A Particle Batch Smoother Approach to Snow Water Equivalent Estimation

STEVEN A. MARGULIS

Department of Civil and Environmental Engineering, University of California, Los Angeles, Los Angeles, California

MANUELA GIROTTO

NASA Goddard Space Flight Center, Greenbelt, and GESTAR, Universities Space Research Association, Columbia, Maryland

GONZALO CORTÉS

Department of Civil and Environmental Engineering, University of California, Los Angeles, Los Angeles, California

MICHAEL DURAND

School of Earth Sciences and Byrd Polar Research Center, The Ohio State University, Columbus, Ohio

(Manuscript received 12 September 2014, in final form 27 March 2015)

ABSTRACT

This paper presents a newly proposed data assimilation method for historical snow water equivalent SWE estimation using remotely sensed fractional snow-covered area fSCA. The newly proposed approach consists of a particle batch smoother (PBS), which is compared to a previously applied Kalman-based ensemble batch smoother (EnBS) approach. The methods were applied over the 27-yr *Landsat 5* record at snow pillow and snow course in situ verification sites in the American River basin in the Sierra Nevada (United States). This basin is more densely vegetated and thus more challenging for SWE estimation than the previous applications of the EnBS. Both data assimilation methods provided significant improvement over the prior (modeling only) estimates, with both able to significantly reduce prior SWE biases. The prior RMSE values at the snow pillow and snow course sites were reduced by 68%–82% and 60%–68%, respectively, when applying the data assimilation methods. This result is encouraging for a basin like the American where the moderate to high forest cover will necessarily obscure more of the snow-covered ground surface than in previously examined, less-vegetated basins. The PBS generally outperformed the EnBS: for snow pillows the PBS RMSE was ~54% of that seen in the EnBS, while for snow courses the PBS RMSE was ~79% of the EnBS. Sensitivity tests show relative insensitivity for both the PBS and EnBS results to ensemble size and fSCA measurement error, but a higher sensitivity for the EnBS to the mean prior precipitation input, especially in the case where significant prior biases exist.

1. Introduction and background

Snow-dominated hydrologic systems are globally significant given the estimate that over one-sixth of the world's population derive the majority of their water resources from basins containing seasonal snowmelt (Barnett et al. 2005). Much of the water supply that comes from snow-dominated basins is stored in regional-scale montane systems, including those in the western United States (including Alaska), Alps, Andes, Himalayas, and Hindu Kush, among others. In many regions, downstream urban and agricultural areas rely almost exclusively on snowmelt runoff from mountainous areas for their water supply (Viviroli et al. 2007).

In mountainous regions, highly variable spatial patterns of snow cover are the result of complex montane topography and orographic effects and equally complex atmospheric circulation patterns (e.g., Dettinger et al. 2004; Lundquist et al. 2010). Snowfall events are often the result of large-scale synoptic meteorological conditions, but heterogeneity in surface conditions are due to a complex mosaic of factors:

Corresponding author address: Steve Margulis, Department of Civil and Environmental Engineering, University of California, Los Angeles, 5732D Boelter Hall, Los Angeles, CA 90095.
E-mail: margulis@seas.ucla.edu

DOI: 10.1175/JHM-D-14-0177.1

1) Variations in elevation contribute directly to snow water equivalent SWE accumulation variations as a result of orographic effects.
2) Vegetation patterns contribute to variability in snowfall interception losses, radiative fluxes, and exposure, which can impact wind-driven redistribution.
3) Slope, aspect, and shading by terrain variability directly impact the incoming shortwave radiation, which causes variability in melt rates and therefore melt-out dates.
4) The albedo of snow, and its dynamics as a result of snow metamorphism under varying meteorological conditions, can further enhance spatial variability in both SWE and the subsequent soil moisture during and after melt.
5) Sloped terrain can influence variability via gravitational redistribution.

Beyond the water–energy cycle in mountainous regions, ecological systems and biogeochemical processes are highly sensitive to variability in the hydrologic cycle (Bales et al. 2006; Trujillo et al. 2012).

Despite the importance of montane snowpacks, fundamental questions remain unanswered to a large degree (Bales et al. 2006; Dozier 2011), including 1) how much water is stored in montane snowpacks, 2) how does this amount vary in space and time and how does it depend on underlying physiographic characteristics, and 3) how is the snowpack changing in time? The lack of insight into these fundamental questions stems primarily from gaps in the available datasets in montane regions. Operational in situ networks in the western United States consist primarily of snow pillow and snow course data. Despite the relatively high sampling density compared to other regions of the globe, these in situ measurements invariably sample only a very small fraction of the snow-covered area in montane regions. For example, in the Sierra Nevada the number of snow pillow sensors yield an average sampling density of one sensor for every 620 km^2 (Guan et al. 2013). Moreover, the in situ sensors generally do not provide a well-distributed sampling versus elevation and other physiographic characteristics, leaving much of the high-elevation regions completely unsampled. This sampling problem is generally exacerbated in drought years when the preponderance of snow may occur at high elevations above the measurement network (e.g., Rice et al. 2011). For all of the above reasons, interpolation of operational in situ data to create spatially continuous snow estimates is difficult and fraught with uncertainty (e.g., Fassnacht et al. 2003; Dozier 2011), making answering scientific questions that have to do with SWE spatial variability difficult. Moreover, in other mountainous regions of the globe, in situ data tend to be even scarcer and more sparsely distributed.

Remote sensing data provide a distinctly different information stream with unique mapping capabilities that provide insight into spatiotemporal snow cover patterns across the relevant physiographic gradients. With respect to snow remote sensing, two satellite-based methods have historically provided the primary mechanisms for characterizing snow. Satellite-based visible and near-infrared (Vis/NIR) measurements provide the ability to map snow-covered area SCA globally at relatively high resolution (30–500 m depending on the sensor; e.g., Rosenthal and Dozier 1996, Hall et al. 2002; Dozier et al. 2008; Painter et al. 2009; Cortés et al. 2014). The direct benefit of such data streams lies primarily with their spatial resolution, since the key modes of spatial variability in snow processes in montane regions are on the order of 100 m (e.g., Clark et al. 2011), while the primary drawback is that the resulting SCA estimates do not provide direct information on SWE. Satellite-based passive microwave (PM) data provide a much more direct link to SWE but are only typically available at resolutions greater than 10 km. This makes estimating SWE in montane regions from PM data difficult because of subgrid variability in snow and vegetation coverage. Efforts to improve retrievals and other estimation frameworks using PM data are an active area of ongoing research (e.g., Kelly et al. 2003; Durand and Margulis 2007; Tedesco et al. 2010; Li et al. 2012; Vander Jagt et al. 2013).

Because of the aforementioned challenges, spatially and temporally continuous estimates of SWE generally require some level of modeling. Snow models have the added benefit of embedding physical laws (i.e., mass and energy balance) and implicit relationships with readily available auxiliary data (topography, gridded meteorological datasets, etc.) into the SWE estimates. That said, uncertainty in these inputs directly propagates into SWE estimate uncertainty. In addition to such forward modeling approaches, which may only use in situ data for calibration or validation, other methods attempt to merge models and measurements to take advantage of information content in both. One such example is the so-called SWE reconstruction approach (e.g., Cline et al. 1998; Molotch et al. 2004; Molotch and Bales 2005; Molotch and Margulis 2008; Molotch 2009; Li and Wang 2011; Rice et al. 2011; Jepsen et al. 2012) that effectively integrates snowmelt estimates backward in time during the ablation season via energy flux estimates scaled by fractional SCA fSCA to get mapped estimates of the peak SWE. One can pose the problem more generally as a (probabilistic) data assimilation framework, which has been shown to be a more robust alternative to traditionally applied deterministic SWE reconstruction (Girotto et al. 2014b). This probabilistic reconstruction was originally outlined in Durand et al. (2008) and Girotto et al. (2014a,b) and amounts to an SWE reanalysis based on precipitation data, meteorological data related to energy balance, remote sensing data, and the physics embedded in snow hydrologic models.

The objective of this paper is to present, apply, and verify a newly proposed data assimilation method for SWE reanalysis, which improves on the methods utilized in Durand et al. (2008) and Girotto et al. (2014a,b). The newly proposed approach is essentially a particle batch smoother (PBS) rather than a Kalman-based ensemble batch smoother (EnBS) and is expected to outperform the Kalman-type approach. Herein, the PBS approach is described in detail, contrasted with the EnBS, and applied in parallel to the EnBS framework. The basic questions to be addressed are 1) does the new PBS SWE reanalysis approach outperform the previously applied EnBS SWE reanalysis approach and 2) how do the sensitivities to input parameters (and therefore robustness to uncertain inputs) for the two methods compare? The goal in answering these questions is to assess what approach is better suited for large-scale applications where inputs are unknown or highly uncertain. The underlying hypothesis is that the PBS is more general and therefore should perform as well as or better than the EnBS approach.

2. Methods

a. Forward and measurement models

The general estimation (data assimilation) problem starts with models for the evolution of the state variables of interest and the measurement process. A general discrete-form state-space (or forward) model formulation can be written as

$$\mathbf{y}_t = \mathbf{a}(\mathbf{y}_{t-1}, \mathbf{u}_t, \boldsymbol{\alpha}); \quad \text{subject to: } \mathbf{y}_{t=0} = \mathbf{y}_0, \tag{1}$$

where $\mathbf{y}_t$ is the system state vector at time t (with an uncertain initial condition $\mathbf{y}_0$), $\mathbf{u}_t$ is a vector of uncertain time-varying inputs, $\boldsymbol{\alpha}$ is a vector of uncertain time-invariant model parameters, and $\mathbf{a}$ is a nonlinear space–time discretized vector model operator. In the context of SWE reanalysis described herein, the state vector is the spatially distributed map of SWE; the time-varying inputs consist of meteorological forcings (radiation, air temperature, precipitation, etc.); the parameters $\boldsymbol{\alpha}$ include a multiplicative precipitation coefficient, the magnitude of subgrid spatial SWE variability, and an albedo decay parameter; and the model operator $\mathbf{a}$ represents the land surface model (LSM) used to predict SWE forward in time. The measurement process is generally written as

$$\mathbf{z}_t = \mathbf{m}(\mathbf{y}_t) + \mathbf{v}_t, \tag{2}$$

where $\mathbf{z}_t$ is the measurement at time t; $\mathbf{m}$ is the measurement operator or model; and $\mathbf{v}_t$ represents an additive measurement error vector, which is assumed to have known error characteristics (i.e., Gaussian with prescribed mean and covariance). In the context of the SWE reanalysis described herein, $\mathbf{z}_t$ is a vector of fSCA measurements and $\mathbf{m}$ represents a nonlinear snow depletion curve (SDC) model (e.g., Liston 2004) that predicts fSCA based on the modeled SWE (and other inputs). The specific forward and measurement models used in this work are described in more detail in section 2d.

b. Bayes's theorem as the basis for data assimilation

It is generally assumed in a data assimilation framework that the uncertain variables ($\mathbf{y}_0$, $\mathbf{u}_t$, $\boldsymbol{\alpha}$, and $\mathbf{v}_t$) can be characterized by postulated prior probability density functions (PDFs) that represent the mean estimate and uncertainty in each quantity. Based on the prior PDFs and the model operator $\mathbf{a}$, one can theoretically define the prior state PDF $p_{\mathbf{y}}(\mathbf{y}_t)$, which provides a full characterization of the state prior to the assimilation of any measurements. The goal of any data assimilation framework is to estimate the state vector conditioned on the measurement vector, which in a probabilistic framework is fully described by the conditional (or posterior) PDF (e.g., Gelman et al. 2004): $p_{\mathbf{y}|\mathbf{z}}(\mathbf{y}_t \,|\, \mathbf{z}_t)$. As described in more detail in Girotto et al. (2014a,b), the SWE reanalysis problem using assimilated fSCA measurements is best suited for a batch estimation (or smoother) approach rather than a sequential (or filtering) approach. This amounts to solving the forward model [Eq. (1)] over the full seasonal cycle and then assimilating all fSCA measurements at once (i.e., in one batch) to estimate SWE. The reason for this has to do with the fact that, in most cases, an instantaneous fSCA measurement has little information on the instantaneous SWE, but the collection of fSCA measurements and their temporal evolution, most notably during the ablation season, are in many cases highly correlated with the amount of accumulated SWE and the melt-season energy input to the snowpack. This is the primary basis of the deterministic SWE reconstruction approaches mentioned above, which can be conceptualized as special cases of the more general SWE reanalysis via a data assimilation technique (Girotto et al. 2014b).

