1 **Short summary.** 250 m estimates of snow water equivalent in the Western US and Canada wereare improved by assimilating
2 observations representative of a future snow-focused satellite mission with a land surface model. Here, by including a gap-
3 filling strategy, snow estimates could be improved in forested regions where remote sensing is challenging. This approach
4 improved estimates of winter maximum snow water volume to within 4%, on average, with persistent improvements to both
5 spring snow and runoff throughout spring snowmelt. in many regions.

Extending the utility of space-borne snow water equivalent observations over vegetated areas with data assimilation

8 Justin M. Pflug[1,2], Melissa L. Wrzesien[1,2], Sujay V. Kumar[2], Eunsang Cho[1,2,a], Kristi R. Arsenault[3,2], Paul
9 R. Houser[4], Carrie M. Vuyovich[2]

10 [1]Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA
11 [2]Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA
12 [3]Science Applications International Corporation, McLean, VA, USA
13 [4]Geography and Geoinformation Science Department, George Mason University, Fairfax,VA, USA

14 [a]current address: Ingram School of Engineering, Texas State University, San Marcos, TX, USA

15 *Correspondence to*: Justin Pflug (Justin.Pflug@nasa.gov)

16 **Abstract.** Snow is a vital component of the Earth system. Yet, no snow-focused satellite remote sensing platform currently
17 exists. In this study, we investigate how synthetic observations of snow water equivalent (SWE) representative of a synthetic
18 aperture radar remote sensing platform could improve spatiotemporal estimates of snowpack. We use an Observationa fraternal
19 twin Observing System Simulation Experiment, specifically investigating how much snow simulated using popularwidely
20 used models and forcing data could be improved by assimilating synthetic observations of SWE. We focus this study across a
21 24°-by-37° domain in the Western United States (US) and Canada, simulating snow at 250 m resolution and hourly timesteps
22 in water-year 2019. We perform two data assimilation experiments, including: 1) a simulation excluding synthetic observations
23 in forests where canopies obstruct remote sensing retrievals, and 2) a simulation inferring snow distribution in forested grid
24 cells using synthetic observations from nearby canopy-free grid cells. Results found that, relative to a nature run, or assumed
25 true simulation of snow evolution, assimilating synthetic SWE observations improved average SWE biases at peakmaximum
26 snowpack timing in shrub, grass, crop, bare-ground, and wetland land cover types from 14%, to within 1%. However, forested
27 grid cells contained a disproportionate amount of SWE volume. In forests, SWE mean absolute errors at peak snowpackthe
28 time of maximum snow volume were 111 mm, and average SWE biases were on the order of 150%. Here, the data assimilation
29 approach that estimated forest SWE using observations from the nearest canopy-free grid cells substantially improved these
30 SWE biases (18%) and the SWE mean absolute error (27 mm). Simulations employing data assimilation also improved
31 estimates of the temporal evolution of both SWE and runoff, even in spring snowmelt periods when melting snow and high
32 snow liquid water content prevented synthetic SWE retrievals. In fact, in the Upper Colorado River basinregion, melt-season

33 SWE biases were improved from 63% to within 1%, and the Nash Sutcliffe Efficiency of runoff improved from –2.59 to 0.22.

34 These results demonstrate the value of data assimilation and a snow-focused globally relevant remote sensing platform for

35 improving the characterization of SWE and associated water availability.

1 Introduction

37 Snow plays important roles in the Earth system~~,~~ by regulating global temperatures and cooling the land surface because of

38 its reflective properties ~~(Barry, 2002). The insulative properties of snow are also vital in ecological balance by protecting plants~~

39 ~~and animals from extreme cold and providing moisture when it melts. Snow is~~(Barry, 2002). Snow is also a major source of

40 water storage for many regions, especially in areas that rely on snowpack to sustain water resources during the dry season. ~~It~~In

41 fact, it has been estimated that more than 2 billion people around the world are reliant on seasonal snow melt for their water

42 supply (Barnett et al., 2005). Snowpack is the natural 'integrator' of climatic conditions and offers more predictability of water

43 availability than variables with shorter memory, such as precipitation and streamflow (Terzago et al., 2023).~~ As a result,~~

44 ~~A~~accurate wintertime estimates of snowpack are therefore critical for water management and agricultural planning ~~(Koster et~~

45 ~~al., 2010). Water managers can prepare mitigation strategies for extreme dry or wet possibilities based on the wintertime~~

46 ~~snowpack.~~(Koster et al., 2010). For example, in the Western ~~U.S.,~~US, where a vast majority of streamflow originates from

47 snow (Li et al., 2017), it is common practice to use the April 1 snowpack, the historic ~~peak for~~date of maximum snowpack in

48 that region, for developing water supply estimates for later in the season. However, climate change impacts have led to

49 increased variability in the snow seasonality (Livneh and Badger, 2020), with warmer temperatures reducing the amount of

50 snow accumulation and seasonal snow storage, and advancing the timing of the spring melt. Therefore, accurate

51 characterization of winter snowpack and its variability is critically important for making informed water supply quantifications.

52 In recognition of the critical need to have spatially distributed measurements of snow mass, there have been several efforts

53 to measure and estimate SWE~~snowpack and snow water equivalent~~ from many different remote sensing platforms in the past

54 several decades. Airborne lidar systems have been able to provide high resolution, accurate measurements of snow mass

55 (Painter et al., 2016), but this approach has significant logistical barriers for global and frequent snow measurements, and the

56 hydrological utility of a practical spaceborne lidar platform is limited (Kwon et al., 2021). In the past three decades, snow

57 depth and SWE estimates have been derived from passive microwave remote sensing measurements, but ~~they tend to be of~~these

58 measurements are at coarse spatial resolution~~s~~, and have limited accuracies over deep and wet snow, complex terrain, and

59 dense vegetation (Derksen et al., 2014; Foster et al., 2005)~~(REF)~~. Active microwave remote sensing instruments such as

60 Synthetic Aperture Radars (SARs) can provide finer spatial resolution measurements to help resolve some of these issues.~~,~~ For

61 example, C-band SAR observations from the Sentinel-1 constellation ~~and~~ have shown promise in obtaining high quality,

62 moderate resolution (1km) observations in deep snow environments (Lievens et al., 2019). A volume scattering radar approach,

63 using X- and Ku-band SAR, has also been demonstrated in several airborne campaigns and proposed for multiple snow mission

64 concepts (Yueh et al 2009, Rott et al 2010) because of its potential to achieve high resolution and global coverage over a range

of snow depths. While these microwave instruments can observe in night-time and cloudy conditions, they are still limited over areas with dense vegetation (Tsang et al., 2022). Further, all spaceborne instruments have inherent coverage gaps due to their orbital and revisit configurations.

