

Water Resources Research®

RESEARCH ARTICLE

10.1029/2021WR030563

The Challenges of Simulating SWE Beneath Forest Canopies are Reduced by Data Assimilation of Snow Depth

Eric J. Smyth¹ , Mark S. Raleigh² , and Eric E. Small¹ 

¹Department of Geological Sciences, University of Colorado, Boulder, CO, USA, ²College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR, USA

Special Section:

Advances in remote sensing, measurement, and simulation of seasonal snow

Key Points:

- Data assimilation improves modeled snow depth, density, and SWE at forested sites in two contrasting climates and over multiple years
- Assimilation reduces apparent model output sensitivity to forest canopy parameters that are difficult to measure or calibrate
- Assimilation increases estimated precipitation in the forest relative to the open, likely due to excessive sublimation of intercepted snow

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

E. J. Smyth,
eric.smyth@colorado.edu

Citation:

Smyth, E. J., Raleigh, M. S., & Small, E. E. (2022). The challenges of simulating SWE beneath forest canopies are reduced by data assimilation of snow depth. *Water Resources Research*, 58, e2021WR030563. <https://doi.org/10.1029/2021WR030563>

Received 6 JUN 2021
Accepted 23 FEB 2022

Abstract Intermittent snow depth observations can be leveraged with data assimilation (DA) to improve model estimates of snow water equivalent (SWE) at the point scale. A key consideration for scaling assimilation to the basin scale is its performance at forested locations—where canopy-snow interactions affect snow accumulation and melt, yet are difficult to model and parameterize. We implement a particle filter (PF) technique to assimilate intermittent snow depth observations into the Flexible Snow Model, and validate model outputs against snow density and SWE measurements across adjacent forest and open sites, at two locations with different climates and forest structures. Assimilation reduces snow depth error by 70%–90%, density error by 5%–30%, and SWE error by 50%–70% at forest locations (relative to controls). The PF simulates the seasonal evolution of the snowpack under forest canopy, including where interception losses decrease SWE in the forest during accumulation, and shading reduces melt during the ablation season (relative to open sites). Model outputs are sensitive to canopy-related parameters, but DA reduces the range in snow depth and SWE estimates resulting from variations or uncertainties in these parameters by over 50%. These results demonstrate that the importance of accurately measuring, estimating, or calibrating canopy-related parameters is reduced when snow depth observations are assimilated. Finally, we assimilate remotely sensed snow depth observations at 50 m resolution across the East River Basin, Colorado (1,000 km²). Across elevations, the PF increases estimated precipitation at forested sites by ~15% relative to open sites, likely compensating for excessive sublimation of intercepted snow.

1. Introduction

Accurately monitoring the seasonal accumulation and melt of snow in mountain basins is essential for managing water resources (Bales et al., 2006; Serreze et al., 1999). However, ground-based observations are insufficient to monitor the spatial and temporal distribution of snow water equivalent (SWE) in mountains (Bales et al., 2006). In addition, no satellite remote sensing platform resolves local scale SWE variations across all environments. A promising approach to calculate SWE is through a hybrid remote sensing/modeling platform: measuring snow depth with remote sensing and estimating bulk snow density with models (e.g., Painter et al., 2016). In this framework, intermittent snow depth mapping is possible with several remote sensing approaches, including lidar, radar, and photogrammetry (e.g., Grünewald et al., 2010; Lievens et al., 2019; Marti et al., 2016; Moller et al., 2017; Painter et al., 2016; Vander Jagt et al., 2015) and at increasing spatial scales (Lievens et al., 2021). Data assimilation (DA) has been shown to be an effective tool for converting remotely sensed snow depth to SWE at the time of observation, and for estimating SWE before and after observations (Hedrick et al., 2018; Magnusson et al., 2017; Margulis et al., 2019; Smyth et al., 2019, 2020). This framework holds promise to guide water resource management decisions.

Remote sensing platforms can be used to accurately map snow depth across landscapes, such as underneath forest canopies with lidar (Deems et al., 2013). However, when these lidar snow depth maps have been used to estimate SWE across deep snowpack in mountainous basins, SWE uncertainty is dominated by uncertainty of modeled snow density (Raleigh & Small, 2017). Model simulation errors stem from the quality of forcing data, the choice of process representations, and parameterizations that vary with landscape factors (e.g., forest canopy), which are often poorly or incompletely characterized.

Of these landscape factors, the processes associated with forest canopy are particularly difficult to parameterize and simulate (Helbig et al., 2020; Mazzotti et al., 2020, 2021). Canopy cover influences both the accumulation (e.g., interception and sublimation) and melt (e.g., shading of shortwave radiation) of snow. Canopy also

modulates meteorological conditions (e.g., wind speed, air temperature) and emits longwave radiation toward the snowpack (Conway et al., 2018; Lundquist et al., 2013; Musselman et al., 2008, 2015; Musselman & Pomeroy, 2017; Varhola et al., 2010). In turn, these changes affect bulk snowpack density in forests. To accurately simulate snowpack with a canopy, models either require explicit parameterizations (e.g., “fractional vegetation” in the Factorial Snow Model, see Essery, 2015) or *a priori* adjustments to meteorological inputs (e.g., Snobal, see Marks & Dozier, 1992). The Snow Model Intercomparison Project 2 (SnowMIP2) experiment found that model performance (in estimating SWE) was inconsistent between open and forested sites. In addition, model divergence at forested locations was primarily due to (a) model-to-model differences in partitioning of precipitation via canopy interception and unloading processes and (b) the parameterization of melt, sublimation, and liquid water drainage from canopy snow (Rutter et al., 2009). More recently, Kim et al. (2021) showed high model uncertainty in forested regions across North America for an ensemble of land surface models in the NASA Snow Ensemble Uncertainty Project.

Many snow models have canopy-related parameters, some of which (e.g., canopy height and density) can be measured with remote sensing (e.g., Broxton et al., 2015; Krogh et al., 2020; Moeser et al., 2015). Still, direct measurement of forest canopy parameters is challenging due to limited data availability, while other parameters are not directly measurable due to limited physical meaning (Günther et al., 2020). In addition, parameter calibration is problematic because field observations are sparse (Elder et al., 2009), automated validation data (e.g., Snow Telemetry, SNOTEL) are preferentially collected in forest clearings (e.g., Gleason et al., 2017; Wetlaufer, et al., 2016), and the benefits of calibration at forest sites may not produce accurate estimates in another basin, in subsequent years, or in real time (e.g., Rutter et al., 2009). In the context of combining spatially distributed snow depth data with models to estimate SWE across watersheds, it is essential to refine SWE estimates by improving model simulations at forested sites. One path forward is with DA techniques that incorporate the aforementioned snow depth observations to constrain uncertain model estimates of snow depth, density, and SWE (e.g., Magnusson et al., 2017).

DA of intermittent snow depth observations at the point scale has been demonstrated in previous studies (e.g., Magnusson et al., 2017; Smyth et al., 2019, 2020). However, these experiments were conducted at flat, open sites—without the mass and energy effects introduced by forest canopies. Other prior studies have assimilated intermittent snow depths at the basin scale (including forested pixels) but did not specifically investigate DA performance at forested pixels, and did not validate density or SWE (Hedrick et al., 2018; Margulis et al., 2019). Lv and Pomeroy (2020) used observed snow depths to assimilate canopy interception data, which improved model representation of interception—but again, did not evaluate density or SWE after DA. Cluzet et al. (2021) assimilated spatially distributed snow depths and evaluated model SWE, but the snow depths, SWE, and terrain were generated in a synthetic experiment that did not consider forest effects. Therefore, prior research has not established the utility of assimilating snow depth observations for the purpose of improving modeled density and SWE at forested sites, which is essential for modeling at the basin scale (where snow density and SWE validation is not possible). Previous research has also not demonstrated the efficacy of implementing the particle filter (PF) assimilation technique at the basin scale; here we demonstrate the viability of a basin-wide PF application. Addressing this research gap holds implications for basic research and science applications in watershed management.

2. Objectives

A key consideration for designing regional and global SWE missions (e.g., NASA SnowEx, Decadal Survey) is the robustness of a snow measurement-modeling strategy across landscapes. In the context of combining high resolution, intermittent (e.g., weekly to monthly, for example from lidar), distributed snow depth data with model simulations of bulk density to estimate SWE, it is essential to show that data assimilation can improve the simulated snowpack variations associated with forest canopy, as 30%–50% of the snow in the western U.S. accumulates in forests (Kim et al., 2021; Oswald & Smith, 2014).

We have three objectives critical to mission design that have not been considered previously. First, we quantify the ability of DA to improve model estimates at forested locations (compared to open sites, typical of previous studies) relative to a set of control model runs without any assimilation (the “open loop,” OL). We validate with snow density and SWE data collected approximately monthly (i.e., similar to lidar retrieval intervals) at paired

forest-open sites over multiple years. Second, we characterize the sensitivity of model outputs, with and without DA, to canopy-related parameters that are difficult to measure across a basin (e.g., snow interception capacity). The parameterization of forest-snow interactions is not trivial: canopy parameter values can be challenging to measure, may have no physical meaning, and are difficult to calibrate (see Introduction). With the goal of improving modeled SWE, if DA reduces the apparent sensitivity of model outputs to canopy parameters (by guiding model outputs with observations), the importance of correctly measuring and/or calibrating these parameters is reduced. Third, we implement DA at the basin scale, assimilating distributed snow depth retrievals from airborne lidar, in a mountain basin with ~30% forested pixels. We investigate differences in model errors at forested and open sites in the basin, and attribute possible sources of those errors.

3. Data, Methods, and Experiments

3.1. Spatial Characteristics, Assimilation Data, and Validation Data at Paired Forest-Open Sites

In our first set of experiments (Section 3.6) we assimilate snow depth observations to investigate improvements of SWE estimates at forested locations. While assimilation of remotely sensed, spatially extensive snow depth data (across both open and forested sites) is the end goal (Section 3.6.3), these data are not accompanied by sufficient observations of snow density and SWE to permit effective—a limitation recognized by prior studies (e.g., Margulis et al., 2019). Therefore, for these experiments we need to use field observations of snow depth as a proxy for remotely sensed observations, and the snow depth observations must be collocated with density and SWE data. We selected two field locations each with extensive snow depth, density, and SWE measurements, collected at both forested and open sites, approximately monthly, over multiple water years. The locations are both in mountainous regions, but provide contrasting snowpack regimes (Trujillo & Molotch, 2014), snow density characteristics (Mizukami & Perica, 2008), and forest types.

Snow depth observations collected in situ (details below) are not a perfect proxy for remotely sensed snow depth. The observations cover a smaller spatial support than gridded remotely sensed datasets. Our snow depth measurements have lower measurement error than remote sensing observations. However, in the DA setup we assume that our snow depth observations have uncertainties typical of remote sensing (Section 3.5). Also, the field measurements are approximately monthly—similar to intermittent remotely sensed snow depth data, which could represent any type of snow depth data (e.g., lidar, stereo imaging, and structure from motion) that is not continuous through time.

3.1.1. Colorado Snodgrass Field Site

Snodgrass Mountain (East River Basin, ERB, Colorado) is located in the continental climate zone (Serreze et al., 1999). The site is characterized by low average winter (November–March) precipitation (621 mm) and cold temperatures (−3.7 C, Table S1 in Supporting Information S1). Forested areas are dominated by subalpine conifers, mostly Engelmann spruce and subalpine fir, with some mixed stands of aspen (Crawford et al., 1998). Typical tree height in the study area is 9 m and vegetation area index (VAI) is approximately 4.0 m² m⁻². The study area was established in December 2018 to provide time series of detailed SWE and snow density data from snow pits, manual snow depth measurements, and meteorological data at forested and open locations. The East River is one of the time-series intensive study areas in the 2019–2020 NASA SnowEx campaign (snow.nasa.gov).

Here we use repeated measurements of snow depth, density, and SWE field measurements taken along a 1 km circumference oval-shaped transect at 3,100 m elevation (Figure 1). In water year 2019, snow pit observations were collected at seven open and seven forested sites along the transect. The measurements were made approximately monthly between February and June (five total dates, 90 total snow pits). In water year 2020, data collection focused on three open and three forested sites from the previous year—but with three snow pits logged at each site (Figure 1). Observations were made on four dates between February and May. In cases where multiple snow pits were logged at a given site (i.e., in 2020), the bulk density and SWE values from the individual pits at each site were averaged. Concurrently, 24 snow depth measurements were collected with a manual probe in a 250 m² area centered on each snow pit. These were averaged to produce a time series of (intermittent) snow depth for each site. A comparison between these depth measurements and available airborne lidar snow depth observations (from the Airborne Snow Observatory, Painter et al., 2016) at Snodgrass demonstrates that the manual probe measurements are a good proxy for remotely sensed snow depth observations—with a 0.99 correlation in the open and 0.98 correlation in the forest (Figure S1 in Supporting Information S1).

