High-Elevation Evapotranspiration Estimates During Drought: Using Streamflow and NASA Airborne Snow Observatory SWE Observations to Close the Upper Tuolumne River Basin Water Balance

Brian Henn[1,2], **Thomas H. Painter[3]**, **Kat J. Bormann[3]**, **Bruce McGurk[4]**, **Alan L. Flint[5]**, **Lorraine E. Flint[5]**, **Vince White[6]**, and **Jessica D. Lundquist[2]**

[1]Currently, Center for Western Weather and Water Extremes, Scripps Institution of Oceanography, UC San Diego, La Jolla, CA, USA, [2]Department of Civil and Environmental Engineering, University of Washington, Seattle, WA, USA, [3]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA, [4]McGurk Hydrologic, Orinda, CA, USA, [5]United States Geological Survey, Sacramento, CA, USA, [6]Southern California Edison, Bishop, CA, USA

Abstract Hydrologic variables such as evapotranspiration (ET) and soil water storage are difficult to observe across spatial scales in complex terrain. Streamflow and lidar-derived snow observations provide information about distributed hydrologic processes such as snowmelt, infiltration, and storage. We use a distributed streamflow data set across eight basins in the upper Tuolumne River region of Yosemite National Park in the Sierra Nevada mountain range, and the NASA Airborne Snow Observatory (ASO) lidar-derived snow data set over 3 years (2013–2015) during a prolonged drought in California, to estimate basin-scale water balance components. We compare snowmelt and cumulative precipitation over periods from the ASO flight to the end of the water year against cumulative streamflow observations. The basin water balance residual term (snow melt plus precipitation minus streamflow) is calculated for each basin and year. Using soil moisture observations and hydrologic model simulations, we show that the residual term represents short-term changes in basin water storage over the snowmelt season, but that over the period from peak snow water equivalent (SWE) to the end of summer, it represents cumulative basin-mean ET. Warm-season ET estimated from this approach is 168 (85–252 at 95% confidence), 162 (0–326) and 191 (48–334) mm averaged across the basins in 2013, 2014, and 2015, respectively. These values are lower than previous full-year and point ET estimates in the Sierra Nevada, potentially reflecting reduced ET during drought, the effects of spatial variability, and the part-year time period. Using streamflow and ASO snow observations, we quantify spatially-distributed hydrologic processes otherwise difficult to observe.

Plain Language Summary The amount of evapotranspiration in the Sierra Nevada mountains is important because this water is not available for downstream uses, supports alpine ecosystems, and may change in a future climate. Currently there are few measurements of evapotranspiration in the Sierra Nevada across a diverse landscape. We use a high-resolution snow data set (NASA's Airborne Snow Observatory) with multiple stream gauge observations from Yosemite National Park to estimate evapotranspiration using a water balance approach. Over 2013–2015 during the California drought, we find that evapotranspiration averages 162–191 mm per year, over the time period from peak snowpack in the spring to the end of the summer. Compared with other estimates of evapotranspiration, we find that the estimates are smaller, perhaps due to the diverse spatial terrain sampled by this approach. We also find that the estimates vary only slightly from year to year during the California drought. Our study may help understand how evapotranspiration, and thus available water supply, may change in a warmer future climate.

1. Introduction

Understanding the spatiotemporal scaling of hydrologic variables in heterogeneous terrain is a major research challenge. In mountainous areas, complex spatial patterns and relatively sparse observations typically prevent observing spatially distributed fluxes and storages of water and energy. Spatially distributed basin water storage, snow water equivalent (SWE), and evapotranspiration (ET) are key hydrologic variables

that often exhibit high variability, challenging our ability to represent them in hydrologic models and to validate remotely sensed observations.

Distributed measurements of streamflow offer a means by which to observe the spatially integrated response of a basin to input of water through rain and snowmelt. Streamflow observations allow for evaluation of trends in precipitation that are not observed directly (Luce et al., 2013), the evaluation of precipitation input forcing schemes (Wayand et al., 2013), and the inference of biases in precipitation input forcing (Henn et al., 2015, 2018a; Renard et al., 2011). Distributed streamflow measurements allow for the observations of hydrologic fluxes at scales that may otherwise be difficult to capture, such as localized convective precipitation (Lundquist et al., 2009). In the upper Tuolumne River basin of Yosemite National Park in the Sierra Nevada mountain range of California, USA, distributed streamflow observations made since 2001 (Lundquist et al., 2016) are unique in that they provide unimpaired information about the hydrology of multiple nested and adjacent high-elevation basins in a critical water supply region.

Recent advancements in remote sensing of the snowpack have also improved understanding of snow processes at a range of spatial scales. Lidar measurement of snow depth can provide observations at spatial resolutions and scales that would not be feasible with conventional manual and point measurements (Deems et al., 2013). Airborne lidar systems can cover spatial scales of greater than 1,500 km^2 in a single flight, while still sampling multiple returns per square meter, producing observational data sets with rich spatial coverage of the land surface. Since the spring of 2013, the NASA Airborne Snow Observatory mission (ASO; Painter et al., 2016) has combined an airborne scanning lidar to measure snow depth, an imaging spectrometer to measure snow albedo and identify snow-covered areas, and distributed modeling constrained by in situ measurements to estimate snow density, to produce 50 m gridded estimates of SWE for basins across the Western United States, including in the Tuolumne River.

The combined availability of distributed streamflow and high-resolution SWE observations over an alpine wilderness region without extensive ground observations allows for comparison of the two data sets in order to advance the understanding of basin-scale hydrologic processes. Water balance components may be used to infer unobserved quantities, such as basin soil storage (e.g., Sayama et al., 2011). Studies of snowmelt and streamflow generation at the point scale have illustrated the timing and physical mechanisms by which snowmelt interacts with the soil to produce streamflow and provide water for ET later in the warm season (Bales et al., 2011; Flint et al., 2008; Harpold et al., 2015; Rose et al., 2003). However, extending understanding of these processes and their magnitudes to larger spatial scales has been challenging due to lack of observations. Previous studies estimating basin storage and ET using water balance components have generally followed two approaches. First, ET can be estimated from the difference between precipitation and streamflow over a watershed (e.g., Christensen et al., 2008; Lundquist & Loheide, 2011), though uncertainties in estimating basin-mean precipitation in complex terrain (Henn et al., 2018b) can confound this approach. Second, where snowpack plays a significant role in the water balance in alpine watersheds on the scale of several km^2, others have used intensive manual snow observations to estimate basin-mean SWE and compare against observed streamflow volumes (Hood & Hayashi, 2015; Kattelmann & Elder, 1991; Marks & Dozier, 1992). The spatial pattern of snowpack (and thus basin-mean SWE) can be measured more precisely than spatially-integrated precipitation accumulations over complex terrain (see precipitation uncertainty estimates in section 2), and so this approach may yield more accurate estimates of basin storage and ET. However, practical considerations limit manual snow survey studies to smaller areas or fairly coarse sampling density, and inaccessible terrain often prevents representative sampling of SWE across basins.

Observations of high-elevation ET in the Sierra Nevada are extremely limited. Operating eddy covariance flux tower instrumentation in snow-dominated areas of complex topography and limited maintenance access is challenging. Goulden et al. (2012) used observations from flux towers along a large elevation gradient to show an ET maximum around 2,000 m in elevation of the southern Sierra Nevada. Bair et al. (2015) have used a snow pillow and drainage lysimeters to estimate snowpack sublimation since 2013. However, these point observations of ET and sublimation are too sparse to sample the heterogeneity of vegetative cover, local water availability, topographic aspect, and other small-scale features. Remote sensing approaches based on land surface temperature or vegetative greenness can identify regions associated with higher or lower ET, but may not capture ET well in energy-limited Alpine environments (Goulden et al., 2012). Thus, we are not aware of existing approaches that can measure area-averaged evaporative fluxes

across the mountain landscape. We propose that the combination of spatially-distributed lidar snow and streamflow observations may be used advance understanding of basin-scale hydrologic processes in the mountains.

In this paper, we estimate and compare water budget components for eight alpine basins in and near Yosemite National Park (Table 1 and Figure 1) over the snow ablation and summer seasons of 2013, 2014, and 2015. We use distributed streamflow observations (Lundquist et al., 2016), a large-scale lidar/spectrom-eter snow data set (ASO; Painter et al., 2016), and estimates of precipitation from a network of regional gauges to constrain the season-scale water budgets for each basin and to estimate the unmeasured terms: ET and basin surface and subsurface water storage. The observations used were made during historic drought conditions in California that began during water year 2012. Thus, the findings here may represent abnormally dry conditions rather than conditions typical of the Sierra Nevada climate. In fact, this period coincides with a major drought-induced tree mortality event (Bales, 2015), suggesting that the magnitudes of ET during the drought have been depressed relative to nondrought periods.

For each basin, we estimate cumulative ET over the period from peak SWE and the initiation of snowmelt to the approximate end of the growing season (30 September). Here we define ET as the sum of land surface evaporation, vegetative transpiration and sublimation from snow cover occurring between springtime peak SWE and 30 September; the analysis period does not include any wintertime ET, including sublimation. In the cold-winter, high-elevation Mediterranean climate, warm season ET likely accounts for the majority of the landscape productivity (Kurpius et al., 2003). We choose 30 September as the end of the analysis period because it is approximately the end of the growing season when ET slows as soil moisture and tempera-tures decline. Changes in basin water storage are also likely minimized over this period, as soil moisture is typically low before the onset of major snowmelt, and declines to low levels again by the end of summer (see section 4).

We first discuss the SWE, streamflow and precipitation data sets used to quantify the water balance of each alpine basin (section 2). Then we describe the methods used to estimate the terms of the water balance and to estimate unmeasured quantities (section 3). In section 4, we present the estimates of ET and basin storage across the range of basins and years, and compare them against available observations of soil moisture stor-age and modeled ET amounts. We then discuss the variability and uncertainty of the ET estimates and com-pare them against other Sierra Nevada ET observations (section 5), and summarize key findings (section 6).

2. Data

2.1. Distributed Streamflow Observations

Since 2001, a network of sensors have recorded stream stage in the upper Tuolumne River basin in Yosem-ite National Park (Lundquist et al., 2016). Pressure transducers measure stream stage in tributaries and the main stem of the Tuolumne River (Figure 1) in the vicinity of Tuolumne Meadows, a large subalpine meadow surrounded by extensive high-elevation wilderness in the Sierra Nevada. At each site, the stage-discharge rating curve was established using the depth and discharge observations, surveyed channel cross-sectional profiles, and a Bayesian parameter estimation technique (Le Coz et al., 2014). This method provides credible intervals of the rating curve discharge uncertainty as a function of stream stage, calcu-lated using a 500-sample ensemble of rating curves. See Lundquist et al. (2016) for details on data collection and processing.

We use times series of streamflow from six sites (Figure 1 and Table 1), with basin areas ranging from 7.6 to 190 km^2, and elevations ranging from 2,600 m to just under 4,000 m. We also use two full natural flow time series, from the Tuolumne River at Hetch Hetchy Reservoir (also described in Lundquist et al. (2016)), and Rush Creek at Agnew Lake, just outside the Tuolumne basin east of the crest of the Sierra Nevada. Both of these discharge estimates are derived from measurements of reservoir releases and storage changes made each day, and were made available by the operators of the reservoirs (Hetch Hetchy Water and Power and Southern California Edison, respectively). Uncertainties in full natural flow estimates are not known; we assume that they are of similar magnitude to the distributed streamflow observations ($\pm$5% for seasonal streamflow volumes). The Tuolumne River basin above Hetch Hetchy Reservoir is the largest (1,180 km^2) of the basins and includes lower terrain between 2,000 m and the reservoir elevation at about 1,150 m.

