Received: 12 September 2023 | Accepted: 17 June 2024

DOI: 10.1111/1752-1688.13221

AMERICAN WATER RESOURCES ASSOCIATION JAWRA JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION

RESEARCH ARTICLE

Sources of seasonal water supply forecast uncertainty during snow drought in the Sierra Nevada

Elijah N. Boardman[1,2] | Carl E. Renshaw[2] | Robert K. Shriver[3] | Reggie Walters[4] | Bruce McGurk[5] | Thomas H. Painter[6] | Jeffrey S. Deems[6,7] | Kat J. Bormann[6] | Gabriel M. Lewis[3,8] | Evan N. Dethier[2,9] | Adrian A. Harpold[1,3]

[1]Graduate Program of Hydrologic Sciences, University of Nevada, Reno, Reno, Nevada, USA

[2]Department of Earth Sciences, Dartmouth College, Hanover, New Hampshire, USA

[3]Department of Natural Resources and Environmental Science, University of Nevada, Reno, Reno, Nevada, USA

[4]Hetch Hetchy Water and Power, San Francisco Public Utilities Commission, San Francisco, California, USA

[5]McGurk Hydrologic, Orinda, California, USA

[6]Airborne Snow Observatories, Inc., Mammoth, California, USA

[7]National Snow and Ice Data Center, Boulder, Colorado, USA

[8]Center for Western Weather and Water Extremes, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California, USA

[9]Geology Department, Colby College, Waterville, Maine, USA

Correspondence
Elijah N. Boardman, Graduate Program of Hydrologic Sciences (Boardman), University of Nevada, Reno, Reno, Nevada, USA.
Email: eli.boardman@mountainhydrology.com

Funding information
National Science Foundation, Grant/Award Number: 1937966 and 1723990

Abstract

Uncertainty attribution in water supply forecasting is crucial to improve forecast skill and increase confidence in seasonal water management planning. We develop a framework to quantify fractional forecast uncertainty and partition it between (1) snowpack quantification methods, (2) variability in post-forecast precipitation, and (3) runoff model errors. We demonstrate the uncertainty framework with statistical runoff models in the upper Tuolumne and Merced River basins (California, USA) using snow observations at two endmember spatial resolutions: a simple snow pillow index and full-catchment snow water equivalent (SWE) maps at 50 m resolution from the Airborne Snow Observatories. Bayesian forecast simulations demonstrate a nonlinear decrease in the skill of statistical water supply forecasts during warm snow droughts, when a low fraction of winter precipitation remains as SWE. Forecast skill similarly decreases during dry snow droughts, when winter precipitation is low. During a shift away from snow-dominance, the uncertainty of forecasts using snow pillow data increases about 1.9 times faster than analogous forecasts using full-catchment SWE maps in the study area. Replacing the snow pillow index with full-catchment SWE data reduces statistical forecast uncertainty by 39% on average across all tested climate conditions. Attributing water supply forecast uncertainty to reducible error sources reveals opportunities to improve forecast reliability in a warmer future climate.

KEYWORDS
drought, snow hydrology, uncertainty analysis, water supply, forecasts, runoff efficiency, climate non-stationarity

Paper No. JAWR-23-0070-P of the *Journal of the American Water Resources Association* (JAWR).

Discussions are open until six months from publication.

Research Impact Statement

Quantifying uncertainty sources in seasonal water supply forecasts during snow drought can justify the adoption of snow measurement and runoff modeling approaches that improve forecast reliability.

1 | INTRODUCTION

Mountains are natural reservoirs with great importance for water resources globally and in the western United States, accumulating cold-season precipitation as snow and releasing it for downstream consumption throughout the growing season (Bales et al., 2006; Immerzeel et al., 2020; Viviroli et al., 2007). In many areas, the timing of downstream water availability is subject to further regulation by artificial storage, a practice necessary to buffer the effects of meteorological droughts and peak flows on daily to decadal time scales, though the historic balance between runoff timing and storage capacity is threatened by climate change (Barnett et al., 2005). Decisions about how to schedule water releases, diversions, and other management activities depend on managers' understanding of future water availability, including uncertainty (Pagano et al., 2004; Stillinger et al., 2021).

Statistical relationships derived from precipitation time series and other environmental data are one of the most widespread tools for anticipating water availability, exemplified by the wide variety of drought indices (Heim, 2002; Zargar et al., 2011). In snow-dominated environments, explicitly simulating the temporal lag effect of snowpack accumulation and ablation can better characterize the timing of drought emergence (Staudinger et al., 2014). Similar to drought indices, statistical water supply forecasts leverage time series of environmental observations to characterize water balance deviations, typically providing predictions of runoff volume over a temporal period of interest. Seasonal water supply forecasts based on snow and precipitation are used to anticipate future runoff volumes in mountain watersheds and thus inform water storage and release schedules (Hartmann et al., 2002). Snow datasets are particularly useful for water supply forecasting because they integrate precipitation throughout the winter season, and snow data have historically provided one of the largest contributions to seasonal water supply forecast skill in the western United States (Koster et al., 2010).

Volumetric water supply forecasts are a crucial component of water management in environments like the California Sierra Nevada mountain range in the western United States (Dettinger & Anderson, 2015). However, forecasts always contain some degree of uncertainty, often represented as cumulative exceedance probabilities for a particular period, such as the spring/summer runoff season. Like all models, probabilistic forecasts are vulnerable to model misspecification, faulty assumptions, and data errors, giving rise to an issue of uncertainty in the uncertainty. As such, water management decisions are often hedged against the possibility of worse-than-expected droughts (Stillinger et al., 2021). Identifying sources of uncertainty and propagating their effects can improve the statistical robustness and transparency of seasonal water supply forecasts, increasing confidence in forecast-informed water management decisions.

Climate change poses challenges to water management paradigms that have been historically successful by shifting the timing of downstream water availability earlier in the season (Barnett et al., 2005; Knowles et al., 2006). The western United States, including the Sierra Nevada, faces a particularly great risk of unmet warm-season water demand due to climate-induced snow drought (Li et al., 2017; Mankin et al., 2015). The high-precipitation winters and low-precipitation springs typical of California's Mediterranean climate have historically enabled seasonal planning using snow-based runoff forecasts in the Sierra Nevada (Pagano et al., 2004). However, significant discrepancies (40%–129% error) have also been noted between forecast values and observed water yields in certain years (Dozier, 2011). For example, during the low runoff season of 2021, surprising forecast overpredictions led to calls for an investigation of "missing" streamflow (Lapides et al., 2022).

As anthropogenic climate warming continues to alter the Sierra snowpack, temperature-related impacts are expected to disproportionately affect mountain regions due to elevation-dependent warming (Mountain Research Initiative EDW Working Group, 2015). Temperature increases on the order of 1°C are likely to lead to increased probabilities of below-average snowpack storage in the Sierra (Huning & AghaKouchak, 2018), and the western United States has experienced a greater percentage increase in snow drought duration compared to other mountain regions over recent decades (Huning & AghaKouchak, 2020). Models anticipate a decline in the winter snowpack in California and across much of western North America (Mote et al., 2018), with periods of low-to-no snow winters possible in California within decades (Siirila-Woodburn et al., 2021). Furthermore, legacy in situ snow monitoring technologies, such as snow pillows and snow courses, will lose predictive power as the snowpack recedes above existing monitoring stations (Livneh & Badger, 2020). Snowpack quantification techniques incorporating remote sensing can potentially mitigate this degradation of station-based snow information by observing the full-catchment snowpack.

Considerable effort has been invested in improving current and future forecast skill with more accurate and representative snow measurements. For example, the NASA/JPL Airborne Snow Observatory (ASO, now Airborne Snow Observatories, Inc.) improved the quantification

of spatially distributed, full-catchment snow water equivalent (SWE) by combining snow depth maps acquired through airborne light detection and ranging (lidar) with spectrometer measurements and physical snowpack modeling using iSnobal (Marks et al., 1999; Painter et al., 2016). Assimilation of ASO data into snow models has resulted in large increases in model skill in snowmelt-dominated basins (Hedrick et al., 2018). Quantifying the reduction in uncertainty associated with ASO snow monitoring techniques could justify the adoption of new forecast methods where appropriate. However, to attribute forecast improvements to a particular dataset, it is necessary to isolate the snowpack quantification method from other sources of uncertainty.

In hydrological modeling, uncertainty arises in many forms. Beven (2016) categorizes uncertainty as aleatory, or reducible to a stationary statistical distribution; epistemic, from a lack of knowledge about the appropriate system dynamics or forcing/response data; semantic, from ambiguities in the concepts used to discuss processes or models; and ontological, from differences in beliefs about appropriate frameworks or assumptions. Statistical approaches to water supply forecasting have typically focused on aleatory (random) uncertainty, estimating total forecast uncertainty through cross-validation standard error or similar techniques (Garen, 1992). In his introduction to forecasts based on principal components regression, Garen (1992) argues that "the source of forecast error does not matter in decision making as long as the error can be quantified and described probabilistically." However, this assertion overlooks the importance of separating uncertainty sources to identify opportunities for improvement. Moreover, not all sources of uncertainty are best treated statistically, particularly those arising from "insidious" data errors (Lundquist et al., 2015) or non-stationarity (Milly et al., 2008).

Although the broad areas of uncertainty in water supply forecasts are understood conceptually, historical conceptualizations have seen limited investigation in the context of climate nonstationarity or recent advances in snow monitoring technology. Schaake and Peck (1985) broadly characterized the sources of forecast uncertainty as climatological uncertainty, model uncertainty, and data error. Although uncertainty in the future climate dominates hydrological uncertainty on multi-decadal time periods (Her et al., 2019), in the context of seasonal water supply forecasts, climatological uncertainty refers to variability in post-forecast weather, such as the amount of spring precipitation. In operational forecasting, post-forecast forcings are often constrained by historical climate variability or numerical weather predictions (e.g., ensemble streamflow prediction: Day, 1985; Troin et al., 2021). Model uncertainty arises from structural biases including the misrepresentation of physical processes and incomplete parameterization (e.g., missing snowmelt processes in a physical model: Webb et al., 2022), in addition to random errors from imperfect calibration of parameter values and uncertainty from equifinality (e.g., Beven, 2006). Data errors can accumulate from many sources, categorized by McMillan et al. (2018) as errors in the measurement of physical quantities, the calculation of derived quantities, interpolation between sparse spatiotemporal measurements, spatial scaling of properties, and human error associated with data management. The variety of climatological, model, and data uncertainty sources each offer a potential avenue to understand total forecast uncertainty and perhaps improve water supply forecast skill by ameliorating a particular uncertainty source. Schaake and Peck (1985) argued that the potential for improvement in seasonal water supply forecast skill has a lower bound defined by the skill of existing forecasts and an upper bound defined by the unpredictable climate variability inherent in natural systems.

The snow drought concept was developed to help understand the connections between meteorological conditions and hydrological drought during years with lower than normal snowpack storage. Snow drought refers to a lack of snow compared to a historical baseline and can be classified as "dry" or "warm" (Harpold et al., 2017; Hatchett & McEvoy, 2018). "Dry snow drought" results from dry conditions when a lack of precipitation (meteorological drought) causes below-normal snow accumulation, and "warm snow drought" results from warm conditions when the snowpack melts earlier than normal or a larger fraction of precipitation falls as rain. High interannual precipitation variability in the Sierra Nevada mountain range results in management challenges from alternating drought and flood years. Moreover, snow storage in western North America is vulnerable to an increasing frequency of warm snow droughts as global temperatures increase (Dierauer et al., 2019). The frequency of consecutive snow drought years resulting from warmer conditions is expected to increase by almost an order of magnitude by mid to late century (Marshall et al., 2019). Investigating the dynamics of snowpack measurement error, hydrological model uncertainty, and post-forecast precipitation variability during snow droughts is necessary to examine vulnerabilities in existing forecast methods and identify opportunities to make water supply forecasts more resilient to climate change.