To overcome these limitations, modeling and data assimilation systems are needed that can extend the coverage and utility of available measurements to areas, times, and variables that are not directly observed. In this article, we present a novel approach through data assimilation, designed specifically to improve the usefulness of spaceborne SWE retrievals over forested areas. The approach is demonstrated using an observing system simulation experiment (OSSE; e.g., Cho et al., ~~2022~~2023; Errico et al., 2007) which is an approach used to formally assess the impact of the data to be collected from an anticipated mission. Several prior studies have examined the use of OSSEs for snow mission studies ~~(Garnaud et al., 2019; Kwon et al., 2021; Wrzesien et al., 2022).~~(Garnaud et al., 2019; Kwon et al., 2021; Wrzesien et al., 2022). Among them, SAR-focused OSSEs have been conducted by Garnaud et al. (2019) and Cho et al. (~~2022~~2023) to assess the utility of hypothetical snow observations. Garnaud et al. (2019) focused on a Ku-band SAR to quantify trade-offs between sensor configurations (e.g., various spatial resolutions and revisit frequencies) with retrieval algorithm accuracy and SWE performance in southern Quebec, Canada, where temperate forests are dominant with shallow and moderate snowpack conditions. Cho et al. (~~2022~~2023) conducted a X-/Ku-band SAR OSSE with an achievable sensor configuration (1 km spatial resolution, 7-day revisit frequency, and orbital configurations ~~using the Trade-space Analysis Tool for Constellations [TAT-C] simulator~~) focusing on mountainous environments in a western Colorado~~. To comprise diverse snow environments, specifically for~~ and testing the degree to which various ~~forested landscapes in this study, we focus on~~SAR retrieval capabilities in different forest densities and snow volumes could improve observationally-based SWE estimates. Here, we build on this prior research by developing an OSSE covering the entire western ~~United States domain along with the proposed, more realistic constellations using TAT-C.~~US and portions of Canada. We simulate finer-scale (250 m) synthetic SWE observations that could be provided from a future X-/Ku-band SAR mission, which are then incorporated within a land surface model through data assimilation to assess their capability to improve snow state estimates, and the integrated impact on hydrologic states in space and time. The assimilation experiments here are conducted with and without a novel strategy to extend SAR-based ~~observational coverage of snow mass to various~~SWE estimates from unforested regions into forested landscapes where SAR retrievals of the snowpack may be obscured by the forest canopy.

The primary contribution of this paper is the development of a viable strategy for extending ~~space-borne remote sensing~~hypothetical remotely sensed SWE ~~measurements in preparation of~~retrievals from a ~~future snow mission concept (a~~ volume-scattering X-/Ku--band SAR~~).~~ satellite mission into difficult-to-observe forest landscapes. We specifically focus on addressing the following research questions: 1) what is the added utility of spaceborne active remote sensing SWE retrievals across the ~~western United States?~~Western US and Canada? 2) how much can spatiotemporal representations of SWE be improved by focusing on developing observationally based snow estimates over areas with dense vegetation, where SAR

97 sensors may be limited? 3) How much added hydrological utility can be obtained through spaceborne active remote sensing

98 measurements and data assimilation approaches, particularly when coverage over forested areas is improved?

99   Section 2 describes the study domain and OSSE modeling setup. This is followed by the description of the results (Sect. 3),

100 a discussion of the findings (Sect. 4), and the study's conclusions (Sect. 5).

2.1 Study domain and OSSE setup

102 ~~As noted above, an~~An OSSE is ~~a typical approach~~ used to assess the value of the data to be collected from an anticipated

103 mission~~, and consists~~. OSSEs often consist of the following steps: 1) ~~Development of an assumed "true state" of the relevant~~

104 ~~system is produced by~~ Developing a "nature run" that ~~involves the use of~~uses a state-of-the-art model employed with the best

105 available boundary conditions (Sect. 2.1); 2) ~~simulation of synthetic observations from~~ using the nature run to generate

106 simulated remote sensing observations, accounting for sources of sensing limitations, sensing uncertainties, and orbital

107 configurations (Sect. 2.2); 3) ~~incorporation of~~incorporating the simulated observations (often through data assimilation, Sect.

108 2.3) in a ~~different,~~separate, "open loop" model configuration with accuracies representative of common modeling biases and

109 uncertainties~~, often called the 'open loop'~~; and 4) ~~evaluation of the impact that~~evaluating how much the simulated remote

110 sensing data ~~have on improving~~improve the open loop model performance relative to the nature run. ~~The details of the OSSE~~

111 ~~setup used in this manuscript, are described below,~~In addition to this OSSE approach, this study goes further by 1) testing the

112 degree of improvement to both the remotely-sensed variable (i.e., SWE) and the resulting changes to land surface runoff in

113 snowy regions, and 2) developing two separate data assimilation experiments, one which masks simulated observations in

114 forested pixels where SAR retrievals may be most challenging, and the other including a novel approach for inferring SWE in

115 forested ~~landscapes~~pixels using simulated observations from nearby, unforested ~~simulated observations~~pixels (Sect. 2.4). The

116 details of the OSSE setup used in this study are described in more depth throughout this section.

117   We employ the NASA Land Information System (LIS; Kumar et al., 2006), an infrastructure for high performance,

118 ensemble-based land surface modelling and data assimilation to enable this OSSE. LIS encompasses several advanced land

119 surface models that can simulate terrestrial water, energy, and carbon balances and related states such as soil moisture, land

120 surface temperature, and SWE, among others. These include different versions of community models such as Noah (Ek et al.,

121 2003), Variable Infiltration Capacity (VIC; Liang et al., 1994), Catchment (Koster et al., 2000), Joint UK Land Environment

122 Simulator (JULES; Best et al., 2011), and Noah-MP (Niu et al., 2011). The LIS framework also includes support for specialized

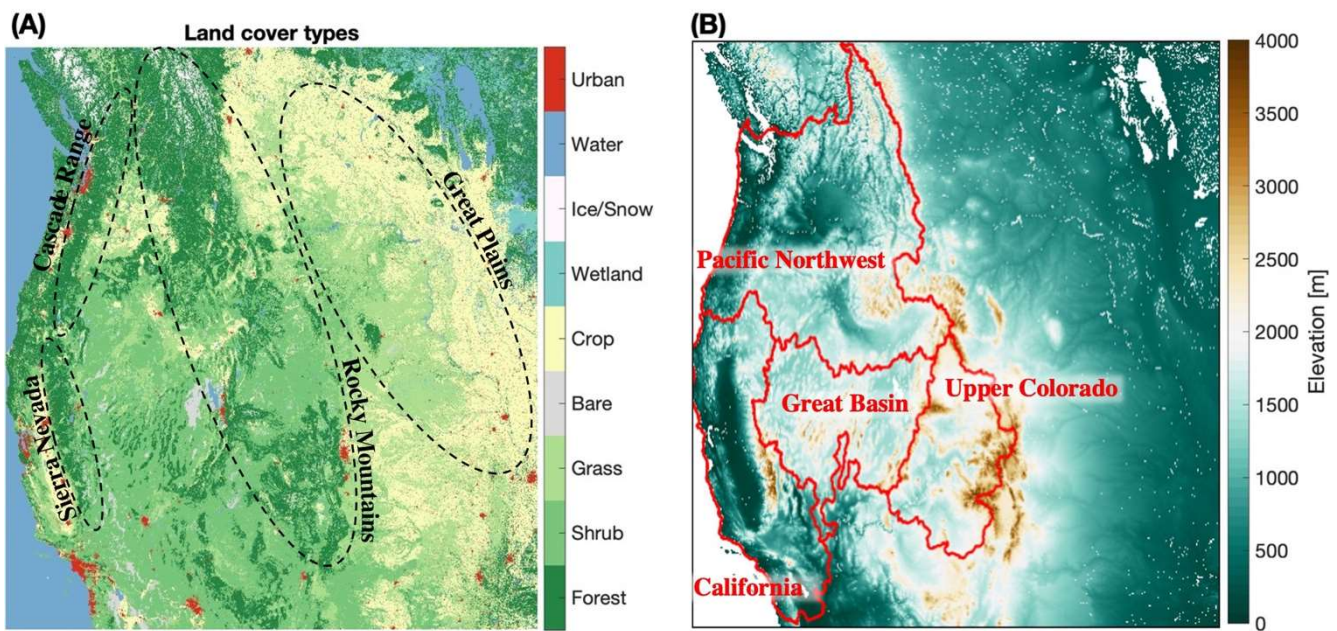
123 models that are designed to provide more detailed representations of certain land surface processes (e.g. snow), while enabling

124 interaction with LSMs that solve for water, energy, and carbon balances at a macroscale. For example, the advanced snow

125 physics model called SnowModel (Liston and Elder, 2006) has been incorporated within LIS in a manner that allows coupling

126 to existing LSMs. This structure allows the use of the advanced snow physics from SnowModel while leveraging the existing

127 process schemes (e.g., sub-surface, groundwater, canopy) within the LSMs. Here we utilize these unique capabilities for
128 enabling the OSSE integrations.