Table 1
Summary of Stream Observations Used in This Study With and Basin Characteristics

	Gauge						Basin				Land cover (NLCD 2011, Jin et al. 2013)			
Number	Name	Latitude	Longitude	Elevation (m)	Type	Area (km²)	Nested basins	10th–50th–90th percentile El. (m)	2013–2015 normalized April–September streamflow (mm)	2013–2015 ASO peak SWE (mm)	Forested (%)	Shrub/ scrub/ grassland (%)	Barren/ ice/ water (%)	Notes
1	Lyell Fk. Tuolumne R. below Maclure Cr.	37.778	−119.261	2940	Solinst in stilling tube	15.6		3080–3335–3623	666	434	5.4%	24.9%	69.7%	
2	Lyell Fk. Tuolumne R. at Twin Bridges	37.869	−119.334	2640	Solinst in stilling tube	109.0	1	2721–3122–3479	327	260	35.3%	36.8%	27.9%	
3	Dana Fk. Tuolumne R. at Bug Camp	37.876	−119.338	2640	Solinst in stilling tube	74.3		2851–3124–3497	214	213	42.3%	34.1%	23.6%	
4	Tuolumne R. at Highway 120	37.875	−119.354	2600	Solinst in stilling tube	189.7	1, 2, 3	2747–3111–3481	269	238	39.5%	35.0%	25.5%	
5	Budd Cr. at Tuolumne Meadows	37.873	−119.382	2600	Solinst in stilling tube	7.6		2669–2915–3130	360	326	42.4%	29.7%	27.8%	
6	Delaney Cr. at Tuolumne Meadows	37.883	−119.381	2600	Solinst in stilling tube	15.6		2720–2975–3332	285	222	55.8%	23.1%	21.1%	
7	Tuolumne R. at Hetch Hetchy Reservoir	37.971	−119.788	1067	Full natural flows estimated at dam	1181.2	1, 2, 3, 4, 5, 6	1991–2756–3200	284	214	35.1%	37.9%	27.0%	
8	Rush Cr. at Agnew Lake	37.759	−119.130	2585	Full natural flows estimated at dam	60.0	-	2819–3105–3457	364	253	21.1%	41.1%	37.9%	ASO data in 2015 only
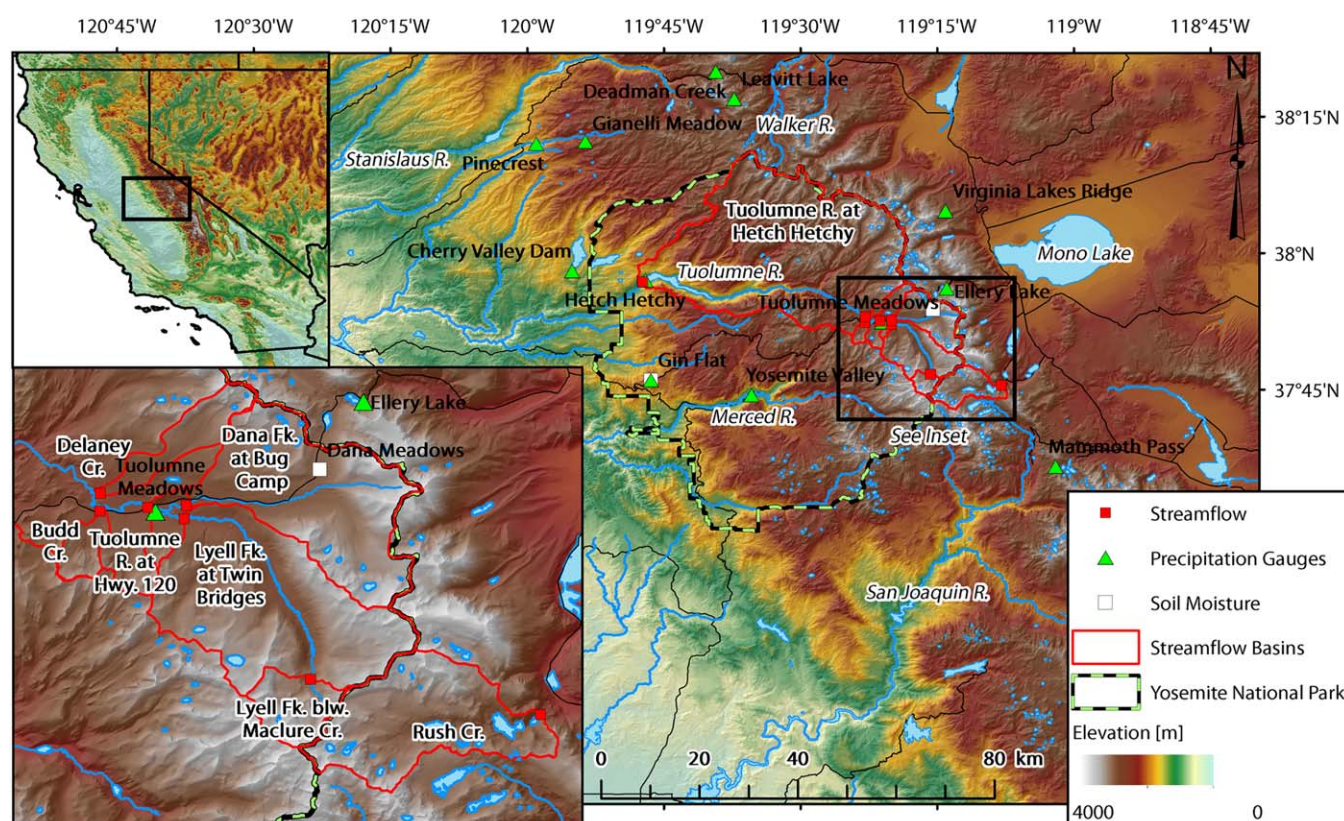
Figure 1. Topographic map of the upper Tuolumne River basin, showing the boundaries of Yosemite National Park in the central Sierra Nevada mountain range of California. The stream gauges used in this study are shown, along with the basin draining to each gauge. Also shown are regional precipitation, temperature, and soil moisture observation sites.

The Lyell and Maclure glaciers, remnants of earlier, more extensive glaciation across the Tuolumne River basin (Basagic & Fountain, 2011), occupied less than 0.5 km^2 of the highest portions of the Lyell Fork below Maclure Creek basin in 2013 (Stock et al., 2017). Glacier areas thus account for less than 3.2% of the basin of the Lyell Fork below Maclure Creek gauge, 1% of the Lyell Fork at Twin Bridges basin, and 0.1% of the Tuolumne River basin above Hetch Hetchy Reservoir. No ice of similar magnitude has been mapped elsewhere in the basins considered in this study.

2.2. ASO Snow Observations

Since water year 2013, NASA's ASO mission has made coupled lidar and spectrometer observations on regular aircraft flights over the upper Tuolumne River basin in Yosemite National Park. As described in Painter et al. (2016), for each flight ASO produces gridded estimates of snow depth at 3 m spatial resolution across the 1,180 km^2 watershed that drains to Hetch Hetchy Reservoir (Figure 1), using data processing algorithms that consider the height difference between snow-on and snow-off lidar surface products, as well as spectrometer identification of snow cover to discriminate against vegetation and other surfaces producing lidar returns. The snow depth product at 3 m resolution is aggregated to 50 m resolution, and merged with distributed estimates of snow density derived from in situ observations and a physically-based snowpack model (iSNOBAL; Marks et al., 1999). This procedure produces 50 m resolution gridded SWE estimates for each flight. Flights begin near the time of peak SWE during late winter or early spring, and proceed over the snowpack ablation season, with average intervals between flights of 5–12 days. The timing and number of flights were different in each water year: 6 flights from 3 April to 8 June in 2013, 11 flights from 23 March to 5 June in 2014, and 10 flights from 17 February to 8 June in 2015. Additionally, 6 flights were conducted over the Rush Creek basin in water year 2015, from 26 March to 9 June, and gridded SWE estimates were provided for this basin using the same methods described above for the Tuolumne basin.

Ground validation of the ASO lidar snow observations showed that uncertainty in snow depth averaged ± 8 cm at 3 m resolution, with nearly zero mean bias (Painter et al., 2016), and less than 1 cm uncertainty when aggregated to 50 m resolution. Errors in density from the distributed snow model as compared to ground validation varied with elevation, and so an elevation-dependent bias correction was applied to the modeled densities in developing the gridded SWE data sets. ASO SWE data validated against weighing snow pillows in the Tuolumne watershed suggested unbiased estimates of SWE per grid cell with uncertainties of 5% for moderate to deeper snowpacks and up to 20% for shallow snowpacks (Painter et al., 2016). Prior work comparing ASO SWE estimates against a regional data set of snow pillows and courses (Henn et al., 2016) showed that the spatially-averaged ASO SWE amounts tend to be less than that of pillows and courses at a given elevation interval, a finding that is likely due to the preferential siting of these types of in situ observations in locations with greater snowpack (Molotch & Bales, 2005; Rice et al., 2011).

2.3. Precipitation and Temperature Observations

In each streamflow basin, we estimate basin-mean precipitation falling over the snow ablation and summer seasons in each water year. We use a network of 12 precipitation gauges located around the Yosemite region (Figure 1). The gauges are of different types (weighing and tipping bucket) and part of a variety of observational networks (National Weather Service Cooperative Observer, Natural Resources Conservation Service Snow Telemetry, and California Department of Water Resources (CDWR) networks), and their coverage is sparse relative to the complex terrain of the upper Tuolumne River basin. All of the gauges are at elevations of 1,000–2,800 m within a region spanning the western foothills of the Sierra Nevada to its eastern slope near Mono Lake. Each gauge provided daily precipitation accumulations, though about 7 percent of the daily data were missing and were not filled. The estimation of daily basin-mean precipitation accumulations from this network of gauges uses Parameter-elevation Regression on Independent Slopes Model (PRISM) 1971–2000 monthly precipitation climatology grids (Daly et al., 1994, 2002, 2008), and is described in section 3.1. Daily high and low temperature observations from the same set of meteorology sites are also used to force hydrologic models of the streamflow basins. More detail on the models and the meteorological data used to drive them is provided in section 3.3.

2.4. Soil Moisture Observations

Soil moisture probes measured vertical profiles of volumetric water content (VWC) at two sites in or near the upper Tuolumne River basin, Dana Meadows and Gin Flat (Figure 1). Dana Meadows is located near the headwaters of the Dana Fork of the Tuolumne River at an elevation of just under 3,000 m, while Gin Flat is located outside of the basin to the southwest, at an elevation of 2,149 m. At both sites, the soil moisture probes are located under a grass surface near CDWR snow pillows, with nearby conifer over-story, at depths of 10, 36, 53, and 79 cm (Dana Meadows) and 10, 36, and 71 cm (Gin Flat) (Flint et al., 2008). Hourly observations of soil moisture were made at each sensor. We averaged the soil moisture observations to daily values, and estimated the soil moisture storage in the top 100 cm via an average of each probe's VWC weighted by the vertical distance between probes.

3. Methods

3.1. Calculation of Basin-Mean SWE, Precipitation, and Temperature

We consider quantities averaged spatially within each of the eight basins shown in Figure 1. We do this to apply the water balance approach using control volumes established by the drainage areas of each stream gauge. We assume that the observed streamflow represents nearly all of the liquid water leaving the basin, as the region is underlain by a major granite batholith (United States Department of Agriculture Natural Resources Conservation Service, 2007) that prevents significant percolation of groundwater into deep aquifer systems. For ASO SWE, we average all of the 50 m grid cells centered within the watershed boundaries shown in Figure 1, including cells with zero SWE, providing estimates of basin-mean SWE.

Precipitation estimates for each basin are calculated in the following manner. First, the 800 m resolution PRISM monthly precipitation climatology is extracted for the cells with precipitation gauges. Second, the PRISM climatology cells located entirely or partially within the boundaries of each basin are used to compute basin-mean precipitation for each month, with fractional weights given to cells partially within the

basin. Third, the ratio of the basin-mean PRISM precipitation to the PRISM precipitation at the gauge grid cell is calculated for each combination of gauges and basins, for each month:

$$R_{i,j,k} = \frac{B_{i,k}}{G_{j,k}} \tag{1}$$

where $B_{i,k}$ is the basin-mean PRISM precipitation for basin i, and $G_{j,k}$ is the PRISM precipitation for the grid cell containing gauge j, all for month k.

The PRISM-scaled estimate of basin-mean precipitation (for basin i, gauge j, and month k) is calculated by multiplying accumulations at the gauge by $R_{i,j,k}$. As we use 12 precipitation gauges (Figure 1), this produces an ensemble of PRISM-scaled estimates of basin-mean precipitation. We take the mean of the 12 estimates to produce single PRISM-scaled precipitation estimates. Because it is difficult to quantify the representativeness and reliability of each gauge, we weight each equally. We also calculate the standard deviation of the ensemble of PRISM-scaled estimates, using an assumption that precipitation accumulations are log-normally distributed (to avoid negative accumulations within the 95% bounds and to reflect their skewed nature).