In a batch smoother context, real-time estimates are not provided as they are in a filtering context, but otherwise the goals of the assimilation are similar. For batch smoothing, the relevant vectors at discrete times can simply be collected into larger vectors, that is, $\mathbf{Y}$, $\mathbf{Z}$, and $\mathbf{M}$ representing the state, measurement, and predicted measurement vectors, respectively. The uppercase notation is used to represent the collection of vectors over the whole seasonal cycle and the t subscript is dropped since this includes all states across the assimilation window, not just those at a particular time. As mentioned above, the goal of the estimation problem is to characterize the conditional PDF, which in the smoother case is given by $p_{\mathbf{Y}|\mathbf{Z}}(\mathbf{Y}\,|\,\mathbf{Z})$. Formally, this PDF is given by Bayes's theorem (e.g., Gelman et al. 2004):

$$p_{\mathbf{Y}|\mathbf{Z}}(\mathbf{Y}\,|\,\mathbf{Z}) = \frac{p_{\mathbf{Z}|\mathbf{Y}}(\mathbf{Z}\,|\,\mathbf{Y})p_{\mathbf{Y}}(\mathbf{Y})}{p_{\mathbf{Z}}(\mathbf{Z})} = c_0 p_{\mathbf{Z}|\mathbf{Y}}(\mathbf{Z}\,|\,\mathbf{Y})p_{\mathbf{Y}}(\mathbf{Y}), \tag{3}$$

where $p_{\mathbf{Z}|\mathbf{Y}}(\mathbf{Z}\,|\,\mathbf{Y})$ is the likelihood function for the state $\mathbf{Y}$, $p_{\mathbf{Y}}(\mathbf{Y})$ is the prior state PDF, and c_0 is simply a normalization constant to ensure that the conditional PDF integrates to one.

c. Ensemble methods

For practical reasons, prior and posterior PDFs are often approximated using ensemble-based methods, where an ensemble of model inputs is randomly sampled from the prior PDFs:

$$\begin{aligned} &p_{\mathbf{y}_0}(\mathbf{y}_0) \rightarrow \mathbf{y}_{0,j}; \quad p_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) \rightarrow \boldsymbol{\alpha}_j; \\ &p_{\mathbf{u}}(\mathbf{u}_t) \rightarrow \mathbf{u}_{t,j}; \quad \text{for} \quad j = 1, \ldots, N, \end{aligned} \tag{4}$$

where j represents an individual ensemble member (or replicate) out of an ensemble of chosen size N. The ensemble of prior inputs can then be used in the forward and measurement models, that is,

$$\mathbf{y}_{j,t} = \mathbf{a}(\mathbf{u}_{j,t}, \boldsymbol{\alpha}_j, \mathbf{y}_{j,t-1}) \Rightarrow \mathbf{Y}_j \quad \text{and} \quad \mathbf{m}(\mathbf{y}_{j,t}) \Rightarrow \mathbf{M}_j. \tag{5}$$

The end result of this Monte Carlo approach is a prior (or open loop) ensemble estimate of SWE ($\mathbf{Y}_j$) and predicted fSCA ($\mathbf{M}_j$) over the full seasonal cycle. An ensemble method amounts to a discrete approximation of the continuous prior PDF (i.e., using Dirac delta function δ):

$$p_{\mathbf{Y}}(\mathbf{Y}^-) \approx \sum_{j=1}^{N} w_j^- \delta(\mathbf{Y} - \mathbf{Y}_j^-), \tag{6}$$

where the superscript minus sign is used to emphasize that the variable is a prior estimate. As discussed in Zhou et al. (2006), this discrete estimate of the prior PDF implicitly assigns an equal discrete probability or weight ($w_j^- = 1/N$) to each replicate. Such ensemble approaches provide significant flexibility in how uncertainties are handled and allows for the nonlinear propagation of uncertainties through the forward model. The posterior PDF can be approximated similarly as

$$p_{\mathbf{Y}|\mathbf{Z}}(\mathbf{Y}^+ = \mathbf{Y}\,|\,\mathbf{Z}) \approx \sum_{j=1}^{N} w_j^+ \delta(\mathbf{Y} - \mathbf{Y}_j^+), \tag{7}$$

where the superscript plus sign is used to emphasize that the variable is a posterior estimate. Different algorithms can be used to approximate this conditional PDF as described in more detail below.

Previous SWE reanalysis efforts have focused on introducing uncertainty in the key meteorological forcing variables (Durand et al. 2008; Girotto et al. 2014a,b), with a particular emphasis on precipitation, that is,

$$P_{j,t}^- = b_j^- P_{\mathrm{nom},t}, \tag{8}$$

where P_{nom} is a nominal precipitation input. The variable P_t can be thought of as one of the dynamic inputs in the forcing vector $\mathbf{u}$ and/or the variable b can be thought of as one of the parameters in the vector $\boldsymbol{\alpha}$. This formulation was proposed because of the typical high uncertainty in snowfall inputs in montane environments and because precipitation provides the first-order direct control on the accumulated SWE in the absence of significant snow redistribution mechanisms. Durand et al. (2008) and Girotto et al. (2014a,b) chose to directly estimate the uncertain b parameter (rather than the states directly) as a means for ultimately improving the prior estimate of SWE through conditioning on fSCA measurements. This was done using an EnBS approach where the ensemble of parameters is updated via

$$\log(b_j^+) = \log(b_j^-) + \mathbf{K}[(\mathbf{Z} + \mathbf{V}_j) - \mathbf{M}_j^-], \tag{9}$$

where $\mathbf{K}$ is the Kalman gain, which is computed from the sample covariances estimated from the ensemble:

$$\mathbf{K} = \mathbf{C}_{bM}(\mathbf{C}_M + \mathbf{C}_V)^{-1}. \tag{10}$$

Equation (9) conditions the parameter b on the measured fSCA values (i.e., $\mathbf{Z}$) over the full temporal assimilation window. A posterior simulation using the posterior ensemble estimate of b from Eq. (9) in the forward model [Eq. (1)] is used to obtain the posterior estimate for $\mathrm{SWE}_{j,t}^+$.

In such Kalman-type update approaches, each posterior replicate is still implicitly assigned an equal weight ($w_j = 1/N$; Zhou et al. 2006), but with a different value assigned to the state–parameter. In this context, one can describe the Bayesian estimation process schematically as

$$\text{Prior: } b_j^- \Rightarrow \mathbf{Y}_j^- = \text{SWE}_j^- \quad \text{with} \quad w_j^- = \frac{1}{N} \quad \text{and}$$
$$\text{Posterior: } b_j^+ \Rightarrow \mathbf{Y}_j^+ = \text{SWE}_j^+ \quad \text{with} \quad w_j^+ = \frac{1}{N}. \tag{11}$$

An important caveat is that the Kalman update is only optimal for Gaussian prior PDFs, which will only strictly be maintained in linear models with Gaussian inputs. Hence, the approach described above is generally suboptimal, yet has been shown to be a useful approach for providing improved posterior estimates of SWE by extracting information from fSCA measurements (Durand et al. 2008; Girotto et al. 2014a,b). The suboptimality comes from the implicit Kalman update assumption that the prior and measurement PDFs can be adequately approximated solely by their means and covariances, neglecting higher-order moments. The computational effort of the EnBS-based SWE reanalysis method described above consists of $2N$ (i.e., N prior plus N posterior) forward model simulations along with evaluation of the update in Eq. (9).

In this paper, we propose a new method for SWE reanalysis that is more general than the previously described EnBS approach. The new approach uses a PBS formulation to estimate the SWE state vector directly. Particle filters (e.g., Arulampalam et al. 2002; Moradkhani et al. 2005; Zhou et al. 2006) attempt to more directly approximate Bayes's theorem for the case where non-Gaussian PDFs and nonlinear models make the ensemble Kalman smoother suboptimal. Extending such methods to a batch smoother, the prior PDF $p_{\mathbf{Y}}(\mathbf{Y})$ in Eq. (3) is approximated in the same way as described above and shown in Eq. (7). This Monte Carlo step involves no loss of generality in that the prior PDFs $p_{\mathbf{y}_0}(\mathbf{y}_0)$, $p_{\mathbf{u}}(\mathbf{u}_t)$, and $p_{\boldsymbol{\alpha}}(\boldsymbol{\alpha})$ can be non-Gaussian and the forward model can be nonlinear yielding a non-Gaussian prior state PDF. Unlike the EnBS, however, the PBS specifies that the prior and posterior state replicates are the same and instead updates the analysis weights (probabilities). This can be formalized by substituting the prior and posterior approximations from Eqs. (6) and (7) into Bayes's theorem [Eq. (3)], which yields

$$\mathbf{Y}_j^+ = \mathbf{Y}_j^- \quad \text{and}$$
$$w_j^+ = c_0 p_{\mathbf{Z}|\mathbf{Y}}(\mathbf{Z}\,|\,\mathbf{Y}) w_j^- = \frac{c_0}{N} p_{\mathbf{Z}|\mathbf{Y}}(\mathbf{Z}\,|\,\mathbf{Y}), \tag{12}$$

where the likelihood function can be expressed as [for the additive error case assumed in Eq. (2)]

$$p_{\mathbf{Z}|\mathbf{Y}}(\mathbf{Z}\,|\,\mathbf{Y}) = p_{\mathbf{V}}(\mathbf{Z}-\mathbf{M}_j^-) = \frac{1}{\sqrt{(2\pi)^{N_{\text{obs}}}|\mathbf{C}_V|}} \times \exp[-0.5(\mathbf{Z}-\mathbf{M}_j^-)^{\mathrm{T}}\mathbf{C}_V^{-1}(\mathbf{Z}-\mathbf{M}_j^-)], \tag{13}$$

where $p_{\mathbf{V}}(\mathbf{V})$ is the specified (i.e., Gaussian) PDF for the measurement error vector $\mathbf{V}$ (typically assumed as zero mean with an error covariance $\mathbf{C}_V$) and N_{obs} is the number of observations in the assimilation window. Equations (12) and (13) provide the mechanism for approximating the posterior PDF, which can be described schematically as

$$\text{Prior: } \mathbf{Y}_j^- = \text{SWE}_j^- \quad \text{with} \quad w_j^- = \frac{1}{N} \quad \text{and}$$
$$\text{Posterior: } \mathbf{Y}_j^+ = \mathbf{Y}_j^- = \text{SWE}_j^- \quad \text{with}$$
$$w_j^+ = \frac{c_0}{N} p_{\mathbf{V}}(\mathbf{Z}-\mathbf{M}_j^-), \tag{14}$$

where the posterior estimate is obtained simply by computing updated probabilities (weights) from the prior information. To make the posterior a valid PDF, one can simply compute the normalization constant as $c_0 = N/\sum_j p_{\mathbf{V}}(\mathbf{Z}-\mathbf{M}_j^-)$. The process will more heavily weigh those replicates (particles) that are more likely (i.e., predictions closer to the observations) and reduce the weight for those that are less likely (i.e., predictions farther from the observations) based on the likelihood function.

The updated weights provide a discrete estimate of the posterior PDF, as shown in Eq. (7), which can be used to determine posterior statistics, for example, those representing the central tendency (e.g., mean and median) and dispersion [e.g., standard deviation and interquartile range (IQR)] of the estimate. Note that while the mean and standard deviation can be estimated directly, the median and IQR for the posterior PBS estimates are an approximation because of the discrete nature of the weights (probabilities) w_j^+. The posterior PBS median (at a given time) is determined by sorting the SWE replicates and summing up the corresponding weights to get the replicate with an integrated value closest to 50%. Similarly, the posterior PBS IQR limits (at a given time) are approximated by determining the replicates with the integrated probabilities closest to 25% and 75%.

It is important to note that the above implementation of the PBS estimates the SWE state directly (and the b parameter implicitly), while the EnBS estimates only the b parameter directly, where SWE is then estimated via a posterior forward model simulation. However, because the b parameter and SWE are highly correlated (results not shown), the implicit joint state–parameter estimate provided by the PBS provides limited additional benefit over the EnBS implementation; instead, the differences are primarily driven by information extraction from the fSCA measurements, as described in more detail below. With respect to computational demand, the above PBS implementation implies that no additional forward simulations need to be performed as shown in Eq. (11), as this articulation of the PBS is a "weak constraint" assimilation operation, while the EnBS is a "strong constraint" operation (Talagrand 2003). Some additional important comments should also be mentioned here. In particle filtering applications, a resampling scheme is often required after each conditioning step (e.g., sequential importance resampling; Moradkhani et al. 2005) to avoid degeneracy (collapse) of the posterior weights after several updates. In the PBS implementation proposed herein, the single update avoids the degeneracy problem in most cases. Additionally, in some applications the particle filter requires a larger ensemble size than the Kalman filter. In the specific SWE reanalysis application shown herein, an ensemble size that is large enough for the EnBS appears to generally be large enough for the particle smoother (PS). When true, the PBS approach proposed herein has about half of the computational expense of the EnBS implementation shown above (i.e., N forward model simulations compared to $2N$ simulations), along with less restrictive assumptions. Implementing the EnBS in a state estimation mode could eliminate this difference in computational expense.

d. *LSM–SDC used in this study*

The specific forward and measurement models used in this study are almost identical to those described in more detail in Girotto et al. (2014a,b; additional details provided therein). The LSM used was the Simplified Simple Biosphere model, version 3 (SSiB3; Xue et al. 2003), which contains a three-layer snow scheme along with vegetation canopy and soil submodules. Static model parameters are defined via lookup tables (LUT) based on vegetation and soil type in a given pixel. The model is forced by hourly meteorological inputs including precipitation, radiation, and reference-level air temperature, humidity, wind speed, and pressure.