129 . The study is conducted over a large domain (Fig. 1), covering the Western ~~U.S.~~US and southern Canada from 31-55N and
130 93-130W at a 250 m spatial resolution. As shown in Fig. 1, this modeling domain encompasses a broad range of vegetation
131 types, topographical regimes and ~~river basins~~water resources regions of the Pacific Northwest, California, Great Basin, and
132 Upper Colorado. 22% of the domain is covered by forests, with grasslands, croplands, and shrublands accounting for 20%,
133 23%, and 26% of the domain, respectively. Forests dominate the coverage of areas with significant snowpack, occupying 58%
134 of regions that are in the mid-elevation range of 2500-3500m, and 15% of the areas with elevations over 3500m. From a
135 ~~modelling~~modeling perspective, the domain extent of Fig. 1 (~83 million land grid points) is computationally challenging. The
136 scalable high performance computational and parallel inputting and outputting capabilities of NASA LIS were leveraged to
137 enable these simulations. A multiprocessor configuration involving approximately 1000 processors was employed to facilitate
138 large model simulations for the nature run, open loop simulation, and two simulations with data assimilation.

139
140 **Figure 1:** Maps of the land and vegetation classes (A; left panel) and elevation (in meters) (B; right panel) used in the
141 simulations. Outlines and labels in the left panel indicate regions discussed in the Results (Sect. 4). Red contours in the right
142 panel indicate ~~basins~~hydrologic regions used in the analysis.

143 ~~The "nature run"~~ Simulations in this study ~~was performed using a model configuration that coupled a state-of-the-art and~~
144 ~~physically based snow model (named SnowModel, discussed later in this section) with the Noah model with multi-~~
145 ~~parameterization (Noah-MP) version 4.0.1. The simulations~~ are conducted by forcing ~~this coupled model setup~~LIS LSMs with

146 the surface meteorology from NASA's Modern Era Retrospective Reanalysis, version 2 (~~MERRA2~~MERRA-2; Gelaro et al.,
147 2017) ~~product, with LIS-provided downscaling to 250 m resolution using lapse rates and topographical downscaling~~
148 ~~approaches, like SW corrections based on topographical shading (Arsenault~~and ECMWF Reanalysis, version 5 (ERA5;
149 Hersbach et al., 2020) products. The model integrations were conducted for the water year 2019 (September 2018 – September
150 2019), which was a wetter than normal year based on the long-term average meteorological conditions over this domain. ~~The~~
151 ~~open loop and data assimilation integrations performed in this study were conducted using Noah-MP alone.~~

152 The open loop and data assimilation integrations performed in this study are conducted using the Noah land surface model
153 with multi-parameterization (Noah-MP ~~model discussed above~~) version 4.0.1 (Niu et al., 2011) and forcing from ERA5. The
154 Noah-MP model evolved from the Noah LSM, with multiple options for various land surface processes. It represents ~~processes~~
155 ~~related to~~ energy, water, and carbon balances at the land surface by accounting for processes related to infiltration, evaporation,
156 transpiration, runoff generation and groundwater recharge. A TOPMODEL-based runoff model (Beven et al., 2021) is used to
157 calculate surface runoff and groundwater discharge. Options for prognostic vegetation dynamics models that represent the
158 growth and senescence of vegetation are also available within Noah-MP. A two-stream radiative transfer approach is employed
159 to calculate surface energy processes. ~~A~~Finally, a multilayer snowpack model (with up to three layers) ~~that can~~is used to
160 account for snow melt metamorphisms, compaction by overlying snow, sublimation of canopy intercepted snow, and
161 snowmelt-refreeze cycles ~~is available within Noah-MP (Niu and Yang, 2004).~~within Noah-MP (Niu and Yang, 2004).

162 Snow states like snow depth and SWE were also model~~l~~ed across the Western U~~S~~nited States (domain highlighted in Fig.
163 1) at 250 m resolution and hourly time steps using ~~a~~the state-of-the-art and physically based single-layer snow
164 ~~model~~implementation of SnowModel (named SnowModel; Liston and Elder, 2006), provided forcing from MERRA-2 with
165 LIS-provided lapse rates and topography-based meteorological downscaling approaches, like incoming shortwave corrections
166 based on topographical shading (Cosgrove et al., 2003; Kumar et al., 2013). SnowModel ~~This model~~ has seen widespread use
167 in the snow community, demonstrating the capability to resolve snow evolution in a variety of landscapes and complex snow
168 processes like the redistribution of snow via wind, and the resulting impact on snow distribution, melt season snow duration,
169 glacier mass balance, and snow habitat for species like polar bears and Dall sheep (Hiemstra et al., 2002; Liston et al., 2016;
170 Mahoney et al., 2018; Mernild et al., 2017; Sturm and Wagner, 2010). In addition to wind redistribution, ~~S~~snow evolution
171 within SnowModel ~~at each grid cell accounted~~ accounts for a wide set of snow processes, including snow sublimation, snow
172 grain size evolution, solar topographical shading, canopy shading, and canopy snow interception~~, and wind redistribution~~.
173 Through the coupling within LIS, Noah-MP snow states and the resulting snow-driven runoff were updated using the
174 SnowModel outputs at hourly ~~daily~~ timesteps for each grid cell.

175 Preliminary research has shown that relative to Noah-MP, LIS simulations coupling Noah-MP with SnowModel have
176 improved the volume and spatial distribution of simulated snow depth and SWE (Arsenault et al., 2021; Wrzesien et al., 2022).

177 Therefore, the coupled SnowModel and Noah-MP model was a prime candidate for the "nature run" in this study, or the
178 simulation most representative of the true underlying spatiotemporal snow states from which simulated observations were
179 derived (Sect. 2.2), and the assimilated model was compared against. Here, the nature run and open loop simulations detailed
180 above were compared to a widely-used Western US snow reanalysis product (Fang et al., 2022) to ensure that 1) the nature
181 run exhibited reasonable model accuracy, and 2) the departure between the open loop simulation and nature run are
182 representative of common regional, continental, and global modeling efforts (Figure S1 and S2). The OSSE developed for this
183 study is a "fraternal twin" OSSE, wherein two different models are used to simulate snow in the open loop (Noah-MP) and
184 nature run (SnowModel) simulations. This approach is selected since "identical twin" OSSEs, which use the same model, can
185 result in less divergence in model states and information content, biasing the degree of model improvement that could come
186 from assimilating an observation (e.g., Yu et al., 2019). More information on the difference between the open loop and nature
187 run models can be found in Table S1.

