JGR Atmospheres

RESEARCH ARTICLE
10.1029/2023JD039780

Evaluating Noah-MP Simulated Runoff and Snowpack in Heavily Burned Pacific-Northwest Snow-Dominated Catchments

Ronnie Abolafia-Rosenzweig[1], Cenlin He[1], Fei Chen[1], Yongxin Zhang[1], Aubrey Dugger[1], Ben Livneh[2], and David Gochis[3]

[1]NSF National Center for Atmospheric Research, Boulder, CO, USA, [2]Department of Civil, Environmental and Architectural Engineering, University of Colorado Boulder, Boulder, CO, USA, [3]Airborne Snow Observatories, Inc., Crowley Lake, CA, USA

Key Points:

- Noah-MP underestimates annual runoff anomalies relative to observations following fires in the Pacific-Northwest
- In post-fire years, Noah-MP fails to simulate deeper snowpacks that ablate faster and the associated enhanced spring runoff
- Informing the model with satellite-observed vegetation characteristics (including fire effect) did not resolve post-fire model deficiencies

Supporting Information:

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

Correspondence to:

R. Abolafia-Rosenzweig,
abolafia@ucar.edu

Citation:

Abolafia-Rosenzweig, R., He, C., Chen, F., Zhang, Y., Dugger, A., Livneh, B., & Gochis, D. (2024). Evaluating Noah-MP simulated runoff and snowpack in heavily burned Pacific-Northwest snow-dominated catchments. *Journal of Geophysical Research: Atmospheres*, *129*, e2023JD039780. https://doi.org/10.1029/2023JD039780

Received 6 AUG 2023
Accepted 7 APR 2024

Author Contributions:

Conceptualization: Cenlin He
Data curation: Yongxin Zhang, Aubrey Dugger
Formal analysis: Cenlin He
Funding acquisition: Cenlin He
Investigation: Cenlin He, Aubrey Dugger
Methodology: Cenlin He, Fei Chen, Yongxin Zhang, Aubrey Dugger, Ben Livneh, David Gochis
Project administration: Cenlin He
Resources: Fei Chen, David Gochis
Supervision: Cenlin He, Fei Chen
Validation: Cenlin He
Writing – review & editing: Cenlin He, Fei Chen, Aubrey Dugger, Ben Livneh, David Gochis

Abstract Terrestrial hydrology is altered by fires, particularly in snow-dominated catchments. However, fire impacts on catchment hydrology are often neglected from land surface model (LSM) simulations. Western U.S. wildfire activity has been increasing in recent decades and is projected to continue increasing over at least the next three decades, and thus it is important to evaluate if neglecting fire impacts in operational land surface models (LSMs) is a significant error source that has a noticeable signal among other sources of uncertainty. We evaluate a widely used state-of-the-art LSM (Noah-MP) in runoff and snowpack simulations at two representative fire-affected snow-dominated catchments in the Pacific Northwest: Andrew's Creek in Washington and Johnson Creek in Idaho. These two catchments are selected across all western U.S. fire-affected catchments because they are snow-dominated and experienced more than 50% burning in a single fire event with minimal burning outside of this event, which allows analyses of distinct pre- and post-fire periods. There are statistically significant shifts in model skills from pre-to post-fire years in simulating runoff and snowpack. At both study catchments, simulations miss enhancements in early-spring runoff and annual runoff efficiency during post-fire years, resulting in persistent underestimates of annual runoff anomalies throughout the 12-year post-fire analysis periods. Enhanced post-fire snow accumulation and melt contributes to observed but unmodeled increases of spring runoff and annual runoff efficiency at these catchments. Informing simulations with satellite observed land cover classifications, leaf area index, and green fraction do not consistently improve the model ability to simulate hydrologic responses to fire disturbances.

Plain Language Summary Western U.S. fire activity has been increasing in recent decades and is expected to continue to increase in coming decades. Fires remove vegetation and alter soils which in turn alters terrestrial hydrology. Fire effects on hydrology are particularly significant over snowy catchments that serve as natural water towers for major western U.S. rivers. Sophisticated models often neglect or underrepresent fire effects on land surface properties, and thus are susceptible to larger errors after fires. This study compares runoff and snow simulations from a widely used land surface model (LSM) with observations to quantify fire-induced changes in model accuracy over two snow dominated catchments in the Pacific Northwest. Simulations persistently underestimate enhanced spring runoff and annual runoff anomalies in post-fire years. These underestimates are consistent with observed enhancements in post-fire snow accumulation and melt, which the model mostly failed to capture. The finding that post-fire model errors are consistent with previously published fire impacts on hydrology supports that fire is an important error source in LSM simulations that should be accounted for.

1. Introduction

Seasonal snowpack serves as a natural water tower in the western United States (WUS) where winter precipitation is stored as snow and released as liquid in spring and summer when water demands are greatest (Viviroli et al., 2007). Water managers rely on accurate simulations of winter snowpack accumulation and spring-summer snow melt from land surface models (LSMs) to forecast and manage water resources in the snow-reliant WUS, which receives most of its streamflow from snowmelt (Kapnick et al., 2018; D. Li et al., 2017). However, LSMs often neglect fire-induced changes to vegetation, snow and soil properties in weather and hydrological applications despite analyses that show fire disturbances significantly alter land surface properties and the terrestrial water budget, particularly in snow-dominated catchments (Hampton & Basu, 2022; Harpold et al., 2014; Koshkin

et al., 2022; Smoot & Gleason, 2021; Williams et al., 2022). This underrepresentation of fire effects in land surface and hydrologic models is a growing threat to the accuracy of water supply forecasts in the WUS where burned area has increased by 255 km^2/year from 1984 to 2020 in snowy regions of WUS (Abolafia-Rosenzweig, He, & Chen, 2022) and projections suggest climate will be twice as conducive for wildfire in the WUS from 2021 to 2050 compared to 1991–2020 (Abatzoglou et al., 2021; Williams et al., 2022). These projections indicate that fire effects on water supply are expected to continue growing over the next few decades such that streamflow will be noticeably altered at regional scales in the WUS (Williams et al., 2022). An important step to addressing this source of uncertainty in weather and hydrologic predictions is to quantify fire-induced errors from operationally-used LSM simulations over snow-dominated catchments.

Cumulative effects of fire on land surface hydrology are often associated with significantly enhanced streamflow (Williams et al., 2022) and reduced evapotranspiration (ET) (Ma et al., 2020; Maina and Siirila-Woodburm, 2020). However, fire impacts on hydrology are heterogenous at the scale of individual catchments because fire-induced changes to hydrologic processes depend on complex interactions among many factors including: burn area and severity, catchment size, human management (e.g., of reservoirs and forests), vegetation, soil type, meteorology, and topography (Atchley et al., 2018; Goeking & Tarboton, 2020, 2022; Niemeyer et al., 2020; Partington et al., 2022; Pugh & Gordon, 2013; Spence et al., 2020). Effects of fire on soil—increased bulk density by decreasing macropores, reduced infiltration capacity by sealing pore space with ash and sediment, and formation of a hydrophobic layer at the soil surface which tends to reduce hydraulic conductivity and sorptivity —are generally associated with higher runoff efficiency (the ratio of runoff (Q) to precipitation (P); that is, Q/P), drier top soils, and wetter subsoils (Ebel, 2020; Ebel & Martin, 2017; Martin & Moody, 2001; Moody et al., 2008; Shakesby & Doerr, 2006; Stoof et al., 2012). The response of vegetation to fires have competing effects on hydrologic states and fluxes. For instance, vegetation removal is associated with reduced downward longwave radiation and precipitation interception, both of which contribute to greater ground snow accumulation (Burles & Boon, 2011; Harpold et al., 2014; Seibert et al., 2010). Conversely, vegetation burning is also associated with decreased snow albedo (i.e., snow darkening) due to burned woody debris shedding from standing trees onto the snowpack, reduced shading and increased wind speeds and associated turbulent heat flux to the snowpack, all of which favor more winter ablation and earlier and faster spring snowmelt (Burles & Boon, 2011; Campbell & Morris, 1988; Gleason et al., 2013, 2019; Harpold et al., 2014; Kampf et al., 2022; Maxwell & St Clair, 2019; Niemeyer et al., 2020; Pugh & Small, 2012; Seibert et al., 2010). Furthermore, vegetation removal also has competing effects on ET: increased soil evaporation but decreased transpiration and canopy evaporation (Bond-Lamberty et al., 2009; Maina & Siirila-Woodburm, 2020).

Previous research has shown that physically-based hydrologic model simulations designed to account for these fire impacts provide significantly different simulations of runoff, snowpack and evapotranspiration relative to simulations that do not account for fire impacts (Maina & Siirila-Woodburm, 2020), and conceptual hydrologic models that do not account for fire impacts have larger runoff biases post-fire than pre-fire (Seibert et al., 2010; Williams et al., 2022). These studies which quantify fire effects on hydrology by differencing conceptual models that do not account for fire effects with post-fire observations approach the problem in a manner akin to earlier paired catchment studies, where models allow water yield from the post-fire watershed to be compared to the same watershed prior to the fire disturbance. Importantly, these studies indicate that neglecting fire effects on land surface properties in physically-based LSMs may be an important source of uncertainty in post-fire hydrology simulations.

In this study, we are motivated to evaluate a state-of-the-art physically-based LSM, Noah-MP, which is a key land component of widely-used operational and research weather and hydrological modeling systems such as the Weather Research and Forecasting model (WRF) (Powers et al., 2017), the WRF-Hydro/NOAA National Water Model (Gochis et al., 2020), the NASA Land Information System (LIS) (Kumar et al., 2006; Peters-Lidard et al., 2007) and the NOAA Unified Forecast System (UFS). Because the Noah-MP LSM was not originally designed to account for fire effects on hydrology, we hypothesize that simulations of runoff and snowpack will be less skillful after fires. Prior to this analysis, it was uncertain if model errors that are attributable to the neglect of fire-perturbations could be identifiable, or if other well-established sources of LSM uncertainty (i.e., model structure, parameter uncertainty, meteorological forcing uncertainty or observational noise) would mask this additional source of uncertainty (Cuntz et al., 2016; He et al., 2021; Zhang et al., 2016). Resolving this uncertainty is valuable to prioritize model development strategies for LSMs. This research is becoming increasingly important over the U.S. which experienced 799,412 km^2 of burning (nearly 1.9 times the area of California) from

1983 to 2021 (https://www.nifc.gov/fire-information/statistics/wildfires) and is expected to experience accelerated burning over at least the next three decades (Abatzoglou et al., 2021).

This study focuses on WUS snow dominated catchments that experienced more than 50% burning in a single fire event. This research uniquely quantifies fire-induced uncertainties in sophisticated LSM simulations of snowpack at daily to annual time scales and streamflow at monthly and annual scales while considering the relationship between simulated errors in snowpack with errors in streamflow. Furthermore, this study is unique in that we use differing vegetation configurations to evaluate potential challenges in accounting for fire impacts on hydrology in LSM simulations. Overall, this study is motivated to better understand the magnitude and nature of fire-induced LSM uncertainties to inform future research that plans to implement a fire module within Noah-MP to explicitly account for fire disturbances. Specifically, this study seeks to answer the following science questions: (a) Are differences in skill scores of modeled runoff and snow fluxes and states between pre- and post-fire periods statistically significant? (b) Are model skill changes in streamflow physically consistent with skill changes in snowpack from pre-to post-fire years? (c) Does informing LSM simulations with remotely sensed vegetation properties help capture the effects of fire on hydrology?

2. Materials and Methods

2.1. Study Domain

We perform analyses at two study catchments—Andrew's Creek and Johnson Creek in the Pacific Northwest (PNW; Figure 1)—which are selected based on the following criteria: (a) catchments that are selected for analysis of fire effects on water supply across the entire WUS by Williams et al. (2022), (b) catchments that are snow-dominated (peak flows are from snow melt), (c) catchments that had a single major fire event that occurred during the MODIS-era that burned at least 50% of the watershed area, and (d) catchments that had no other significant fire events (i.e., burning more than 15% of the catchment) in the analysis periods. The last two criteria are to ensure a clear separation of pre- and post-fire conditions.

The Andrew's Creek catchment, approximately 58 km^2 with elevations ranging from 1,532 to 2,247 m, is in Okanogan County in northern Washington. 96% of the Andrew's Creek catchment burned during the Fawn Peak Complex in June 2003, with 67% of the catchment experiencing high-severity burning, 18% of the catchment experiencing mid-severity burning and 11% of the catchment experiencing low-severity burning (based on the Relative differenced Normalized Burn Ratio index; Eidenshink et al., 2007) (Figure 1). The Andrew's Creek catchment received an average of 1,031 mm of precipitation each water year, with approximately 63% of precipitation falling as snow during the 1990–2019 analysis period. Andrew's Creek was comprised of 46% evergreen needle leaf forest, 41% woody savannas, and 13% grasslands pre-fire and transitioned to over 90% grasslands across all post-fire years (2004–2019).

The Johnson Creek catchment, approximately 562 km^2 with elevations ranging from 1,539 to 2,577 m, is in Central Idaho within the Boise National Forest. Sixty percent of the Johnson Creek catchment burned during the Cascade Complex in July 2007, with 22% of the catchment experiencing high-severity burning, 15% experiencing mid-severity burning and 23% experiencing low-severity burning. The Johnson Creek catchment received approximately 1,069 mm of precipitation each water year, with approximately 66% of precipitation falling as snow during the 1995–2019 analysis period. Johnson Creek was comprised of 82% woody savanna, 8% savanna, 7% grasslands and 3% evergreen needleleaf forest pre-fire, and had increased grassland area post-fire (ranging from 31%–59% from 2008 to 2019) and decreased woody savanna area (ranging from 19%–54% from 2008 to 2019) and a modest increase to savanna area (ranging from 14%–24% from 2008 to 2019).