The LSM is coupled to the Liston (2004) snow depletion curve model as described in Girotto et al. (2014b). The SDC serves as the primary component of the measurement model since it predicts the fSCA dynamics on the ground over time as a function of evolving snowpack properties. Using the physiographic and meteorological inputs for a given pixel, the bare soil and forested portions of the pixel are simulated separately, yielding separate fSCA and SWE estimates for each fraction of the pixel. Rather than make assumptions about the subcanopy snow cover (as commonly done using vegetation gap fraction approaches; e.g., Molotch and Margulis 2008), the predicted fSCA is taken from the bare soil (nonforested) simulation. Specifically, the bare soil fSCA is transformed to the at-sensor fSCA via

$$\mathbf{m}(\mathbf{y}_{j,t}) = (1 - f_{\mathrm{veg},t}^{\mathrm{TM}})\mathrm{fSCA}_{j,t}, \tag{15}$$

where $\mathrm{fSCA}_{j,t}$ is the predicted bare soil fSCA from the Liston SDC and $f_{\mathrm{veg},t}^{\mathrm{TM}}$ is the estimated forest fraction that obscures the ground from the Landsat sensor. Note that the measurement model is only applied at the times of Landsat fSCA measurements. The dynamical $f_{\mathrm{veg},t}^{\mathrm{TM}}$ retrieved from Landsat was used in Eq. (15) to maintain internal consistency between the vegetation cover and fSCA at each measurement time. Although the retrieved vegetation fraction from Landsat can include shrubs and other "nonobscuring" types of vegetation (when not buried by snow), over this region it is expected that most (if not all) of the vegetation seen by Landsat during ablation consists of mostly trees and forested areas.

3. Study site, data, and experimental design

a. *Study site characteristics*

The American River watershed located in the northern Sierra Nevada of California (United States) shown in Fig. 1 (left; total drainage area of 4819 km^2) was chosen as the test bed for this study. The choice was made because of its representativeness with respect to other northern Sierra Nevada basins, because of the relatively large amount of in situ data that are used for verification in this study, and because the basin has a higher forest cover fraction than in the previous EnBS applications shown in Girotto et al. (2014a,b). It is expected a priori that more forest cover will make for a more challenging estimation of SWE from fSCA data.

The American River watershed drains mostly from east to west and feeds the Folsom Lake reservoir (~1.4 cubic kilometers capacity), which is one of the large reservoirs that make up the California (CA) water system. The runoff from the watershed is driven largely by snowmelt. The watershed area above the nominal rain–snow transition of 1500 m above mean sea level (MSL; Rice et al. 2011) is 2124 km^2 and ranges up to elevations of almost 3100 m MSL (Fig. 1). The area above the rain–snow transition in the American watershed contains six land-cover types according to the National Land Cover Database (NLCD): open water (~1.0%), developed

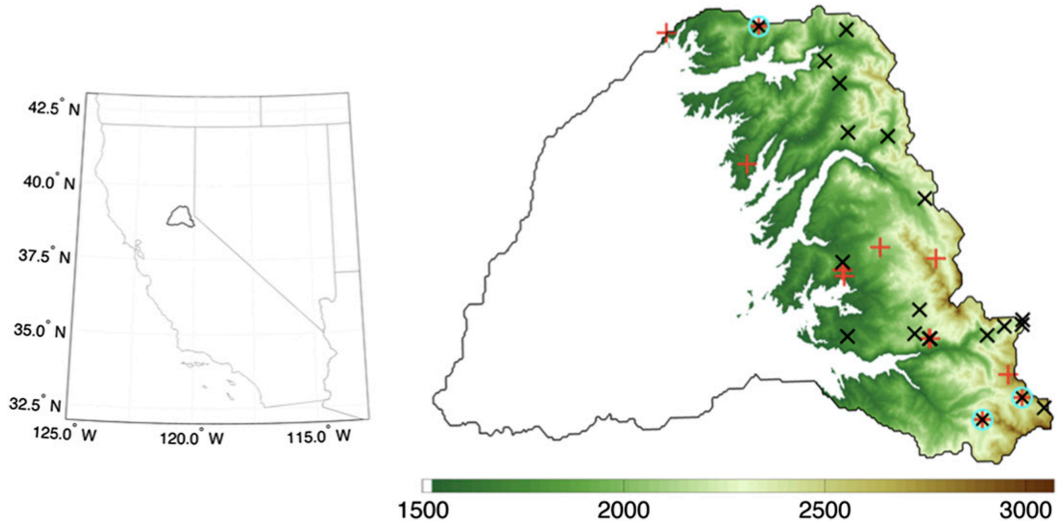

FIG. 1. (left) Site map showing the location of the American River basin (black outline) in the Sierra Nevada, CA. (right) DEM for the American River basin showing the elevation distribution above 1500 m MSL. In situ snow pillow and snow course sites are shown using red plus signs and black crosses, respectively. Four of the sites have collocated pillows and courses as described in Table 1. The three sites highlighted by the cyan circles are the locations of the stations with illustrative results shown in Figs. 3–8.

areas (~0.7%), bare soil (~2.0%), forest (including evergreen, deciduous, and mixed; ~67.3%), short vegetation (including shrubs and grass; 28.8%), and wetlands (~0.1%). Fractional forest coverage ranges from 0% to above 90% across the basin, with forest fraction generally decreasing with increasing elevation.

b. *Reanalysis input data*

1) In situ data for verification

The American River watershed contains 28 in situ monitoring locations with direct measurements of SWE, as shown in Fig. 1. All results and analyses presented herein are at these specific sites. The in situ network is a mix of 12 telemetered snow pillow sites that provide daily SWE estimates and 20 snow course sites that provide SWE estimates around the first of the month, typically during the months of January–May. Four of the sites have collocated pillows and courses. All data were acquired from the CA Department of Water Resources Data Exchange Center (CDEC; http://cdec.water.ca.gov/). The physiographic characteristics corresponding to the model pixel centered on the provided coordinates for each in situ location are shown in Table 1. Figure 2 shows the comparison of the distribution of elevation, land-cover type, and forest cover fraction of the pixels centered on the in situ sites relative to the full domain (i.e., above 1500 m MSL). The in situ sites have a comparable physiographic distribution to that of the full domain except that they do not sample the highest (or lowest) elevations (Fig. 2, top) or the highest forest cover fractions (above 80% forest cover; Fig. 2, bottom). It should be noted that the snow pillows themselves are often sited in flat clearings even when in a forested region, which invariably leads to potential representativeness issues when comparing with a pixel-averaged SWE estimate.

2) Model input data

The LSM–SDC applied in this study (described in section 2d), uses the same input data and is run at the same spatial resolution (90 m) as in Girotto et al. (2014a,b). Hence, for brevity, only the key points are repeated here. The static datasets needed by the forward model, which include topography, land-cover type, and forest cover fraction, are derived from the ASTER global DEM (http://asterweb.jpl.nasa.gov) and the National Land Cover Database (Homer et al. 2007). The data shown in Figs. 1 and 2 come from these sources. Ancillary inputs like slope, aspect, shading, and sky-view factor are derived from the ASTER DEM.

The dynamic meteorological inputs were taken from phase 2 of the North American Land Data Assimilation System (NLDAS-2) dataset (Xia et al. 2012), which is available over the United States from 1979 to present. The coarse-scale (1/8°) inputs are downscaled using similar techniques as those described in Girotto et al. (2014a,b), which make use of topographic corrections to estimate the hourly meteorological forcings at a given pixel. The primary differences have to do with the

TABLE 1. List of stations (sorted by NLCD fractional forest cover) with in situ snow pillows and/or snow courses in the American River basin and their corresponding key physiographic characteristics. Stations shown in boldface are those highlighted for illustrative purposes in Figs. 3–8. (Full names and information about each station are available online: http://cdec.water.ca.gov/cgi-progs/staSearch.)

CDEC station identification	Pillow and/or course	NLCD f_{veg} (%)	Elev (m)
LOS	Pillow	0	2538
WRG	Course	5	2115
VVL	Pillow	8	2042
SIL	**Pillow and course**	**14**	**2168**
ONN	Course	16	1859
ABN	Course	28	2224
FRN	Pillow	29	2301
ALP/APH	Pillow and course	36	2307
DMN	Course	40	1846
SCN	Pillow	44	2668
BLC	Pillow	45	1593
RBB	Pillow	45	1798
RBP	Pillow	46	1578
ECS	Course	47	2270
LCP	Course	49	2570
WBM	Course	50	1919
MCB	Course	50	1889
CAP	**Pillow and course**	**52**	**2439**
PHL	Course	60	2093
LYN	Course	60	2039
CRF	Course	65	2223
IHS	Course	65	1612
TMF	Course	70	1999
TBC	Course	71	1753
LCF	Course	72	2283
GKS	Pillow	73	1707
HYS	**Pillow and course**	**76**	**2028**
RBV	Course	77	1708

long- and shortwave parameterization. The new longwave model used is the clear-sky model of Prata (1996) with a cloudy-sky correction based on a solar index as described in Crawford and Duchon (1999). The choice of this model was based on the results shown in Flerchinger et al. (2009), which compared a multitude of models from sites across the globe and found this combination to be among the best. As part of the downscaling procedure [appendix A in Girotto et al. (2014b)], uncertainty is added to the inputs to generate an ensemble of meteorological forcings as indicated schematically in Eq. (4). For consistency with the longwave model, the shortwave uncertainty and bias-correction model was rederived based on the solar index (instead of relative humidity). The use of these static and dynamic variables as inputs to the LSM–SDC generate prior estimates of SWE and fSCA needed in the reanalysis [i.e., $\mathbf{Y}_j^-$ and $\mathbf{M}_j^-$, respectively, in Eq. (14)].

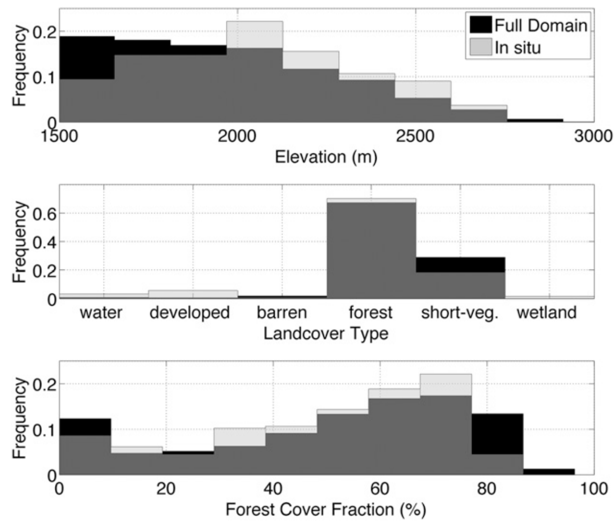

FIG. 2. Relative frequency plots showing the distribution above 1500 m MSL of (top) elevation, (middle) land-cover type, and (bottom) forest cover fraction in the full domain (black) or at the in situ snow pillow and course sites (light gray). The dark gray color represents the overlap between the full domain and in situ distribution.

3) fSCA INPUT DATA

The reanalysis methods described above condition prior SWE estimates on remotely sensed fSCA estimates. The retrieved fSCA estimates used in this study are derived from *Landsat 5* Thematic Mapper (TM) data (covering 1985–2011) as described in Girotto et al. (2014a) and using the methodology shown in Cortés et al. (2014). Specifically, estimates at a given satellite overpass time are estimated from Landsat reflectances using a spectral end-member unmixing approach. This results in mapped *Landsat 5* TM estimates of fractional snow-covered area $\mathrm{fSCA}^{\mathrm{TM}}$ and fractional vegetation cover f_{veg}^{TM} at each overpass time. The pixel-wise fSCA values are screened to include only reliable (cloud free) measurements during a window defined by a specified number of days (60 days) before the peak SWE in the prior estimate going forward through the end of the water year (WY). This screening implicitly uses different windows for each pixel since the peak SWE is a function of elevation (and other factors). The screened $\mathrm{fSCA}^{\mathrm{TM}}$ data are collected in the measurement vector $\mathbf{Z}$ for assimilation.

c. Experimental design

The experimental setup for this study was designed to elucidate how well the newly proposed PBS SWE reanalysis method performs (compared to the open-loop and EnBS estimates) and to assess its sensitivity to key input parameters. The PBS and EnBS reanalysis methods were applied at each of the 28 in situ station sites for all available years spanning the *Landsat 5* record (1984–2011), which amounts to 27 water years (WY 1985–2011) tested. All told, the analysis described herein amounted to 756 station years, although a small fraction of stations are missing in situ data in some years (presumably because of pillow or telemetry failures or inability to perform the snow course surveys).