To quantify the anticipated impact of climate change on the sources of uncertainty in snow-driven seasonal water supply forecasts, we develop a statistical framework to attribute total forecast uncertainty to different sources. We demonstrate one application of this framework using an ensemble of simple statistical runoff models and two different snow observation-based datasets, a snow pillow index and ASO SWE maps, selected as endmembers of spatial resolution and technological sophistication. Our investigation of forecast uncertainty in this test case addresses the following research question: what fraction of seasonal water supply forecast uncertainty is attributable to snowpack quantification methods, post-forecast precipitation variability, and statistical runoff model error, and how do these uncertainty sources change during snow drought?

2 | METHODS

2.1 | Study area

We model the impact of snow drought on the sources of water supply uncertainty by propagating model error, snowpack quantification uncertainty, and post-forecast precipitation uncertainty through a set of statistical runoff models in the Tuolumne River basin (1184 km^2, mean elevation 2721 m) and the Merced River basin (471 km^2, mean elevation 2753 m), located on the west side of the southern California Sierra Nevada mountain range (Figure 1). The Tuolumne River provides water to the City of San Francisco and other Bay Area communities, and both rivers provide water to agricultural districts in California's Central Valley. Water supply forecasts for the Tuolumne and Merced Rivers are operated by the California Nevada River Forecast Center using process-based runoff models and by the California Department of Water Resources using statistical methods. In-house forecasts are also operated by multiple reservoir management groups using a variety of physical and statistical models (e.g., as described by Chris Graham for Hetch Hetchy Reservoir, 2018 Yosemite Hydroclimate Meeting, https://cafiresci.com/s/Graham_ASO_inTuolumneForecastingComp.pdf). These models typically report uncertainty using probabilistic exceedance levels derived from an ensemble of model runs using variable post-forecast weather conditions.

We use a Bayesian statistical framework to evaluate a suite of simplified statistical runoff models for the Tuolumne and Merced watersheds across a range of snow drought scenarios. The models were developed in collaboration with the management and forecasting team at Hetch Hetchy Reservoir, part of the San Francisco Public Utilities Commission, to represent a range of structural complexities and data streams that are analogous to operational statistical forecast methods.

2.2 | Runoff models

To analyze forecast behaviors under a range of hydrologic and climatic conditions, we implement four statistical runoff models for the Tuolumne and Merced Rivers and propagate uncertainty through the parameters and inputs of each model. The conceptual structure of many

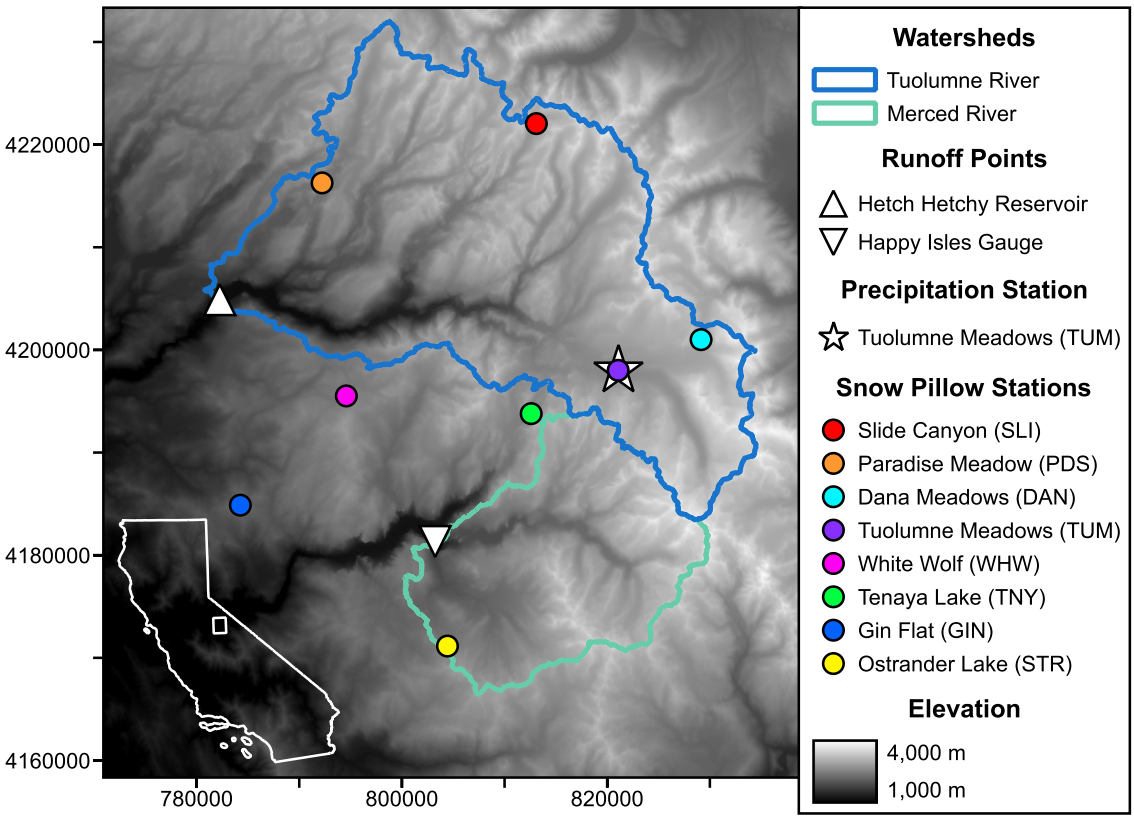

FIGURE 1 Map of the upper Tuolumne and Merced River watersheds. A snow pillow index is calculated from eight stations spanning the study domain (listed from north to south). Precipitation is measured at a central station. Runoff is measured at a reservoir on the Tuolumne River and a stream gauge on the Merced River. The inset map shows the location of the study area in the State of California.

lumped runoff models consists of an algebraic expression relating runoff, Q, to water inputs from precipitation, P, and snowpack ablation, ΔSWE, under the influence of one or more additional terms. We generalize this framework to the expression:

$$Q = P + \Delta\text{SWE} - L, \tag{1}$$

where L is a generalized loss function, and all terms are in area-normalized water depth units. The loss function, also known as the water balance residual, includes the lumped effects of evapotranspiration (ET), surface or subsurface storage changes, and groundwater import or export. There is negligible artificial regulation or anthropogenic consumption of water upstream from the runoff measurement points in the study watersheds, and groundwater effects are expected to be small in this region of the Sierra Nevada due to the glaciated granitic lithology. To model cumulative water yield over an extended forecast period of several months, the terms Q, P, and L are further simplified from instantaneous rates to time-integrated volumes, or equivalently, area-normalized specific yields in units of water depth.

We define runoff equations applicable to the time interval from October 1 through July 31. This time interval includes the regional snowpack accumulation and ablation seasons and captures the target April–July forecast period in addition to the October–April look-back window used to generate forecasts. We define a water balance on this time interval,

$$L_{\text{Total}} \equiv P_{\text{Total}} - Q_{\text{Total}} \tag{2}$$

such that Q_{Total} is the total runoff measured between October 1 and July 31, P_{Total} is the cumulative precipitation on the same time interval, and L_{Total} is defined as the difference between the former terms. Since the full-season period includes the normal snowpack accumulation and ablation seasons, ΔSWE is assumed to be zero and is excluded from Equation 2, although this assumption may contribute to model error in very high snow years if some snow persists into the late summer (e.g., 2017 in the Sierra). In different cases, the terms constituting L_{Total} may have the same or opposite sign; for instance, groundwater contributions to streamflow associated with a reduction in groundwater storage during droughts could cause L_{Total} to be smaller than ET. We develop four models for L_{Total} to represent a variety of potential assumptions that can underpin empirical runoff models, defined by the functional forms of L_{Total}:

$$L_1 = a, \tag{3.1}$$

$$L_2 = b \times P_{\text{Total}}, \tag{3.2}$$

$$L_3 = a + b \times P_{\text{Total}}, \tag{3.3}$$

$$L_4 = a + b \times P_{\text{Total}} - c \times Q_{\text{LastYear}} \tag{3.4}$$

Model 1 (Equation 3.1) assumes L_{Total} is a constant, a, where the difference P–Q is the same each year. This model is assumed to approximate a storage-limited system resulting in constant ET demand that can be satisfied even during drought conditions.

Model 2 (Equation 3.2) calculates L_{Total} by applying a scale factor, b, to P_{Total}, resulting in a constant runoff efficiency, Q/P. This model framework approximates a purely supply-limited system where ET increases in direct proportion to seasonal precipitation.

Model 3 (Equation 3.3) provides a linear structure for the loss term, with L_{Total} dependent on the current year's precipitation and a constant offset. The architecture of this model still assumes a closed water balance on seasonal time scales but allows for the relationship of ET and precipitation to vary between the endmember scenarios of Model 1 and Model 2.

Model 4 (Equation 3.4) adds a second slope term, c, to L_{Total} that is applied to the previous year's total streamflow, Q_{LastYear}, as a proxy for antecedent deficit or carryover effects. This framework implies an open water balance with a sponge-like storage system. In this case, L_{Total} is recursively dependent on each prior year's water balance. We note that the antecedent streamflow, Q_{LastYear}, is calculated for the full October 1 to September 30 period, while the model only predicts runoff up to the end of the forecast period (July 31).

To enable analysis of water supply forecasts on variable time scales, which is necessary to accommodate the opportunistic and irregular timing of ASO SWE acquisitions, we develop equations to scale L_{Total} over subsets of the runoff season. Runoff efficiency, R,

$$R \equiv \frac{Q_{\text{Total}}}{P_{\text{Total}}} \tag{4}$$

can be used to establish a scalable relationship between Q and P on the variable-length periods from each ASO acquisition through the end of the forecast period. Although R varies depending on the time period under consideration, our uncertainty simulations are restricted to a single date (April 1), so the assumption of a constant value for R in a given runoff season for each watershed should not impact the intercomparison of uncertainty. We calculate cumulative runoff during a forecast period as R multiplied by a water balance input, which is the complete ablation of the current snowpack plus any precipitation falling during the post-forecast period, P_{Spring}:

$$Q = R \times (\text{SWE} + P_{\text{Spring}}). \tag{5}$$

2.3 | Data sources

Evaluating Equation 5 requires measuring cumulative runoff, SWE storage, and post-forecast precipitation accumulation. Cumulative runoff is calculated from the day of each ASO flight through the end of July by summing daily inflow estimates to Hetch Hetchy Reservoir, which were smoothed to prevent unphysical negative values while preserving total volumes (McGurk, personal communication, 2022), and hourly discharge values at the Happy Isles stream gauge (station number 11264500, U.S. Geological Survey, 2022) for the Tuolumne and Merced Rivers, respectively. Since the lag between snowmelt and runoff is assumed to be an inconsequential fraction of the seasonal forecast period in these steep, mountainous study watersheds, we neglect runoff travel time in our calculations for the sake of parsimony.

Our primary approach to high-resolution snowpack quantification leverages a dataset of 59 ASO SWE maps across winter/spring of 10 years (2013–2022) in the Tuolumne and 21 ASO SWE maps across winter/spring of years (2014–2015, 2018–2022) in the Merced (Painter et al., 2018). For each ASO flight over the study basins between the months of January and June, inclusive, we aggregate the SWE maps from the 50 m scale to mean SWE over the watershed. Additional July flights are excluded because they are too close to the end of the runoff season to give a representative estimate of seasonal runoff dynamics around the April 1 forecast date. We note that the study period includes two notable multi-year droughts (2013–2015 and 2020–2022), which increases the robustness of our subsequent snow drought analysis.