Hence, both Andrew's Creek and Johnson Creek experienced persistent landcover transitions from forested and wooded areas to grasslands in post-fire years, with more dramatic transitions at Andrew's Creek which experienced more severe burning. MODIS observes post-fire reductions in vegetation characteristics at both catchments, where post-fire leaf area index (LAI) and green fractions are lower than any pre-fire years in the MODIS era (Figures 1f and 1i).

2.2. Selection of Pre- and Post-Fire Periods

To quantify changes in model uncertainty following fires, model skill scores are calculated based on comparisons with observations of streamflow and snow during distinct pre- and post-fire periods. In annual analyses of

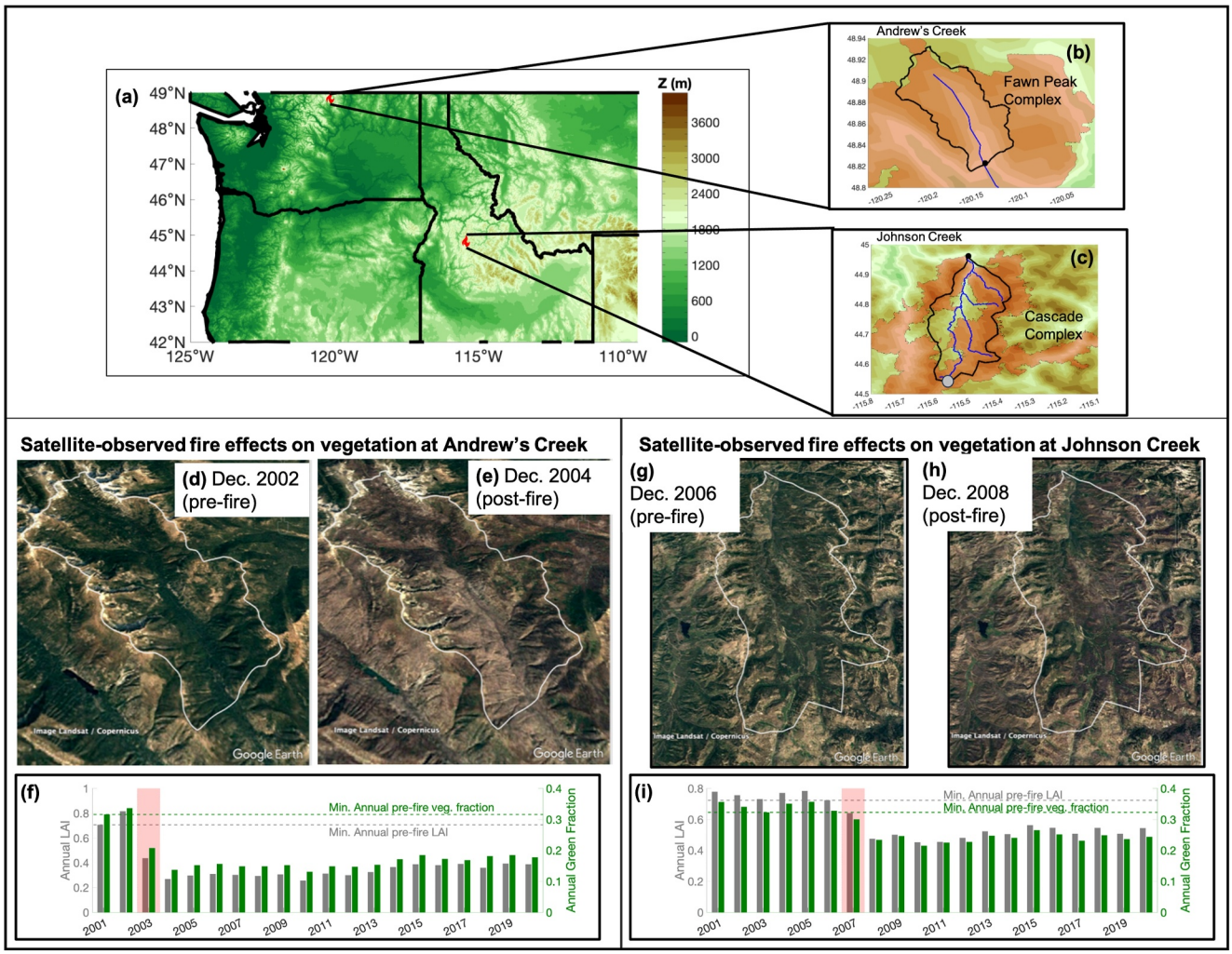

Figure 1. Analyses of fire effects on Noah-MP skill are conducted at the Andrew's Creek and Johnson Creek catchments shown in the (a) site vicinity map with topography, (b) zoom in of Andrew's Creek catchment overlain on the topographic map, with the blue line representing Andrew's Creek, red shading showing the area of the study fire event (Fawn Peak Complex: June 2003) and the black dot showing the location of the USGS stream gage at the watershed outlet, and (c) same as (b) but for the Johnson Creek catchment which experienced the Cascade Complex fire event in July 2007. The gray dot shows the location of the SNOTEL station (Deadwood Summit) used to evaluate simulated snowpack. (d), (e) Google Earth images of Andrew's Creek pre- and post-fire shows satellite-observed decreases in vegetation in response to the fire event. (f) Bar graph of mean annual MODIS leaf area index (LAI) and green fraction (8-day 500-m MOD15A2 data), with red shading marking the fire-year and dashed gray and green reference lines marking the minimum pre-fire LAI and green fraction, respectively. (g)–(i) Same as (d)–(f) but for Johnson Creek. Fire perimeters are from the Landsat-based Monitoring Trends in Burn Severity data set (Eidenshink et al., 2007).

streamflow, data are aggregated by water year. Water years that do not have observations from at least 70% of days are excluded to reduce impacts of data availability. To ensure consistency between temporally continuous model results and discontinuous observations, model outputs are matched temporally with observations so that we only account for simulated values on days when observations are available. At Andrew's Creek, the selected pre-fire years are 1990–1999 and 2001–2002 and post-fire years are 2004–2009, 2011, 2015–2019, while 2003 is the fire event year. At Johnson Creek, pre-fire years are 1995–2006 and post-fire years are 2008–2019 (Figure S1 in Supporting Information S1), while 2007 is the fire event year. Pre- and post-fire periods for monthly streamflow analyses evaluate the 120 pre- and post-fire months with available streamflow observations closest to the fire ignition month (June 2003 at Andrew's Creek and July 2007 at Johnson Creek). The 120-month period is selected to be consistent in duration with the 12-year pre- and post-fire periods in annual streamflow analyses.

The snowpack analysis at Johnson Creek uses 11 pre- and post-fire years, constrained by SNOTEL snow water equivalent (SWE) observations which are through the 2018 water year (11-years post-fire). Daily SNOTEL snow observations are available during all days in the 11 pre- and post-fire years (1996–2006 and 2008–2018,

respectively). For snow disappearance analyses relying on MODIS snow cover data (Section 3.3), only burned pixels are considered. Pre- and post-fire data sets are set to have an equal number of data points to enforce consistency in pre-to post-fire comparisons. The length of pre- and post-fire periods for these analyses are constrained by the number of pre-fire retrievals: two pre-fire years at Andrew's Creek (51 1-km pixels × 2 pre-fire years = 102 total data points) and six pre-fire years at Johnson Creek (354 1-km pixels × 6 pre-fire years = 2,124 data points). Thus, all pre-fire data are included, and the following 2 and 6 years of data after fire events are selected for Andrew's Creek and Johnson Creek analyses, respectively.

We consider if changes in meteorological conditions at each of the study catchments from pre-to post-fire periods may contribute to significant changes in LSM uncertainty between the two periods (Figure S2 in Supporting Information S1). At Andrew's Creek, the median annual precipitation (1,104 mm) and surface air temperature (274.5 K) during the 12 pre-fire years were not significantly different ($p > 0.46$, based on 2-tailed Wilcoxon rank sum tests) from the median annual precipitation (1,179 mm) and surface air temperature (274.8 K) during the 12 post-fire years. At Johnson Creek, the pre-fire median annual precipitation (1,075 mm/year) and surface air temperature (275.2 K) tended to be cooler and wetter than post-fire values (968 mm/year and 275.7 K); however, these differences are not statistically significantly ($p > 0.12$; based on 2-tailed Wilcoxon rank sum tests). We found similar annual cycles for these meteorological conditions between the two periods, particularly for surface air temperature. Minor differences in meteorological conditions between pre-to post-fire years, relative to the persistent reductions in vegetation characteristics (Figure 1), supports that significant changes in LSM skill scores from pre-to post-fire years are more likely to be attributable to difficulties in modeling fire impacts than changes in meteorological regimes.

2.3. Noah-MP Simulations

Noah-MP (Niu et al., 2011) is a widely-used state-of-the-art LSM, which serves as a key land component in WRF-Hydro/National Water Model (NWM), WRF, NASA LIS, and NOAA UFS. Noah-MP is designed to simulate land surface water and energy budgets, including soil and snow water balances. Noah-MP considers a three-layer snowpack depending on snow depth, which allows it to simulate snow variables more accurately than its predecessor (Noah; Chen & Dudhia, 2001; Chen et al., 1996, 1997; Ek et al., 2003). Comprehensive details of Noah-MP snowpack treatment can be found in He et al. (2023). The open-source community Noah-MP LSM is maintained by the National Center for Atmospheric Research (NCAR) and is available on GitHub: https://github.com/NCAR/noahmp.

Noah-MP simulations are run at a 1-km spatial resolution forced with eight surface meteorological variables (precipitation, temperature, east-to-west and north-to-south wind speed, long- and short-wave radiation, air pressure, and specific humidity) from the 1-km hourly observation-constrained Analysis of Record for Calibration (AORC; Fall et al., 2023) data set that is used to drive NWM. We bias-correct AORC daily surface air temperature and precipitation to match observationally-based values for these variables from the widely evaluated 4-km PRISM data set (Parameter-elevation Relationships on Independent Slopes Model) (Daly et al., 2008) (adjustments are shown in Figure S3 of Supporting Information S1). Specifically, hourly AORC data are adjusted to maintain daily mean temperature and precipitation values from PRISM, while maintaining the diurnal pattern from AORC. Model topography is based on 30-m data from the Shuttle Radar Topography Mission (https://doi.org/10.5066/F7K072R7). Simulations use the land model physics options from the WRF/Noah-MP options used in the continental-scale convection-permitting regional climate simulations (He et al., 2019; Liu et al., 2017; Rasmussen et al., 2023) that successfully captured precipitation and snow conditions in WUS, with updated snow cover parameters that improve surface albedo and temperature (He et al., 2021). Quantitative results presented in this study may change depending on different model physics options and parameters; however, comprehensively assessing model sensitivity to various combinations of physics options available within Noah-MP such as in Zhang et al. (2016), J. Li et al. (2019), and Zhang et al. (2020) is beyond the scope of this study. All model simulations are generated without river routing, which we assume does not introduce substantial uncertainty at monthly to annual time scales that streamflow validations are conducted at.

Dominant (i.e., mode) pre- and post-fire vegetation classifications are selected for each 1-km model pixel allowing simulations to account for land cover classification conversions following fire events. Simulations are forced with remotely sensed green fraction and LAI from 8-day MODIS retrievals (Section 3.3) which are linearly interpolated to hourly forcing. During pre-MODIS times, simulations are forced with a pre-fire multiyear average

monthly climatology of green fraction and LAI, computed for each vegetation classification in the study domain based on MODIS observations during the total pre-fire periods.

It is valuable to determine if the model configuration detailed above—which is designed to be partially fire-aware by accounting for fire impacts on vegetation characteristics and classifications—is more skillful than a no-fire model configuration. Therefore, we conduct supplementary no-fire twin-simulations for comparison. Namely, we evaluate if simulations using the partially fire-aware model configuration has relatively greater skill than no-fire simulations in post-fire years. No-fire simulations are configured as discussed above, except these twin-simulations use pre-fire vegetation classifications and are forced with the pre-fire monthly MODIS green fraction and LAI climatology during the entire analysis period. Vegetation forcing from the fire-aware and no-fire simulations are compared in Figure S4 of Supporting Information S1.

2.4. Comparing Simulated and Observed Monthly Flow and Relationships to Post-Fire Snowmelt and Accumulation

It is important to evaluate LSM biases at both annual and sub-annual (e.g., monthly) time scales to understand the bulk impact of fire-induced uncertainties to the LSM as well as when (i.e., which season) uncertainties are most prevalent. In monthly streamflow evaluations, we use a normalized streamflow (Q') which is defined as monthly Q divided by annual P throughout pre- and post-fire years. This metric allows for comparison of changes in monthly streamflow from pre-to post-fire years from the LSM and observations while controlling for the primary meteorologic control (P). Annual precipitation is selected for this normalization, because normalizing monthly Q by monthly P is not physically meaningful in snow-dominated catchments where there are multi-month lags between P and Q (e.g., winter P contributes to spring Q).

We hypothesized that a deeper snowpack that melts faster would contribute to LSM streamflow uncertainties in post-fire years based on previous research which found removal of vegetation in snowy areas favors greater ground snow accumulation and often greater ablation rates when temperatures exceed freezing (Gleason et al., 2019; Harpold et al., 2014). A key assumption in this hypothesis is that a deeper snowpack that melts faster has a robust relationship with runoff efficiency. We evaluate this assumption using Noah-MP simulations by quantifying correlations between simulated annual runoff efficiency with simulated ablation rates and peak SWE. Correlations between these snow-related metrics with runoff efficiency are used to determine if the aforementioned hypothesis is consistent with relationships depicted by the sophisticated physically-based model. Statistically significant positive correlations between Q/P with peak SWE and ablation rates would indicate that a deeper snowpack that melts faster is physically consistent with enhanced runoff efficiency.

