Supporting Information

Text S1. rdNDVI and backscatter processing methods

We examined relative differences in normalized difference vegetation index (rdNDVI) calculated from Sentinel-2 satellite data using the HazMapper v1.0 Google Earth Engine application (see full description in 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.

We also 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 Iratio = $10*\log_{10}(\sigma^0_{pre}$ / (σ^0_{post}), where σ^0_{pre} is a pre-event image stack (defined as the temporal median) of SAR backscatter and σ^0_{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 Iratio < 99th percentile value.

Text S2. The default soil hydraulic and land surface properties in WRF-Hydro

By default, WRF-Hydro uses Moderate Resolution Imaging Spectroradiometer (MODIS) Modified International Geosphere-Biosphere Program (IGBP) 20-category land cover product as land cover (Figure S2) and 1-km Natural Resources Conservation Service State Soil Geographic (STATSGO) database for soil type classification (Figure S3; Miller & White, 1998). Land surface properties including canopy height, maximum carboxylation rate, and overland flow roughness are functions of land cover type (Table S1& Figure S2). Default soil hydraulic parameters in WRF-Hydro (i.e., soil porosity and saturated hydraulic conductivity) are based on Cosby et al.'s (1984) soil analysis (Table S2) and are used to map onto the STATSGO 16 soil texture types (Figure S3).

Text S3. The governing equation in the Noah-MP LSM to calculate rate of infiltration excess:

$$\frac{\partial h}{\partial t}=\frac{\partial P_d}{\partial t}\left\{1-\frac{\left[\sum_{i=1}^{4}\Delta Z_i(\theta_s-\theta_i)\right]\left[1-exp\left(-k\frac{K_s}{K_{ref}}\frac{\delta_t}{86400}\right)\right]}{P_d+\left[\sum_{i=1}^{4}\Delta Z_i(\theta_s-\theta_i)\right]\left[1-exp\left(-k\frac{K_s}{K_{ref}}\frac{\delta_t}{86400}\right)\right]}\right\}$$

where P_d(m) is the precipitation not intercepted by the canopy; ΔZ_i (m) is the depth of soil layer i; θ_i is the soil moisture in soil layer i; θ_s is the porosity; K_s (m/s) is the saturated hydraulic conductivity; K_{ref} is 2×10^{-6} m/s which represents the saturated hydraulic conductivity of the silty–clay–loam soil texture chosen as a reference; δ_t (s) is the model time step; and k is the runoff–infiltration partitioning parameter [k represents kdt_{ref} in Chen & Dudhia (2001)].

Text S4. Calculation of evaluation metrics

The Nash-Sutcliffe efficiency (NSE; Nash & Sutcliffe, 1970) and Kling-Gupta efficiency (KGE; Gupta el al., 2009; Kling et al., 2012) are the two most widely used metrics for calibration and evaluation of hydrologic models.

The NSE has been used to calibrate streamflow (e.g., Xia et al., 2012; Bitew & Gebremichael, 2011), and it is calculated following:

$$NSE=1-\frac{\sum_{t=1}^{t=T}(Q_{sim}(t)-Q_{obs}(t))^2}{\sum_{t=1}^{t=T}(Q_{obs}(t)-\overline{Q_{obs}})^2}$$

where T is the length of the time series, $Q_{sim}(t)$ and $Q_{obs}(t)$ are the simulated and observed discharge at time t, respectively, and $\overline{Q_{obs}}$ 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).

The KGE has been used for soil moisture calibration (e.g., Lahmers et al., 2019; Vergopolan et al., 2020) and the calculation of the KGE follows:

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

where r is the correlation coefficient between the observation and simulation, α is the ratio of the standard deviation of simulation to the standard deviation of observation, and β 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 bad model performance (Schönfelder et al., 2017; Andersson et al., 2017).

