https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

Improving Model Calibrations in a Changing World: Controlling for Nonstationarity After Mega Disturbance Reduces Hydrological Uncertainty

Elijah N. Boardman[1,2], Gabrielle F. S. Boisramé[3], Mark S. Wigmosta[4], Robert K. Shriver[5], Adrian A. Harpold[1,5]

[1]Graduate Program of Hydrologic Sciences, University of Nevada, Reno, Reno, Nevada, 89557, USA
[2]Mountain Hydrology LLC, Reno, Nevada, 89503, USA
[3]Hydrologic Sciences Division, Desert Research Institute, Reno, Nevada, 89512, USA
[4]Energy and Environment Directorate, Pacific Northwest National Laboratory, Richland, Washington, 99354, USA
[5]Department of Natural Resources and Environmental Science, University of Nevada, Reno, Reno, Nevada, 89557, USA

Correspondence to: Elijah N. Boardman (eli.boardman@mountainhydrology.com)

Abstract. Model simulations are widely used to understand, predict, and respond to environmental changes, but uncertainty in these models can hinder decision-making. The simulation of hydrological changes after a forest fire is a typical example where process-based models with uncertain parameters may inform consequential predictions of water availability. Different parameter sets can yield similarly realistic simulations during model calibration but generate divergent predictions of change, a problem known as "equifinality." Despite longstanding recognition of the problems posed by equifinality, the implications for environmental disturbance simulations remain largely unconstrained. Here, we demonstrate how equifinality in water balance partitioning causes compounding uncertainty in hydrological changes attributable to a recent 1,540 km^2 megafire in the Sierra Nevada mountains (California, USA). Different sets of calibrated parameters generate uncertain predictions of the four-year post-fire streamflow change that vary up to six-fold. However, controlling for nonstationary model error (e.g., a shift in the model bias after disturbance) can significantly ($p < 0.01$) reduce both equifinality and predictive uncertainty. Using a statistical metamodel to correct for bias shift after disturbance, we estimate a streamflow increase of 11% ±1% in the first four years after the fire, with an 18% ±4% increase during drought. Our metamodel framework for correcting nonstationarity reduces uncertainty in the post-fire streamflow change by 80% or 82% compared to the uncertainty of pure statistical or pure process-based model ensembles, respectively. As environmental disturbances continue to transform global landscapes, controlling for nonstationary biases can improve process-based models that are used to predict and respond to unprecedented hydrological changes.

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

1 Introduction

Calibration – systematic adjustment of model parameters to improve simulation accuracy.

Disturbance – an event that changes an environmental system from one state to another.

Equifinality – the production of similar results for different reasons.

Stationarity – the invariance of a statistical property across different time periods.

Environmental disturbances (e.g., forest fires, other vegetation mortality events, floods, anthropogenic land cover conversion, etc.) can alter the structure and function of ecohydrological systems (Zehe and Sivapalan 2009, Ebel and Mirus 2014, Buma 2015, Johnstone et al. 2016). Climate change and environmental disturbances introduce nonstationarity into the hydrological cycle, which is disrupting longstanding statistical approaches to water resource and risk management (Milly et al. 2008, 2015, Hirsch 2011, Salas et al. 2012, Yang et al. 2021).

Pure statistical methods (e.g., regression models lacking an explicit physical foundation) can sometimes detect streamflow changes attributable to environmental disturbance by comparing measurements to a stationary model, which represents a no-disturbance counterfactual. Statistical change attribution is generally applied across many years and numerous sites (e.g., Goeking and Tarboton 2022a, Hampton and Basu 2022, Williams et al. 2022) or in careful paired watershed studies to overcome climate/weather variability (e.g., Bart 2016, Manning et al. 2022, Johnson and Alila 2023, Kang and Sharma 2024). However, in a single watershed with a short post-disturbance record, pure data-driven statistical approaches are inherently limited. Crucially, many water management decisions (e.g., reservoir release schedules) are made on a per-watershed and per-year basis, so large-scale retrospective statistical assessments of disturbance effects may not provide actionable insights in any particular watershed.

Spatially distributed process-based hydrological models, and related land surface or Earth system models, are a widely accepted tool that can overcome some limitations of statistical disturbance attribution (Fatichi et al. 2016, Pongratz et al. 2018, Fisher and Koven 2020). Since interannual climate variability often obscures hydrological changes caused by disturbance, counterfactual model experiments using an undisturbed control are a cornerstone of ecohydrological disturbance attribution studies (e.g., Moreno et al. 2016, Saksa et al. 2017, Boisramé et al. 2019, Meili et al. 2024). Moreover, key process representations (e.g., flow routing and the snowpack energy balance) are expected to generalize beyond observed conditions, providing a basis for the prediction of hydrological responses to out-of-sample events including extreme storms (e.g., Huang and Swain 2022), decadal-scale climate change (e.g., Tague et al. 2009), and unprecedented "megafires" (e.g., Abolafia-Rosenzweig et al. 2024).

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

Since we lack landscape-scale observations of many important environmental properties, model parameters are often estimated through calibration. Equifinality arises during calibration when different parameter sets yield similar realizations of observable phenomena (Beven 1993, 2006, Ebel and Loague 2006). Recognizing that equifinality may preclude the possibility of picking a single "best" parameter set, some modelers advocate for using a "behavioral" ensemble based on subjective goodness-of-fit criteria in a generalized likelihood uncertainty estimation (GLUE) framework (Spear and Hornberger 1980, Beven and Binley 1992, Her and Chaubey 2015, Vrugt and Beven 2018).

Equifinality implies process uncertainty (Grayson et al. 1992, Khatami et al. 2019). For example, total evapotranspiration (ET) is the sum of overstory and understory transpiration, interception loss, soil evaporation, snow sublimation, and other vapor fluxes; equifinal parameter sets may produce the same total ET with different partitioning between constituent fluxes (Franks et al. 1997, Birkel et al. 2024). Since each vapor flux component can respond differently to disturbance (Goeking and Tarboton 2020), we hypothesize that equifinal parameter sets may produce divergent predictions when the model is perturbed beyond the calibration space.

We illustrate the hypothesized interaction of equifinality, disturbance, and bias (non)stationarity using a conceptual water balance model (Fig. 1). Example models of the pre-disturbance water balance each achieve the same mean pre-disturbance streamflow (Q), which is forced to approximately match Q observations through model calibration. Due to equifinality, there is residual uncertainty in the bias-corrected total precipitation (P) and the partitioning of ET between transpiration and interception from tree canopies (ET_{Tree}) and other vapor fluxes (ET_{Other}, e.g., understory ET and soil evaporation). When a disturbance such as a fire reduces ET_{Tree}, the streamflow response is sensitive to the initial ET_{Tree} magnitude (and hence the potential ET reduction) as well as the degree to which ET_{Other} responds to increased soil water availability. Over- or under-estimation of the resultant streamflow change (ΔQ) manifests as a positive or negative "bias shift" after disturbance. The bias shift metric, as defined here, is a special discrete case of the more general concept of nonstationarity. In a system with changes that occur over longer time periods (in contrast to the discrete disturbance shown in Fig. 1), a different stationarity metric would be necessary to account for incremental changes. In the present study, zero bias shift after disturbance implies stationary error overall.

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

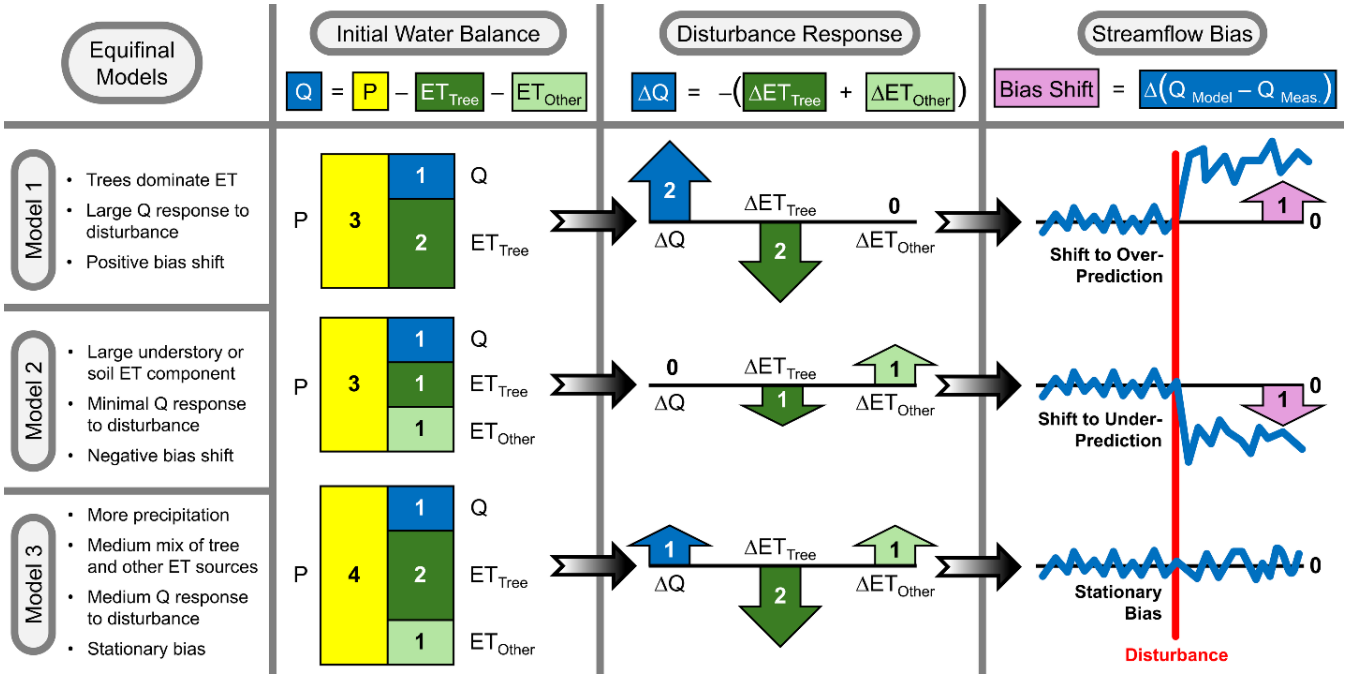

Figure 1: Conceptual model illustrating how equifinality in the modeled water balance may lead to uncertainty in the streamflow response to disturbance, and how we expect this to manifest in a measurable "bias shift" after a discrete disturbance. Numbers are indicative and not intended to represent actual disturbance magnitudes.

We build on this conceptual example of the interaction between equifinality, disturbance, and nonstationarity (Fig. 1) to consider how the bias shift metric can help select parameter sets with enhanced physical fidelity and greater predictive confidence. The initial water balance of Model 1 is dominated by ET_{Tree}, leading to a large streamflow gain and a positive bias shift (tendency toward over-prediction of post-disturbance streamflow). Conversely, Model 2 has a large ET_{Other} component, which compensates for the comparatively small reduction in ET_{Tree}, leading to a negligible streamflow gain and a negative bias shift (tendency toward under-prediction of post-disturbance streamflow). Finally, Model 3 has more precipitation than the other models and a more balanced combination of ET_{Tree} and ET_{Other}, leading to a medium streamflow gain and stationary bias. In this case, Model 3 should be preferred due to its negligible bias shift, which would help achieve a better prediction of ΔET and ΔQ and also help constrain uncertainty in the underlying parameterization.

The interaction of equifinality and disturbance is rarely addressed in process-based simulations. In contemporary studies, single parameter sets are sometimes used with or without calibration (e.g., Furniss et al. 2023, Abolafia-Rosenzweig et al. 2024). When calibrated ensembles are used, uncertainty propagation is commonly limited to subsurface parameters and meteorological biases (e.g., Shields and Tague 2012, Saksa et al. 2017, Boisramé et al. 2019). We expect that latent uncertainty in vegetation parameters may contribute an unconstrained source of uncertainty in studies of ecohydrological

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

disturbance that do not account for vegetation parameter equifinality. Conversely, model equifinality can be reduced by leveraging additional types of information beyond traditional streamflow calibration metrics (Kelleher et al. 2017). One unexplored approach to equifinality reduction is evaluating the stationarity of model biases after environmental disturbance, which we consider here.