To evaluate the effect of the variable snowpack distribution on station-based forecasts, we use an index of eight snow pillow stations (Figure 1) as a secondary method of snowpack quantification (California Department of Water Resources, 2022). For methodological consistency, we calculate the snow pillow index only on the dates of ASO flights. Due to the nature of remote mountain infrastructure, not all snow pillows are online at any given time, causing missing data that could bias a multi-station index. To correct for this, we use the time series of all non-missing snow pillows and a seasonal sinusoid to fit Gaussian Process regression models (implemented in the R DiceKriging package, Roustant et al., 2012) in order to impute missing measurements. This process attempts to emulate the effect of making inferences about missing data in real-time using all available information; that is, separate imputation models are constructed for every combination of missing and non-missing stations that arises over the study period. A sensitivity test using the pre-imputation mean across all available stations on each flight date shows that the imputation has minimal effect on the snow pillow index.

We calculate cumulative precipitation using the Tuolumne Meadows meteorological station (TUM) precipitation gauge (California Department of Water Resources/DFM-Hydro-SMN, 2022). The station is centrally located within the study area at an elevation of 2635 m, approximately 100 m below the mean watershed elevations of 2721 m (Tuolumne) and 2753 m (Merced). Unlike many other precipitation gauges in mountainous regions, the TUM station is maintained year-round by National Park staff living nearby, so snow bridging and similar problems are expected to be minimal. A sensitivity test shows that cumulative post-forecast TUM precipitation values from each ASO flight date through the end of the forecast period (July 31) are consistent with gridMET (Abatzoglou, 2013) basin-average precipitation fields (Tuolumne correlation = 0.95, RMSE = 5 cm; Merced correlation = 0.92, RMSE = 7 cm). Station-based precipitation measurements are susceptible to under-catch and do not capture variability in the spatial pattern of precipitation, but the availability of near real-time data can support the use of station measurements in forecast operations. The uncertainties introduced by assuming that the TUM station is representative of watershed precipitation are propagated through the Bayesian framework and contribute to the model backcast error, similar to the treatment of uncertainty in other data sources such as reservoir inflows computed from release rates and water level measurements.

2.4 | Bayesian framework

A Bayesian framework is well-suited to uncertainty propagation from random and systematic sources because it quantifies our degree of certainty in beliefs about future outcomes. A set of posterior parameter samples provides an estimate of uncertainty in the mathematical shape of the model and thus quantifies uncertainty in the model's mean output, μ_{QTotal}, while the standard deviation of the model's residual errors, σ, quantifies the effect of random process variability. Each model is calibrated independently in a particular watershed using a given snow data source (ASO SWE or the snow pillow index). To calibrate the runoff models, we choose the parameters of L_{Total} to minimize the percent uncertainty of modeled seasonal runoff, $Q_{TotalModeled}$, which is sampled from a distribution of possible runoff values (Equation 6), where μ_{QTotal} is the mean of the Q_{Total}, σ is the standard deviation, and the tilde symbol means "distributed as:"

$$Q_{TotalModeled} \sim \text{normal}\left(\mu_{QTotal}, \sigma\right). \tag{6}$$

However, since there are multiple irregular ASO SWE observations within most years, each defining a unique model run, we cannot directly sample the Q_{Total} distribution. Instead, we normalize the difference between measured post-flight runoff values, $Q_{Measured}$, and modeled post-flight runoff values, $Q_{Modeled}$, with the full-season runoff value, $Q_{TotalMeasured}$, measured from October 1 through July 31, which gives an estimate of the error in each runoff period as a fraction of the full-season runoff. We then update the target log probability density function (lpdf) of the Bayesian sampler by sampling from a normal distribution with mean seasonal error of 0 and standard deviation σ according to Equation 7:

$$\text{target lpdf} = \frac{1}{N_i}\text{normallogPDF}\left(\frac{Q_{\text{Modeled}} - Q_{\text{Measured}}}{Q_{\text{TotalMeasured}}}\middle|0,\sigma\right), \tag{7}$$

where N_i is the number of flights (January–June) in the year currently being sampled. This approach has the effect of weighting each year equally in the calibration, so that years with several ASO flights are not over-fit relative to years with fewer flights. All runoff periods end on July 31, but since each runoff period begins with a particular ASO flight date, some runoff periods are longer than others. By normalizing the model error by $Q_{\text{TotalMeasured}}$, we weight longer runoff periods more heavily in the calibration since longer periods more closely approximate the full seasonal water balance. A simulated forecast using a given set of model parameters can then be reconstructed for a subseasonal timeframe, such as the period of interest from April 1 to July 31, by calculating a deterministic value for post-forecast runoff, μ_{QModeled}, and sampling the model error as a fraction of $Q_{\text{TotalModeled}}$, the total modeled runoff from October 1 to July 31:

$$Q_{\text{Modeled}} \sim \mu_{\text{QModeled}} + \text{normal}(0,\sigma) \times Q_{\text{TotalModeled}} \tag{8}$$

This unique approach to model calibration is designed to handle multiple nested runoff intervals without biasing the model toward over-representation of years with more ASO flights. Weighting each interval by the fraction of total seasonal runoff that it represents results in consistent relative uncertainty between different years while minimizing the absolute uncertainty in each year's total water balance. In this study, we are interested in seasonal water supply uncertainty on fixed April–July intervals, so we focus on reducing uncertainty in the total water balance instead of optimizing error in runoff timing.

We use the No U-Turn Sampler (NUTS) implementation of Hamiltonian Monte Carlo in the Stan Bayesian language (Stan Development Team, 2023) to obtain 1000 parameter samples on each of two Markov chains (2000 samples total) for each of the four models in a particular basin with a particular snow data source. Within each basin the models using snow pillow or SWE map data are calibrated separately since any data errors in these snow quantification methods must be propagated into the model's standard error distribution. The runoff models are purely empirical, so we lack informative priors on the model parameters and thus sample from flat priors since the models are linear functions. Each Markov chain undergoes 5000 warmup iterations to tune the Hamiltonian Monte Carlo parameters before drawing the desired 1000 parameter samples. We observe that effective sample sizes and plots of parameter samples are consistent with convergence.

Using the calibrated models, we analyze model behaviors by calculating median values of L_{Total}, Q_{Total}, and R for each model across a P_{Total} range of 0.4 to 1 m, which is subset from the observed range of 0.43 to 1.89 m to highlight the nonlinear behavior of R during drought conditions. Additionally, we calculate median runoff values for each of the ASO flight-based calibration periods using ASO SWE and post-flight precipitation measurements to evaluate the historical performance of the hydrological models.

2.5 | Forecast simulations

To analyze the runoff models in forecast mode, that is, only using information that would be available in real-time, we replace measured values of P_{Spring} with a distribution of possible values than can be propagated through the models to generate probabilistic runoff exceedance levels. The TUM station precipitation record provides 37 years of precipitation records over water years 1986–2022. We calculate the total precipitation recorded by the gauge from the April 1 forecast date through the end of the forecast period for each of the years of record, producing a historical distribution for P_{Spring}. We similarly calculate a P_{Winter} distribution for the pre-forecast period from October 1 through April 1. A different cutoff date separating P_{Spring} and P_{Winter} could be chosen to investigate uncertainty in forecasts issued at other times of year. One sample each from the P_{Winter} and P_{Spring} distributions together define a unique seasonal precipitation scenario. The assumption that post-forecast precipitation is statistically independent from pre-forecast precipitation is supported by the low correlation ($r = 0.17$) between the P_{Spring} and P_{Winter} distributions in the historical record.

Our statistical forecast models only predict runoff and do not have an explicit representation of snow processes, so even with a defined value of P_{Winter}, the SWE volume at an arbitrary point in time is undetermined. We leverage this innate model flexibility to investigate the effects of warm snow drought on forecast behavior. We define the SWE fraction,

$$\text{SWE}_{\text{frac}} \equiv \frac{\text{SWE}}{P_{\text{Winter}}}, \tag{9}$$

which relates the amount of SWE measured on the day of the forecast to the amount of pre-forecast precipitation measured at the meteorological station, similar to the SWE/P metric of Heldmyer et al. (2023). By iterating model runs over a range of SWE fractions from 10% (minimal SWE) to 90% (almost all P_{Winter} remaining as SWE), we characterize model behaviors across a range of low- to high-SWE conditions within one precipitation scenario, representing different warm snow drought severities. We note that the historically observed range of SWE_{frac} varies from ~30% to ~100% for ASO flights in the March–April time period. In the study watersheds, snow cover can extend as low as the basin outlets where runoff

is measured, so we compute SWE_{frac} as a catchment average; in other watersheds, SWE_{frac} might be more robustly computed above an elevation threshold defining the typical snow line.

To analyze the contribution of the post-forecast precipitation distribution to total forecast uncertainty, we evaluate the ASO-based runoff models in both forecast and backcast modes. Forecast mode requires propagation of uncertainty in post-forecast precipitation. In contrast, backcast mode uses only a single value of P_{Spring} which, together with a single value of P_{Winter} and a given value of the SWE fraction, constitutes a unique sample of the model space:

$$\text{Backcast } Q \sim Q\left(SWE_{frac} \times P_{Winter}, P_{Spring}\right), \tag{10.1}$$

$$\text{Forecast } Q \sim Q\left(SWE_{frac} \times P_{Winter}, \begin{matrix} P_{Spring_1} \\ \vdots \\ P_{Spring_N} \end{matrix}\right). \tag{10.2}$$

We run forecast and backcast modes for synthetic "years" consisting of precipitation values sampled from P_{Spring} and P_{Winter}. Two sets of parameters and error distributions (Equation 8) are applied in each watershed, one set calibrated to ASO data and one set calibrated using the snow pillow index. Given a particular model and a single value of X, generalized backcast distributions are created by iterating all 2000 model parameter sets over all 37 spring precipitation samples for each of the 37 winter precipitation samples (Equation 10.1). This results in 1369 posterior predictive distributions of 2000 samples each, which constrains the backcast uncertainty under a particular combination of winter and spring precipitation conditions. Next, we create forecast distributions by iterating the model parameter sets over the same 1369 winter–spring precipitation combinations representing unique "years." The forecast distributions are additionally iterated over all 37 randomized spring precipitation samples (Equation 10.2), which are all propagated through the model in place of the single P_{Spring} value to simulate real-time uncertainty in post-forecast precipitation. We thereby obtain a forecast distribution of 74,000 samples (37 precipitation values times 2000 parameter sets) for each of the 1369 "years" (101,306,000 total forecast samples), which constrains uncertainty in the real-time forecast. For both backcast and forecast scenarios, the prior year's runoff value, used in Model 4, is randomly sampled from the observed yearly runoff values over the calibration period to capture a range of wet and dry antecedent conditions.