Because of potential geolocation errors in the in situ coordinates and potential representativeness issues in comparing in situ point-scale versus single-pixel results, simulations for the nine (90 m) model pixels surrounding each in situ location were performed. The motivation for looking at neighboring pixel estimates stems from inherent problems in comparing pixel-averaged estimates to in situ point-scale data. In the analysis described in sections 4a and 4b, the SWE estimates from the neighboring pixel with the smallest errors relative to the snow pillow or course data are used for computing the bulk error statistics (Rittger 2012; Girotto et al. 2014a). This is an attempt to hedge against representativeness and geolocation errors. The primary focus is on the comparison of error metrics between the two reanalysis methods.

Three ensemble-based estimates of SWE were generated for analysis: 1) an open-loop or prior estimate (i.e., those estimates prior to assimilation), 2) a posterior estimate using the newly proposed PBS method, and 3) a posterior estimate using the previously applied EnBS approach. Such a comparison allows for assessing 1) whether the fSCA data provide a benefit to the SWE estimate (i.e., do the posterior estimates outperform the prior estimate in a highly vegetated basin?) and 2) whether the PBS method outperforms the EnBS method. A nominal set of results is analyzed first, followed by sensitivity tests to key input parameters.

The prior ensemble simulations require specification of prior PDF input parameters. The perturbed parameters are the same as those in Girotto et al. (2014a) and include, most notably, the parameter b shown in Eq. (8) (assumed lognormally distributed), along with the subpixel SWE coefficient of variation β needed in the Liston SDC model (assumed uniformly distributed) and the snow albedo decay parameter C_{VIS} (assumed uniformly distributed). Nominal values for the parameters of these distributions are shown in Table 2. In particular, the prior mean value of the parameter b is specified as 2.5, which is based on previous literature estimates (Pan et al. 2003; DeLannoy et al. 2010), is the same value used in Girotto et al. (2014a), and indicates a prior belief in a general underestimation of snowfall in the NLDAS-2 precipitation data. The nominal SWE reanalysis estimates are obtained using an ensemble size of $N = 100$ replicates, which is the order of magnitude shown to perform well previously in the EnBS applications of Durand et al. (2008) and Girotto et al. (2014a,b). A nominal fSCA measurement error standard deviation of $\sigma_Z = 15\%$ was used based on previous work (e.g., Cortés et al. 2014) and defines the measurement error covariance $\mathbf{C}_V$ used in Eqs. (9), (10), and (13), where it is assumed that measurement errors are independent in time.

TABLE 2. Key nominal input parameters used in the SWE reanalysis. The postulated distribution for b is lognormal, and the postulated distributions for $\beta_{\min}$ and C_{VIS} are uniform.

Parameter name	Variable	Value
Ensemble size	N	100
Prior b mean	μ_b	2.5
Prior b coefficient of variation	CV_b	0.25
Prior b max value	$b_{\max}$	5
Prior β min	$\beta_{\min}$	0.05
Prior β max	$\beta_{\max}$	0.80
Prior C_{VIS} min	$C_{\mathrm{VIS}_{\min}}$	0.20
Prior C_{VIS} max	$C_{\mathrm{VIS}_{\max}}$	0.45
fSCA measurement error std dev	σ_Z	15%

To assess the robustness of the nominal case SWE estimates, sensitivity tests were also performed with respect to key input parameters. In the sensitivity tests the focus was on those parameters most closely connected to the data assimilation framework, with an emphasis on the ensemble size, the fSCA measurement error standard deviation, and the prior mean value of the coefficient b μ_b to understand how these inputs affect the two methods. The parameter N must be large enough to accurately represent the statistics (i.e., covariances) or likelihood needed in the EnBS or PBS, respectively, but larger values directly increase the computational expense of the assimilation methods. The parameter σ_Z dictates how much the fSCA measurements are trusted relative to the prior estimate of fSCA in the assimilation step. Finally, the parameter μ_b determines how much the nominal precipitation is scaled on average in the prior estimate. Ideally, the posterior would be insensitive to this prior parameter specification. The specific perturbation sensitivity experiment cases examined were with 1) ensembles of size $N = 50$ and 200, 2) a reduction in fSCA measurement error standard deviation of $\sigma_Z = 10\%$ [used in Girotto et al. (2014a,b)], and 3) $\mu_b = 2.0$ and 3.0. The set of sensitivity experiments are identified in Table 3.

4. Results and discussion

a. *Illustrative nominal results for individual in situ stations*

For illustration, representative individual station-year results are presented, which allow for an understanding

TABLE 3. Identification of sensitivity experiments and their corresponding perturbed parameters. The single perturbed parameter and its value are shown while all other parameters are the same as the nominal set shown in Table 2.

Sensitivity expt name	Perturbed variable	Perturbed variable value
Case 0 (Nominal)	—	—
Case 1a	N	50
Case 1b	N	200
Case 2	σ_Z	10%
Case 3a	μ_β	2.0
Case 3b	$\sigma_{C_{\text{VIS}}}$	3.0

of how the methods work. To simultaneously show results and the typical similarity and disparity between snow pillow and snow course data, the illustrative results were taken from sites where snow pillows and snow courses are nominally collocated. Of the four sites with collocated pillows/courses, three were chosen for their forest cover fraction that span the range of forest cover seen in the basin (Table 1): station SIL ($f_{\text{veg}} = 14\%$; elevation: 2168 m MSL), station CAP ($f_{\text{veg}} = 52\%$; elevation: 2439 m MSL), and station HYS ($f_{\text{veg}} = 76\%$; elevation: 2028 m MSL). The driest year (WY 1988) and wettest year (WY 2011) over the 27-yr record examined are shown to illustrate the performance for different SWE conditions. The subsample of station years described above provides a representative sample of results across all 756 station years. Bulk results across all stations are presented in the following section.

Figure 3 illustrates the two reanalysis estimates relative to the prior estimate and observations for the SIL station in the dry WY 1988. The observations shown consist of the Landsat-derived fSCA data (which are assimilated) and the independent ground verification data from snow pillows and snow courses (which are not assimilated). Note that in the fSCA figures (top panels), the data shown are the Landsat-derived observations scaled by vegetation [i. e., $\text{fSCA}_t^{\text{TM}}/(1 - f_{\text{veg},t}^{\text{TM}})$] to make the measurements comparable to the bare ground SDC fSCA estimates; it is the $\text{fSCA}_t^{\text{TM}}$ data (not the scaled version) that are assimilated. This inversion of the Landsat data in the figures is done simply for visualization purposes to allow for a comparison with the continuous time series of the prior fSCA prediction.

The prior and posterior estimates are those before and after assimilation of the $\text{fSCA}_t^{\text{TM}}$ data. The prior estimate shows the ensemble median and IQR of predicted SWE and fSCA over the water year with prior SWE accumulation beginning near 1 December, peaking at just under 0.6 m in February, and melting out around 1 May (Fig. 3, bottom). The corresponding prior (on the ground) fSCA from the SDC model shows intermittent peaks early in the accumulation season, followed by fSCA quickly reaching 100% once the primary accumulation begins, where it remains at 100% until ~1 April and reduces to 0% by mid-May (followed by some intermittent events late in the season; Fig. 3, top). The observed SWE shows a smaller peak (~0.3 m) and an earlier melt-out date (~1 April) compared to prior (Fig. 3, bottom). It should be noted that there are some discrepancies between the in situ snow pillow and snow course data for the SIL station. This is not uncommon and is likely the result of representativeness and/or geolocation differences between the pillow site and course locations. Such differences should be kept in mind when comparing estimates to the observations, that is, it is possible to fit one and not the other in many cases.

The fSCA observations show an earlier melt than the prior (Fig. 3, top), which is consistent with the prior overestimation in SWE. It is precisely this signal that is used by the assimilation step to provide an improved posterior estimate. The number of available fSCA observations depends on the cloud conditions and generally results in 5–10 observations from before peak SWE through the end of the WY. The number of observations that effectively cover the ablation season is also dependent on the amount of peak SWE and the energy regime (i.e., cloudiness) during the ablation season (i.e., where an above-average cloudy ablation season will generally slow the melt process).

The assimilated $\text{fSCA}_t^{\text{TM}}$ data result in an updated ensemble SWE-fSCA estimate. This is where the PBS and EnBS posterior estimates can depart from each other. By construct, both methods aim to improve the fit with $\text{fSCA}_t^{\text{TM}}$ observations and thereby implicitly improve the SWE estimates. Both posterior fSCA estimates move toward the observations, with the PBS providing a slightly better fit in this case (Fig. 3, top). The PBS accomplishes this by more heavily weighting replicates that more closely match the observations [Eq. (12)], while the EnBS does so by updating the b coefficients using the Kalman update shown in Eq. (9).

With respect to SWE, both posterior estimates also move toward the independent observations, with the PBS providing a better estimate than the EnBS (Fig. 3, bottom). The PBS matches the pillow and course data well in terms of both the peak SWE and the melt-out date, while the EnBS improves over the prior, but still overestimates peak SWE and melt-out date. The IQR of the posterior estimates is reduced with respect to the prior. This is simply a reflection of the information content in the $\text{fSCA}_t^{\text{TM}}$ measurements propagating to the posterior estimates. For this station year, the posterior IQR for the PBS estimate is smaller than for the EnBS estimate,

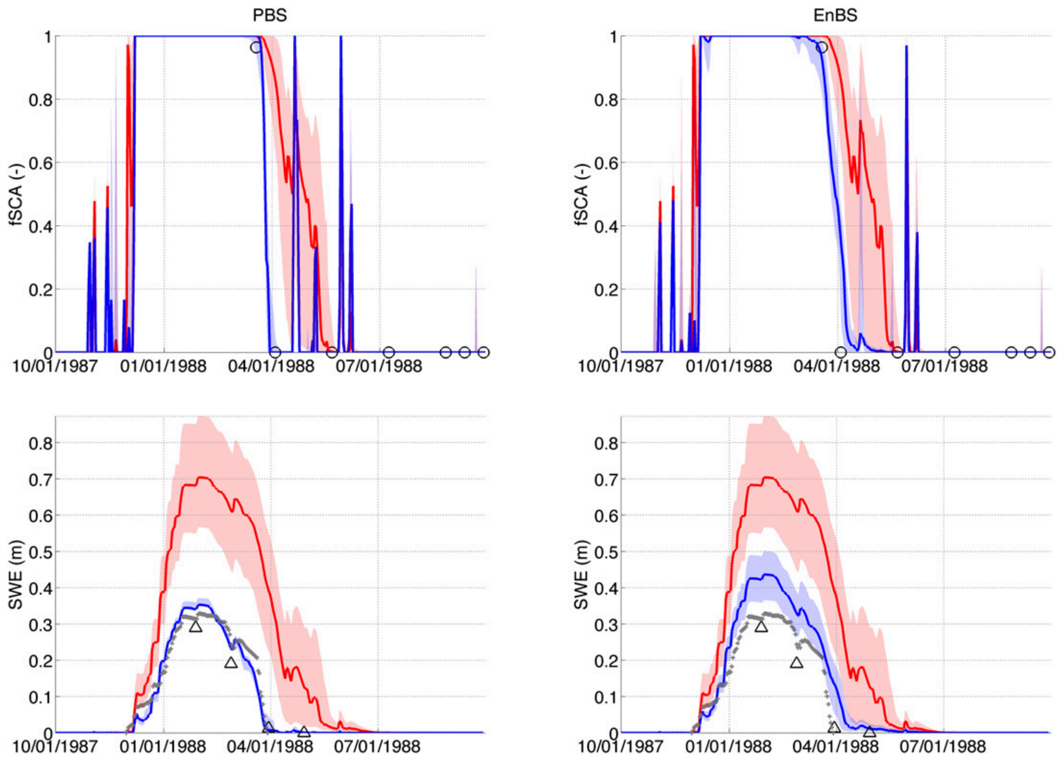

FIG. 3. Comparison of (left) PBS and (right) EnBS reanalysis results at a low forest cover (NLCD f_{veg} = 14%) station (SIL) during a dry year (WY 1988). (top) Prior (red) and posterior (blue) predictions of fSCA vs scaled fSCA observations; (bottom) the prior (red) and posterior (blue) predictions of SWE vs SWE observations from snow pillows (gray plus signs) and snow courses (black triangles). The solid lines are the ensemble median and the shaded region represents the ensemble IQR. Note that the fSCA observations (black circles) are Landsat derived and scaled by vegetation [$\mathrm{fSCA}_t^{\mathrm{TM}}/(1 - f_{\mathrm{veg},t}^{\mathrm{TM}})$] to make the measurements comparable to the bare ground fSCA predictions. This inversion of the Landsat data is for visualization purposes to generate a continuous time series of the prior fSCA prediction. The $\mathrm{fSCA}_t^{\mathrm{TM}}$ data (not the scaled data) are assimilated.

indicating more information is extracted from the $\mathrm{fSCA}_t^{\mathrm{TM}}$ measurements. Assuming the observed SWE is representative of the reanalysis estimate, the fact that the PBS results match both the independent fSCA and SWE measurements is indicative of a reasonable estimate of the ablation-season melt fluxes (at least for this pixel). It is important to keep in mind that both the PBS and EnBS will generally attempt to fit the fSCA data by construct; whether this yields an improved estimate of SWE depends mostly on the LSM–SDC accurately estimating the ablation-season melt fluxes, which are primarily dependent on disaggregated meteorological forcings and snow albedo.