Daily high and low temperatures at the 12 precipitation sites are also converted to estimates of the basin-mean daily high and low temperature for each basin. This is done in the same manner described above for precipitation, except that temperature is adjusted via an additive PRISM temperature offset instead of multiplicative scaling. The offset for high and low temperature is calculated based on PRISM monthly climatologies, and the PRISM-adjusted estimates produced at each observation site are averaged to produce daily high and low temperature series for each basin.

3.2. Estimation of Water Balance Terms and Residuals

We estimate terms of the water balance for each of the 8 basins in Figure 1 and Table 1, over the snow ablation and summer seasons over 2013–2015. We can write the water balance for each basin over a time period as:

$$\Delta SWE + \sum P - \sum Q - \sum ET = \Delta S \tag{2}$$

where SWE is snow water equivalent, P is precipitation, Q is streamflow, ET is evapotranspiration, and S is basin water storage (e.g., soil moisture, surface depression, lake and channel storage). ET as defined here includes all losses to the atmosphere (surface evaporation, plant transpiration and sublimation from snow). We calculate the water balance from the date of the lidar flight to 30 September of the same year. Because all snow measured during the lidar flight is assumed to have melted by 30 September, snowmelt for the full period (ΔSWE) is simply the amount of SWE measured by ASO. ET in (2) is calculated over the warm season and excludes wintertime snow sublimation. Glacier mass balance change is neglected in (2); Table 1 shows that mapped glacier fractions are small within the basins, with the exception of the Lyell Fork below Maclure Creek. Estimates of recent glacier mass loss in this basin were not yet available but are being produced (Stock et al., 2017).

Over each period, we compute PRISM-scaled precipitation accumulations ($\sum P$) and streamflow accumulations ($\sum Q$). We sum the streamflow time series to produce seasonal streamflow volumes ($\sum Q$), and we sum the daily variances of the streamflow ensemble to estimate the uncertainty of $\sum Q$:

$$\sigma^2_{\sum Q} = \sum_{i=1}^{n} var(Q_i) \tag{3}$$

where n is the number of days in the seasonal period, and $var(Q_i)$ is the variance of the ensemble of daily streamflow volumes, assuming that $\sum Q$ is normally distributed and daily streamflow volume errors are independent.

The other terms in (2) are the sum of evapotranspiration from the basin over the period ($\sum ET$) and the change in soil moisture from the beginning of the period to the end (ΔS). As the first two terms in (2) are water inputs to the basin, and the second two are outputs, the difference must equal changes in storage. All variables in (2) are basin-mean quantities and have units of mm.

To illustrate the estimation of the terms of the basin water balance under this framework, we provide an example using snow, streamflow and precipitation data for the basin of the Tuolumne River at Highway 120 in water year 2013 (Figure 2). The ASO SWE map for the 3 April flight is shown in Figure 2a. In Figure 2b,
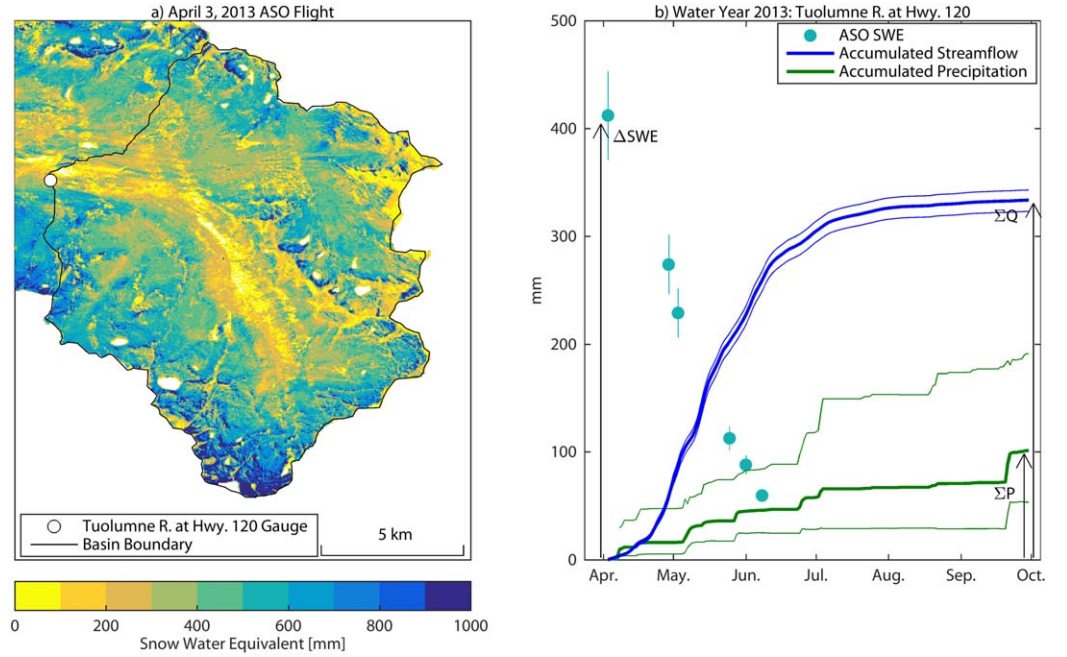
Figure 2. (a) ASO lidar-derived 50 m SWE map for 3 April 2013, over the basin of the Tuolumne River at Highway 120. (b) Example plot for this ASO flight, showing how the basin's water balance is quantified. All SWE from the 3 April flight is assumed to melt by 30 September (ΔSWE); cumulative streamflow ($\sum Q$) and precipitation ($\sum P$) between the flight date and 30 September are then calculated. Uncertainty bounds at 95% confidence are shown for each variable.

time series of SWE, streamflow and precipitation are plotted over the snow ablation and summer period for this basin. The basin-mean SWE for the 3 April flight (ΔSWE) is highlighted, and the trace of cumulative streamflow beginning on that date and ending 30 September ($\sum Q$) is shown. We also show the upper and lower 95% intervals on the cumulative streamflow amounts. A similar trace of cumulative precipitation ($\sum P$) is also plotted in Figure 2b. The 95% bounds on cumulative precipitation show that the uncertainty produced by PRISM-scaled precipitation gauge estimates is large compared to the uncertainty in SWE and streamflow. Qualitatively, Figure 2b shows that over the snow ablation season, streamflow responds to snowmelt with higher flows during the spring; there is much less streamflow once snowmelt ends as precipitation is limited during the region's relatively dry summers.

We rearrange the terms of the water balance equation (3) to find the residual of snowmelt, precipitation and streamflow:

$$\Delta SWE + \sum P - \sum Q = \sum ET + \Delta S. \tag{4}$$

We calculate the residual for the time period from each ASO flight to 30 September. We estimate the uncertainty in $\sum ET + \Delta S$ due to uncertainties in SWE, streamflow and precipitation by summing their variances:

$$\sigma_{\sum ET + \Delta S} = \sqrt{var\left(\sum ET + \Delta S\right)} = \sqrt{var\left(\Delta SWE + \sum P - \sum Q\right)} = \sqrt{\sigma_{\Delta SWE}^2 + \sigma_{\sum P}^2 + \sigma_{\sum Q}^2} \tag{5}$$

where $\sigma_{\Delta SWE}$, $\sigma_{\sum P}$ and $\sigma_{\sum Q}$ are the standard deviations of the ASO SWE, basin-mean PRISM-scaled precipitation accumulations, and streamflow accumulations, respectively. $\sigma_{\Delta SWE}$ is derived from the $\pm 10\%$ uncertainty for SWE at 95% confidence estimated in Painter et al. (2016). As we assume $\sum P$ to be log-normally distributed, its standard deviation is given by:

$$\sigma_{\sum P} = \sqrt{var\left(\sum P\right)} = \sqrt{\left(e^{[\sigma^2]} - 1\right)\left(e^{[2\mu + \sigma^2]}\right)} \tag{6}$$

where μ and σ are the mean and standard deviation, respectively, of the log-transformed precipitation accumulations. Equation (5) assumes that the errors in ASO SWE, the precipitation

estimates and the streamflow volumes are independent, as each comes from an independent data source.

3.3. Hydrologic Model Simulations

In (5), the residual is the sum of two terms that lack observations at the scale of the basins: ET and basin storage. Therefore, in order to partition the residual term between $\sum ET$ and ΔS, we use hydrologic models to simulate basin water storage and evapotranspiration, given the same observations of streamflow, snow and precipitation used in the water balance calculations. We use lumped, conceptual hydrologic models to represent the basin, with the goal of qualitatively understanding the timing and magnitude of changes in soil storage and in evapotranspiration. The modeling setup follows that in Henn et al. (2015), and we refer readers there for full details.

We use the Framework for Understanding Structural Errors hydrologic model (FUSE; Clark et al., 2008, 2011), which represents basin storage as upper and lower zones. An elevation-banded, temperature index snow model based on Snow-17 (Anderson, 2006) is used to simulate SWE; the temperature index parameters include the rain-snow partition temperature, the base melting temperature, and winter and summer temperature melt factors. Snow accumulation and melt are simulated over 100 m elevation bands, and basin-mean quantities are the average of each band weighted by its fraction of basin area. We use a configuration of the model (FUSE-070; Clark et al., 2008) that has relatively few tunable parameters and was previously found to reproduce daily streamflow and ASO SWE observations in the Tuolumne River basin with Nash-Sutcliffe coefficients of 0.8–0.9 (Henn et al., 2015, 2016).

Snowmelt and rain are partitioned into surface runoff and infiltration as a function of upper zone soil storage similar to that of the VIC model (Liang et al., 1994). Percolation from the upper to lower zone is a function of upper zone storage above field capacity, and baseflow from the lower zone is a linear function of storage, similar to that of the PRMS model (Leavesley et al., 1983). Evapotranspiration is allowed only from the upper zone and is a function of potential evapotranspiration (PET) scaled by the saturation of the upper zone. Streamflow is routed using a single delay parameter. For each basin the FUSE model is driven by daily time series of temperature, precipitation and PET. The precipitation and temperature data have been described in section 2.3; the PET forcing series is estimated from the daily minimum and maximum temperature series using a combination of the Bristow-Campbell approximation (Bristow & Campbell, 1984) to estimate available shortwave radiation and the Makkink equation (e.g., Cristea et al., 2013) to estimate PET from shortwave and temperature. Calibration coefficients for both equations were estimated from meteorological data at Dana Meadows, as described in Henn et al. (2015).

We calibrate the FUSE model for each basin using a Bayesian parameter estimation framework (BATEA; Kavetski et al., 2006a, 2006b). BATEA estimates the probability distributions of the model parameters given observations of streamflow and SWE, with the assumption that both forcing (i.e., precipitation) and response (streamflow and SWE) data may have biases. Specifically, BATEA accounts for precipitation forcing biases with a multiplicative error parameter, which is constant with each water year but is independent for each year. BATEA uses prior distributions on the parameters; in the case of the multipliers, these are normal distributions with mean of 1 and standard deviations derived from the PRISM-scaled precipitation gauge ensemble. The goal of the multiplicative error parameters is to allow for adjustment in the highly uncertain basin-mean winter precipitation, such that model simulated SWE can match ASO observations, allowing for more realistic simulation of water balance terms (ET and storage) during the melt season.

The upper and lower soil storage capacities are each fixed at 200 mm for all basins, which approximates the shallow, relatively porous soils of the upper Tuolumne basin (United States Department of Agriculture Natural Resources Conservation Service, 2007). For each basin, the model is calibrated to maximize the parameter posteriors, which combine goodness of fit of the model to streamflow and SWE and the priors of the parameters. Observational uncertainties in SWE and streamflow are set using subjective weights shown to produce model matches to both in Henn et al. (2016).

We calibrate the FUSE models for each of the seven watersheds in the Tuolumne River basin, for water years 2014 and 2015; Rush Creek is not simulated as only 1 year of ASO flights is available, and water year 2013 is not included in the modeling due to uncertainties in the ASO SWE products in that year (see section 5.1). The models are calibrated to the period from the first ASO flight to 30 September of each year; daily

streamflow observations and all ASO flights during that period are included in the calibration. Spin-up is accomplished using water years 2012 and 2013. Due to the climate of the Sierra Nevada (cold, wet winters and warm, dry summers), we hypothesize that there is only minimal carry-over of soil moisture and SWE from year to year, making each year's simulation relatively insensitive to its initial conditions on 1 October.

