A High-Resolution Data Assimilation Framework for Snow Water Equivalent Estimation across the Western United States and Validation with the Airborne Snow Observatory

CATALINA M. Oaida,^a JOHN T. REAGER, KONSTANTINOS M. ANDREADIS,^b CÉDRIC H. DAVID,
STEVE R. LEVOE, THOMAS H. PAINTER, KAT J. BORMANN, AND AMY R. TRANGSRUD

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

MANUELA GIROTTO

*Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, and GESTAR,
Universities Space Research Association, Columbia, Maryland*

JAMES S. FAMIGLIETTI^c

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

(Manuscript received 20 January 2018, in final form 12 October 2018)

ABSTRACT

Numerical simulations of snow water equivalent (SWE) in mountain systems can be biased, and few SWE observations have existed over large domains. New approaches for measuring SWE, like NASA's ultra-high-resolution Airborne Snow Observatory (ASO), offer an opportunity to improve model estimates by providing a high-quality validation target. In this study, a computationally efficient snow data assimilation (DA) approach over the western United States at 1.75-km spatial resolution for water years (WYs) 2001–17 is presented. A local ensemble transform Kalman filter implemented as a batch smoother is used with the VIC hydrology model to assimilate the remotely sensed daily MODIS fractional snow-covered area (SCA). Validation of the high-resolution SWE estimates is done against ASO SWE data in the Tuolumne basin (California), Uncompahgre basin (Colorado), and Olympic Peninsula (Washington). Results indicate good performance in dry years and during melt, with DA reducing Tuolumne basin-average SWE percent differences from −68%, −92%, and −84% in open loop to 0.6%, 25%, and 3% after DA for WYs 2013–15, respectively, for ASO dates and spatial extent. DA also improved SWE percent difference over the Uncompahgre basin (−84% open loop, −65% DA) and Olympic Peninsula (26% open loop, −0.2% DA). However, in anomalously wet years DA underestimates SWE, likely due to an inadequate snow depletion curve parameterization. Despite potential shortcomings due to VIC model setup (e.g., water balance mode) or parameterization (snow depletion curve), the DA framework implemented in this study shows promise in overcoming some of these limitations and improving estimated SWE, in particular during drier years or at higher elevations, when most in situ observations cannot capture high-elevation snowpack due to lack of stations there.

Denotes content that is immediately available upon publication as open access.

^a Current affiliation: Raytheon Company, Pasadena, California.

^b Current affiliation: Department of Civil and Environmental Engineering, University of Massachusetts Amherst, Amherst, Massachusetts.

^c Current affiliation: Global Institute for Water Security, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

Corresponding author: Catalina M. Oaida, oaidac@gmail.com

DOI: 10.1175/JHM-D-18-0009.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](https://www.ametsoc.org/PUBSReuseLicenses) (www.ametsoc.org/PUBSReuseLicenses).

1. Introduction

Mountain systems with snowmelt-driven hydrologic cycles are crucial to more than one-sixth of the world's population that depends on seasonal runoff as its primary water source (Barnett et al. 2005). Such systems include the western United States, the Alps, and the Hindu Kush, among others, many of which have agricultural and urban areas that depend heavily on the water supply provided by seasonal snowmelt. In addition, the timing of snow ablation plays an important role in soil and vegetation conditions in these mountainous regions, affecting ecosystems and phenomena such as wildfire timing, extent, and length of season (Westerling et al. 2006).

Because of the complexity of mountainous terrain, snow mass and snow extent exhibit great spatial heterogeneity, with their variability being influenced by a range of factors such as elevation, vegetation cover, slope and aspect, topography, orographic effects, wind redistribution, and atmospheric circulation (e.g., Dettinger et al. 2004; Clark et al. 2011; Giroto et al. 2014a). Given the importance and highly variable nature of the mountain snowpack, our ability to accurately estimate how much snow water equivalent (SWE) is stored in these mountainous regions is not only greatly desirable, but also a major scientific challenge (e.g., Dozier 2011; Margulis et al. 2015; Dozier et al. 2016; Painter et al. 2016). In situ observations, remotely sensed data, modeling, or a combination of these data sources have been used to address fundamental questions such as how much water is stored in mountain snowpacks, how does the snowpack vary in space and time, and how various physiological and climatological factors impact snow accumulation and ablation (e.g., Dozier 2011; Margulis et al. 2015). Answers to such questions still have significant uncertainty associated with them. This is due in great part to the fact that mountain snow is one of the most seasonally changing components of the Earth system, and there is presently no one single method that can accurately and fully capture the extent and heterogeneity of mountain snow and the processes that control the system dynamics.

In the western United States, in situ measurements such as snow pillows and snow courses have long been used by water resource managers to estimate seasonal water storage, yet these left runoff forecasts with significant errors (Dozier 2011). While the American West is a densely sampled mountain region, point-scale measurements cover only a very small portion of the snow-dominated areas. For example, in the Sierra Nevada the snow sensor density is 1 for every 700 km² (Guan et al. 2013), with most sensors covering the lower to

mid-elevations, 6500–9500 ft (~1900–2900 m), leaving the higher elevations undersampled (Dozier et al. 2016). This can be a problem in particular during drier years or toward the end of the snow season, when most snow is concentrated at the higher elevations and the snow has melted from many in situ measurement locations, as well as in a warming climate.

Remote sensing can provide useful measurements for mapping snow, including the undersampled high-elevation locations. There are two main methods of collecting snow information using spaceborne instruments. Passive microwave (PM) has been explored by past and current studies (e.g., Durand and Margulis 2007; Tedesco et al. 2010; Li et al. 2012), primarily because it can retrieve snow depth, which is a more direct and often more precise way of obtaining SWE, and it experiences no interference from clouds. However, PM data are often too coarse in spatial resolution—often greater than 10 km—making it difficult to use to estimate SWE given the subpixel spatial variability of snow and factors controlling snow and snowmelt. In addition, PM signal can saturate over deep snow. Studies have found this threshold to range from 100 to 150 mm, and to be as high as 200 mm (e.g., Takala et al. 2011; Liu et al. 2014). The other main technique is using visible and near-infrared (Vis/NIR; e.g., Painter et al. 2009; Rittger et al. 2013), which can provide information at better spatial resolution (30–500 m), but which is hindered by cloud cover, and perhaps more problematically, only retrieves snow-covered area (SCA) and not depth-integrated metrics. Because instantaneous snow cover does not provide ample information on instantaneous snow depth or SWE, using remotely sensed snow extent for water estimation requires more indirect methods such as reconstruction to determine SWE (e.g., Molotch and Margulis 2008; Schneider and Molotch 2016).

Recent advancements in airborne retrievals of snow depth, snow albedo, and SWE across watershed-scale regions have been providing unparalleled spatial distributions at relatively high resolution. The Airborne Snow Observatory (ASO; Painter et al. 2016) provides SWE estimates at 50-m spatial resolution, taken at approximately monthly intervals during accumulation and approximately weekly during ablation. ASO data are currently available over a growing number of watersheds in California and Colorado, with limited retrievals over the Olympic Peninsula in Washington State. Despite the high accuracy and rich spatial distribution of snow information it collects, at this time ASO does not yet provide these measurements over entire mountain ranges, at regional or subcontinental scales, or during the full duration of the water year (WY). In this study we use it as a validation dataset but recognize its value in

informing current basin-scale operational water management practices, as well as its potential in data assimilation (DA) techniques.

Given the above challenges and shortcomings in observations when it comes to estimating SWE across mountainous terrain, using models is necessary to bridge temporal and spatial gaps in measurements. Snow models, as components in hydrologic, land surface, or climate models, are based on physical processes, and in combination with additional physiological data such as land cover (soil and vegetation type) and elevation maps, and forced with meteorological data, can provide spatially and temporally continuous SWE estimates at varying horizontal scales. With the advancement in computing resources in recent years, we can simulate SWE (and other hydrologic variables) at increasingly higher spatial resolutions. However, uncertainties in the input data, model parameters, and/or model design can lead to uncertainties in model output.