In this study, we leverage a large wildfire as a "natural experiment" to test the hypothesis that quantifying stationarity across pre- and post-disturbance periods can reduce equifinality and improve the predictive confidence of a process-based hydrological model. Specifically, we apply the Distributed Hydrology Soil Vegetation Model (DHSVM, Wigmosta et al. 1994) to simulate streamflow changes attributable to the Creek Fire in the Sierra Nevada mountains (California, USA), which burned 56% of the forested area in our 4,244 km^2 study watershed (Stephens et al. 2022, Ayars et al. 2023). We expect that this drastic landscape-scale environmental disturbance should have a clear impact on regional-scale water fluxes, providing an opportunity to test whether model process representations are robust to disturbance. We leverage a multi-objective calibration of vegetation, snow, subsurface, and meteorological bias-correction parameters to address two research questions:

(1) How does calibration equifinality impact process-based simulations of the hydrological response to a megafire?

(2) Can we reduce equifinality and uncertainty by testing the model's representation of hydrological change?

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

2 Methods

2.1 Study Area and Data

Our study watershed encompasses the Upper San Joaquin River Basin above Millerton Lake, a total of 4,244 km^2 with an elevation range of 100 to 4,200 m (Fig. 2A). The 2020 Creek Fire burnt 1,540 km^2 of mixed conifer and scrub forest, including 1,481 km^2 within the study watershed (56% of the forested watershed area). Landsat-based data from Monitoring Trends in Burn Severity (MTBS, Fig. 2B) indicate that 16% of the Creek Fire exhibited high burn severity and 30% exhibited moderate severity (Eidenshink et al. 2007, MTBS Project 2022). However, using a longer time period for pre- and post-fire imagery, Stephens et al. (2022) estimate 41% high severity and 35% moderate severity, illustrating the proliferation of uncertainty in disturbance assessments.

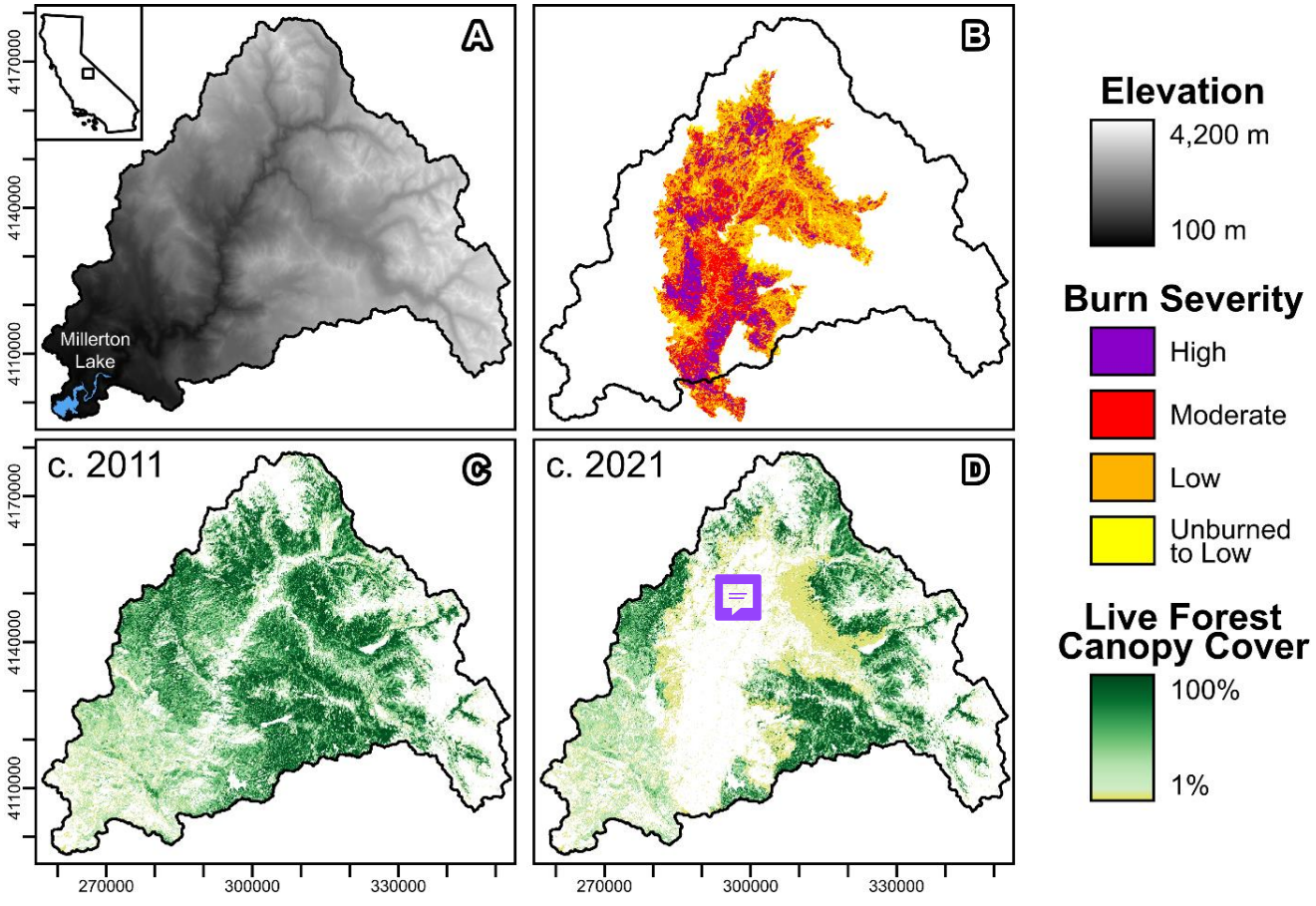

Figure 2: Maps of the study watershed: (A) elevation and watershed location in the U.S. State of California, (B) 2020 Creek Fire burn severity from MTBS, (C) pre-fire and (D) post-fire forest canopy cover from RCMAP. Tick marks indicate 10 km intervals in UTM zone 11N.

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

We represent fire disturbance in DHSVM by adjusting maps of vegetation properties. All maps are projected to our selected DHSVM resolution of 90 m. The Landsat-based RCMAP data provide yearly fractional cover estimates for trees and shrub/herbaceous vegetation at 30 m resolution (Rigge et al. 2021a,b). We use the 2011-era RCMAP data as a pre-fire baseline and the 2021-era RCMAP data to capture the effects of the 2020 Creek Fire (Fig. 2C-D). We also update the vegetation maps in 2013, 2014, and 2018 to reflect smaller fires in those years. The DHSVM vegetation maps are updated on October 1st in the year of a fire, i.e., about one month after the September 2020 Creek Fire ignition. Vegetation is classified based on the species (when available) or functional type (e.g., mixed conifer forest) using Landfire data (2022), and abiotic land surface classes are derived from NLCD (Dewitz and U.S. Geological Survey 2019). Landfire and RCMAP provide tree and shrub height data, respectively. Tree leaf area index (LAI) is estimated empirically from fractional cover following Pomeroy et al. (2002), which is reproduced as Eq. (1) of Goeking and Tarboton (2022b). Vegetation transpiration is calculated by DHSVM based on the vegetation type and local weather, soil moisture, and light in each grid cell (Wigmosta et al. 1994). Baseline values of minimum stomatal resistance are estimated from species-level field studies as detailed in the Supporting Information of Boardman et al. (2025). Fractional cover, LAI, and stomatal resistance are refined by calibration relative to baseline (Sect. 2.2).

Spatial maps and parameter values for DHSVM are collated from a wide range of literature and field studies, as detailed in Boardman et al. (2025) and Boardman (2023). We briefly summarize key setup procedures here. Subsurface properties are estimated by disaggregating regional soil survey databases (Gupta et al. 2022, Soil Survey Staff 2022) using Random Forest models trained on topographic metrics (Breiman et al. 2002). In the updated version of DHSVM used here, streamflow in channels is bidirectionally coupled to the groundwater level in each grid cell, and the maximum network extent is derived from the National Hydrography Dataset (U.S. Geological Survey 2019) with channel geometry from regional regressions (Bieger et al. 2015). Meteorological data from gridMET (Abatzoglou 2013) are disaggregated to a 3-hour timestep using MetSim (Bennett et al. 2020). Modeled snowfall is distributed in proportion to the pixel-wise maximum observed snow water equivalent (SWE) pattern derived from Airborne Snow Observatory (ASO) data in the study watershed (Painter et al. 2016), which implicitly accounts for snow transport (Vögeli et al. 2016). Regional snow/rain partitioning parameters are adopted from Sun et al. (2019).

2.2 Model Calibration

We calibrate 14 sensitive and uncertain parameters in DHSVM that control aspects of the meteorology, vegetation, subsurface, and snowpack dynamics (Table 1). While most of these parameters are widely recognized as suitable for calibration (Cuo et al. 2011, Du et al. 2014), precipitation and temperature biases are less frequently included in the calibration of distributed process-based models despite considerable uncertainty in gridded meteorological data (Henn et al. 2018). We expect that interactions between meteorological uncertainty and parameter equifinality may contribute to the

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

overall uncertainty of disturbance simulations (Fig. 1), but this uncertainty would remain hidden if meteorological biases were assumed zero.

Category	Parameter	Range	Primary Process Controls
Meteorology	Precipitation Bias	±25%	Net water balance input, interannual variability in water yield
	Temperature Bias	±4 °C	Snow/rain partitioning, potential evapotranspiration (PET)
	Temperature Lapse Rate	-8 to -2 °C/km	Spatial distribution of snow and PET
Vegetation	Tree Fractional Cover	50% to 200% of baseline, each cell ≤ 100% cover	Canopy interception and transpiration, understory and snowpack shading
	Tree Leaf Area Index (LAI)	50% to 200% of baseline	Canopy interception capacity, overstory transpiration
	Stomatal Resistance	50% to 200% of baseline	Overstory and understory transpiration
Subsurface	Soil Depth	1 to 10 m	Lateral transmissivity, root zone groundwater access, storage capacity
	Hydraulic Conductivity	10^{-5} to 10^{-2} m/s	Lateral transmissivity, vertical recharge rate (by anisotropy ratio), surface/subsurface flow partitioning
	Exponential Decrease in Conductivity	10^{-3} to 10^{0} [unitless]	Lateral transmissivity, vertical conductivity profile, baseflow recession, water table depth
	Porosity	0.3 to 0.6 [fractional]	Dynamic storage range, water table response to infiltration
	Field Capacity	0.1 to 0.4 [fractional]	Dynamic storage range, soil water retention, plant available water
Snow	Albedo Decay (Accumulation)	0.7 to 0.99	Snowpack energy balance, maximum accumulation
	Albedo Decay (Melt Season)	-0.3 to 0.0 relative to accumulation	Snowpack energy balance, snowmelt rate, snow cover duration
	Albedo Reset Snowfall Scale	10^{-4} to 10^{0} m SWE	Albedo increase associated with new snowfall of a given depth

Table 1: Prior ranges and process controls of DHSVM parameters calibrated in this study. All vegetation and subsurface parameters listed here are defined by spatially variable maps, and calibration ranges determine the area-average value around which the pattern is rescaled.