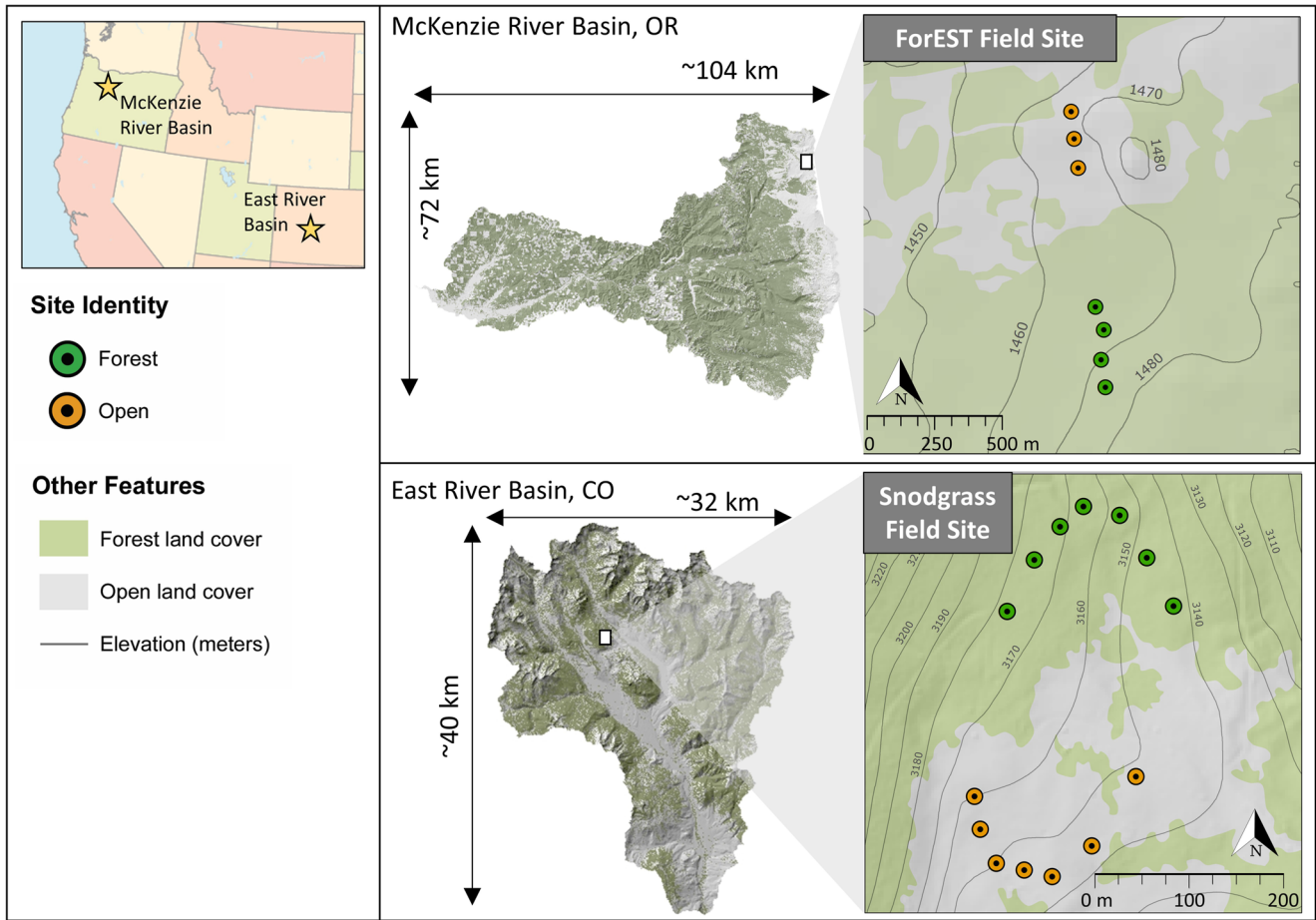


Figure 1. Map of the McKenzie River Basin, with the ForEST field site (top) and East River Basin, with the Snodgrass field site (bottom). Each field site has measurement locations in open and forested areas.

The time series of intermittent snow depth data at each site was used for assimilation into the snow model (described below). The snow depth, density, and SWE data at each site (seven open, seven forested in 2019 and three open, three forested in 2020) were all used for validation. We restrict our modeling to these two water years (WY, 2019–2020). Relative to recent years, 2020 was slightly drier than average, and 2019 was wetter, with peak SWE ~40% above the 10-year average at the nearby Butte SNOTEL site.

3.1.2. Oregon ForEST Field Site

The second site is within The Oregon Forest Elevation Snow Transect (ForEST) network (McKenzie River Basin, MRB), in the maritime climate zone, with wetter, warmer winters than Snodgrass (Table S1 in Supporting Information S1). The forest is composed of tall, thick, coniferous trees—dominantly Douglas Fir (Roth & Nolin, 2017), with an average canopy height of 13 m. We performed our analysis at the upper-elevation paired open-forest site in the study area (1,465 m elevation) to minimize complications with the rain-snow transition zone. SWE measurements (federal sampler) are available every 100 m, and snow depth measurements (manual depth probe) every 5 m, along a 900 m transect between open and forested endpoints (Figure 1). Data collection along the transect was conducted monthly during the accumulation period, and then bi-weekly during ablation, between water years 2012–2015 (Roth & Nolin, 2017). Therefore, we restrict our modeling to these 4 years. Based on data at the nearby Hogg Pass SNOTEL site, water year 2012 was wet (peak SWE ~60% above the 10-year average), 2015 was abnormally dry (peak SWE ~85% below average), and the other two years were near-normal.

We restricted the available snow depth measurements to those (100 m spaced) locations that also had SWE observations, to calculate bulk snow density values for validation. We averaged the snow depth, density, and SWE values for those locations along the transect classified by Roth and Nolin (2017) as “forest” (>60% canopy

cover) and “open” (<20% canopy cover) to produce a single time series of (intermittent) snow depth, density, and SWE for each. The generation of these two (forest and open) time series is comparable in methodology to any of the time series from the Snodgrass site, where we averaged several snow depth and SWE measurements at each individual pit location (see Section 3.1.1). The snow depth data were used for assimilation, and all data were used for validation.

3.2. Spatial and Assimilation Data for Basin Scale Modeling

For our basin scale experiment, we implemented the model across the ERB only (see Section 3.1.1 and Figure 1) at a 50 m resolution, in water year 2019. Two airborne lidar snow depth maps were retrieved by the Airborne Snow Observatory (ASO, Painter et al., 2016) on 4/7/19 and 6/10/19. We used these snow depths (the 50 m product) for assimilation and analysis. We also used canopy height and elevation data provided by ASO—and the latter to calculate slope and aspect across the basin.

3.3. Model Forcing Data

Hourly forcing data for the snow model (Section 3.4) were sourced from 1/8-degree data from the North American Land Data Assimilation System Phase 2 (NLDAS-2) grid cells within, and directly bordering, the ERB and MRB, respectively (Xia et al., 2012). The data were then downscaled to 50 m using the MeteIO preprocessing library (Bavay & Egger, 2014). The preprocessor treats each NLDAS cell/data stream as a point source, and then interpolates fields (e.g., precipitation) across the basin, using lapse rates over a supplied digital elevation model (DEM), with further adjustments (e.g., shortwave radiation) for terrain shading, slope, aspect etc. For modeling at the Snodgrass and ForEST sites, the model was forced with the downscaled forcing data associated with the 50 m pixel containing each individual pit location. Although the Snodgrass and ForEST campaigns have local meteorological stations, these were not used for analysis for several reasons. First, using downscaled NLDAS yields consistent input data between the two sites. Second, the local meteorological data were very similar to the downscaled NLDAS forcing, but the local data has gaps that require filling for continuous simulations. Finally, this methodology more closely replicates a current approach to estimate SWE at watershed and larger scales: combining snow depth with modeled density using non-local meteorological forcing (e.g., Margulis et al., 2019).

Digital elevation models were required for downscaling and modeling—these were sourced from ASO (50 m resolution) in the ERB and the 1/3 arc-second 3D Elevation Program (USGS) in the MRB.

3.4. Snow Model

For this study we used the Flexible Snow Model (FSM2; Essery, 2015), a multi-physics energy balance model of snow accumulation and melt. Model physics and configuration options are documented at <https://github.com/RichardEssery/FSM2> (version released 2/5/19). We selected FSM2 for this analysis for its speed and explicit representation of forest canopy processes. It has also been used in recent modeling studies of snow-forest interactions (e.g., Mazzotti et al., 2020, 2021; Moeser et al., 2016). We ran FSM2 at the point scale for all experiments at both sites, and in the ERB at the basin scale (i.e., all pixels are simulated independently). All inputs and outputs were hourly.

We compiled FSM2 with the default physics options, except for the parameterization of snow density, which we set to evolve as a function of overburden compaction (option 2). We set the number of snow layers to 15 (from default of 3) because offline tests show that additional snow layers produce more realistic snow temperatures. Our version of FSM2 has a one-layer canopy module, where the basic canopy structure is defined by its height, VAI, and fractional vegetation (percent of pixel with forest cover). Snow is intercepted as a function of fractional vegetation and the canopy carrying capacity (per unit VAI). Sublimation from the canopy occurs until snow is unloaded, with different time scales for cold (sloughing) and melting (dripping) snow. The sloughed snow is assumed to be the same density as the underlying snowpack. Incident radiation is transmitted through the canopy—where the amount transmitted decreases with increasing VAI and is modified by leaf angle distribution parameters. The canopy also reflects a portion of incoming shortwave radiation, with reflection increasing as more snow is intercepted by the canopy. Diffuse radiation is transmitted based on a fractional sky view parameter

(an upward-looking, hemispherical portion of the sky without canopy). As the fractional sky view increases (i.e., less canopy cover), less longwave radiation is radiated from the canopy to the underlying snowpack.

In the ERB (including at Snodgrass), fractional vegetation for forested sites was derived from a 3 m resolution canopy height data set provided by ASO. We categorized canopy heights above 2 m to classify forested and open pixels at the 3 m scale, then calculated the forested area fraction within each 50 m grid cell (corresponding to the forcing data grid, Section 3.3). At the ForEST site, we allowed FSM2 to estimate fractional vegetation from our literature-derived VAI value. All other parameters (i.e., except those discussed in Sections 3.1 and 3.2) were set to FSM2 default values (Table S3 in Supporting Information S1).

3.5. Data Assimilation Methodology

The DA technique used is the particle filter (PF, Smyth et al., 2019; van Leeuwen, 2009). The PF estimates the true probability density of a state variable (e.g., snow depth) over time by producing an ensemble of model simulations or “particles” (Figures 2a–2c), which are generated with perturbed meteorological forcings and parameter values, to reflect their respective measurement uncertainties (see below). This causes particle estimates to diverge through time, until confronted with an observation (e.g., a snow depth observation, Figure 2a). The PF then assigns a weight to each particle by comparing its state to the observation (i.e., particles with estimates close to the observation get higher weights) while also accounting for observational uncertainty (Figure 2d). To avoid filter degeneracy (van Leeuwen, 2009), the particles are resampled after each assimilation timestep: those with very low weights are eliminated, while particles with high weights are duplicated and advanced to the next time-step with an observation. This is why the range between particles contracts at assimilation points in Figures 2a–2c.

In calculating weights for particles at each assimilation step, the PF considers the uncertainty of the snow depth observation. Although our assimilated snow depths are collected with manual probes, for the purpose of this analysis we assumed an observation uncertainty relevant to snow depth remote sensing platforms, because we treat our snow depth measurements as a proxy for remotely sensed data. Here we assumed observation uncertainty of 10 cm in the open and 20 cm in the forest, in line with the uncertainty of airborne lidar (Deems et al., 2013; Grünwald et al., 2010; Harpold et al., 2014; Painter et al., 2016), as other techniques (e.g., synthetic aperture radar, photogrammetry) do not have demonstrated capacity for snow depth mapping under forest canopies. This assumed snow depth uncertainty is higher than the true measurement uncertainty associated with manual probes. With a higher observation uncertainty, fewer particles are eliminated at each assimilation step, and DA adjustments to model states are less severe.

We calculated the PF “best estimate” of snowpack depth, density, and SWE as the weighted average of all particle simulations over time—with weights generated by the PF at every assimilation timestep and carried backwards in time through the preceding interval (as in Smyth et al., 2020). We also generated an OL control run for comparison at all sites and years—the equivalent of the simulation one would use in the absence of DA (Figures 2a–2c).

We propagated 75 particles between assimilation timesteps, because previous research has established diminishing returns (for reducing model snow depth and SWE errors) when implementing the PF with more than 50–75 particles (Magnusson et al., 2017). We used the stochastic universal (re)sampling algorithm, which Kitagawa (1996) showed to have the lowest resampling noise among several methods (van Leeuwen, 2009). Weighting and resampling do not alter model states—similar to the Particle Batch Smoother (e.g., Margulis et al., 2015), but unlike direct insertion (e.g., Hedrick et al., 2018) and Kalman update approaches. For a more comprehensive description of the particle filter, including its Bayesian motivation and formula derivations, see Arulampalam et al. (2007); Dong et al. (2015); Smyth et al. (2019); and van Leeuwen (2009).

Following procedures from Smyth et al. (2020), we generated ensembles of particles by perturbing the model forcing inputs and by varying a sensitive snow compaction parameter—to represent sources of model uncertainty.