To address some of these deficiencies, a combination of observed data and modeling can be employed. DA is one such technique, which, generally speaking, uses observed data (e.g., remotely sensed) to correct or improve model estimates, and provides the associated uncertainty. This results in a spatially and temporally continuous SWE estimation that leverages the benefits of both observed data and modeling capabilities, which, when used over a relatively long period of time (e.g., multiyear), leads to the development of reanalysis datasets. Such datasets include the North American Regional Reanalysis (NARR; 32 km; [Mesinger et al. 2006](#)), NCEP–NCAR reanalysis (2.5°; [Kanamitsu et al. 2002](#)), and Modern-Era Retrospective Analysis for Research and Applications (MERRA; 50 km; [Rienecker et al. 2011](#)). While these existing data can be a useful resource for larger-scale regional or global climate studies, their relatively coarse spatial resolution makes them inadequate for subgrid, basin-scale SWE estimation across mountainous terrain. The Snow Data Assimilation System (SNODAS; [Barrett 2003](#)) does produce snow data over the United States and parts of southern Canada at a relatively high horizontal resolution of $1\text{ km} \times 1\text{ km}$, by using a forward modeling approach that is updated via nudging from a variety of remote sensing and in situ data. However, it has been shown to either underestimate or overestimate SWE when compared with other SWE estimation methods such as reconstruction, interpolation, or recent airborne SWE measurements ([Dozier 2011](#); [Rittger 2012](#); [Bair et al. 2016](#); [Dozier et al. 2016](#)). Independent validation against ground-based observations is difficult since SNODAS assimilates the vast majority of these data as part of the SWE estimation process. Higher-resolution

reanalyses have been recently developed, but they only cover certain river basins or mountain ranges (e.g., Sierra Nevada reanalysis; [Guan et al. 2013](#); [Rittger et al. 2013](#); [Margulis et al. 2016](#)). As a result, there are no high-resolution, spatially distributed, temporally and spatially continuous SWE datasets over mountainous terrain across regional to continental scales and spanning multiple years (with the exception of SNODAS, whose shortcomings are presented above).

While significant strides have been made in combining models and observations of snow in complex topographic terrain, accurately and comprehensively estimating spatially distributed SWE remains one of the biggest challenges in mountain hydrology (e.g., [Dozier et al. 2016](#)). One reason for this can be traced to the fact that in assimilation, most approaches integrate snow cover products into models to estimate SWE. Because an instantaneous SCA measurement does not inform how much instantaneous SWE there is in the snowpack, previous studies that used sequential filtering schemes such as ensemble Kalman filters (EnKFs) to assimilate snow cover data have only attained limited improvements in SWE estimates (e.g., [Andreadis and Lettenmaier 2006](#); [De Lannoy et al. 2012](#)). If the goal of snow DA is to produce an SWE reanalysis—a multiyear retrospective snow dataset as opposed to short-term seasonal streamflow forecast—one can use a batch estimation approach (or smoothing), which is ideal for historic data ([Dunne and Entekhabi 2005](#)). In this method, the model is first run forward over the entire water year to estimate prior SCA, and then observed SCA values from throughout the seasonal cycle are assimilated into the hydrologic model in batch (all at once), such that estimated SWE at a given time is determined by including both prior and subsequent SCA observations. This is effective because in many cases SCA evolution throughout the snow season is highly correlated with SWE and in particular in the ablation period (e.g., [Durand et al. 2008](#); [Giroto et al. 2014a](#); [Margulis et al. 2015](#)). This method is different from the more typically used sequential filtering, where the state variable (e.g., SWE) is updated sequentially as (SCA) observations become available (and therefore the updates do not include information on the future state of the observed or modeled quantity). In other words, an advantage to using a batch smoother approach to assimilate SCA in order to update SWE is that it alleviates shortcomings related to how much information instantaneous SCA can providing for estimating instantaneous SWE (i.e., snowpack depth), by incorporating longer-term, seasonal SCA evolution behavior, which is more closely tied to SWE evolution. The batch smoother approach has

been used successfully with assimilation of Landsat-derived fractional SCA data over the Sierra Nevada mountain range (Margulis et al. 2016). Because mountain snow changes rapidly during the ablation season, assimilating daily MODIS Vis/NIR—and in particular the fractional SCA (fSCA) MODIS Snow Covered Area and Grain Size (MODSCAG; Painter et al. 2009) product—using a batch smoothing approach could prove to be beneficial in improving SWE estimates across continental-scale domains.

Another challenge is the computational power and resources required to simulate and produce data at higher spatial resolution over extended periods of time and space. Members of the hydrologic community have acknowledged that in order to effectively tackle both research and applied water cycle challenges, higher-resolution (on the order of 1 km) hydrologic data are necessary across continental to global domains (e.g., Wood et al. 2011). As computational capabilities increase with the growth in computer power, hydrological models are run at increasingly fine resolutions and across larger spatial scales. These types of simulations require a DA approach that is computationally efficient and easy to understand mechanistically.

In this study, we test and validate a snow DA method that aims at simulating a continental-scale domain at 1.75-km resolution, to provide a spatially and temporally complete estimate of mountain SWE across the western United States over the past 17 years. The local ensemble transform Kalman filter (LETKF; Hunt et al. 2007), employed here as a batch smoother, is chosen for its relative simplicity, ease of implementation, and computational efficiency. The Variable Infiltration Capacity (VIC) hydrologic model (Liang et al. 1994) is used to compute the water and energy budgets and produces a first estimate of SCA and SWE. We then assimilate the MODSCAG (Painter et al. 2009) daily satellite SCA product to constrain an updated model SWE estimate. MODSCAG has a daily temporal resolution, provides subpixel (fractional) SCA, and offers the most accurate snow products available from MODIS (Rittger et al. 2013). One of the major novelties of this study is that our high-resolution SWE estimates (1.75 km) are validated against the ASO SWE product (50 m), the only ultra-high-resolution, spatially distributed SWE observations, which facilitates a more accurate assessment of modeled SWE spatial variability. In section 2 we describe the method for this study, including the hydrologic model, the data assimilation technique, the assimilation and validation datasets, and experimental design. In section 3 we present results, followed by discussion and summary in section 4.

2. Methods

a. Hydrologic model

The model used here is the VIC model (Liang et al. 1994), a robust and widely used macroscale hydrologic model that provides estimates of hydrologic state variables. The meteorological forcings used here to run VIC version 4.2.d include the PRISM daily precipitation and minimum and maximum temperature (<http://prism.oregonstate.edu>), and MERRA-2 wind (Gelaro et al. 2017; Global Modeling and Assimilation Office 2015; data can be accessed through https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/data_access/). These datasets were chosen because of their long temporal span, adequate spatial coverage over the western United States, daily temporal frequency, and in the case of PRISM precipitation, a high enough spatial resolution to capture complex mountain precipitation regimes.

In this realization we employ VIC in the water balance mode only, not full energy. This means that while the model does not solve the surface energy balance directly, allowing it to be computationally faster, in the snow module the surface energy balance is still computed explicitly in order to estimate the fluxes that drive snow accumulation and ablation (<http://vic.readthedocs.io>). The snow model, described in detail by Andreadis et al. (2009), accounts for snow on the ground as well as snow in the vegetation canopy. Ground snow has two layers, a thin surface layer and a pack layer with accumulation being controlled by snowfall and ablation that are calculated by solving the energy balance at the snow–air interface. Snow albedo is assumed to decay with age using an established parameterization, while densification results from compaction and snow metamorphism. Snow cover fraction (i.e., SCA) is calculated as a function of snow depth based on a probability distribution, being equal to 100% for depths greater than a critical depth (Cherkauer and Lettenmaier 2003). This relationship (SCA–snow depth) is often referred to as the snow depletion curve (SDC). The SDC algorithm in VIC was developed by Cherkauer and Lettenmaier (2003) based on deriving uniform probability distribution from observed snow depth at the University of Minnesota’s Rosemount Agricultural Experiment Station during the winters of 1997/98 and 1998/99. This fitted uniform distribution ignores extreme minimum and maximum observed values ($P < 0.1$ and $P > 0.9$) and is able to represent most of the spatial variability in daily snow depth observations, largely because there is little spatial variability in observations in this Cherkauer and Lettenmaier (2003) study, which is to be expected given the relatively flat terrain in Minnesota. In the parameterization, when the mean snow depth is greater than or

FIG. 1. Schematic diagram describing the process of data assimilation in this study. See section 2b in the main text for a description of the LETKF batch smoother and how it was implemented in the modeling framework.

equal to a critical depth, the snow coverage is 100% for that model grid. If the mean snow depth is less than the critical depth, snow-free areas exist in the grid. The algorithm can also account for cases when there is melting of thin accumulating snowpacks, or when there is fresh snow on a partially snow-free grid cell. More details about this algorithm can be found in Cherkauer and Lettenmaier (2003). It is important to note that because these parameterizations were derived based on snow depth observations over relatively flat surfaces, using uniform distribution, the resulting SDC is shallow, meaning once the snow is melting and reaches a certain depth, fSCA will quickly go to zero. Our western U.S. domain is mountainous, often with heterogeneous terrain and therefore snow depth, implying the current VIC SDC parameterization is not entirely appropriate for this study. Ideally, an SDC parameterization should be derived for respective mountain ranges or basins, since the snowpack is heterogeneous in mountain environments; however, this is an undertaking worthy of a study in and of itself, and was thus beyond the scope of this study. We hope to address this shortcoming in future work. In this instance of the model, the frozen soil option is set to "FALSE." While frozen soil is an important consideration for springtime runoff and streamflow (e.g., Cherkauer et al. 2003) (from a water balance perspective), the focus of this study is on improving SWE estimates via DA rather than downstream components. Furthermore, frozen soil state can contribute to ground heat-snow fluxes, but these are often small compared to the snow-atmosphere interface flux exchange. Therefore, having frozen soil set to FALSE is not expected to have a significant effect on SWE accumulation or ablation processes.