Multiple parameters combine to control simulated processes. For example, area-average LAI (related to total interception loss) is the product of tree-scale LAI with grid-scale fractional cover. Tree transpiration is determined by fractional cover, LAI, stomatal resistance, available soil water (related to subsurface parameters), and other factors. Lateral transmissivity in the saturated subsurface is controlled by three parameters: soil depth, surface hydraulic conductivity, and the exponential decrease in conductivity with depth. Cross-compensation among interrelated parameters thus contributes to equifinality. Within our 14-dimensional calibration space, 23 parameter pairs have correlations that are significant at $p < 0.05$

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

(Supplemental Fig. S1). Furthermore, perturbing one aspect of the model can lead to cascading effects due to the coupling of ecohydrological processes and spatial water connectivity in the model. For example, lateral hydraulic conductivity is coupled to vertical conductivity by anisotropy ratios dependent on the soil textural classification (Fan and Miguez-Macho 2011), so calibrating lateral conductivity also influences groundwater recharge rates from losing stream reaches, which in turn can affect soil evaporation and transpiration from riparian trees. Spatial heterogeneity in modeled soil and vegetation properties (Sect. 2.1) further complicates all of these interactions, e.g., different parts of the landscape are relatively more sensitive to calibration of different parameters depending on the baseline map patterns.

Given the complexity of expected interactions, we define seven objective functions to constrain parameters based on different hydrological signatures (Table 2). Three objectives are based on daily streamflow, which is reconstructed at Millerton Lake (Fig. 2) to remove the effects of upstream reservoirs and diversions (California Department of Water Resources 2024). Two objective functions similarly target annual percent error in the annual water yield and the April-July water yield, which is a well-established benchmark for snowmelt runoff modeling in the Sierra Nevada (Pagano et al. 2004). Two objective functions are based on the eight-year (2017-2024), 30-survey database of ASO SWE maps in the study area, targeting both the spatial distribution at the 90 m grid scale and the percent error in total volume across surveys. Hydrograph and water yield objectives are calculated for water years 2015-2024, which includes six years before and four years after the Creek Fire. By calibrating across this disturbance (vegetation maps updated during calibration), we automatically reject parameter sets that fail to provide reasonably accurate estimates of both pre- and post-fire streamflow.

Category	Objective Function	Best Value	Worst Value	Target Hydrological Signatures
Daily Streamflow (2015-2024)	NSE	0.89	0.80	Hydrograph shape (high flows), rainfall-runoff response, snow/rain partitioning, peak flow and recession timing
	Log-Scaled NSE	0.85	0.80	Hydrograph shape (low flows), baseflow recession characteristics, multi-year storage/deficit effects
	>95th-Percentile RMSE	26 m^3/s	74 m^3/s	High flow magnitude, shape of flow duration curve independent of timing
Water Yield (2015-2024)	Yearly MAPE	4%	9%	Bulk water balance, interannual variability across wet and dry years
	April-July MAPE	7%	10%	Interannual variability in snowmelt runoff efficiency and timing
Snow Maps (30 ASO Surveys, 2017-2024)	Pixel-Wise SWE RMSE	0.23 m	0.25 m	Spatial distribution of snow accumulation and ablation, absolute magnitude of SWE in different years
	Total SWE Volume MAPE	18%	32%	Evolution of snowpack volume between surveys, interannual variability

Table 2: Calibration objective functions used in this study with descriptions of the primary hydrological signatures constrained by each objective. The best (worst) value given here is the lowest (highest) error achieved by any of the Pareto-efficient parameter sets in our calibrated 30-member behavioral ensemble. NSE = Nash Sutcliffe Efficiency (identical to R^2 for statistical models), RMSE = root mean square error, MAPE = mean absolute percent error.

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

To efficiently sample behavioral parameter sets from the 14-dimensonal space of potential interactions, we apply a multi-objective Bayesian optimization scheme (Jones et al. 1998). After an initial Latin hypercube sample of 320 parameter sets (Dupuy et al. 2015), we perform parallel particle swarm optimization (Kennedy and Eberhart 1995, Zambrano-Bigiarini et al. 2013) using the expected hypervolume indicator (Emmerich et al. 2011, Binois and Picheny 2019) to sample promising parameter sets based on Gaussian Process surrogate models of the objective function response surfaces (Roustant et al. 2012). After six optimization generations, we have tested a total of 600 parameter sets (n.b. this requires ~950 days of CPU time on 2.5 GHz servers, and the elapsed wall-clock time is several weeks since multiple parameter sets are tested in parallel). Of all the tested parameter sets, 48 qualify as "behavioral" by satisfying the following subjective criteria: daily NSE > 0.8, daily log NSE > 0.8, yearly MAPE < 10%, April-July MAPE < 10%, and Pareto-efficient across all objectives. Some parameter samples are similar, so for efficiency we further select 30 diverse samples by iteratively choosing the behavioral parameter set with the maximum mean parameter separation from previously selected samples. These 30 parameter sets define the behavioral DHSVM ensemble referenced hereafter. We note that our conclusions are robust to random sub-selection of fewer models, as long as at least ~10 parameter sets are used (Supplemental Fig. S2).

2.3 Disturbance Simulations

We investigate the ecohydrological effects of the Creek Fire by comparing model simulations using dynamic and static vegetation maps to quantify the fire effect relative to a no-fire control scenario. For each of the 30 DHSVM parameter sets, we simulate streamflow for the past 20 years (water years 2005-2024) with either static 2011-era vegetation maps or dynamic vegetation maps updated in 2013, 2014, 2018, and 2020. The 2020 Creek Fire accounts for most of the vegetation disturbance in the study area, with a 42% reduction in watershed-average RCMAP tree fractional cover compared to 2-3% reductions associated with the 2013, 2014, and 2018 fires. Differences between fire-aware (dynamic vegetation) and no-fire control (static vegetation) simulations define the modeled disturbance effect. In addition to comparing daily streamflow, we also compare annual water yield and ET fluxes between fire-aware and no-fire control scenarios.

2.4 Detecting and Correcting Nonstationarity

We calculate a "bias shift" metric by comparing observed streamflow with modeled streamflow from the fire-aware (dynamic vegetation) simulations. The 30-member behavioral DHSVM ensemble has a reasonably small mean streamflow bias for the overall 2005-2024 evaluation period (interquartile range among parameter sets of ±2%). However, some parameter sets have different mean streamflow biases on pre- and post-fire periods, congruent with our conceptual model in Fig. 1. We theorize that over- or under-estimation of the disturbance effect on streamflow may result in a matching positive or negative bias shift after disturbance, defined as the difference in mean streamflow bias between post-fire and pre-fire periods:

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

$$Bias\ Shift = (\overline{Q_{Model}} - \overline{Q_{Meas.}})_{Post-Fire} - (\overline{Q_{Model}} - \overline{Q_{Meas.}})_{Pre-Fire} \quad (1)$$

We correct for the bias shift of different parameter sets by developing a "metamodel," i.e., a statistical model trained on DHSVM outputs. The bias shift metric, Eq. (1), is averaged across multiple years, whereas we expect that each individual year may have a larger or smaller streamflow response due to variable interactions between climate and vegetation. In the case that the streamflow response is purely energy-limited (P >> ET), we would expect the same post-fire streamflow gain in all years; conversely, in a water-limited case (P closer to ET magnitude) we would expect a 1:1 scaling between annual precipitation and the post-fire streamflow gain. Between these two endmember scenarios, we expect that the magnitude of the simulated streamflow change in any particular year may be offset and/or fractionally re-scaled relative to the mean multi-year streamflow change. Thus, we posit a linear relationship between the multi-year bias shift and the simulated streamflow response to fire in any particular year, ΔQ_{Fire}.

In a Bayesian statistical framework, we treat each DHSVM parameter set as an independent realization of the possible post-fire response, with a stochastic error term describing scatter in the hypothesized linear relationship between bias shift and ΔQ_{Fire}. We define the metamodel using a normal distribution with mean determined by the linear bias shift vs. ΔQ_{Fire} relationship and uncertainty defined by the sample standard deviation σ, which can be expressed in Bayesian sampling notation as:

$$\Delta Q_{Fire} \sim normal(c_0 + c_1 * Bias\ Shift, \sigma) \quad (2)$$

To estimate the values of c_0, c_1, and σ (with quantified uncertainty in all three parameters), we generate 1,000 Bayesian samples using the Hamiltonian Monte Carlo algorithm with two chains (500 samples per chain) after 10,000 warmup iterations (Stan Development Team 2023). The metamodel is fit using all 30 pairs of bias shift and ΔQ_{Fire} values calculated for each parameter set in the behavioral DHSVM ensemble, with c_0, c_1, and σ re-fit for each of the four post-fire years. We subsequently generate a conditional prediction of ΔQ_{Fire} in each year by setting the bias shift equal to zero in Eq. (2), which yields a normal distribution with mean c_0 and standard deviation σ. Unlike simple least-squares linear regression, uncertainty in the metamodel parameters (c_0, c_1, and σ) is propagated into our conditional predictions through the Bayesian sampling routine, which considers 1,000 different combinations of plausible c_0, c_1, and σ values. Sampling the posterior distribution of Eq. (2) with bias shift set to zero yields a conditional distribution describing the expected post-fire streamflow change and uncertainty of a hypothetical DHSVM simulation with zero bias shift.

2.5 Empirical Regression Model

To compare statistical and process-based approaches to ecohydrological disturbance attribution, we also apply an empirical annual water balance model using Bayesian multiple linear regression. We posit a simple four-parameter lumped empirical model that estimates the annual runoff efficiency (Q / P) as a linear function of annual precipitation (P), the prior year's streamflow ($Q_{LastYear}$) to account for multi-year storage or deficit effects, and the aridity index calculated from annual

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

potential evapotranspiration (PET / P). The model structure is adapted from a similar regression approach applied to analyze seasonal water supply in adjacent watersheds (Boardman et al. 2024). We assume that each year's actual runoff efficiency is randomly sampled from a normal distribution with standard deviation σ and mean defined by the linear model, expressed analogously to Eq. (2) in Bayesian sampling notation:

$$\frac{Q}{P} \sim normal\left(c_0 + c_1 * P + c_2 * Q_{LastYear} + c_3 * \frac{PET}{P}, \sigma\right) \quad (3)$$

We constrain the empirical model using pre-fire data and compare its post-fire predictions with measured post-fire streamflow. Meteorological data required for Eq. (3) are aggregated from the same gridMET data used for DHSVM (Abatzoglou 2013) over water years 1980-2020. As for Eq. (2), we generate 1,000 Bayesian samples of the empirical model parameters (c_0, c_1, c_2, and σ) using Hamiltonian Monte Carlo (Stan Development Team 2023). The empirical model achieves R^2 = 0.91 for annual variations in runoff efficiency across the 41-year fitting period. By sampling the model's posterior predictive distribution using meteorological data from 2021-2024, we generate 1,000 counterfactual estimates of annual streamflow in each of the post-fire years. The difference between measured post-fire streamflow and predicted streamflow from the stationary statistical model provides an estimate of the streamflow change attributable to disturbance.

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

3 Results

The behavioral ensemble of 30 calibrated DHSVM parameter sets all satisfactorily reproduce observed streamflow hydrographs (Table 2, Fig. 3). Daily NSE values for the 2015-2024 calibration period vary between 0.80 and 0.89 (log NSE 0.80-0.85), with similar statistics on the 2005-2014 validation period (NSE 0.76-0.88, log NSE 0.80-0.89). All behavioral parameter sets also achieve satisfactory NSE (0.80-0.87) and log-scale NSE (0.76-0.84) considering just the four years after the Creek Fire. The mean post-fire bias varies by -9% to +6%. Comparing fire-aware and no-fire control scenarios, the behavioral ensemble indicates a bulk streamflow increase of +2 to +17% after the Creek Fire (median +12%). DHSVM also indicates a shift towards earlier snowmelt runoff after the Creek Fire, particularly in the snowy 2023 water year.

