>
<td>0.450</td>
<td></td>
</tr>
<tr>
<td>Loam</td>
<td>5.25</td>
<td>0.439</td>
<td>3.38E-6</td>
<td>4.73</td>
<td>0.408</td>
<td></td>
</tr>
<tr>
<td>Sandy clay loam</td>
<td>6.77</td>
<td>0.404</td>
<td>4.45E-6</td>
<td>6.09</td>
<td>0.376</td>
<td></td>
</tr>
<tr>
<td>Silty clay loam</td>
<td>8.72</td>
<td>0.464</td>
<td>2.03E-6</td>
<td>7.85</td>
<td>0.432</td>
<td></td>
</tr>
<tr>
<td>Clay loam</td>
<td>8.17</td>
<td>0.465</td>
<td>2.45E-6</td>
<td>7.35</td>
<td>0.432</td>
<td></td>
</tr>
<tr>
<td>Sandy clay</td>
<td>10.73</td>
<td>0.406</td>
<td>7.22E-6</td>
<td>9.66</td>
<td>0.378</td>
<td></td>
</tr>
<tr>
<td>Silty clay</td>
<td>10.39</td>
<td>0.468</td>
<td>1.34E-6</td>
<td>9.35</td>
<td>0.435</td>
<td></td>
</tr>
<tr>
<td>Clay</td>
<td>11.55</td>
<td>0.468</td>
<td>9.74E-7</td>
<td>10.40</td>
<td>0.435</td>
<td></td>
</tr>
<tr>
<td>Organic material</td>
<td>5.25</td>
<td>0.439</td>
<td>3.38E-6</td>
<td>4.73</td>
<td>0.408</td>
<td></td>
</tr>
<tr>
<td>Bedrock</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>4.26</td>
<td>0.421</td>
<td>1.41E-5</td>
<td>3.83</td>
<td>0.392</td>
<td></td>
</tr>
<tr>
<td>Playa</td>
<td>11.55</td>
<td>0.468</td>
<td>9.74E-7</td>
<td>10.40</td>
<td>0.435</td>
<td></td>
</tr>
<tr>
<td>Lava</td>
<td>2.79</td>
<td>0.200</td>
<td>1.41E-4</td>
<td>2.51</td>
<td>0.186</td>
<td></td>
</tr>
<tr>
<td>White sand</td>
<td>2.79</td>
<td>0.339</td>
<td>4.66E-5</td>
<td>2.51</td>
<td>0.315</td>
<td></td>
</tr>
</tbody>
</table>

Soil parameters in default and calibrated WRF-Hydro. Default soil parameters in WRF-Hydro are adapted from the soil analysis by Cosby et al. (1984). Grain size distribution index and soil porosity are altered from default values during the global soil moisture calibration. Saturated hydraulic conductivity is altered from default values during the streamflow calibration.
Fig. B6 WRF-Hydro simulated discharge time-series at four debris flow source areas. (a)–(c) MRMS precipitation (green bars) and simulated discharge time-series for January 26th 00:00 to 31st 23:00 at Mill Creek, Big Creek, and Nacimiento debris flow source areas (black circles in Fig. 7b–d) in baseline (purple dashed line) and burn scar simulation (red line).
<table>
<thead>
<tr>
<th>Site name</th>
<th>Baseline simulation</th>
<th>Burn scar simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total volume (m³)</td>
<td>Peak discharge (m³ s⁻¹)</td>
</tr>
<tr>
<td>Mill Creek</td>
<td>10,023</td>
<td>0.23</td>
</tr>
<tr>
<td>Big Creek</td>
<td>11,611</td>
<td>0.71</td>
</tr>
<tr>
<td>Nacimiento</td>
<td>3,031</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Table B4 The total runoff volume, peak discharge, and peak timing in the baseline and burn scar simulations from January 27th 00:00 to 31st 23:00 at source areas of Rat Creek, Mill Creek, Big Creek, and Nacimiento debris flows (black circles in Fig. 7b–d). The percent change of the total volume and peak discharge in the burn scar simulation relative to the baseline simulation are shown in parentheses.
Fig. B7 Discharge volume-based runoff-generated debris flow hazard at catchment level in the (a) baseline simulation, (b) burn scar simulation, and (c) the difference between the burn scar and baseline simulations. For each catchment, the hazard is assessed by computing the total discharge volume at the catchment outlet from January 27th 00:00 to 28th 12:00.

**Data availability statement**

The modified WRF-Hydro Fortran code and instructions to output the overland flow at terrain routing grid can be downloaded at https://github.com/NU-CCRG/Modified-WRF-Hydro.

HazMapper v1.0 is available at https://hazmapper.org/. The SAR backscatter change method code is available at https://github.com/MongHanHuang/GEE_SAR_landslide_detection.

Author contribution
Conceptualization: CL, ALH, & DEH; Simulation and model analysis: CL; JW & WY model methodological development. Remote sensing analysis: ALH; Field Observations: NJF; GIS assistance: YX; Funding acquisition: GB & DH; CL wrote the original draft and all authors reviewed and edited the manuscript.

Competing interests
The authors declare that they have no conflict of interest.

Acknowledgments
C.L., A.L.H., J.W., X.L., G.B., and D.E.H. acknowledge support from NSF PREEVENTS #1848683. We acknowledge high-performance computing support from Cheyenne (doi:10.5065/D6RX99HX) provided by NCAR’s Computational and Information Systems Laboratory, sponsored by the National Science Foundation.

References


Preprint. Discussion started: 3 December 2021
© Author(s) 2021. CC BY 4.0 License.


Huang, X., Swain, D. L., & Hall, A. D. (2020). Future precipitation increase from very high resolution ensemble downsampling of extreme atmospheric river storms in California. 6(29), eaba1323. doi:10.1126/sciadv.aba1323 %J Science Advances


Santi, P. M., & Morandi, L. (2013). Comparison of debris flow volumes from burned and unburned areas. Landslides, 10(6), 757-769.