b. Data assimilation approach: LETKF

Data assimilation techniques aim to estimate the state of a dynamic system by using both models

and observations. There are several methods for approaching DA, ranging from direct insertion to a simple Kalman filter, extended Kalman filter, or EnKF, to batch smoothers like the Kalman-based ensemble batch smoother (EnBS) or a particle batch smoother. Examples of such DA work include Liston et al. (1999), Rodell and Houser (2004), Slater and Clark (2006), Andreadis and Lettenmaier (2006), De Lannoy et al. (2012), and Margulis et al. (2015). Ensemble-based assimilation techniques (e.g., Evensen 1994; Evensen 2003; Hamill 2006) are used for nonlinear dynamic systems and are often preferred since they can represent complex error structures without being computationally cumbersome (i.e., calculates and propagates error estimates from the variance of the ensemble of model states). In this study we use a particular version of an EnKF, the LETKF (Hunt et al. 2007), and we do so within the context of a batch smoother. This technique is relatively easy to implement, is computationally efficient, and has the flexibility to be adapted to a variety of applications while maintaining formal simplicity (of other DA methods). The LETKF DA method used here is adapted from an implementation in the Regional Hydrologic Extremes Assessment System (RHEAS) software (Andreadis et al. 2017), which also employs the VIC hydrologic model to produce hydrologic estimates.

The LETKF decomposes the estimation problem into a set of local domains (spatially or temporally discretized) where the analysis equations are solved independently using only nearby observations. Defining the background (i.e., open loop) state vector as $\mathbf{x}^b$ (here, SWE), the predicted measurements (here, modeled SCA) as $\mathbf{y}^b$, and the observations as $\mathbf{y}^o$ (here, observed SCA), we can derive the analysis (post-DA SWE) mean $\mathbf{x}^a$ as

$$\mathbf{x}^a = \mathbf{X}^b \bar{\mathbf{w}}^a + \mathbf{x}^b, \quad (1)$$

where

$$\bar{\mathbf{w}}^a = \mathbf{P}^a \mathbf{Y}^{bT} \mathbf{R}^{-1} (\mathbf{y}^o - \bar{\mathbf{y}}^b), \quad (2)$$

and

$$\mathbf{P}^a = [(N-1)\mathbf{I}/\rho + \mathbf{Y}^{bT} \mathbf{R}^{-1} \mathbf{Y}^b]^{-1}. \quad (3)$$

where N is the ensemble size (here set to equal 10); $\mathbf{X}^b$ is the matrix of background perturbations (i.e., a priori modeled SWE); $\mathbf{Y}^b$ is the matrix of the predicted measurement perturbations (i.e., a priori modeled SCA); $\mathbf{R}$ is the measurement (i.e., observed SCA) error covariance matrix; ρ is a multiplicative covariance inflation factor, here set to 1.05; and $\mathbf{I}$ is the identity matrix. The ρ value is chosen based on recommendation of [Hunt et al. \(2007\)](#) and can be thought of as adjusting the influence of observations on the current analysis, such that the influence of an observation on future analyses decays exponentially with time, and thus localizes the analysis in time. Given that the assimilation is applied locally for each pixel, our relatively small ensemble size can capture the variability (e.g., [Ott et al. 2004](#)). The analysis (state) ensemble perturbations $\mathbf{X}^a$ (i.e., post-assimilation modeled SWE) can be updated as

$$\mathbf{X}^a = \mathbf{X}^b \mathbf{W}^a, \quad (4)$$

where

$$\mathbf{W}^a = [(N-1)\mathbf{P}^a]^{1/2}, \quad (5)$$

and where $\mathbf{W}^a$ can be thought of as the optimal weights between model and observed states given model and observed errors. The ensembles of background $\mathbf{X}^b$ and predicted measurement $\mathbf{Y}^b$ perturbations are randomly sampled from a 30-yr modeled SWE and SCA climatology (1981–2010) that is only generated once. This avoids producing meteorological forcing perturbation every time, making the algorithm more efficient. The observational error $\mathbf{R}$ is computed by randomly sampling a Gaussian distribution with mean zero and standard deviation set based on the observed data used. Here we set a standard deviation observation error of 0.10, corresponding to MODSCAG data.

As mentioned in [section 1](#), if SCA observations are assimilated in order to obtain retrospective SWE estimates, using a batch smoother method is more appropriate because it includes both prior and subsequent observations in the estimation of SWE at a given time, thus capturing the stronger SCA–SWE seasonal temporal correlation not often seen for instantaneous

SCA–SWE relationships (which is what a sequential filter would capture). In other words, SCA DA methods that employ sequential filtering only use the observed SCA value at a given point in time (instantaneous SCA) to update (instantaneous) SWE, which typically limits the efficiency of this method to shallower snowpack. On the other hand, assimilating SCA in batch (using a batch smoother) includes an entire snow season's SCA observations to update SWE, taking advantage of the stronger seasonal correlation between SCA and SWE evolution. For this study, we use the LETKF as a batch smoother ([Fig. 1](#)): the model is first run forward in time for each water year (from 1 October to 1 October) to estimate prior SCA, and then observed SCA over a specified interval (an assimilation window) is assimilated into the hydrologic model in batch (all at once), retroactively updating SWE estimates for that period. More specifically, VIC is first run in open loop, estimating prior SCA and SWE. When the modeling system is run in the assimilation scenario (see [section 2d](#)), we restrict the assimilation window to 1 January–1 August of each water year in order to avoid spurious correlations during the beginning of the accumulation season. Outside of that window the model is run in open-loop mode, and assimilation of observed SCA is excluded. All observations within this assimilation window are incorporated at once, as a batch, and not sequentially. This approach has been used by other batch smoother DA studies, such as [Giroto et al. \(2014a,b\)](#). Because SCA is quite variable during the early accumulation season, going from 0% to 100% quickly, it can lead to positive, but spurious, correlations that would misleadingly update the weights in the filter (i.e., $\mathbf{W}^a$). Furthermore, based on our sensitivity studies (not shown), the choice of length and time of the year for the assimilation window also matters: if too much of the accumulation season is excluded from the assimilation (e.g., assimilation window is 1 March–1 July or 1 February–1 August), the updated SWE will be underestimated at the end of the snow season because we would miss capturing part of the period during which SCA evolution is highly correlated with SWE. This sensitivity highlights the time of the year remote sensing snow cover observations are most needed.

c. Observational datasets

1) MODSCAG

The data assimilation method described above uses remotely sensed fractional SCA to update modeled SWE. The observational dataset employed here is MODSCAG, a MODIS-derived product that uses the MODIS Snow Covered Area and Grain size retrieval algorithm ([Painter et al. 2009](#)) to produce a per-pixel

fractional SCA and snow grain size from MOD09GA surface reflectance. In this work, we use the daily fractional SCA data, which range from 0% to 100% and account for any vegetation and soil within each pixel. MODSCAG fractional SCA has a lower root-mean-square error (RMSE) compared to MOD10A1 fractional cover (0.10 vs 0.23, respectively), consistently maintaining its retrieval capability during crucial transition periods of accumulation and melt, across all land cover classes and land surface types (Rittger et al. 2013).

To retrieve and process the MODSCAG data, the RHEAS software downloads daily geotiffs from the MODSCAG data portal (<https://snow-data.jpl.nasa.gov/>) via ftp for the time period indicated by the user, stitches the native ~ 500 -m tiles into one daily geotiff, regrids the data to a user-specified spatial resolution (here 1 km^2 , in this case using the nearest neighbor method), and trims the data for a region of interest (here, the western United States). During assimilation the resampled 1-km^2 MODSCAG data get reaggregated to a user-specified model resolution, here 1.75 km , by selecting the nearest MODSCAG 1-km^2 pixel to the center of the model grid. At the time of this study only the noncanopy adjusted daily MODSCAG SCA data were available for the full length of DA model simulations (WYs 2001–17), which means the observed SCA used is the viewable rather than the ground observation, likely underestimating true SCA. Due to this factor, and given this is a fractional SCA which adjusts for vegetation cover area (but does not see under vegetation canopy), the MODSCAG fractional SCA over a model pixel can often range between zero and a value less than 100%. This becomes important when comparing to modeled SCA, which is generally either zero or 100% (due to a shallow snow depletion curve parameterization). In addition, while MODSCAG has a high daily temporal resolution, it is affected by the instrument's viewing angle, causing snow cover measurements to often jump from full coverage to little to no coverage within the span of a few days. The implications of this model-observed snow cover discrepancy will be further discussed in sections 3 and 4.

2) AIRBORNE SNOW OBSERVATORY

Because this study aims to break new ground in terms of running at high spatial resolution over a large domain to produce better SWE estimates, our framework needs to use a high-resolution SWE dataset for validation. The one dataset that currently exists, and which can be exploited effectively for that use, is the NASA Airborne Snow Observatory SWE data.