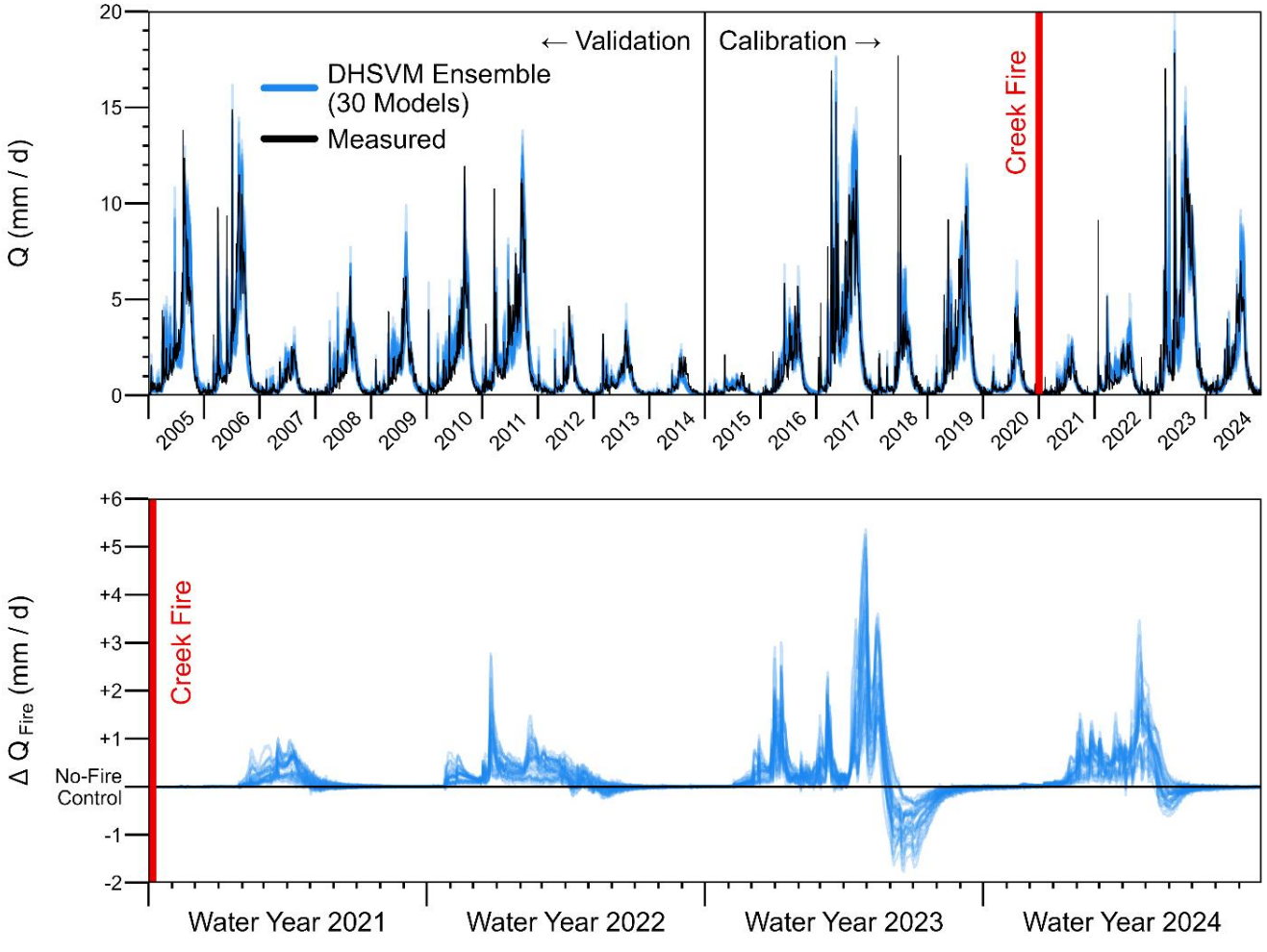

Figure 3: Modeled and measured daily streamflow hydrographs (top panel) and streamflow differences between fire-aware and no-fire control simulations (bottom panel). Both panels show results from 30 calibrated "behavioral" parameter sets (Sect. 2.2).

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

Uncertainty in the streamflow response to disturbance is large relative to the size of the effect, even after a megafire. The difference in total post-fire streamflow volume between fire-aware and no-fire control scenarios has a coefficient of variation of 41%. Some parameter sets predict up to a 650% larger streamflow response than other parameter sets (inter-model range of +13 to +97 mm/yr). Relative uncertainty is higher in dry years, with the simulated streamflow response in 2021 varying between +3 mm/yr and +47 mm/yr across different parameter sets (1,400% range). The predicted streamflow change after the Creek Fire is on the same order of magnitude as stochastic error in the annual water balance (Supplemental Fig. S3), which intuitively explains why the disturbance response remains uncertain despite direct calibration of pre- and post-fire streamflow (Fig. 3).

Uncertainty in the post-fire streamflow response is linked to equifinality in modeled water balance fluxes (Figs. 1, 4). To qualify as "behavioral," parameter sets must satisfactorily estimate the annual water balance (MAPE < 10%), but the model can achieve this in different ways. Some parameter combinations suggest that transpiration and interception loss from vegetation accounts for up to 95% of total pre-fire ET, while others suggest a vegetation contribution as low as 77%, with the balance contributed by evaporation from abiotic surfaces (stream channels and soil, including rock above treeline). Relatively dense initial forests (high area-average LAI) are associated with large decreases in post-fire transpiration and interception loss (Pearson $r = -0.92$, $p < 0.01$). Low transmissivity is associated with increases in post-fire soil evaporation and channel evaporation ($r = -0.99$, $p < 0.01$, both variables log-transformed).

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

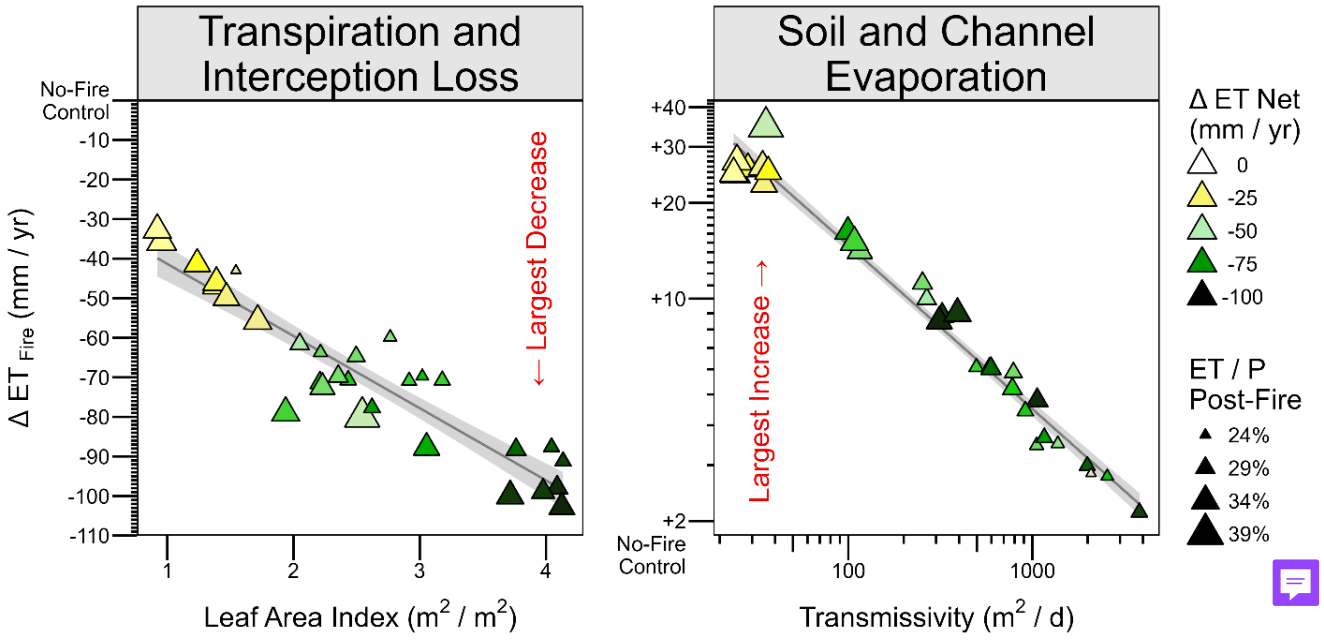

Figure 4: Difference in ET fluxes between fire-aware and no-fire control simulations visualized relative to model parameter uncertainty. The area-average leaf area index (LAI) is aggregated within the pre-fire forested area from maps of tree-scale LAI and grid-scale fractional cover, and the area-average transmissivity is aggregated from maps of soil depth, conductivity, and exponential decrease using the DHSVM transmissivity equations (Table 1). Trend lines indicate the least-squares fit and 90% confidence interval of the best-fit linear estimator.

Compensating errors in equifinal parameter sets can produce compounding discrepancies after disturbance. Not only do some parameter sets indicate much larger changes in individual fluxes, those with the smallest reductions in vegetation ET also exhibit the largest fractional compensation (up to 76%) from increased abiotic evaporation ($r = 0.71$, 0.01). A similar compensation between modeled overstory and understory ET components is illustrated by Boardman et al. (2025). Low calibrated transmissivity implies slower groundwater recharge and shallower flowpaths, contributing to higher soil evaporation, which compensates for low vegetation ET. These parameter sets are primed for large increases in evaporation when soil moisture increases in de-forested areas after fire. Consequentially, there is a negative correlation ($r = -0.93$, $p < 0.01$) between the fraction of pre-fire ET contributed by abiotic evaporation and the magnitude of the post-fire net ET reduction.

Evaluating the model bias shift (Eq. 1) can help escape this morass of uncertainty. Across the 30-member behavioral ensemble, there is a strong correlation ($r = 0.96$-0.99 depending on year, $p < 0.01$) between the mean streamflow bias shift after disturbance and the annual streamflow change attributable to fire (Fig. 5). Lines in Fig. 5 correspond to Eq. (2), and the horizontal axis is defined by Eq. (1). Bayesian sampling of a linear model conditioned on zero bias shift yields an estimate of

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

the uncertainty in the vertical intercept (Sect. 2.4), which is the predicted streamflow change of models with stationary bias. Comparing the annual streamflow errors of models with positive or negative bias shift (Supplemental Fig. S3), we note that the positive-shift models tend to have more-positive errors on the pre-fire period compared to negative-shift models, but this stratification reverses after the Creek Fire. This reversal of model over- and under-prediction after disturbance is consistent with our conceptual model in Fig. 1. Additionally, as shown by the shape-size in Fig. 5, the models with the largest over-prediction have anomalously high overstory LAI, and vice versa, which is similarly consistent with the conceptualization of ET_{Tree} and ET_{Other} equifinality in Fig. 1.