After evaluating this assumption, we then evaluate if Noah-MP has a shift from pre-to post-fire biases in snow accumulation and ablation metrics that are consistent with this hypothesis. First, we perform catchment-wide comparisons of simulated and satellite-observed day of snow disappearance (DSD) to determine if there is a systematic shift in DSD biases from pre-to post-fire years across all burned pixels. If observations capture the signal of fire-induced ablation acceleration, but the model fails to simulate this effect, then Noah-MP post-fire biases (Noah-MP DSD—observed DSD) will be larger than pre-fire biases because observed DSD would occur earlier under the same meteorological conditions whereas Noah-MP would remain constant. DSD from Noah-MP and MODIS observations are identified as the first day when there is a transition from snow-covered to no snow conditions during the spring melt season. The snow/no-snow threshold used for Noah-MP and MODIS are 0% and 50% snow cover area, respectively (O'Leary et al., 2018, 2019). The stricter no-snow threshold for Noah-MP favors relatively later disappearance from simulations; however, preliminary analyses show simulations generally estimate earlier snow disappearance than MODIS observes indicating that the differing no snow assumptions do not impose a systematic bias. However, differences between LSM and observational uncertainties can impact the DSD bias analyses so we primarily focus on evaluating the change in bias from pre-to post-fire periods under consistent assumptions. Additional details for MODIS DSD observations are provided in Section 3.3.

Final snow analyses use in situ observations from the Deadwood SNOTEL station in the Johnson Creek catchment (the only SNOTEL station in either study fire perimeters; Figure 1c) to compare simulated and observed snow accumulation and melt. In these analyses, we compare multiyear mean daily averages of pre- and post-fire SWE to determine differences in magnitude and timing of simulated and observed SWE. We also compare simulated and observed ablation rates and peak SWE to quantify changes in Noah-MP skill for these

snow evolution metrics after the Cascade Complex fire. Ablation rates are calculated following the methodology presented in Xiao et al. (2021), as the rate of change in SWE between 80% and 20% of peak SWE during the falling limb of the annual SWE cycle. Noah-MP simulations for the in situ analysis are identical to the description in Section 2.3, except in situ observed daily precipitation and temperature from the SNOTEL site are used to bias correct these meteorological forcing inputs for simulations (following the same forcing bias-correction method described in Section 2.3). These meteorological adjustments are shown in Figure S5 of Supporting Information S1.

2.5. Evaluation Metrics and Significance Testing

We use bias and Spearman's correlation coefficient (ρ) as skill metrics throughout the analysis. Spearman's correlation coefficient is used instead of Pearson's because Pearson's correlation coefficient assumes normality, and there are multiple instances of non-gaussian data in this study. Additionally, Spearman's correlation coefficient is relatively more robust to outliers. Biases are calculated as the difference between the mean of simulated and observed time series, during respective pre- and post-fire periods. Streamflow time series are converted to anomalies (discussed in Section 3) prior to comparison.

Statistical differences in pre- and post-fire bias and ρ are computed with the widely used 2-tailed Wilcoxon rank sum test and the permutation test, respectively. Permutation tests are based on resampling the original data without replacement to test the hypothesis of no statistical difference between ρ calculated during pre- and post-fire periods, respectively. This is performed by: (a) calculating the absolute difference between post- and pre-fire ρ, (b) pooling data from pre- and post-fire periods together, (c) shuffling the data randomly and calculating differences in ρ between random samples equivalent to the length of pre- and post-fire periods (i.e., 12-year), (d) repeat step (c) 10,000 times, and (e) calculate the proportion of shuffled absolute differences in ρ that exceed the originally calculated ρ when data was organized in pre- and post-fire years.

3. Observations for Model Evaluations

3.1. In Situ Streamflow Observations

Simulations are compared with gauged streamflow from USGS stations at the outlets of water catchment areas (Figures 1b and 1c; catchment polygons are from the USGS GAGES-II data set: https://doi.org/10.3133/70046617). Streamflow model evaluations are conducted at monthly and annual timescales because simulated streamflow uncertainty at sub-monthly timescales may be large due to the neglect of river routing in model simulations. Following Williams et al. (2022), observed and simulated time series of Q are converted to the standardized anomaly space (Q-anomalies) to isolate fire-induced discrepancies rather than analyzing results confounded by systematic biases between simulations and observations that are not related to fire-disturbances. For pre- and post-fire comparisons, Q-anomalies are computed across the entire 25-year analysis period (12 pre-fire years, the fire year, and 12 post-fire years):

$$Q_{\text{anomaly},i} = \frac{Q_i - \overline{Q}}{\sigma_Q} \tag{1}$$

where Q_i is Q from the ith month or year in the time series (depending on if comparisons are monthly or annual), $\overline{Q}$ is the total time series mean, and σ_Q is the total time series standard deviation of Q. Because Q time series are not normally distributed, the anomaly conversion is exclusively used to enforce a zero bias between simulated and observed time series (mean and variance), rather than interpretation of anomaly magnitude in a z-score context. Although this methodology removes biases for total time series comparisons, biases from distinct pre- and post-fire periods remain allowing quantification of bias shifts from pre-to post-fire periods.

3.2. In Situ Snowpack Observations

We use quality-controlled bias-corrected ground observations of SWE processed and bias-corrected by the Pacific Northwest National Laboratory (Sun et al., 2019; Yan et al., 2018) to evaluate snow simulations at the SNOTEL station (Deadwood Summit) in Johnson Creek (blue dot in Figure 1c). Pacific Northwest National Laboratory has made these bias corrected observations publicly available: https://www.pnnl.gov/data-products.

3.3. MODIS Vegetation and Snow Disappearance Observations

Vegetation classifications used in analyses and model simulations are based on the annual 500 m MCD12Q1V6 MODIS data product (https://doi.org/10.5067/MODIS/MCD12Q1.006). Green fraction and LAI come from 8-day 500-m MODIS retrievals (MOD15A2; Myneni et al., 2015). In this study we assume that MODIS-observed Fraction of Photosynthetically Active Radiation (FPAR) is equal to green fraction used by the LSM as done by the widely used WRF Preprocessing System (WPS) which is frequently used to derive Noah-MP LSM inputs (Lu et al., 2021). Using FPAR as a proxy for vegetation fraction in LSM simulations can introduce uncertainties (Filipponi et al., 2018; Myneni & Williams, 1994) so future modeling research should consider using other vegetation cover fraction data products (e.g., DiMiceli et al., 2015; Liang et al., 2021).

We use snow disappearance maps from the Oak Ridge National Laboratory (O'Leary et al., 2019), derived from MODIS observations (MOD10A2), to evaluate simulated snow disappearance. These maps identify the day of year that each MODIS 500-m pixel transitions from snow-covered to no snow conditions during the spring melt season. To account for cloud-obscured images, these data interpolate the dates between the last observed snow cover and the first observed snow-free date using a 50% snow cover area threshold (O'Leary et al., 2018, 2019). The 500-m MODIS data are aggregated to the 1-km model grid prior to model evaluations such that the latest DSD is selected among MODIS pixels which are contained by each model pixel for each calendar year.

4. Results

4.1. Pre-Fire Streamflow Validation

We first evaluate Noah-MP simulations of monthly and annual Q-anomalies in pre-fire years because high accuracy of the modeled runoff in pre-fire years is essential to discern fire-induced modeling errors from other sources of uncertainty in post-fire years. There are robust pre-fire correlations ($\rho = 0.89$–0.92) between simulated and observed Q-anomalies at monthly and annual time scales at both catchments (Figures 2–4). Despite uncertainties that may be attributable to neglecting river-routing in Noah-MP simulations, we proceed under the reasonable assumption that Noah-MP can accurately translate meteorological conditions to monthly and annual Q-anomalies in both catchments in the absence of severe land disturbances. Because all comparisons are conducted in the anomaly space, forcing a zero-bias, the bias analyses focus on changes in bias from pre-to post-fire years as an indication of shifts in model skill. Note, LSM biases in one period (either pre- or post-fire) will impact anomaly biases in both periods (Equation 1). For example, failure to simulate enhanced post-fire streamflow would impose a systematic high Q-anomaly bias in the pre-fire period.

4.2. Comparing Model Skills for Monthly and Annual Streamflow From Pre-Fire To Post-Fire Years

Model simulations of Q-anomalies have a statistically significant ($p < 0.01$) shift in biases from pre-to post-fire years and modestly different correlations in monthly and annual comparisons (Figures 2–4). Simulated Q-anomaly biases tend to be positive in pre-fire years and negative in post-fire years. At Andrew's Creek, pre-fire annual Q-anomaly biases have a mean of 0.34 and interquartile range (IQR) of 0.11–0.55; whereas the corresponding post-fire biases have a mean of −0.33 and an IQR of −0.65 to 0.10 (Figures 2a and 2b). Likewise, pre- and post-fire monthly Q-anomaly biases at Andrew's Creek have means 0.08 and −0.08, respectively (Figure 3b). Similarly at Johnson Creek, pre-fire annual Q-anomaly biases have a mean of 0.22 and IQR of 0.02–0.43; whereas the corresponding post-fire biases have a mean of −0.26 and IQR of −0.36 to −0.16 (Figures 2c and 2d). Pre- and post-fire monthly mean Q-anomaly biases at Johnson Creek are 0.05 and −0.05, respectively (Figure 4b). The shift in annual Q-anomaly biases is more consistent at Johnson Creek where 11 out of the 12 post-fire years have negative biases, whereas 8 out of the 12 post-fire years at Andrew's Creek have negative biases. Q-anomaly overestimates in pre-fire years are at least partially attributable to LSM anomalies being calculated during the entire 25-year analysis period covering both pre- and post-fire years, and thus pre-fire anomaly overestimates are inflated by Noah-MP failing to simulate the observed enhanced runoff in post-fire years. There are minor and inconsistent differences between pre- and post-fire correlations (Figures 2–4), which indicates the LSM has similar skill in simulating the runoff temporal variability during pre- and post-fire years. These comparisons support that fire contributes to an altered relationship between meteorology and land surface hydrology that the model does not capture.

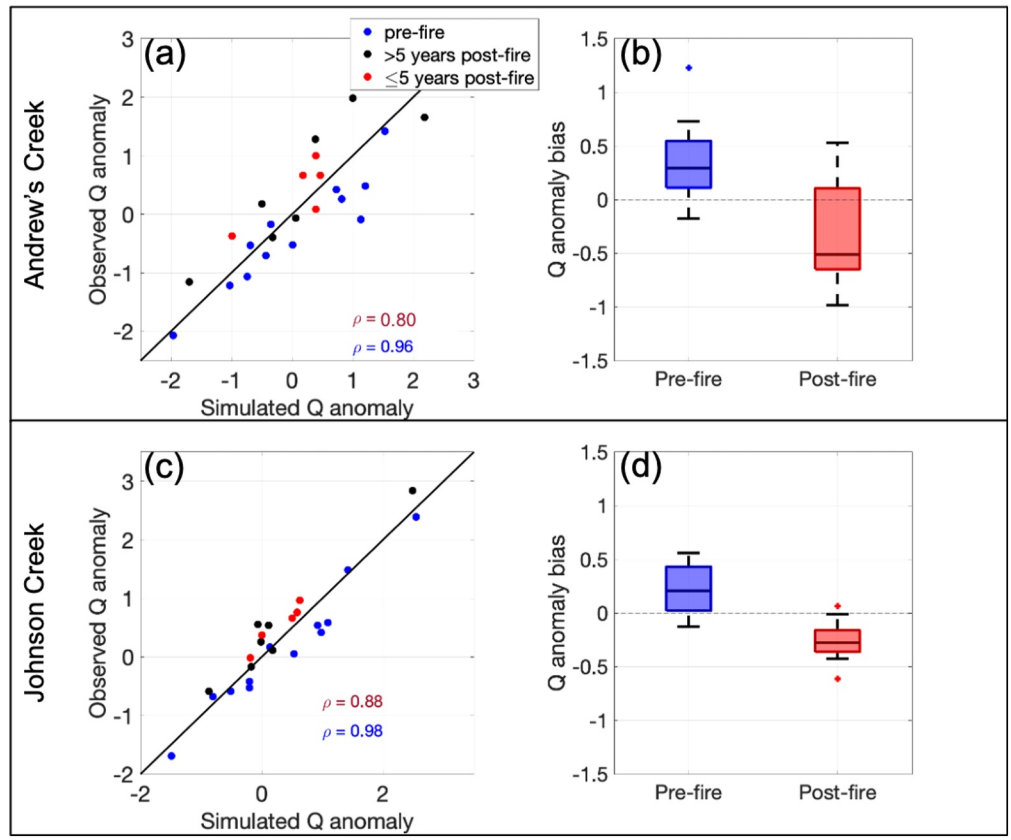

Figure 2. Comparisons of Noah-MP simulated annual catchment runoff (Q) anomalies compared to in situ observed anomalies at (a) and (b) Andrew's Creek and (c) and (d) Johnson Creek. Biases in Q-anomalies (in b, d) are calculated by subtracting observed runoff anomalies from simulated anomalies. Black 1:1 lines in (a) and (c) are provided for reference. For boxplots, central marks indicate the median, and the bottom and top edges of boxes indicate the 25th and 75th percentiles, respectively. Whisker lengths are equal to the interquartile range and outliers are plotted as blue and red dots. Correlations in blue text are for all pre-fire years and those in red text are for all post-fire years.