In summary, the goal of modeling in this study is to differentiate between soil storage changes and ET at different times over the snow ablation season. By forcing the models to match the observed terms of the water balance (SWE, streamflow and precipitation), they should estimate the relative magnitudes of the unobserved terms ($\sum ET$ and ΔS).

4. Results

### 4.1. Comparison of ASO SWE and Distributed Streamflow Observations

Figures 3–5 show ASO SWE plotted against distributed streamflow across the basins over 2013, 2014, and 2015, respectively. In each subpanel, an individual basin is shown, with each point representing a time period beginning with an individual lidar flight. Cumulative streamflow is on the y-axis, and ASO SWE (open circles), or ASO SWE plus cumulative precipitation (filled circles), is on the x-axis. Time advances toward the origin from the upper right, as the SWE and the cumulative precipitation and streamflow between the lidar flight date and 30 September all decline over the ablation season. The 1:1 line indicates where cumulative streamflow is equal to SWE plus cumulative precipitation, and the distances from the 1:1 line to the SWE plus precipitation data points indicate the residual term from (2), $\sum ET + \Delta S$.

The data for water year 2013 (Figure 3) show that the water inputs and outputs are strongly correlated. The R^2 values of linear relationships fitted between ASO SWE plus precipitation and streamflow range from 0.87 to nearly 1 across the seven basins with data in 2013. Most of the data points fall near the 1:1 line, with $\Delta SWE + \sum P > \sum Q$ for nearly all basins and ASO flights. The progression of points toward the origin for each basin in Figure 3 suggests that the basins store only small quantities of water outside of the snow-pack. As periods beginning later in the melt season are considered, the SWE, precipitation, and stream-flow volumes all decline toward zero. However, Figure 3 also shows that precipitation falling during the snow ablation and summer seasons has a nontrivial role in the basins' water balance. The open circles in each panel show that SWE, while much greater than spring and summer precipitation in 2013, is at times less than observed streamflow. Thus, consideration of both ASO SWE and precipitation falling after the flight date is necessary to estimate total water inputs to each basin.

One basin shows markedly different results from the others in terms of the sign of $\sum ET + \Delta S$ in 2013, the Lyell Fork below Maclure Creek. Most or all of the ASO flight dates produce negative $\sum ET + \Delta S$, i.e., stream-flow volumes are greater than the sum of SWE and cumulative precipitation. In the Lyell Fork below Maclure Creek, receding glaciers are likely providing this "extra" water input to the basin; we do not quantify melting ice in this study, but we discuss its role in the water balance in section 5. We also plot the error bounds of the sum of ASO SWE and cumulative precipitation (horizontal), as well as of cumulative streamflow (vertical) in Figures 3–5. The error bars indicate that there is substantial uncertainty in the magnitude of the residual $\sum ET + \Delta S$ (even in its sign in some cases). We discuss uncertainties in the different data types, and the resulting effects on estimating ET, in section 5.

We next consider the same water balance comparison in the snow ablation season of 2014 (Figure 4). As seen in 2013, there are high correlations across all of the basins between SWE plus cumulative precipitation and cumulative streamflow between the flight date and 30 September (R^2 from 0.87 to 0.96). The contribution of precipitation relative to SWE in the total water input of each basin is greater in 2014 than 2013. Peak SWE amounts are less in 2014 than in 2013 across most basins, and cumulative spring and summer precipitation is greater, consistent with decreased wintertime precipitation, increased temperatures, and increased spring precipitation in 2014. A consistent shift is also apparent in the data points, with the residual $\sum ET + \Delta S$ greater in water balance estimates based on the earlier ASO flights (23 March, 7 April and 13 April, in particular). Data points from later flight dates approach, but do not reach, the 1:1 line indicating $\sum ET + \Delta S = 0$.

In water year 2015, a substantially different hydrologic regime is seen in the relative magnitudes of water balance terms (Figure 5). The winter was the warmest on record in the Sierra Nevada, with mean January, February and March temperatures 3–13°C warmer than climatology (PRISM Climate Group, 2015).
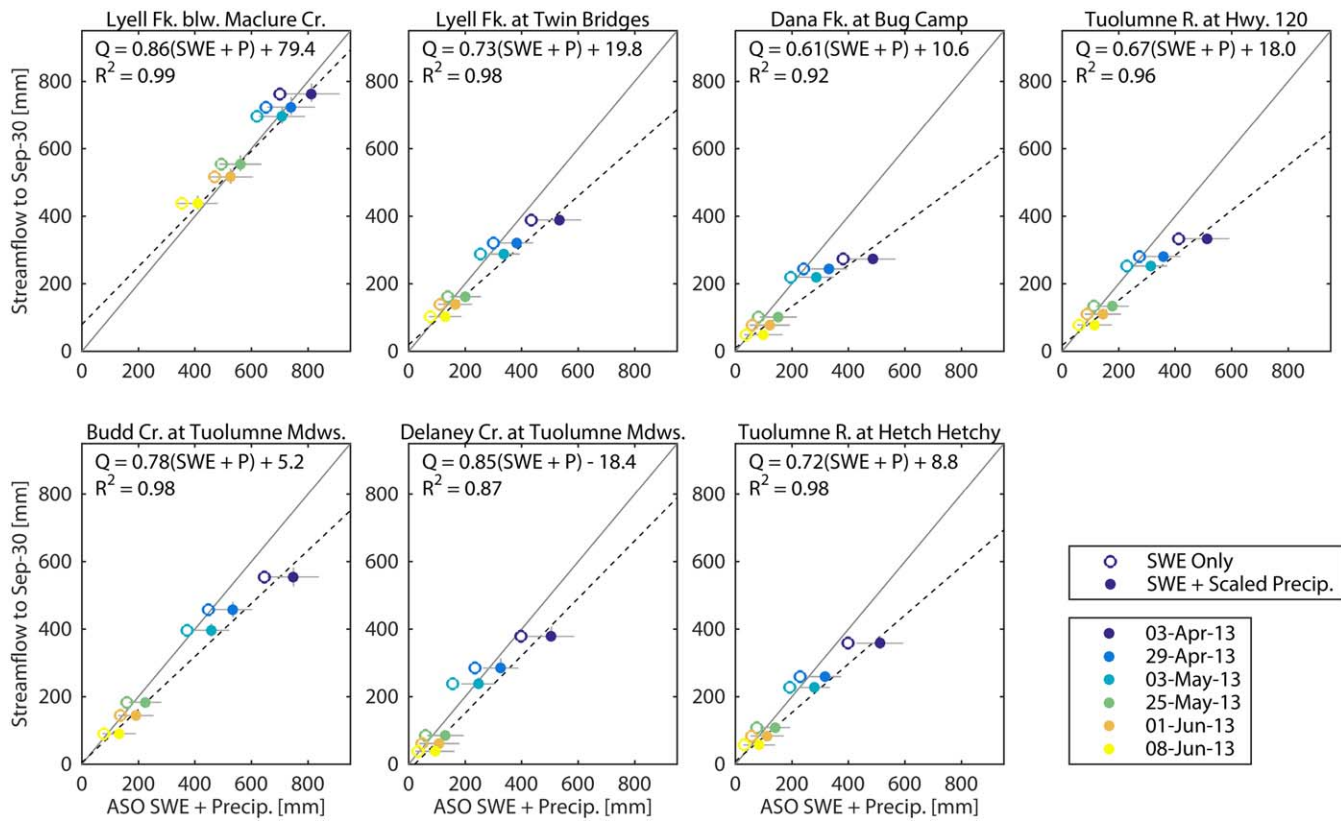
Figure 3. 2013 comparison of water inputs (ΔSWE and $\sum P$) and streamflow ($\sum Q$), computed from each ASO flight date to 30 September, for each of the seven basins with available streamflow and SWE data. The color of the point indicates the ASO flight date used to begin the estimation period; open circles show SWE only, and filled circles show SWE plus precipitation estimated from PRISM-scaled gauge observations. The 1:1 line and a best-fit line are shown for each basin, along with 95% confidence intervals derived from SWE and precipitation uncertainty (horizontal) and streamflow uncertainty (vertical, where available).

Combined with below-average precipitation, the lack of normal Sierra Nevada snowpack in 2015 was estimated to have at least a 500 year return period (Belmecheri et al., 2016). The lack of snow is evident across all of the study basins in Figure 5; however, in terms of streamflow, it is partially offset by spring precipitation falling after the initiation of ASO flights on 17 February. Much of this precipitation fell as rain, and uncertainties in estimating precipitation across high-elevation basins are seen in the larger horizontal uncertainty bounds in 2015.

Nonetheless, 2015 shows many similarities with the prior 2 years in terms of the relationships between water inputs and outputs. The correlations between SWE plus precipitation and streamflow remain high ($R^2 > 0.87$ except for the Lyell Fork below Maclure Creek). The sum of SWE and precipitation is greater than streamflow for all basins and flight dates (again excluding the Lyell below Maclure basin). The residual $\sum ET + \Delta S$ is largest for time periods beginning with the earlier ASO flights, and declines as the beginning of time period extends through the snow ablation season. The results are qualitatively similar across all basins, including the Rush Creek basin located just east of the Tuolumne in the eastern Sierra Nevada.