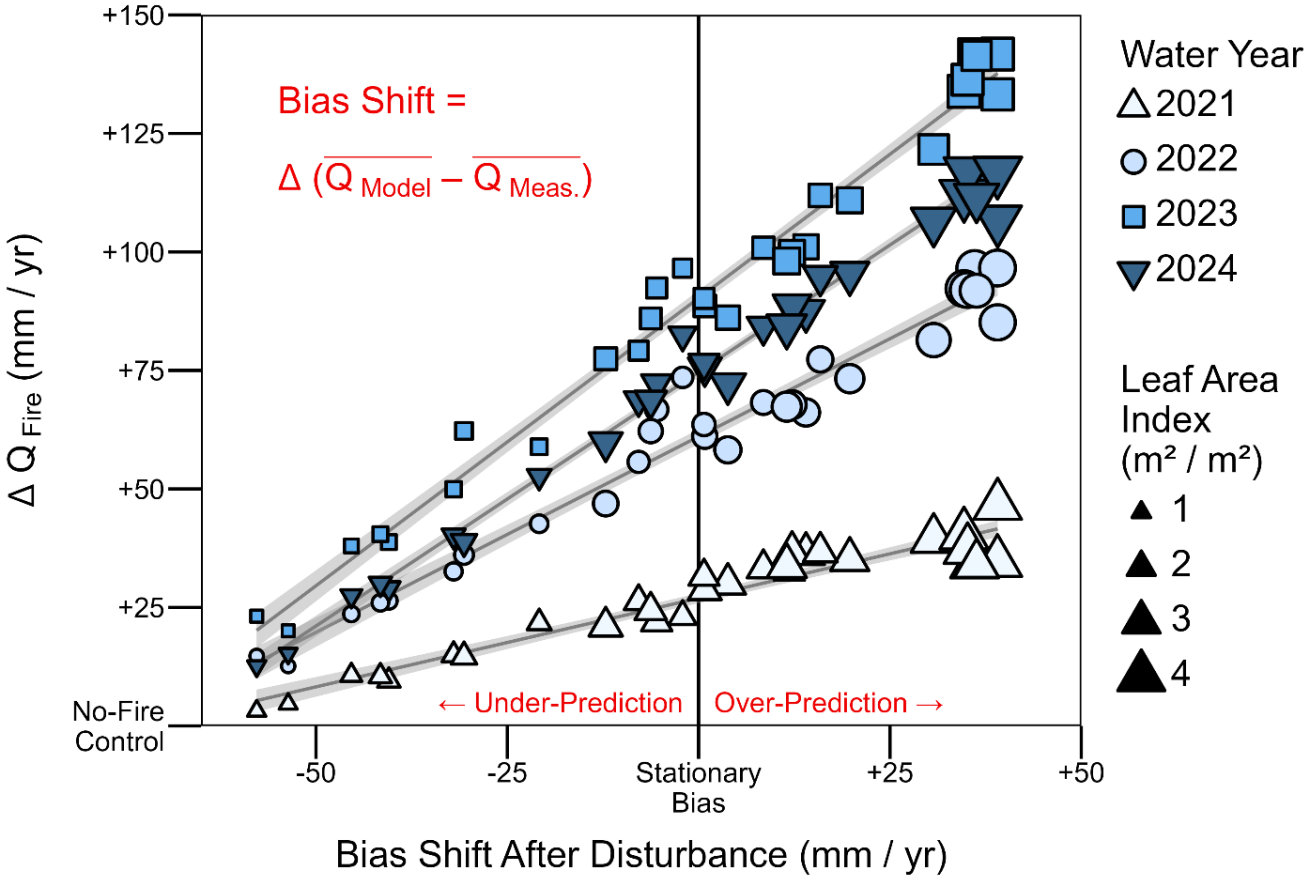

Figure 5: Annual post-fire streamflow change visualized relative to the mean bias shift after disturbance for all 30 parameter sets. Parameter sets with a shift towards overestimation predict a relatively large streamflow response to disturbance, and vice versa. Parameter sets with near-stationary bias are assumed to give the most accurate estimate of changes due to disturbance. Trend lines indicate the least-squares fit and 90% confidence interval of the best-fit linear estimator, distinct from the analogous Bayesian regression in Eq. (2), which also propagates parameter uncertainty.

Eight parameter sets result in near-stationary bias (shift less than ±10 mm/yr). This eight-member "stationary sub-ensemble" demonstrates how considering bias shift after disturbance can reduce equifinality. Compared to 10^4 alternative sub-

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

ensembles of eight parameter sets each randomly selected from the 30-member ensemble, the stationary sub-ensemble has significantly reduced uncertainty in LAI ($p < 0.01$) and transmissivity ($p < 0.02$), calculated from the cumulative distribution function for the fractional uncertainty reduction of all 10^4 sub-ensembles. This uncertainty reduction associated with the stationary sub-ensemble is 72% larger for LAI and 285% larger for log-transformed transmissivity compared to the median uncertainty reduction of same-sized sub-ensembles selected randomly. The stationary sub-ensemble has mean 2011-era LAI between 1.9-2.8 m^2/m^2, which is 74% less uncertainty compared to the full behavioral ensemble range of 0.9-4.1 m^2/m^2 (Fig. 5). Analogously, the stationary sub-ensemble has 50% less uncertainty in log-transformed mean transmissivity despite order-of-magnitude residual uncertainty (108-1378 m^2/d). Meteorological uncertainty remains mostly unchanged in the stationary sub-ensemble, with 3.7 °C and 15% uncertainties in temperature and precipitation biases, respectively, compared to 4.4 °C and 15% for the behavioral ensemble (no significant change relative to random sub-ensemble selection). This leads to similar uncertainty in the stationary sub-ensemble's post-fire evaporative index (ET / P) compared to the full ensemble (25-34% vs. 24-39% respectively). Among all 14 calibrated parameters (Table 1), only the melt-season albedo decay rate has a statistically significant difference in the mean ($p < 0.05$, Welch two-sample t-test) between the stationary sub-ensemble and the full 30-member ensemble. Instead, most of the equifinality reduction arises from shrinking the uncertainty of the parameter distributions rather than changing their mean (Supplemental Fig. S4) and/or from constraining multi-dimensional parameter interactions (Supplemental Fig. S1).

Compared to the full behavioral ensemble, the stationary sub-ensemble has slightly sub-optimal hydrograph fit (NSE 0.80-0.85 vs. 0.89 max), but generally better SWE volume error (MAPE of 18-27% across 30 ASO surveys vs. 32% for the highest-NSE parameter set). The stationary sub-ensemble has statistically lower (worse) mean NSE values compared to the 30-member ensemble ($p < 0.05$) and approaches the threshold for significantly lower (better) mean SWE volume percent error ($p = 0.055$). The >95th-percentile peak flow RMSE is also significantly worse ($p < 0.05$) for the stationary sub-ensemble. Differences in log-scale NSE and annual or April-July water yield error are not statistically significant. The improvement in snow skill despite a slight worsening of streamflow skill (Supplemental Fig. S5) may arise from overfitting during calibration, which leads to a tradeoff between enhanced model physical fidelity (represented by the near-zero bias shift and better snow performance of the stationary sub-ensemble) and minor degradation in the streamflow performance metrics.

Empirical regression and process-based simulations both suggest an increase in streamflow after the Creek Fire, albeit with different uncertainty ranges (Fig. 6). An empirical model (Eq. 3) fit to pre-fire data predicts relatively less post-fire streamflow than observed, implying a total streamflow increase of +12% with a 90% credible interval of +5 to +18% assuming that each year's error distribution is independent. (The 90% credible interval represents the 5th-95th percentiles.) Although the four-year total streamflow increase is significant at $p < 0.01$, for individual post-fire years we cannot reject the null hypothesis (no change after disturbance) at the $p < 0.01$ level, and we cannot even reject the no-change hypothesis at the

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

p < 0.1 level in 2021 or 2024. Compared to the pure statistical model based on the same meteorological data, our process-based modeling approach yields remarkably similar uncertainty. All 30 behavioral DHSVM parameter sets indicate at least some streamflow increase in each post-fire year, with a similar mean increase of +12% and a marginally wider 90% range of +3 to +17% across the ensemble. Although the 90% uncertainty ranges are similar between DHSVM and the empirical regression, all of the DHSVM parameter sets show at least some increase. Additionally, within individual years, the DHSVM uncertainty can be much lower (e.g., in 2023, the DHSVM 90% range is +2 to +11%, while the empirical regression 90% credible interval is +3% to +23%). The uncertainty of the empirical model benefits from considering all four years simultaneously, since the empirical model assumes that each year's fire effect is independent, while different DHSVM parameter sets are systematically biased high or low across all post-fire years.

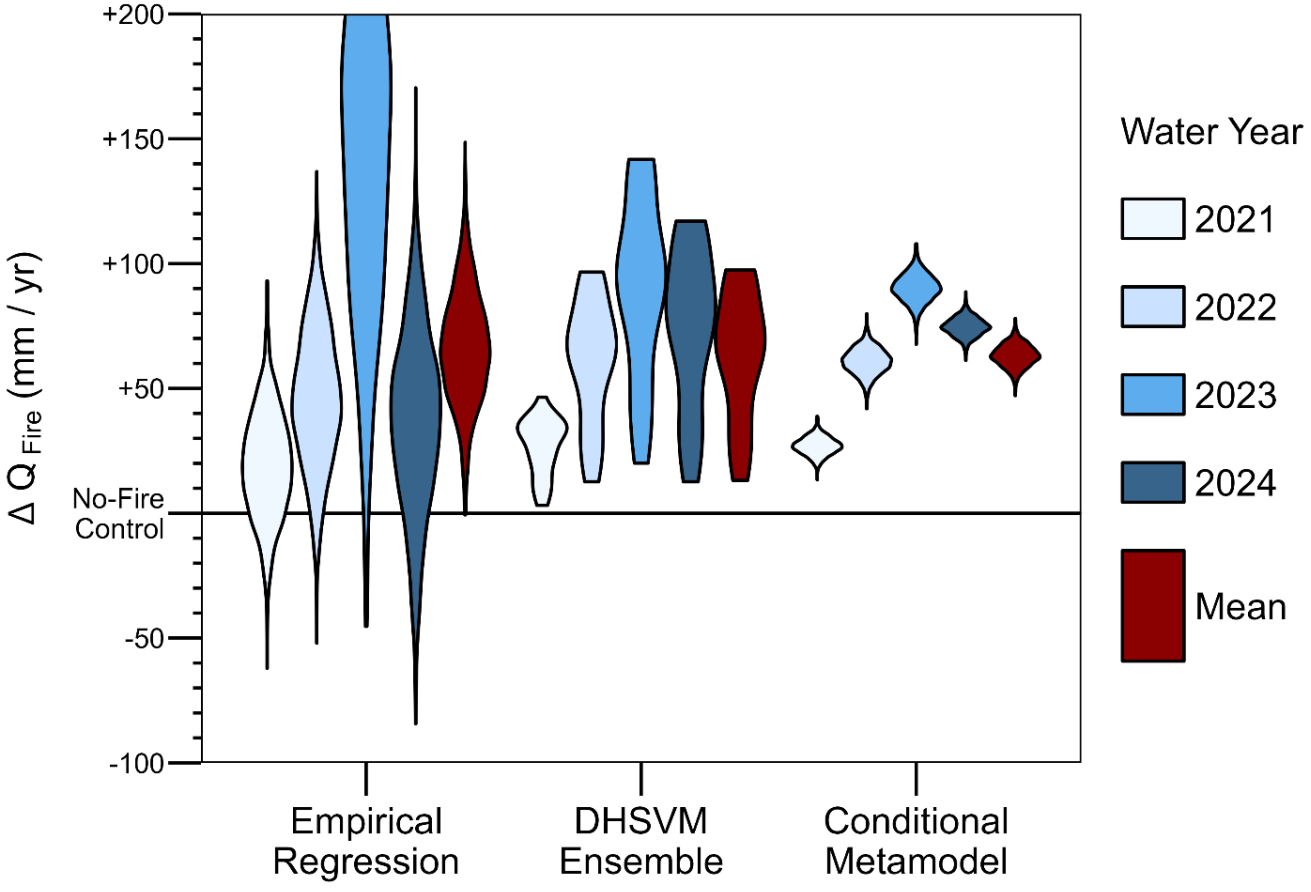

Figure 6: Uncertainty distributions for the annual post-fire streamflow change relative to a control scenario with no fire. Empirical regression results are estimated by comparing post-fire measurements with 1,000 random samples of a pre-fire multi-linear regression model (Eq. 3). The DHSVM ensemble represents the difference between fire-aware and no-fire control simulations using 30 different calibrated parameter sets. The conditional metamodel predicts the DHSVM response subject to the requirement of stationary bias using 1,000 random samples of the Bayesian regression in Eq. (2). (Note that the vertical axis is truncated at +200 mm/yr for increased visibility of most results.)