Monthly and annual Q-anomaly bias shifts are similar from the twin-simulations designed to neglect fire impacts on vegetation (Figures S6–S8 in Supporting Information S1). In the twin no-fire simulation at Andrew's Creek, pre-fire annual Q-anomaly biases have a mean of 0.35 and IQR of 0.13–0.57; whereas post-fire biases have a mean of −0.35 and an IQR of −0.67 to 0.10 (Figure S6 in Supporting Information S1). Likewise, pre- (post-) fire mean monthly Q-anomaly biases at Andrew's Creek are 0.08 (−0.08) in the no-fire simulation (Figure S7b in Supporting Information S1). Similarly, at Johnson Creek, pre-fire annual Q-anomaly biases have a mean of 0.25 and IQR of 0.05–0.47; whereas corresponding post-fire biases have a mean of −0.30 and IQR of −0.40 to −0.20 (Figure S6 in Supporting Information S1). Likewise, pre- (post-) fire mean monthly Q-anomaly biases at Johnson Creek are 0.05 (−0.06) in the no-fire simulation (Figure S8b in Supporting Information S1). Fire-aware simulations are not more skillful in simulating post-fire annual Q-anomalies in terms of correlation than the twin no-fire simulations. These results show that accounting for fire impacts on vegetation in the Noah-MP system as discussed in Section 2.3 only slightly reduces the pre-to post-fire bias shift, and therefore is not an adequate approach to resolve fire-related uncertainties in Noah-MP for these study catchments.

Further analyses support that the systematic shifts in Q-anomaly biases from Noah-MP (Figures 2–4) are mainly attributable to the observed, but unmodeled, enhanced runoff efficiency (i.e., Q/P) after fires (Figures 5a and 5b). Observed increases in mean annual runoff efficiency from pre-to post-fire years is statistically significant ($p \leq 0.01$): a 54% increase at Andrew's Creek (0.40–0.61) and a 22% increase at Johnson Creek (0.49–0.59). However, simulated increases in runoff efficiency are much smaller: a 22% increase at Andrew's Creek ($p = 0.09$) and a 3% increase at Johnson Creek ($p = 0.31$). Simulated increases in runoff efficiency at Andrew's Creek suggests that meteorology contributes to the observed enhanced runoff efficiency, but the model which primarily

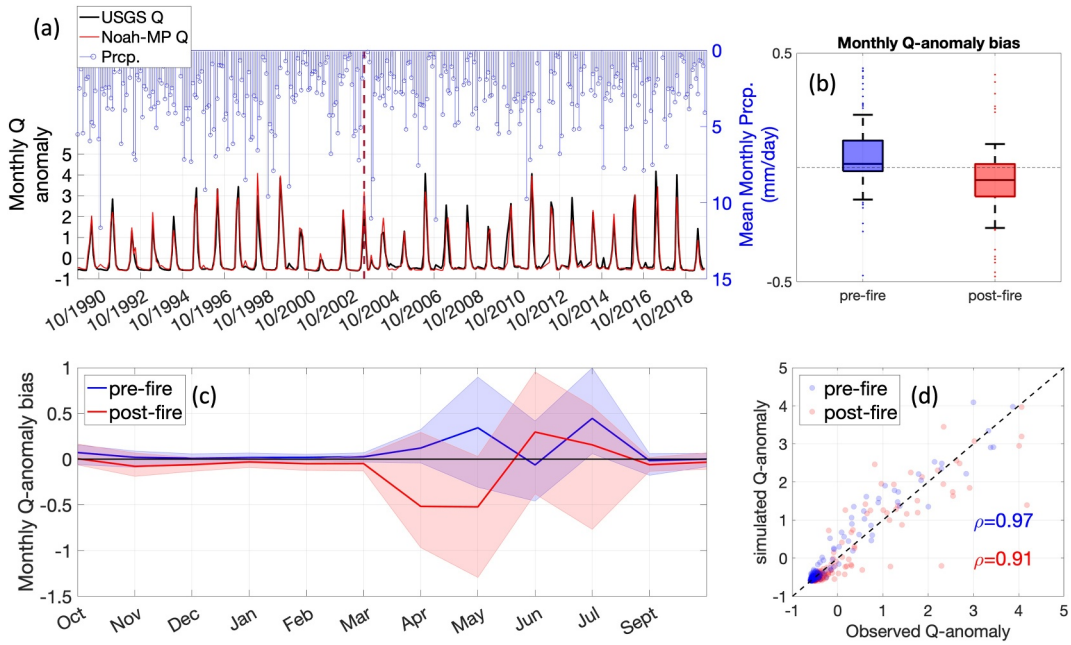

Figure 3. Comparisons of Noah-MP simulated monthly catchment runoff (Q) anomalies compared to in situ observed anomalies at Andrew's Creek. (a) Time series comparison of simulated (red) and observed (black) monthly *Q*-anomalies; vertical dashed red line represents the date of the fire event. Vertical blue lines on top show monthly precipitation. (b) Boxplots of pre- and post-fire monthly *Q*-anomaly biases. (c) Multiyear mean monthly pre- and post-fire *Q*-anomaly biases; shading represents the 80% range from interannual variability. (d) Simulated and observed *Q*-anomalies with correlations; dashed black line represents 1:1 line for reference.

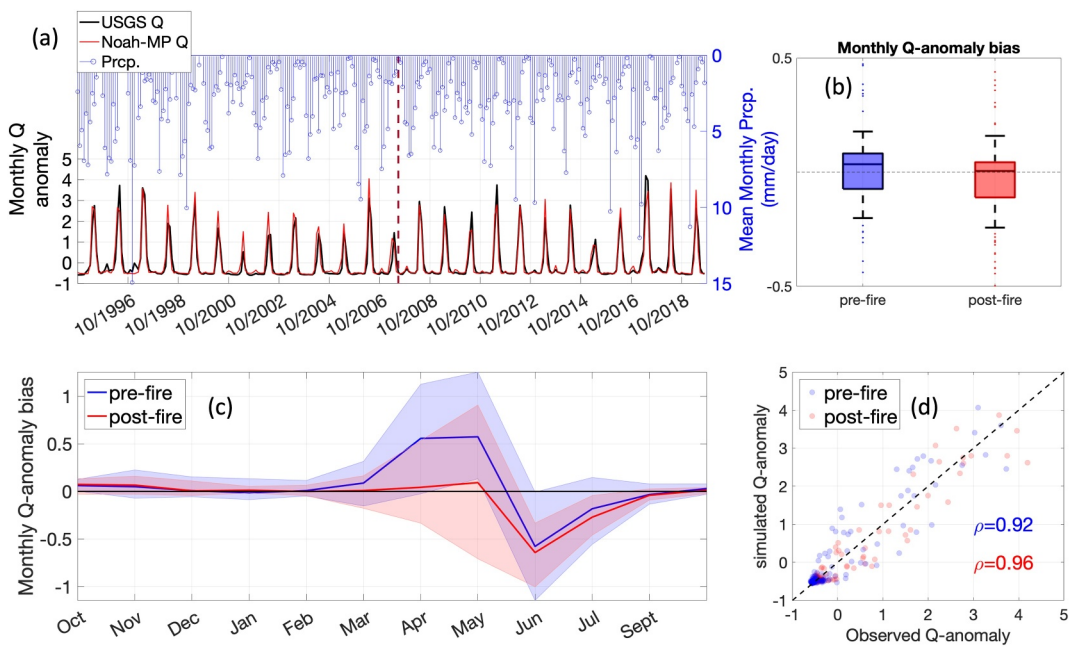

Figure 4. Comparisons of Noah-MP simulated monthly catchment runoff (*Q*) anomalies compared to in situ observed anomalies at Johnson Creek. (a) Time series comparison of simulated (red) and observed (black) monthly *Q*-anomalies; vertical dashed red line represents the date of the fire event. Vertical blue lines on top show monthly precipitation. (b) Boxplots of pre- and post-fire monthly *Q*-anomaly biases. (c) Multiyear mean monthly pre- and post-fire *Q*-anomaly biases; shading represents the 80% range from interannual variability. (d) Simulated and observed *Q*-anomalies with correlations; dashed black line represents 1:1 line for reference.

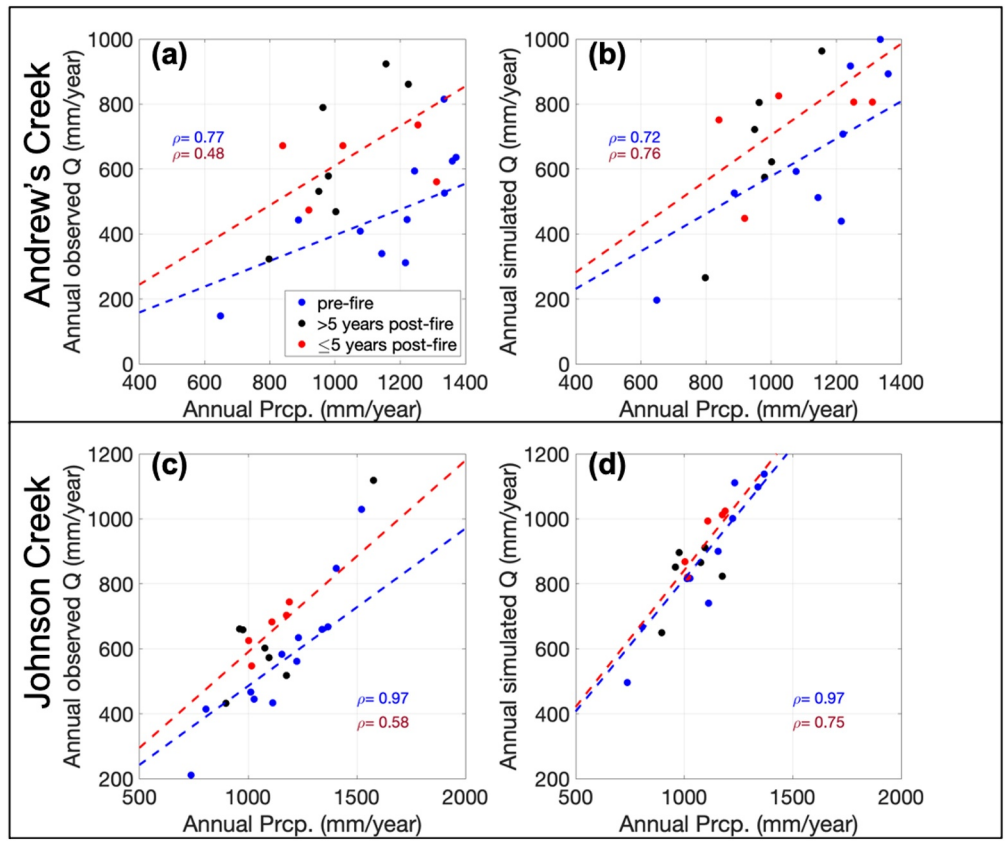

Figure 5. Comparison of annual catchment runoff (Q) and precipitation at (a) and (b) Andrew's Creek and (c) and (d) Johnson Creek. Dashed blue and red lines represent the slopes of pre- and post-fire runoff efficiencies, respectively. Correlations in blue text are for all pre-fire years and those in red text are for all post-fire years.

accounts for meteorologic controls and partially accounts for vegetation disturbances underestimates a large portion of the increase in observed post-fire runoff efficiency at this catchment (Figures 5a and 5b). Similarly, the observed but unmodeled enhancement in post-fire runoff efficiency at Johnson Creek (Figures 5c and 5d) results in underestimates of Q-anomalies in post-fire years (Figures 2c and 5d). Observed correlations between Q and P are reduced during post-fire years at both catchments relative to pre-fire years; however, these reductions in correlation are not statistically significant ($p = 0.40$ and $p = 0.10$ at Andrew's Creek and Johnson Creek, respectively). This indicates that fire-induced land cover changes may cause environmental factors other than precipitation to exert modest increases in control on annual discharge after fire events, which is consistent with prior research that suggests wind and radiation exert a greater control on post-fire annual runoff efficiency in snow-dominated catchments due to enhanced winter sublimation and faster melt (Harpold et al., 2014). Qualitative results from Figure 5 are similar to analyses using the no-fire twin-simulations (Figure S9 in Supporting Information S1), further supporting that this study's approach to account for fire impacts on vegetation in the Noah-MP system is not an adequate approach to resolve fire-related model uncertainties in these study catchments.