The ASO is a coupled scanning lidar and imaging spectrometer observational airborne platform (Painter

et al. 2016). It retrieves snow depth directly by flying over the region of interest during the snow-free season to capture the baseline topography, and then again when snow is present to measure the snow surface elevation. Differencing retrievals when snow is melted away and when snow is accumulated determines the snowpack depth. Using spatially distributed, physically based energy and mass balance models that are constrained by measured densities at the time of flight to estimate snow density, SWE can be deduced. ASO SWE product is available at 50-m spatial resolution, with a 5%–13% uncertainty, which is based on measured snow depth and modeled snow density uncertainties. ASO acquisitions are roughly weekly (weather permitting) during the ablation period, and only one snow-free retrieval is necessary. ASO can measure SWE under vegetation canopy.

In this study, we use the ASO SWE data product from several basins, including the Tuolumne basin in the Sierra Nevada (California), Uncompahgre basin in the Rocky Mountains (Colorado), and Olympic Peninsula in Washington State. Each basin has a varying number of observations and has different observational dates. Table 1 summarizes the available dates that were used in this study. The Tuolumne basin was the first basin observed by ASO, and therefore has the longest record, starting in 2013 and available each year to the present. This longer record is useful for assessing modeled SWE during a variety of snow years, including the lowest snow on record for the area (WY 2015), a roughly average snow year (WY 2016) and one of the highest snow years on record (WY 2017).

To have a more direct comparison between ASO and modeled SWE, ASO data are reprojected from the native transverse Mercator (m) projection to VIC's geographic coordinate system (latitude–longitude; $^\circ$), and regridded from the original 50-m resolution to the model's 1.75 km by aggregating ASO pixels within each VIC grid using area averaging. The modeled data are trimmed to the ASO basin shape (e.g., see Fig. 5).

3) IN SITU SNOW MEASUREMENTS (SNOTEL)

Previous modeling studies have used point measurements for validation (e.g., Guan et al. 2013; Rittger et al. 2016), so we include them here for completeness. However, we note the limitations that come with comparing them against a spatially distributed dataset. In situ measurements such as snow pillows do not explicitly or implicitly capture SWE spatial variability because of their location, or lack thereof—snow pillows do not exist at the highest elevations and are often placed on flat ground, which may not represent the surrounding complex terrain (Dozier et al. 2016). These types of observations also

TABLE 1. ASO data description, including basins and dates used in this study. Values in parentheses in the table header indicate total number of retrievals for each basin.

	Tuolumne, CA (48)	Olympic Peninsula, WA (2)	Uncompahgre basin, CO (2)
WY 2013	3 Apr, 29 Apr, 3 May, 25 May, 1 Jun, 8 Jun		
WY 2014	23 Mar, 7 Apr, 13 Apr, 20 Apr, 28 Apr, 2 May, 11 May, 17 May, 27 May, 31 May, 5 Jun		
WY 2015	18 Feb, 6 Mar, 25 Mar, 3 Apr, 9 Apr, 15 Apr, 27 Apr, 1 May, 28 May, 8 Jun		20 Apr, 3 Jun
WY 2016	26 Mar, 1 Apr, 7 Apr, 16 Apr, 26 Apr, 9 May, 28 May, 7 Jun, 14 Jun, 20 Jun, 25 Jun, 1 Jul, 9 Jul	8 Feb, 29 Mar	
WY 2017	29 Jan, 3 Mar, 1 Apr, 2 May, 4 Jun, 9 Jul, 17 Jul, 27 Jul		

present a scale discrepancy, since a “point” measurement cannot fully capture SWE amount over an area (like the model grid space). In addition, interpolation of such point-based data can have large uncertainties (e.g., Fasnacht et al. 2003; Dozier 2011) or insufficiently explain SWE variability (e.g., Molotch and Bales 2005). Because of these reasons, we focus our validation on the more direct comparison of modeled data to ASO SWE observations, especially when assessing spatial variability, but do include snow pillow data for reference, and because its daily resolution allows the evaluation of the modeled SWE temporal evolution. In this study, we use snow course and snow pillow [i.e., Snowpack Telemetry (SNOTEL)] data from the National Resources Conservation Service and National Water and Climate Center website (https://www.wcc.nrcs.usda.gov/snow/snow_map.html) for the Tuolumne basin, averaging over the available stations within the validation domain.

d. Experimental design

The VIC model is run over the western U.S. domain (see Fig. 3) at a spatial resolution of 1.75 km (3 km²) and a daily time step. This domain is composed of five Hydrologic Unit Code 2 (HUC-2) areas (the regional-level USGS-defined hydrologic units): California, Upper Colorado, Lower Colorado, Great Basin, and Pacific Northwest, covering an area roughly bounded between 31° and 53°N and between 125° and 105°W. A significant portion of this region is characterized by complex mountainous terrain, with three main mountain ranges: the Sierra Nevada, the Cascades, and the Rocky Mountains, each defined by their own climate characteristics. Two datasets are produced over the western United States, a 37-yr open-loop simulation covering WYs from 1981 to present (1 October 1980–1 August 2017) where no observed data are assimilated, and a 17-yr assimilation simulation for WYs from 2001 to present (1 January 2000–1 August 2017), where MODSCAG SCA is assimilated during the 1 January–1 August window only, for

each year (with the rest of the water year being run in open loop; see section 2b for more in-depth discussion of implementation of the assimilation window). In both cases we initialize the model at the start date (1 October 1980 for open loop and 1 January 2000 for DA) and use the restart functionality [at 1 October of each year in the open-loop case, and at the beginning (1 January) and end (1 August) of the assimilation window in the DA case], to provide continuous, consistent energy and water balance simulations.

To validate the snow DA framework employed here, we take advantage of the ASO observational data. Here we focus most of our model validation over the Tuolumne basin in California’s Sierra Nevada during the recent severe drought and recovery the region has experienced, WYs 2013–17 (a note that WY2017 is defined here as 1 January 2016–1 August 2017 due to the timing of forcing data availability, the writing of the study, and time needed to rerun WY2017). The Tuolumne basin is the main source of water for the city of San Francisco and other Bay Area communities and is also a source of hydroelectric power. The watershed reaches elevations as high as 3998 m (~13 100 ft) above sea level and is characterized by granite bedrock and steep slopes in the highest parts of the watershed, as well as alpine meadows and forested areas at mid- and lower elevations. ASO retrievals over the Uncompahgre basin in Colorado’s Rocky Mountains and the Olympic Peninsula in Washington are also used in the validation process, assessing model ability across a variety of regional climates and mountain ranges. The location and general coverage of these basins are marked in Fig. 3d with red boxes.

3. Results

a. Overall results in western United States

Figure 2 shows western U.S. domain-averaged SWE for WYs 2001–17. Comparing the open-loop and assimilation SWE results, in general the assimilation of

FIG. 2. Time series of full western U.S. domain-averaged SWE (mm) for open-loop (blue) and DA (green) model simulations: (a) WYs 2001–09 and (b) WY 2010–May 2017.

MODSCAG fractional snow cover lowers modeled SWE. This pattern is evident across all mountain ranges for springtime (April) modeled SWE (Fig. 3), with the exception of the Sierra Nevada, where some areas see an increase in SWE post-assimilation. This difference is largest over deeper snowpacks, such as those in the Cascade Mountains. The main reason for this is that MODSCAG SCA for a given pixel ranges from zero to a value often less than 100% (as detailed in section 2c). By contrast, modeled snow cover area is essentially binary, either 0% or 100% (due to a shallow snow depletion curve parameterization). (Here we refer to the value of SCA as being mostly binary because it changes from 0%

to 100% relatively quickly; it does not represent a flag of snow cover existing or not existing.) In deeper snow, modeled SCA will be 100%, therefore adjusting VIC snow cover with MODSCAG fractional SCA observations will generally tend to lower post-assimilation SWE estimates. Exceptions to this occur during extreme snow years or late in the melt season. For example, during winter and spring of WY 2015 when drought dominated most of the American West (e.g., <https://www.ncdc.noaa.gov/sotc/drought/201504>), snow cover was lower, and so modeled SCA was mostly 0% (as opposed to 100%), whereas observations varied from 0% to 100%, therefore causing post-assimilation SWE to increase.

FIG. 3. Maps showing western U.S. mean April SWE (mm) for years 2000–17 in the (a) DA case, (b) open-loop case, and (c) DA minus open loop, along with (d) model domain extent and elevation (m). Red boxes in (d) indicated the general location of validation basins.