2.6 | Forecast skill

It is impossible to directly determine the accuracy of modeled runoff in our full suite of sampled scenarios since observed runoff values are not available across all possible combinations of meteorological and snowpack conditions. However, our Bayesian forecast framework provides an opportunity to understand the uncertainty of a given scenario by considering the spread of the posterior predictive distribution. We estimate expected forecast accuracy, or skill, by calculating the mean relative deviation around the median:

$$\text{Skill} \equiv \max\left\{0, \sum_{i=1}^{n}\left(1 - \left[\frac{|Q_i - Q_{Median}|}{Q_{Median}}\right]\right)\frac{1}{n}\right\}. \tag{11}$$

Here, Q_{Median} is the median of either a backcast or forecast distribution in a particular "year" using a particular model and snow data source, which represents the central tendency of the model, or the 50% exceedance level. The spread of the distribution around its median characterizes the forecast or backcast uncertainty, with each sample, Q_i, representing a single possible value that could be the "true" future Q, with the index i ranging from 1 to n, the number of samples in a forecast distribution ($n = 74{,}000$) or backcast distribution ($n = 2000$). Therefore, the term inside the square brackets represents the absolute relative error between the constant modeled value, Q_{Median}, and a particular possible true value, Q_i. Subtracting the absolute relative error from unity yields an estimate of accuracy, presented here as a percentage. For example, 80% forecast accuracy corresponds to an uncertainty of ±20% on average. Equation 11 gives the mean expected accuracy across all samples in a particular forecast or backcast distribution, which we define as the model skill. Although uncertainty can theoretically increase without bound, we constrain skill to the practical range [0%,100%], where a skill of 100% results from perfect predictions and a skill of 0% indicates a model with ±100% mean error.

Our methods ultimately yield an "uncertainty function" that quantifies the uncertainty in our statistical runoff forecasts under arbitrary precipitation and SWE conditions (Equation 11). The contribution of each uncertainty source, U, can be separated by comparing the skill of the three model modes, as diagrammed in Figure 2 and summarized here:

$$U_{Model} \equiv 1 - \text{Skill}_{Backcast_{ASO}}, \tag{12.1}$$

$$U_{Precipitation} \equiv \text{Skill}_{Backcast_{ASO}} - \text{Skill}_{Forecast_{ASO}}, \tag{12.2}$$

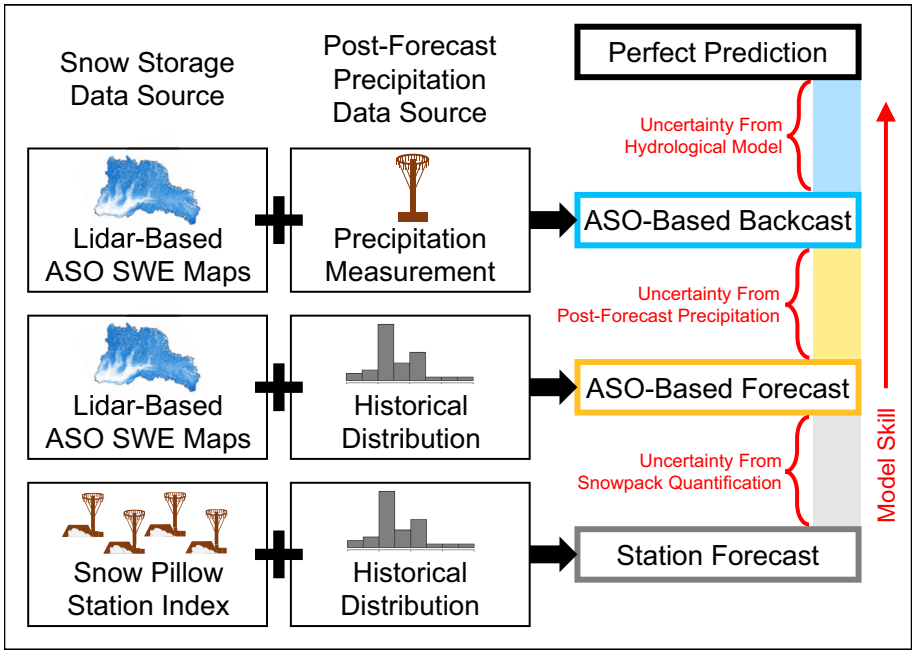

FIGURE 2 Flowchart of workflow and conceptual framework. We estimate forecast uncertainty contributed by errors in the hydrological model, variability in post-forecast precipitation, and uncertainty in station-based snowpack quantification methods by comparing three different simulation modes that are distinguished by their unique combinations of data sources. ASO, Airborne Snow Observatories; SWE, snow water equivalent.

$$U_{\text{Snowpack}} \equiv \text{Skill}_{\text{Forecast}_{\text{ASO}}} - \text{Skill}_{\text{Forecast}_{\text{Station}}}. \tag{12.3}$$

A theoretical "perfect model" would have 100% skill, so subtracting the skill of an actual model provides an estimate of that model's uncertainty (Equation 12.1). The difference between the ASO-based forecast and backcast simulations is attributable to uncertainty from post-forecast precipitation variability (Equation 12.2). The difference between ASO-based forecasts and station-based forecasts captures the uncertainty contributed by snowpack variability that is not captured by the snow pillow index (Equation 12.3).

We analyze the sources of water supply forecast uncertainty under warm snow drought conditions by calculating the average contribution of each uncertainty source for all discrete combinations of SWE_{frac}, P_{Spring}, and P_{Winter} in both watersheds for all four models. Additionally, we repeat these calculations using only the driest 25% of P_{Winter} and driest 25% of P_{Spring} samples to simulate forecast uncertainty during dry snow droughts. In the dry snow drought scenario, the size of forecast and backcast distributions is kept consistent with the baseline simulation, which uses the full 37-year record, by drawing 37 samples of P_{Winter} and P_{Spring} with replacement from the lower quartile of the historical record (9 years). In forecast mode, the P_{Spring} distribution always includes the full historical range of variability, even under dry snow drought conditions, because information about post-forecast precipitation is assumed to be unavailable in real-time. We evaluate the effect of dry snow drought conditions (lower-quartile P_{Winter} and backcast P_{Spring}) by computing the percent change in U_{Snowpack}, $U_{\text{Precipitation}}$, and U_{Model} between the baseline and drought simulations.

3 | RESULTS

3.1 | Runoff models

Bayesian sampling results in 2000 calibrated parameter sets for each model, in each basin, for each of the two snow data sources. Since we assume that the distributed SWE volumes are a more reliable estimate of the snowpack conditions than the station-based measurements, we use the ASO-based backcast model as a benchmark to examine the characteristics of each runoff model. Validation of the models on the runoff intervals from each ASO flight through July 31 produces an average Nash-Sutcliffe model efficiency (NSE) of 0.93 across all four models in both watersheds. We note that this high NSE value is computed for the nested intervals from each ASO flight through July 31, which is an unconventional runoff time series. Regardless, the models achieve a satisfactory correspondence between measured and observed runoff values

(Figure 3). Since we use ASO flights from a range of dates to calibrate the runoff models, points in Figure 3 represent variable time periods, so low-runoff values represent a combination of snow drought conditions and late-season periods in wetter years.

3.2 | Water balance characteristics

The calibrated loss functions indicate that the water balance residual, $P-Q$, is nearly constant over the observed range of precipitation inputs, consistent with observations of nearly constant yearly ET in the Sierra Nevada by Goulden et al. (2012). The exception to this behavior is Model 2, which forces a constant runoff efficiency and a linearly increasing $P-Q$ residual by design. In the other three models, runoff efficiency is nonlinearly dependent on inputs to the water balance ($SWE+P_{Spring}$), with the basin-wide seasonal runoff efficiency, Q/P, varying over a maximum range of 56% during droughts to 83% under wet conditions (Figure 4).

We find that the constant-efficiency loss function (Model 2) introduces a bias in runoff estimates (slope of best-fit line between measured and observed = 1.20, average NSE over both watersheds = 0.89) by not accounting for the nonlinear decrease in runoff efficiency during low precipitation years. In contrast, the constant loss function (Model 1) produces a lower bias and better fit to the observed runoff data (slope = 0.91, NSE = 0.95), suggesting that the difference between precipitation and runoff in this system is better modeled as an offset than a ratio. This result is consistent with other work in our study area: Hedrick et al. (2020) found that evapotranspiration losses were relatively constant in the Tuolumne watershed across a sample of dry, average, and wet years, and Henn et al. (2018) concluded that the seasonal water balance residual in the Tuolumne watershed represented mean evapotranspiration. The relative constancy of $P-Q$ over a range of wet and dry years could represent the partial decoupling of plant water use from precipitation variability as observed elsewhere in California by Hahm et al. (2019), although drought tree mortality in the Sierra Nevada suggests that this decoupling, if present in the study watersheds, may break down in the driest years. Alternatively, the efficacy of the constant-loss model could stem from structural factors in the analysis by compensating a latent bias in the meteorological or snowpack data or limiting the impact of extreme conditions on uncertainty propagation.

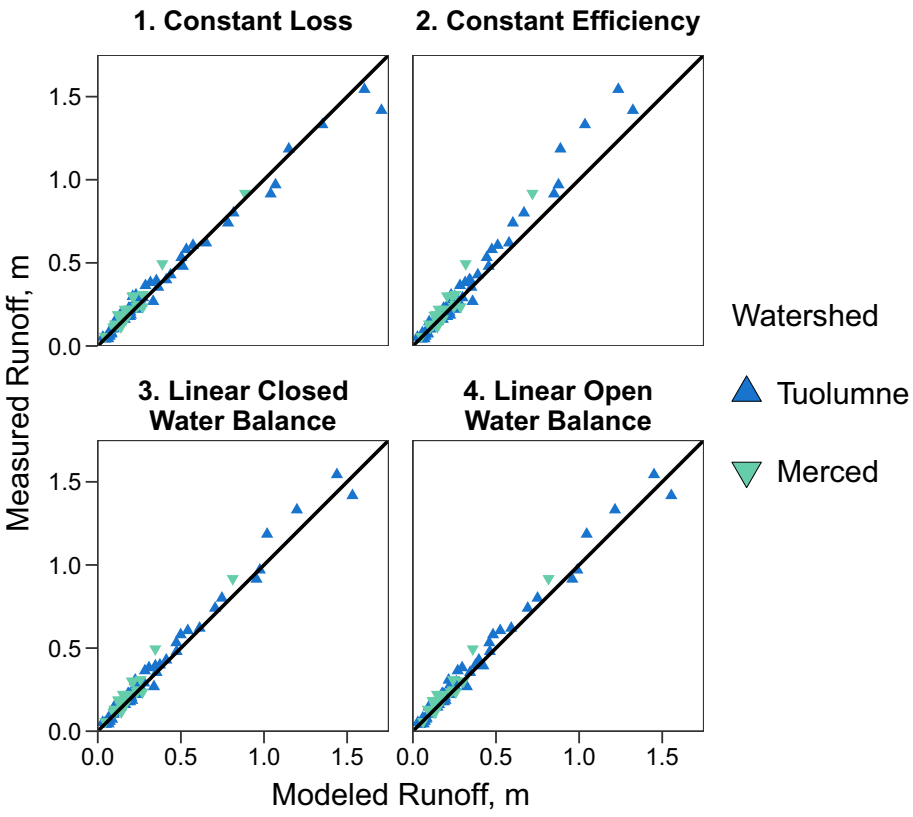

FIGURE 3 Scatter plots of modeled versus measured runoff values. Each panel represents a different empirical model parameterization (Equations 3.1–3.4). The models are evaluated using basin-average SWE and post-flight gauged precipitation measurements for all available January–June ASO acquisitions to predict runoff on the interval from each flight date through July 31.