4.3. Comparing Simulated and Observed Monthly Runoff Variation

Comparisons between monthly changes in Q' (monthly Q divided by annual P; Section 2.4) from pre-to post-fire years are consistent with previous research that found fires accelerate spring snow melt (Koshkin et al., 2022; Smoot & Gleason, 2021). Namely, at both study catchments, there are observed increases from pre-to post-fire spring Q' which are underestimated in model simulations (Figure 6). Observations show a mean 0.03 increase in Q' at Andrew's Creek and Johnson Creek in April (111% and 79% increase, respectively) and a 0.13 and 0.08 increase in May (97% and 46% increase, respectively) (Figure 6). However, Noah-MP simulates small decreases to April Q' from pre-to post-fire years by 6% and 8% at Andrew's Creek and Johnson Creek, respectively, and

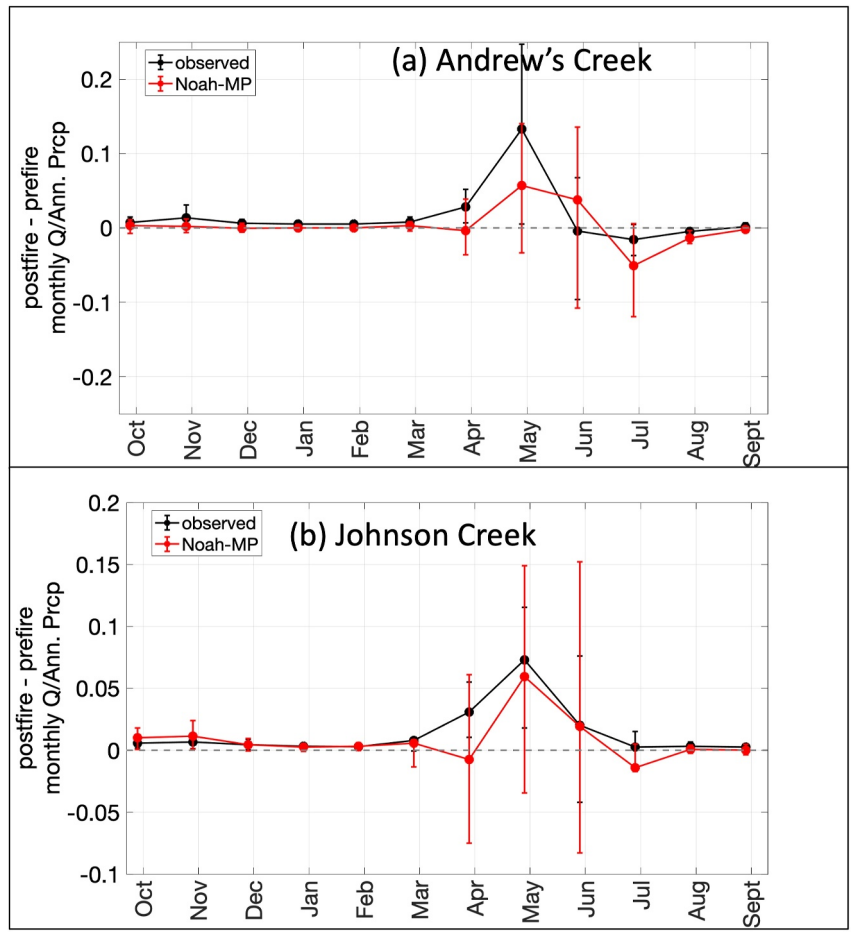

Figure 6. Mean monthly runoff divided by annual precipitation (Q') from (a) Andrew's Creek and (b) Johnson Creek. Solid dots represent the multiyear mean and uncertainty bars represent the interquartile range based on temporal variability across the 12 pre- and post-fire periods.

simulates relatively minor increases to peak Q' in May (16% and 26% at Andrew's Creek and Johnson Creek, respectively). These comparisons are consistent with the monthly bias analyses which show the most substantial shift in Q-anomaly biases occur during spring months (Figures 3c and 4c). Namely, at Andrew's Creek, the largest pre-to post-fire shift in the multiyear mean monthly Q-anomaly biases occurs in April and May when pre-fire Q-anomaly biases range from 0.12 to 0.34 and post-fire Q-anomaly biases are −0.52. Likewise, at Johnson Creek, April-May biases in pre-fire years are much larger (0.56–0.57) than post-fire biases in April–May (0.04–0.09). These results support that LSM simulations are missing key processes related to enhanced post-fire spring runoff, although there is substantial temporal variability in the results (Figure 6).

Differences between observed and simulated peak Q timing support that LSM simulations are more prone to estimating too-late peak Q in post-fire years (Figure S11 in Supporting Information S1), which is consistent with observed, but unmodeled, fire-induced acceleration of snowmelt (Section 4.4). At Andrew's Creek, observations show post-fire peak monthly Q in May during 3 years when Noah-MP simulates later peak Q in June, while Noah-MP accurately simulates peak Q timing in all other post-fire years. During all pre-fire years at Andrew's Creek, Noah-MP accurately simulates the month of peak Q except in 2000 when Noah-MP simulated too-early peak Q. Likewise, at Johnson creek, observations show peak Q in May in 1 year when Noah-MP simulates later peak Q in June, while accurately simulating peak Q timing in all other post-fire years. This is counter to the Noah-MP pre-fire bias at Johnson Creek, where the simulations show too-early peak Q in 2 years, with matching month of peak Q for all other pre-fire years. This indicates that although the Noah-MP LSM has a slight early peak Q bias in pre-fire years at both study catchments, Noah-MP has a greater tendency to simulate too-late peak Q in post-fire years.

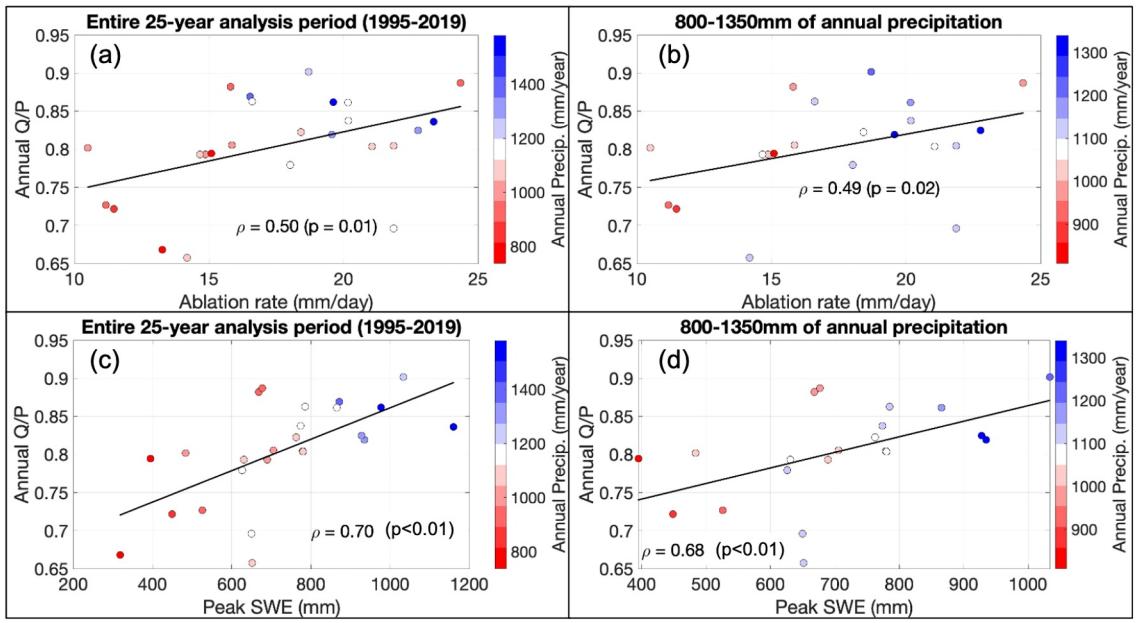

Figure 7. Comparisons of simulated annual runoff efficiency (*Q*/*P*) with (a) and (b) simulated ablation rates and (c) and (d) simulated peak snow water equivalent at Johnson Creek. Dots are colored by annual precipitation, and black lines represent the best-fit line. Relationships between variables across all 25-years in the analysis are shown in (a) and (c). Relationships between variables during a shortened window with less variable precipitation (800–1,350 mm; 21-years) are shown in (b) and (d).

4.4. Evaluating Consistency Between Runoff Biases With Snowpack Biases

We hypothesize that a large contributing factor to the observed post-fire enhanced *Q*/*P* and spring *Q*′ is that reduced vegetation favors deeper snowpack that melts out faster. Correlations between Noah-MP simulated annual *Q*/*P* with peak SWE and ablation rates are used to determine if the aforementioned hypothesis is consistent with relationships depicted by the sophisticated physically-based model. Indeed, Noah-MP simulates these linkages—greater volumes of snow that melt faster favor higher runoff efficiency—which supports that the hypothesis is physically consistent with state-of-the-art physical hydrologic modeling in these catchments. Specifically, Noah-MP simulates significant and positive correlations between *Q*/*P* with ablation rates and peak snowpack volume (i.e., peak SWE) at both study catchments ($\rho = 0.43$–0.80; $p < 0.05$; Figures 7a, 7c, 8a, and 8c). To evaluate the robustness of these correlations, we consider that interpretations of these comparisons across the entire 25-year analysis period may be convoluted by annual precipitation having significant correlations with all the variables compared: ablation rates, peak SWE and annual *Q*/*P* ($\rho = 0.48$–0.87 ($p \leq 0.02$) at Johnson Creek and $\rho = 0.29$–0.73 ($p = 0.16$ to <0.001) at Andrew's Creek). Therefore, we also quantify these correlations between *Q*/*P* with peak SWE and ablation rates during years when precipitation is less variable (800–1,350 mm), while still maintaining 68%–84% of the analysis period (17–21 years) (Figures 7b, 7d, 8b, and 8d). In this narrowed window, annual precipitation has insignificant relationships with *Q*/*P* ($\rho = 0.23$–0.32; $p > 0.14$) while *Q*/*P* remains positively correlated with peak SWE and ablation rates at Johnson Creek ($\rho = 0.49$–0.68; $p < 0.05$). In this narrowed window, *Q*/*P* is significantly correlated with peak SWE at Andrew's Creek ($\rho = 0.73$; $p < 0.01$), but not significantly correlated with ablation rates ($\rho = 0.24$; $p = 0.34$). However, at Andrew's Creek, there is an insignificant relationship between ablation rates and annual precipitation across the total 25-year analysis period ($\rho = 0.29$; $p = 0.16$), which supports that the significant relationship between *Q*/*P* and ablation rates during this same 25-year period ($\rho = 0.43$; $p = 0.03$) is likely linked to a connection between ablation rates with *Q*/*P* at this catchment.

Next, we evaluate if Noah-MP has shifts in snow ablation biases that are consistent with the hypothesis noted in the paragraph above. We use catchment-wide comparisons of snow disappearance between Noah-MP with MODIS snow cover observations to determine if there are significant shifts in Noah-MP biases for simulated DSD from pre-to post-fire years, relative to MODIS observations (Figures 9 and 10). In pre-fire years, Noah-MP simulates a persistent early-melt bias relative to satellite observations at both catchments (median underestimates = 11–14-day) (Figures 9b and 10b). These pre-fire biases are consistent with Section 4.3 and

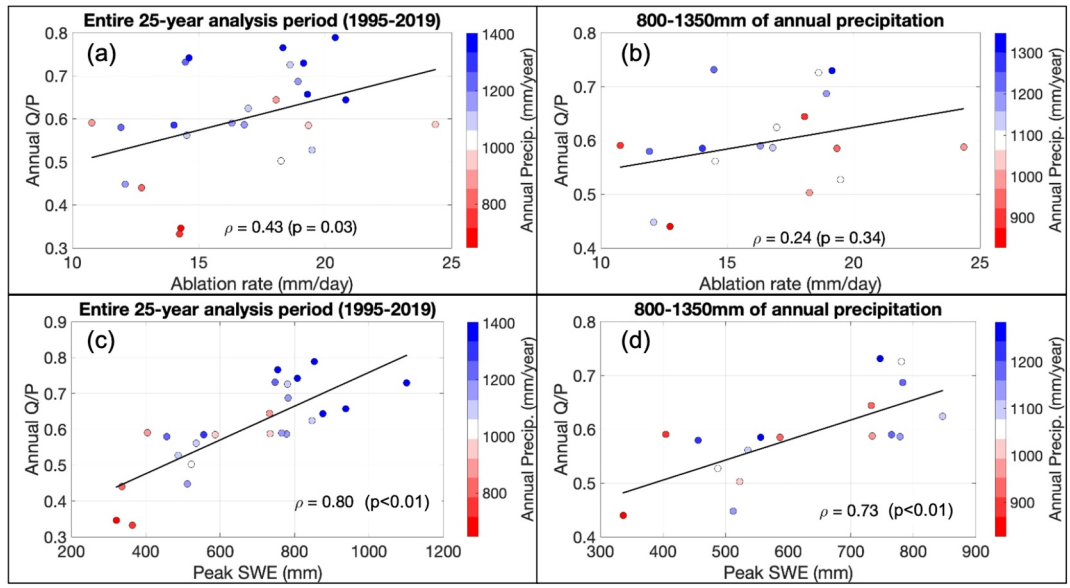

Figure 8. Comparisons of simulated annual runoff efficiency (*Q*/*P*) with (a) and (b) simulated ablation rates and (c) and (d) simulated peak snow water equivalent at Andrew's Creek. Dots are colored by annual precipitation, and black lines represent the best-fit line. Relationships between variables across all 25-years in the analysis are shown in (a) and (c). Relationships between variables during a shortened window with less variable precipitation (800–1,350 mm; 17-years) are shown in (b) and (d).

previous research which found Noah-MP underestimates peak SWE and melts snow too-early over the western US (He et al., 2019) due to the model deficiencies in representing vegetation-snow-radiation/turbulence interactions (He et al., 2021). In post-fire years, the too-early DSD bias is reduced (median underestimates = 4–10.5-day). This is consistent with observations capturing fire-induced snowmelt acceleration which Noah-MP does not accurately simulate, which happens to compensate for a portion of pre-fire early melt biases in Noah-MP in these catchments. The twin no-fire simulations show later post-fire DSD than simulations informed

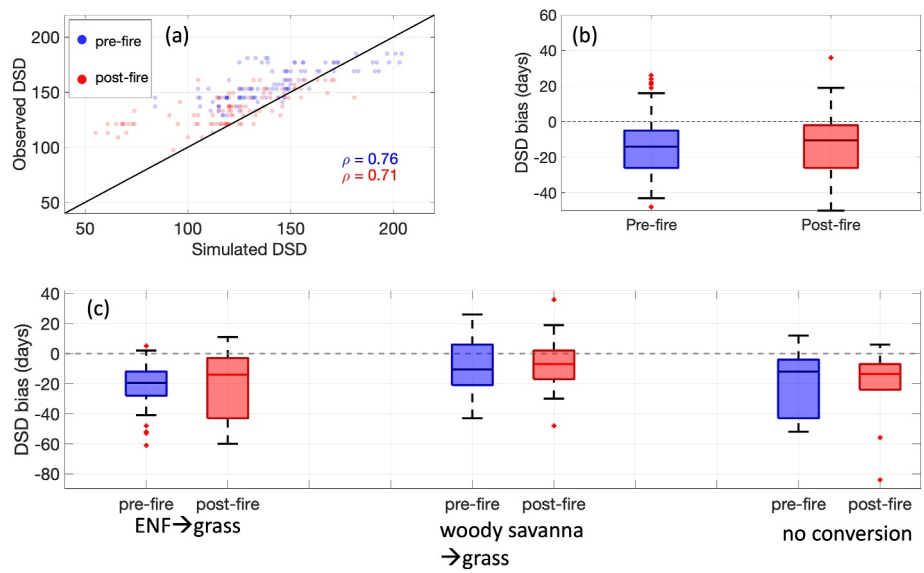

Figure 9. Comparison of simulated and remotely sensed day of snow disappearance (DSD) for each 1-km burned pixel in Andrew's Creek. (a) Scatter plot comparing MODIS-observed and simulated DSD based on the calendar year at Andrew's Creek across all burned pixels. (b) Biases in DSD in pre- and post-fire years at Andrew's Creek across all burned pixels. (c) Same as (b), but data are separated based on MODIS observed pre-to post-fire landcover conversions.