188 **2.2 Observation simulator**

189 ~~From the nature run integrations conducted with Noah-MP coupled to SnowModel in LIS, synthetic SWE retrievals at 250~~
190 ~~m spatial resolution, expected from a hypothetical SAR mission, were simulated. First, the orbital swaths were simulated using~~
191 ~~TAT-C (Le Moigne et al., 2017). TAT-C, a NASA software system designed for future Distribution Spacecraft Missions~~
192 ~~(DSM),.~~ Synthetic SWE retrievals at 250 m spatial resolution, representative of a hypothetical X- and Ku-band SAR mission,
193 are simulated from the nature run. To do this, the orbital swaths are simulated using TAT-C (Le Moigne et al., 2017). TAT-C
194 is a NASA software system designed for future Distribution Spacecraft Missions (DSM), which enables us to explore a range
195 of feasible design options (e.g., constellation vs. single, geostationary vs. polar-orbiting, low vs. high temporal frequencies) to
196 estimate optimal gains for the given mission configuration. ~~Using an assumed 10-14 day revisit frequency, these~~Previous
197 OSSEs have been conducted to test the impact from different snow mission configurations (e.g. Garnaud et al 2019). Here we
198 instead focus on demonstrating the value of a gap-filling approach (Sect. 2.4) for estimating snow in forested landscapes where
199 SAR retrievals may be most challenging. Therefore, we used TAT-C to design a conservative mission configuration consisting
200 of a small constellation of X- and Ku-band SAR satellites. Using a 10-14 day revisit frequency, depending on latitude, TAT-
201 C orbital swaths were applied to the nature run outputs to simulate the satellite viewing area. The ~~realistic~~remote sensing spatial
202 coverage is simulated by extending the ground track to a swath width ~~(i.e., 50 km)~~. The daily viewing extents are then simulated
203 as a daily binary map (so-called "cookie cutter") masking the surface as viewed or not at a 250 m spatial resolution.

204 ~~Additionally, b~~Based on an error level of 20%, spatially and temporally uncorrelated random errors drawn from a Gaussian
205 distribution ~~we~~are added to the synthetic SWE retrievals. ~~Note that here we use an optimistic~~This 20% error level is selected
206 using a conservative estimate of SWE measurement uncertainty for a volume-scattering X-/Ku -band SAR mission based on
207 developed mission design concepts and ground validation. For example, the ESA Cold Regions Hydrology High-Resolution
208 Observatory (CoREH2O) mission expected to meet instrument and retrieval requirements of ±30 mm accuracies for SWE of

209   300 mm, ±10% for SWE greater than 300 mm (Rott et al 2010, 2012). Similarly, the Canadian Terrestrial Snow Mass Mission

210   (TSMM) concept that is currently under development aims to achieve better than 20% measurement uncertainty for SWE

211   greater than 50 mm, though it is limited to SWE less than 200 mm due to the dual Ku-band system (Garnaud et al. 2019).

212   Airborne and tower-based field data have demonstrated that a combination X- and Ku-band system can provide SWE retrievals

213   over a range of snow conditions at accuracies better than 20% (Zhu et al. 2018, 2021, Tsang et al 2022, Durand et al. 2023,

214   Singh et al. 2023). However, we use an assumption of uniform error levels throughout the domain, whereas in reality, the

215   errors are likely to be dependent on other factors, including the terrain characteristics, ~~vegetation, and precipitation regime.~~

216   snow characteristics, and vegetation. This is discussed more in Section 4.

217   **2.3 Data assimilation setup**

218   A one-dimensional ensemble Kalman Filter (EnKF; Reichle et al., 2002) ~~was~~is used to assimilate the synthetic observations

219   within the open loop configuration of the model. EnKF is widely used for land data assimilation studies (Kumar et al., 2022),

220   as it provides a flexible approach for the treatment of model and observation errors and non-linear models. An ensemble of

221   model realizations is used by EnKF to assess and propagate model errors. In this instance, the ensemble requirement further

222   adds to the significant computational requirements of the large model domain (Fig. 1) and fine spatial resolution of the

223   simulations (250 m). Therefore, a 5-member ensemble with perturbations applied to the meteorological variables and model

224   prognostic fields are used for ~~-~~simulating uncertainty in the mode~~l~~lled estimates. Table 1 details the parameters for

225   meteorological and model state perturbations, which are based on recent snow data assimilation studies (Lahmers et al., 2022;

226   Kwon et al., 2021). ~~-~~Though a larger ensemble size is better for ensuring sufficient sampling density, our choice of five

227   ensembles is reasonable given that the model state vector used in the assimilation only consists of two variables; the total SWE

228   and snow depth. The assimilation setup employs a sequential update strategy, where at each time step an ensemble of model

229   forecasts is propagated forward in time, followed by an update based on observational inputs. The model states are updated

230   toward the observations based on the relative uncertainties in the model and observations using the following formulation, at

231   a certain time k.

232
$$x_k^{i+} = x_k^{i-} + K_k \left[y_k^i - H_k x_k^{i-} \right]$$
Eq. (1)

233   Where x_k and y_k are the model and observation state vectors, respectively. The term H_k represents the observation operator

234   that maps the model states to the observed variables. The superscripts $i-$ and $i+$ represent the ith ensemble member before

235   and after the update, respectively. K_k is the "Kalman gain" term, that allows the weighting of the observations and model

236   forecasts is a function of the model and observation error covariances.

237   **Table 1.** Model forcing and state-variable perturbations used by the 5-member ensemble of LIS simulations

Variable	Perturbation Type	Cross Correlation across variables

Meteorological Forcing		Std. Dev.	SW corr	LW corr	PCP corr	T corr
Downward Shortwave (SW)	Multiplicative	0.2	1	−0.3	−0.5	0.3
Downward Longwave (LW)	Additive	30	−0.3	1	0.5	0.6
Precipitation (PCP)	Multiplicative	0.5	−0.5	0.5	1	−0.1
Near surface Air Temperature (T)	Additive	0.5	0.3	0.6	−0.1	1
Noah-MP LSM snow states			SWE	Snow depth		
SWE	Multiplicative	0.01	1	0.9		
Snow depth	Multiplicative	0.01	0.9	1		

The data assimilation procedure detailed here assimilated the synthetic SWE retrievals (Sect. 2.2) with the open loop simulation. The degree to which the simulation with data assimilation approached SWE simulated by the nature run is intended to represent the extent to which a SAR remote sensing platform with the SWE retrieval characteristics from Sect. 2.2 could be combined with a land surface model to provide near real-time estimates of SWE at 250 m resolution. However, the SAR observations synthesized in this study have known issues with observing snow with high liquid water contents and dense forest cover. Therefore, synthetic observations at each timestep were masked at grid cells ~~with~~where the most-dominant landcover type from the North American Land Change Monitoring System (NALCMS; Latifovic et al., 2017) was forested, including deciduous, evergreen, and mixed forest cover (Fig. ~~1), and~~1). Synthetic observations were also masked at grid cells where and when snow was experiencing melt~~.~~, identified by the presence of liquid water in the snowpack from the nature run. Although limited in area, grid cells with "ice" landcover (Fig. 1) were also excluded. In this study, this simulation which used assimilation only in unforested, non-melting, and ice-free grid cells is termed "Data Assimilation, without the forest strategy" (DA). In Sect. 2.4 below, we present a novel approach used to infer SWE in grid cells with forests using the nearest canopy-free synthetic observations.

2.4 Extending observations over forests

The 1-d EnKF approach employed here updated each model grid SWE from the open loop simulation based on the observations available at that grid point. Though studies have employed 3-d EnKF approaches to spatially propagate

255 observational information to neighbouring grid cells (De Lannoy et al., 2012)-, here we relied on 1-d updates due to several

256 factors. First, a 2-d update requires the knowledge of spatial error correlations and their variability, which is challenging to

257 specify (Ying, 2020). Most prior studies using such schemes employ uniform specifications and are limited to small domains.

258 Second, a 2-d update increases the size of the state vector and consequently requires the use of a larger ensemble. This,

259 combined with the added computational expense of a 2-d analysis significantly increases the computational cost. Therefore,

260 we employed an alternate approach that is computationally more efficient while allowing the extension of observations to

261 nearby areas.