This kind of impact from assimilating MODSCAG can be further understood by looking at Fig. 4, which shows maps of modeled and observed snow cover for 1 February and 1 May 2013. During the middle of the snow season (e.g., 1 February 2013) when snow coverage tends to be greater, the model shows a value of 100% everywhere there is snow, while MODSCAG SCA varies between 0% and 100% (Figs. 4a,c), therefore leading to a lower post-assimilation modeled SWE. One the other hand, toward the end of the melt season (e.g., 1 May 2013), or when snow coverage is low, the model will have much less snow (Figs. 4b,d), and inclusion of MODSCAG will tend to increase post-assimilation modeled SWE. (The impact to SWE can be seen in Fig. 7a below.) This behavior is further discussed in the next section.

b. California: Tuolumne basin, Sierra Nevada

To evaluate whether assimilating MODSCAG SCA using the LETKF batch smoother method is effective in updating SWE, both open-loop and post-assimilation SWE are compared to ASO-observed SWE. ASO retrieved spatially distributed snow observations during both the lowest snowpack (2015) and one of the largest snowpack (2017) years seen in California, as well as an average water year (2016), allowing for an unusually comprehensive evaluation of model results and snow DA technique across space and time and across a wide range of snow conditions. Figure 5 shows modeled and observed springtime SWE (mm) in the Tuolumne basin (California), averaged over ASO available dates for each given water year (see Table 1), which are mostly during the spring snow ablation period.

FIG. 4. SCA maps from (a),(b) model (VIC open loop) and (c),(d) MODSCAG satellite observations for (left) 1 Feb 2013 and (right) 1 May 2013 over a bounding box surrounding the Tuolumne basin. Units of percent (%) coverage. Lower-left values indicate SCA averaged over the region shown.

In general, the model (open loop) tends to underestimate SWE with respect to ASO observations (Figs. 5, 6), anywhere from a -36% to -92% difference, but assimilation of MODSCAG reduces that bias (Table 2). In drier years such as WYs 2013–15, assimilation of MODSCAG SCA adds snow (Fig. 5), reducing the basin bias from -68% , -92% , and -84% in the open-loop scenario to 0.6% , 25% , and 3% , respectively (Fig. 6). Under these drought conditions, when the snowpack is already low, the open-loop model scenario removes snow too quickly, underestimating peak SWE and reaching complete melt-out early compared to ASO retrievals (Figs. 7a–c). By adjusting the model snow to a higher value, DA causes the model to maintain a relatively deeper pack (compared to open loop) for longer (Figs. 7a–c), more closely matching ASO observations (albeit with a small

positive bias). This DA-model behavior can be seen in Figs. 4b and 4d where modeled and MODSCAG observed spatial SCA are plotted for a melting season date, 1 May 2013. The open-loop modeled SCA is almost zero basinwide, with a 5% average snow cover, while MODSCAG basin-average snow cover is 18% with greater snow distribution. The inclusion of these observed MODSCAG values through assimilation nudges the model closer to observations, as seen in Fig. 7a. This improvement during drier years occurs at all elevations, with DA assimilation increasing SWE magnitude throughout all elevation zones (Fig. 8), and more importantly at higher ones (above 2400 m , i.e., $\sim 8000\text{ ft}$) where most snow tends to exist during drought years. The RMSE metric also reflects an improvement in modeled SWE via DA during these years (Table 2), with the DA

FIG. 5. Modeled and observed springtime SWE (mm) in the Tuolumne basin averaged over ASO-available dates for a given water year: (from top to bottom) WYs 2013–17 for (left) open loop, (center) DA, and (right) ASO observations. Model data were trimmed to the ASO domain. (See Table 1 for exact ASO dates.) The upper right-hand corner indicates number of days included in averaging.

FIG. 6. Difference (mm) and RMSE SWE spatial maps in the Tuolumne basin averaged over ASO-available dates for a given water year: (from top to bottom) WYs 2013–17 for (from left to right) model open loop minus ASO, model DA minus ASO, RMSE (open loop, ASO), and RMSE (DA, ASO). Model data were trimmed to the ASO domain. (See Table 1 for exact ASO dates.) The upper right-hand corner indicates number of days included in averaging, lower-left corner in SWE difference maps indicates basin-average SWE percent difference for model with respect to ASO, and lower-left corner in RMSE maps indicates basin-average RMSE for model with respect to ASO.

scenario having a lower RMSE throughout most of the Tuolumne basin (Fig. 6) and as a basin average, and in particular at mid- to high elevations and along ridgelines. One exception is WY 2015, where DA had a higher basin-average RMSE value compared to open loop, despite the fact that basin-average SWE bias is greatly improved in the DA case. It appears that while DA increases SWE estimates in WY 2015 over open loop, generally improving SWE estimates, in some areas of the basin it increases a bit too much, leading to a slightly higher RMSE compared to open loop (Fig. 6).

In an average year (WY 2016), data assimilation reduces the negative bias (Fig. 6) relative to ASO by 28% (Table 2). During the late melt season (Fig. 7d) the improvement after DA is even greater, the percent

difference being -27% compared to open loop's -96% . The DA allows the model to capture the final melt-out during June 2016 compared to open loop, which removes the snowpack about a month earlier. During peak SWE, however, the model underestimates snow in both scenarios, with the DA case performing slightly better (-55%) than the open loop (-74%) (Fig. 7d). While DA adds some snow throughout all elevation zones (Fig. 8), it is not enough compared to ASO observations. This deficiency is likely due to the snow cover–SWE relationship parameterized in the model through the snow depletion curve. Overall, DA does improve SWE estimates throughout most of the Tuolumne basin compared to the open loop by increasing SWE and decreasing SWE bias and RMSE (Fig. 6, Table 2).

TABLE 2. Model and observed basin-average SWE (mm), SWE difference (model minus ASO), and SWE RMSE (model with respect to ASO) in the Tuolumne basin (TUO), Olympic Peninsula (OP), and Uncompahgre basin (UB), time-averaged for ASO available dates only. Values in parentheses indicate percent difference [(model – ASO)/ASO].

	Model: open loop (OL)	Model: DA	Obs: ASO	D_{OL-ASO} (% difference)	D_{DA-ASO} (% difference)	RMSE (OL, ASO)	RMSE (DA, ASO)
WY 2013 TUO	54	169	168	–114 (–68%)	1 (0.6%)	159	147
WY 2014 TUO	11	178	142	–131 (–92%)	36 (25%)	147	144
WY 2015 TUO	18	98	95	–80 (–84%)	3 (3%)	90	108
WY 2016 TUO	85	199	395	–310 (–78%)	–196 (–50%)	392	318
WY 2017 TUO	559	449	872	–313 (–36%)	–423 (–49%)	537	571
WY 2016 OP	236	187	188	48 (26%)	–0.4 (–0.2%)	228	203
WY 2015 UB	16	35	101	–85 (–84%)	–66 (–65%)	98	92

During an anomalously wet year (WY 2017), open-loop SWE has a –36% change bias (–313 mm), with assimilation of MODSCAG SCA leading to a greater SWE bias, a basinwide –49% difference (–423 mm; Fig. 6; Table 2). This lack of improvement in WY 2017 by the DA scheme is also reflected in the RMSE metrics (Table 2) and can be seen throughout all elevation zones (Fig. 8). While open-loop simulations tend to underestimate SWE with respect to observations (both in situ SNOTEL and ASO) in all years modeled here, particularly during peak snow, the difference in WY2017 is that DA does not lead to an increase in SWE as it does in other WYs (and therefore an improvement in the negative bias); rather, it further increases the negative bias with respect to observations. In this anomalously wet scenario in 2017, during the main part of the snow season (from January to mid-May), the snow depletion curve keeps VIC (open loop) modeled SCA at 100%; in other words, the snow depth is well above the rather shallow critical depth and SCA is then at 100%. On the other hand, MODSCAG SCA varies between 0% and 100% and tends to be less than 100% even when the snowpack is that large, due to vegetation, rocks, and instrument viewing angle lowering the overall pixel SCA (e.g., Figs. 4a,c exemplify this behavior). The difference in SCA between model (open loop) and assimilated data therefore leads to a lower post-DA SWE, both compared to open loop, and to ASO observations. Late in the season, during peak ablation (from mid-May through June) when snow depth is lower, the shallow snow depletion curve causes open-loop modeled SCA to jump to low SCA (near zero), and therefore assimilating MODSCAG leads to a post-DA SWE slightly greater than open loop (e.g., Figs. 4b,d exemplify this behavior), but still not reaching observed ASO values. In other words, during peak accumulation season when the snowpack is deep and snow coverage is relatively large, modeled SCA is already at 100% for a significant period of time (albeit biased low) and DA cannot further increase that, while during melt season DA, despite

increasing SWE over open loop, is simply not able to compensate for the open-loop biases. The model SWE underestimation in this work is consistent with other current SWE model estimates (e.g., Raleigh et al. 2015, Broxton et al. 2016) and likely includes contributions from precipitation forcing biases; given that 2017 was an anomalously wet year, it is possible the precipitation gauges that inform the model forcing under capture the full extent of the snowfall, causing a negative precipitation forcing bias and an underestimated open-loop SWE.