Meteorological measurements are subject to stochastic and systematic errors. To capture stochastic errors, we perturb each particle’s hourly precipitation, radiation, and temperature inputs with additive stochastic noise, pulled from normal distributions with bounds given in Table S2 in Supporting Information S1. To capture systematic errors (e.g., low bias in NLDAS precipitation in mountainous regions), we also apply a systematic bias to each forcing—a single bias perturbation applied to an entire “window” between subsequent assimilation timesteps,

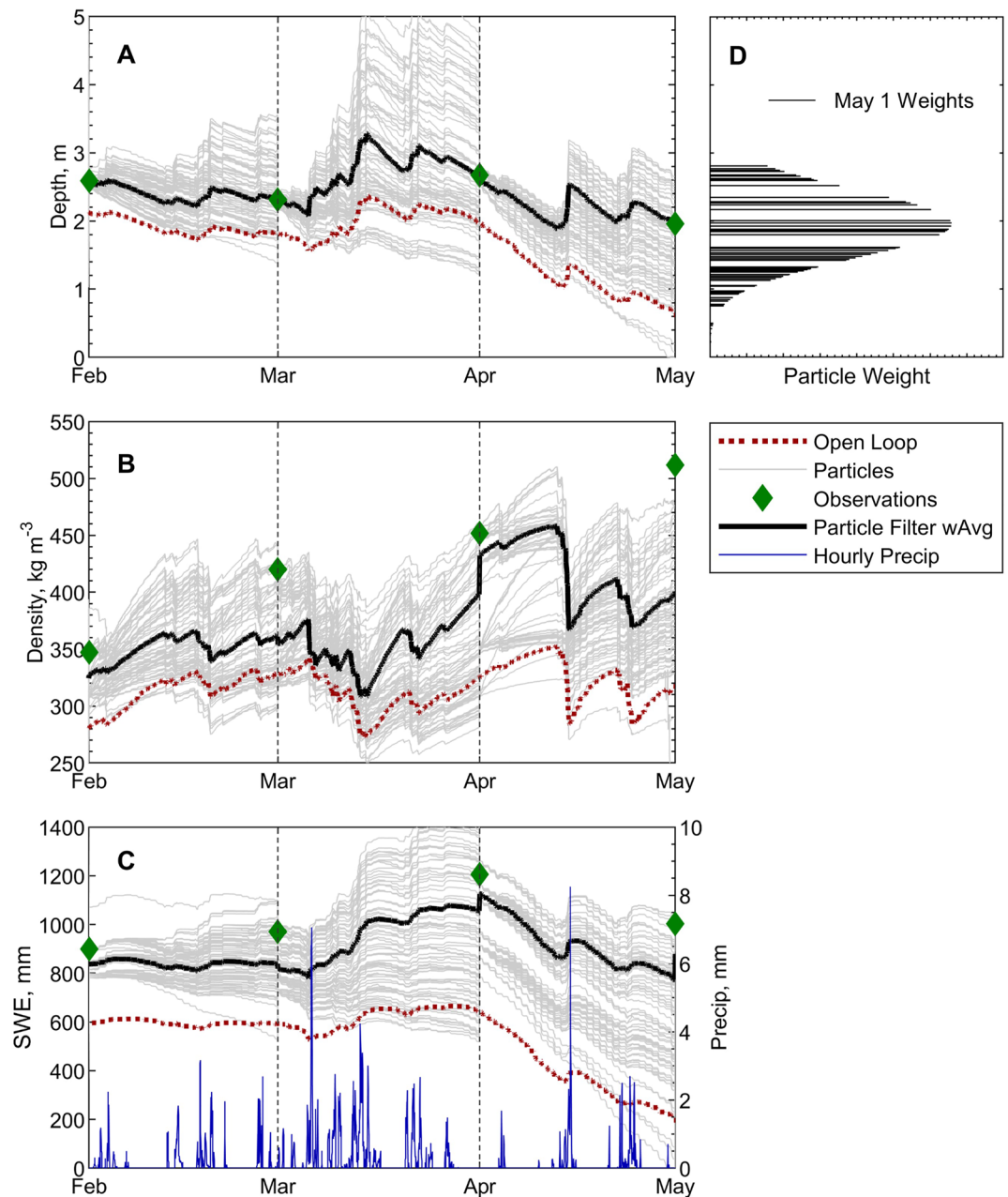


Figure 2. Example implementation of PF improvements to (a) snow depth, (b) density, and (c) snow water equivalent (SWE). The “Open Loop” is the control model run without any assimilation. The particle filter “wAvg” is the weighted average of particle values, with weights generated by the PF. Precipitation values are daily sums. Panel (d) shows the distribution of relative particle weights at the end of the example window, based on the distance between each particle’s snow depth estimate and the observation on May 1.

following ranges applied in Raleigh et al. (2015). This two-step approach to generate a forcing ensemble follows the methodology of previous work (Magnusson et al., 2017; Smyth et al., 2019, 2020).

For example, precipitation is adjusted with an additive, hourly, random error pulled from a normal distribution bound between $\pm 25\%$ of the observed hourly data. A multiplicative window-wide bias termed the “Snowfall Correction Factor” (SCF) is pulled from a uniform distribution between (-100%) and $(+100\%)$. All additive perturbations are positive, that is, the perturbed value is the observed value + perturbation. To bound each set of normally distributed perturbations in a desired range (e.g., $\pm 80 \text{ W/m}^2$ for hourly shortwave radiation), we pull 75

(i.e., one for each particle) values from a standard gaussian distribution (mean of zero, standard deviation of one) and then rescale and translate the set of numbers to have desired maximum and minimum values.

We vary the snow viscosity parameter (called “etab” in FSM2 code) around the commonly used default value of $21 \text{ cm}^3 \text{ g}^{-1}$ (Kojima, 1967) in a uniform range ($\pm 6 \text{ cm}^3 \text{ g}^{-1}$) to represent uncertainty in model parameter values. This single parameter was selected based on a parameter sensitivity analysis, which showed variations in etab accounted for almost all of variations in SWE at an open site when all parameters were varied and no forcing data uncertainty was considered (Smyth et al., 2020).

3.6. Experiments

3.6.1. Error Reductions With DA at Forested and Open Sites

First, we compared the performance of the PF relative to the OL at the Snodgrass and ForEST sites, to evaluate whether the PF reduces model errors at forested sites across two different climates (Objective 1). Specifically, we generated an OL run using default model parameters and downscaled meteorological forcing (Section 3.3). Then, we assimilated intermittent snow depth observations available at each site. To compare the PF and OL, we calculated root mean squared error (RMSE) between model estimates and observations of snow depth, density, and SWE. Thus, validation was based on monthly (or biweekly in OR) data, not continuous observations of SWE and snow depth as in Smyth et al. (2020) or other efforts. Because the same snow depth data were used for assimilation and validation, the validation of snow depth was not independent. In the absence of continuous observations at our sites, we did not evaluate model errors between observation times. However, previous work has demonstrated the PF’s ability to reduce snow depth, density, and SWE errors between assimilation timesteps, with DA of intermittent observations (e.g., Smyth et al., 2019, 2020). We also include a leave-one-out validation experiment, systematically omitting snow depth observations for assimilation and evaluating the open loop and PF estimates on those dates (Figure S3 in Supporting Information S1).

3.6.2. Model Sensitivity With and Without Assimilation

To find the most sensitive canopy-related model parameters, we performed a Sobol’ sensitivity analysis, which can be used to attribute variation in the output of a non-linear model (like FSM2) to variations in model parameters (Günther et al., 2019; Raleigh et al., 2015). Specifically, we selected a single typical snow pit in the forest at Snodgrass for one year, and completed a Sobol’ global sensitivity analysis on the model’s SWE output, varying all canopy-related parameters $\pm 50\%$ from their default values, without any assimilation. We then calculated and compared the total-order sensitivity indices of these parameters: the percent variance in SWE due to variations in each given parameter, including both unique (i.e., first-order) effects and all interactions with other parameters. We identified the most sensitive parameters, and also grouped a collection of (individually, less sensitive) parameters to vary and evaluate together for the remainder of the analysis.

Specifically, the canopy-related parameters we considered were snow-free and snow-covered vegetation albedo (“avg0” and “avgs,” respectively), normal and dense canopy turbulent transfer coefficient (“cveg” and “cden,” respectively), canopy snow capacity per unit VAI (“cvai”), roughness length to canopy height ratio (“rchz”), canopy unloading time scale for cold and melting snow (“tcnc” and “tcnm,” respectively), canopy heat capacity (“canh”), fractional sky view (“fsky”), fractional vegetation (“fveg”), canopy transmissivity (“trcn”), and vegetation area index (“VAI”). Default parameter values and units are detailed in Table S3 in Supporting Information S1.

Next, we performed an experiment to evaluate whether assimilation of intermittent snow depth reduced the apparent sensitivity (explained below) of model outputs to these parameters (Objective 2). The purpose of this test is to demonstrate how the PF can compensate for incorrect or uncertain parameterizations. For each parameter/collection identified in the previous step, we:

1. Ran the model without any assimilation (OL) 20 times at all forested sites and years, varying the parameter/collection in a uniform range $\pm 50\%$ from default values
2. Repeated step 1, but for each of the 20 vegetation parameter sets, snow depth observations at the given site/year were assimilated with the PF. For each of these PF runs, a full set of 75 particles was used to generate a PF best estimate

3. Compared the resulting range (maximum–minimum) between the 20 OL runs and the 20 PF “best estimates” (the weighted average of each particle ensemble) throughout the water year to evaluate the sensitivity of model outputs to these parameters, with and without assimilation

With the OL experiment, we are characterizing the sensitivity of model state estimates (e.g., snow depth) to individual parameters, all else equal. With the PF experiment, we are investigating the apparent sensitivity of the model/assimilation framework to these same parameters. Assimilation does not influence the model itself and does not change the sensitivity of model processes to the given parameters. But, assimilation does influence the range in model state estimates resulting from uncertain parameter values.

Note that in step 2, we ran the PF 20 separate times, each instance with a different value for the canopy-related parameter in question. This is not to be confused with the snow viscosity and meteorological input perturbations that are applied when generating the ensemble of particles for each instance of the PF (Section 3.5). The canopy-related parameters were not perturbed as part of that PF ensemble generation process for two reasons. First, perturbing these canopy-related parameters would obfuscate the influence of each individual parameter we systematically varied in this experiment. Second, offline tests show that excluding these parameter perturbations does not impact the ability of the PF to operate successfully—model errors are similar with and without these perturbations. In the PF setup we aim to capture sources of model uncertainty, and the model is very sensitive to snow viscosity and meteorological forcings (see Smyth et al., 2020 and Figure S2 in Supporting Information S1).

The experiment is designed to compare the impact of uncertain parameters on model simulations, with and without DA. We discuss our arbitrary choice of perturbation range ($\pm 50\%$) in the discussion, and also show that the results do not change with a different assumed range ($\pm 25\%$, Figure S8 in Supporting Information S1).

3.6.3. Implementing the PF at the Basin Scale

Finally, we implemented the PF at the basin scale, assimilating distributed, and remotely sensed snow depth observations. For each pixel in the basin, we ran an OL simulation, and then assimilated both available snow depth measurements with the PF. At 50 m resolution, there were over 300,000 pixels to simulate, and each pixel was modeled separately. Snow density and SWE measurements were not available across the basin for validation, so we evaluate the performance of OL and PF snow depth only. This analysis compares open and forested pixels.

The PF guides model estimates by calculating weights for each particle in the ensemble, and then resampling model states based on those weights. It is also possible to track the parameter values associated with the highest-weighted particles over time. This process is often used for parameter estimation (e.g., Clark & Vrugt, 2006) which we do not perform here. However, in this experiment we do track the snowfall correction factor (SCF, Section 3.5) associated with the best (highest weighted) particles, for each pixel in the basin. Specifically, similar to how we calculate a weighted average of state variables (e.g., snow depth) over time (Section 3.5 and Figure 2), we also calculate a weighted average of the SCF applied to particles, as the PF guides model estimates toward observations. The resulting, best SCF values for each pixel give us an (imperfect) idea of adjustments to precipitation needed to match snow depth observations.

4. Results

4.1. Error Reductions With the PF

First, we compared OL and PF performance relative to in-situ measurements at the Snodgrass and ForEST sites (Objective 1). Sample time series are shown in Figure 3 for the ForEST open and forested sites in WY 2014. We assimilated intermittent snow depth observations (diamonds in Figures 3a and 3d) and then evaluated the accuracy of the model (with and without assimilation) relative to these same snow depth observations—and relative to density and SWE observations (diamonds in Figures 3b–3f). Figure 3 illustrates the PF framework in an average year, and where the canopy affects the seasonal snowpack evolution relative to an adjacent open site. For example, note that the open site melted out by the beginning of May, while the forested site still had >500 mm of SWE at that date.