262 Assuming that the SWE retrievals from the hypothetical SAR instrument are limited over ~~forested areas, here~~areas where

263 the dominant vegetation type are forests (Fig. 1a), we employ a novel approach to extend the observations obtained in non-

264 forested areas (Fig 2). For every forested location, valid retrievals over nearby non-forested locations within a radius of

265 influence of 750 m are identified. An observation at the forested pixel is then estimated by scaling the model SWE by the ratio

266 of the average observed SWE to modeled SWE over the 'clearing' areas (Fig. 2). This scaled observation is then used for

267 assimilation over the forested pixel. Here we implicitly use the spatial correlations inherent in the model between forested and

268 clearing areas to extend observational coverage over the clearing to forested locations. This simulation is termed "Data

269 Assimilation, with the forest strategy" (symbolized by DA+F in Section 3). To evaluate the accuracy and added value of this

270 scaling approach, we compare SWE and runoff from the nature run simulation, versus simulations with data assimilation both

271 1) employing the forest scaling strategy discussed here, and 2) masking synthetic observations in forested grid cells (Sect. 2.3).

$$SWE_{obs,f} = SWE_{model,f} \times \frac{\overline{SWE_{obs,c}}}{\overline{SWE_{model,c}}}$$

$SWE_{obs,f}$: scaled SWE observation value at the forested location

$SWE_{model,f}$: model SWE value at the forested location

$\overline{SWE_{model,c}}$: average model SWE value at the clearing locations within a search radius

$\overline{SWE_{obs,c}}$: average observation SWE value at the clearing locations within a search radius

272

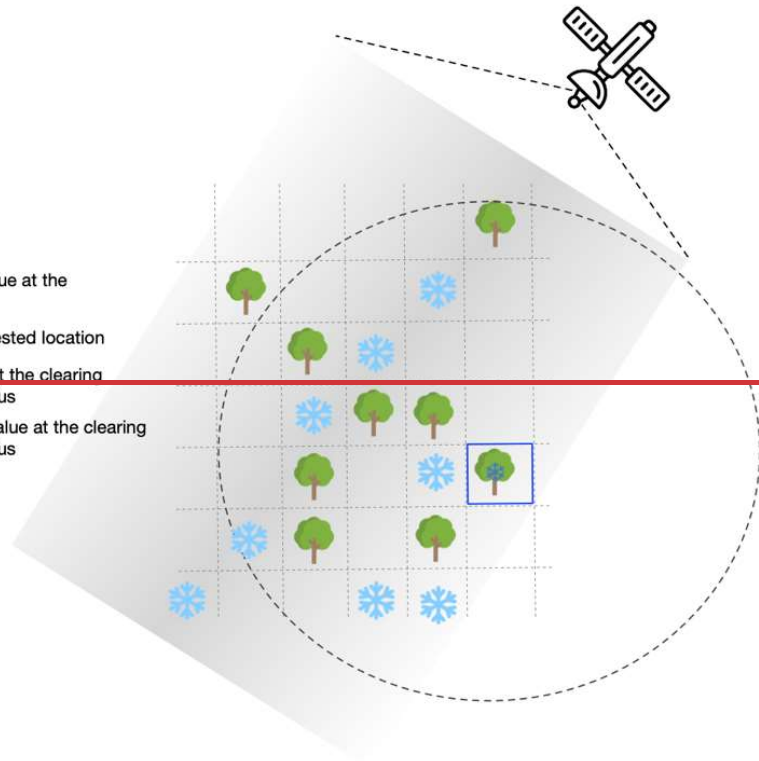
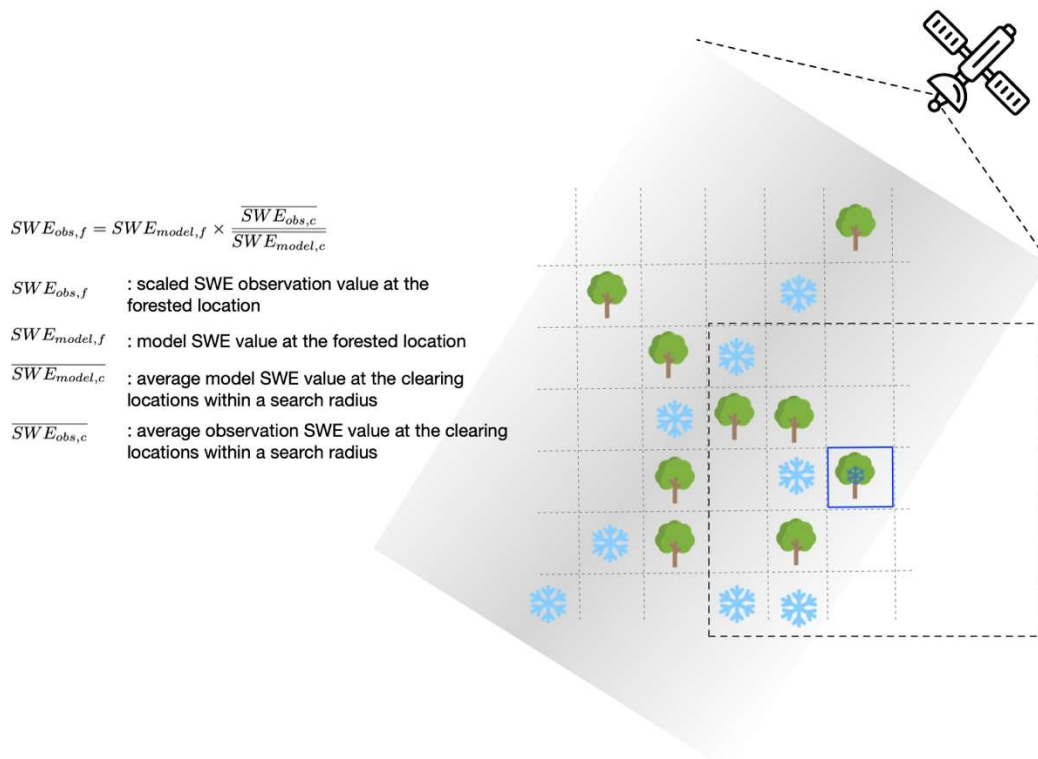



Figure 2. Conceptual depiction and equations demonstrating the forest strategy used here, which estimates a SWE observation at a given grid cell (outlined box in blue color) based on the modeled SWE ($SWE_{model,f}$) and the ratio between the average synthetic SWE observations ($\overline{SWE_{obs,c}}$) and average modeled SWE ($\overline{SWE_{model,c}}$) from grid cells within a 750 m radius (dashed box~~circle~~). The light gray shading represents the satellite swath, the tree icons indicate forested locations, and the snowflake icons represent grid cells with valid SWE retrievals~~measurements~~ at non-forested locations. The grid cell from this example is near the satellite swath edge, so observations are unavailable in the nearby regions South and East of this pixel.

3 Results

In this section, we compute the difference between the open loop simulation, nature run, and the two open loop simulations with data assimilation, one masking synthetic observations over regions with forests, and time periods with melting snow, and ice, and the other applying the same data assimilation but extending snow estimates in forested regions using the strategy from Sect. 2.4 and Fig. 2. The differences between these simulations are detailed in Section 2 and Table S1. We focus on the differences between these four simulations using: 1) average SWE from the winter snow accumulation season (December, January, and February; DJF), when snowmelt is minimized and synthetic observations are masked by grid cells with liquid water content to the smallest degree, 2) spatially distributed SWE on 13 March, the date corresponding to the timing of maximum SWE volume in water-year 2019, and 3) daily average SWE and total runoff for each day in water-year 2019 over a number of selected ~~river basins~~hydrologic regions including the Pacific Northwest, California, Great Basin, and Upper Colorado (Fig. 1b).