4.2. Time Series of $\Delta SWE+P-Q$ and Soil Moisture

In Figures 3–5, we plot SWE and cumulative precipitation against cumulative streamflow from the flight date to 30 September, but we can also consider changes in these variables between successive ASO flight dates. Figure 6 shows these progressions for a subset of four basins (Lyell Fork at Twin Bridges, Dana Fork at Bug Camp, Tuolumne River at Highway 120, and Tuolumne River at Hetch Hetchy). For each basin, we calculate the quantity $\Delta SWE+ \sum P- \sum Q$ from the date of one ASO flight to the next, taking into account the change in SWE between flights (defined to be positive for declines in SWE, i.e., snowmelt, and negative for increases in SWE), as well as the accumulated precipitation (whether snow or rain) and accumulated streamflow over the time between flights. (Note that precipitation falling as snow should be captured by increases
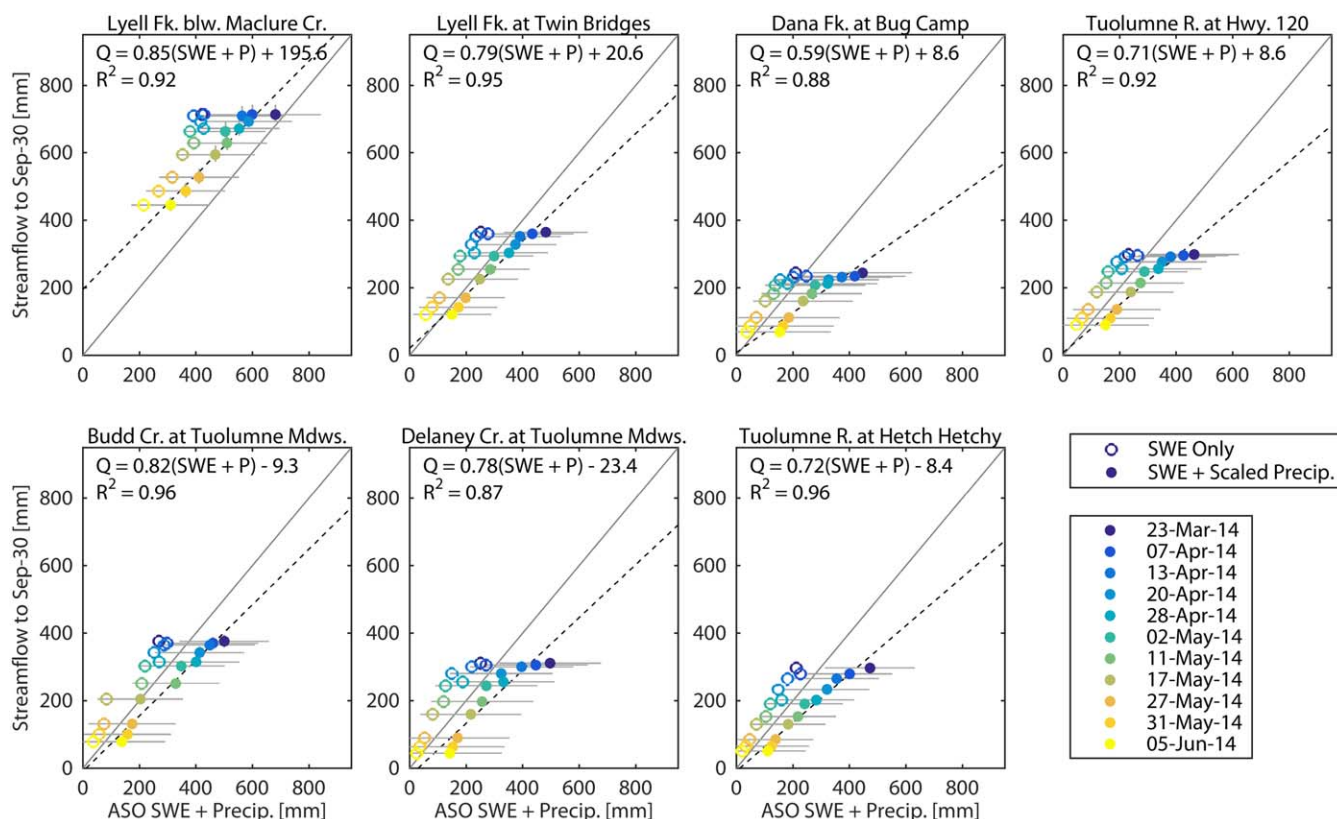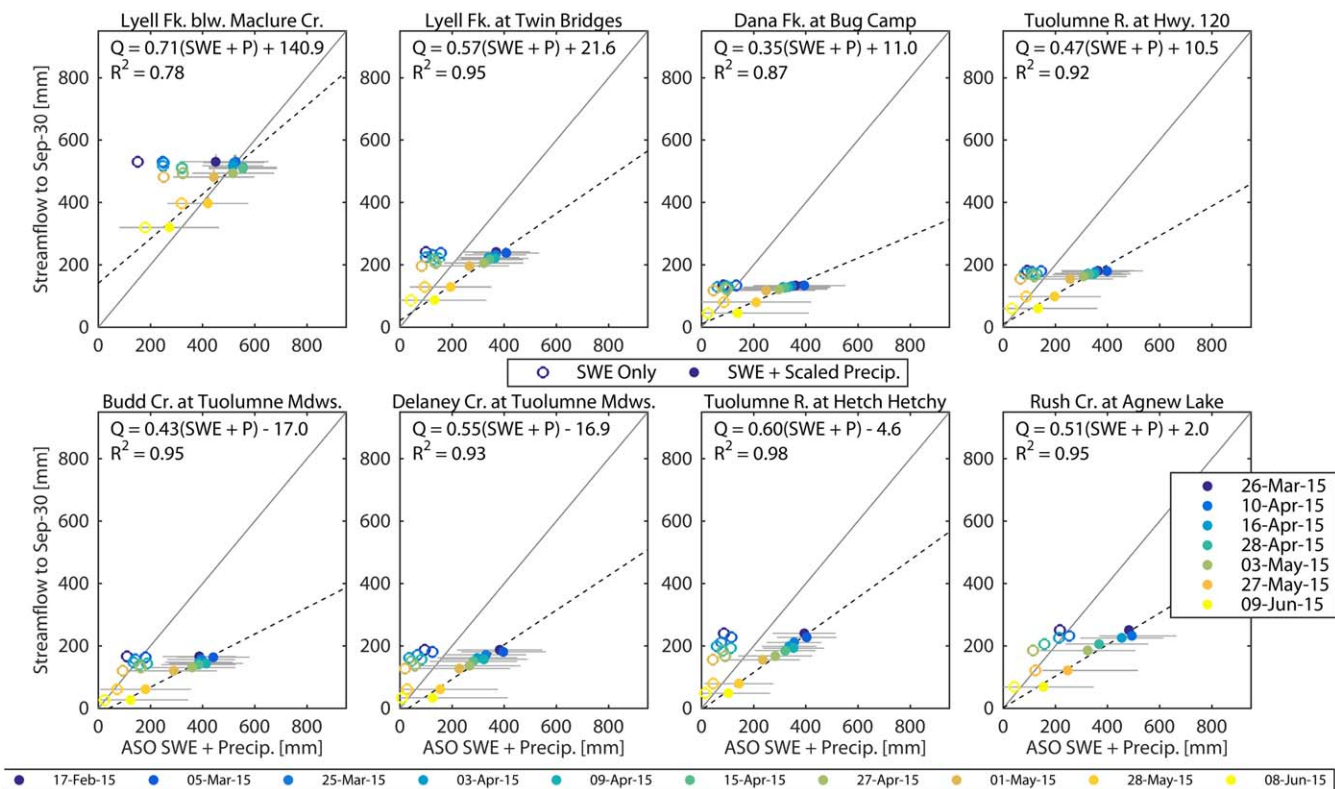
Figure 4. Same as Figure 3, but for 2014.

in ASO SWE, defined here as negative, whereas precipitation falling as rain may produce a streamflow response, so in calculating changes between ASO flights of the residual $\Delta SWE + \sum P - \sum Q$, it is not necessary to discriminate between rain and snow). For the last ASO flight of each water year, we calculate these quantities from the flight date to 30 September. We divide this quantity by the number of days between flights to yield a daily rate of water that is made available to the basins but does not exit via streamflow, i.e., a rate of water flux to the residual term, $\sum ET + \Delta S$.

Figures 6a–6d show that the time series of this flux are quite consistent across the different basins in each of the 3 years. In 2013, each basin had positive rates of change of the residual between the first several ASO flights (3 April, 29 April, and 3 May), meaning that more water was provided to the basins by snowmelt and precipitation than left them by streamflow. This was followed by smaller positive and negative values between the next three flights. From the last flight on 5 June to 30 September shows rates of $\Delta SWE + \sum P - \sum Q$ of 0.23–0.44 mm d^{-1} across the four basins over this summertime period. In 2014, the periods between ASO flights during snowmelt show larger and more variable rates of the residual, but similar agreement on sign and timing between basins. The four basins show positive values between the 23 Mar, 7 April, and 13 April flights, followed by small or negative values between the next three flights, and then positive values of 7–10 mm d^{-1} between 28 April and 2 May. The 2014 summertime (5 June to 30 September) values range from 0.33–0.81 mm d^{1}. Finally, in 2015, the basins show a similar pattern of alternating between low negative rates of $\Delta SWE + \sum P - \sum Q$ and brief periods of higher values (reaching 11 mm d^{-1}) likely associated with rapid snowmelt and infiltration in spring. All basins show small (0.35–0.76 mm d^{-1}) residual values averaged between the last ASO flight (9 June) and 30 September. The agreement between basins in the timing of the residual rates (estimated from data that is largely independent between basins) suggests that the water balance approach is capturing physical signals of the region's hydrology, not simply noise in the observational data.

To estimate how the timing of soil moisture changes may influence the residual term, we compare the time series of the residual with the in situ measurements of soil moisture made at Dana Meadows and Gin Flat

Figure 5. Same as for Figures 3 and 4, but for 2015. Rush Creek is included as ASO flights are available for this basin in 2015, although for a different set of dates than in the Tuolumne basins.

over the same time periods (Figures 6e and 6f). We plot both the daily soil column moisture storage (lines, Figures 6e and 6f), as well as the change in soil moisture storage between flights divided by the number of days between them (bars, Figures 6e and 6f), an equivalent to the average residual changes plotted in Figures 6a–6d.

First, we note the differences in the soil moisture observational time series between Dana Meadows, located at 2,988 m elevation in the sub-alpine zone, and Gin Flat, located at 2,149 m elevation in coniferous forest. The seasonal cycle of soil moisture at Dana Meadows shows low values during the fall and winter, followed by a rapid increase and peak in late spring associated with snowmelt, and a fairly rapid decline over the summer. At Gin Flat, there is higher wintertime soil moisture, presumably due a mix of rain and snow during the winter at the lower site. Next, during the early ablation season of each year (March and April), the positive changes in $\Delta SWE + \sum P - \sum Q$ across the basins approximately correspond to periods of increases in soil moisture at the Dana Meadows site. This suggests that when snowmelt and precipitation exceed streamflow during spring, the difference is being stored in the soil column, i.e., that the residual $\sum ET + \Delta S$ is dominated by the latter term. However, as the year progresses, the relationship between the water balance residual and changes in soil moisture shifts. Changes in soil moisture over the summer are negative, while $\sum ET + \Delta S$ is positive across all basins. In other words, $\Delta S < 0$ and $\sum ET + \Delta S > 0$, and so $\sum ET$ exceeds the residual over late summer by drawing down soil moisture.

It should be noted that the soil moisture sites are not representative of the basin storage as a whole. For example, the magnitude of the in situ soil moisture changes is larger than the changes in the basins' residual terms (note different axis scale between Figures 6a–6d and Figures 6e and 6f), which is likely due to the siting of the soil moisture observations in meadows with deeper soils than the typical steep, rocky terrain. However, the soil moisture observations are likely indicative of the relative seasonal patterns of basin water storage, which is largely driven by snowmelt and winter rains. In the section 4.4, we present modeling results of basin-mean storage and ET.
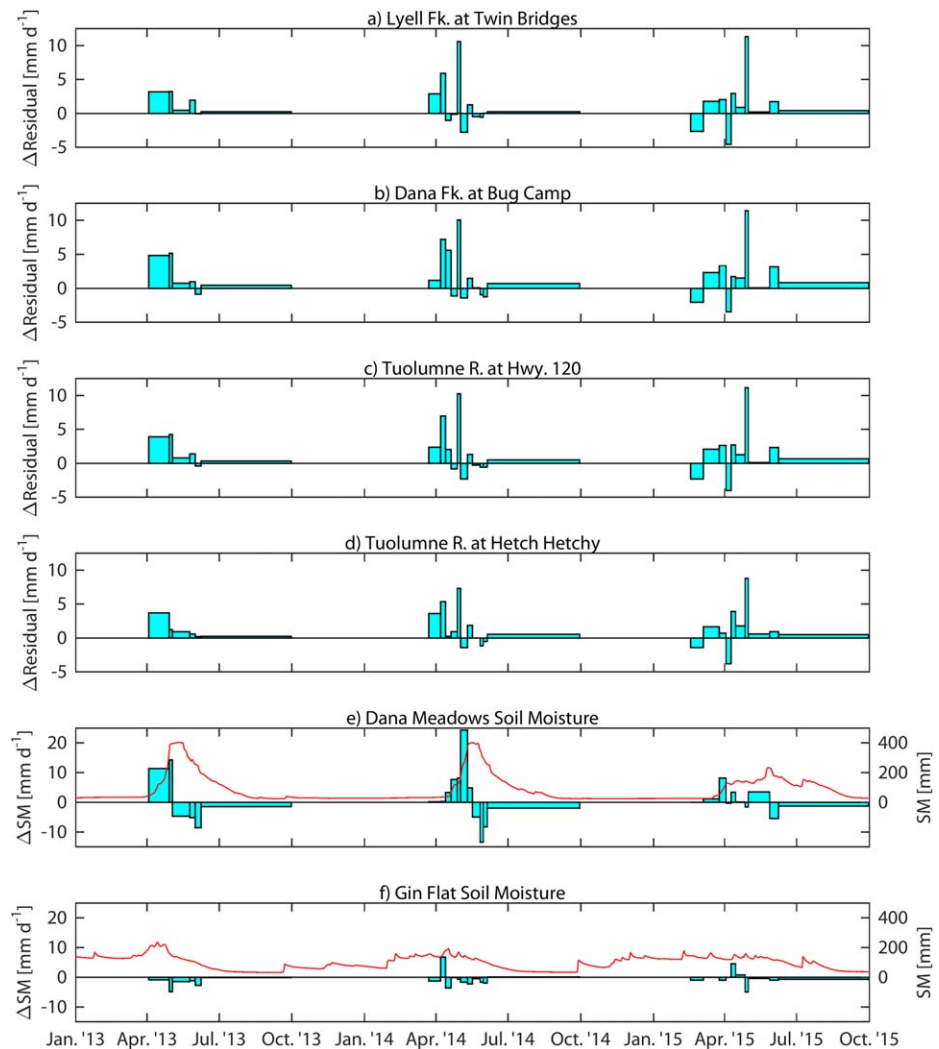
Figure 6. (a–d) Times series of the residual term $\Delta SWE + \sum P - \sum Q$ between ASO flight dates, divided by the number of days between flights to yield a rate of change of the residual, for four basins over 2013–2015. (e and f) Time series of column-integrated soil moisture storage (red lines) at the Dana Meadows and Gin Flat sites, respectively; the bar plots show the rate of changes in soil moisture divided by the number of days between ASO flights.