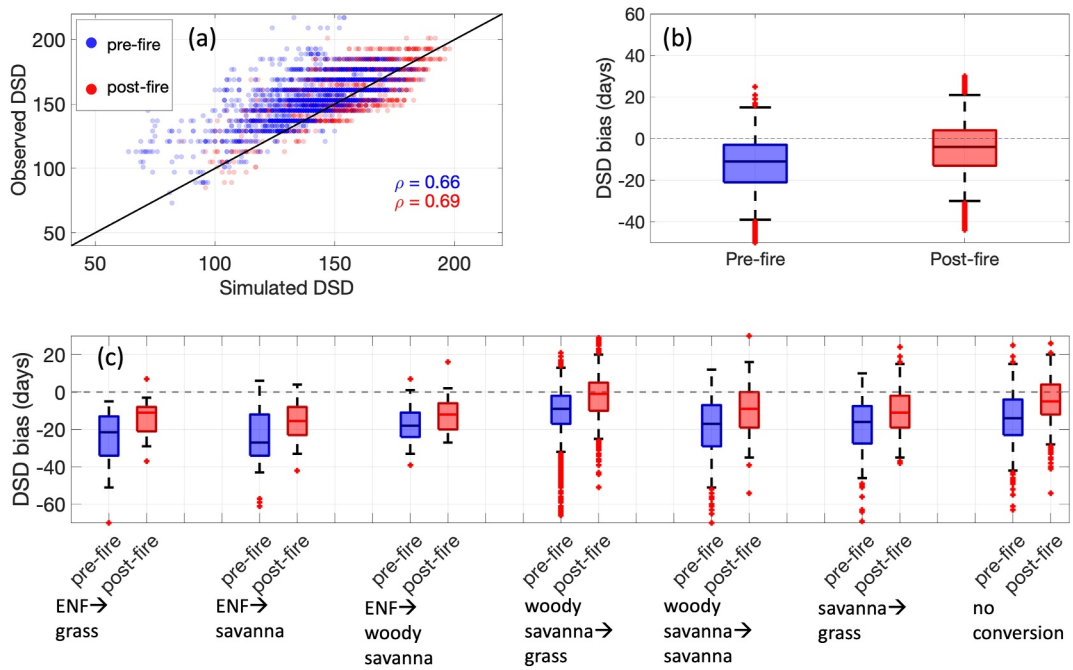

Figure 10. Same as Figure 9 but for Johnson Creek. (a) Scatter plot comparing MODIS-observed and simulated day of snow disappearance (DSD) based on the calendar year at Johnson Creek across all burned pixels. (b) Biases in DSD in pre- and post-fire years at Johnson Creek across all burned pixels. (c) Same as (b), but data are separated based on MODIS observed pre-to post-fire landcover conversions.

with post-fire vegetation changes by an average of 3-day at Johnson Creek and 1-day at Andrew's Creek (Figures S12b and S12e in Supporting Information S1). This indicates that informing the LSM with fire-induced changes to green fraction, LAI and classifications accelerates post-fire snow disappearance in the model; however, the aforementioned shift in DSD biases support that further model enhancements are required to comprehensively capture the associated fire-vegetation-snow interactions.

In Andrew's Creek, pixels which experience land cover conversions, particularly conversions from forested land to grassland, tend to experience greater impacts associated with observed, but unmodeled, post-fire acceleration of ablation. Namely, the too-early DSD bias tends to be reduced from pre-to post-fire years by 5.5- and 3.5-days across burned pixels which have observed pre-to post-fire land cover conversions from evergreen need leaf forest (ENF) to grasslands and from woody savanna to grasslands, respectively (Figure 9c). Conversely, burned pixels which do not have observed land cover conversions at Andrew's Creek tend to have a small increase to the negative bias magnitude (by 1.5-day). This result is mostly consistent with analysis of the no-fire simulation from Andrew's Creek (Figures S12a–S12c in Supporting Information S1), except the pre-to post-fire bias change is largest over pixels converted from woody savanna to grasslands rather than pixels converted from ENF to grasslands. The median pre-to post-fire bias change over pixels converted from woody savanna to grassland are twice as large in the fire blind simulation relative to the simulation informed with fire-impacted vegetation characteristics which indicates that informing the Andrew's Creek simulation with post-fire vegetation data helps the model capture post-fire snow ablation acceleration, particularly across pixels which experienced the observed land cover conversion from woody savanna to grassland.

In Johnson Creek, pixels which experienced land cover conversions from ENF to grassland or savannas experienced marginally larger impacts associated with observed, but unmodeled, post-fire acceleration of ablation relative to other burned pixels. Namely, the too-early DSD model bias tends to be reduced from pre-to post-fire years by 10.5- and 11.5-days over burned pixels which were converted from ENF to grassland and savannas, respectively. The too-early DSD bias tends to be reduced from pre-to post-fire years by 6–8-days across other burned pixels in the Johnson Creek catchment where land cover conversions were observed. Burned pixels at Johnson Creek which did not have observed land cover classification conversions—consisting of woody savanna, savanna and grassland classifications in pre- and post-fire years—tend to have the too-early DSD bias reduced by

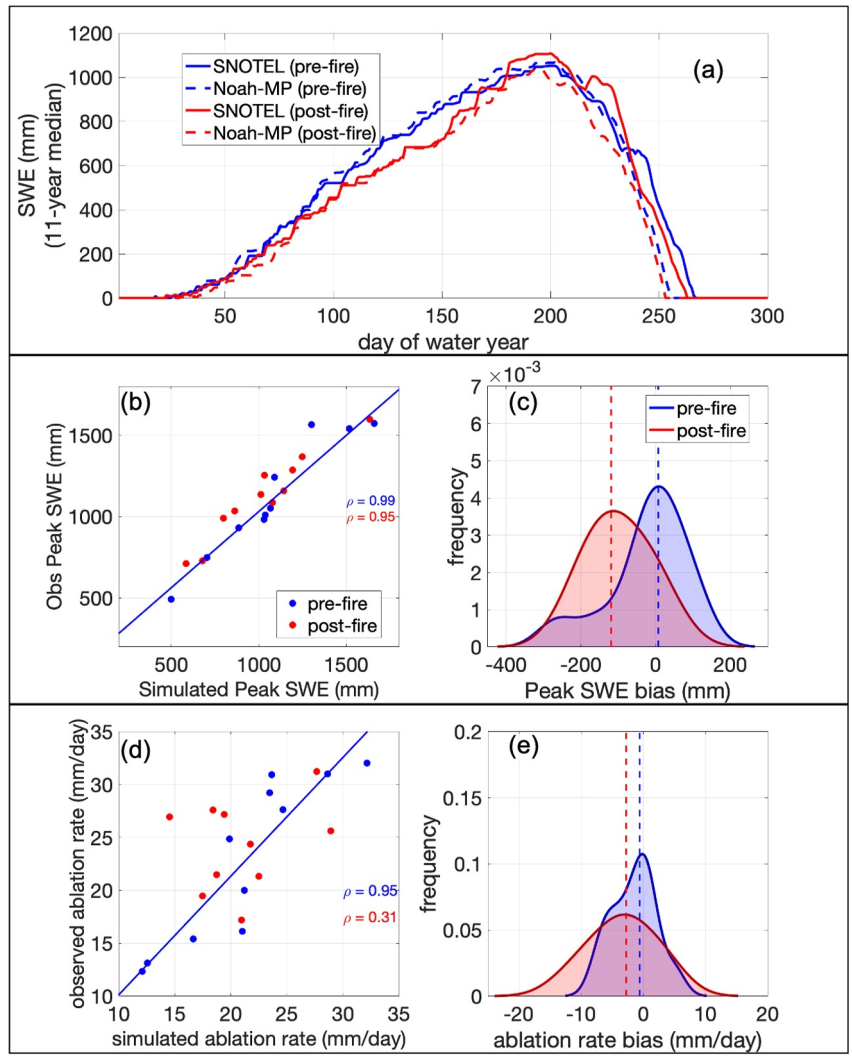

Figure 11. Comparison of simulated and observed snow water equivalent (SWE) at the Deadwood SNOTEL station in Johnson Creek. (a) Mean multiyear SWE for 11 pre- and post-fire years from observations and Noah-MP simulations. (b) Scatter plot comparison of simulated and observed peak SWE during pre- and post-fire years. (c) Kernel density estimator for peak SWE bias during pre- and post-fire years. (d) Scatter plot comparison of simulated and observed ablation rates. (e) Kernel density estimator for ablation rate bias. Vertical blue and red lines in (c) and (e) represent medians of pre- and post-fire distributions, respectively.

9-days from pre-to post-fire years, which falls within the range of pre-to post-fire bias shifts across burned pixels where land cover conversions were observed. This indicates that fire impacts on Noah-MP skill in simulating DSD are generally spatially consistent at Johnson Creek catchment, with little sensitivity to land cover classification conversions. This finding is consistent with the corresponding analysis of the no-fire simulation (Figures S12d–S12f in Supporting Information S1) which also shows the too-early DSD bias is reduced by 9-days from pre-to post-fire years across pixels where no land cover conversion was observed.

Finally, we compare simulated and in situ observed snowpack at the Deadwood SNOTEL station in Johnson Creek. Noah-MP has degraded skill in simulating snow accumulation and spring ablation in post-fire years, with underestimates of peak SWE and spring ablation rates (Figure 11a). In pre-fire years, Noah-MP accurately simulates peak SWE ($\rho = 0.99$ and bias = −22 mm (−2%)) (Figures 11b and 11c). In post-fire years, Noah-MP maintains accurate representation of peak SWE interannual variability ($\rho = 0.95$) but has larger underestimates of peak SWE magnitude (bias = −99 mm (−9%)). Observations show a lower multiyear average peak SWE in pre-fire years (1,052 mm) than post-fire years (1,107 mm), whereas Noah-MP shows the opposite: a decrease from

pre-fire peak SWE of 1,067 mm to post-fire value of 1,033 mm. This supports that fire impacts favor greater snow accumulation in post-fire years at this location, an effect which the simulation informed with post-fire vegetation characteristics failed to capture. Noah-MP also has degraded skill in simulating post-fire ablation rates, with pre- and post-fire correlations of 0.95 and 0.31, respectively (Figure 11d). Furthermore, observations show higher post-fire ablation rates compared to pre-fire years, resulting in a 4-day earlier post-fire SWE disappearance despite greater observed SWE accumulation during post-fire years (Figure 11a). Noah-MP fails to simulate the enhanced post-fire ablation, resulting in a larger model underestimate in mean ablation rates in post-fire years (bias of −3.4 mm/day) relative to pre-fire years (bias of −1.5 mm/day); however, there is substantial variability in the results with the model overestimating ablation rates in some pre- and post-fire years (Figure 11e). Results from Figure 11 are consistent with analyses using the twin-simulation designed to neglect fire impacts on vegetation (Figure S13 in Supporting Information S1) which indicates that updating vegetation classifications, LAI and green fraction model inputs to reflect fire impacts does not resolve fire-induced snow modeling uncertainties in Noah-MP at this SNOTEL site. This is partially expected because the Deadwood SNOTEL site is in a canopy gap, and thus the bulk of the vegetated-related fire impacts to the snowpack at this site are likely from interactions with the neighboring forested area (e.g., interception, wind patterns, and ash deposition onto snow) which is not captured by Noah-MP.