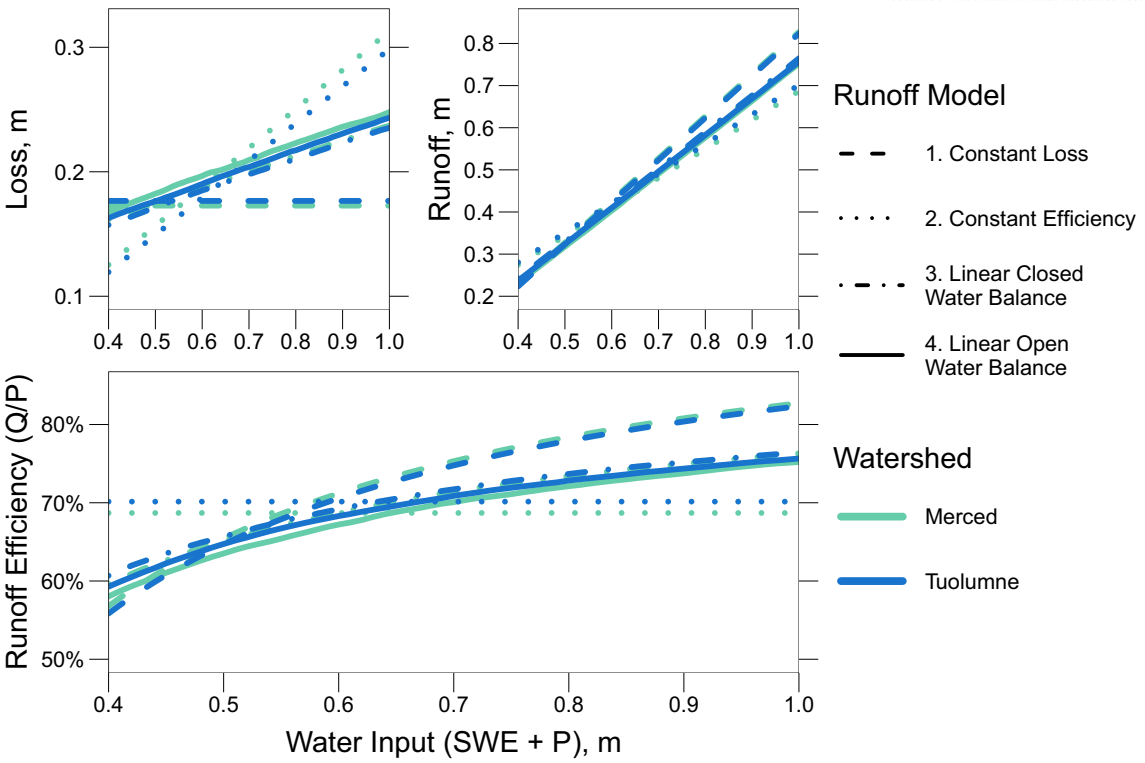

FIGURE 4 Median characteristics of the calibrated Bayesian runoff models. The water balance residual, snow and precipitation (SWE + *P*) minus runoff (*Q*), is modeled using four different loss functions (Equations 3.1–3.4). Runoff efficiency, the ratio of *Q* to (SWE + *P*), is calculated using runoff outputs from the calibrated models. Models with sufficient flexibility (1, 2, and 4) indicate nonlinear decreases in runoff efficiency during drought.

3.3 | Uncertainty sources

Running the hydrologic models in three different forecast and backcast modes allows us to partition the major sources of uncertainty during warm snow droughts. The backcast mode, using a station-measured precipitation value and basin-wide average SWE, is always more skillful than the forecast modes, because the latter propagate uncertainty in post-forecast precipitation. Similarly, the forecast using ASO SWE is always more skillful than the forecast using a snow pillow index since the full-catchment SWE maps capture year-to-year variability in the spatial snowpack distribution. Figure 5 illustrates the average skill of the four models across both watersheds, with uncertainty sources partitioned analogously to the conceptual model in Figure 2. Under historical conditions with 50% to 90% of pre-forecast precipitation remaining as SWE on April 1, on average we find that 46% of total water supply uncertainty comes from the hydrological model, 13% from post-forecast precipitation variability, and 41% from uncertainty in station-based snowpack quantification methods. The relative proportion of total forecast uncertainty attributable to each source is relatively constant, varying by 1%–7% across the full range of tested SWE_{frac} values (Figure S1). The total magnitude of the station-based forecast uncertainty (one minus model skill) varies between 33% when SWE_{frac} = 0.9% and 77% during extreme warm snow drought when SWE_{frac} = 0.1.

3.4 | Snow drought effects

We observe a nonlinear decline in forecast skill during warm snow droughts, when the fraction of pre-forecast precipitation remaining as SWE is less than 50% (Figure 5). The forecast using snow pillow data experiences a 59% decrease in mean skill between the historic (SWE_{frac} = 50%) and anticipated future (SWE_{frac} = 10%) snow drought regimes, while the forecast using ASO SWE undergoes a 31% decrease in mean skill for the same transition. In our snow-driven modeling framework, the fraction of pre-forecast precipitation remaining as SWE is the primary driver of changes in overall fractional uncertainty. However, we find that the forecasts using ASO SWE are about half as sensitive to warm snow drought conditions as forecasts using an index of snow pillows, since the station-based forecast skill decreases 1.9 times faster for the 50% to 10% SWE_{frac} transition. The heightened sensitivity of the station-based forecast aligns with expectations informed by Livneh and Badger (2020), since remotely sensed snowpack data are less affected by a receding snowline than station-based data.

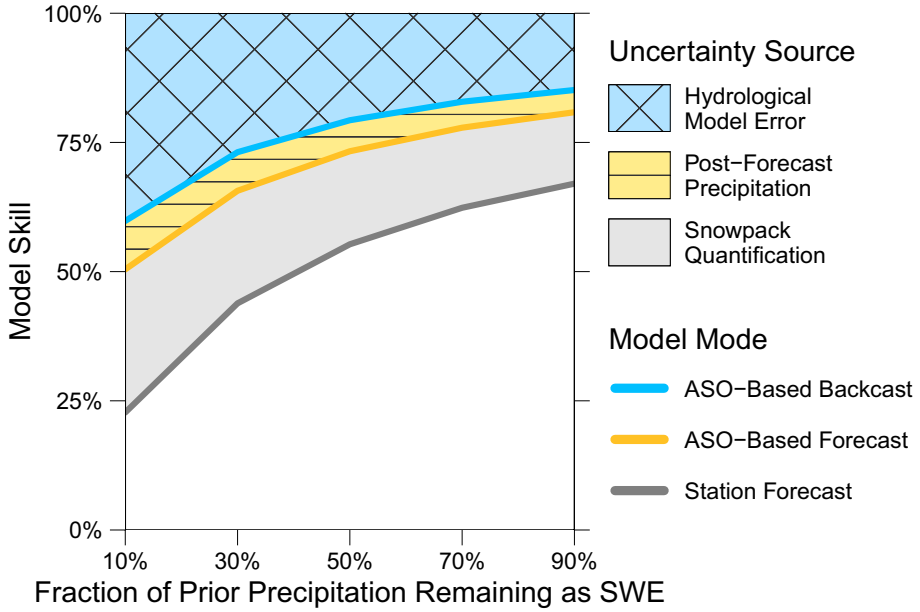

FIGURE 5 Mean forecast skill across both watersheds and all four hydrological models for the April 1 forecast date. We compare the skill of three simulation modes to isolate uncertainty arising from errors in the hydrological model, variability in post-forecast precipitation, and heterogeneity in the snowpack distribution that is not captured by station-based snowpack quantification methods.

Increased model uncertainty during warm snow drought conditions is more substantive than differences between models (Figure 6). In the most extreme case, when SWE_{frac} decreases from 90% to 10%, the uncertainty contributed by errors in the statistical runoff model increases by 165% in the Merced and 178% in the Tuolumne. This extraordinary increase in model uncertainty during warm snow drought suggests that the models are not uniformly applicable across all values of SWE_{frac}, and an alternative parameterization may be preferable for simulating runoff under low-snow conditions. However, since operational statistical models are commonly assumed valid across varied climatological conditions, it is important to quantify the breakdown of statistical snowmelt runoff assumptions in a warmer climate.

Differences between the skill of the four models in backcast mode are on the order of 2%–16%, with the largest differences between models in the Merced during extreme warm snow drought. Overall, all four models produce comparable uncertainty ranges, suggesting

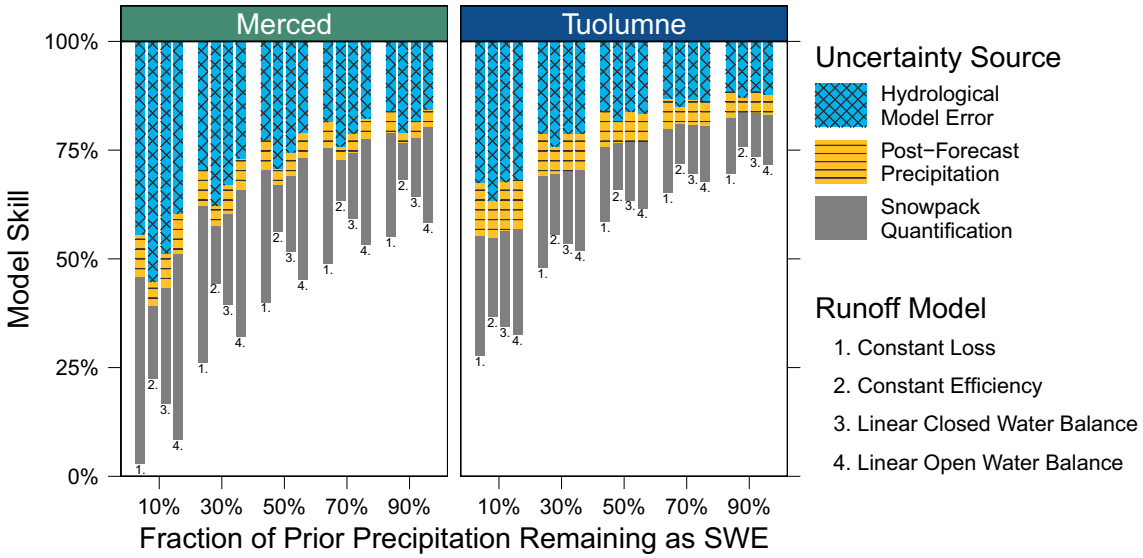

FIGURE 6 Effect of warm snow drought on uncertainty sources for the April 1 forecast date. As the climate becomes less snow-dominated (smaller fraction of precipitation remaining as SWE), uncertainty in the hydrological model and station-based snowpack quantification increase nonlinearly, rapidly outweighing the differences between model parameterization and watersheds. Note that increasing uncertainty (longer stacked bars) causes lower forecast skill, defined as one minus the total uncertainty (c.f. Figure 5).

that the differences in parameterization between lumped statistical models are not a dominant driver of forecast uncertainty in this system. Interestingly, the station-based forecast using snow pillow data is most skillful when paired with the constant efficiency runoff model, even though the constant efficiency model is biased (Figure 3) and consistently achieves the lowest backcast skill (Figure 6). In contrast, the ASO-based forecasts are typically most skillful when paired with the constant loss or linear runoff models (Figure 6), particularly in the Merced River basin, suggesting that the reversal in the constant efficiency model's relative performance could derive in part from an opposing bias in the snow pillow index.

The runoff models perform similarly across both study watersheds, though forecast uncertainty is marginally higher (6–14%) in the Merced catchment (Figure 6). Possible reasons for the slightly lower performance of the Merced forecast models include the smaller catchment size, leading to a greater dependence on localized precipitation events, and greater geographic separation from the TUM precipitation gauge, potentially leading to less-representative precipitation estimates. Additionally, there are fewer ASO flights in the Merced compared to the Tuolumne (and none during 2017, the wettest year of the study period), so the Merced models could be more loosely constrained by the historic calibration.