4.3. Observed and Modeled Estimated *ET* Across Years and Basins

We present the terms of the water balance for each basin and year in Table 2, and plot the residual term for each basin and year in Figure 7, calculated from the first ASO flight date of each year to 30 September. In the observations, the residual values range from below zero (in the Lyell Fork below Maclure Creek, likely due to glacier melt; see section 5), to 227 mm in the Dana Fork basin and 231 mm in the Rush Creek basin in 2015. The differences in the residual term are fairly consistent across the basins in for each year, when the Lyell Fork below Maclure Creek basin is excluded. In 2013, the residuals range from 125 to 213 mm, in 2014, they range from 118 to 203 mm, and in 2015 they range from 127 to 231 mm. Note the scale of the 95% confidence uncertainty bounds for the residual terms (Figure 7 and Table 2), estimated from uncertainties in SWE, cumulative streamflow and cumulative precipitation. The uncertainties have the greatest magnitudes in 2014, largely due to greater spring and summer precipitation. Most of the combined uncertainty is due to the estimation of basin-mean precipitation from gauge observations and PRISM. In 2014, the bounds suggest that the true values may fall from near zero to nearly double the estimated values, whereas in 2013 and 2015, the bounds are relatively smaller (about ±75% of the estimated values). We discuss the robustness of the water balance terms, and the implications for the spatial and annual variability of basin-mean ET, in section 5.

Table 2
*Summary of Observed and Modeled Water Balance Terms for Each Basin and Year*

Basin	Year	Observations			Residual			Model		
		SWE	Precipitation	Streamflow	Mean	95% CI-low	95% CI-high	Streamflow bias (%)	Storage change	ET
Lyell Fk. Tuolumne R. below Maclure Cr.	2013	701	112	763	49	−55	153	-		
	2014	421	260	714	−33	−196	130	−14	28	106
	2015	152	299	531	−80	−213	52	6	−19	104
Lyell Fk. Tuolumne R. at Twin Bridges	2013	434	100	389	145	67	222	-		
	2014	252	230	365	118	−31	266	−2	2	115
	2015	100	269	242	127	−3	257	10	−16	106
Dana Fk. Tuolumne R. at Bug Camp	2013	382	105	274	213	131	294	-		
	2014	209	238	245	203	27	378	−4	10	127
	2015	84	277	134	227	68	386	18	−14	128
Tuolumne R. at Highway 120	2013	412	101	334	180	101	259	-		
	2014	232	232	299	165	7	324	1	3	131
	2015	93	270	182	181	41	321	20	−26	132
Budd Cr. at Tuolumne Meadows	2013	646	103	555	194	99	289	-		
	2014	270	231	376	124	−35	284	18	2	210
	2015	111	277	167	221	80	362	52	−51	205
Delaney Cr. at Tuolumne Meadows	2013	398	107	379	125	41	210	-		
	2014	251	246	311	186	5	367	15	7	121
	2015	95	288	188	196	31	360	9	−28	117
Tuolumne R. at Hetch Hetchy Reservoir	2013	398	114	359	153	69	238	-		
	2014	212	261	297	176	16	335	5	−20	179
	2015	86	308	241	153	33	273	7	−41	189
Rush Cr. at Agnew Lake	2015	219	264	252	231	83	379	-		

4.4. Hydrologic Model Results

In sections 4.1–4.3, we infer the basins' water balance residual, $\sum ET + \Delta S$, by comparing ASO SWE, distributed streamflow, and precipitation observations. However, we lack methods of observing basin-mean soil moisture and ET, and so we also use the FUSE hydrologic model simulations to distinguish between the two over the course of the snow ablation and summer seasons. The model simulations across the basins are qualitatively similar, and so we present detailed model output for one basin (Lyell Fork at Twin Bridges), while presenting a summary of the modeled and observed water balance terms.

Figure 8 shows model output for the Lyell Fork basin. SWE observations over the ablation season, and the associated streamflow peak, are simulated well by the calibrated model (Figure 8a). The Nash-Sutcliffe of the streamflow simulation over 2014 and 2015 is 0.91, and the SWE simulation shows little bias, though the rate of snowmelt is somewhat overestimated each year in the model. The match to observed streamflow and SWE suggests that the FUSE model reproduces the magnitudes of the water balance over the snow ablation season. Thus, we consider the simulation of basin water storage and ET (Figure 8b). The simulated storage shows a strong seasonal peak in May of 2014; storage peaks in early June 2015 but at a lower value. It ranges from less than 10 mm in the winter to 76 mm around peak snowmelt in 2014. The seasonal cycle of modeled basin storage matches the timing of the soil moisture observations at Dana Meadows. Note, however, that the observed soil moisture is plotted at 1/5 of its magnitude for better visual comparison to the observations; as was previously noted, the soil moisture sites are likely not representative of the basin-mean storage magnitude.

The model estimates of storage and ET allow for evaluation of the relative magnitudes of these terms in the basin's water balance estimates. For the Lyell Fork at Twin Bridges basin, modeled changes in storage between the first ASO flight and 30 September are 2 and −51 mm over 2014 and 2015. In contrast, modeled $\sum ET$ is 211 and 205 mm for those years. Thus, the residual term $\sum ET + \Delta S$ is likely dominated by $\sum ET$ when integrated from peak SWE to 30 September, as net differences in soil moisture are small in comparison, despite larger changes within the period. The negative values of ΔS suggest that $\sum ET$ may be slightly larger than $\sum ET + \Delta S$, as the net decline in storage provides additional water for ET. The modeled ET time series in Figure 8b shows peaks around 1 mm d^{-1} approximately corresponding to peak storage in spring, declining with soil moisture by an order of magnitude by late summer.
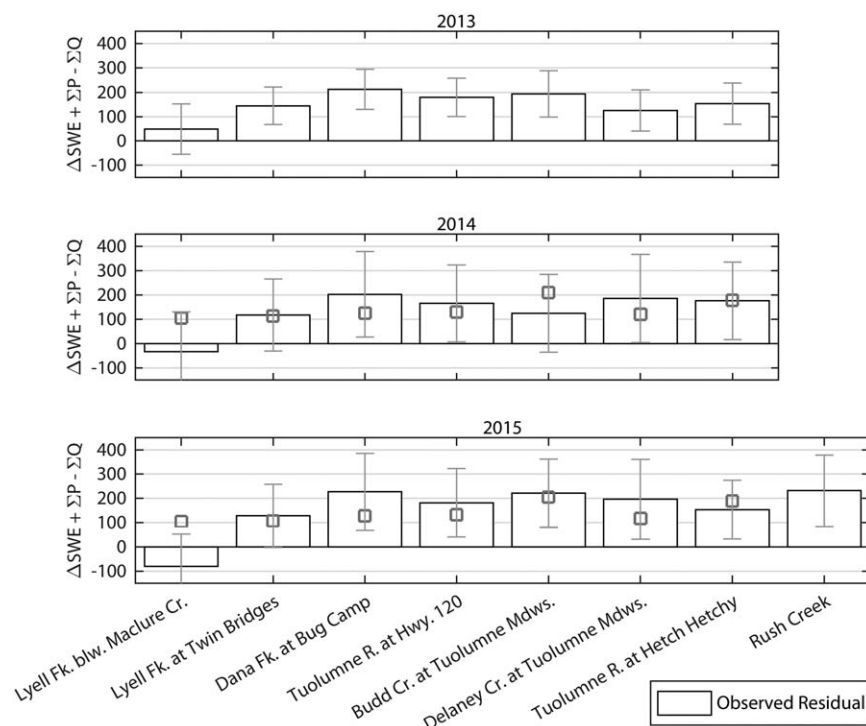
Figure 7. 2013–2015 basin estimates of the residual term ($\Delta SWE + \sum P - \sum Q$) calculated over the period from the first ASO flight date of each year to 30 September. Error bars indicate uncertainties in the residual. Modeled ET amounts are shown with squares for 2014 and 2015. Error bars for Budd Creek, the Tuolumne at Hetch Hetchy and Rush Creek are calculating using assumed ±5% uncertainty in streamflow volumes, as no estimates of streamflow volume uncertainty are available for those basins. See Table 2 for numerical values of the observed and modeled water balance terms.

We present the modeled basin storage and cumulative ET over seven basins in Table 2 (modeled ET also shown as squares in Figure 7). Net changes in modeled soil moisture average −12 mm across the basins over 2014 and 2015, and range from −51 mm to +28 mm. Modeled ET, in comparison, averages 141 mm. Thus, the modeling also supports the hypothesis that the majority of the water balance residual $\sum ET + \Delta S$ from peak SWE to 30 September comprises $\sum ET$. The average modeled residual $\sum ET + \Delta S$ is 129 mm over the simulated basins in 2014 and 2015; compared to the averaged observed residual of 140 mm over these basins and years, it suggests that the models generally capture the magnitudes of the water balance terms. The streamflow volume biases in the models across basins and years (bias magnitudes average ±13%, Table 2) also suggest that the model capture the magnitudes of the water balance terms.

Despite the fairly consistent storage and ET amounts, the calibrated model parameters do show variability between basins. In particular, the precipitation multiplier ranges from 0.59 to 1.23; differences between basins in this parameter are fairly consistent between 2014 and 2015. For example, the Lyell Fork at Twin Bridges basin had multipliers of 0.82 and 0.64 in 2014 and 2015, while Budd Creek had multipliers of 1.23 and 0.92. This suggests that PRISM-based precipitation estimate biases may differ from basin to basin but be consistent from year to year. Other hydrologic model parameters show variability between basins, including the percolation rates between the upper and lower soil zones and the snowmelt factors. We discuss the variability in model parameters between basins further in section 5.3.