Figures and tables

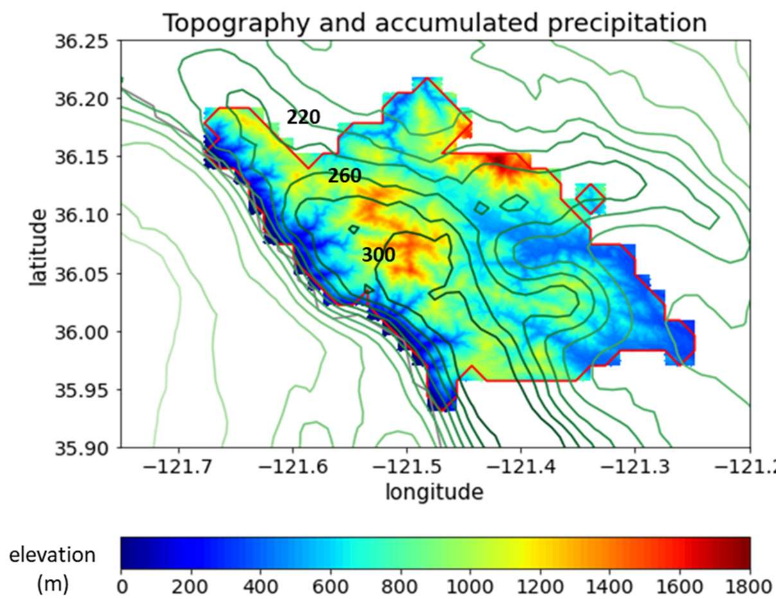

Figure S1. The topography (shading; m) and MRMS accumulated precipitation (contour lines; mm) during the AR event from January 26th 00:00 to 29th 23:00 in the Dolan burn scar. Contour line interval for accumulated precipitation is 20 mm, and lines of 220, 260, and 300 mm are labeled. The red polygon outlines the perimeter of the Dolan burn scar.