Overall, when compared to spatially distributed ASO SWE observations over the Tuolumne basin in the Sierra Nevada, assimilating observed SCA improves modeled SWE when the snowpack is low, such as during dry years or toward the end of snow season, which is significant given this is a time when snow estimates are most uncertain given the lack of in situ point-based measurements at high elevations. Figure 8 supports this finding, showing DA increases snow at all elevation bands, but particularly important above 2400 m (~8000 ft). Having the ASO spatially distributed SWE data has enabled a more direct evaluation of model and DA performance in a spatial context. Temporally, DA improves correlation with respect to the daily SNOTEL measurements in the Tuolumne for all years, but most notably during WYs 2014–16, with open-loop Pearson correlation coefficients of 0.76, 0.78, and 0.85 and post-assimilation coefficients of 0.87, 0.89, and 0.91, respectively. If we consider the anomaly correlation coefficient, which calculates correlations with respect to climatology, DA improves SWE estimates for WYs 2015 (0.98 for DA and 0.95 in open loop) and 2017 (0.98 for DA and 0.96 in open loop). Temporal anomaly correlation is lower in the DA case for WYs 2013 (0.78 for DA, 0.94 for open loop), 2014 (0.77 for DA, 0.95 for open loop), and 2016 (0.44 for DA, 0.56 for open loop). [We calculated the anomaly correlation coefficients by computing the daily modeled SWE anomaly (for both open loop and DA) with respect to

FIG. 7. SWE (mm) time series for Tuolumne basin for WYs 2013–17: open loop (gray line), DA (blue line), ASO (red triangles), and SNOTEL (green line). Modeled data are averaged over the ASO domain only (see Fig. 4 for spatial coverage). ASO data were regridded to VIC pixel resolution, then averaged over the domain. SNOTEL data are averaged over stations located within Tuolumne basin domain (no interpolation). The two vertical orange lines in plot WY 2013 indicate the two dates (1 Feb, 1 May) for which SCA maps are shown in Fig. 3.

the corresponding monthly modeled climatology, the daily observed (SNOTEL) anomaly with respect to the corresponding monthly observed climatology, then computing the correlation coefficient of these anomalies (e.g., $r(\text{SWE}'_{\text{DA}, \text{WY}2013}, \text{SWE}'_{\text{SNOTEL}, \text{WY}2013})$). We used years 2001–16 to calculate monthly climatologies, since that was the period for which both open-loop and DA SWE were simulated.] It is important to note here that the improvements seen by assimilating MODSCAG are over the VIC model open loop, which perhaps produces suboptimal SWE estimates given that the model was run in water balance mode and not full energy mode. Employing the VIC model in its water balance mode means the skin temperature to air temperature and frozen soil parameterization is turned off, which can affect the accuracy of the snowpack simulation (implicitly via the ground heat flux). Nevertheless, despite these limitations of the open-loop simulations, assimilation of MODSGAC actually overcomes most of these model physics simplifications, thus

adding value and the potential for improvement in estimating SWE.

c. Washington: Olympic Peninsula

Over the Pacific Northwest's Olympic Peninsula, for the two dates in 2016 that ASO data are available, the model has a relatively low bias compared to observed SWE, with the open loop having a domain-average percent difference of 26%, while assimilation decreases that bias to a nearly 0% change (Table 2), albeit the biases have more spatial variability in their sign relative to the other basins studies. The decrease in the RMSE score in the DA case, both spatially throughout the basin (Fig. 9) and as a basin average (Table 2), supports the argument for DA improving SWE simulation in the Olympic Peninsula. Spatially, the model is able to capture the SWE distribution across this complex terrain (Figs. 9a–c). The model–ASO difference maps show an SWE bias pattern corresponding to the rain shadow typical of this region, with the model underestimating SWE in the wetter

FIG. 8. Modeled and observed springtime SWE (mm) in the Tuolumne basin averaged by elevation band (ft), for ASO-available dates for each given water year WY 2013–WY 2017 (noted in upper left corner). SWE sorted by respective (model, ASO) elevation maps.

regions of the peninsula (southwest part), and overestimating in the drier areas (northeast part), to varying degree for open loop and assimilation (Figs. 9d,e). Comparing the model cases in Fig. 9f, DA tends to add snow in the midelevation regions, where snow might be more variable, and removes snow in areas where the snowpack is deeper (higher elevations), that is, where model snow cover tends to be 100% more consistently throughout the season. This behavior is consistent with the analysis over the Tuolumne basin in California, where incorporating observed SCA nudges the model toward a higher SWE value in drier, more variable snow conditions (or places), and toward a lower snow for deeper snowpacks. While it is difficult to draw any robust conclusions about model and DA performance given there are only two ASO measurements, Fig. 11a displays the temporal variability of modeled SWE in the Olympic Peninsula mountain range along with ASO observations, for completeness.

d. Colorado: Uncompahgre basin, Rocky Mountains

As with the other regions of the western United States, the model (open loop) underestimates SWE across the Uncompahgre basin in the Colorado Rockies. However, assimilating observed SCA does improve the spatial distribution of snow, as well as SWE magnitude (Fig. 10), reducing basin bias by 19% and decreasing the RMSE (Table 2). DA increases SWE across the higher model elevation bands in the southern part of the basin sampled

here (Fig. 10f), bringing snow distribution closer to that observed by ASO. This is also reflected in the lower RMSE values seen in that part of the basin in the DA scenario compared to the open loop (Figs. 10g,h). With only two ASO observational dates available over the Uncompahgre, it is difficult to draw any significant conclusions, especially about the temporal evolution of the modeled snow, but as Fig. 11b suggests, DA improves late spring SWE values, keeping a deeper snowpack compared to the open loop. This too is consistent with the previous basins, with DA acting to increase SWE for conditions when model snowpack is lower. Just as we have seen in the Tuolumne basin, toward the end of the melt season, informing the model with observed MODSCAG SCA leads to a relatively deeper snowpack, bringing it closer to ASO SWE measurements (Fig. 11b).

It is also important to note that some of the spatial SWE differences between the model and observations for any of the regions analyzed could be due to differing topographic maps, that is, certain features such as high ridges, deep valleys, or peaks might not line up exactly in the model and ASO data, as well as differing elevation maps, that is, the model does not capture the highest of the elevation zones.

4. Summary and discussion

The capability to estimate time- and space-continuous SWE in mountainous terrain with snowmelt-dominated

FIG. 9. Modeled and observed springtime SWE (mm) in the Olympic Peninsula, WA for WY 2016, averaged over ASO-available dates for given water year: (a) open loop, (b) DA, (c) ASO observations, (d) open loop minus ASO, (e) DA minus ASO, (f) DA minus open loop, (g) RMSE for open loop and ASO, and (h) RMSE for DA and ASO. (See text for exact ASO dates.)

systems is greatly needed, but remains a significant challenge. In situ observations can provide direct SWE information throughout the water year, but fall short on capturing SWE spatial distribution. Observing the snow mountain systems remotely presents some opportunities, but clouds, temporal frequency or spatial resolution, and satellite instrument viewing angle can interfere with reliable, complete snow retrievals. Data assimilation, which combines modeling and observations, is one approach to achieve a spatially and temporally complete SWE state estimate. In this study, we have expanded upon previous efforts by implementing and testing a DA method for long-term snow estimation over a large domain at a high spatial resolution (1.75 km). We employ the VIC hydrologic model and the LETKF as a batch smoother over the western U.S. region and assimilate MODSCAG SCA to obtain SWE estimates, which we then validate against the unique ASO data. This is the first time a snow DA technique is assessed against spatially distributed SWE measurements, which provides a more reliable way to evaluate model performance when it comes to snow distribution.

General conclusions taken from our results indicate that our DA approach performs best under relatively low(er) snow conditions, whether that is a dry year or the end of a melt season. Post-assimilation Tuolumne basin-average model bias with respect to ASO is lowest in 2013 (0.6%) and 2015 (3%), in the midst of the California drought. In general, the DA SWE had a lower bias with respect to ASO than open loop, for all years and all basins analyzed, with the exception of Tuolumne basin in 2017, which is one of the wettest years on record. This different model behavior for low versus high snow conditions is also evident in 2016 over the Tuolumne. During peak SWE and early snowmelt (April–May) the model post-DA underestimates SWE, but at the end of the melt season (June) the data-assimilated SWE matches observations very closely both spatially and temporally, with a basin-average percent difference of -18% . These results are generally speaking consistent with previous Kalman-based ensemble batch smoother (EnBS) studies, where the DA improved SWE estimates significantly over open loop, especially for drier years, and sees a more modest improvement during wet years (Margulis et al. 2015),

FIG. 10. Modeled and observed springtime SWE (mm) in the Uncompahgre basin, CO for WY 2015, averaged over ASO-available dates for given water years for (a) open loop, (b) DA, (c) ASO observations, (d) open loop minus ASO, (e) DA minus ASO, (f) DA minus open loop, (g) RMSE for open loop and ASO, and (h) RMSE for DA and ASO. (See Table 1 for ASO dates.)

when as mentioned earlier, the DA leads to a smaller update when prior model bias is significant. For drier years, the LETKF batch smoother biases with respect to observations in this study are comparable (~ 13 mm) to those seen in the EnBS study [~ 200 mm estimated from Figs. 3–5 in Margulis et al. (2015)]. Of course, it is important to note that there are several differences between these studies aside from the batch smoother used, such as observational datasets used for validation (spatial SWE from ASO in current work and point-based snow pillows in EnBS study), the time period and domain of the simulations (albeit both looked at basins in the Sierra Nevada), model resolution, and assimilated data, which would contribute to some of the differences we see in the respective results. Temporally, we found DA improves SWE estimates and time evolution for all (five) water years when compared to daily snow pillow data, with Pearson correlation coefficients ranging 0.87 (WY 2014) to 0.97 (WY 2017) over Tuolumne. This is comparable to the temporal Pearson correlations of 0.89 achieved by other EnBS studies (e.g., Margulis et al. 2015) when compared to snow

pillow and snow course data in a basin in the Sierra Nevada. In general, assimilation of MODSCAG reduces the RMSE in all the basins we looked at in this study, further supporting the added values of employing such a DA framework and technique, with one exception being the extreme snow years in the Tuolumne basin, WYs 2015 (dry) and 2017 (wet).