291   The open loop and nature run simulations exhibited differences in both the volume and spatial distribution of average winter

292   (December, January, and February; DJF) SWE (Fig. 3a and 3b). Relative to the nature run, the open loop simulation tended to

293   simulate lower elevation winter SWE that was both larger in magnitude and persisted for longer before melting. In the Pacific

294   Northwest domain (Fig. 4), DJF average snow cover (defined as grid cells with mean DJF SWE exceeding 5 mm), was

295   approximately 12% larger for the open loop simulation than the nature run (Table 2). These snow extent biases were also

296   apparent in the other ~~basins~~hydrologic regions (Figs. ~~S1~~ S3 – S5), where open loop snow extents exceeded snow extents

297   from the nature run by 26% in the Upper Colorado, 45% in the Great Basin, and 6% in California. Visually, the nature run had

298   significant increases in the spatial variability of winter SWE, better representing the differences in SWE between mountain

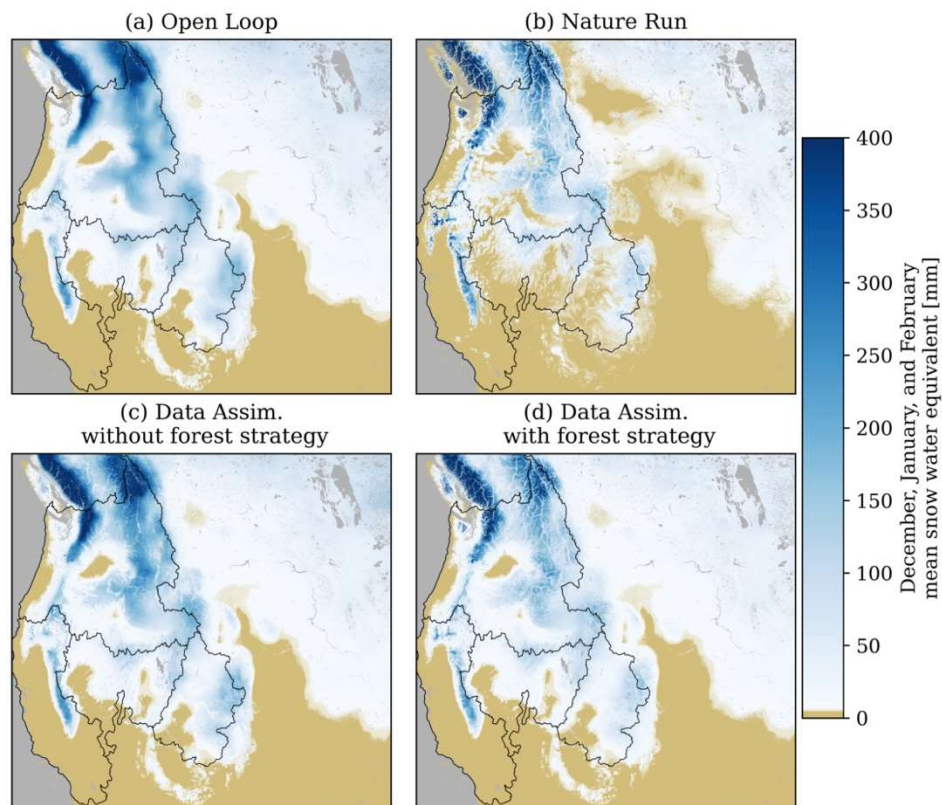
299   peaks and valleys, and the patchiness of snow cover in regions with winter snowmelt and ephemeral snow cover (e.g., Fig. 4,

300   Fig. S1). Relative to the nature run, DJF SWE from the open loop simulation was biased high across the full modeling region

301   (Fig. 3) by approximately 26%, on average, with a mean absolute error of 41 mm and spatial coefficient of correlation of

302   approximately 0.74. Across ~~snowy~~ the Pacific Northwest ~~basin~~ (Fig. 4), DJF mean SWE biases were approximately 37%, with

303   a mean absolute error of 55 mm. Open loop model performance for the other ~~basins~~hydrologic regions can be found in Table

304   2.

305

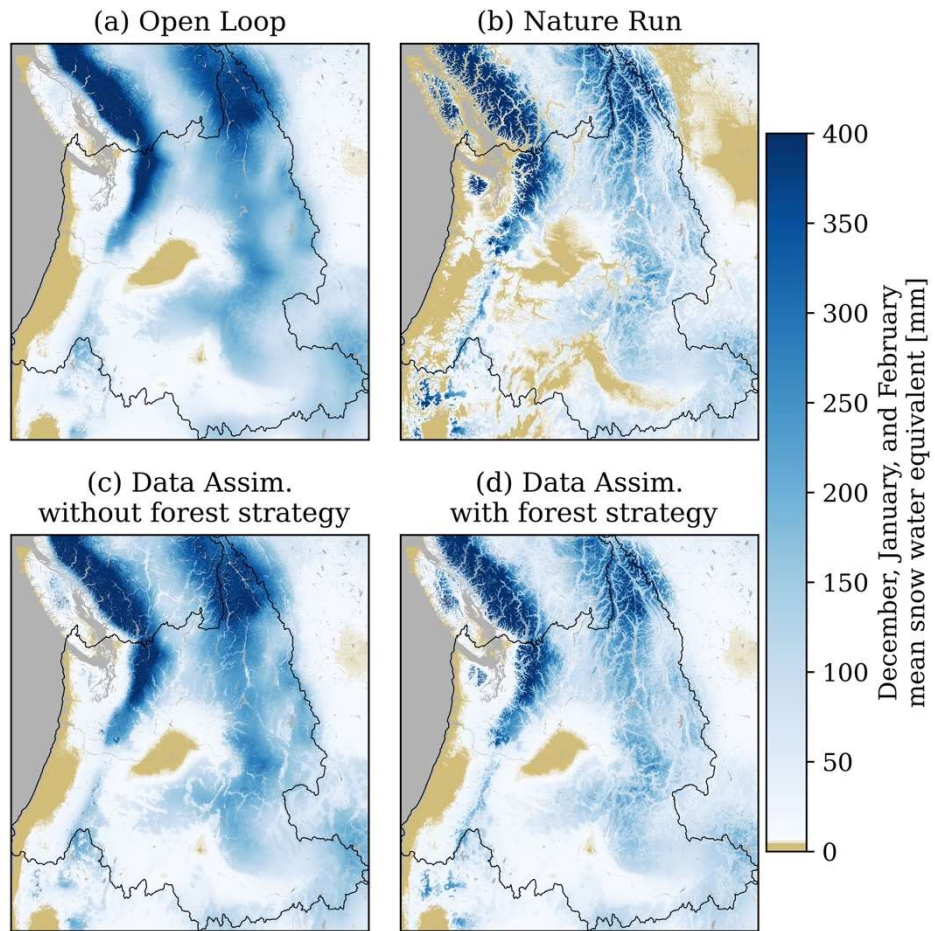
306   **Figure 3.** Winter (December, January, and February) mean SWE simulated at 250 m resolution from the open loop (a), nature
307   run (b), and data assimilation simulations, both with (d) and without (c) the forest strategy presented in Sect. 2.4.

308
309   **Figure 4.** Winter (December, January, and February) mean SWE in the Pacific Northwest region simulated at 250 m
310   resolution from the open loop (a), nature run (b), and data assimilation simulations, both with (d) and without (c) the forest
311   strategy.