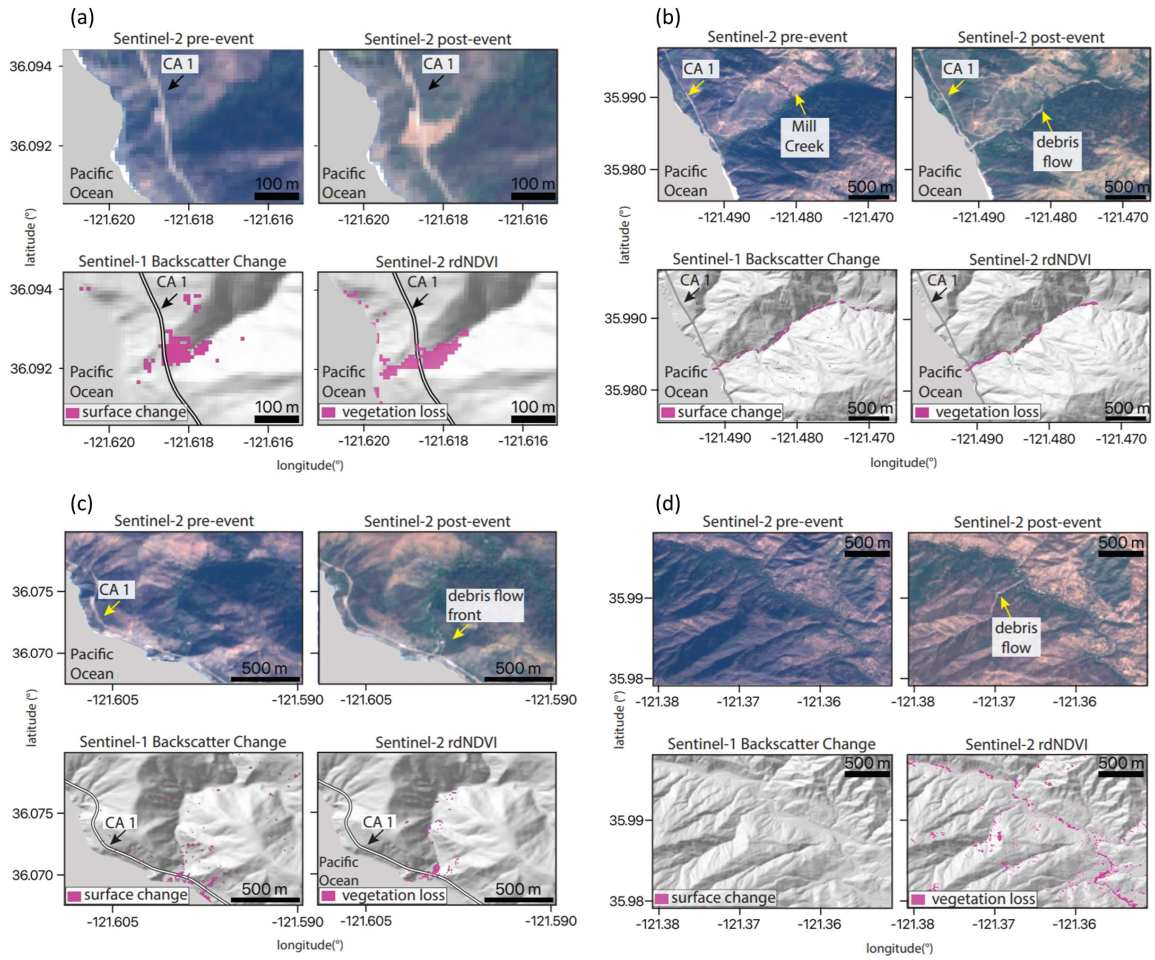

Figure S2. Optical- and SAR-based remote sensing data of four debris flows. Optical data from Sentinel-2 show pre- and post-debris flow imagery in real color. rdNDVI calculated from the Sentinel-2 data show a decrease in vegetation corresponding to debris flow locations. Sentinel-1 backscatter change 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.

Table S1

MODIS IGBP 20-category land cover type and properties in Noah-MP LSM

Land cover code	Land cover type	Canopy height (m)	Max carboxylation rate at 25°C ($\mu mol\ CO_2/(m^2 \cdot s)$)	Overland flow roughness
1	Evergreen Needleleaf Forest	20	50	0.2
2	Evergreen Broadleaf Forest	20	60	0.2
3	Deciduous Needleleaf Forest	18	60	0.2
4	Deciduous Broadleaf Forest	16	60	0.2
5	Mixed Forests	16	55	0.2
6	Closed Shrublands	1.1	40	0.055
7	Open Shrublands	1.1	40	0.055
8	Woody Savannas	13	40	0.055
9	Savannas	10	40	0.055
10	Grasslands	1	40	0.055
11	Permanent wetlands	5	50	0.07
12	Croplands	2	80	0.035
13	Urban and Built-Up	15	0	0.025
14	Cropland/natural vegetation mosaic	1.5	60	0.035
15	Snow and Ice	0	0	0.01
16	Barren or Sparsely Vegetated	0	0	0.035
17	Water	0	0	0.005
18	Wooded Tundra	4	50	0.055
19	Mixed Tundra	2	50	0.055
20	Barren Tundra	0.5	50	0.055

Table S1. MODIS IGBP 20-category land cover type and properties in Noah-MP LSM.

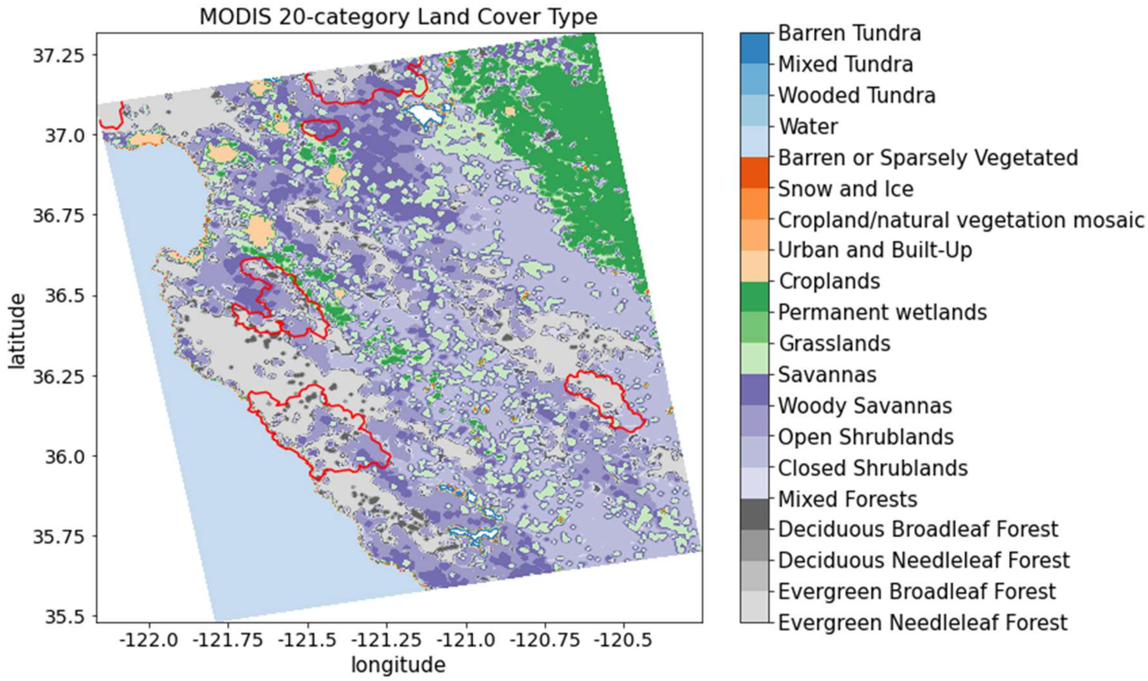

Figure S3. MODIS IGBP 20-category land cover type in the model domain. Red polylines are 2020 wildfire burn scar perimeters.