Water supply uncertainty increases during dry snow droughts due to the increased importance of post-forecast precipitation relative to snowpack storage when P_{Winter} is small (Figure 7). The reduced skill of all three model modes under lower-quartile precipitation conditions demonstrates the sources of increased water supply forecast uncertainty during dry snow droughts. Under typical historic climate conditions (SWE_{frac} 50–90%) during a lower-quartile precipitation drought, fractional uncertainty from the snowpack distribution increases by 6%–29%, hydrological model fractional uncertainty increases by 10%–19%, and fractional uncertainty from post-forecast precipitation increases by 16%–46%, with the effect magnitude depending on the watershed and SWE_{frac}. The Tuolumne watershed is more sensitive to dry snow droughts than the Merced, likely due to a combination of higher baseline forecast skill in the Tuolumne (Figure 6) and the Merced watershed's smaller area and greater separation from the precipitation station (Figure 1).

When both precipitation and the SWE_{frac} are low, that is during coincident warm and dry snow droughts, overall uncertainty increases, but the relative partitioning of uncertainty sources remains relatively consistent. Due to the method we use to propagate random error, namely a scale factor on the entire seasonal water balance, the model can counterintuitively predict a relative decrease in precipitation and snow uncertainty sources when an extremely low precipitation amount coincides with a warm snow drought. This result follows from the lack of correspondence between runoff and the snowpack mass balance in the limit where SWE approaches zero. Since it is illogical to use a snow-based runoff model in a nearly snow-free environment, we limit our exploration of dry snow drought effects on individual uncertainty sources to the case where >50% of pre-forecast precipitation remains as SWE on the forecast date (Figure 7). However, the relative partitioning of total forecast uncertainty remains interpretable even during coincident warm and dry snow droughts (SWE_{frac} = 10% with lower-quartile P_{Winter} and P_{Spring}). In this extreme snow drought case, the total water supply forecast uncertainty can still be reduced by 33% in the Tuolumne and 38% in

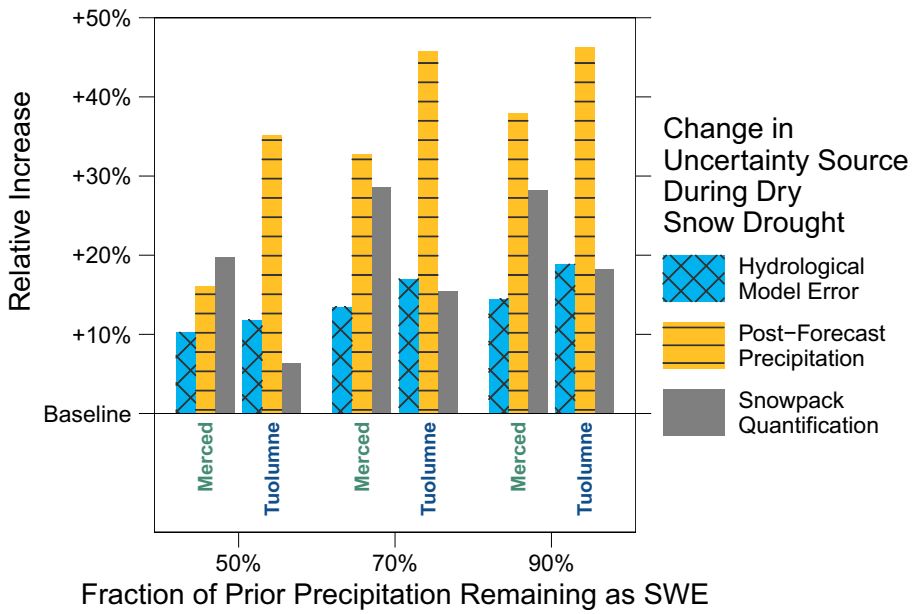

FIGURE 7 Increase in uncertainty sources during dry snow droughts on the April 1 forecast date, caused by lower-quartile winter and spring precipitation, relative to mean fractional uncertainty across all precipitation conditions. Dry snow droughts have the largest relative effect on the uncertainty from post-forecast precipitation variability, but the severity of this effect depends on the watershed and becomes relatively less important during coincident warm snow droughts when overall uncertainty is already high.

the Merced by replacing the snow pillow index with full-catchment ASO SWE, comparable to the 39% average improvement when considering the full range of drought and non-drought scenarios across both watersheds.

4 | DISCUSSION

During dry snow droughts, post-forecast precipitation becomes an increasingly important component of runoff uncertainty. Since dry snow droughts are characterized by a low volume of pre-forecast precipitation (meteorological drought), the pre-forecast catchment water input is relatively small compared to the potential variability in post-forecast precipitation, so the fractional uncertainty in post-forecast runoff increases. During a dry snow drought with lower-quartile winter and spring precipitation, the water supply uncertainty contributed by variability in post-forecast precipitation can increase by as much as 46%, depending on the watershed and any coincident warm snow drought conditions (Figure 7). This finding is intuitive but highlights a potential weakness of forecasting paradigms that rely on black-box statistics to derive runoff exceedance levels without considering the difference in precipitation uncertainty during wet and dry years. For example, a Principal Components Regression (PCR) model using April 1 SWE and October–March precipitation (Garen, 1992), widely used for forecasting by the Natural Resources Conservation Service (NRCS), could underestimate the true uncertainty range during dry snow droughts if the uncertainty is based on fractional errors across all years. In other words, a small fractional uncertainty on average across the historical record (including wet years) does not guarantee a small fractional error during droughts. For example, Goble and Schumacher (2023) found that incorporating spring and summer precipitation data improved the accuracy of water supply forecasts in Colorado during the snow droughts of 2020 and 2021, when forecasts using snowpack data alone overpredicted runoff volumes. Analogously, our results demonstrate that the issue of unrealistic exceedance levels, driven by uncertainty in the uncertainty, can be addressed by propagating precipitation variability separately from other uncertainty sources.

Forecast paradigms that do not explicitly propagate post-forecast precipitation variability, such as the NRCS's 2-component PCR model, can gain increased interpretability by incorporating a separate precipitation term and calibrating the model with historic spring precipitation data to directly constrain the uncertainty introduced by post-forecast precipitation variability. By calibrating runoff models with explicit post-forecast precipitation values instead of absorbing precipitation variability into a black-box error term, we can better understand the effect of dry snow droughts on water supply uncertainty. Our methods utilize records from a single precipitation station, but other approaches could leverage an index of multiple stations or combine station measurements with a spatially distributed meteorological product.

While explicitly propagating post-forecast precipitation is necessary to adequately constrain uncertainty ranges and inform management during dry snow droughts, the underlying component of uncertainty is analogous to Schaake and Peck (1985) description of "Unpredictable Climate Error," as it derives from weather phenomena on seasonal time scales, which remains difficult to forecast. The future impact of precipitation variability on runoff forecast uncertainty is unclear because advances in subseasonal-to-seasonal weather forecasts may reduce uncertainty (e.g., DeFlorio et al., 2019; Waliser et al., 2003) while the introduction of stochastic weather generators that account for climate non-stationarity (e.g., Steinschneider et al., 2019) may suggest increased precipitation variability. As a result, our methods highlight a latent uncertainty source, but we do not attempt to further extrapolate the effects of dry snow drought into the future climate. In future decades, uncertainty from post-forecast precipitation variability could be affected by the countervailing effects of improved weather forecasts and climate nonstationarity.

Inaccurate quantification of snowpack storage is a key factor causing runoff uncertainty during both dry and warm snow droughts. Although the snowpack represents a diminished fraction of the seasonal water balance under snow drought conditions, accurate measurement of near-zero SWE is itself important for constraining the lower bound of water supply guidance. By estimating the SWE volume at a small value, snowpack measurements still provide an important lower bound on runoff during low-snow periods such as the 2013–2015 and 2020–2022 Sierra Nevada snow droughts (e.g., Hetch Hetchy Reservoir operations team, personal communication, 2021–2022). However, quantifying the snowpack volume with conventional in situ station-based methods becomes problematic when the snow line exceeds station elevations. When snow observation stations are concentrated at mid-elevations, as in the study region, significant snow may remain even after most stations melt out, leading to uncertainty in the actual SWE volume that corresponds to a near-zero measurement. While this phenomenon has been raised by Livneh and Badger (2020) as an important concern for water managers, ours is one of the first studies to quantify its importance during historical snow drought conditions in a data-driven framework.

We partially ameliorate this problem by incorporating remotely sensed, spatially distributed snowpack observations and physical model outputs from ASO. Switching from a snow pillow index to ASO SWE maps reduced the expected water supply forecast uncertainty by an average of 39% across all models, meteorological conditions, and both watersheds (average across SWE_{frac} in Figure 6). Consequently, the uncertainty of the station-based forecast increases about 1.9 times as quickly as the uncertainty of the ASO-based forecast during warm snow droughts. This finding supports our hypothesis that remotely sensed snowpack observations will become increasingly important in a warmer future climate, when a smaller fraction of winter precipitation persists as snow. We could expect additional room for improvement in watersheds with sparser snow monitoring station networks or greater variability in the spatial distribution.

In our comparison of snow quantification methods, we consider only the snow pillow index and the ASO SWE maps because these datasets are both used operationally by local water managers and represent endmembers of spatial resolution and sophistication. However, there is a growing array of snow data products that are available for water supply forecasting. Physically based snow models can represent spatial heterogeneities and estimate storage even when the snowline recedes above station elevations, particularly when combined with remote sensing data assimilation (e.g., Hedrick et al., 2018). Similarly, satellite-derived snow covered area maps are widely available and can discriminate varying levels of snow drought, which has proven useful for runoff forecasting (e.g., Martinec et al., 2008). Finally, statistical relationships between station measurements and remote sensing or model outputs can be used to construct synthetic datasets leveraging the unique combination of available snow data in a particular watershed or a particular year through various combinations of bias correction, weighting, and machine learning techniques (e.g., Pflug & Lundquist, 2020).

For this study, we focus on different snow measurement resolutions using two methods that are based directly on physical measurements. The snow pillow and ASO datasets represent endmember snow measurement methods, both in terms of technological complexity and spatial sampling resolution, and both methods are operationally deployed in the study region. We would expect that an intermediate-fidelity product, such as satellite-derived snow covered area or open-loop modeled SWE, would likely produce an intermediate forecast uncertainty between the snow pillow and ASO endmembers. Future research could use a similar uncertainty quantification framework for an intercomparison of the diverse snow data products that are becoming available in the Sierra Nevada and elsewhere, both from measurement campaigns and physical or statistical models (e.g., Deschamps-Berger et al., 2023; Guan et al., 2013; Tarricone et al., 2023; Yang et al., 2022). Similarly, future work could apply our comparative uncertainty framework to process-based water supply forecast models to evaluate potential improvements in model uncertainty and the sensitivity of models to different datasets or disturbances.

During warm snow droughts, when relatively low snow storage results from warmer climate conditions, regardless of the total precipitation magnitude, the empirical hydrological models become the dominant source of water supply forecast uncertainty (Figure 6). The models presented here strongly depend on the snowpack mass balance, with the variable effects of evapotranspiration, storage change, and other fluxes lumped together as regression coefficients. Moreover, the statistical error distributions associated with each Bayesian runoff model subsume uncertainty in the data used to calibrate the models, so when a particular dataset becomes more important, the uncertainties associated with that dataset can have a larger impact on the backcast model skill. For instance, some or all of the higher fractional model uncertainty observed during dry snow drought (Figure 7) could result from greater uncertainty in relatively small precipitation measurements. Similarly, the density field used to generate SWE maps could increase the inferred statistical runoff uncertainty if the snow model uncertainty increases during anomalous climatological conditions. With lumped empirical runoff models, it is not possible to further attribute uncertainty sources to stochastic errors in the forcing data or shortcomings in the model, since the model itself is effectively derived from the data. Future analysis using process-based hydrological models could potentially separate structural model errors from errors in the forcing data by applying a similar comparative framework to an ensemble of model parameterizations.