By comparison against the spatially distributed ASO observed SWE, we show the model is generally able to correctly capture snow spatial distribution in both open-loop and assimilation cases—at least with respect to elevation, one of the main factors influencing snow distribution. During dry and average snow years, while the open loop correctly reproduces snow distribution with elevation, it fails to capture the SWE magnitude throughout all elevation bands. Constraining the model with MODSCAG SCA improves SWE magnitude estimates during these years. This is an encouraging result given that most in situ snow observations do not capture high-elevation snowpack due to lack of stations at these locations (e.g., Dozier et al. 2016), yet this is where most of the snow exists at the end of the snow season, or

FIG. 11. SWE (mm) time series for select WYs and basins: (a) Olympic Peninsula for WY 2016 and (b) Uncompahgre basin for WY 2015. Open loop (gray line), DA (blue line), and ASO (red triangles) are shown. Modeled data are averaged over the ASO domain boundary box. ASO data were regridded to VIC pixel resolution and averaged over the domain.

during drought years. In abnormally wet years, when there is full snow coverage for an extended amount of time, and when the snowpack is deep, assimilation acts to increase the bias by lowering SWE. Our results suggest that this is because snow cover parameterization in the model allows for snow cover to have values of either 0% or 100%, while MODSCAG snow cover, being a fractional SCA product that accounts for vegetation and soil, has snow cover values ranging from 0% to 100%, most often not reaching or staying at full coverage. Thus, during dry periods when the snowpack

is smaller, MODSCAG SCA is larger than the open-loop modeled snow cover, leading to a greater SWE post-assimilation, whereas during wet years like 2017, modeled snow is at 100% coverage, and MODSCAG SCA would cause a decrease in post-assimilation SWE. It is important to note that while DA in this case can lead to a decrease in SWE for pixels with deeper snow, there are locations where DA appropriately adds snow where snow may have not existed at all in open-loop simulation, or there was too little snow. An example of this is the 2016 Olympic Peninsula snowpack, decreasing

the model bias with respect to ASO. The same is true in Colorado's Uncompahgre basin, where assimilation of MODSCAG SCA resulted in a lowering of the negative model bias by increasing SWE for most grids in the domain (Fig. 10f).

Looking at the results for the western United States, DA generally tends to lower SWE compared to open loop, which at first might appear contradictory. However, these are consistent with the validation basins results. Areas that have deeper, more consistently snow-covered mountains (e.g., Cascade Mountains, northern U.S. Rocky Mountains) see a reduction in SWE after DA, given the model versus MODSCAG SCA behavior explained above (binary vs fractional). However, there are regions or times where DA adds snow, for example in the Sierra Nevada in April (e.g., Fig. 3, blue areas), or during dry years or late in the season (e.g., Fig. 2, years 2009, 2013, 2015, and 2016), consistent with ASO basin validation results. Because we validate with ASO, which is mostly available over the Sierra Nevada, we end up assessing the model and DA behavior over what might be an exception to the western United States as a whole rule, that is, the Sierra Nevada are different from the rest of the western U.S. mountain ranges. As for the western U.S. behavior, the fact that DA underestimates SWE over this area has also been found by other studies, where it was concluded that most reanalysis (DA) datasets have too much snow ablation at near-freezing temperatures (e.g., Broxton et al. 2016).

The data assimilation method implemented and tested in this study shows promise, in particular during critical times such as drought years or at the end of the snow season. However, the model, even prior to DA, shows a negative SWE bias with respect to observations, not capturing peak SWE during average to wet water years. This is likely due to an underestimated precipitation forcing, an inadequate snow cover–SWE parameterization in the model, and/or limitations from running VIC in water balance mode. For future work, improving the representation of the relationship between SWE and SCA (i.e., the snow depletion curve) in the hydrologic model should improve SWE estimates. Furthermore, the different behavior of DA results for low versus high snow conditions, rising in part from the discrepancy between modeled SCA estimation (0% or 100%) and observed fractional snow cover (ranges from 0% to 100%, but often less than 100%), points to the fact that if we want to take full advantage of the fractional SCA information MODSCAG provides (which captures subpixel snow variability and impact of vegetation and rocks), the model needs to be improved to also capture that subpixel snow distribution, even when the model is run at a relatively high

resolution such as $1.75 \text{ km} \times 1.75 \text{ km}$. Another improvement in post-assimilation SWE estimates could come from adjusting the assimilation window based on length of snow season. In other words, instead of performing assimilation during 1 January–1 August for each water year, the end date for assimilation should be more dynamic, ending at about the time of the last snow.

Despite the potential shortcomings due to the VIC model setup (e.g., water balance mode instead of full energy; precipitation input) or parameterization (snow depletion curve), the DA method implemented and tested in this study shows promise in overcoming some of these limitations and improving estimated SWE using statistical methods, in particular during drier years or at higher elevations, that is, in situations when snowpack would be lower. Here we have described developing a modeling and DA framework that is suitable for high spatial and temporal resolution and over a large spatial extent while remaining computationally effective, which meant making some perhaps less than ideal choices when it comes to snow and surface energy and physics modeling (e.g., by running the VIC hydrologic model in water balance mode as opposed to full energy). To improve upon the promising framework and resulting snow estimates shown in this study, model settings such as SDC, precipitation input, and water versus energy balance mode need to be further tested through sensitivity studies in future efforts, and adjusted accordingly.

Connecting SCA and SWE, in particular when assimilating SCA to update SWE (i.e., the state estimate is not the same as the observed assimilated estimate), remains a challenge because the relationship between these two variables is not well represented in models (nor is it well defined in the observations, as the processes controlling SCA and SWE are different). This highlights the need to assimilate observed SWE where present in order to compute SWE estimates that are continuous in space and time over continental scale. The modeling and DA framework developed here for the western United States provides a path forward for assimilating such data, at a high spatial resolution, and continuous in time. A next step would be to assimilate SWE or snow depth measurements from the ASO program, which is expanding to the entire Sierra Nevada of California and into the Colorado River basin and Rio Grande basin. The ultimate goal is to assimilate remotely sensed SWE from spaceborne instruments, which would cover a greater area than airborne instruments can, and ASO can serve as pathfinding for spaceborne missions with the promise of its operational version providing robust calibration of the spaceborne retrievals.

Acknowledgments. This work was supported by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the U.S. National Aeronautics and Space Administration. We thank the three anonymous reviewers for their constructive feedback and suggestions, which improved this paper.