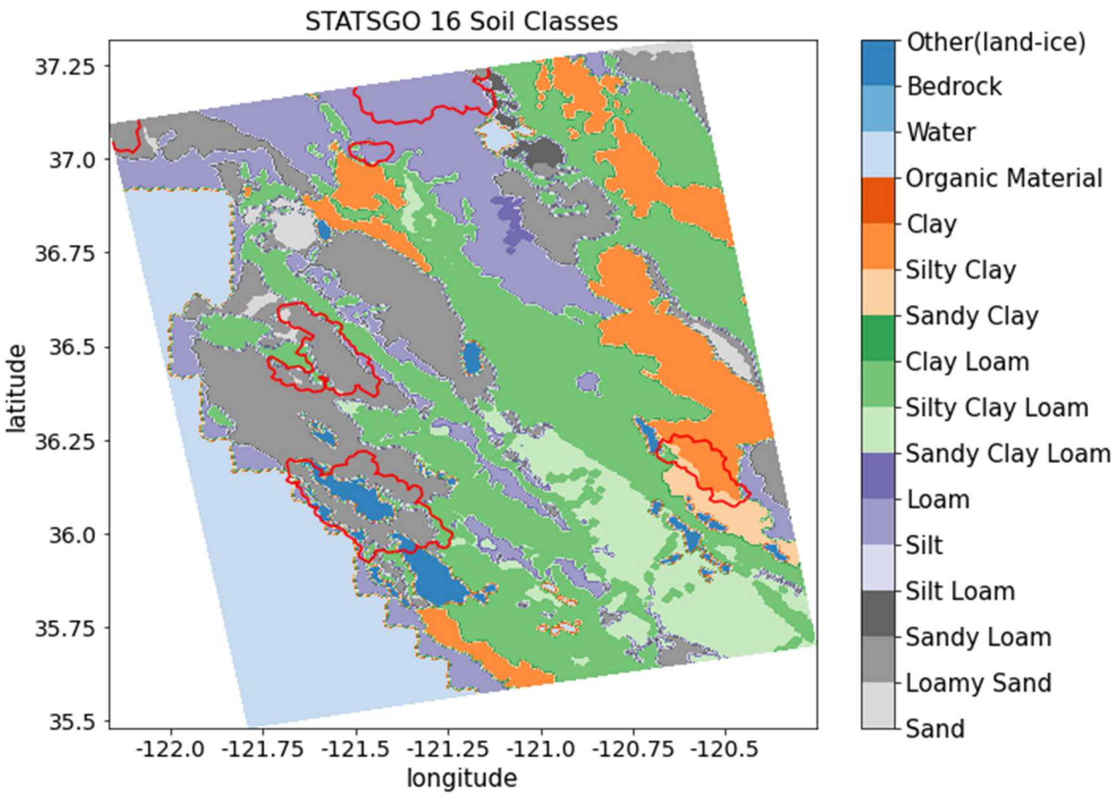

Figure S4. 1-km STATSGO data with 16 soil texture types. Red polylines are 2020 wildfire burn scar perimeters.

Table S2
Default and calibrated soil parameters in WRF-Hydro

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

Table S2. 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.