Analogous estimates are shown in Figs. 4 and 5 for the CAP and HYS stations, respectively. Qualitatively, the results are similar in the sense that the prior overestimates the observed SWE and predicts a later melt out than the Landsat fSCA observations would suggest. For all three pixels, the EnBS posterior does not deviate as much from the prior as the PBS posterior does. While in some cases this could simply be a reflection of a good prior estimate, here it is indicative of a smaller (suboptimal) update. In all three pixels, the EnBS generally overestimates the peak SWE, while the PBS generally has smaller errors with respect to peak SWE. Note that both the PBS and EnBS estimates for melt out are later than those shown by the snow pillow data. This is consistent with the fact that the pillow data are point-scale estimates, while the posterior estimates are a pixel average, where subgrid SWE may generally melt out later in the ablation season. Also note that for the HYS station year (Fig. 5, bottom) there is a significant difference between the snow pillow and snow course SWE compared to SIL and CAP. This discrepancy and the smaller updates may be partly explained by the fact that the HYS pixel has the largest forest cover fraction (76%). As such, the actual Landsat $\mathrm{fSCA}_t^{\mathrm{TM}}$ data span a smaller range [i.e., between 0 and $(1 - f_{\mathrm{veg},t}^{\mathrm{TM}})$]. This necessarily makes the model–measurement misfit in Eqs. (9) and (14) smaller than for a nonforested pixel despite having

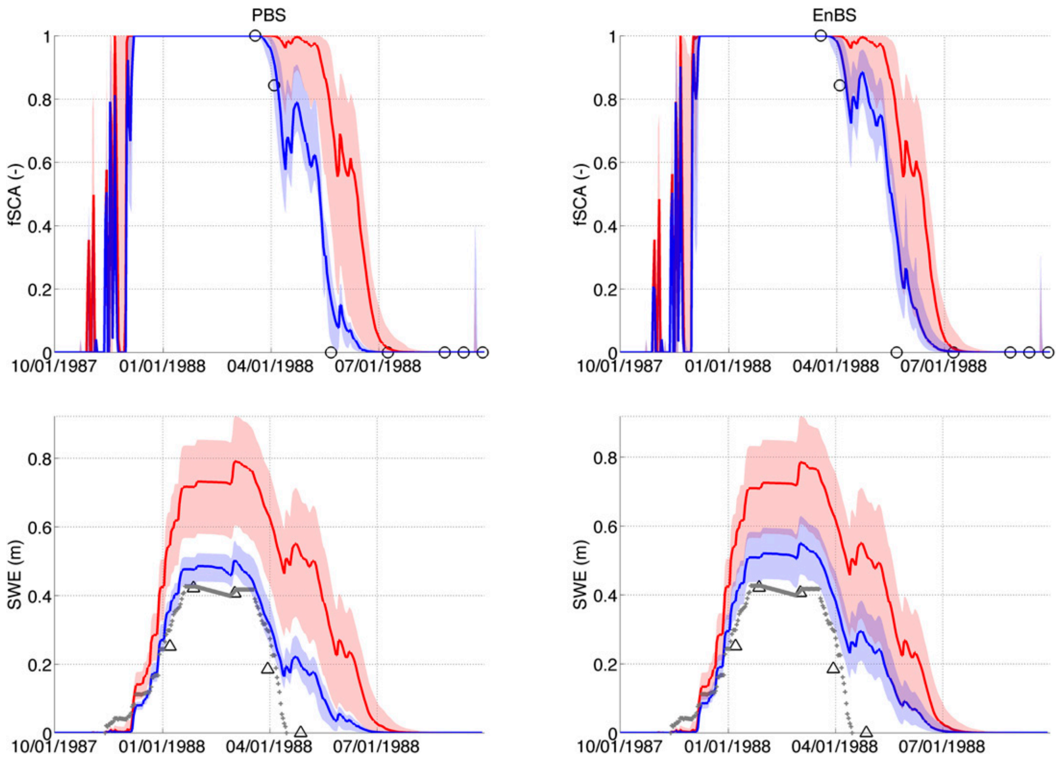

FIG. 4. As in Fig. 3, but for results at a medium forest cover (NLCD f_{veg} = 52%) station (CAP).

the same measurement error (i.e., lower signal-to-noise ratio). In general, the PBS performs better than the EnBS at removing prior biases; this is more evident (and discussed in more detail) in the context of the WY 2011 results shown below.

Results from the same three stations are shown in Figs. 6–8 for WY 2011, which was the wettest year of the 27-yr period. For all three stations, the peak SWE was 3–4 times larger in 2011 than it was in 1988. This directly leads to a large difference in the melt-out date, which is approximately 2–3 months later than it was in 1988. Note, however, that even in this wet year the prior overestimated the peak SWE (Figs. 6–8, bottom). This is indicative of a likely overestimate of μ_b. While the value of $\mu_b = 2.5$ is based on previously used values in the literature, it appears to yield an overestimate of peak SWE for these stations. It is also possible (if not likely) that there is interannual variability in b, which is not being represented in the prior by using a constant mean value across all years. Sensitivity to this parameter is explored more below.

The PBS estimates are able to overcome the prior bias more effectively than the EnBS in WY 2011. For all three stations, the fSCA observations during ablation are outside of the prior IQR (Figs. 6–8, top). The PBS is able to overcome this by using the replicates that pass near the observations and weighting them heavily, thereby fitting the observations within the expected measurement error (Figs. 6–8, top). The EnBS relies on covariances in the update [Eq. (10)], which may not be capable of providing a full update in cases with significant prior biases. Specifically, at times when most of the ensemble has predicted values of $\mathrm{fSCA}_{j,t} = 1$, the covariance in the ensemble [$\mathbf{C}_{bM}$ in Eq. (10)] at these times will be near zero. This is the case here where the bulk of the prior ensemble (as represented by the IQR) is showing values of $\mathrm{fSCA}_{j,t} = 1$ while the observations are showing ablation [e.g., Fig. 7 (top, right)]. So despite there being large model–measurement misfits $\{\text{i.e., } [(\mathbf{Z} + \mathbf{V}_j) - \mathbf{M}_j^-]\}$ with respect to some of the early ablation measurements, the contribution to the update [in Eq. (9)] by these measurements is small because of the small covariance terms in the Kalman gain. In contrast, the PBS uses the measurement–model misfit at these times in the likelihood function that contributes to the weight updates [Eq. (12)]. This represents a potential mechanism for the EnBS to have the undesirable property of having the posterior estimate be sensitive to the prior input parameters. In other words, if the prior is too far from the underlying true fSCA, the covariance needed for the EnBS may be inappropriately small leading to a smaller update. One approach for obtaining

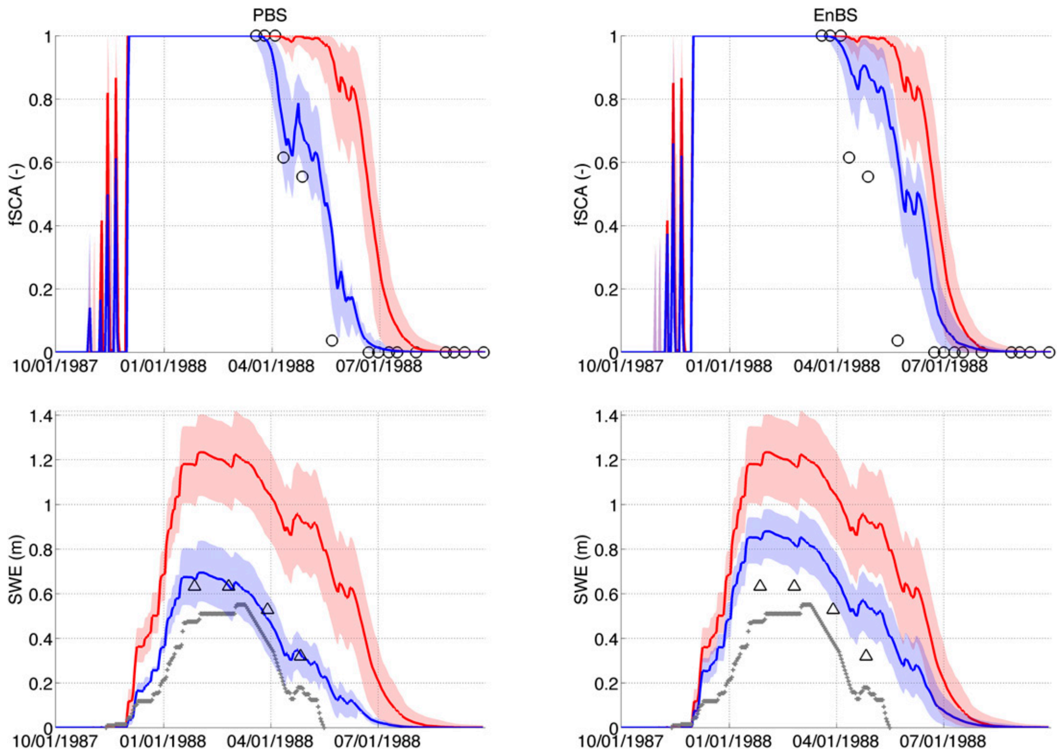

FIG. 5. As in Fig. 3, but for results at a high forest cover (NLCD f_{veg} = 76%) station (HYS).

better coverage of the observations by some ensemble members would be to increase the coefficient of variation of the prior *b* coefficients. The PBS can also have sensitivity to the prior, but in a potentially less damaging way. As long as some replicates cover the fSCA depletion record, there will be potential to extract that information from the measurement sequence. If the prior is so far biased (or with unrealistically low uncertainty) as to not cover the measurements, then the PBS may also suffer from sensitivity to the prior mean *b* parameter by most heavily weighting a small number of replicates at the extreme tail of the prior distribution that are closest to the measurements.

In the nominal case examined here, the PBS reanalysis shows almost no bias in peak SWE across the three sites, while the EnBS generally overestimates peak SWE (Figs. 6–8, bottom). The other primary distinction between the PBS and EnBS estimates is the ensemble spread, where the IQR for the PBS is generally smaller than for the EnBS. This is not inherently a positive trait, as it is possible that the spread could be unrealistically small. Given the relatively close fit between the observations and the PBS posterior estimates, the PBS posterior uncertainty seems reasonable. As in WY 1998, the posterior estimates show a later melt out when compared to the snow pillow data, which we hypothesize is due to representativeness issues between point-scale and pixel-averaged estimates.

b. Bulk nominal results for the full 27-yr reanalysis record

The bulk nominal case results for all station years are shown in Figs. 9 and 10. Figure 9 shows scatterplots of predicted SWE (both prior and posterior) for the PBS and EnBS relative to the observed snow pillow and snow course data. Specifically, all course data are used in the comparison and the peak SWE is compared to the corresponding snow pillow data at the time of predicted peak SWE. Note that the prior results are the same in both cases; only the posterior estimates differ between the PBS and EnBS.