As winter temperatures continue to increase, causing precipitation regimes to shift from snow to rain and snowmelt to occur earlier in the runoff season, the assumption that post-forecast runoff depends primarily on April 1 SWE could become invalid. We anticipate a nonlinear increase in the uncertainty of snow-based statistical runoff models as SWE_{frac} decreases on the April 1 forecast date (Figure 6). Accordingly, there is a shift toward water supply forecasting with physical hydrological models that can account for non-snow processes. In the Tuolumne River basin, managers are testing the Precipitation Runoff Modeling System (PRMS, Markstrom et al., 2015), and ASO is testing deployment of the WRF-Hydro model (Gochis et al., 2020) using assimilation of ASO SWE data. Transferring the statistical benefits of the Bayesian uncertainty framework presented here to more complex physical models will likely necessitate innovative methods to propagate hydrological model uncertainties into probabilistic forecasts.

5 | CONCLUSION

We have explored three key effects of snow drought on water supply forecast uncertainty: the increased importance of post-forecast precipitation variability during dry snow droughts, the breakdown of empirical model assumptions during extreme warm snow droughts, and the value of distributed SWE data. Furthermore, we have considered several avenues to ameliorate anticipated increases in forecast uncertainty in a warmer future climate, namely the incorporation of remotely sensed SWE data and the deployment of hydrological models that account for liquid water storage cycles. When considering the investment required to incorporate new technologies into existing water management systems, the comparative forecasting framework developed here could be applied to identify the most promising areas for improvement in a cost–benefit analysis. Separating sources of uncertainty can help quantify the value of deploying updated snow monitoring technologies such as the ASO SWE maps, which are now part of the State of California's Department of Water Resources forecast infrastructure. Warm and dry snow droughts are expected to increasingly challenge water supply forecasting, so a rigorous treatment of uncertainty propagation during snow drought in more complex hydrological models and operational forecasting systems is crucial to developing resilient strategies for climate change adaptation.

AUTHOR CONTRIBUTIONS

Elijah N. Boardman: Conceptualization; data curation; formal analysis; investigation; methodology; visualization; writing – original draft. **Carl E. Renshaw:** Conceptualization; methodology; project administration; supervision; writing – review and editing. **Robert K. Shriver:** Methodology; supervision; validation; writing – review and editing. **Reggie Walters:** Conceptualization; methodology; writing – review and editing. **Bruce McGurk:** Conceptualization; data curation; writing – review and editing. **Thomas H. Painter:** Conceptualization; data curation; funding acquisition; writing – review and editing. **Jeffrey S. Deems:** Conceptualization; data curation; funding acquisition; writing – review and editing. **Kat J. Bormann:** Conceptualization; data curation; funding acquisition; writing – review and editing. **Gabriel M. Lewis:** Methodology; writing – review and editing. **Evan N. Dethier:** Methodology; writing – review and editing. **Adrian A. Harpold:** Conceptualization; funding acquisition; methodology; project administration; supervision; writing – review and editing.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. 1937966. We gratefully acknowledge NSF EAR 1723990, which supported submission of this manuscript. We thank the Hetch Hetchy Water and Power hydrology team for helping to contextualize this project in the Tuolumne River basin. We acknowledge the NASA Terrestrial Hydrology program, NASA Applied Sciences program, JPL/Caltech Western Water Applications Office, and the California Department of Water Resources for funding the ASO program while at JPL and since it moved to a public-benefit corporation across the last decade.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest in this study or manuscript.

DATA AVAILABILITY STATEMENT

The ASO SWE data that support the findings of this study are available from the National Snow and Ice Data Center at https://doi.org/10.5067/M4TUH28NHL4Z or from the Airborne Snow Observatories, Inc. website at https://data.airbornesnowobservatories.com. Precipitation and SWE records for the meteorological and snow pillow stations used in this study are available from the California Data Exchange Center at https://cdec.water.ca.gov. Streamflow records for the Merced River at Happy Isles (U.S. Geological survey station number 11264500) are available from the National Water Information System at https://waterdata.usgs.gov/nwis. Reservoir inflow records for the Hetch Hetchy Reservoir are available from the corresponding author upon reasonable request. Computer code and associated outputs that support the findings of this study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT

Authors Painter, Deems, and Bormann acknowledge an interest in Airborne Snow Observatories, Inc., a public benefit corporation that produced SWE maps used in this study. Ethics review is not applicable to this study because it does not involve live subjects.

ORCID

Elijah N. Boardman https://orcid.org/0009-0009-1979-6954
Carl E. Renshaw https://orcid.org/0000-0002-5022-4307
Kat J. Bormann https://orcid.org/0000-0003-3697-6550

REFERENCES

Abatzoglou, J.T. 2013. "Development of Gridded Surface Meteorological Data for Ecological Applications and Modelling." *International Journal of Climatology* 33(1): 121–31. https://doi.org/10.1002/joc.3413.

Bales, R.C., N.P. Molotch, T.H. Painter, M.D. Dettinger, R. Rice, and J. Dozier. 2006. "Mountain Hydrology of the Western United States." *Water Resources Research* 42: W08432. https://doi.org/10.1029/2005WR004387.

Barnett, T.P., J.C. Adam, and D.P. Lettenmaier. 2005. "Potential Impacts of a Warming Climate on Water Availability in Snow-Dominated Regions." *Nature* 438: 303–9. https://doi.org/10.1038/nature04141.

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. 2016. "Facets of Uncertainty: Epistemic Uncertainty, Non-stationarity, Likelihood, Hypothesis Testing, and Communication." *Hydrological Sciences Journal* 61(9): 1652–65. https://doi.org/10.1080/02626667.2015.1031761.

California Department of Water Resources. 2022. "Database of Daily Snow Water Equivalent Records for the DAN, GIN, PDS, SLI, STR, TNY, TUM, WHW Snow Pillow Stations: California Data Exchange Center." August 2, 2022. https://cdec.water.ca.gov.

California Department of Water Resources/DFM-Hydro-SMN. 2022. "Database of Daily Precipitation Records for the TUM Meteorological Station: California Data Exchange Center." August 15, 2022. https://cdec.water.ca.gov/dynamicapp/staMeta?station_id=TUM.

Day, G.N. 1985. "Extended Streamflow Forecasting Using NWSRFS." *Journal of Water Resources Planning and Management* 111(2): 157–70. https://doi.org/10.1061/(ASCE)0733-9496(1985)111:2(157).

DeFlorio, M.J., D.E. Waliser, F.M. Ralph, B. Guan, A. Goodman, P.B. Gibson, S. Asharaf, et al. 2019. "Experimental Subseasonal-to-Seasonal (S2S) Forecasting of Atmospheric Rivers over the Western United States." *Journal of Geophysical Research: Atmospheres* 124(21): 11242–65. https://doi.org/10.1029/2019JD031200.

Deschamps-Berger, C., S. Gascoin, D. Shean, H. Besso, A. Guiot, and J.I. López-Moreno. 2023. "Evaluation of Snow Depth Retrievals from ICESat-2 Using Airborne Laser-Scanning Data." *The Cryosphere* 17(7): 2779–92. https://doi.org/10.5194/tc-17-2779-2023.

Dettinger, M.D., and M.L. Anderson. 2015. "Storage in California's Reservoirs and Snowpack in this Time of Drought." *San Francisco Estuary and Watershed Science* 13(2): 1-5. https://doi.org/10.15447/sfews.2015v13iss2art1.

Dierauer, J.R., D.M. Allen, and P.H. Whitfield. 2019. "Snow Drought Risk and Susceptibility in the Western United States and Southwestern Canada." *Water Resources Research* 55(4): 3076–91. https://doi.org/10.1029/2018WR023229.

Dozier, J. 2011. "Mountain Hydrology, Snow Color, and the Fourth Paradigm." *Eos, Transactions, American Geophysical Union* 92(43): 373–4. https://doi.org/10.1029/2011EO430001.

Garen, D.C. 1992. "Improved Techniques in Regression-Based Streamflow Volume Forecasting." *Journal of Water Resources Planning and Management* 118(6): 654–70. https://doi.org/10.1061/(ASCE)0733-9496(1992)118:6(654).

Goble, P.E., and R.S. Schumacher. 2023. "On the Sources of Water Supply Forecast Error in Western Colorado." *Journal of Hydrometeorology* 24: 2321–32. https://doi.org/10.1175/JHM-D-23-0004.1.

Gochis, D.J., M. Barlage, R. Cabell, M. Casali, A. Dugger, K. FitzGerald, M. McAllister, et al. 2020. "The WRF-Hydro® Modeling System Technical Description, Version 5.2.0. NCAR Technical Note." https://ral.ucar.edu/sites/default/files/public/projects/wrf-hydro/technical-description-user-guide/wrf-hydrov5.2technicaldescription.pdf.

Goulden, M.L., R.G. Anderson, R.C. Bales, A.E. Kelly, M. Meadows, and G.C. Winston. 2012. "Evapotranspiration Along an Elevation Gradient in California's Sierra Nevada." *Journal of Geophysical Research: Biogeosciences* 117(G3). https://doi.org/10.1029/2012JG002027.

Guan, B., N.P. Molotch, D.E. Waliser, S.M. Jepsen, T.H. Painter, and J. Dozier. 2013. "Snow Water Equivalent in the Sierra Nevada: Blending Snow Sensor Observations with Snowmelt Model Simulations." *Water Resources Research* 49(8): 5029–46. https://doi.org/10.1002/wrcr.20387.

Hahm, W.J., D.N. Dralle, D.M. Rempe, A.B. Bryk, S.E. Thompson, T.E. Dawson, and W.E. Dietrich. 2019. "Low Subsurface Water Storage Capacity Relative to Annual Rainfall Decouples Mediterranean Plant Productivity and Water Use from Rainfall Variability." *Geophysical Research Letters* 46(12): 6544–53. https://doi.org/10.1029/2019GL083294.

Harpold, A.A., M. Dettinger, and S. Rajagopal. 2017. "Defining Snow Drought and Why it Matters." *Eos* 98. https://doi.org/10.1029/2017EO068775.

Hartmann, H.C., R. Bales, and S. Sorooshian. 2002. "Weather, Climate, and Hydrologic Forecasting for the US Southwest: A Survey." *Climate Research* 21(3): 239–58. https://doi.org/10.3354/cr021239.

Hatchett, B.J., and D.J. McEvoy. 2018. "Exploring the Origins of Snow Drought in the Northern Sierra Nevada, California." *Earth Interactions* 22(2): 1–13. https://doi.org/10.1175/EI-D-17-0027.1.

Hedrick, A.R., D. Marks, S. Havens, M. Robertson, M. Johnson, M. Sandusky, H.P. Marshall, P.R. Kormos, K.J. Bormann, and T.H. Painter. 2018. "Direct Insertion of NASA Airborne Snow Observatory-Derived Snow Depth Time Series into the iSnobal Energy Balance Snow Model." *Water Resources Research* 54(10): 8045–63. https://doi.org/10.1029/2018WR023190.

Hedrick, A.R., D. Marks, H.P. Marshall, J. McNamara, S. Havens, E. Trujillo, M. Sandusky, K.J. Bormann, and T.H. Painter. 2020. "From Drought to Flood: A Water Balance Analysis of the Tuolumne River Basin during Extreme Conditions (2015–2017)." *Hydrological Processes* 34(11): 2560–74. https://doi.org/10.1002/hyp.13749.

Heim, R.R. 2002. "A Review of Twentieth-Century Drought Indices Used in the United States." *Bulletin of the American Meteorological Society* 83(8): 1149–66. https://doi.org/10.1175/1520-0477-83.8.1149.