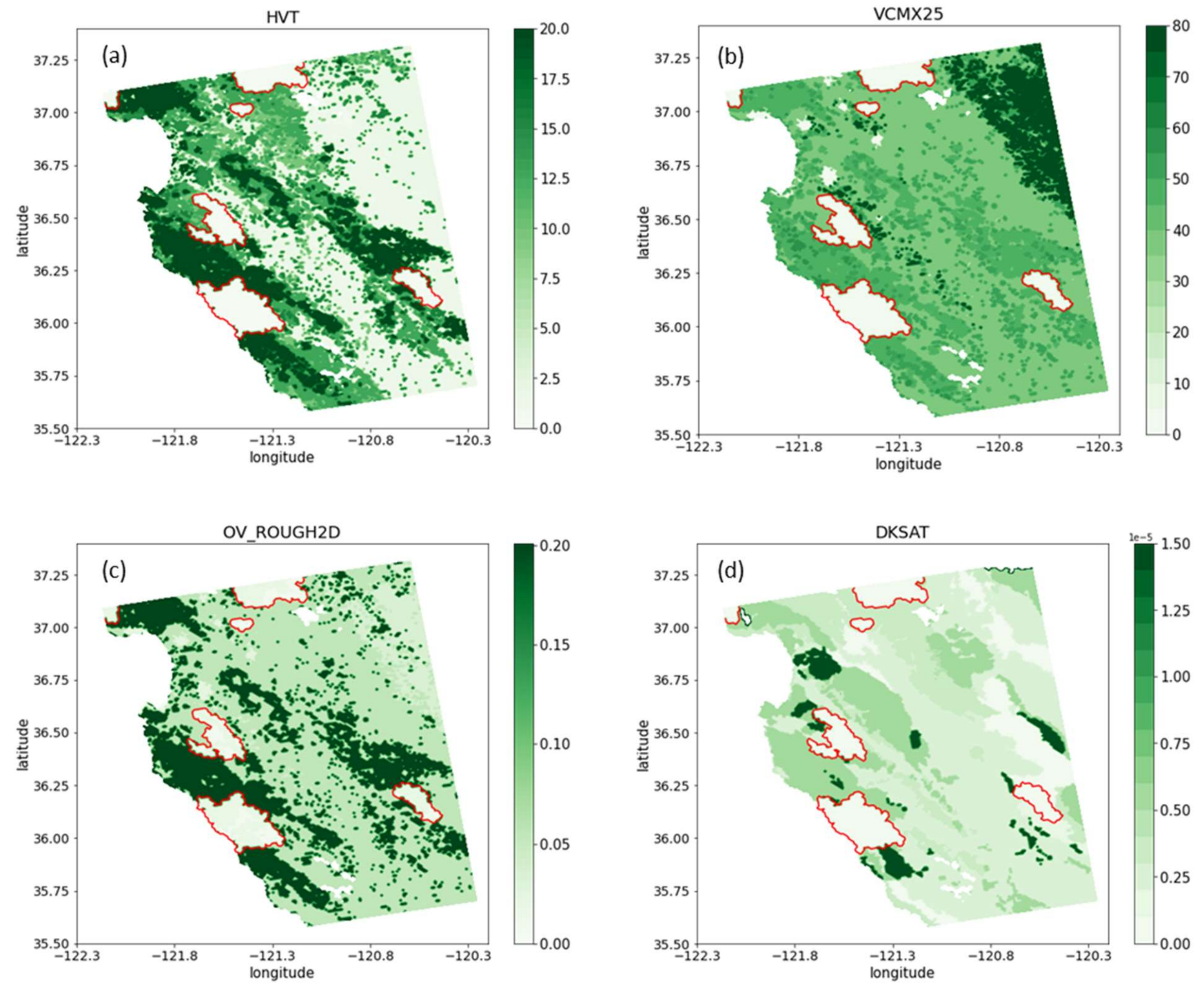

Figure S5. Parameter setting in the WRF-Hydro burn scar simulation. (a) The height of the canopy (HVT; m; shading), (b) maximum rate of carboxylation at 25°C (VCMX25; $\mu mol\ CO_2/(m^2 \cdot s)$; shading), (c) overland flow roughness coefficient (OV_ROUGH2D; shading), and (d) saturated hydraulic conductivity (DKSAT; m/s; shading) in the burn scar simulation.

Table S3

Evaluation metrics of simulated soil moisture and streamflow

Soil moisture (Default / Baseline)

Station	*r*	RMSE	Bias	KGE
lwd	0.97 / 0.98	7.06 / 4.32	5.21 / 4.16	0.10 / 0.72
gry	0.94 / 0.94	5.19 / 2.53	-4.79 / -1.66	0.80 / 0.88

Streamflow (Baseline / Burn scar)

Station	*r*	RMSE	Bias	NSE
1870	0.28 / 0.93	39.29 / 14.69	1.65 / 3.36	-0.17 / 0.84
2000	0.26 / 0.86	51.22 / 24.92	2.47 / 4.81	-0.15 / 0.73
2050	0.25 / 0.81	49.96 / 27.43	5.70 / 8.24	-0.38 / 0.53

Table S3. 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 Figure 1c). For streamflow, the numbers in front of "/" are computed between the baseline run (purple dashed line in Figure 1d) and the observations, while the numbers behind "/" are for burn scar simulation (red line in Figure 1d). If the model performance regarding a certain metric is enhanced in the burn scar simulation, the number after "/" is underlined.