5. Implications for Future Model Improvement

Fire-aware simulations informed with satellite-observed fire-effects on vegetation classifications, green fraction and LAI have similar skill in simulating post-fire runoff and snowpack as the no-fire twin-simulations that do not account for fire effects on these vegetation inputs across the study catchments. Hence, Noah-MP uncertainties associated with simulating post-fire land surface hydrology in these snow-dominated and heavily burned catchments are not suitably reduced by informing simulations with the satellite observed vegetation data using the modeling approach presented in Section 2.3. Namely, Noah-MP has unresolved uncertainties in representing fire-snow-hydrology interactions which need to be improved to accurately reflect fire impacts on runoff and snowpack. This finding is supported by streamflow and snowpack comparisons which reveals the LSM fails to simulate the observed increase in ground snow accumulation and enhanced ablation rates post-fire (Figures 9–11) which is physically and statistically consistent with errors in simulating enhanced post-fire runoff efficiency, enhanced spring runoff and earlier peak runoff (Figures 2–8; Figure S11 in Supporting Information S1). Because fire impacts on snow are physically consistent with vegetation disturbances, rather than soil disturbances, these findings support that future modeling efforts aimed at improving Noah-MP skill in simulating post-fire hydrology in snow-dominated areas may benefit from using different physics options (e.g., for dynamic vegetation) and incorporating model developments that more accurately represent interactions between vegetation and snowpack. This is further supported by the supplementary analyses in Figures S14 of Supporting Information S1 which shows correcting post-fire model SWE to be consistent with in situ observed SWE (which implicitly observes fire impacts) can substantially reduce the shift in systematic biases in annual Q-anomalies at Johnson Creek. These results are consistent with previous research which found interactions between snow, meteorology, and vegetation are important sources of uncertainty in Noah-MP snow simulations over undisturbed landscapes (Abolafia-Rosenzweig et al., 2021; Abolafia-Rosenzweig, He, McKenzie Skiles, et al., 2022; Chen et al., 2014; He et al., 2019, 2021). We note that explicitly evaluating error sources from fire disturbances on vegetation and soil separately is out of the scope of this analysis, and thus the relative contributions from these sources of uncertainty remains unknown.

Despite only minor differences between fire-aware and no-fire simulations in this study; there are recent studies which found physically-based model sensitivities to fire-adjusted vegetation and soil properties in runoff simulations (Atchley et al., 2018; Maina & Siirila-Woodburn, 2020; Wang et al., 2020). More distinct differences from pre-to post-fire simulated hydrology reported in these previous studies, relative to those reported herein, are likely attributable to differing study domains and potentially due to differences in land model physics. For example, failure to simulate enhanced post-fire streamflow by Noah-MP in this study is partially due to compensatory effects of vegetation reductions on ET, which is domain specific. Namely, in both study catchments, simulations that account for fire impacts on vegetation show reductions in canopy evaporation and transpiration, relative to no-fire simulations; however, these reductions are mostly accounted for by increases in bare ground evaporation (Figure S15 in Supporting Information S1). Thus, additional research using Noah-MP in other fire-effected regions across a range of climates, elevations and vegetations are required to determine the

extent of which the results presented herein are applicable to a diverse array of fire-effected areas. For instance, Atchley et al. (2018) and Wang et al. (2020) examine simulated fire impacts over catchments that do not receive heavy snowfall, and thus inter-study model results are not directly comparable. However, Maina and Siirila-Woodburn (2020) show that accounting for fire impacts on vegetation yields greater winter snowpack accumulation in ParFlow—Community Land Model (Community Land Model) simulations, a fire impact which Noah-MP does not accurately show in this study. This may be attributable to differences between the ParFlow-CLM and the Noah-MP model physics configuration used herein; however, inter-model comparisons that are out of the scope of this analysis are required to determine specific land model differences in simulating fire-related sensitivities to land cover disturbance.

Our analysis did not account for fire perturbations to soil and runoff parameters (e.g., saturated hydraulic conductivity, porosity, residual soil moisture content, field capacity and surface roughness) which are important to consider when modeling fire impacts on hydrology. The choice to not perturb these parameters in this study was made because fire impacts on snowpack evolution are likely dominated by vegetation impacts, which was a large focus of this analysis. Updating soil and runoff parameters would not noticeably impact post-fire snow model deficiencies presented herein which contributed largely to the model runoff uncertainties in this study. However, fire-impacts on soil may play an important role in post-fire hydrology in these catchments, particularly at timescales finer than those considered herein (e.g., daily), and thus will be important to evaluate in future modeling studies that employ river routing. Augmenting Noah-MP soil and runoff parameters to reflect fire impacts may improve model accuracy (i.e., by enhancing streamflow), but until vegetation-related uncertainties are resolved, the associated accuracy increases are likely to be a result of compensatory errors.

Based on the results from this study, we suggest future efforts to improve post-fire simulated hydrology with Noah-MP should attempt to resolve vegetation-related uncertainties that allow the LSM to have higher sensitivity to fire perturbations. Then, it is necessary to quantify sensitivities of simulations to model parameters that are altered by fires, including: (a) soil hydraulic conductivity, (b) soil bulk density, (c) infiltration/runoff partitioning, (d) LAI, (e) green fraction, (f) vegetation classifications, (g) vegetation height, and (h) snow albedo aging and impurity parameters. Adjustments to soil hydrologic properties can be informed by observationally-based formulas that relate fire severity to post-fire hydrologic properties (e.g., Ebel, 2020). Fire-induced changes to vegetation classifications, LAI and green fraction are remotely sensed and can be used as direct inputs to Noah-MP simulations. Noah-MP vegetation height is inferred from a look-up table based on vegetation classifications; however, following fires these classification-based parameters should be considered adjustable. Finally, post-fire snow albedo enhanced darkening (Gleason et al., 2013, 2019), should be accounted for either implicitly (e.g., using an empirical increase in snow darkening with the BATS or CLASS scheme) or explicitly (e.g., with the SNICAR model which is being coupled with Noah-MP).

This analysis focused on two case studies where fire burned most of the study catchment areas in a single event. Thus, although Noah-MP fire-related uncertainties reported herein were distinguishable from other sources of uncertainty, additional analysis beyond this study is required to determine if catchments with smaller burned areas have fire signals in uncertainty analyses. It is likely that a threshold of burn area and severity exists in which fire-induced uncertainties are masked by other uncertainty sources. Therefore, we note value in future research that explores how large and severe a catchment must be burned in order for LSM uncertainties from fire to be noticeable and significant. Furthermore, future model development activities aimed toward accurately modeling post-fire hydrology should consider Andrew's Creek and Johnson Creek as a starting point for analysis but should use data from other fire-impacted catchments as well.

6. Conclusions

We evaluated a widely-used state-of-the-art LSM (Noah-MP) in runoff and snowpack simulations at two representative fire-affected snow-dominated Pacific Northwest catchments: Andrew's Creek in Washington and Johnson Creek in Idaho. These two catchments are selected across all western U.S. fire-affected catchments, because they are snow-dominated and experienced more than 50% burning in a single fire event with minimal burning outside of this event, allowing analyses of distinct pre- and post-fire periods. Noah-MP simulations have a statistically significant shift to model underestimates in monthly and annual runoff anomalies from pre-to post-fire years at both catchments. Simulations underestimate the observed increases in runoff efficiency from pre-to post-fire years, resulting in systematic underestimates of post-fire runoff anomalies. Monthly model evaluations

revealed key uncertainties in post-fire simulations including observed but unmodeled: increases to spring flow, earlier peak flows, accelerated ablation rates and increased snow accumulation. We conclude that a large contributing factor to the observed but unmodeled enhanced annual Q/P and spring flows are partially attributable to the LSM failing to simulate deeper snowpacks that melt faster in post-fire years. In this study, simulations designed to include fire-effects on vegetation characteristics (classification, green fraction and LAI) performed similarly to simulations designed to neglect fire effects on vegetation, indicating that additional model enhancements or physics option configurations aimed at increasing LSM sensitivity to fire-perturbations to vegetation over snow-dominated areas are likely a valuable research endeavor.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

Noah-MP input and output data used to support the findings of this study are publicly available on Mendeley Data (Abolafia-Rosenzweig & He, 2023). The widely-used Noah-MP land surface modeling system, which is used in this study, is publicly available on GitHub here: https://github.com/NCAR/hrldas (He et al., 2023).

Acknowledgments
We thank Dr. A. Park Williams and Tyler B. Hampton for helpful discussions and data sharing. Ronnie Abolafia-Rosenzweig and Cenlin He were supported by NOAA's Climate Program Office's Modeling, Analysis, Predictions, and Projections Program (MAPP), Grant NA20OAR4310421. The NSF National Center for Atmospheric Research is a major facility sponsored by the National Science Foundation (NSF) under Cooperative Agreement #1852977. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Abatzoglou, J. T., Battisti, D. S., Williams, A. P., Hansen, W. D., Harvey, B. J., & Kolden, C. A. (2021). Projected increases in western US forest fire despite growing fuel constraints. *Communications Earth & Environment*, *2*(1), 227. https://doi.org/10.1038/s43247-021-00299-0

Abolafia-Rosenzweig, R., & He, C. (2023). Evaluating Noah-MP simulated runoff and snowpack in heavily burned Pacific-Northwest snow-dominated catchments [Dataset]. *Mendeley Data, V1*. https://doi.org/10.17632/3vpjk28z39.1

Abolafia-Rosenzweig, R., He, C., Burns, S. P., & Chen, F. (2021). Implementation and evaluation of a unified turbulence parameterization throughout the canopy and roughness sublayer in Noah-MP snow simulations. *Journal of Advances in Modeling Earth Systems*, *13*(11), e2021MS002665. https://doi.org/10.1029/2021MS002665

Abolafia-Rosenzweig, R., He, C., & Chen, F. (2022). Winter and spring climate explains a large portion of interannual variability and trend in western U.S. Summer fire burned area. *Environmental Research Letters*, *17*(5), 054030. https://doi.org/10.1088/1748-9326/ac6886

Abolafia-Rosenzweig, R., He, C., McKenzie Skiles, S., Chen, F., & Gochis, D. (2022). Evaluation and optimization of snow albedo scheme in Noah-MP land surface model using in situ spectral observations in the Colorado Rockies. *Journal of Advances in Modeling Earth Systems*, *14*(10), e2022MS003141. https://doi.org/10.1029/2022MS003141

Atchley, A. L., Kinoshita, A. M., Lopez, S. R., Trader, L., & Middleton, R. (2018). Simulating surface and subsurface water balance changes due to burn severity. *Vadose Zone Journal*, *17*, 1–13. https://doi.org/10.2136/vzj2018.05.0099

Bond-Lamberty, B., Peckham, S. D., Gower, S. T., & Ewers, B. E. (2009). Effects of fire on regional evapotranspiration in the central Canadian boreal forest. *Global Change Biology*, *15*(5), 1242–1254. https://doi.org/10.1111/j.1365-2486.2008.01776.x

Burles, K., & Boon, S. (2011). Snowmelt energy balance in a burned forest plot, Crowsnest Pass, Alberta, Canada. *Hydrological Processes*, *25*(19), 3012–3029. https://doi.org/10.1002/hyp.8067

Campbell, W. G., & Morris, S. E. (1988). Hydrologic response of the Pack river, Idaho, to the sundance fire. *Northwest Science*, *62*(4), 165–170.

Chen, F., Barlage, M., Tewari, M., Rasmussen, R., Jin, J., Lettenmaier, D., et al. (2014). Modeling seasonal snowpack evolution in the complex terrain and forested Colorado headwaters region: A model intercomparison study. *Journal of Geophysical Research: Atmospheres*, *119*(24), 13795–13819. https://doi.org/10.1002/2014JD022167

Chen, F., & Dudhia, J. (2001). Coupling an advanced land surface–hydrology model with the Penn state–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Monthly Weather Review*, *129*(4), 17–585. https://doi.org/10.1175/1520-0493(2001)129<0569:caalsh>2.0.co;2

Chen, F., Janjić, Z., & Mitchell, K. (1997). Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP mesoscale Eta model. *Boundary-Layer Meteorology*, *85*(3), 391–421. https://doi.org/10.1023/A:1000531001463

Chen, F., Mitchell, K., Schaake, J., Xue, Y., Pan, H.-L., Koren, V., et al. (1996). Modeling of land surface evaporation by four schemes and comparison with FIFE observations. *Journal of Geophysical Research*, *101*(D3), 7251–7268. https://doi.org/10.1029/95JD02165

Cuntz, M., Mai, J., Samaniego, L., Clark, M., Wulfmeyer, V., Branch, O., et al. (2016). The impact of standard and hardcoded parameters on the hydrologic fluxes in the Noah-MP land surface model. *Journal of Geophysical Research: Atmospheres*, *121*(18), 10676–10700. https://doi.org/10.1002/2016JD025097

Daly, C., Halbleib, M., Smith, J. I., Gibson, W. P., Doggett, M. K., Taylor, G. H., et al. (2008). Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology*, *28*(15), 2031–2064. https://doi.org/10.1002/joc.1688

DiMiceli, C., Carroll, M., Sohlberg, R., Kim, D., Kelly, M., & Townshend, J. (2015). MOD44B MODIS/Terra vegetation continuous fields yearly L3 global 250m SIN grid V006 [Dataset]. *NASA EOSDIS Land Processes Distributed Active Archive Center*. https://doi.org/10.5067/MODIS/MOD44B.006

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*, 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*, *31*(21), 3682–3696. https://doi.org/10.1002/hyp.11288

Eidenshink, J., Schwind, B., Brewer, K., Zhu, Z.-L., Quayle, B., & Howard, S. (2007). A project for monitoring trends in burn severity. *Fire Ecology*, *3*(1), 3–21. https://doi.org/10.4996/fireecology.0301003

Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., et al. (2003). Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *Journal of Geophysical Research*, *108*(D22), 2002JD003296. https://doi.org/10.1029/2002JD003296

Fall, G., Kitzmiller, D., Pavlovic, S., Zhang, Z., Patrick, N., St. Laurent, M., et al. (2023). The Office of water prediction's analysis of Record for calibration, version 1.1: Dataset description and precipitation evaluation. *Journal of the American Water Resources Association*, *59*(6), 1246–1272. https://doi.org/10.1111/1752-1688.13143

Filipponi, F., Valentini, E., Nguyen Xuan, A., Guerra, C. A., Wolf, F., Andrzejak, M., & Taramelli, A. (2018). Global MODIS fraction of green vegetation cover for monitoring abrupt and gradual vegetation changes. *Remote Sensing*, *10*(4), 653. https://doi.org/10.3390/rs10040653