As expected, DA removed nearly all snow depth errors at the example open site (Figure 3a), reducing average RMSE by 73% (from 0.26 to 0.07 m). In guiding the model toward a more realistic simulation of the snowpack, the PF also reduced snow density RMSE by 16% (168–142 kg/m³, Figure 3b) and SWE RMSE by 57% (125–

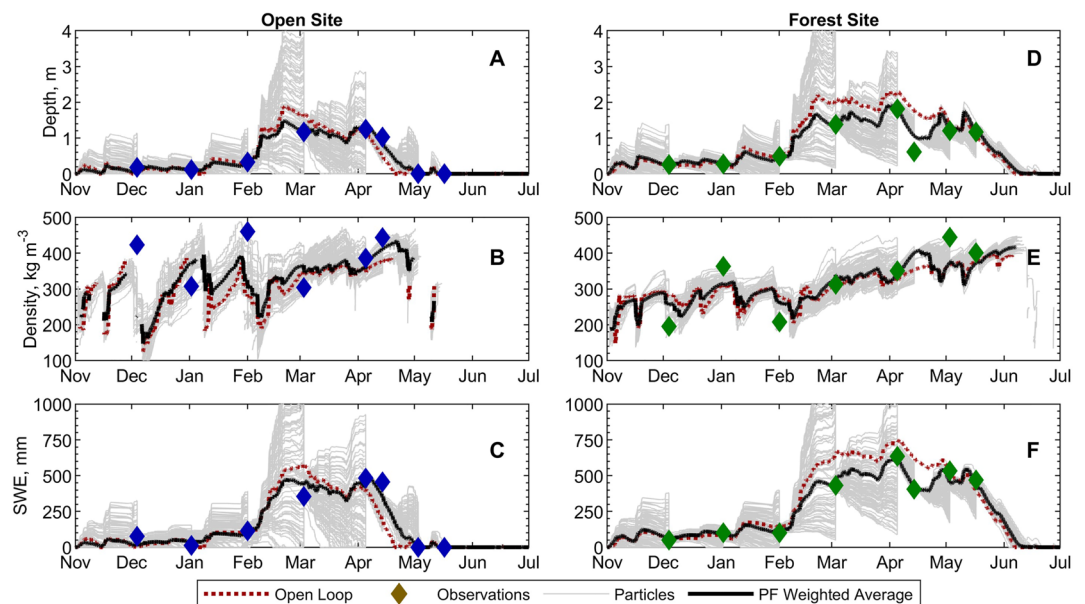


Figure 3. Comparison of observed data, open loop (OL), and particle filter (PF) estimates with assimilation of snow depth observations at the ForEST open (a–c) and forested (d–f) sites in water year 2014. At each site, the PF is able to closely match observed snow depth (a and d), and SWE (c and f)—with more modest improvements to snow density (b and e) relative to the OL.

54 mm, Figure 3c). At the example forested site, assimilation led to similar reductions in error (Figures 3d–3f). Snow depth RMSE was lowered by 68% (from 0.5 to 0.16 m), snow density RMSE by 15% (123–104 kg/m³) and SWE RMSE by 70% (118–35 mm).

Across all years, we found consistent improvements to modeled snow depth, density, and SWE estimates at the ForEST site—at both open and forested locations (Figure 4). Over 4 years at the ForEST transect, the PF reduced snow depth RMSE by 83% and 74% at open and forested areas, respectively (Figure 4a). Snow density error (Figure 4b) was reduced by 4% (open) and 8% (forest). These lesser improvements to modeled density (compared to snow depth improvements) point toward possible deficiencies in the snow model, which DA is unable to overcome—or that DA of snow depth alone is not sufficient to correct density by itself at this site (see discussion in Section 5). Finally, the PF lowered SWE RMSE by 46% in the open and by 55% in the forest (Figure 4c). The larger error reduction at forested sites was mostly due to the relatively poor performance of the OL at those locations—assimilation brought SWE errors in-line with simulations at adjacent open locations.

Results are similar at Snodgrass across our two model years (WY 2019–2020): DA reduced snow depth RMSE by 74% and 93% at open and forested sites, respectively (Figure 4d). Improvements to modeled density were greater than at the Oregon ForEST sites: density RMSE was lowered by 31% in the open and 37% in the forest (Figure 4e). Finally, the PF lowered SWE errors by 52% (open) and 70% (forest)—and although the magnitude of errors in the OL simulations differ between open and forest locations, after assimilation the errors were similar (Figure 4f).

In the leave-one-out validation experiment (evaluating model estimates at dates when snow depth observations were not assimilated), DA reduced snow depth, density, and SWE errors across the forested and open locations at Snodgrass and the ForEST site (Figure S3 in Supporting Information S1). We found smaller improvements to snow depth RMSE, compared to the case where all observations were assimilated (Figure 4)—for example, 18% and 58% snow depth RMSE reduction in the open and forest (respectively) at Snodgrass in the leave-one-out validation, compared to 74% and 93% improvement when all snow depth observations were assimilated. Still, the PF lowered density RMSE by 30%–60% and SWE errors by 15%–60% (Figure S3 in Supporting Information S1).

The full-season average metrics shown in Figure 4 do not reveal how the PF yields a more realistic simulation of forest canopy effects on accumulation and melt. Thus, we focus on example years (Figure 5). During the accumulation season at Snodgrass, the forest canopy intercepts snowfall, which is partially sublimated before

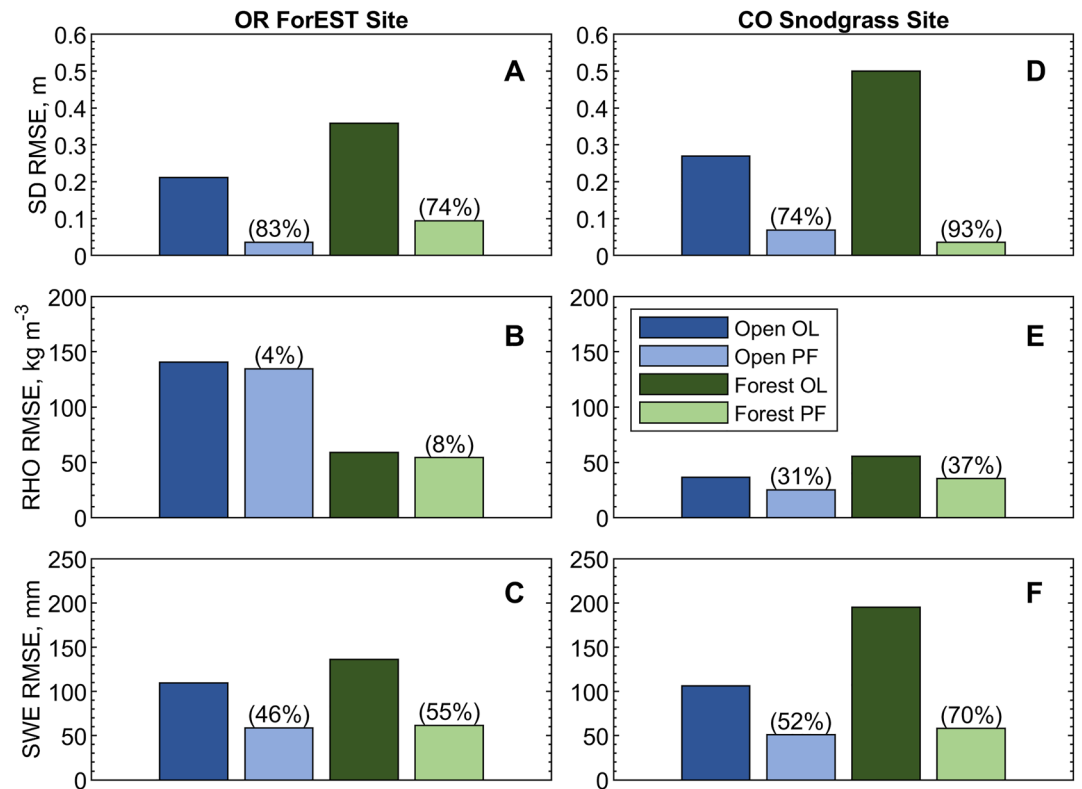


Figure 4. Comparison of open loop (OL) and particle filter (PF) RMSE in snow depth (top row), density (middle row), and SWE (bottom row) estimates across all modeled years at the ForEST (a–c) and Snodgrass (d–f) sites. The blue bars show average RMSE at open sites, and the green bars average errors at forest sites. The darker bars are for OP (no assimilation) errors, and the lighter bars are RMSE after implementing the PF.

unloading. As the amount of total precipitation is similar between the adjacent forest and open plots, this causes peak snow depth and SWE to be lower at the forest sites than in the open. However, during the melt season, the tree canopy shades shortwave radiation, causing the snow to persist later compared to adjacent open sites (Figure S4 in Supporting Information S1). This interchange in the observed magnitude of snow depth and SWE between the open and forest was not predicted by the OL—the model showed consistently lower snow depth and SWE in the forest (Figures 5a and 5b). Clearly, there is a deficiency in the OL model, parameters selected, or input data. After DA, the interchange is simulated (Figures 5c and 5d), so the inclusion of observations via DA compensates for this model deficiency (at least partially). In some years at the ForEST site, the impact of interception was weaker, and there was no observed interchange in the relative magnitude of snow depth/SWE between open and forested locations—Figures 5e–5h shows a case (WY 2013) where observed snow depth and SWE was higher in the forest compared to the open throughout the entire year. Again, the OL model did not capture this trend (Figures 5e and 5f) and incorrectly suggested an interchange from higher snow depth in the open in winter and in the forest in spring. This is another case where the model, parameters, or inputs are deficient—but where PF correctly predicted the observed pattern (Figures 5g and 5h).

4.2. Model Sensitivity to Canopy-Related Parameters

We next evaluate the sensitivity of outputs to the model’s canopy-related parameters by performing a Sobol’ global sensitivity analysis (Figure 6). This was completed in the absence of DA. Results show that the model’s SWE outputs were most sensitive to variations in the fractional sky view (“fsky”) and fractional vegetation (“fveg”) parameters. Fsky accounted for over 50% of modeled SWE variance (including interactions with other parameters) and close to 90% more typically. Beyond these two, we identified a second tier of moderately sensitive parameters (~1%–5% of SWE variance explained): vegetation area index (“VAI”), the ratio of snow roughness length to canopy height (“rchz”), the snow-free vegetation albedo (“avg0”), and the vegetation turbulent transfer

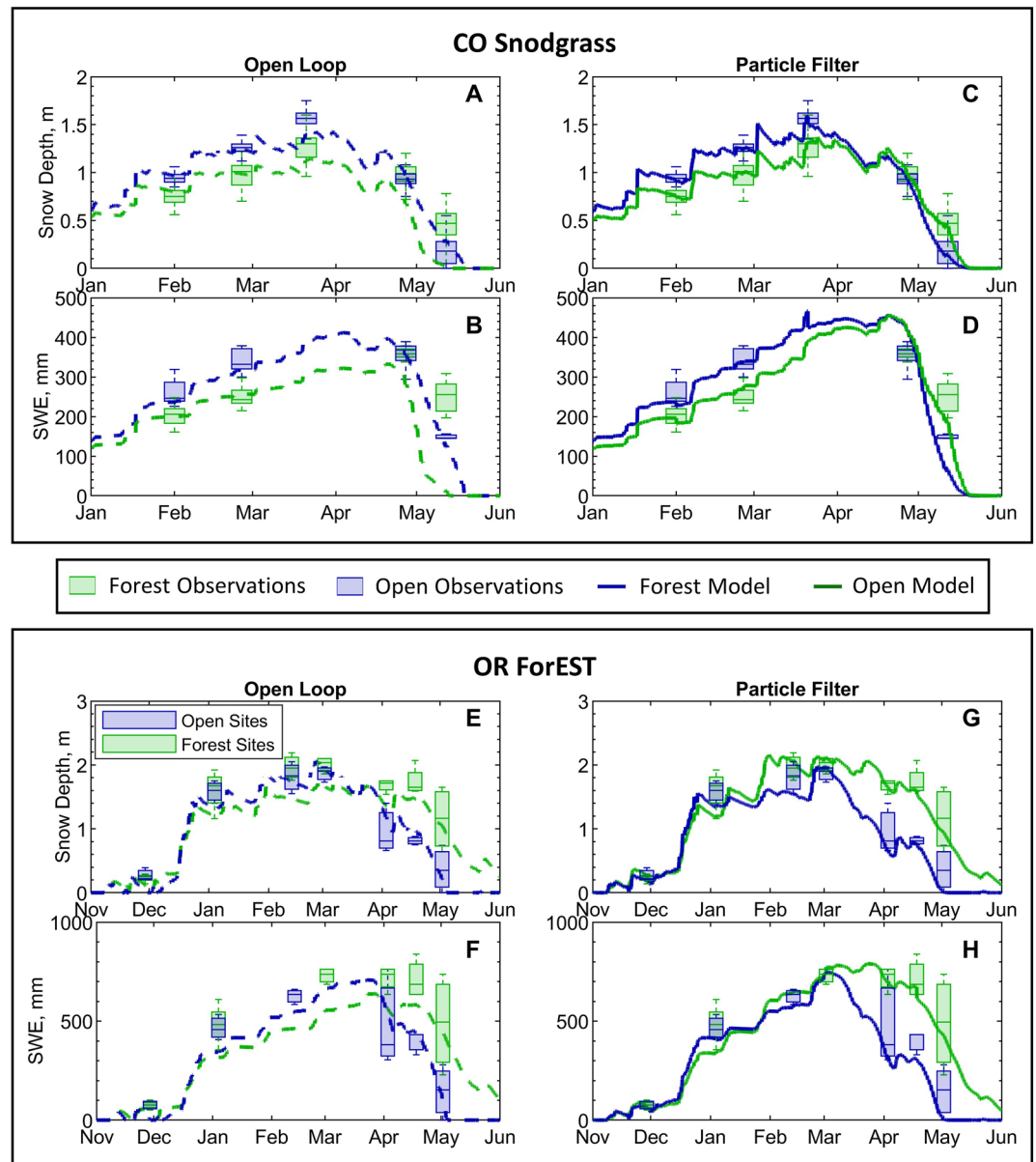


Figure 5. Box-and-whisker plots aggregating snow depth and SWE observations at the Snodgrass site (top section) and ForEST site (bottom section) during example years. The box extent indicates 25th and 75th percentiles, with a median bar within. The whiskers contain all points. At Snodgrass, the observations are from WY 2020, and show a case when the relative magnitudes of snow depth/SWE in the open versus the forest swap during the melt season. The OL model (dotted lines, A-B) does not capture this pattern, but the PF (solid lines, c and d) does. At the ForEST site, the observations are from WY 2013 and demonstrate a case when observed snow depth/SWE is higher in the forest than in the open during the entire year (i.e., no swap in relative magnitudes). Again, the OL does not predict this trend (e and f) but the PF does (g and h).

coefficient (“cveg”). A sensitivity analysis on the model’s snow depth output yields similar results (Figure S5 in Supporting Information S1).