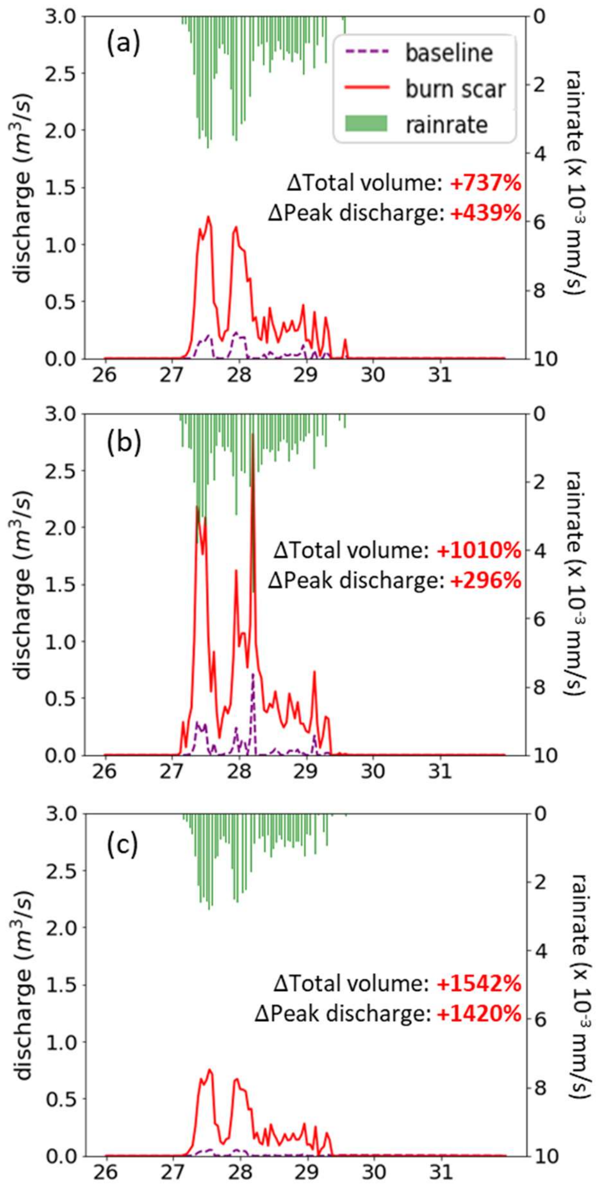

Figure S6. (a)–(c) MRMS precipitation and baseline and burn scar simulated discharge time series for January 26th 00:00 to 31st 23:00 at Mill Creek, Big Creek, and Nacimiento debris-flow source areas (black circles in Figures 2f–h).

Table S4

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

Site name	Baseline simulation			Burn scar simulation		
	Total volume (m^3)	**Peak discharge (m^3/s)**	**Peak timing**	**Total volume (m^3)**	**Peak discharge (m^3/s)**	**Peak timing**
Mill Creek	10,023	0.23	27th 23:00	83,853 (+737%)	1.24 (+439%)	27th 13:00
Big Creek	11,611	0.71	28th 05:00	128,879 (+1010%)	2.81 (+296%)	28th 05:00
Nacimiento	3,031	0.05	27th 13:00	49,792 (+1542%)	0.76 (+1420%)	27th 13:00

Table S4. 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 Figures 2f–h). The percent change of the total volume and peak discharge in the burn scar simulation relative to the baseline simulation are shown in parentheses.

Table S5

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

Site name	Baseline simulation			Burn scar simulation			
	Total volume (m^3)	Peak discharge (m^3/s)	Highest peak timing	Total volume (m^3)	Peak discharge (m^3/s)	1st Peak timing	2nd Peak timing
Rat Creek	6,897	0.54	28th 05:00	61,425 (+791%)	1.73 (+220%)	27th 09:00	28th 05:00
Mill Creek	312,925	13.10	29th 08:00	2,347,457 (+650%)	45.21 (+245%)	27th 13:00	27th 23:00
Big Creek	842,808	46.10	29th 16:00	8,354,095 (+891%)	154.10 (+234%)	27th 10:00	28th 05:00
Nacimiento	743,531	33.15	29th 16:00	6,904,706 (+829%)	135.41 (+308%)	27th 14:00	28th 00:00

Table S5. 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 deposition sites of Rat Creek, Mill Creek, Big Creek, and Nacimiento debris flows (black triangles in Figures 2e–h). The peak timing shown in the baseline simulation is for the highest peak. The percent change of the total volume and peak discharge in the burn scar simulation relative to the baseline simulation are shown in parentheses.