5. Discussion

5.1. Data Uncertainty and Estimates of ET

The uncertainties in the different data sources used to estimate the water balance terms impact the robustness of the estimates of ET and storage. As shown by the error bars in Figures 2−5 and 7, we estimate uncertainty in ASO SWE, basin-mean precipitation accumulations, and cumulative streamflow volumes. For precipitation, the uncertainties are on the order of ±75% of the cumulative amounts over the ablation and
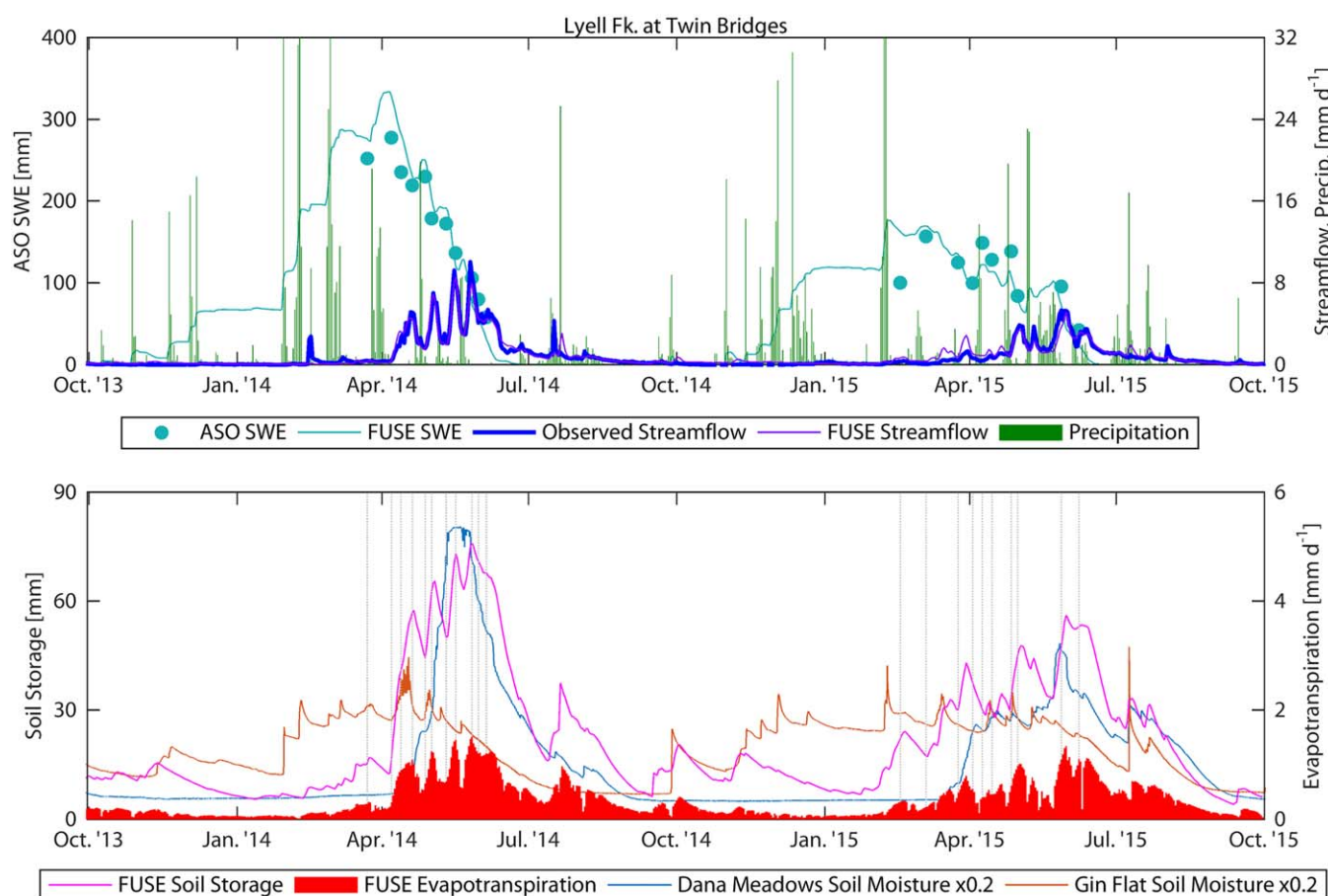
Figure 8. Observations and model simulations for the Lyell Fork at Twin Bridges over 2014 and 2015. (a) SWE, precipitation and streamflow data: ASO and FUSE modeled SWE are shown in light blue; modeled and observed streamflow are shown in thin and heavy dark blue lines; precipitation shown with green bars. (b) Basin water storage and ET simulation: FUSE model basin-mean storage shown against the Dana Meadows and Gin Flat soil moisture storage observations; note that the soil moisture observations are shown at 1/5 of their values for visibility. FUSE simulated basin-mean ET is shown in red bars.

summer seasons, a result of the disagreements between the different PRISM-scaled gauge estimates of basin-mean precipitation in a region of topographic complexity and orographic effects. In particular, summer convective precipitation over high terrain is poorly captured by the sparse observational network (Figure 1; also see Lundquist et al. (2009)), which produces divergent estimates of basin-mean precipitation. For streamflow, we use the gauges' rating curve uncertainty estimates to generate bounds on the cumulative volumes, which produces uncertainties of ±3 to 6% at 95% confidence. The uncertainties in the ASO SWE estimates are spatially variable, but here are assumed to be $\pm10\%$ based on ground validation from meadow areas presented in Painter et al. (2016), where snow depth errors were small and errors in the modeled snow density estimates were found to be 8–13% at the 50 m grid cell. The uncertainty for the residual is the sum of the variances of the other terms of the water balance in equation (5), and the results of these methodological choices are the bounds shown in Figure 7. It is important to rigorously quantify the uncertainties in the terms of the basins' water balance in order to apply the residual approach used here, in particular the uncertainty in basin-mean precipitation.

To some extent, we can attribute the relative contributions of the uncertainties from the streamflow, snow, and precipitation observations. For example, in 2013 relatively little precipitation was observed during the spring and summer months, and so nearly all of the water input to the basins over this period came from snowmelt. Thus, the uncertainty contribution from precipitation is smaller compared to 2014 and 2015 (though still significant), and this is reflected in the smaller uncertainty magnitudes in estimated ET. Additionally, there are uncertainties in lidar estimates of snow depth and modeled snow density under the forest canopy. ASO used a lower quality lidar instrument in 2013 than the Riegl Q1560 dual laser scanning system

used in subsequent years, and methods for snow/no-snow discrimination under the forest canopy, as well as in in the snow density modeling methodology, have been evolving based on experience. These uncertainties may impact the ET estimates presented here. For example, the Delaney Creek basin, which had the lowest ET estimate in 2013 outside of the glacier-affected Lyell Fork below Maclure Creek basin, has the highest fraction of forest cover among all of the study basins (Table 1). Explaining inter-annual variability of the ET estimates may be aided by observations from more basins and from additional years with greater variability in precipitation, snow and streamflow magnitudes.

5.2. Glacier Melt Effects on Lyell Basin Water Balance

Extensive mass loss from the Lyell and Maclure glaciers resulted from the California drought that began in 2012. During the low-snow winters of water years 2014 and 2015, multiple meters of ice downwasting were reported each year (Stock et al., 2017). The additional streamflow is apparent in the points above the 1:1 line in the Lyell Fork below Maclure Creek each year (Figures 3–5), i.e., where $\Delta SWE + \sum P < \sum Q$. The effect of the additional melt grows smaller as the basins' areas increase moving downstream, but may slightly impact the results in other basins: for example, if 1 m of ice melted over 0.5 km^2 during the warm season, it would result in up to 32 mm of streamflow in the Lyell Fork below Maclure Creek basin, 5 mm in the Lyell Fork at Twin Bridges basin, and <1 mm for the basin of the Tuolumne River at Hetch Hetchy.

The Lyell Fork below Maclure Creek basins is also subject to greater uncertainty in precipitation than the other basins. Due to its high elevation, the PRISM-derived estimates diverge more in this basin (Figures 3–5), and so it is possible that the low residual values are partly due to precipitation uncertainty, if precipitation amounts are strongly underestimated at high elevations. An additional, smaller uncertainty occurs when changes in the glacier ice surface elevation produce biased lidar snow-off surface elevations, and thus errors in ASO snow depths. We attribute the inconsistent results in the Lyell Fork below Maclure Creek basin primarily to glacier melt effects, and argue that these effects are smaller in the other basins.

5.3. ET Climatic and Spatial Distributions

Assuming that the basins' water balance residual terms in Figure 7 approximate basin-mean ET over the spring and summer seasons, this study can assess the variability of ET between different basins and years. First, we note that the magnitude of variations in the residual between basins each year are relatively small and inconsistent from year to year, though some basin differences do appear robust, such as the Dana Fork basin having greater residual values than the Lyell and combined Tuolumne at Highway 120 basins. This could be due to the Dana Fork having a greater fraction of forest cover (Table 1; see previous discussion of the lidar uncertainties under the canopy), to the Dana Fork having greater area underlain by metamorphic rock rather than granodiorite that results in deeper and more developed soils, or to the Lyell Fork having greater glacial melt contributions. As noted above regarding data uncertainty, we do not find a positive relationship between the degree of forest cover in each basin and its ET estimate, as more heavily forested basins (e.g., Delaney Creek, Table 1) do not correspond to higher ET estimates (Figure 7).

Additionally, the residuals are not obviously correlated with elevation, though most of the basins here span relatively similar elevations (Table 1). The larger basin of the Tuolumne River at Hetch Hetchy, which unlike the other basins contains terrain below 2,600 m, does not have significantly different ET magnitudes (Figure 7 and Table 2), and so we do not find evidence of a strong ET-elevation gradient. However, the mean elevation of the Hetch Hetchy basin is only slightly lower than the other basins, which makes it difficult to assess any relationship; the Hetch Hetchy basin also includes significant exposed rock (Table 1). The differences between 2013, 2014, and 2015 are relatively small as well, though the residual values are slightly greater in 2015, a year with lower precipitation than 2013 and 2014 but warmer temperatures in the central Sierra Nevada. Thus, the increase in estimated ET from 2014 to 2015 could support the hypothesis that higher temperatures drive greater ET in high-elevation Sierra Nevada ecosystems due to an increase in the growing season length (Goulden & Bales, 2014). Overall, however, an extension of this approach to a wider range of basins and years would better estimate inter-annual and elevation-driven differences in ET.

The FUSE model simulations presented in section 4.4 suggest that uncertainty remains in the water balance terms and their differences across basins. While the observed and simulated ET is fairly consistent across years and basins, model parameters such as the precipitation multiplier and percolation rates vary between basins. The differences in the multiplier between basins may be attributed to discrepancies between the

basin-mean precipitation inputs to the model derived from gauges and PRISM, and the lidar-derived ASO SWE observations; in these case the multiplier allows the model to match SWE observations given uncertain precipitation. Given these uncertainties, it is difficult to interpret hydrologic model parameters in terms of physical differences between basins. Instead, we use the model findings to support our assumptions about the regional timing and magnitude of basin storage and ET.

5.4. Comparisons to Other ET Estimates

Finally, we place our findings in the context of other observations of ET in the Sierra Nevada. Goulden et al. (2012) found ET of 350–550 mm yr^{-1} at a flux tower site at 2,700 m elevation over 2 water years with above-average precipitation, and estimated mean ET of 400–450 mm yr^{-1} over the Kings River basin using a NDVI-based extrapolation of flux tower observations. Kurpius et al. (2003) found ET of 600–800 mm over a year at a forest canopy flux tower and sap flow site (Blodgett Forest, 1,300 m elevation). Using potential ET and water availability accounting to estimate actual ET, Kattelmann and Elder (1991) found ET of 400 mm yr^{-1} over an above-treeline alpine basin, and found that the basin's water balance closed when also considering streamflow, snow and precipitation observations. Marks and Dozier (1992) used a Monin-Obukhov approach to estimate snowpack sublimation losses from two snow sites at 500 mm. Bair et al. (2015) used a snow pillow and drainage lysimeters to estimate that sublimation averaged 477 mm yr^{-1} over 2013–2015 at Mammoth Mountain, approximately 40 km southeast of Tuolumne Meadows. That study is perhaps most comparable to this one, in location and methodology.

Here we estimate ET from peak SWE through September, and so those wintertime sublimation amounts would not be included in our estimates. However, our results may show lower amounts of ET, even accounting for the partial-year period of our study. This may be due to a combination of factors, including the spatial variability of ET and sublimation. The flux tower sites are located above forest canopies, whereas forests cover less than half of the basins here; (Table 1). The drought conditions over 2013–2015 that may have reduced ET as evidenced by large-scale drought-induced tree mortality in the Sierra Nevada over this time (Bales, 2015). The partial-year analysis conducted here that does not include wintertime sublimation losses from the snowpack, such as those observed by Bair et al. (2015). Further work is needed to constrain ET magnitudes and variability over the alpine basins of the Sierra Nevada, given their crucial water supply role and susceptibility to changes in climate. This includes comparisons of ET estimates derived from basin water balance, flux towers and other point measurements, and remotely-sensed energy balance approaches.

6. Conclusions

In this study, we compare water balance components based on streamflow observations and ASO SWE in the upper Tuolumne River basin of Yosemite National Park, during drought conditions for water years 2013–2015. We estimate the difference between cumulative streamflow and the sum of ASO SWE and cumulative precipitation from the date of the first ASO flight near peak SWE to 30 September. The sum of water inputs (snowmelt and precipitation) exceeds outputs (streamflow) by amounts that range from 118 to 231 mm. By considering soil moisture observations and a hydrologic model that simulates basin water storage over the snowmelt and summer seasons, we show that this residual term in the water balance largely represents cumulative ET from the basin, as changes in basin water storage are relatively small from the time of peak SWE to late summer.

We find that high-elevation ET amounts are fairly robust from basin to basin, and that the amounts are greater in 2015 (average 191 mm, range 48–334 mm) relative to 2013 (average 168 mm, range 85–252 mm) and 2014 (average 162 mm, range 0–326 mm). The reasons for a shift between years is not clear, as it does not correspond to the assumption that ET is water-limited during drought years, and could be due to observational uncertainties in the data. Uncertainties in the estimated ET amounts, primarily due to uncertainty in precipitation, are significant, with 95% bounds on the ET estimates ranging from near zero to nearly double the values presented here. The presence of small, receding glaciers in one basin is seen in the basin water balance in the form of additional streamflow.