Table 2. Simulation performance, relative to the nature run simulation, for the open loop simulation (OL) and the simulations with data assimilation, both with (DA+F) and without (DA) the forest strategy. Statistics are presented for the full domain, the four hydrologic ~~basins~~regions, and all forested and unforested grid cells.

		DJF* snow-extent biases	13 March 2019 SWE			Seasonal SWE and runoff	
			Mean bias	SWE abs. error [mm]	Coeff. of corr.	MAM* mean SWE bias	Nash-Sutcliffe Efficiency
Full study domain	OL^	+22%	+26%	41	0.74	-	-
	DA^	+23%	+9%	36	0.79	-	-
	DA+F^	+22%	+4%	17	0.91	-	-
Upper Colorado	OL	+26%	+37%	55	0.74	+63%	-2.59
	DA	+28%	+27%	50	0.74	+86%	-3.71
	DA+F	+28%	+8%	23	0.90	< 1%	0.22
Pacific Northwest	OL	+12%	+42%	89	0.69	+44%	-0.17
	DA	+13%	+32%	80	0.74	+80%	-0.34
	DA+F	+13%	+6%	35	0.89	+15%	0.39
Great Basin	OL	+45%	+35%	38	0.62	-29%	0.58
	DA	+46%	+46%	32	0.75	+10%	0.58
	DA+F	+46%	+28%	23	0.83	-38%	0.53
California	OL	+6%	-34%	50	0.64	-50%	0.92
	DA	+8%	-6%	40	0.79	-15%	0.88
	DA+F	+8%	-6%	28	0.88	-26%	0.89
Unforested	OL	+19%	+14%	22	0.83	-	-
	DA	+20%	< 1%	14	0.91	-	-
	DA+F	+20%	< 1%	14	0.91	-	-
Forested	OL	+29%	+150%	111	0.67	-	-
	DA	+30%	+150%	111	0.67	-	-
	DA+F	+30%	+18%	27	0.93	-	-

*DJF = December, January, and February; MAM = March, April, and May (averages)
^OL = open loop simulation; DA = ~~simulation with~~ data assimilation without the forest strategy; DA+F = ~~simulation with~~ data assimilation ~~and~~with the forest strategy

As expected, the simulations assimilating the synthetic SWE observations agreed with the nature run better than the open loop simulation. However, on 13 March 2019 (the date of maximum domain SWE volume), the simulation with data assimilation without the forest strategy had high-biased SWE across large portions of the Rocky Mountains and the Cascade Mountain range (Fig. 1, Fig. 5b and Fig. 5e). Low biased SWE was more common in Northernmost Canadian portions of the Rocky Mountains and Cascade Range, the Western montane regions in Washington State, the Northern portions of the Great Basin, and the lower-lying elevations of the California Sierra Nevada. Additionally, despite the assimilation, snow extents were still biased high relative to the nature run (Fig. 3) at magnitudes similar to the open loop simulation (Table 2). This was driven by the expansive snow extents of the open loop simulation, which were decreased by data assimilation, but still resulted in widespread early-season SWE increases for short periods of time between synthetic observations (at 10 – 14 day

324 frequencies), increasing to the number of grid cells with DJF SWE exceeding 5 mm (threshold used to define average winter

325 snow extents in Fig. 3).

326   Assimilating the synthetic SAR observations without the forest strategy best improved SWE in shrub, grass, crop, bare, and

327 wetland landcover types (Fig. 6b and 6c). For example, relative to the open loop simulation (Fig. 5a and 5d), data assimilation

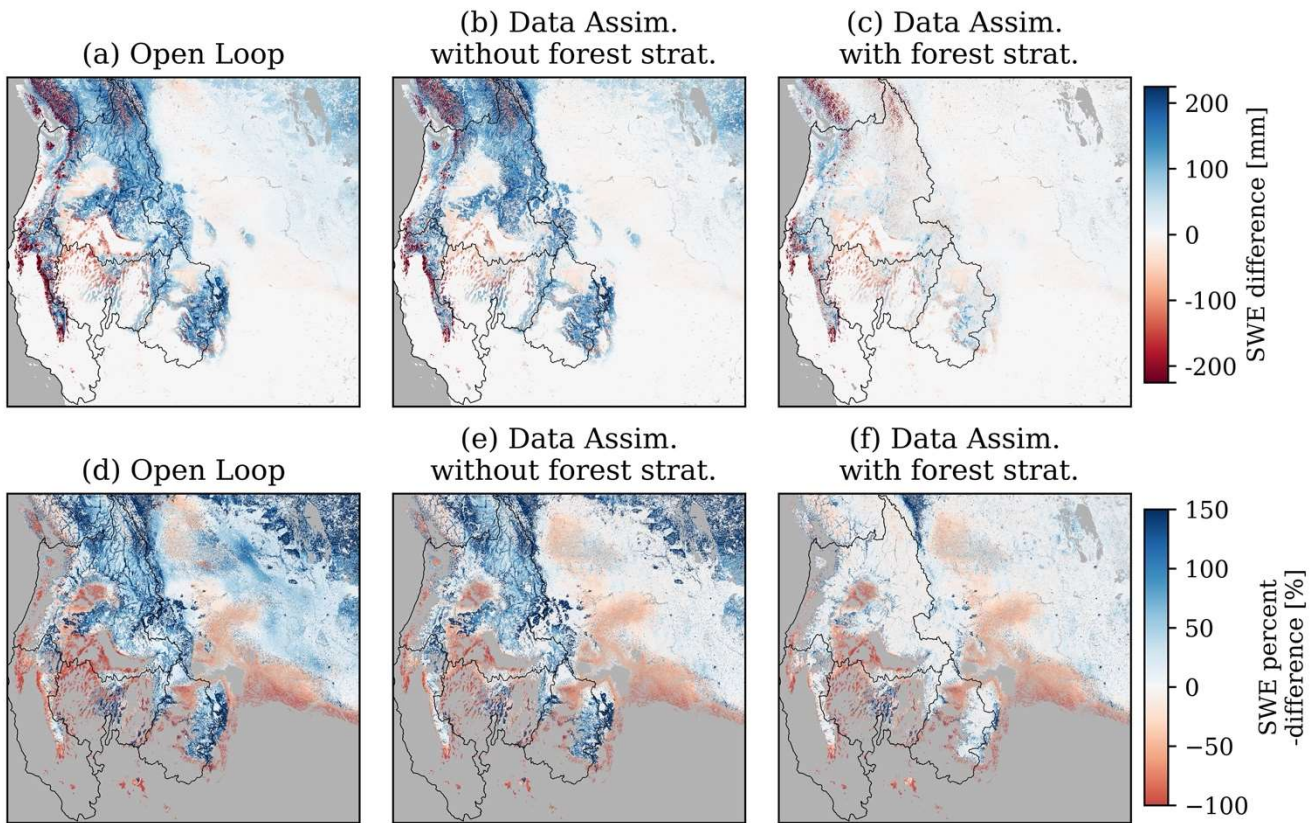
328 without the forest strategy (Fig. 5b and 5e) corrected the high SWE biases in the Great Plains (Fig. 1). While 13 March SWE

329 in shrub, grass, crop, bare-ground, and wetland regions was typically small in magnitude, these landcover types accounted for

330 77% of the modeling domain area, and 61% of the domain total SWE volume on 13 March (Fig. 6a). In these regions, SWE

331 from the open loop simulation had a mean absolute error of 22 mm, and a mean bias of approximately 14%, relative to the

332 nature run (Table 2). Data assimilation significantly improved the SWE bias in these land cover types to within 1%, on average

333 (Fig. 6b), with a mean absolute error of 14 mm, relative to the nature run.