Gleason, K. E., McConnell, J. R., Arienzo, M. M., Chellman, N., & Calvin, W. M. (2019). Four-fold increase in solar forcing on snow in western U.S. burned forests since 1999. *Nature Communications*, *10*(1), 2026. https://doi.org/10.1038/s41467-019-09935-y

Gleason, K. E., Nolin, A. W., & Roth, T. R. (2013). Charred forests increase snowmelt: Effects of burned woody debris and incoming solar radiation on snow ablation: Charred forests increase snowmelt. *Geophysical Research Letters*, *40*(17), 4654–4661. https://doi.org/10.1002/grl.50896

Gochis, D. J., Barlage, M., Dugger, A., FitzGerald, K., Karsten, L., McAllister, M., et al. (2020). *The WRF-Hydro® modeling system technical description, (Version 5.2.0)* (pp. 108). NCAR Technical Note. Retrived from https://ral.ucar.edu/sites/default/files/public/projects/wrf-hydro/technical-description-user-guide/wrf-hydrov5.2technicaldescription.pdf

Goeking, S. A., & Tarboton, D. G. (2020). Forests and water yield: A synthesis of disturbance effects on streamflow and snowpack in western coniferous forests. *Journal of Forestry*, *118*(2), 172–192. https://doi.org/10.1093/jofore/fvz069

Goeking, S. A., & Tarboton, D. G. (2022). Variable streamflow response to forest disturbance in the western US: A large-sample hydrology approach. *Water Resources Research*, *58*(6), e2021WR031575. https://doi.org/10.1029/2021WR031575

Hampton, T. B., & Basu, N. B. (2022). A novel Budyko-based approach to quantify post-forest-fire streamflow response and recovery timescales. *Journal of Hydrology*, *608*, 127685. https://doi.org/10.1016/j.jhydrol.2022.127685

Harpold, A. A., Biederman, J. A., Condon, K., Merino, M., Korgaonkar, Y., Nan, T., et al. (2014). Changes in snow accumulation and ablation following the Las Conchas Forest Fire, New Mexico, USA: Changes in snow following fire. *Ecohydrology*, *7*(2), 440–452. https://doi.org/10.1002/eco.1363

He, C., Chen, F., Abolafia-Rosenzweig, R., Ikeda, K., Liu, C., & Rasmussen, R. (2021). What causes the unobserved early-spring snowpack ablation in convection-permitting WRF modeling over Utah Mountains? *Journal of Geophysical Research: Atmospheres*, *126*(22), e2021JD035284. https://doi.org/10.1029/2021JD035284

He, C., Chen, F., Barlage, M., Liu, C., Newman, A., Tang, W., et al. (2019). Can convection-permitting modeling provide decent precipitation for offline high-resolution snowpack simulations over mountains? *Journal of Geophysical Research: Atmospheres*, *124*(23), 12631–12654. https://doi.org/10.1029/2019JD030823

He, C., Valayamkunnath, P., Barlage, M., Chen, F., Gochis, D., Cabell, R., et al. (2023). The community Noah-MP land surface modeling system technical description version 5.0 (No. NCAR/TN-575+STR). https://doi.org/10.5065/ew8g-yr95

Kampf, S. K., McGrath, D., Sears, M. G., Fassnacht, S. R., Kiewiet, L., & Hammond, J. C. (2022). Increasing wildfire impacts on snowpack in the western U.S. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(39), e2200333119. https://doi.org/10.1073/pnas.2200333119

Kapnick, S. B., Yang, X., Vecchi, G. A., Delworth, T. L., Gudgel, R., Malyshev, S., et al. (2018). Potential for western US seasonal snowpack prediction. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(6), 1180–1185. https://doi.org/10.1073/pnas.1716760115

Koshkin, A. L., Hatchett, B. J., & Nolin, A. W. (2022). Wildfire impacts on western United States snowpacks. *Front. Water*, *4*, 971271. https://doi.org/10.3389/frwa.2022.971271

Kumar, S. V., Peterslidard, C., Tian, Y., Houser, P., Geiger, J., Olden, S., et al. (2006). Land information system—An interoperable framework for high resolution land surface modeling. *Environmental Modelling & Software*, *21*(10), 1402–1415. https://doi.org/10.1016/j.envsoft.2005.07.004

Li, D., Wrzesien, M. L., Durand, M., Adam, J., & Lettenmaier, D. P. (2017). How much runoff originates as snow in the western United States, and how will that change in the future? *Geophysical Research Letters*, *44*(12), 6163–6172. https://doi.org/10.1002/2017GL073551

Li, J., Zhang, G., Chen, F., Peng, X., & Gan, Y. (2019). Evaluation of land surface subprocesses and their impacts on model performance with global flux data. *Journal of Advances in Modeling Earth Systems*, *11*(5), 1329–1348. https://doi.org/10.1029/2018MS001606

Liang, S., Cheng, J., Jia, K., Jiang, B., Liu, Q., Xiao, Z., et al. (2021). The global land surface satellite (GLASS) product suite. *Bulletin America Meteorology Social*, *102*(2), E323–E337. https://doi.org/10.1175/BAMS-D-18-0341.1

Liu, C., Ikeda, K., Rasmussen, R., Barlage, M., Newman, A. J., Prein, A. F., et al. (2017). Continental-scale convection-permitting modeling of the current and future climate of North America. *Climate Dynamics*, *49*(1–2), 71–95. https://doi.org/10.1007/s00382-016-3327-9

Lu, B., Zhong, J., Wang, W., Tang, S., & Zheng, Z. (2021). Influence of near real-time green vegetation fraction data on numerical weather prediction by WRF over North China. *Journal of Meteorological Research*, *35*(3), 505–520. https://doi.org/10.1007/s13351-021-0163-6

Ma, Q., Bales, R. C., Rungee, J., Conklin, M. H., Collins, B. M., & Goulden, M. L. (2020). Wildfire controls on evapotranspiration in California's Sierra Nevada. *Journal of Hydrology*, *590*, 125364. https://doi.org/10.1016/j.jhydrol.2020.125364

Maina, F. Z., & Siirila-Woodburn, E. R. (2020). Watersheds dynamics following wildfires: Nonlinear feedbacks and implications on hydrologic responses. *Hydrological Processes*, *34*(1), 33–50. https://doi.org/10.1002/hyp.13568

Martin, D. A., & Moody, J. A. (2001). Comparison of soil infiltration rates in burned and unburned mountainous watersheds. *Hydrological Processes*, *15*, 2893–2903. https://doi.org/10.1002/hyp.380

Maxwell, J., & St Clair, S. B. (2019). Snowpack properties vary in response to burn severity gradients in montane forests. *Environmental Research Letters*, *14*(12), 124094. https://doi.org/10.1088/1748-9326/ab5de8

Moody, J. A., Martin, D. A., Haire, S. L., & Kinner, D. A. (2008). Linking runoff response to burn severity after a wildfire. *Hydrological Processes*, *22*(13), 2063–2074. https://doi.org/10.1002/hyp.6806

Myneni, R., Knyazikhin, Y., & Park, T. (2015). *MOD15A2H MODIS leaf area index/FPAR 8-day L4 global 500m SIN grid V006*. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MOD15A2H.006

Myneni, R. B., & Williams, D. L. (1994). On the relationship between FAPAR and NDVI. *Remote Sensing of Environment*, *49*(3), 200–211. https://doi.org/10.1016/0034-4257(94)90016-7

Niemeyer, R. J., Bladon, K. D., & Woodsmith, R. D. (2020). Long-term hydrologic recovery after wildfire and post-fire forest management in the interior Pacific Northwest. *Hydrological Processes*, *34*(5), 1182–1197. https://doi.org/10.1002/hyp.13665

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., et al. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research*, *116*(D12), D12109. https://doi.org/10.1029/2010JD015139

O'Leary, D., Hall, D., Medler, M., & Flower, A. (2018). Quantifying the early snowmelt event of 2015 in the Cascade Mountains, USA by developing and validating MODIS-based snowmelt timing maps. *Frontiers of Earth Science*, *12*(4), 693–710. https://doi.org/10.1007/s11707-018-0719-7

O'Leary, III, D., Hall, D. K., Medler, M., Matthews, R., & Flower, A. (2019). *Snowmelt timing maps derived from MODIS for North America, version 2, 2001–2018*. ORNL DAAC. https://doi.org/10.3334/ORNLDAAC/1712

Partington, D., Thyer, M., Shanafield, M., McInerney, D., Westra, S., Maier, H., et al. (2022). Predicting wildfire induced changes to runoff: A review and synthesis of modeling approaches. *WIREs Water*, *9*(5),e1599. https://doi.org/10.1002/wat2.1599

Peters-Lidard, C. D., Houser, P. R., Tian, Y., Kumar, S. V., Geiger, J., Olden, S., et al. (2007). High-performance Earth system modeling with NASA/GSFC's land information system. *Innovations in Systems and Software Engineering*, *3*(3), 157–165. https://doi.org/10.1007/s11334-007-0028-x

Powers, J. G., Klemp, J. B., Skamarock, W. C., Davis, C. A., Dudhia, J., Gill, D. O., et al. (2017). The weather research and forecasting model: Overview, system efforts, and future directions. *Bulletin of the American Meteorological Society*, *98*(8), 1717–1737. https://doi.org/10.1175/BAMS-D-15-00308.1

Pugh, E., & Gordon, E. (2013). A conceptual model of water yield effects from beetle-induced tree death in snow-dominated lodgepole pine forests: A conceptual model of the water yield effects of tree death. *Hydrological Processes*, *27*(14), 2048–2060. https://doi.org/10.1002/hyp.9312

Pugh, E., & Small, E. (2012). The impact of pine beetle infestation on snow accumulation and melt in the headwaters of the Colorado River. *Ecohydrology*, *5*(4), 467–477. https://doi.org/10.1002/eco.239

Rasmussen, R. M., Chen, F., Liu, C., Ikeda, K., Prein, A., Kim, J., et al. (2023). The NCAR/USGS 4-km long-term regional hydroclimate reanalysis over the CONUS. *Bulletin of the American Meteorological Society*, *104*(8), E1382–E1408. https://doi.org/10.1175/BAMS-D-21-0326.1

Seibert, J., McDonnell, J. J., & Woodsmith, R. D. (2010). Effects of wildfire on catchment runoff response: A modelling approach to detect changes in snow-dominated forested catchments. *Hydrology Research*, *41*(5), 378–390. https://doi.org/10.2166/nh.2010.036

Shakesby, R., & Doerr, S. (2006). Wildfire as a hydrological and geomorphological agent. *Earth-Science Reviews*, *74*(3–4), 269–307. https://doi.org/10.1016/j.earscirev.2005.10.006

Smoot, E. E., & Gleason, K. E. (2021). Forest fires reduce snow-water storage and advance the timing of snowmelt across the western U.S. *Water*, *13*(24), 3533. https://doi.org/10.3390/w13243533

Spence, C., Hedstrom, N., Tank, S. E., Quinton, W. L., Olefeldt, D., Goodman, S., & Dion, N. (2020). Hydrological resilience to forest fire in the subarctic Canadian shield. *Hydrological Processes*, *34*(25), 4940–4958. https://doi.org/10.1002/hyp.13915

Stoof, C. R., Vervoort, R. W., Iwema, J., van den Elsen, E., Ferreira, A. J. D., & Ritsema, C. J. (2012). Hydrological response of a small catchment burned by experimental fire. *Hydrology and Earth System Sciences*, *16*(2), 267–285. https://doi.org/10.5194/hess-16-267-2012

Sun, N., Yan, H., Wigmosta, M., Skaggs, R., Leung, R., & Hou, Z. (2019). Regional snow parameters estimation for large-domain hydrological applications in the western United States. *Journal of Geophysical Research: Atmospheres*, *124*(10), 5296–5313. https://doi.org/10.1029/2018JD0301

Viviroli, D., Dürr, H. H., Messerli, B., Meybeck, M., & Weingartner, R. (2007). Mountains of the world, water towers for humanity: Typology, mapping, and global significance. *Water Resources Research*, *43*(7). https://doi.org/10.1029/2006WR005653

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

Williams, A. P., Livneh, B., McKinnon, K. A., Hansen, W. D., Mankin, J. S., Cook, B. I., et al. (2022). Growing impact of wildfire on western US water supply. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(10), e2114069119. https://doi.org/10.1073/pnas.2114069119

Xiao, M., Mahanama, S. P., Xue, Y., Chen, F., & Lettenmaier, D. P. (2021). Modeling snow ablation over the mountains of the western United States: Patterns and controlling factors. *Journal of Hydrometeorology*, *22*(2), 297–311. https://doi.org/10.1175/JHM-D-19-0198.1

Yan, H., Sun, N., Wigmosta, M., Skaggs, R., Hou, Z., & Leung, R. (2018). Next-generation intensity-duration-frequency curves for hydrologic design in snow-dominated environments. *Water Resources Research*, *54*(2), 1093–1108. https://doi.org/10.1002/2017wr021290

Zhang, G., Chen, F., Chen, Y., Li, J., & Peng, X. (2020). Evaluation of Noah-MP land-model uncertainties over sparsely vegetated sites on the Tibet Plateau. *Atmosphere*, *11*(5), 458. https://doi.org/10.3390/atmos11050458

Zhang, G., Chen, F., & Gan, Y. (2016). Assessing uncertainties in the Noah-MP ensemble simulations of a cropland site during the Tibet Joint International Cooperation program field campaign. *Journal of Geophysical Research: Atmospheres*, *121*(16), 9576–9596. https://doi.org/10.1002/2016JD024928