The prior estimates show a systematic positive SWE bias relative to observations (Fig. 9, left), which is consistent with the individual station results discussed above. The IQR for the prior is generally large, as shown with the uncertainty bars (Fig. 9, left), because of the large unconditional uncertainty in model inputs. The assimilation of fSCA measurements can generally have two main impacts on the prior: shifting of the prior mean (i.e., reduction of bias) and/or reduction of the prior uncertainty. The posterior estimates for both the PBS and EnBS (Fig. 9, middle and right) generally show a

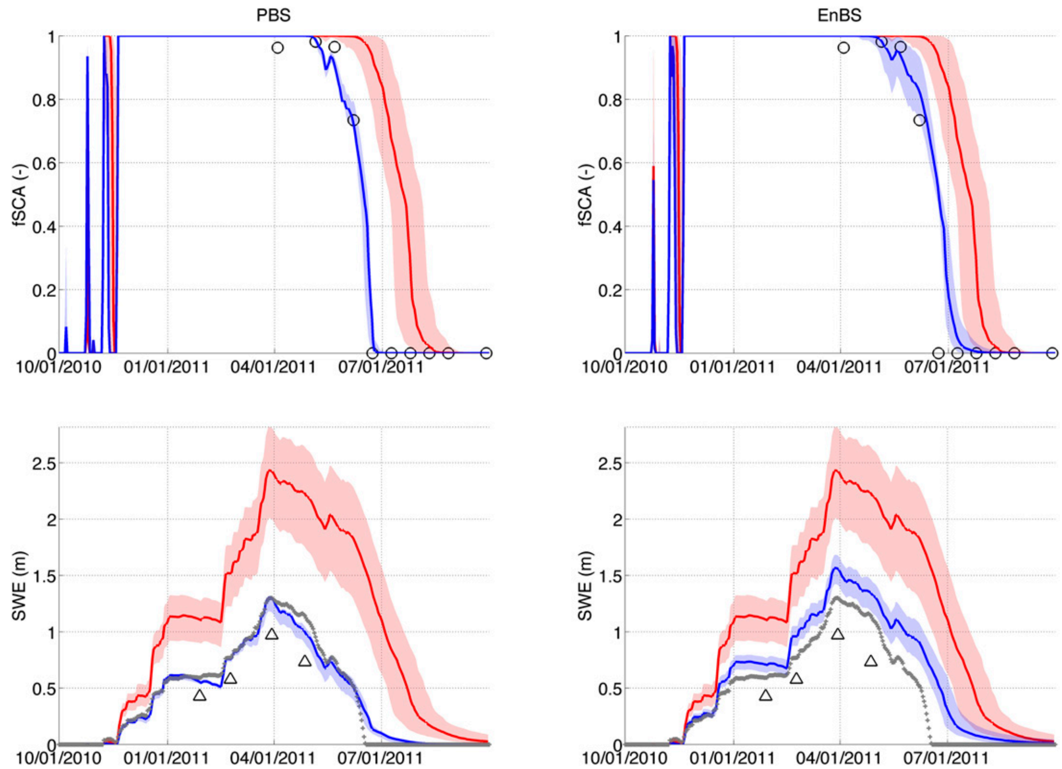

FIG. 6. As in Fig. 3, but for a wet year (WY 2011).

large improvement in bias relative to the prior. The PBS shows less scatter in the posterior estimate and the IQR values are reduced considerably relative to the prior (Fig. 9, middle). The EnBS also removes most of the prior bias but has higher IQR values compared to the PBS (Fig. 9, right).

Figure 10 shows the distribution of prior and posterior snow pillow and snow course errors. The prior snow pillow errors are biased (with individual errors ranging between −0.5 and 1.5 m), while the posterior PBS SWE errors are centered nearly symmetrically around zero and limited primarily to the range of −0.4 and 0.4 m [Fig. 10 (top, left)]. The prior snow course errors are less biased, but show a broader range of values compared to the snow pillow errors. The PBS posterior SWE snow course errors are also centered on zero, with most error values again limited to −0.4 and 0.4 m [Fig. 10 (bottom, left)]. The posterior EnBS results show SWE errors that are also centered near zero. The pillow errors are a bit more skewed toward positive values [Fig. 10 (top, right)], while the snow course errors are more symmetric about zero [Fig. 10 (bottom, right)]. For both snow pillow and snow course data, the range of EnBS errors is slightly larger than for the PBS errors. The skewness and bias in the EnBS posterior is largely due to the prior bias, which the EnBS appears to be more sensitive to. This is discussed in more detail in the context of the sensitivity experiment results below.

Table 4 shows bulk SWE error metrics for both reanalysis methods in terms of the mean error ME, root-mean-square error RMSE, and the correlation coefficient between estimates and measurements. The prior mean errors for the snow pillows and snow courses were 0.58 and 0.30 m, respectively, indicating the prior estimate was positively biased relative to the observations. The differences in prior ME values between the snow pillows and snow courses is also indicative of some inconsistency between the two observed SWE datasets. The prior RMSE values (0.74 and 0.47 m) indicate larger errors for the snow pillows compared to the snow courses, while the prior correlations were similar (0.82 and 0.81). The assimilation of fSCA data improved the prior estimate for both the PBS and EnBS for both snow pillows and snow courses in terms of all error and correlation metrics. The posterior ME for the PBS is small for both the pillows and courses (0.02 and −0.01 m) and is relatively small for the posterior EnBS estimates for the snow courses (0.05 m). The ME for the EnBS estimates compared to snow pillows is larger compared to the other cases (0.13 m). The posterior RMSE is reduced in all cases relative to the prior. In general, the PBS RMSE is less than the EnBS RMSE, that is, ~54% and

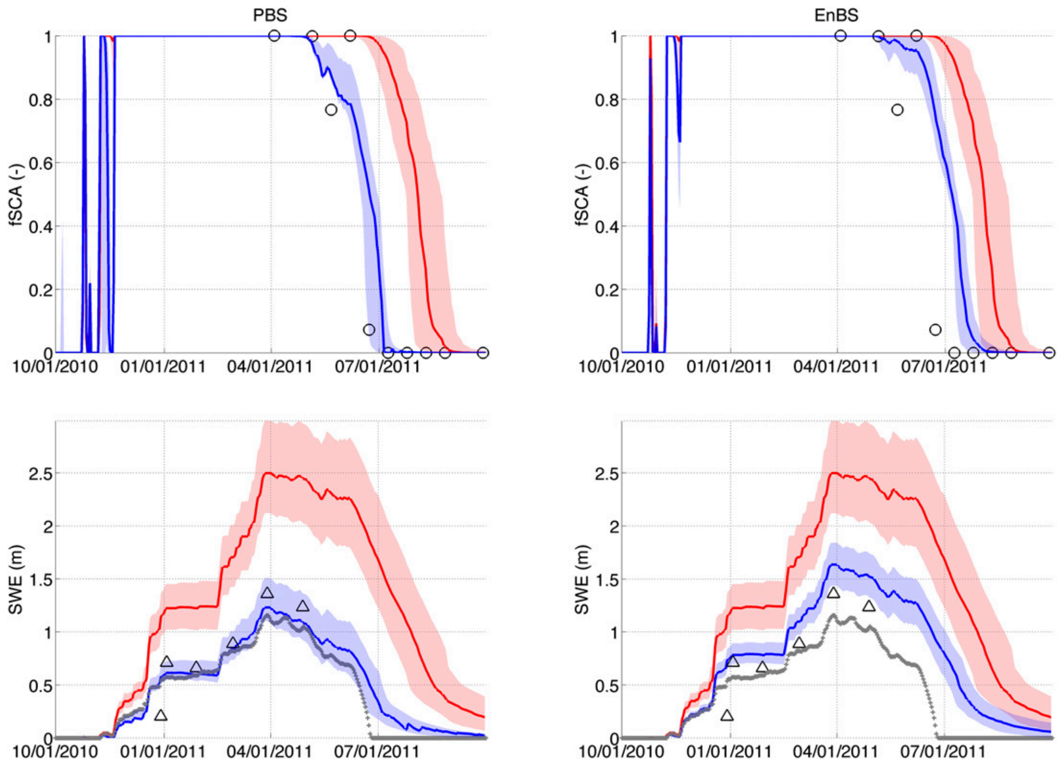

FIG. 7. As in Fig. 4, but for a wet year (WY 2011).

79% of the EnBS RMSE for the snow pillows and snow courses, respectively. The correlations are improved for the posterior in both cases, with slightly higher values for the PS. These SWE error metrics compare favorably to the EnBS implementation applied to the Kern River basin in Girotto et al. (2014a), where the pillow and course ME values were 0.03 and 0.05 m and RMSE values were 0.14 and 0.21 m, respectively. This is notably positive since the American River generally has higher forest cover than Kern, which invariably will obscure more of the snow-covered area on the ground. This implies that the data assimilation methods can work well even in less ideal basins than those applied to previously.

Table 4 also shows the averaged ensemble IQR of the prior and posterior estimates. The IQR cannot be validated in this context, but it is provided as a reference for the internal estimate of uncertainty in the ensemble prior–posterior estimates. The IQR will generally decrease after assimilation of measurements, but by how much depends on the information content of the observations and how efficient the assimilation method is at extracting the information. For the snow pillows and snow courses, the PBS IQR is about 63% and 85%, respectively, of the IQR for the EnBS. Since the two methods assimilate identical observations, this indicates that the PBS is able to extract more information under the nominal setup. One important caveat is that if the number of replicates near the measurements is artificially small (i.e., because of a poor prior distribution), the PBS IQR could be artificially small as well.

The results presented in Table 4 are based on using the smallest error from the nine neighboring pixels around a given in situ site to alleviate representativeness and geolocation errors. Analysis was performed (not shown) that alternatively used the nearest pixel SWE or the average SWE from all neighboring pixels. While the values changed, the qualitative comparison between the PBS and EnBS was essentially the same as those discussed above.

c. *Sensitivity experiment results*

Results for the sensitivity tests identified in Table 3 are shown in Fig. 11 in terms of the ME, RMSE, and IQR over the full 27-yr reanalysis at the snow pillow/snow course sites. The nominal results shown in Fig. 11 correspond to those shown in Table 4. The goal of these tests was to identify how sensitive (if at all) the two methods were to key input variables.

Cases 1a and 1b correspond to the smaller ($N = 50$) and larger ($N = 200$) ensemble size and show limited differences in ME or RMSE for either the PBS or EnBS relative to the nominal case. This supports the argument

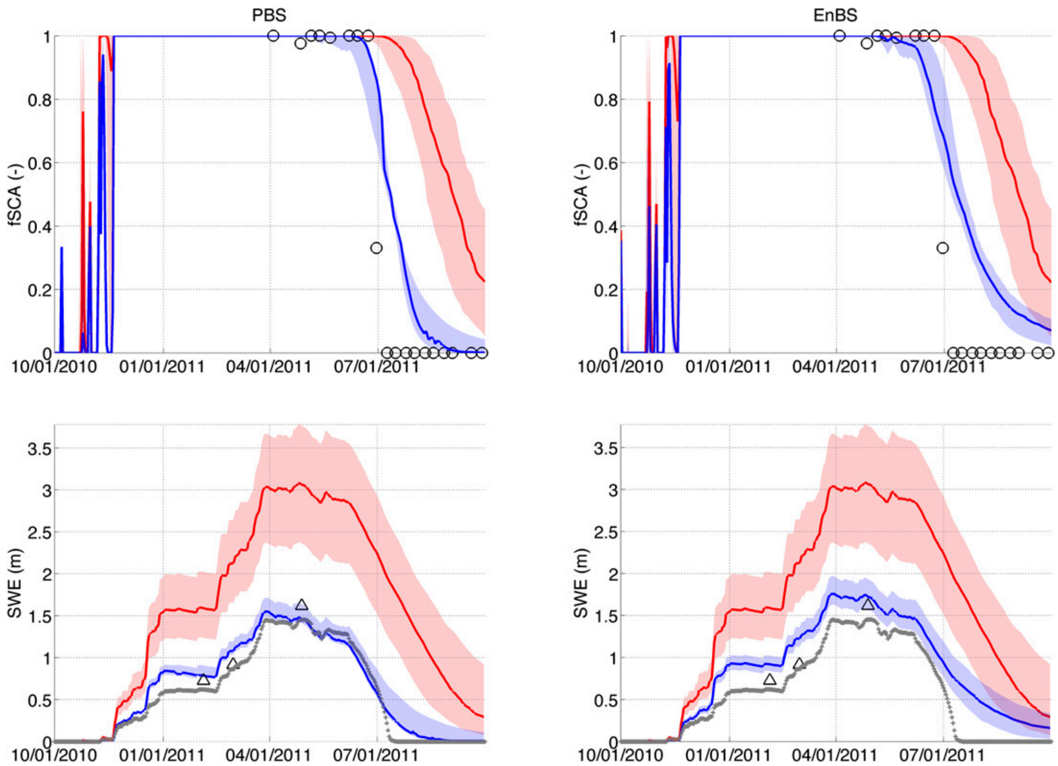

FIG. 8. As in Fig. 5, but for a wet year (WY 2011).

that the nominal ensemble size of $N = 100$ seems sufficient for both methods in the cases tested and, if computational savings is desired, an ensemble size of $N = 50$ may even be acceptable based on these results. For the IQR, there are also only small differences, but the PBS shows an increase in IQR for the larger ensemble size. This may be indicative of an underestimation of ensemble spread in the nominal case, which had prior overestimation in SWE and therefore may have had fewer ensemble members in the neighborhood of the measurements.

Case 2 corresponds to the case of fSCA measurement error being decreased from the nominal value of $\sigma_Z = 15\%$ to $\sigma_Z = 10\%$. The differences in ME and RMSE (relative to the nominal) for pillows and courses are small for both the PBS and EnBS methods. Except in the case of the PBS at the snow course locations, the RMSE becomes slightly smaller when using $\sigma_Z = 10\%$, which may indicate that $\sigma_Z = 10\%$ is more optimal; however, the differences are very small, indicating that using values between 10% and 15% will yield similar SWE estimates in either the PBS or EnBS. The reduction in IQR for both the PBS and EnBS (relative to the nominal) is an expected result since the measurements are trusted more and therefore the posterior spread will be less.