Model SWE is moderately sensitive to the canopy snow capacity per unit VAI (“cvai”) parameter (Figure 6). However, its impact to modeled SWE is limited by the unloading time scales for cold (“tcnc”) and melting (“tcnm”) snow. Hereafter, we group these three parameters—all of which affect the interception and unloading of snow—to better evaluate the sensitivity of model outputs to canopy interception, which is an important canopy process. We call this collection of interception parameters “intcpt.”

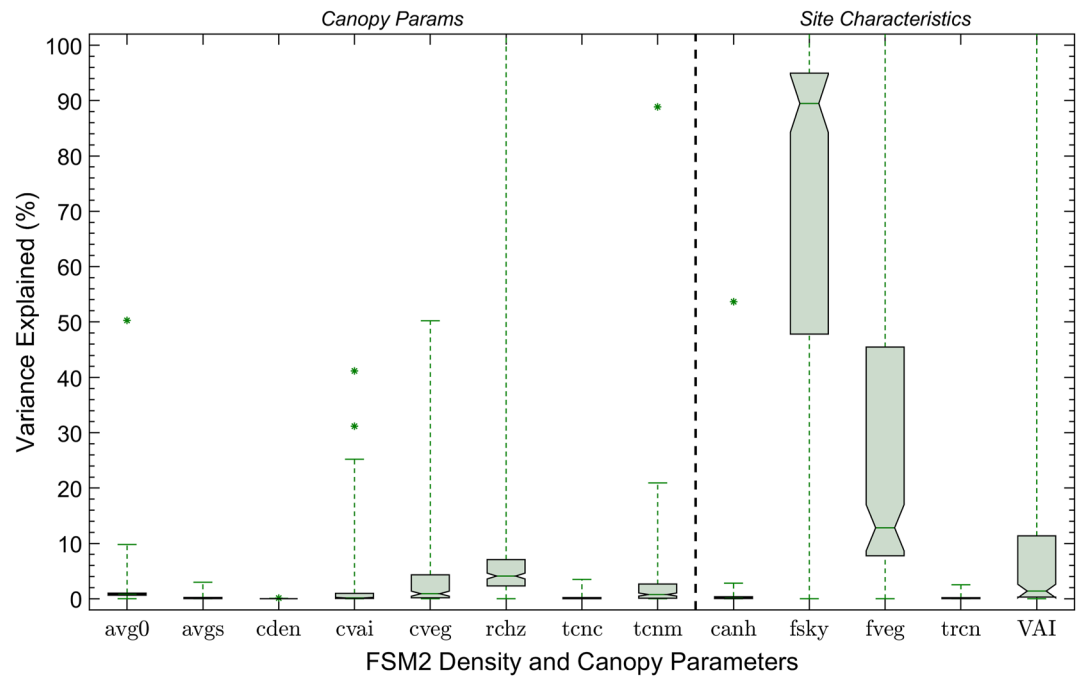


Figure 6. Box-and-whisker plot aggregating the Sobol’ total-order sensitivity values for canopy-related parameters in FSM2. All parameters are varied $\pm 50\%$ from their default values. The plot values represent the percent of modeled SWE variance explained by each parameter, calculated from over 25,000 individual model runs. The percentages (e.g., bar medians) do not add to 100% because the total-order sensitivity indices allow for interactions between parameters. For explanations of parameter name abbreviations in the figure, see Table S3 in Supporting Information S1.

To evaluate whether DA of intermittent snow depths affects the apparent sensitivity of model outputs to these parameters, we varied the parameters one at a time (with the exception of the “intcpt” collection in which the three associated parameters were all varied together in the same direction—for example, a given perturbation would increase or decrease *cvai*, *tcnc*, and *tcnm* at the same time) and evaluated the resulting range between snow depth and SWE estimates throughout the water year—with and without implementing the PF (Objective 2).

Figure 7 shows an example comparison at the Oregon (ForEST) forested site in 2012, when the snow roughness length to canopy height ratio (*rchz*) parameter (and only this parameter) was varied $\pm 50\%$ from its default value of 0.1 (unitless). We highlight the model’s response to this parameter because it is difficult to measure in the field—results are similar for the fractional sky view (*fsky*) and fractional vegetation (*fveg*) parameters (Figures S6 and S7 in Supporting Information S1). The average range between OL snow depth estimates was 0.47 m through the WY, and the average range in SWE estimates was 162 mm (Figures 7a and 7b). For both snow depth and SWE, the model range increased through time. However, after assimilation the range between PF estimates averaged 0.09 m (79% reduction) for snow depth and 52 mm (68% reduction) for SWE (Figures 7c and 7d).

Not only did the PF reduce the apparent sensitivity to the *rchz* parameter, but it also improved performance compared to the OL (validated against observations, diamonds in Figure 7). Across the 20 PF simulations (each with different values for the *rchz* parameter), the average snow depth error was just 0.08 m and the average SWE error was 99 mm. This is an 84% reduction in snow depth error, which is an expected result because we are assimilating snow depth observations. The PF also reduces SWE error by 63% relative to the OL. This example demonstrates that the importance of choosing the best *rchz* parameter via calibration is lessened with the assimilation of snow depth observations.

These results were consistent across other canopy-related parameters in the model: the range between snow depth and SWE estimates was reduced after DA (Figure 8). As identified above, the range between OL runs was highest when varying fractional sky view parameter values. The PF reduced variations in both snow depth (from a 67% variation to a 10% spread between model estimates, Figure 8a) and SWE (from a 62% spread to 23%, Figure 8b) estimates stemming from this parameter. There are a few exceptions: for example, modeled SWE was more

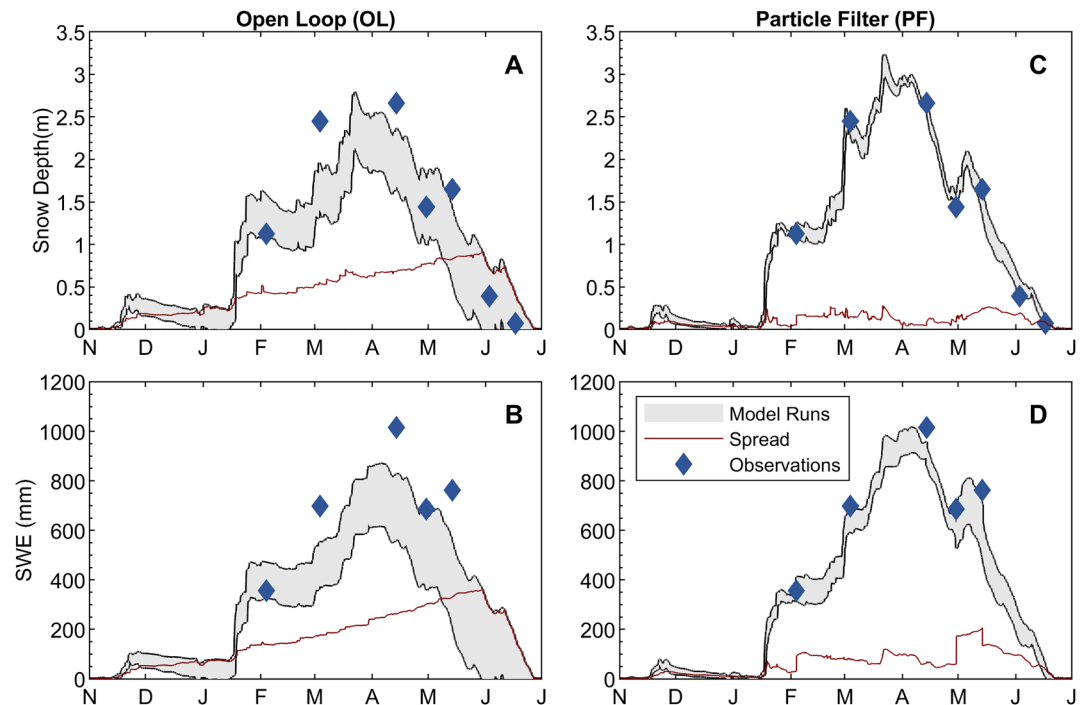


Figure 7. Range of model outputs (maximum-minimum) of snow depth (top row) and snow water equivalent (bottom row) for an example WY (2012) at the Oregon ForEST forested site, generated by varying the “snow roughness length to canopy height” model parameter (“rchz”) $\pm 50\%$ from its default value of 0.1 (unitless). The range between OL model runs (a and b) is much greater than the model range after assimilation of available snow depth observations (c and d).

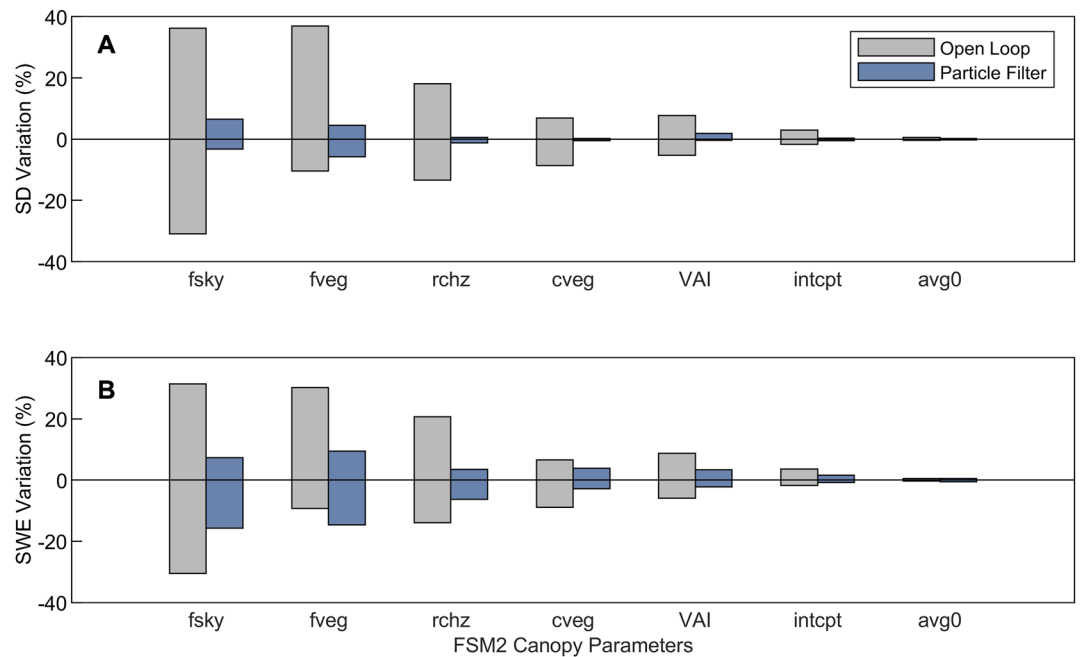


Figure 8. Bar chart comparing the sensitivity of the snow model (FSM2) snow depth (a) and snow water equivalent (SWE) (b) outputs to variations in canopy-related parameters, before (open loop) and after (particle filter) assimilation of snow depth observations. In each case, the given parameter is varied $\pm 50\%$ from its default value, and the output range is calculated at the Oregon ForEST forested site in a typical year. The middle (zero) line is the snow depth/SWE value when all parameters are set to default values.