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

Compared to pure statistical or pure process-based approaches, a statistical metamodel trained on DHSVM results and conditioned on bias stationarity can drastically reduce uncertainty. Using 1,000 random samples of the metamodel (Eq. 2), we find a +11% increase in total post-fire streamflow with a 90% credible interval of +10 to +12%. The mean streamflow response is 14 standard deviations above zero, confidently rejecting the no-change hypothesis. Moreover, interannual variability in the conditional streamflow response is separable between all pairs of years at the $p < 0.01$ level. In contrast, raw DHSVM simulations of the streamflow response in 2022, 2023, and 2024 are not mutually separable at the $p < 0.05$ level. Comparing 90% credible intervals, the conditional approach reduces uncertainty in the total post-fire streamflow change by 80% compared to the empirical regression and 82% compared to the DHSVM ensemble. Of course we cannot know precisely what the true streamflow would have been without a fire, so some uncertainty must always remain, but our metamodel results suggest that we can substantially reduce this uncertainty by fusing process-based and statistical approaches.

4 Discussion

Hydrological models must include forest fires and other environmental disturbances to provide robust predictions for water resource management, risk assessment, and operational planning. In 2021, the first year after the Creek Fire, our hybrid modeling approach estimates that the additional streamflow attributable to forest disturbance provided 0.11 ±0.03 km^3 (92,000 ac-ft.) of extra water to Millerton Lake (a major regional reservoir, Fig. 2), which is 18% ±4% of the total water yield in a year where drought conditions caused curtailment of downstream water rights (California DWR 2021). In the wet 2023 water year, extra streamflow attributable to the fire totaled ~0.38 ±0.04 km^3 (310,000 ac-ft., 7% of total water yield), equivalent to an extra 60% ±7% of the reservoir storage capacity in a year with widespread flooding (California DWR 2023). (All uncertainty ranges indicate the 90% credible interval.) These examples illustrate the potential for major forest disturbances like forest megafires to enhance water resources and/or exacerbate water risks (e.g., Boardman et al., 2025). Accurately representing disturbance and accounting for other sources of nonstationarity should be a priority of ecohydrological modeling.

Our results suggest that equifinality demands more thoughtful consideration in hydrological model-based studies of disturbance. At the same time, studies investigating disturbance have a unique and underutilized opportunity to reduce model equifinality. Much of the spread in the DHSVM ensemble (Fig. 6) could be eliminated by reducing the number of calibrated parameters or narrowing their prior range (Table 1). However, in typical landscape-scale simulations, we do not know the "correct" parameter values. For example, there is considerable uncertainty in vegetation properties derived from satellite imagery (Garrigues et al. 2008, Tang et al. 2019) or extrapolated from sparse field data (Meyer et al. 2016). Moreover, in modeling applications, "effective" parameters may subsume additional sources of structural or data uncertainty (Dolman and

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

Blyth 1997, Vázquez 2003, Were et al. 2007), and some quasi-empirical parameters (e.g., grid-scale hydraulic conductivity) do not have single "correct" values (Beven 1993). Uncertainty in vegetation properties like LAI can produce significantly different streamflow predictions (e.g., Bart et al. 2016 and Fig. 5 of this study), and latent uncertainty could cause systematic biases (Fig. 4). Furthermore, even if vegetation properties could be tightly constrained, introducing parameter variability into model experiments can reveal compensating ecohydrological processes (Figs. 1, 4), counterintuitively leading to higher confidence in the statistical metamodel by providing datapoints for regression in Eq. (2). Nevertheless, reducing model parameter uncertainty is generally desirable when justified, and our results show that a sub-ensemble of parameter sets with near-stationary bias after disturbance can significantly reduce uncertainty in LAI ($p < 0.01$) and saturated transmissivity ($p < 0.02$).

Using process-based models for post-disturbance predictions based on traditional streamflow calibration metrics can be dangerously misleading. Describing model performance with simple goodness-of-fit metrics (e.g., NSE) is problematic in general due to sampling uncertainty and other issues (Clark et al. 2021), but these metrics remain ubiquitous in modeling studies (including this study) due to their ease of application and simplicity of interpretation. Although these metrics are useful for loosely identifying an initial ensemble of behavioral models, our results provide a clear example of the pitfalls in blindly trusting NSE-based (or similar) calibration strategies. In particular, the four models with the highest daily NSE (0.88-0.89) have anomalously small disturbance effects, and the parameter set with the absolute highest NSE underestimates the post-fire streamflow change by 79% relative to the metamodel mean (Supplemental Fig. S6). These outlying parameter sets are probably compensating for unknown deficiencies in the model structure and/or forcing data, leading the model to get a slightly higher NSE for what are apparently the wrong reasons. These undesirable and yet numerically optimal solutions are endemic to high-dimensional optimization problems, an issue known as "reward hacking" (Amodei et al. 2016). Although log-transformed NSE appears less vulnerable to reward hacking (Supplemental Fig. S6), the parameter set with the absolute highest log NSE still underestimates the metamodel-based mean streamflow change by 58%. It is noteworthy that our ultimate model evaluation metric (bias shift, Fig. 5) is not included in the calibration. If this metric were directly calibrated, it might be susceptible to reward hacking, leading to unreliable inference in Eq. (2).

With care, process-based models can remain powerful tools for hydrological investigation. Despite a recent focus on machine learning approaches to hydrological prediction (Ardabili et al. 2020, Xu and Liang 2021), purely empirical methods are limited by the amount of available data, the ability to assign clear process attribution, and the potentially ambiguous interpretation of nonstationarity (Slater et al. 2021). In the years immediately after a large forest disturbance (e.g., megafire), water managers may require rapid estimates of the potential hydrological impacts without the luxury of waiting for more data to accrue. Counterfactual model simulations can help isolate disturbance effects from stochastic weather and other variability, but it is notoriously challenging to quantify the uncertainty of distributed process-based models. Direct uncertainty estimation based on subjective likelihood metrics (e.g., GLUE, Beven and Binley 1992) is statistically

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

unjustified (Mantovan and Todini 2006, Stedinger et al. 2008). Nevertheless, by using a subjective sample of models to constrain sensitivity (Fig. 5), we can perform classical Bayesian inference (Eq. 2) to sample the relationship between an unknown outcome (e.g., the streamflow response to fire) and an observational constraint (e.g., bias stationarity). This framework overcomes the statistical limitations of process-based models by treating the results from equifinal parameter sets as independent data points that constrain a statistical metamodel (Eq. 2), which should be transferable to other disturbance studies. Moreover, the statistical metamodel provides an 82% uncertainty reduction compared to the raw DHSVM ensemble with just four years of post-fire streamflow data in a single watershed. Further calibration tests in other watersheds could inform a transferrable understanding of equifinality that might help constrain the post-fire streamflow response with fewer years of post-fire data, and could even help constrain predictions for possible future disturbances before they happen.

Our findings suggest a generally applicable conceptual framework for hydrological model simulations of many types of change beyond just environmental disturbance. In brief, it is important to ensure that our models are stationary with respect to the variability they are used to investigate. For our present study of the streamflow response to a forest fire, the model should have stationary bias across pre- and post-fire periods. For a study on climate warming effects, modelers could test the stationarity of model biases in warmer or cooler years or locations. For a drought study, model bias stationarity could be evaluated between wet and dry periods. Many analogous examples are easily imagined. We anticipate that testing bias stationarity across different types of change can help reduce both equifinality and uncertainty, as shown here.

5 Conclusion

In light of our results from this test application, we offer some general recommendations for process-based simulations of environmental disturbance in a hydrological context.

- Parameter uncertainty extends beyond the subsurface. We suggest calibrating or at least testing the sensitivity of parameters controlling the partitioning of ET fluxes (e.g., overstory/understory transpiration, interception loss, soil evaporation, etc.). Ideally, uncertainty in meteorological biases should also be propagated during calibration.

- Nonstationary error metrics (e.g., positive or negative bias shift) can indicate a failure to adequately represent change. Rejecting parameter sets with a bias shift after disturbance (or with respect to some other change) can help reduce equifinality and thus also reduce uncertainty in hydrological changes.

- Modelers should beware of “reward hacking” (i.e., overfitting) during calibration. In this study, selecting the highest-NSE parameter sets would lead to a 79% underestimation of the streamflow response to disturbance relative to the mean value of the conditional metamodel (analogous 58% underestimation using log NSE).

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

- We caution against direct derivation of uncertainty ranges from subjective parameter ensembles, as this could lead to unnecessarily high uncertainty with poor statistical justification (Fig. 5). Instead, model results can inform a statistical metamodel conditioned on observation-based metrics related to stationarity (e.g., bias shift), enabling the derivation of uncertainty from classical Bayesian inference (Eq. 2). The statistical metamodel structure will necessarily depend on the study objectives, but the principle is generally applicable to other models and types of environmental change.

Data and Code Availability

All model inputs/outputs, model code, analysis scripts, processed data, and other materials needed to understand and reproduce the results of this study will be archived on Zenodo after acceptance. In particular, the R script "5_BayesianModeling.R" may be useful for anyone interested in adapting our hybrid Bayesian framework (Eq. 2) to other projects.

Author Contribution

ENB and AAH formulated the project conceptualization. ENB developed the methodology and lead data curation, software development, visualization, and writing. GFSB contributed to methodology, visualization, and writing. MSW contributed to methodology and writing. RKS contributed to visualization and writing. AAH contributed to supervision, methodology, visualization, and writing.

Competing Interests

The authors declare that they have no conflict of interest relevant to this manuscript.

Financial Support

E. N. Boardman was partially supported by the NSF Graduate Research Fellowship Program under Grant #1937966. G. F. S. Boisrame was partially supported by NSF EAR #2011346 A. A. Harpold was partially supported by NSF EAR #2012310.

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

References

Abatzoglou, J. T. (2013). Development of gridded surface meteorological data for ecological applications and modelling. *International Journal of Climatology*, *33*(1), 121–131. https://doi.org/10.1002/joc.3413

Abolafia-Rosenzweig, R., Gochis, D., Schwarz, A., Painter, T. H., Deems, J., Dugger, A., Casali, M., & He, C. (2024). Quantifying the Impacts of Fire-Related Perturbations in WRF-Hydro Terrestrial Water Budget Simulations in California's Feather River Basin. *Hydrological Processes*, *38*(11), e15314. https://doi.org/10.1002/hyp.15314

Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., & Mané, D. (2016). *Concrete Problems in AI Safety* (arXiv:1606.06565). arXiv. https://doi.org/10.48550/arXiv.1606.06565

Ardabili, S., Mosavi, A., Dehghani, M., & Várkonyi-Kóczy, A. R. (2020). Deep Learning and Machine Learning in Hydrological Processes Climate Change and Earth Systems a Systematic Review. In A. R. Várkonyi-Kóczy (Ed.), *Engineering for Sustainable Future* (Vol. 101, pp. 52–62). Springer International Publishing. https://doi.org/10.1007/978-3-030-36841-8_5

Ayars, J., Kramer, H. A., & Jones, G. M. (2023). The 2020 to 2021 California megafires and their impacts on wildlife habitat. *Proceedings of the National Academy of Sciences*, *120*(48), e2312909120. https://doi.org/10.1073/pnas.2312909120

Bart, R. R. (2016). A regional estimate of postfire streamflow change in California. *Water Resources Research*, *52*(2), 1465–1478. https://doi.org/10.1002/2014WR016553

Bart, R. R., Tague, C. L., & Moritz, M. A. (2016). Effect of Tree-to-Shrub Type Conversion in Lower Montane Forests of the Sierra Nevada (USA) on Streamflow. *PLOS ONE*, *11*(8), e0161805. https://doi.org/10.1371/journal.pone.0161805