REFERENCES

- Andreadis, K., and D. P. Lettenmaier, 2006: Assimilating remotely sensed snow observations into a macroscale hydrology model. *Adv. Water Resour.*, **29**, 872–886, <https://doi.org/10.1016/j.advwatres.2005.08.004>.
- , P. Storck, and D. P. Lettenmaier, 2009: Modeling snow accumulation and ablation processes in forested environments. *Water Resour. Res.*, **45**, W05429, <https://doi.org/10.1029/2008WR007042>.
- Andreadis, K. M., and Coauthors, 2017: The Regional Hydrologic Extremes Assessment System: A software framework for hydrologic modeling and data assimilation. *PLOS ONE*, **12**, e0176506, <https://doi.org/10.1371/journal.pone.0176506>.
- Bair, E. H., K. Rittger, R. E. Davis, T. H. Painter, and J. Dozier, 2016: Validating reconstruction of snow water equivalent in California's Sierra Nevada using measurements from the NASA Airborne Snow Observatory. *Water Resour. Res.*, **52**, 8437–8460, <https://doi.org/10.1002/2016WR018704>.
- 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, <https://doi.org/10.1038/nature04141>.
- Barrett, A., 2003: National Operational Hydrologic Remote Sensing Center Snow Data Assimilation System (SNODAS) products at NSIDC. Spec. Rep. 11, 19 pp., https://nsidc.org/sites/nsidc.org/files/files/nsidc_special_report_11.pdf.
- Broxton, P. D., X. Zeng, and N. Dawson, 2016: Why do global reanalyses and land data assimilation products underestimate snow water equivalent? *J. Hydrometeorol.*, **17**, 2743–2761, <https://doi.org/10.1175/JHM-D-16-0056.1>.
- Cherkauer, K. A., and D. P. Lettenmaier, 2003: Simulation of spatial variability in snow and frozen soil. *J. Geophys. Res.*, **108**, 8858, <https://doi.org/10.1029/2003JD003575>.
- , L. C. Bowling, and D. P. Lettenmaier, 2003: Variable infiltration capacity cold land process model updates. *Global Planet. Change*, **38**, 151–159, [https://doi.org/10.1016/S0921-8181\(03\)00025-0](https://doi.org/10.1016/S0921-8181(03)00025-0).
- 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, <https://doi.org/10.1029/2011WR010745>.
- De Lannoy, G. J. M., R. H. Reichle, K. R. Arsenault, P. R. Houser, S. Kumar, N. E. C. Verhoest, and V. R. N. Pauwels, 2012: Multiscale assimilation of Advanced Microwave Scanning Radiometer–EOS snow water equivalent and Moderate Resolution Imaging Spectroradiometer snow cover fraction observations in northern Colorado. *Water Resour. Res.*, **48**, W01522, <https://doi.org/10.1029/2011WR010588>.
- Dettinger, M. D., K. Redmond, and D. Cayan, 2004: Winter orographic precipitation ratios in the Sierra Nevada—Large-scale atmospheric circulations and hydrologic consequences. *J. Hydrometeorol.*, **5**, 1102–1116, <https://doi.org/10.1175/JHM-390.1>.
- Dozier, J., 2011: Mountain hydrology, snow color, and the fourth paradigm. *Eos, Trans. Amer. Geophys. Union*, **92**, 373–375, <https://doi.org/10.1029/2011EO430001>.
- , E. H. Bair, and R. E. Davis, 2016: Estimating the spatial distribution of snow water equivalent in the world's mountains. *Wiley Interdiscip. Rev.: Water*, **3**, 461–474, <https://doi.org/10.1002/wat2.1140>.
- Dunne, S., and D. Entekhabi, 2005: An ensemble-based reanalysis approach to land data assimilation. *Water Resour. Res.*, **41**, W02013, <https://doi.org/10.1029/2004WR003449>.
- 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, <https://doi.org/10.1029/2006JD008067>.
- , N. P. Molotch, and S. A. Margulis, 2008: A Bayesian approach to snow water equivalent reconstruction. *J. Geophys. Res.*, **113**, D20117, <https://doi.org/10.1029/2008JD009894>.
- Evensen, G., 1994: Sequential data assimilations with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162, <https://doi.org/10.1029/94JC00572>.
- , 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367, <https://doi.org/10.1007/s10236-003-0036-9>.
- 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, <https://doi.org/10.1029/2002WR001512>.
- Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>.
- Giroto, M., G. Cortes, 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, <https://doi.org/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, <https://doi.org/10.1002/hyp.9887>.
- Global Modeling and Assimilation Office, 2015: MERRA-2 tavg1_2d_slv_Nx: 2d,1-hourly, time-averaged, single-level, assimilation, single-level diagnostics V5.12.4. Goddard Earth Sciences Data and Information Services Center, accessed 15 February 2019, <https://doi.org/10.5067/VJAFPLI1CSIV>.
- 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, <https://doi.org/10.1002/wrcr.20387>.
- Hamill, T. M., 2006: Ensemble-based atmospheric data assimilation. *Predictability of Weather and Climate*, T. N. Palmer and R. Hagedorn, Eds., Cambridge University Press, 712 pp.
- Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126, <https://doi.org/10.1016/j.physd.2006.11.008>.
- Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83**, 1631–1644, <https://doi.org/10.1175/BAMS-83-11-1631>.
- Li, D., M. Durand, and S. A. Margulis, 2012: Potential for hydrologic characterization of deep mountain snowpack via passive microwave remote sensing in the Kern River basin, Sierra Nevada, USA. *Remote Sens. Environ.*, **125**, 34–48, <https://doi.org/10.1016/j.rse.2012.06.027>.

- Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges, 1994: A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *J. Geophys. Res.*, **99**, 14 415–14 428, <https://doi.org/10.1029/94JD00483>.
- Liston, G. E., R. A. Pielke, and E. M. Greene, 1999: Improving first-order snow-related deficiencies in a regional climate model. *J. Geophys. Res.*, **104**, 19 559–19 567, <https://doi.org/10.1029/1999JD900055>.
- Liu, J., Z. Li, L. Huang, and B. Tian, 2014: Hemispheric-scale comparison of monthly passive microwave snow water equivalent products. *J. Appl. Remote Sens.*, **8**, 084688, <https://doi.org/10.1117/1.JRS.8.084688>.
- Margulis, S. A., M. Girotto, G. Cortés, and M. Durand, 2015: A particle batch smoother approach to snow water equivalent estimation. *J. Hydrometeorol.*, **16**, 1752–1772, <https://doi.org/10.1175/JHM-D-14-0177.1>.
- , G. Cortés, M. Girotto, and M. Durand, 2016: A Landsat-Era Sierra Nevada Snow Reanalysis (1985–2015). *J. Hydrometeorol.*, **17**, 1203–1221, <https://doi.org/10.1175/JHM-D-15-0177.1>.
- Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360, <https://doi.org/10.1175/BAMS-87-3-343>.
- Molotch, N. P., and R. C. Bales, 2005: Scaling snow observations from the point to the grid element: Implications for observation network design. *Water Resour. Res.*, **41**, W11421, <https://doi.org/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, <https://doi.org/10.1016/j.advwatres.2008.07.017>.
- Ott, E., and Coauthors, 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, **56A**, 415–428, <https://doi.org/10.3402/tellusa.v56i5.14462>.
- 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, <https://doi.org/10.1016/j.rse.2009.01.001>.
- , and Coauthors, 2016: The Airborne Snow Observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo. *Remote Sens. Environ.*, **184**, 139–152, <https://doi.org/10.1016/j.rse.2016.06.018>.
- Raleigh, M. S., J. D. Lundquist, and M. P. Clark, 2015: Exploring the impact of forcing error characteristics on physically based snow simulations within a global sensitivity analysis framework. *Hydrol. Earth Syst. Sci.*, **19**, 3153–3179, <https://doi.org/10.5194/hess-19-3153-2015>.
- Rienecker, M. M., and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, **24**, 3624–3648, <https://doi.org/10.1175/JCLI-D-11-00015.1>.
- Rittger, K., 2012: Spatial estimates of snow water equivalent in the Sierra Nevada. Ph.D. thesis, University of California, Santa Barbara, 245 pp.
- , T. H. Painter, and J. Dozier, 2013: Assessment of methods for mapping snow cover from MODIS. *Adv. Water Resour.*, **51**, 367–380, <https://doi.org/10.1016/j.advwatres.2012.03.002>.
- , E. H. Bair, A. Kahl, and J. Dozier, 2016: Spatial estimates of snow water equivalent from reconstruction. *Adv. Water Resour.*, **94**, 345–363, <https://doi.org/10.1016/j.advwatres.2016.05.015>.
- Rodell, M., and P. R. Houser, 2004: Updating a land surface model with MODIS-derived snow cover. *J. Hydrometeorol.*, **5**, 1064–1075, <https://doi.org/10.1175/JHM-395.1>.
- Schneider, D., and N. P. Molotch, 2016: Real-time estimation of snow water equivalent in the upper Colorado River basin using MODIS-based SWE reconstructions and SNOTEL data. *Water Resour. Res.*, **52**, 7892–7910, <https://doi.org/10.1002/2016WR019067>.
- Slater, A., and M. P. Clark, 2006: Snow data assimilation via an ensemble Kalman filter. *J. Hydrometeorol.*, **7**, 478–493, <https://doi.org/10.1175/JHM505.1>.
- Takala, M., K. Luojus, J. Pulliainen, C. Derksen, J. Lemmetyinen, J.-P. Kärnä, J. Koskinen, and B. Bojkov, 2011: Estimating Northern Hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements. *Remote Sens. Environ.*, **115**, 3517–3529, <https://doi.org/10.1016/j.rse.2011.08.014>.
- 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, <https://doi.org/10.1109/TGRS.2009.2036910>.
- Westerling, A. L., H. G. Hidalgo, D. R. Cayan, and T. W. Swetnam, 2006: Warming and earlier spring increase western U.S. forest wildfire activity. *Science*, **313**, 940–943, <https://doi.org/10.1126/science.1128834>.
- Wood, E. F., and Coauthors, 2011: Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water. *Water Resour. Res.*, **47**, W05301, <https://doi.org/10.1029/2010WR010090>.