Cases 3a and 3b correspond to the smaller ($\mu_b = 2.0$) and larger ($\mu_b = 3.0$) prior precipitation coefficient mean values. For the PBS, the magnitude of the ME is larger for the two perturbation cases (for both pillows and courses) relative to the nominal, with a more negative ME for $\mu_b = 2.0$ and a positive ME for $\mu_b = 3.0$. For the PBS RMSE comparison, the nominal case has the lowest errors, with the largest errors corresponding to case 3b. This would seem to indicate that the optimal value for μ_b in the PBS lies between 2.0 and 3.0, making the nominal value of 2.5 a reasonable choice. For the EnBS case, the ME is smallest for both the snow pillows and snow courses for case 3a and largest for case 3b. This indicates that for a smaller value of μ_b, the EnBS would have performed significantly better than the nominal case and comparable to that of the PBS. This also supports the conclusion that the optimal value of μ_b in the EnBS lies between 2.0 and 3.0, but that the posterior EnBS results are more sensitive to the chosen value. Specifically, a lower value would lead to a smaller prior bias that would more easily be corrected. In particular for the snow pillow comparison, the EnBS errors are smaller for $\mu_b = 2.0$ than for the nominal case, while much larger for the $\mu_b = 3.0$ case compared to the nominal. The key point is that the EnBS posterior estimates are somewhat dependent on the prior values of μ_b and that the PBS is

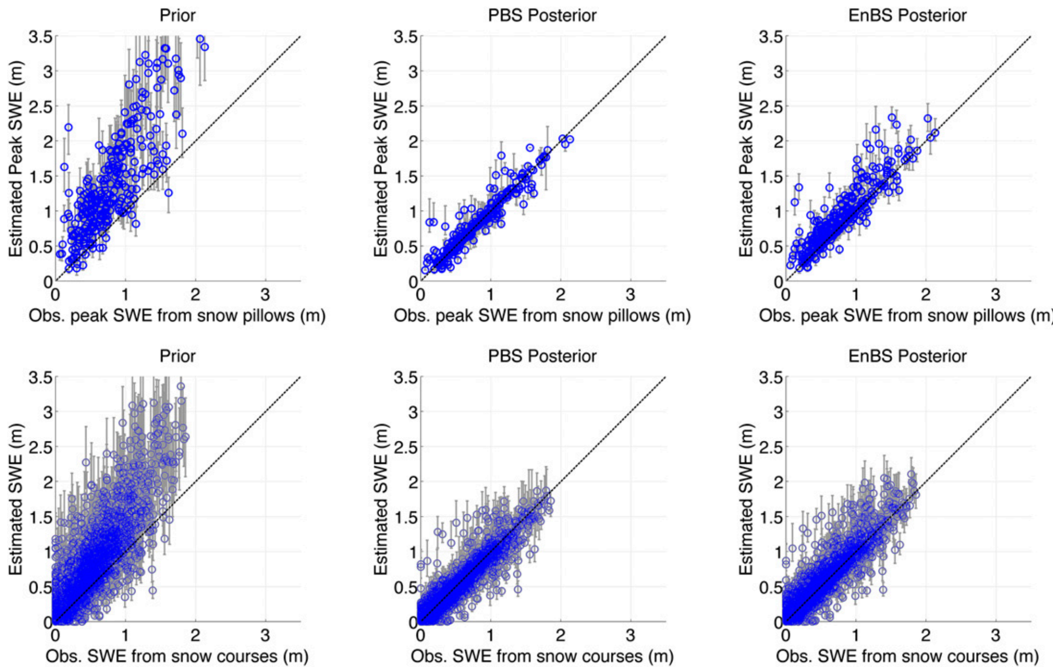

FIG. 9. Scatterplots of (left) prior and SWE reanalysis estimates from (middle) PBS and (right) EnBS vs observations from (top) snow pillows and (bottom) snow courses over all station years. The pillow estimates are those at the time of peak (posterior) SWE. The open circles represent the ensemble median and the error bars represent the ensemble IQR.

less sensitive to the choice for the prior than the EnBS is. The prior for μ_b in this case was time invariant; a more sophisticated prior would likely lead to improved posterior results, especially for the EnBS. For both the PBS and EnBS, the IQR results for the nominal case are bracketed by the two different priors, but in different ways. For the PBS, case 3b yields a smaller IQR. Since this corresponds to a more biased prior, the lower IQR is not necessarily indicative of a more confident estimate, but rather likely a result of fewer replicates near the measurements. For the EnBS, the IQR is largest for case 3b, which corresponds to the fact that the method is unable to extract as much information when the prior is farther from the underlying truth.

5. Conclusions

A previously developed ensemble batch smoother (EnBS) and newly developed particle batch smoother (PBS) data assimilation (reanalysis) method for estimating SWE from historical Landsat-derived fSCA data were applied over the 27-yr *Landsat 5* record to in situ verification sites in the American River basin in the northern Sierra Nevada. The study basin was more highly forested than in previous applications of the EnBS, which allowed for a more robust test of both the old and new methods in a more challenging regime. The key conclusions from this study are as follows:

1) Both data assimilation methods provided significant improvement over the prior (modeling only) estimates. In particular, the prior estimate showed a general overestimation of SWE at the in situ sites examined, which both the EnBS and PBS methods largely removed. The prior RMSE values at the snow pillow and snow course sites were reduced by 68%–82% and 60%–68%, respectively. This result is encouraging for a basin like the American, where the moderate to high forest cover will necessarily obscure more of the snow-covered ground surface than in previously examined, less-vegetated basins.
2) The PBS outperformed the EnBS in the nominal case using nominal input parameters. Specifically, the PBS generally had lower ME and RMSE values and higher correlations in all cases. Differences in ME values between snow pillows and snow courses indicate some inconsistency between the two verification datasets. For snow pillows, the PBS RMSE was ~54% of that seen in the EnBS, while for snow courses the PBS RMSE was ~79% of the EnBS. In general, the EnBS showed smaller (suboptimal) updates relative to the prior (and to the PS) leading to larger posterior estimation errors.

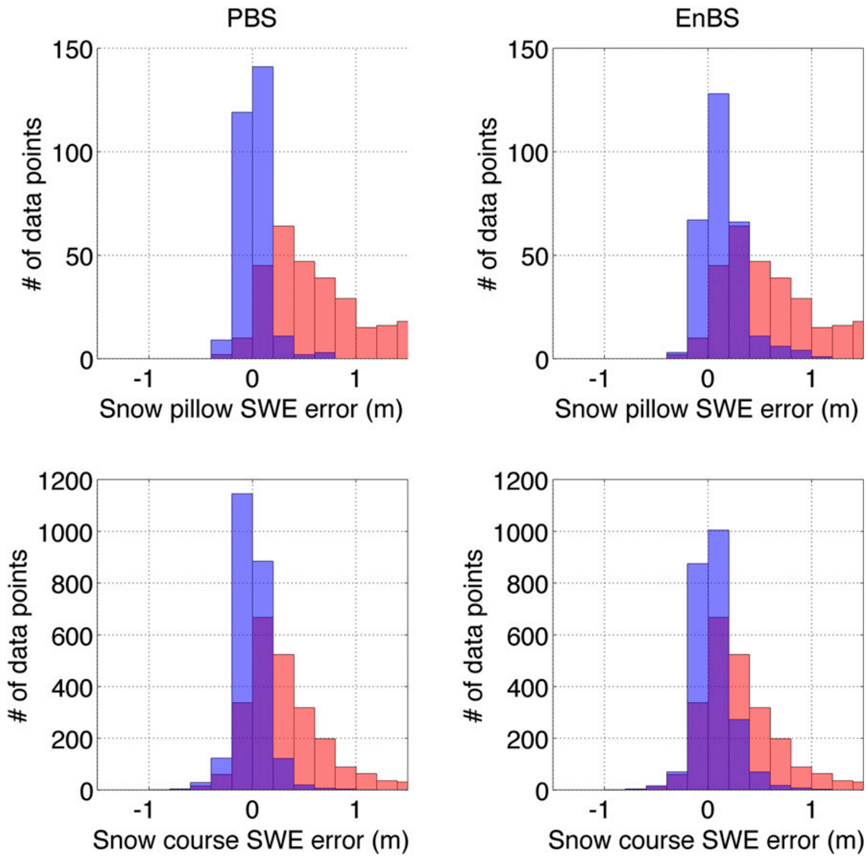

FIG. 10. Error histograms of (left) PBS and (right) EnBS for (top) snow pillow and (bottom) snow course SWE over all station years. The prior errors are shown in red and the reanalysis posterior errors are shown in blue.

3) A sensitivity test to key inputs to the data assimilation framework show relative insensitivity of both the PBS and EnBS results to ensemble size and fSCA measurement error around the nominal values used. However, the results showed that, in comparison to the PS, the EnBS was much more sensitive to the mean prior precipitation input. This sensitivity was directly tied to the suboptimal updates in the nominal case: when the prior estimate is highly erroneous (i.e., positively biased), the EnBS is likely to suffer more from the prior errors. This is likely exacerbated in the case of a highly vegetated basin where the dynamic range of the fSCA measurements is smaller, leading to a smaller signal-to-noise ratio in the assimilated observations. The PBS has less restrictive assumptions on the update, and as long as some replicates in the prior ensemble are close to the observations, the posterior estimate is still reasonably

TABLE 4. Bulk error statistics (ME and RMSE) and correlation coefficient between prior and posterior for the PBS and EnBS estimates across all in situ (snow pillow and snow course) sites for WY 1985–2011. The comparison with snow pillow and snow course data is based on the neighboring pixel with the smallest mean errors to overcome potential representativeness and geolocation errors. In this case, the prior and posterior estimate used in the calculations is the ensemble median. The ME is in reference to the observations so that a positive ME refers to an overestimate (positive bias) relative to the observations. Also shown is the averaged ensemble IQR for each case.

	ME (m)	RMSE (m)	Pearson correlation coef	IQR (m)
Prior pillows	0.58	0.74	0.82	0.53
PBS posterior pillows	0.02	0.13	0.95	0.15
EnBS posterior pillows	0.13	0.24	0.89	0.24
Prior courses	0.30	0.47	0.81	0.38
PBS posterior courses	−0.01	0.15	0.92	0.17
EnBS posterior courses	0.05	0.19	0.89	0.20

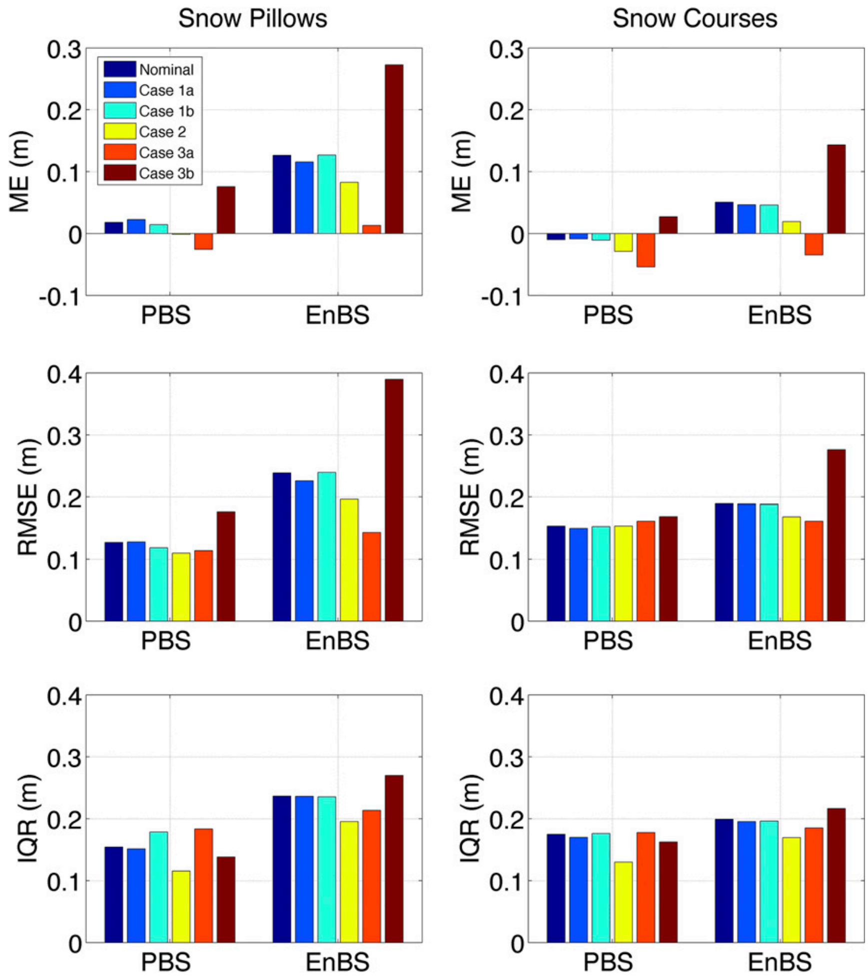

FIG. 11. Comparison of posterior SWE (top) ME, (middle) RMSE, and (bottom) IQR for (left) snow pillows and (right) snow courses for the various sensitivity cases described in Table 3. The nominal case is shown for reference.

accurate. In the sensitivity test case where the prior estimate used in the EnBS was less biased, the estimation errors were comparable to that of the PBS.