Heldmyer, A.J., N.R. Bjarke, and B. Livneh. 2023. "A 21st-Century Perspective on Snow Drought in the Upper Colorado River Basin." *Journal of the American Water Resources Association* 59(2): 396–415. https://doi.org/10.1111/1752-1688.13095.

Henn, B., T.H. Painter, K.J. Bormann, B. McGurk, A.L. Flint, L.E. Flint, V. White, and J.D. Lundquist. 2018. "High-Elevation Evapotranspiration Estimates During Drought: Using Streamflow and NASA Airborne Snow Observatory SWE Observations to Close the Upper Tuolumne River Basin Water Balance." *Water Resources Research* 54(2): 746–66. https://doi.org/10.1002/2017WR020473.

Her, Y., S.-H. Yoo, J. Cho, S. Hwang, J. Jeong, and C. Seong. 2019. "Uncertainty in Hydrological Analysis of Climate Change: Multi-Parameter vs. Multi-GCM Ensemble Predictions." *Scientific Reports* 9: 4974. https://doi.org/10.1038/s41598-019-41334-7.

Huning, L.S., and A. AghaKouchak. 2018. "Mountain Snowpack Response to Different Levels of Warming." *Proceedings of the National Academy of Sciences* 115(43): 10932–7. https://doi.org/10.1073/pnas.1805953115.

Huning, L.S., and A. AghaKouchak. 2020. "Global Snow Drought Hot Spots and Characteristics." *Proceedings of the National Academy of Sciences* 117(33): 19753–9. https://doi.org/10.1073/pnas.1915921117.

Immerzeel, W.W., A.F. Lutz, M. Andrade, A. Bahl, H. Biemans, T. Bolch, S. Hyde, et al. 2020. "Importance and Vulnerability of the world's Water Towers." *Nature* 577: 364–9. https://doi.org/10.1038/s41586-019-1822-y.

Knowles, N., M.D. Dettinger, and D.R. Cayan. 2006. "Trends in Snowfall Versus Rainfall in the Western United States." *Journal of Climate* 19(18): 4545–59. https://doi.org/10.1175/JCLI3850.1.

Koster, R.D., S.P.P. Mahanama, B. Livneh, D.P. Lettenmaier, and R.H. Reichle. 2010. "Skill in Streamflow Forecasts Derived from Large-Scale Estimates of Soil Moisture and Snow." *Nature Geoscience* 3(9): 613–6. https://doi.org/10.1038/ngeo944.

Lapides, D.A., W.J. Hahm, D.M. Rempe, J. Whiting, and D.N. Dralle. 2022. "Causes of Missing Snowmelt Following Drought." *Geophysical Research Letters* 49(19): e2022GL100505. https://doi.org/10.1029/2022GL100505.

Li, D., M.L. Wrzesien, M. Durand, J. Adam, and D.P. Lettenmaier. 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): 6163–72. https://doi.org/10.1002/2017GL073551.

Livneh, B., and A.M. Badger. 2020. "Drought Less Predictable under Declining Future Snowpack." *Nature Climate Change* 10: 452–8. https://doi.org/10.1038/s41558-020-0754-8.

Lundquist, J.D., N.E. Wayand, A. Massmann, M.P. Clark, F. Lott, and N.C. Cristea. 2015. "Diagnosis of Insidious Data Disasters." *Water Resources Research* 51(5): 3815–27. https://doi.org/10.1002/2014WR016585.

Mankin, J.S., D. Viviroli, D. Singh, A.Y. Hoekstra, and N.S. Diffenbaugh. 2015. "The Potential for Snow to Supply Human Water Demand in the Present and Future." *Environmental Research Letters* 10: 114016. https://doi.org/10.1088/1748-9326/10/11/114016.

Marks, D., J. Domingo, D. Susong, T. Link, and D. Garen. 1999. "A Spatially Distributed Energy Balance Snowmelt Model for Application in Mountain Basins." *Hydrological Processes* 13(12–13): 1935–59. https://doi.org/10.1002/(SICI)1099-1085(199909)13:12/13<1935::AID-HYP868>3.0.CO;2-C.

Markstrom, S.L., R.S. Regan, L.E. Hay, R.J. Viger, R.M.T. Webb, R.A. Payn, and J.H. LaFontaine. 2015. "PRMS-IV, The Precipitation-Runoff Modeling System, Version 4. *U.S. Geological Survey Techniques and Methods* Book 6, Chapter B7." https://doi.org/10.3133/tm6B7.

Marshall, A.M., J.T. Abatzoglou, T.E. Link, and C.J. Tennant. 2019. "Projected Changes in Interannual Variability of Peak Snowpack Amount and Timing in the Western United States." *Geophysical Research Letters* 46(15): 8882–92. https://doi.org/10.1029/2019GL083770.

Martinec, J., A. Rango, and R. Roberts. 2008. *Snowmelt Runoff Model (SRM) User's Manual.* Agricultural Experiment Station Special Report 100. Las Cruces, New Mexico: New Mexico State University College of Agriculture and Home Economics, https://pubs.nmsu.edu/water/SRMSpecRep100.pdf.

McMillan, H.K., I.K. Westerberg, and T. Krueger. 2018. "Hydrological Data Uncertainty and its Implications." *WIREs Water* 5: e1319. https://doi.org/10.1002/wat2.1319.

Milly, P.C.D., J. Betancourt, M. Falkenmark, R.M. Hirsch, Z.W. Kundzewicz, D.P. Lettenmaier, and R.J. Stouffer. 2008. "Stationarity Is Dead: Whither Water Management?" *Science* 319: 573–4. https://doi.org/10.1126/science.1151915.

Mote, P.W., S. Li, D.P. Lettenmaier, M. Xiao, and R. Engel. 2018. "Dramatic Declines in Snowpack in the Western US." *npj Climate and Atmospheric Science* 1(2): 1–6. https://doi.org/10.1038/s41612-018-0012-1.

Mountain Research Initiative EDW Working Group. 2015. "Elevation-Dependent Warming in Mountain Regions of the World." *Nature Climate Change* 5(5): 424–30. https://doi.org/10.1038/nclimate2563.

Pagano, T., D. Garen, and S. Sorooshian. 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., D.F. Berisford, J.W. Boardman, K.J. Bormann, J.S. Deems, F. Gehrke, A. Hedrick, et al. 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(October): 139–52. https://doi.org/10.1016/j.rse.2016.06.018.

Painter, T.H., D.F. Berisford, J.W. Boardman, K.J. Bormann, J.S. Deems, F. Gehrke, A. Hedrick, et al. 2018. "Database of ASO L4 Lidar Snow Water Equivalent 50 m UTM Grid, Version 1: NASA National Snow and Ice Data Center Distributed Active Archive Center." February 11, 2021. https://nsidc.org/data/aso https://doi.org/10.5067/M4TUH28NHL4Z.

Pflug, J.M., and J.D. Lundquist. 2020. "Inferring Distributed Snow Depth by Leveraging Snow Pattern Repeatability: Investigation Using 47 Lidar Observations in the Tuolumne Watershed, Sierra Nevada, California." *Water Resources Research* 56(9): e2020WR027243. https://doi.org/10.1029/2020WR027243.

Roustant, O., D. Ginsbourger, and Y. Deville. 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.

Schaake, J.C., Jr., and E.L. Peck. 1985. "Analysis of Water Supply Forecast Accuracy." In *Proceedings of the 53rd Annual Western Snow Conference*, edited by B.A. Shafer, 44–53. Boulder, Colorado: Western Snow Conference. https://westernsnowconference.org/bibliography/1985Schaake.pdf.

Siirila-Woodburn, E.R., A.M. Rhoades, B.J. Hatchett, L.S. Huning, J. Szinai, C. Tague, P.S. Nico, et al. 2021. "A Low-to-No Snow Future and Its Impacts on Water Resources in the Western United States." *Nature Reviews Earth and Environment* 2(11): 800–19. https://doi.org/10.1038/s43017-021-00219-y.

Stan Development Team. 2023. "Stan Modeling Language Users Guide and Reference Manual, 2.32.5." https://mc-stan.org.

Staudinger, M., K. Stahl, and J. Seibert. 2014. "A Drought Index Accounting for Snow." *Water Resources Research* 50(10): 7861–72. https://doi.org/10.1002/2013WR015143.

Steinschneider, S., P. Ray, S.H. Rahat, and J. Kucharski. 2019. "A Weather-Regime-Based Stochastic Weather Generator for Climate Vulnerability Assessments of Water Systems in the Western United States." *Water Resources Research* 55(8): 6923–45. https://doi.org/10.1029/2018WR024446.

Stillinger, T., C. Costello, R.C. Bales, and J. Dozier. 2021. "Reservoir Operators React to Uncertainty in Snowmelt Runoff Forecasts." *Journal of Water Resources Planning and Management* 147(10): 06021010. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001437.

Tarricone, J., R.W. Webb, H.-P. Marshall, A.W. Nolin, and F.J. Meyer. 2023. "Estimating Snow Accumulation and Ablation with L-Band Interferometric Synthetic Aperture Radar (InSAR)." *The Cryosphere* 17(5): 1997–2019. https://doi.org/10.5194/tc-17-1997-2023.

Troin, M., R. Arsenault, A.W. Wood, F. Brissette, and J.-L. Martel. 2021. "Generating Ensemble Streamflow Forecasts: A Review of Methods and Approaches over the Past 40 Years." *Water Resources Research* 57(7): e2020WR028392. https://doi.org/10.1029/2020WR028392.

U.S. Geological Survey. 2022. "Database of discharge records for USGS 11264500 MERCED R A HAPPY ISLES BRIDGE NR YOSEMITE CA: National Water Information Center." August 9, 2023. https://waterdata.usgs.gov/nwis/dv/?site_no=11264500.

Viviroli, D., H.H. Dürr, B. Messerli, M. Meybeck, and R. Weingartner. 2007. "Mountains of the World, Water Towers for Humanity: Typology, Mapping, and Global Significance." *Water Resources Research* 43: W07447. https://doi.org/10.1029/2006WR005653.

Waliser, D.E., K.M. Lau, W. Stern, and C. Jones. 2003. "Potential Predictability of the Madden–Julian Oscillation." *Bulletin of the American Meteorological Society* 84(1): 33–50. https://doi.org/10.1175/BAMS-84-1-33.

Webb, R.W., K.N. Musselman, S. Ciafone, K.E. Hale, and N.P. Molotch. 2022. "Extending the Vadose Zone: Characterizing the Role of Snow for Liquid Water Storage and Transmission in Streamflow Generation." *Hydrological Processes* 36(3): e14541. https://doi.org/10.1002/hyp.14541.

Yang, K., K.N. Musselman, K. Rittger, S.A. Margulis, T.H. Painter, and N.P. Molotch. 2022. "Combining Ground-Based and Remotely Sensed Snow Data in a Linear Regression Model for Real-Time Estimation of Snow Water Equivalent." *Advances in Water Resources* 160: 104075. https://doi.org/10.1016/j.advwatres.2021.104075.

Zargar, A., R. Sadiq, B. Naser, and F.I. Khan. 2011. "A Review of Drought Indices." *Environmental Reviews* 19: 333–49. https://doi.org/10.1139/a11-013.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Boardman, Elijah N., Carl E. Renshaw, Robert K. Shriver, Reggie Walters, Bruce McGurk, Thomas H. Painter, Jeffrey S. Deems et al; 2024. "Sources of Seasonal Water Supply Forecast Uncertainty During Snow Drought in the Sierra Nevada." *JAWRA Journal of the American Water Resources Association* 00 (0): 1–19. https://doi.org/10.1111/1752-1688.13221.