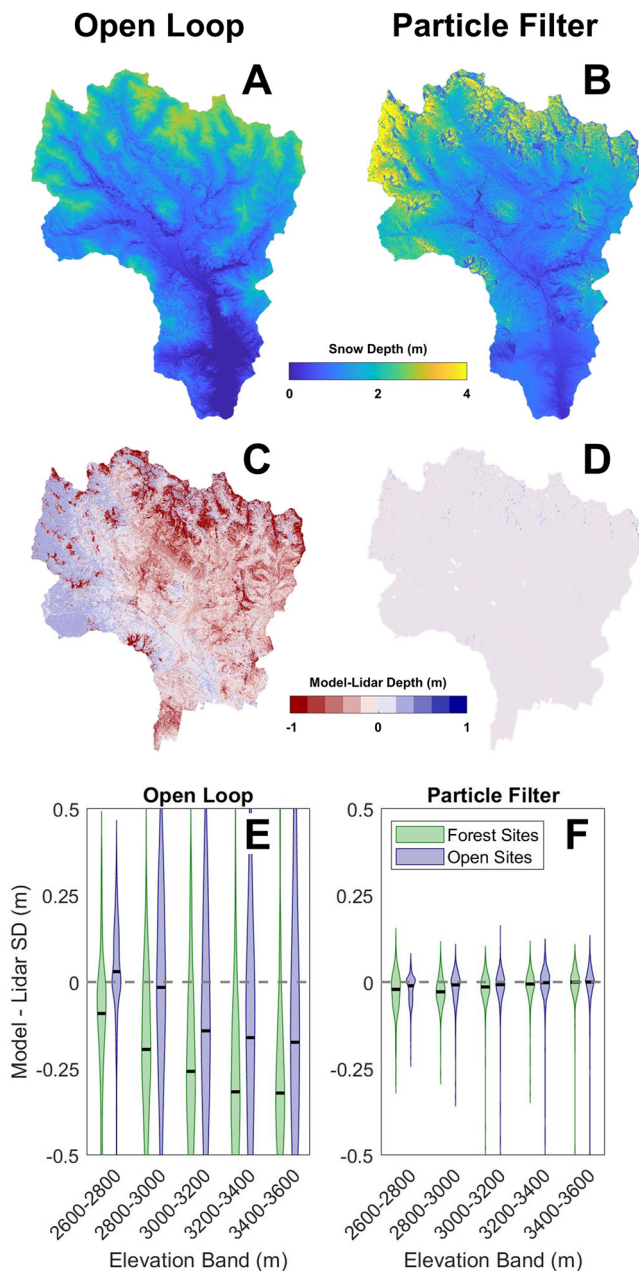


Figure 9. Comparison of modeled snow depth for the open loop (OL; a) and particle filter (PF; b) at 4/7/19, after assimilating lidar snow depth observations. PF estimates are much closer to observations (d) compared to the OL (c). The violin plots of observed-modeled snow depth (e and f) show that the OL snow depth estimates have higher errors in the forest (compared to open locations) and at higher elevations (e), while PF errors are similar across the terrain (f).

sensitive to negative changes in the f_{veg} parameter (i.e., reducing the parameter value from a given default) after DA, compared to the OL. However, even in this case, the total range between SWE estimates (i.e., with both positive and negative perturbations to the parameter) was 15% points lower than the OL. The PF constrained snow depth apparent sensitivity more than SWE because the PF assimilates snow depth—so by design directly forces simulations toward the same values (at least intermittently).

4.3. Model Errors at Forested Versus Open Sites at the Basin Scale

Finally, to assess the performance and viability of the PF at the basin scale, we implemented the OL and the PF at all pixels in the ERB, in water year 2019 (Figure 9). The OL model tends to underestimate snow depth (Figure 9c), and model errors are not consistent across elevations in the basin: the OL greatly underestimates snow depth at the highest elevations (the dark red areas to the North and Northeast of the basin, Figure 9c) and has lower errors at lower elevations (the light blue and light red areas in the South and West). In addition, the OL performs worse at forested pixels—for example, average OL error at 3,400–3,600 m elevation is 0.31 m in the forest and 0.24 m in the open. Overall, on 4/7/19 (the first observation/assimilation date) the OL average snow depth error is -0.25 m in the forest and -0.10 m in the open.

Not surprisingly, after assimilating distributed snow depth observations, the PF reduces snow depth errors across elevations and in both forested and open pixels (Figures 9d and 9f). The PF still slightly underestimates snow depth, with average errors of -0.02 m in the forest and the open. DA lowers errors by increasing precipitation at the highest elevations in the basin and in forested pixels (Figure 10). This is accomplished (in part) with the SCF: in matching snow depth observations in the forest, the best particles generally have SCF values above one (average of 1.2, Figure 10). This indicates that the PF is compensating for precipitation and/or model errors by assigning higher weights to particles with more precipitation. The average SCF is 1.0 at open sites—with SCF values below one at lower elevations. Across open and forest sites, SCF values are highest at the highest elevations in the basin (e.g., 1.2–1.3 at 3,400 to 3,600 m), which is consistent with the elevations where the OL most underestimated snow depth (Figure 9e). Likewise, SCF is higher at forested compared to open pixels (e.g., 1.30 vs. 1.15 at 3,200 to 3,400 m elevation). The difference between SCF values at forested versus open pixels is consistent across elevations (~ 0.15 difference). The difference in SCF values indicates that model errors are not the same in forested versus open locations, and thus are likely due to the simulation of canopy processes.

5. Discussion and Conclusions

We considered three objectives critical to the design of regional and global SWE missions. First, we quantified the ability of DA to improve model estimates at forested locations, and compared to improvements at open locations. Second, we characterized the apparent sensitivity of model outputs, with and without DA, to canopy-related parameters that are difficult to measure across

a basin. Third, we implemented the PF at the basin scale and investigated differences in model errors at forested and open locations across the basin.

The PF improves simulations of snow depth, density, and SWE at both forested and open sites in two disparate climates (Objective 1). Of these three variables, the improvements to modeled snow depth are the greatest (Figure 4). This is partly an artifact of the DA framework—in assimilating snow depth, we expect to improve

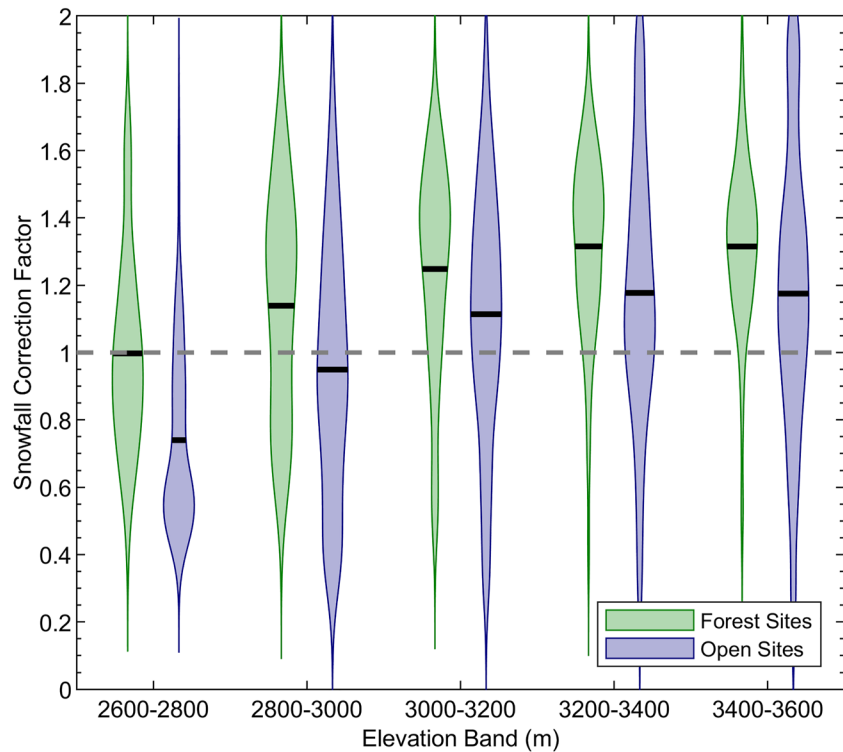


Figure 10. Violin plot comparison of snowfall correction factors applied to the highest weighted particles, between forested and open sites and across elevations in the ERB. Most values (especially at higher elevations) are above 1.

estimates of snow depth directly (e.g., Magnusson et al., 2017). The improvements to SWE are more modest than improvements to snow depth, because the model relates snow depth and SWE through snow density, which has the smallest reduction in error with DA. Improvements to snow depth only indirectly affect density, through processes such as overburden compaction (Smyth et al., 2019). In addition to forcing data and parameter values, density errors can stem from deficient model process—Terzago et al. (2020) tested several models of varying complexity and found that all demonstrated weakness in the representation of snow density. Still, DA improvements to density simulations were not always—the PF reduced density RMSE by 8%–37% at our forested sites, with the larger error reduction at the Colorado study site (Figures 3 and 4). Assimilating density observations would improve modeled density, but such observations are not available at the basin scale.

Forest canopy affects the rates of snow accumulation and melt (e.g., Varhola et al., 2010), relative to open sites with similar meteorological conditions (Figure 5). However, every forested site is different. For some (e.g., at Snodgrass, Figures 5a–5d), the interception and subsequent sublimation of snowfall leads to lower observed snow depth and SWE at forested (compared to open) sites during the accumulation season. Then, canopy shading of shortwave radiation causes snow to melt more slowly and persist longer in the forest through the melt season, leading to an interchange in the magnitudes of snow depth/SWE between forested and open sites. Other locations exhibit different dynamics: the ForEST area experienced a fire in 2003 (B&B Complex fires), creating some areas with a mix of standing dead trees and live—potentially lowering the impact of canopy interception on snow accumulation at the stand scale. In addition, possible mid-winter melt and wind scouring at the open site leads to roughly equal snow depth and between open and forest areas through the accumulation season (Figures 5e–5h). Snow depth observations help elucidate these patterns, and DA improves simulations of accumulation and melt across forest types and throughout the water year, including between observations.

At the open sites, we do not include any forest-edge effects (e.g., radiation shading, canopy longwave emissions from nearby trees) in the model implementation. All open sites are over 25 m from the forest edge (approximately 3x the average tree height at Snodgrass), so we assume these impacts are minimal.

In assessing model sensitivity (Objective 2) we found that two canopy-related parameters, f_{sky} and f_{veg} , are most important. While the specific parameter names and formulations (e.g., fractional sky view) are in some cases unique to FSM2, they capture physical processes built into other snow models that simulate forest canopy (e.g., infiltration of diffuse radiation). Some of these parameters can be estimated via remote—including some of the most sensitive model parameters. For example, fractional vegetation can be estimated with lidar (as in this paper) or other techniques (e.g., multispectral Normalized Differential Vegetation Index [NDVI] derived values, see Gao et al., 2020), and fractional sky view can be estimated from lidar (Kidd & Chapman, 2012). However, the estimates resulting from these methods are uncertain (e.g., NDVI-derived fractional vegetation RMSE of ~ 0.18 , see Johnson et al., 2012), which in turn can generate considerable variations in OL simulations of snow depth and SWE (Figures 7 and 8). Other canopy parameters (e.g., unloading time scales) are challenging to measure at a single site (e.g., Lv & Pomeroy, 2020), and cannot be measured at the basin scale. Others are not measurable even at a single site (e.g., r_{chz}). Given these challenges, reducing the effective sensitivity to canopy parameters is essential. Our results demonstrate that the PF does reduce the apparent sensitivity of model outputs to these parameters, and that PF estimates are robust to uncertainties in canopy—a key consideration for designing SWE missions at regional and global scales.

We choose to vary parameters $\pm 50\%$ from default values, and while the resulting range between model estimates is a reasonable proxy for the uncertainty stemming from the given parameter, the magnitude (50%) of the perturbation itself does not necessarily reflect the true remote sensing uncertainty range for each parameter. However, fractional vegetation measurement errors (RMSE up to 0.2) are larger in sparse vegetation areas (i.e., low fractional vegetation), which corresponds to a large percentage error (Song et al., 2017). Remote sensing errors for leaf area index (LAI) have errors of 25%–46% depending on methodology (Jensen & Binford, 2004). Further, the experiment is designed to compare the apparent sensitivity of model outputs to uncertain parameters, with and without assimilation—and demonstrate how the PF can compensate for such parameterization errors, even when the errors are quite large. The results are consistent when considering other assumed parameter uncertainty ranges: for example, with a $\pm 25\%$ uncertainty range, the spread between OL runs is still highest when varying the fractional sky view parameter, and the PF significantly reduces variations in snow depth (from a 28% variation to a 1% spread between model estimates) and SWE (from 27% to 5%) estimates stemming from this parameter (Figure S8 in Supporting Information S1).

We assimilated all available snow depth observations at the ForEST and Snodgrass sites—generally 5–7 for each site/year. We also assume a 10 cm observation error in the open and 20 cm in the forest, in line with the uncertainty of airborne lidar (Deems et al., 2013; Grünwald et al., 2010; Harpold et al., 2014; Painter et al., 2016). Previous results from Smyth et al. (2020) demonstrate that model SWE error is not sensitive to the number of assimilated observations, whereas SWE uncertainty is sensitive to both the number/timing of observations and measurement uncertainty. Therefore, the degree to which DA reduces model sensitivity to canopy-related parameters may vary with the platform used to retrieve snow depth measurements for assimilation. Additionally, the minimum number of snow depth samples needed to constrain the model to simulate the seasonal interchange in SWE and snow depth (Figure 5, higher SWE in the open in winter and in the forest in spring) was not evaluated here but could be established in future experiments.

Assimilation results in a reduced range between model simulations because the PF compensates for parameterization errors by adjusting meteorological inputs. For example, in an otherwise perfect OL model, if simulated canopy interception is too high relative to nature, the OL will predict snow depth and SWE values that are low relative to observations, because of high amounts of sublimation from intercepted snow prior to unloading. However, in implementing the PF, individual particles are perturbed with (among other things) different amounts of precipitation forcing (Section 3.5). When confronted by a snow depth observation that is higher than the OL, the PF places higher weights on those particles that are closer to the observation—which tend to be those forced with greater amounts of precipitation (Figure S9 in Supporting Information S1). In this example, the PF compensates for an error in interception parameterization (not enough snowfall reaching the ground) by increasing the amount of snowfall. This is similar to the method used by Lv and Pomeroy (2020) to estimate canopy interception.