We also find that ET amounts are less than those found from point observations of ET over forested areas of the Sierra Nevada, a finding that may be due to a combination of the spatially variable nature of ET and drought effects over 2013–2015. We also find that basin-mean ET from peak SWE to the end of the water

year is less (with at least 95% confidence across all of the basins studied) than the 400–500 mm per full year found in high-elevation Sierra Nevada basins in prior studies.

Finally, the findings above show that distributed streamflow and ASO snow observations can be used to advance understanding of hydrologic processes across a range of scales in complex terrain. In this study, we produce physically-based estimates of the temporal evolution of basin water storage and basin-mean ET in alpine watersheds, quantities that are otherwise difficult to observe and are crucial for improving hydrologic process understanding.

References

Anderson, E. (2006). *Snow accumulation and ablation model–SNOW-17*, Silver Spring, MD.

Bair, E. H., Dozier, J., Davis, R. E., Colee, M. T., & Claffey, K. J. (2015). CUES—A study site for measuring snowpack energy balance in the Sierra Nevada. *Frontiers of Earth Science*, *3*(September). https://doi.org/10.3389/feart.2015.00058

Bales, R. C. (2015). *Drought effects on evapotranspiration and subsurface water storage in the southern Sierra Nevada*, Abstract #H33M-06 presented at 2015 Fall Meeting, AGU, San Francisco, CA.

Bales, R. C., Hopmans, J. W., O'geen, A. T., Meadows, M., Hartsough, P. C., Kirchner, P., et al. (2011). Soil moisture response to snowmelt and rainfall in a Sierra Nevada mixed-conifer forest. *Vadose Zone Journal*, *10*(3), 786. https://doi.org/10.2136/vzj2011.0001

Basagic, H. J., & Fountain, A. G. (2011). Quantifying 20th century glacier change in the Sierra Nevada, California. *Arctic, Antarctic, and Alpine Research*, *43*(3), 317–330. https://doi.org/10.1657/1938-4246-43.3.317

Belmecheri, S., Babst, F., Wahl, E. R., Stahle, D. W., & Trouet, V. (2016). Multi-century evaluation of Sierra Nevada snowpack. *Nature Climatic Change*, *6*(1), 2–3. https://doi.org/10.1038/nclimate2809

Bristow, K., & Campbell, G. (1984). On the relationship between incoming solar radiation and daily maximum and minimum temperature. *Agricultural and Forest Meteorology*, *31*(427), 159–166.

Christensen, L., Tague, C. L., & Baron, J. S. (2008). Spatial patterns of simulated transpiration response to climate variability in a snow dominated mountain ecosystem. *Hydrological Processes*, *22*(18), 3576–3588. https://doi.org/10.1002/hyp.6961

Clark, M. P., McMillan, H. K., Collins, D. B. G., Kavetski, D., & Woods, R. A. (2011). Hydrological field data from a modeller's perspective: Part 2: Process-based evaluation of model hypotheses. *Hydrological Processes*, *25*(4), 523–543. https://doi.org/10.1002/hyp.7902

Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., et al. (2008). Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models. *Water Resources Research*, *44*, W00B02. https://doi.org/10.1029/2007WR006735

Cristea, N., Kampf, S., & Burges, S. (2013). Revised coefficients for Priestley-Taylor and Makkink-Hansen equations for estimating daily reference evapotranspiration. *Journal of Hydrologic Engineering*, *18*(October), 1289–1300. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000679

Daly, C., Gibson, W., & Taylor, G. (2002). A knowledge-based approach to the statistical mapping of climate. *Climate Research*, *22*, 99–113.

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

Daly, C., Neilson, R., & Phillips, D. (1994). A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology*, *33*.

Deems, J. S., Painter, T. H., & Finnegan, D. C. (2013). Lidar measurement of snow depth: A review. *Journal of Glaciology*, *59*(215), 467–479. https://doi.org/10.3189/2013JoG12J154

Flint, A. L., Flint, L. E., & Dettinger, M. D. (2008). Modeling soil moisture processes and recharge under a melting snowpack. *Vadose Zone Journal*, *7*(1), 350. https://doi.org/10.2136/vzj2006.0135

Goulden, M. L., Anderson, R. G., Bales, R. C., Kelly, A. E., Meadows, M., & Winston, G. C. (2012). Evapotranspiration along an elevation gradient in California's Sierra Nevada. *Journal of Geophysical Research*, *117*, G03028. https://doi.org/10.1029/2012JG002027

Goulden, M. L., & Bales, R. C. (2014). Mountain runoff vulnerability to increased evapotranspiration with vegetation expansion. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(39), 14071–14075. https://doi.org/10.1073/pnas.1319316111

Harpold, A. A., Molotch, N. P., Musselman, K. N., Bales, R. C., Kirchner, P. B., Litvak, M., & Brooks, P. D. (2015). Soil moisture response to snowmelt timing in mixed-conifer subalpine forests. *Hydrological Processes*, *29*(12), 2782–2798. https://doi.org/10.1002/hyp.10400

Henn, B., Clark, M. P., Kavetski, D., & Lundquist, J. D. (2015). Estimating mountain basin-mean precipitation from streamflow using Bayesian inference. *Water Resources Research*, *51*, 8012–8033. https://doi.org/10.1002/2014WR016736

Henn, B., Clark, M. P., Kavetski, D., McGurk, B., Painter, T. H., & Lundquist, J. D. (2016). Combining snow, streamflow, and precipitation gauge observations to infer basin-mean precipitation. *Water Resources Research*, *52*, 8700–8723. https://doi.org/10.1002/2015WR018564

Henn, B., Clark, M. P., Kavetski, D., Newman, A. J., Hughes, M., McGurk, B., & Lundquist, J. D. (2018a). Spatiotemporal patterns of precipitation inferred from streamflow observations across the Sierra Nevada mountain range. *Journal of Hydrology*, *556*, 993–1012. https://doi.org/10.1016/j.jhydrol.2016.08.009

Henn, B., Newman, A. J., Livneh, B., Daly, C., & Lundquist, J. D. (2018b). 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

Hood, J. L., & Hayashi, M. (2015). Characterization of snowmelt flux and groundwater storage in an alpine headwater basin. *Journal of Hydrology*, *521*, 482–497. https://doi.org/10.1016/j.jhydrol.2014.12.041

Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., & Xian, G. (2013). A comprehensive change detection method for updating the National Land Cover Database to circa 2011. *Remote Sensing of Environment*, *132*, 159–175. http://doi.org/10.1016/j.rse.2013.01.012

Kattelmann, R., & Elder, K. (1991). Hydrologic characteristics and water-balance of an alpine basin in the Sierra-Nevada. *Water Resources Research*, *27*(7), 1553–1562. https://doi.org/10.1029/90WR02771

Kavetski, D., Kuczera, G., & Franks, S. W. (2006a). Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. *Water Resources Research*, *42*, W03407. https://doi.org/10.1029/2005WR004368

Kavetski, D., Kuczera, G., & Franks, S. W. (2006b). Bayesian analysis of input uncertainty in hydrological modeling: 2. Application. *Water Resources Research*, *42*, W03408. https://doi.org/10.1029/2005WR004376

Kurpius, M. R., Panek, J. A., Nikolov, N. T., McKay, M., & Goldstein, A. H. (2003). Partitioning of water flux in a Sierra Nevada ponderosa pine plantation. *Agricultural and Forest Meteorology*, *117*(3–4), 173–192. https://doi.org/10.1016/S0168-1923(03)00062-5

Leavesley, G., Lichty, Troutman, R. B., & Saindon, L. (1983). *Precipitation-runoff modeling system: User's manual*. Washington, DC: USGS.

Le Coz, J., Renard, B., Bonnifait, L., Branger, F., & Le Boursicaud, R. (2014). Combining hydraulic knowledge and uncertain gaugings in the estimation of hydrometric rating curves: A Bayesian approach. *Journal of Hydrology*, *509*, 573–587. https://doi.org/10.1016/j.jhydrol.2013.11.016

Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research*, *99*(D7), 14415–14428.

Luce, C. H., Abatzoglou, J. T., & Holden, Z. A. (2013). The missing mountain water: Slower westerlies decrease orographic enhancement in the Pacific Northwest USA. *Science*, *342*(6164), 1360–1364. https://doi.org/10.1126/science.1242335

Lundquist, J., Huggett, B., Roop, H., & Low, N. (2009). Use of spatially distributed stream stage recorders to augment rain gages by identifying locations of thunderstorm precipitation and distinguishing rain from snow. *Water Resources Research*, *45*, W00D25. https://doi.org/10.1029/2008WR006995

Lundquist, J. D., & Loheide, S. P. (2011). How evaporative water losses vary between wet and dry water years as a function of elevation in the Sierra Nevada, California, and critical factors for modeling. *Water Resources Research*, *47*, W00H09. https://doi.org/10.1029/2010WR010050

Lundquist, J. D., Roche, J. W., Forrester, H., Moore, C., Keenan, E., Perry, G., et al. (2016). Yosemite Hydroclimate Network: Distributed stream and atmospheric data for the Tuolumne River watershed and surroundings. *Water Resources Research*, *52*, 7478–7489. https://doi.org/10.1002/2016WR019261

Marks, D., Domingo, J., Susong, D., Link, T., & Garen, D. (1999). A spatially distributed energy balance snowmelt model for application in mountain basins. *Hydrological Processes*, *13*(February), 1935–1959. https://doi.org/10.1002/(SICI)1099-1085(199909)13:12/13<1935::AID-HYP868>3.0.CO;2-C

Marks, D., & Dozier, J. (1992). Climate and energy exchange at the snow surface in the Alpine Region of the Sierra Nevada: 2. Snow cover energy balance. *Water Resources. Research*, *28*(11), 3043–3054. https://doi.org/10.1029/92WR01483

Molotch, N. P., & Bales, R. C. (2005). Scaling snow observations from the point to the grid element: Implications for observation network design. *Water Resources Research*, *41*, W11421. https://doi.org/10.1029/2005WR004229

Painter, T. H., Berisford, D. F., Boardman, J. W., Bormann, K. J., Deems, J. S., Gehrke, F., 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 Environment*, *184*, 139–152. https://doi.org/10.1016/j.rse.2016.06.018

PRISM Climate Group (2015). *Descriptions of PRISM spatial climate datasets for the conterminous United States*.

Renard, B., Kavetski, D., Leblois, E., Thyer, M., Kuczera, G., & Franks, S. W. (2011). Toward a reliable decomposition of predictive uncertainty in hydrological modeling: Characterizing rainfall errors using conditional simulation. *Water Resources Research*, *47*, W11516. https://doi.org/10.1029/2011WR010643

Rice, R., Bales, R. C., Painter, T. H., & Dozier, J. (2011). Snow water equivalent along elevation gradients in the Merced and Tuolumne River basins of the Sierra Nevada. *Water Resources Research*, *47*, W08515. https://doi.org/10.1029/2010WR009278

Rose, K., Graham, R., & Parker, D. (2003). Water source utilization by Pinus jeffreyi and Arctostaphylos patula on thin soils over bedrock. *Oecologia*, *134*(1), 46–54. https://doi.org/10.1007/s00442-002-1084-4

Sayama, T., Mcdonnell, J. J., Dhakal, A., & Sullivan, K. (2011). How much water can a watershed store? *Hydrological Processes*, *25*(25), 3899–3908. https://doi.org/10.1002/hyp.8288

Stock, G., Anderson, R. S., Painter, T. H., Henn, B., & Lundquist, J. D. (2017). Impending loss of little ice age glaciers in Yosemite National Park. *Geological Society of America Abstracts with Programs*, *49*(6). https://doi.org/10.1130/abs/2017AM-299617

United States Department of Agriculture Natural Resources Conservation Service (2007). *Soil survey of Yosemite National Park, California*. Retrieved from http://soils.usda.gov/surve/printed_survey/

Wayand, N. E., Hamlet, A. F., Hughes, M., Feld, S. I., & Lundquist, J. D. (2013). Intercomparison of meteorological forcing data from empirical and mesoscale model sources in the North Fork American River Basin in Northern Sierra Nevada, California. *Journal of Hydrometeorology*, *14*(3), 677–699. https://doi.org/10.1175/JHM-D-12-0102.1

Erratum

The originally-published version of this article included two errors in the way the author list and affiliations were displayed, and the abstract erroneously included a separate statement of research significance. These errors have been corrected, and this may be considered the official version of record.