Bennett, A., Hamman, J., & Nijssen, B. (2020). MetSim: A Python package for estimation and disaggregation of meteorological data. *Journal of Open Source Software*, *5*(47), 2042. https://doi.org/10.21105/joss.02042

Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources*, *16*(1), 41–51. https://doi.org/10.1016/0309-1708(93)90028-E

Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, *320*(1), 18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007

Beven, K., & Binley, A. (1992). The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, *6*(3), 279–298. https://doi.org/10.1002/hyp.3360060305

Bieger, K., Rathjens, H., Allen, P. M., & Arnold, J. G. (2015). Development and Evaluation of Bankfull Hydraulic Geometry Relationships for the Physiographic Regions of the United States. *JAWRA Journal of the American Water Resources Association*, *51*(3), 842–858. https://doi.org/10.1111/jawr.12282

Binois, M., & Picheny, V. (2019). **GPareto**: An *R* Package for Gaussian-Process-Based Multi-Objective Optimization and Analysis. *Journal of Statistical Software*, *89*(8). https://doi.org/10.18637/jss.v089.i08

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

Birkel, C., Arciniega-Esparza, S., Maneta, M. P., Boll, J., Stevenson, J. L., Benegas-Negri, L., Tetzlaff, D., & Soulsby, C. (2024). Importance of measured transpiration fluxes for modelled ecohydrological partitioning in a tropical agroforestry system. *Agricultural and Forest Meteorology*, *346*, 109870. https://doi.org/10.1016/j.agrformet.2023.109870

Boardman, E. N. (2023). *DHSVM-WSF: The Distributed Hydrology Soil Vegetation Model Water Supply Forecast Platform* (Mountain Hydrology LLC White Paper 2). https://mountainhydrology.com/mountainhydrology_wp2_dhsvm-wsf/

Boardman, E. N., Duan, Z., Wigmosta, M. S., Flake, S. W., Sloggy, M. R., Tarricone, J., & Harpold, A. A. (2025). Restoring Historic Forest Disturbance Frequency Would Partially Mitigate Droughts in the Central Sierra Nevada Mountains. *Water Resources Research*, *61*(4), e2024WR039227. https://doi.org/10.1029/2024WR039227

Boardman, E. N., Renshaw, C. E., Shriver, R. K., Walters, R., McGurk, B., Painter, T. H., Deems, J. S., Bormann, K. J., Lewis, G. M., Dethier, E. N., & Harpold, A. A. (2024). Sources of seasonal water supply forecast uncertainty during snow drought in the Sierra Nevada. *JAWRA Journal of the American Water Resources Association*, *60*(5), 972–990. https://doi.org/10.1111/1752-1688.13221

Boisramé, G. F. S., Thompson, S. E., Tague, C. (Naomi), & Stephens, S. L. (2019). Restoring a Natural Fire Regime Alters the Water Balance of a Sierra Nevada Catchment. *Water Resources Research*, *55*(7), 5751–5769. https://doi.org/10.1029/2018WR024098

Breiman, L., Cutler, A., Liaw, A., & Wiener, M. (2002). *randomForest: Breiman and Cutlers Random Forests for Classification and Regression* (p. 4.7-1.2) [Dataset]. https://doi.org/10.32614/CRAN.package.randomForest

Buma, B. (2015). Disturbance interactions: Characterization, prediction, and the potential for cascading effects. *Ecosphere*, *6*(4), art70. https://doi.org/10.1890/ES15-00058.1

California DWR. (2021). *Water Year 2021: An Extreme Year*. https://water.ca.gov/-/media/DWR-Website/Web-Pages/Water-Basics/Drought/Files/Publications-And-Reports/091521-Water-Year-2021-broch_v2.pdf

California DWR. (2023). *Water Year 2023: Weather Whiplash, From Drought To Deluge*. https://water.ca.gov/-/media/DWR-Website/Web-Pages/Water-Basics/Drought/Files/Publications-And-Reports/Water-Year-2023-wrap-up-brochure_01.pdf

California DWR. (2024). *Database of full natural flow records for the SBF station* [Dataset]. California Data Exchange Center. https://cdec.water.ca.gov/index.html

Clark, M. P., Vogel, R. M., Lamontagne, J. R., Mizukami, N., Knoben, W. J. M., Tang, G., Gharari, S., Freer, J. E., Whitfield, P. H., Shook, K. R., & Papalexiou, S. M. (2021). The Abuse of Popular Performance Metrics in Hydrologic Modeling. *Water Resources Research*, *57*(9), e2020WR029001. https://doi.org/10.1029/2020WR029001

Cuo, L., Giambelluca, T. W., & Ziegler, A. D. (2011). Lumped parameter sensitivity analysis of a distributed hydrological model within tropical and temperate catchments. *Hydrological Processes*, *25*(15), 2405–2421. https://doi.org/10.1002/hyp.8017

Dewitz, J., & U.S. Geological Survey. (2019). *National Land Cover Database (NLCD) 2019 Products* [Dataset]. U.S. Geological Survey. https://doi.org/10.5066/P9JZ7AO3

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

Dolman, A. J., & Blyth, E. M. (1997). Patch scale aggregation of heterogeneous land surface cover for mesoscale meteorological models. *Journal of Hydrology*, *190*(3–4), 252–268. https://doi.org/10.1016/S0022-1694(96)03129-0

Du, E., Link, T. E., Gravelle, J. A., & Hubbart, J. A. (2014). Validation and sensitivity test of the distributed hydrology soil-vegetation model (DHSVM) in a forested mountain watershed. *Hydrological Processes*, *28*(26), 6196–6210. https://doi.org/10.1002/hyp.10110

Dupuy, D., Helbert, C., & Franco, J. (2015). DiceDesign and DiceEval: Two R Packages for Design and Analysis of Computer Experiments. *Journal of Statistical Software*, *65*, 1–38. https://doi.org/10.18637/jss.v065.i11

Ebel, B. A., & Loague, K. (2006). Physics-based hydrologic-response simulation: Seeing through the fog of equifinality. *Hydrological Processes*, *20*(13), 2887–2900. https://doi.org/10.1002/hyp.6388

Ebel, B. A., & Mirus, B. B. (2014). Disturbance hydrology: Challenges and opportunities. *Hydrological Processes*, *28*(19), 5140–5148. https://doi.org/10.1002/hyp.10256

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

Emmerich, M. T. M., Deutz, A. H., & Klinkenberg, J. W. (2011). Hypervolume-based expected improvement: Monotonicity properties and exact computation. *2011 IEEE Congress of Evolutionary Computation (CEC)*, 2147–2154. https://doi.org/10.1109/CEC.2011.5949880

Fan, Y., & Miguez-Macho, G. (2011). A simple hydrologic framework for simulating wetlands in climate and earth system models. *Climate Dynamics*, *37*(1), 253–278. https://doi.org/10.1007/s00382-010-0829-8

Fatichi, S., Vivoni, E. R., Ogden, F. L., Ivanov, V. Y., Mirus, B., Gochis, D., Downer, C. W., Camporese, M., Davison, J. H., Ebel, B., Jones, N., Kim, J., Mascaro, G., Niswonger, R., Restrepo, P., Rigon, R., Shen, C., Sulis, M., & Tarboton, D. (2016). An overview of current applications, challenges, and future trends in distributed process-based models in hydrology. *Journal of Hydrology*, *537*, 45–60. https://doi.org/10.1016/j.jhydrol.2016.03.026

Fisher, R. A., & Koven, C. D. (2020). Perspectives on the Future of Land Surface Models and the Challenges of Representing Complex Terrestrial Systems. *Journal of Advances in Modeling Earth Systems*, *12*(4), e2018MS001453. https://doi.org/10.1029/2018MS001453

Franks, S. W., Beven, K. J., Quinn, P. F., & Wright, I. R. (1997). On the sensitivity of soil-vegetation-atmosphere transfer (SVAT) schemes: Equifinality and the problem of robust calibration. *Agricultural and Forest Meteorology*, *86*(1–2), 63–75. https://doi.org/10.1016/S0168-1923(96)02421-5

Furniss, T. J., Povak, N. A., Hessburg, P. F., Salter, R. B., Duan, Z., & Wigmosta, M. (2023). Informing climate adaptation strategies using ecological simulation models and spatial decision support tools. *Frontiers in Forests and Global Change*, *6*. https://doi.org/10.3389/ffgc.2023.1269081

Garrigues, S., Lacaze, R., Baret, F., Morisette, J. T., Weiss, M., Nickeson, J. E., Fernandes, R., Plummer, S., Shabanov, N. V., Myneni, R. B., Knyazikhin, Y., & Yang, W. (2008). Validation and intercomparison of global Leaf Area Index

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

products derived from remote sensing data. *Journal of Geophysical Research: Biogeosciences*, *113*(G2). https://doi.org/10.1029/2007JG000635

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. (2022a). Spatially Distributed Overstory and Understory Leaf Area Index Estimated from Forest Inventory Data. *Water*, *14*(15), Article 15. https://doi.org/10.3390/w14152414

Goeking, S. A., & Tarboton, D. G. (2022b). 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

Goodhart, C. (1984). *Problems of Monetary Management: The U.K. Experience*. Springer.

Grayson, R. B., Moore, I. D., & McMahon, T. A. (1992). Physically based hydrologic modeling: 1. A terrain-based model for investigative purposes. *Water Resources Research*, *28*(10), 2639–2658. https://doi.org/10.1029/92WR01258

Gupta, S., Papritz, A., Lehmann, P., Hengl, T., Bonetti, S., & Or, D. (2022). Global Mapping of Soil Water Characteristics Parameters—Fusing Curated Data with Machine Learning and Environmental Covariates. *Remote Sensing*, *14*(8), 1947. https://doi.org/10.3390/rs14081947

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

Henn, B., Newman, A. J., Livneh, B., Daly, C., & Lundquist, J. D. (2018). An assessment of differences in gridded precipitation datasets in complex terrain. *Journal of Hydrology*, *556*, 1205–1219. https://doi.org/10.1016/j.jhydrol.2017.03.008

Her, Y., & Chaubey, I. (2015). Impact of the numbers of observations and calibration parameters on equifinality, model performance, and output and parameter uncertainty: Parameters, Observations, and Uncertainty. *Hydrological Processes*, *29*(19), 4220–4237. https://doi.org/10.1002/hyp.10487

Hirsch, R. M. (2011). A Perspective on Nonstationarity and Water Management. *JAWRA Journal of the American Water Resources Association*, *47*(3), 436–446. https://doi.org/10.1111/j.1752-1688.2011.00539.x

Huang, X., & Swain, D. L. (2022). Climate change is increasing the risk of a California megaflood. *Science Advances*, *8*(32), eabq0995. https://doi.org/10.1126/sciadv.abq0995

Johnson, R. S. H., & Alila, Y. (2023). Nonstationary stochastic paired watershed approach: Investigating forest harvesting effects on floods in two large, nested, and snow-dominated watersheds in British Columbia, Canada. *Journal of Hydrology*, *625*, 129970. https://doi.org/10.1016/j.jhydrol.2023.129970

Johnstone, J. F., Allen, C. D., Franklin, J. F., Frelich, L. E., Harvey, B. J., Higuera, P. E., Mack, M. C., Meentemeyer, R. K., Metz, M. R., Perry, G. L., Schoennagel, T., & Turner, M. G. (2016). Changing disturbance regimes, ecological memory, and forest resilience. *Frontiers in Ecology and the Environment*, *14*(7), 369–378. https://doi.org/10.1002/fee.1311

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

Jones, D. R., Schonlau, M., & Welch, W. J. (1998). Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global Optimization*, *13*(4), 455–492. https://doi.org/10.1023/A:1008306431147

Kang, T.-H., & Sharma, A. (2024). A state-controlled hypothesis test for paired-watershed experiments. *Journal of Hydrology*, *640*, 131664. https://doi.org/10.1016/j.jhydrol.2024.131664