Our basin scale results suggest that the PF does indeed increase the amount of snowfall as it weights particles to match observed snow depths—more so at higher elevations and at forested versus open locations (Figure 10). With snow depth observations alone (especially with such low temporal resolution), it is not possible to absolutely identify why SCF is typically above one: parameterization, process representation, or forcing errors may all

be important. However, we can hypothesize why SCF is greater than one, contingent on a few reasonable assumptions. First, results from open sites suggest an elevation-dependent precipitation error in the basin: precipitation is too high (low) at low (high) elevations. The magnitude of the elevation dependent precipitation errors ($\pm 20\%$) is consistent with past studies (Dahri et al., 2018; Li et al., 2014; Romilly & Gebremichael, 2011).

Second, given that errors in precipitation forcing should be equivalent for adjacent forest and open sites, the higher SCF at forest sites (~ 0.15 greater, Figure 10) indicates the model forest canopy is leading to an *effective* precipitation—one that is consistent across elevation bands. Our analysis is performed near peak SWE (i.e., before the melt season), which indicates that the effective precipitation loss is more likely due to an error in simulating an accumulation-related processes (e.g., excessive sublimation of intercepted snow), than a melt-related process (e.g., insufficient shading causing rapid melt in forests). Based on this reasoning, canopy sublimation error is approximately 15% (Figure 10). Sublimation is difficult to model and measure: prior research shows that models both underestimate and overestimate sublimation relative to observations (Ellis et al., 2010; Pomeroy et al., 1998) and that observations vary widely with methodology—for example, cumulative sublimation varying from 20% to 70% of total SWE in subalpine forests in Colorado, similar to the ERB (Molotch et al., 2007; Montesi et al., 2004; Sexstone et al., 2018). Thus, an error equivalent to 15% of SWE is not unexpected.

Overestimation of canopy sublimation is not the only possible explanation for elevation and canopy related errors in the basin. An alternate explanation is that the model has an elevation-dependent error at open sites, yielding excessive (insufficient) melt or sublimation at low (high) elevations. Or, the model could overestimate snow density in the forest. Our analysis of SCF cannot be extended to other parameters without including additional observations or changing the experimental design (which we do not attempt here) because of the model's high sensitivity to precipitation forcing relative to parameters (Figure S2 in Supporting Information S1). Smyth et al. (2020) demonstrated that the PF does not systematically shift other meteorological forcings or the snow viscosity parameter when matching depth observations, and an offline test at the ERB showed that the PF also does not consistently shift the best canopy-related parameter values at forested sites (Figure S10 in Supporting Information S1).

For the experiments in this paper, we do not perturb canopy-related parameters in the PF ensemble generation process for several reasons (Section 3.6.2). However, in an operational implementation of the PF, we recommend that all sensitive model parameters, including canopy-related parameters (e.g., fractional sky view, fractional vegetation), should be perturbed when generating particles. This would help to ensure that the spread between particles encompasses observations in all climates and years, and can be accomplished with virtually no additional computational cost.

In implementing the PF across the ERB (50 m resolution) for one water year, we simulated over 300,000 individual points, with 75 particles at each point: over 22 million individual model runs. We leveraged 48 cores on the RMACC Summit Supercomputer (Anderson et al., 2017) and completed the simulation in one week. However, an operational implementation in “update” mode would not need to simulate the entire water year at once. In addition, the use of more cores (for example, RMACC has over 10,000 available) and other code improvements (e.g., daily output frequencies for state variables rather than hourly, moving PF code to the same language as FSM2) will make compute time more manageable in the future. Challenges include data storage and the coupling of forcing data streams (including downscaling) to the DA modeling framework.

In conclusion, the PF improves simulations of snow depth, density, and SWE across open and forested sites in two disparate climates, supporting and extending the results of Smyth et al. (2020), which focused on open SNOTEL sites in forest clearings. At both Snodgrass and ForEST sites, the OL model performs worse at forested sites compared to adjacent open sites (in snow depth and SWE). As the meteorological forcing data is roughly the same between paired forest-open sites at the same elevation, this disparity points toward errors in canopy process representation/ parameterization at the forested sites. There is greater improvement (relative to the OL) at forest sites (Figure 4), so error after DA is roughly equal between forested and open areas. Calibration can be used to lower model errors, but the outcome is specific to individual basins and tree types, can produce errors in subsequent years, and is difficult to implement in real time (Rutter et al., 2009). DA reduces errors without the need for calibration—no adjustments are necessary across climates or landscape factors. The PF is able to overcome the importance of accurate parameterization by reducing the apparent sensitivity of model outputs to forest canopy parameters. The PF code can be successfully implemented at the basin scale as demonstrated with real, remotely

sensed snow depth observations. Across the entire basin, the PF identifies ~15% less precipitation at forested locations compared to the open. The simplest interpretation of this result is excessive interception loss at forest sites. However, measurements of interception loss would be needed to test this hypothesis.

Data Availability Statement

Model inputs and validation data used this study are available for free at: <https://zenodo.org/record/4904848#.YQWTJ15KhPY>.

Acknowledgments

This work was supported by the National Aeronautics and Space Administration (NASA) Terrestrial Hydrology Program under Grant No. NNX17AL41G and NASA FINESST Grant No. 80NSSC19K1374. The authors thank Richard Essery for making FSM2 code available (<https://github.com/RichardEssery/FSM2>). NLDAS-2 forcing data were acquired through NASA GES DISC (<https://ldas.gsfc.nasa.gov/nldas/NLDAS2forcing.php>). This work utilized resources from the University of Colorado Boulder Research Computing Group, which is supported by the National Science Foundation (awards ACL1532235 and ACL1532236), the University of Colorado Boulder, and Colorado State University. Analysis would not be possible without data collection teams in CO (Eric Small Hydrology Group) and OR (Anne Nolin Computational Mountain Studies Research Group). The authors also thank the ASO team for providing the snow depth, canopy and elevation data used in the basin-scale experiment.

References

- Anderson, J., Burns, P. J., Milroy, D., Ruprecht, P., Hauser, T., & Siegel, H. J. (2017). Deploying RMACC Summit: An HPC resource for the Rocky Mountain region. In *Proceedings of PEARC17*. <https://doi.org/10.1145/3093338.3093379>
- Arulampalam, M. S., Maskell, S., Gordon, N., & Clapp, T. (2007). A tutorial on particle filters for online nonlinear/nongaussian bayesian tracking. *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*, 50(2), 723–737. <https://doi.org/10.1109/9780470544198.ch73>
- Bales, R. C., Molotch, N. P., Painter, T. H., Dettinger, M. D., Rice, R., & Dozier, J. (2006). Mountain hydrology of the western United States. *Water Resources Research*, 42(8), 1–13. <https://doi.org/10.1029/2005WR004387>
- Bavay, M., & Egger, T. (2014). MeteIO 2.4.2: A preprocessing library for meteorological data. *Geoscientific Model Development*, 7, 3135–3151. <https://doi.org/10.5194/gmd-7-3135-2014>
- Broxton, P. D., Harpold, A. A., Biederman, J. A., Troch, P. A., Molotch, N. P., & Brooks, P. D. (2015). Quantifying the effects of vegetation structure on snow accumulation and ablation in mixed-conifer forests. *Ecohydrology*, 8(6), 1073–1094. <https://doi.org/10.1002/eco.1565>
- Clark, M. P., & Vrugt, J. A. (2006). Unraveling uncertainties in hydrologic model calibration: Addressing the problem of compensatory parameters. *Geophysical Research Letters*, 33, 1–5. <https://doi.org/10.1029/2005GL025604>
- Cluzet, B., Lafaysse, M., Cosme, E., Albergel, C., Meunier, L. F., & Dumont, M. (2021). CrocO-v1.0: A particle filter to assimilate snowpack observations in a spatialised framework. *Geoscientific Model Development*, 14(3), 1595–1614. <https://doi.org/10.5194/gmd-14-1595-2021>
- Conway, J. P., Pomeroy, J. W., Helgason, W. D., & Kinar, N. J. (2018). Challenges in modeling turbulent heat fluxes to snowpacks in forest clearings. *Journal of Hydrometeorology*, 19(10), 1599–1616. <https://doi.org/10.1175/JHM-D-18-0050.1>
- Crawford, J. L., McNulty, S. P., Sowell, J. B., & Morgan, M. D. (1998). Changes in aspen communities over 30 years in Gunnison County, Colorado. *American Midland Naturalist*, 140, 197–205. [https://doi.org/10.1674/0003-0031\(1998\)140\[0197:ciacoy\]2.0.co;2](https://doi.org/10.1674/0003-0031(1998)140[0197:ciacoy]2.0.co;2)
- Dahri, Z. H., Moors, E., Ludwig, F., Ahmad, S., Khan, A., Ali, I., & Kabat, P. (2018). Adjustment of measurement errors to reconcile precipitation distribution in the high-altitude Indus basin. *International Journal of Climatology*, 38(10), 3842–3860. <https://doi.org/10.1002/joc.5539>
- Deems, J. S., Painter, T. H., & Finnegan, D. C. (2013). Lidar measurement of snow depth: A review. *Journal of Glaciology*, 59(215), 467–479. <https://doi.org/10.3189/2013JG12J154>
- Dong, J., Steele-Dunne, S. C., Judge, J., & van de Giesen, N. (2015). A particle batch smoother for soil moisture estimation using soil temperature observations. *Advances in Water Resources*, 83, 111–122. <https://doi.org/10.1016/j.advwatres.2015.05.017>
- Elder, K., Goodbody, A., Cline, D., Houser, P., Liston, G. E., Mahrt, L., & Rutter, N. (2009). NASA Cold Land Processes Experiment (CLPX 2002/03): Ground-based and near-surface meteorological observations. *Journal of Hydrometeorology*, 10(1), 330–337. <https://doi.org/10.1175/2008JHM878.1>
- Ellis, C. R., Pomeroy, J. W., Brown, T., & MacDonald, J. (2010). Simulation of snow accumulation and melt in needleleaf forest environments. *Hydrology and Earth System Sciences*, 14(6), 925–940. <https://doi.org/10.5194/hess-14-925-2010>
- Essery, R. (2015). A factorial snowpack model (FSM 1.0). *Geoscientific Model Development*, 8(12), 3867–3876. <https://doi.org/10.5194/gmd-8-3867-2015>
- Gao, L., Wang, X., Johnson, B. A., Tian, Q., Wang, Y., Verrelst, J., et al. (2020). Remote sensing algorithms for estimation of fractional vegetation cover using pure vegetation index values: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 159, 364–377. <https://doi.org/10.1016/j.isprsjprs.2019.11.018>
- Gleason, K. E., Nolin, A. W., & Roth, T. R. (2017). Developing a representative snow-monitoring network in a forested mountain watershed. *Hydrology and Earth System Sciences*, 21(2), 1137–1147. <https://doi.org/10.5194/hess-21-1137-2017>
- Grünwald, T., Schirmer, M., Mott, R., & Lehning, M. (2010). Spatial and temporal variability of snow depth and SWE in a small mountain catchment. *The Cryosphere Discussions*, 4(1), 1–30. <https://doi.org/10.5194/tcd-4-1-2010>
- Günther, D., Hanzer, F., Warscher, M., Essery, R., & Strasser, U. (2020). Including parameter uncertainty in an Intercomparison of physically-based snow models. *Frontiers in Earth Science*, 8, 1–15. <https://doi.org/10.3389/feart.2020.542599>
- Günther, D., Marke, T., Essery, R., & Strasser, U. (2019). Uncertainties in snowpack simulations—Assessing the impact of model structure, parameter choice, and forcing data error on point-scale energy balance snow model performance. *Water Resources Research*, 55(4), 2779–2800. <https://doi.org/10.1029/2018WR023403>
- Harpold, A. A., Biederman, J. A., Condon, K., Merino, M., Korgaonkar, Y., Nan, T., et al. (2014). Changes in snow accumulation and ablation following the Las Conchas Forest Fire, New Mexico, USA. *Ecohydrology*, 7(2), 440–452. <https://doi.org/10.1002/eco.1363>
- Hedrick, A. R., Marks, D., Havens, S., Robertson, M., Johnson, M., Sandusky, M., et al. (2018). Direct insertion of NASA Airborne Snow Observatory-derived snow depth time-series into the iSnobal energy balance snow model. *Water Resources Research*, 54, 1–8063. <https://doi.org/10.1029/2018WR023190>
- Helbig, N., Moeser, D., Teich, M., Vincent, L., Lejeune, Y., Sicart, J. E., & Monnet, J. M. (2020). Snow processes in mountain forests: Interception modeling for coarse-scale applications. *Hydrology and Earth System Sciences*, 24(5), 2545–2560. <https://doi.org/10.5194/hess-24-2545-2020>
- Jensen, R. R., & Binford, M. W. (2004). Measurement and comparison of Leaf Area Index estimators derived from satellite remote sensing techniques. *International Journal of Remote Sensing*, 25(20), 4251–4265. <https://doi.org/10.1080/01431160410001680400>
- Johnson, B., Tateishi, R., & Kobayashi, T. (2012). Remote sensing of fractional green vegetation cover using spatially-interpolated endmembers. *Remote Sensing*, 4(9), 2619–2634. <https://doi.org/10.3390/rs4092619>
- Kidd, C., & Chapman, L. (2012). Derivation of sky-view factors from lidar data. *International Journal of Remote Sensing*, 33(11), 3640–3652. <https://doi.org/10.1080/01431161.2011.635163>