Given the increased accuracy and robustness of the PBS over the EnBS method (especially in forested domains) shown in this study, it is suggested as a better alternative to SWE estimation in montane basins where verification data may be limited and input uncertainties can often be high. Some caveats and comments that should be mentioned include the following:

1) The PBS method as implemented herein implicitly updates both states (SWE) and the multiplicative precipitation parameter (b), while the EnBS only updates the multiplicative precipitation parameter directly and then generates a posterior estimate via posterior simulation. However, analysis (not shown herein) found that the high degree of correlation between SWE and b makes the added benefit of dual state–parameter estimation minimal in this context. Rather, it is the generality of the PBS that provides the primary estimation benefit.
2) The likelihood function needed in the PBS as formulated herein may become more difficult to evaluate for higher dimensional measurement vectors. This could become more of an issue when including fSCA estimates from other sensors (e.g., *Landsat 7* and *8* or MODIS) and/or when combining with other data streams (e.g., PM data).
3) The PBS could also likely be made more robust to input errors by using larger prior input uncertainties or different prior distributions to increase the number of replicates that span and cover the measurement space.

Acknowledgments. This was based in part upon work supported by the National Science Foundation under

Grants EAR-0943551, EAR-0943681, and EAR-1246473 and a NASA Earth System Science Fellowship (Grant NNX11AL58H). The NLDAS-2 data used in this study were acquired as part of the mission of NASA's Earth Science Division and are archived and distributed by the Goddard Earth Sciences (GES) Data and Information Services Center (DISC). The raw Landsat data were obtained from the USGS Earth Explorer (http://earthexplorer.usgs.gov/).

REFERENCES

Arulampalam, M. S., S. Maskell, N. Gordon, and T. Clapp, 2002: A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans. Signal Process.,* **50,** 174–188, doi:10.1109/78.978374.

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 Resour. Res.,* **42,** W08432, doi: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,** 303–309, doi:10.1038/nature04141.

Clark, M. P., and Coauthors, 2011: Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review. *Water Resour. Res.,* **47,** W07539, doi:10.1029/2011WR010745.

Cline, D. W., R. C. Bales, and J. Dozier, 1998: Estimating the spatial distribution of snow in mountain basins using remote sensing and energy balance modeling. *Water Resour. Res.,* **34,** 1275–1285, doi:10.1029/97WR03755.

Cortés, G., M. Girotto, and S. A. Margulis, 2014: Analysis of sub-pixel snow and ice extent over the extratropical Andes using spectral unmixing of historical Landsat imagery. *Remote Sens. Environ.,* **141,** 64–78, doi:10.1016/j.rse.2013.10.023.

Crawford, T. M., and C. E. Duchon, 1999: An improved parameterization for estimating effective atmospheric emissivity for use in calculating daytime downwelling longwave radiation. *J. Appl. Meteor.,* **38,** 474–480, doi:10.1175/1520-0450(1999)038<0474:AIPFEE>2.0.CO;2.

De Lannoy, G. J., R. H. Reichle, P. R. Houser, K. R. Arsenault, N. E. Verhoest, and V. R. Pauwels, 2010: Satellite-scale snow water equivalent assimilation into a high-resolution land surface model. *J. Hydrometeor.,* **11,** 352–369, doi:10.1175/2009JHM1192.1.

Dettinger, M., K. Redmond, and D. Cayan, 2004: Winter orographic precipitation ratios in the Sierra Nevada—Large-scale atmospheric circulations and hydrologic consequences. *J. Hydrometeor.,* **5,** 1102–1116, doi:10.1175/JHM-390.1.

Dozier, J., 2011: Mountain hydrology, snow color, and the fourth paradigm. *Eos, Trans. Amer. Geophys. Union,* **92,** 373–375, doi:10.1029/2011EO430001.

——, T. H. Painter, K. Rittger, and J. E. Frew, 2008: Time–space continuity of daily maps of fractional snow cover and albedo from MODIS. *Adv. Water Resour.,* **31,** 1515–1526, doi:10.1016/j.advwatres.2008.08.011.

Durand, M., and S. A. Margulis, 2007: Correcting first-order errors in snow water equivalent estimates using a multifrequency, multiscale radiometric data assimilation scheme. *J. Geophys. Res.,* **112,** D13121, doi:10.1029/2006JD008067.

——, N. Molotch, and S. Margulis, 2008: A Bayesian approach to snow water equivalent reconstruction. *J. Geophys. Res.,* **113,** D20117, doi:10.1029/2008JD009894.

Fassnacht, S. R., K. A. Dressler, and R. C. Bales, 2003: Snow water equivalent interpolation for the Colorado River basin from snow telemetry (SNOTEL) data. *Water Resour. Res.,* **39,** 1208, doi:10.1029/2002WR001512.

Flerchinger, G. N., W. Xaio, D. Marks, T. J. Sauer, and Q. Yu, 2009: Comparison of algorithms for incoming atmospheric long-wave radiation. *Water Resour. Res.,* **45,** W03423, doi:10.1029/2008WR007394.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin, 2004: *Bayesian Data Analysis*. Chapman & Hall/CRC Press, 668 pp.

Girotto, M., G. Cortés, S. A. Margulis, and M. Durand, 2014a: Examining spatial and temporal variability in snow water equivalent using a 27 year reanalysis: Kern River watershed, Sierra Nevada. *Water Resour. Res.,* **50,** 6713–6734, doi:10.1002/2014WR015346.

——, S. A. Margulis, and M. Durand, 2014b: Probabilistic SWE reanalysis as a generalization of deterministic SWE reconstruction techniques. *Hydrol. Processes,* **28,** 3875–3895, doi:10.1002/hyp.9887.

Guan, B., N. Molotch, D. Waliser, S. Jepsen, T. Painter, and J. Dozier, 2013: Snow water equivalent in the Sierra Nevada: Blending snow sensor observations with snowmelt model simulations. *Water Resour. Res.,* **49,** 5029–5046, doi:10.1002/wrcr.20387.

Hall, D. K., G. A. Riggs, V. V. Salomonson, N. E. DiGirolamo, and K. J. Bayr, 2002: MODIS snow-cover products. *Remote Sens. Environ.,* **83,** 181–194, doi:10.1016/S0034-4257(02)00095-0.

Homer, C., and Coauthors, 2007: Completion of the 2001 National Land Cover Database for the conterminous United States. *Photogramm. Eng. Remote Sensing,* **73** (4), 337–341.

Jepsen, S. M., N. P. Molotch, M. W. Williams, K. E. Rittger, and J. O. Sickman, 2012: Interannual variability of snowmelt in the Sierra Nevada and Rocky Mountains, United States: Examples from two alpine watersheds. *Water Resour. Res.,* **48,** W02529, doi:10.1029/2011WR011006.

Kelly, R. E., A. T. Chang, L. Tsang, and J. L. Foster, 2003: A prototype AMSR-E global snow area and snow depth algorithm. *IEEE Trans. Geosci. Remote Sens.,* **41,** 230–242, doi:10.1109/TGRS.2003.809118.

Li, D., M. Durand, and S. A. Margulis, 2012: Potential for hydrologic characterization of deep mountain snowpack via pass microwave remote sensing in the Kern River basin, Sierra Nevada, USA. *Remote Sens. Environ.,* **125,** 34–48, doi:10.1016/j.rse.2012.06.027.

Li, H.-Y., and J. Wang, 2011: Simulation of snow distribution and melt under cloudy conditions in an alpine watershed. *Hydrol. Earth Syst. Sci.,* **15,** 2195–2203, doi:10.5194/hess-15-2195-2011.

Liston, G. E., 2004: Representing subgrid snow cover heterogeneities in regional and global models. *J. Climate,* **17,** 1381–1397, doi:10.1175/1520-0442(2004)017<1381:RSSCHI>2.0.CO;2.

Lundquist, J. D., J. R. Minder, P. J. Neiman, and E. Sukovich, 2010: Relationships between barrier jet heights, orographic precipitation gradients, and streamflow in the northern Sierra Nevada. *J. Hydrometeor.,* **11,** 1141–1156, doi:10.1175/2010JHM1264.1.

Molotch, N. P., 2009: Reconstructing snow water equivalent in the Rio Grande headwaters using remotely sensed snow cover data and a spatially distributed snowmelt model. *Hydrol. Processes,* **23,** 1076–1089, doi:10.1002/hyp.7206.

——, and R. C. Bales, 2005: Scaling snow observations from the point to the grid element: Implications for observational network design. *Water Resour. Res.,* **41,** W11421, doi:10.1029/2005WR004229.

——, and S. A. Margulis, 2008: Estimating the distribution of snow water equivalent using remotely sensed snow cover data and a spatially distributed snowmelt model: A multi-resolution, multi-sensor comparison. *Adv. Water Resour.,* **31,** 1503–1514, doi:10.1016/j.advwatres.2008.07.017.

——, T. H. Painter, R. C. Bales, and J. Dozier, 2004: Incorporating remotely-sensed snow albedo in a spatially-distributed snowmelt model. *Geophys. Res. Lett.,* **31,** L03501, doi:10.1029/2003GL019063.

Moradkhani, H., K.-L. Hsu, H. Gupta, and S. Sorooshian, 2005: Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. *Water Resour. Res.,* **41,** W05012, doi:10.1029/2004WR003604.

Painter, T. H., K. Rittger, C. McKenzie, P. Slaughter, R. E. Davis, and J. Dozier, 2009: Retrieval of subpixel snow-covered area, grain size, and albedo from MODIS. *Remote Sens. Environ.,* **113,** 868–879, doi:10.1016/j.rse.2009.01.001.

Pan, M., and Coauthors, 2003: Snow process modeling in the North American Land Data Assimilation System (NLDAS): 2. Evaluation of model simulated snow water equivalent. *J. Geophys. Res.,* **108,** 8850, doi:10.1029/2003JD003994.

Prata, A. J., 1996: A new long-wave formula for estimating downward clear-sky radiation at the surface. *Quart. J. Roy. Meteor. Soc.,* **122,** 1127–1151, doi:10.1002/qj.49712253306.

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

Rittger, K. E., 2012: Spatial estimates of snow water equivalent in the Sierra Nevada. Ph.D. thesis, University of California, Santa Barbara, 245 pp.

Rosenthal, W., and J. Dozier, 1996: Automated mapping of montane snow cover at subpixel resolution from the Landsat Thematic Mapper. *Water Resour. Res.,* **32,** 115–130, doi:10.1029/95WR02718.

Talagrand, O., 2003: Bayesian estimation, optimal interpolation, statistical linear estimation. *Data Assimilation for the Earth System,* R. Swinbank, V. Shutyaev, and W. A. Lahoz, Eds., NATO Science Series: IV, Vol. 26, Kluwer Academic, 21–35.

Tedesco, M., R. H. Reichle, A. Loew, T. Markus, and J. L. Foster, 2010: Dynamic approaches for snow depth retrieval from spaceborne microwave brightness temperature. *IEEE Trans. Geosci. Remote Sens*., **48,** 1955–1967, doi:10.1109/TGRS.2009.2036910.

Trujillo, E., N. P. Molotch, M. L. Goulden, A. E. Kelly, and R. C. Bales, 2012: Elevation-dependent influence of snow accumulation on forest greening. *Nat. Geosci.,* **5,** 705–709, doi:10.1038/ngeo1571.

Vander Jagt, B. J., M. T. Durand, S. A. Margulis, E. J. Kim, and N. P. Molotch, 2013: The effect of spatial variability on the sensitivity of passive microwave measurements to snow water equivalent. *Remote Sens. Environ.,* **136,** 163–179, doi:10.1016/j.rse.2013.05.002.

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 Resour. Res.,* **43,** W07447, doi:10.1029/2006WR005653.

Xia, Y., and Coauthors, 2012: Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *J. Geophys. Res.,* **117,** D03109, doi:10.1029/2011JD016048.

Xue, Y., S. Sun, D. S. Kahan, and Y. Jiao, 2003: Impact of parameterizations in snow physics and interface processes on the simulation of snow cover and runoff at several cold region sites. *J. Geophys. Res.,* **108,** 8859, doi:10.1029/2002JD003174.

Zhou, Y., D. McLaughlin, and D. Entekhabi, 2006: Assessing the performance of the ensemble Kalman filter for land surface data assimilation. *Mon. Wea. Rev.,* **134,** 2128–2142, doi:10.1175/MWR3153.1.