Kelleher, C., McGlynn, B., & Wagener, T. (2017). Characterizing and reducing equifinality by constraining a distributed catchment model with regional signatures, local observations, and process understanding. *Hydrology and Earth System Sciences*, *21*(7), 3325–3352. https://doi.org/10.5194/hess-21-3325-2017

Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. *Proceedings of ICNN'95 - International Conference on Neural Networks*, *4*, 1942–1948 vol.4. https://doi.org/10.1109/ICNN.1995.488968

Khatami, S., Peel, M. C., Peterson, T. J., & Western, A. W. (2019). Equifinality and Flux Mapping: A New Approach to Model Evaluation and Process Representation Under Uncertainty. *Water Resources Research*, *55*(11), 8922–8941. https://doi.org/10.1029/2018WR023750

LANDFIRE. (2022). *LANDFIRE 2022 Existing Vegetation Type (EVT)* [Dataset]. U.S. Department of the Interior, Geological Survey, and U.S. Department of Agriculture. http://www.landfire/viewer

Manning, A. L., Harpold, A., & Csank, A. (2022). Spruce Beetle Outbreak Increases Streamflow From Snow-Dominated Basins in Southwest Colorado, USA. *Water Resources Research*, *58*(5), e2021WR029964. https://doi.org/10.1029/2021WR029964

Mantovan, P., & Todini, E. (2006). Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. *Journal of Hydrology*, *330*(1), 368–381. https://doi.org/10.1016/j.jhydrol.2006.04.046

Meili, N., Beringer, J., Zhao, J., & Fatichi, S. (2024). Aerodynamic effects cause higher forest evapotranspiration and water yield reductions after wildfires in tall forests. *Global Change Biology*, *30*(1), e16995. https://doi.org/10.1111/gcb.16995

Meyer, C., Weigelt, P., & Kreft, H. (2016). Multidimensional biases, gaps and uncertainties in global plant occurrence information. *Ecology Letters*, *19*(8), 992–1006. https://doi.org/10.1111/ele.12624

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J. (2008). Stationarity Is Dead: Whither Water Management? *Science*, *319*(5863), 573–574. https://doi.org/10.1126/science.1151915

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., Stouffer, R. J., Dettinger, M. D., & Krysanova, V. (2015). On Critiques of "Stationarity is Dead: Whither Water Management?" *Water Resources Research*, *51*(9), 7785–7789. https://doi.org/10.1002/2015WR017408

Moreno, H. A., Gupta, H. V., White, D. D., & Sampson, D. A. (2016). Modeling the distributed effects of forest thinning on the long-term water balance and streamflow extremes for a semi-arid basin in the southwestern US. *Hydrology and Earth System Sciences*, *20*(3), 1241–1267. https://doi.org/10.5194/hess-20-1241-2016

MTBS Project (USDA Forest Service / U.S. Geological Survey). (2022). *MTBS Data Access: Fire Level Geospatial Data* [Dataset]. http://mtbs.gov/direct-download

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

Pagano, T., Garen, D., & Sorooshian, S. (2004). Evaluation of Official Western U.S. Seasonal Water Supply Outlooks, 1922–2002. *Journal of Hydrometeorology*, *5*(5), 896–909. https://doi.org/10.1175/1525-7541(2004)005<0896:EOOWUS>2.0.CO;2

Painter, T. H., Berisford, D. F., Boardman, J. W., Bormann, K. J., Deems, J. S., Gehrke, F., Hedrick, A., Joyce, M., Laidlaw, R., Marks, D., Mattmann, C., McGurk, B., Ramirez, P., Richardson, M., Skiles, S. M., Seidel, F. C., & Winstral, A. (2016). The Airborne Snow Observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo. *Remote Sensing of Environment*, *184*, 139–152. https://doi.org/10.1016/j.rse.2016.06.018

Pomeroy, J. W., Gray, D. M., Hedstrom, N. R., & Janowicz, J. R. (2002). Physically Based Estimation of Seasonal Snow Accumulation in the Boreal Forest. *Proceedings of the 59th Eastern Snow Conference*, 93–108.

Pongratz, J., Dolman, H., Don, A., Erb, K.-H., Fuchs, R., Herold, M., Jones, C., Kuemmerle, T., Luyssaert, S., Meyfroidt, P., & Naudts, K. (2018). Models meet data: Challenges and opportunities in implementing land management in Earth system models. *Global Change Biology*, *24*(4), 1470–1487. https://doi.org/10.1111/gcb.13988

Rigge, M. B., Bunde, B., Shi, H., & Postma, K. (2021). *Rangeland Condition Monitoring Assessment and Projection (RCMAP) Fractional Component Time-Series Across the Western U.S. 1985-2020* (Version 2.0, October 2021) [Dataset]. U.S. Geological Survey. https://doi.org/10.5066/P95IQ4BT

Rigge, M., Homer, C., Shi, H., Meyer, D., Bunde, B., Granneman, B., Postma, K., Danielson, P., Case, A., & Xian, G. (2021). Rangeland Fractional Components Across the Western United States from 1985 to 2018. *Remote Sensing*, *13*(4), Article 4. https://doi.org/10.3390/rs13040813

Roustant, O., Ginsbourger, D., & Deville, Y. (2012). DiceKriging, DiceOptim: Two R Packages for the Analysis of Computer Experiments by Kriging-Based Metamodeling and Optimization. *Journal of Statistical Software*, *51*, 1–55. https://doi.org/10.18637/jss.v051.i01

Saksa, P. C., Conklin, M. H., Battles, J. J., Tague, C. L., & Bales, R. C. (2017). Forest thinning impacts on the water balance of Sierra Nevada mixed-conifer headwater basins. *Water Resources Research*, *53*(7), 5364–5381. https://doi.org/10.1002/2016WR019240

Salas, J. D., Rajagopalan, B., Saito, L., & Brown, C. (2012). Special Section on Climate Change and Water Resources: Climate Nonstationarity and Water Resources Management. *Journal of Water Resources Planning and Management*, *138*(5), 385–388. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000279

Shields, C. A., & Tague, C. L. (2012). Assessing the Role of Parameter and Input Uncertainty in Ecohydrologic Modeling: Implications for a Semi-arid and Urbanizing Coastal California Catchment. *Ecosystems*, *15*(5), 775–791. https://doi.org/10.1007/s10021-012-9545-z

Slater, L. J., Anderson, B., Buechel, M., Dadson, S., Han, S., Harrigan, S., Kelder, T., Kowal, K., Lees, T., Matthews, T., Murphy, C., & Wilby, R. L. (2021). Nonstationary weather and water extremes: A review of methods for their detection,

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

attribution, and management. *Hydrology and Earth System Sciences*, *25*(7), 3897–3935. https://doi.org/10.5194/hess-25-3897-2021

Soil Survey Staff. (2022). *Soil Survey Geographic (SSURGO) Database* [Dataset]. Natural Resources Conservation Service, United States Department of Agriculture. https://sdmdataaccess.sc.egov.usda.gov

Spear, R. C., & Hornberger, G. M. (1980). Eutrophication in peel inlet—II. Identification of critical uncertainties via generalized sensitivity analysis. *Water Research*, *14*(1), 43–49. https://doi.org/10.1016/0043-1354(80)90040-8

Stan Development Team. (2023). *Stan Modeling Language Users Guide and Reference Manual, 2.32.5* [Computer software]. https://mc-stan.org

Stedinger, J. R., Vogel, R. M., Lee, S. U., & Batchelder, R. (2008). Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. *Water Resources Research*, *44*(12). https://doi.org/10.1029/2008WR006822

Stephens, S. L., Bernal, A. A., Collins, B. M., Finney, M. A., Lautenberger, C., & Saah, D. (2022). Mass fire behavior created by extensive tree mortality and high tree density not predicted by operational fire behavior models in the southern Sierra Nevada. *Forest Ecology and Management*, *518*, 120258. https://doi.org/10.1016/j.foreco.2022.120258

Sun, N., Yan, H., Wigmosta, M. S., Leung, L. R., Skaggs, 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/10.1029/2018JD030140

Tague, C., Seaby, L., & Hope, A. (2009). Modeling the eco-hydrologic response of a Mediterranean type ecosystem to the combined impacts of projected climate change and altered fire frequencies. *Climatic Change*, *93*(1–2), 137–155. https://doi.org/10.1007/s10584-008-9497-7

Tang, H., Song, X.-P., Zhao, F. A., Strahler, A. H., Schaaf, C. L., Goetz, S., Huang, C., Hansen, M. C., & Dubayah, R. (2019). Definition and measurement of tree cover: A comparative analysis of field-, lidar- and landsat-based tree cover estimations in the Sierra national forests, USA. *Agricultural and Forest Meteorology*, *268*, 258–268. https://doi.org/10.1016/j.agrformet.2019.01.024

U.S. Geological Survey. (2019). *USGS TNM Hydrography (NHD)* [Dataset]. https://apps.nationalmap.gov/downloader/

Vázquez, R. F. (2003). Effect of potential evapotranspiration estimates on effective parameters and performance of the MIKE SHE-code applied to a medium-size catchment. *Journal of Hydrology*, *270*(3), 309–327. https://doi.org/10.1016/S0022-1694(02)00308-6

Vögeli, C., Lehning, M., Wever, N., & Bavay, M. (2016). Scaling Precipitation Input to Spatially Distributed Hydrological Models by Measured Snow Distribution. *Frontiers in Earth Science*, *4*. https://doi.org/10.3389/feart.2016.00108

Vrugt, J. A., & Beven, K. J. (2018). Embracing equifinality with efficiency: Limits of Acceptability sampling using the DREAM(LOA) algorithm. *Journal of Hydrology*, *559*, 954–971. https://doi.org/10.1016/j.jhydrol.2018.02.026

Were, A., Villagarcıa, L., Domingo, F., Alados-Arboledas, L., & Puigdefabregas, J. (2007). Analysis of effective resistance calculation methods and their effect on modelling evapotranspiration in two different patches of vegetation in semi-arid SE Spain. *Hydrol. Earth Syst. Sci.*

https://doi.org/10.5194/egusphere-2025-1877
Preprint. Discussion started: 21 May 2025

Wigmosta, M. S., Vail, L. W., & Lettenmaier, D. P. (1994). A distributed hydrology-vegetation model for complex terrain. *Water Resources Research*, *30*(6), 1665–1679. https://doi.org/10.1029/94WR00436

Williams, A. P., Livneh, B., McKinnon, K. A., Hansen, W. D., Mankin, J. S., Cook, B. I., Smerdon, J. E., Varuolo-Clarke, A. M., Bjarke, N. R., Juang, C. S., & Lettenmaier, D. P. (2022). Growing impact of wildfire on western US water supply. *Proceedings of the National Academy of Sciences*, *119*(10), e2114069119. https://doi.org/10.1073/pnas.2114069119

Xu, T., & Liang, F. (2021). Machine learning for hydrologic sciences: An introductory overview. *WIREs Water*, *8*(5), e1533. https://doi.org/10.1002/wat2.1533

Yang, Y., Roderick, M. L., Yang, D., Wang, Z., Ruan, F., McVicar, T. R., Zhang, S., & Beck, H. E. (2021). Streamflow stationarity in a changing world. *Environmental Research Letters*, *16*(6), 064096. https://doi.org/10.1088/1748-9326/ac08c1

Zambrano-Bigiarini, M., Rojas, & R. (2013). A model-independent Particle Swarm Optimisation software for model calibration. *Environmental Modelling & Software*, *43*, 5–25. https://doi.org/10.1016/j.envsoft.2013.01.004

Zehe, E., & Sivapalan, M. (2009). Threshold behaviour in hydrological systems as (human) geo-ecosystems: Manifestations, controls, implications. *Hydrology and Earth System Sciences*, *13*(7), 1273–1297. https://doi.org/10.5194/hess-13-1273-2009