- Kim, R. S., Kumar, S., Vuyovich, C., Houser, P., Lundquist, J., Mudryk, L., et al. (2021). Snow Ensemble Uncertainty Project (SEUP): Quantification of snow water equivalent uncertainty across North America via ensemble land surface modeling. *Cryosphere*, *15*(2), 771–791. <https://doi.org/10.5194/tc-15-771-2021>
- Kitagawa, G. (1996). Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics*, *5*(1), 1–25. <https://doi.org/10.1080/10618600.1996.10474692>
- Kojima, K. (1967). Densification of seasonal snow cover. In *Physics of ice and snow, proceedings of an international conference on low temperature science* (pp. 929–952). Sapporo, Japan: Institute of Low Temperature Science, Hokkaido University.
- Krogh, S. A., Broxton, P. D., Manley, P. N., & Harpold, A. A. (2020). Using process based snow modeling and lidar to predict the effects of forest thinning on the Northern Sierra Nevada snowpack. *Frontiers in Forests and Global Change*, *3*. <https://doi.org/10.3389/ffgc.2020.00021>
- Li, X., Zhang, Q., & Xu, C. Y. (2014). Assessing the performance of satellite-based precipitation products and its dependence on topography over Poyang Lake basin. *Theoretical and Applied Climatology*, *115*(3–4), 713–729. <https://doi.org/10.1007/s00704-013-0917-x>
- Lievens, H., Brangers, I., Marshall, H.-P., Jonas, T., Olefs, M., & De Lannoy, G. (2021). Observing snow depth at sub-kilometer resolution over the European Alps from Sentinel-1. *The Cryosphere Discussions*, 618–621. <https://doi.org/10.1109/figrass47720.2021.9553420>
- Lievens, H., Demuzere, M., Marshall, H., Reichle, R. H., Brangers, I., & De Rosnay, P. (2019). Snow depth variability in the Northern Hemisphere mountains observed from space. *Nature Communications*, *10*, 1–33. <https://doi.org/10.1038/s41467-019-12566-y>
- Lundquist, J. D., Dickerson-Lange, S. E., Lutz, J. A., & Cristea, N. C. (2013). Lower forest density enhances snow retention in regions with warmer winters: A global framework developed from plot-scale observations and modeling. *Water Resources Research*, *49*(10), 6356–6370. <https://doi.org/10.1002/wrcr.20504>
- Lv, Z., & Pomeroy, J. W. (2020). Assimilating snow observations to snow interception process simulations. *Hydrological Processes*, *34*(10), 2229–2246. <https://doi.org/10.1002/hyp.13720>
- Magnusson, J., Winstral, A., Stordal, A. S., Essery, R., & Jonas, T. (2017). Improving physically based snow simulations by assimilating snow depths using the particle filter. *Water Resources Research*, *53*, 1125–1143. <https://doi.org/10.1002/2016WR019092>
- Margulis, S. A., Fang, Y., Li, D., Lettenmaier, D. P., & Andreadis, K. (2019). The utility of infrequent snow depth images for deriving continuous space-time estimates of seasonal snow water equivalent. *Geophysical Research Letters*, *46*, 5331–5340. <https://doi.org/10.1029/2019GL025207>
- Margulis, S. A., Giroto, M., Cortés, G., & Durand, M. (2015). A particle batch smoother approach to snow water equivalent estimation. *Journal of Hydrometeorology*, *16*(4), 1752–1772. <https://doi.org/10.1175/JHM-D-14-0177.1>
- Marks, D., & Dozier, J. (1992). Climate and energy exchange at the snow surface in the alpine region of the Sierra Nevada: 2. Snow cover energy balance. *Water Resources Research*, *28*(11), 3043–3054. <https://doi.org/10.1029/92WR01483>
- Marti, R., Gascoin, S., Berthier, E., De Pinel, M., Houet, T., & Laffly, D. (2016). Mapping snow depth in open alpine terrain from stereo satellite imagery. *Cryosphere*, *10*(4), 1361–1380. <https://doi.org/10.5194/tc-10-1361-2016>
- Mazzotti, G., Essery, R., Moeser, C. D., & Jonas, T. (2020). Resolving small-scale forest snow patterns using an energy balance snow model with a one-layer canopy water resources research. *Water Resources Research*, *56*. <https://doi.org/10.1029/2019WR026129>
- Mazzotti, G., Webster, C., Essery, R., & Jonas, T. (2021). Increasing the physical representation of forest-snow processes in coarse-resolution models: Lessons Learned from upscaling hyper-resolution simulations. *Water Resources Research*, *57*. <https://doi.org/10.1029/2020WR029064>
- Mizukami, N., & Perica, S. (2008). Spatiotemporal characteristics of snowpack density in the mountainous regions of the western United States. *Journal of Hydrometeorology*, *9*, 1416–1426. <https://doi.org/10.1175/2008jhm981.1>
- Moeser, D., Mazzotti, G., Helbig, N., & Jonas, T. (2016). Representing spatial variability of forest snow: Implementation of a new interception model. *Water Resources Research*, *52*, 1208–1226. <https://doi.org/10.1002/2015WR017961>
- Moeser, D., Morsdorf, F., & Jonas, T. (2015). Novel forest structure metrics from airborne LiDAR data for improved snow interception estimation. *Agricultural and Forest Meteorology*, *208*, 40–49. <https://doi.org/10.1016/j.agrformet.2015.04.013>
- Moller, D., Andreadis, K. M., Bormann, K. J., Hensley, S., & Painter, T. H. (2017). Mapping snow depth from ka-band interferometry: Proof of Concept and comparison with scanning lidar retrievals. *IEEE Geoscience and Remote Sensing Letters*, *14*(6), 886–890. <https://doi.org/10.1109/LGRS.2017.2686398>
- Molotch, N. P., Blanken, P. D., Williams, M. W., Turnipseed, A. A., Monson, R. K., & Margulis, S. A. (2007). Estimating sublimation of intercepted and sub-canopy snow using eddy covariance systems. *Hydrological Processes*, *21*, 1567–1575. <https://doi.org/10.1002/hyp.6719>
- Montesi, J., Elder, K., Schmidt, R. A., & Davis, R. E. (2004). Sublimation of intercepted snow within a subalpine forest canopy at two elevations. *Journal of Hydrometeorology*, *5*, 7632–773. [https://doi.org/10.1175/1525-7541\(2004\)005<0763:soiswa>2.0.co;2](https://doi.org/10.1175/1525-7541(2004)005<0763:soiswa>2.0.co;2)
- Musselman, K. N., Molotch, N. P., & Brooks, P. D. (2008). Effects of vegetation on snow accumulation and ablation in a mid-latitude sub-alpine forest. *Hydrological Processes*, *22*, 2267–2776. <https://doi.org/10.1002/hyp.7050>
- Musselman, K. N., & Pomeroy, J. W. (2017). Estimation of needleleaf canopy and trunk temperatures and longwave contribution to melting snow. *Journal of Hydrometeorology*, *18*(2), 555–572. <https://doi.org/10.1175/jhm-d-16-0111.1>
- Musselman, K. N., Pomeroy, J. W., & Link, T. E. (2015). Variability in shortwave irradiance caused by forest gaps: Measurements, modelling, and implications for snow energetics. *Agricultural and Forest Meteorology*, *207*, 69–82. <https://doi.org/10.1016/j.agrformet.2015.03.014>
- Oswalt, S. N., & Smith, B. (2014). *U.S. forest resource facts and historical trends*. USDA Forest Service.
- Painter, T. H., Berisford, D. F., Boardman, J. W., Bormann, K. J., Deems, J. S., Gehrke, F., et al. (2016). The Airborne Snow Observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo. *Remote Sensing of Environment*, *184*, 139–152. <https://doi.org/10.1016/j.rse.2016.06.018>
- Pomeroy, J. W., Gray, D. M., Shook, K. R., Toth, B., Essery, R. L. H., Pietroniro, A., & Hedstrom, N. (1998). An evaluation of snow accumulation and ablation processes for land surface modelling. *Hydrological Processes*, *12*, 2339–2367. [https://doi.org/10.1002/\(sici\)1099-1085\(199812\)12:15<2339::aid-hyp800>3.0.co;2-1](https://doi.org/10.1002/(sici)1099-1085(199812)12:15<2339::aid-hyp800>3.0.co;2-1)
- Raleigh, M. S., Lundquist, J. D., & Clark, M. P. (2015). Exploring the impact of forcing error characteristics on physically based snow simulations within a global sensitivity analysis framework. *Hydrology and Earth System Sciences*, *19*(7), 3153–3179. <https://doi.org/10.5194/hess-19-3153-2015>
- Raleigh, M. S., & Small, E. E. (2017). Snowpack density modeling is the primary source of uncertainty when mapping basin-wide SWE with lidar. *Geophysical Research Letters*, *44*(8), 3700–3709. <https://doi.org/10.1002/2016GL071999>
- Romilly, T. G., & Gebremichael, M. (2011). Evaluation of satellite rainfall estimates over Ethiopian river basins. *Hydrology and Earth System Sciences*, *15*(5), 1505–1514. <https://doi.org/10.5194/hess-15-1505-2011>
- Roth, T. R., & Nolin, A. W. (2017). Forest impacts on snow accumulation and ablation across an elevation gradient in a temperate montane environment. *Hydrology and Earth System Sciences*, *21*(11), 5427–5442. <https://doi.org/10.5194/hess-21-5427-2017>
- Rutter, N., Essery, R., Pomeroy, J., Altimir, N., Andreadis, K., Baker, I., et al. (2009). Evaluation of forest snow processes models (SnowMIP2). *Journal of Geophysical Research*, *114*(6). <https://doi.org/10.1029/2008JD011063>

- Serreze, M. C., Clark, M. P., Armstrong, R. L., McGinnis, D. A., & Pulwarty, R. S. (1999). Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data. *Water Resources Research*, *35*(7), 2145–2160. <https://doi.org/10.1029/1999WR900090>
- Sextstone, G. A., Clow, D. W., Fassnacht, S. R., Liston, G. E., Hiemstra, C. A., Knowles, J. F., & Penn, C. A. (2018). Snow sublimation in mountain environments and its sensitivity to forest disturbance and climate warming. *Water Resources Research*, *54*(2), 1191–1211. <https://doi.org/10.1002/2017WR021172>
- Smyth, E. J., Raleigh, M. S., & Small, E. E. (2019). Particle filter data assimilation of monthly snow depth observations improves estimation of snow density and SWE. *Water Resources Research*, *55*(4), 1296–1311. <https://doi.org/10.1029/2018WR023400>
- Smyth, E. J., Raleigh, M. S., & Small, E. E. (2020). Improving SWE estimation with data assimilation: The influence of snow depth observation timing and uncertainty. *Water Resources Research*, *56*(5), 1–17. <https://doi.org/10.1029/2019WR026853>
- Song, W., Mu, X., Ruan, G., Gao, Z., Li, L., & Yan, G. (2017). Estimating fractional vegetation cover and the vegetation index of bare soil and highly dense vegetation with a physically based method. *International Journal of Applied Earth Observation and Geoinformation*, *58*, 168–176. <https://doi.org/10.1016/j.jag.2017.01.015>
- Terzago, S., Andreoli, V., Arduini, G., Balsamo, G., Campo, L., Cassardo, C., et al. (2020). Sensitivity of snow models to the accuracy of meteorological forcings in mountain environments. *Hydrology and Earth System Sciences*, *24*(8), 4061–4090. <https://doi.org/10.5194/hess-24-4061-2020>
- Trujillo, E., & Molotch, N. P. (2014). Snowpack regimes of the western United States. *Water Resources Research*, *50*(5), 4149–4162. <https://doi.org/10.1002/2013wr014753>
- Vander Jagt, B., Lucieer, A., Wallace, L., Turner, D., & Durand, M. (2015). Snow depth retrieval with UAS using photogrammetric techniques. *Geosciences*, *5*(3), 264–285. <https://doi.org/10.3390/geosciences5030264>
- van Leeuwen, P. J. (2009). Particle filtering in geophysical systems. *Monthly Weather Review*, *137*(12), 4089–4114. <https://doi.org/10.1175/2009MWR2835.1>
- Varhola, A., Coops, N. C., Weiler, M., & Moore, R. D. (2010). Forest canopy effects on snow accumulation and ablation: An integrative review of empirical results. *Journal of Hydrology*, *392*(3–4), 219–233. <https://doi.org/10.1016/j.jhydrol.2010.08.009>
- Wetlaufer, K., Hendriks, J., & Marshall, L. (2016). Spatial heterogeneity of snow density and its influence on snow water equivalence estimates in a large mountainous basin. *Hydrology*, *3*. <https://doi.org/10.3390/hydrology3010003>
- Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., et al. (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. *Journal of Geophysical Research*, *117*(D3). <https://doi.org/10.1029/2011JD016048>