334
335 **Figure 5.** 13 March 2019 SWE difference (top row) and percent-difference (bottom row), relative to the nature run, for the
336 open loop simulation (a and d), and simulations with data assimilation, both with (c and f) and without (b and e) the forest
337 strategy. SWE percent-different maps (bottom row) only compare grid cells where SWE from the nature run was greater
338 than 5 mm.
339

340   The data assimilation results discussed above did not use the synthetic observations over forested grid cells, where

341   retrievals from SAR instruments may be either partially or fully occluded by the canopy overstory (Tsang et al., 2022; Ruiz

342   et al., 2022; Huang et al., 2019). However, a significant portion of the snow volume in mid-latitude domains overlaps with

343   forests. For example, although forests only covered approximately 22% of the study region investigated here (Fig. 1a),

344   forested grid cells contained just over 34% of the total 13 March SWE volume, a volume about 10% higher than the snow

345   volume contained in the next-largest landcover type (Fig. 6a). In forested grid cells, SWE simulated by the open loop

346   simulation were biased high by approximately 87 mm (+150%) on average (Fig. 6), with a mean absolute error of 111 mm

347   (Table 2). These errors were propagated into the simulation with data assimilation without the forest strategy. Fortunately,

348   the ratio between modeled SWE and synthetic SWE observations in forested grid cells and the nearest canopy-free grid cells

349   had high levels of similarity. Therefore, estimating snow in forest regions using the nearest canopy-free pixels (Fig. 2)

350   improved snow simulations significantly (Fig. 3d, Fig. 4d, and Fig. 5c and 5f). In fact, snow simulated in forest landscapes

351   using data assimilation with the forest strategy agreed well with the nature run, exhibiting a 13 March SWE average bias in

352   forested grid cells of only 14 mm (+8%) (Fig. 6), and a mean absolute error of 27 mm. This forest strategy resulted in large-

353   scale improvements to total domain SWE (Fig. 5), reducing the 13 March full-domain SWE volume bias by 28%, and

354   improving the spatial coefficient of correlation by 0.12, relative to the data assimilation simulation without the forest

355   strategy.

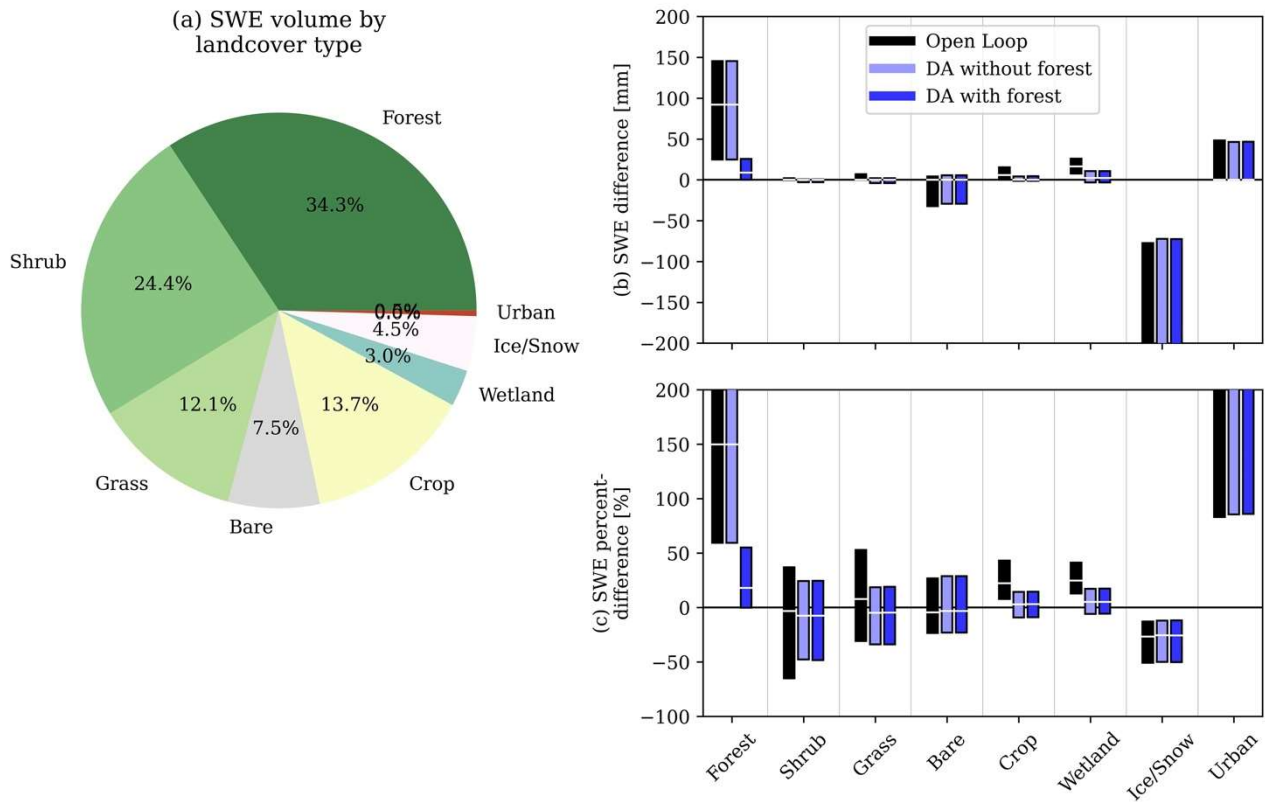



Figure 6. SWE volume on 13 March 2019 broken down by landcover type in subplot a. For each landcover type, the interquartile range and median of SWE differences (b) and SWE percent-differences (c) are calculated for the open loop simulation (black) and each simulation with data assimilation (blue bars). SWE differences (b) and (c) are calculated relative to the nature run.

The comparisons above focused on mean DJF SWE and SWE from the date nearest ~~peak~~maximum snow volume (13 March, 2019). However, assimilating the synthetic SWE data also improved estimates of snow water resources throughout the duration of the water year, even in periods when most snow-covered regions were experiencing snowmelt and synthetic observations were masked. For example, in the Upper Colorado ~~basin~~, approximately 75% of the ~~basin~~region had DJF snow cover with little or no winter snowmelt (Fig. 7). The simulation with data assimilation and the forest strategy substantially improved ~~basin~~ mean SWE evolution in the snow accumulation season in this ~~basin~~hydrologic region (Fig. 7, October - March). However, snowmelt onset in the March, April, and May (MAM) months increased the number of grid cells experiencing snowmelt from the open loop model outputs, reducing the number of grid cells across the full Upper Colorado Region that could be observed by the synthetic SAR observations to approximately 5%, on average, over this period of time. Despite this, since the simulation with data assimilation improved the volume, timing, and spatial distribution of ~~peak~~maximum SWE, mean SWE evolution tracked the nature run simulation significantly better than the open loop simulation in the spring snowmelt period. In fact, relative to the nature run, MAM SWE from the open loop simulations was

373 biased high by approximately 63%, on average, in the Upper Colorado (Table 2). The simulation with data assimilation

374 using the forest strategy improved this bias to less than 1%, on average, over the same period. In this study, simulations

375 using Noah-MP (open loop and data assimilation simulations) melted snow more rapidly in the later-half of the spring

376 snowmelt season than the nature run simulation which evolved SWE using SnowModel (Section 2.1). Therefore, although

377 ~~peak~~maximum SWE volume, ~~peak~~maximum SWE timing, and MAM SWE were improved by data assimilation, the timing

378 of snow disappearance for the simulation with data assimilation using the forest strategy was approximately 18 days earlier

379 than the nature run in the Upper Colorado.

380

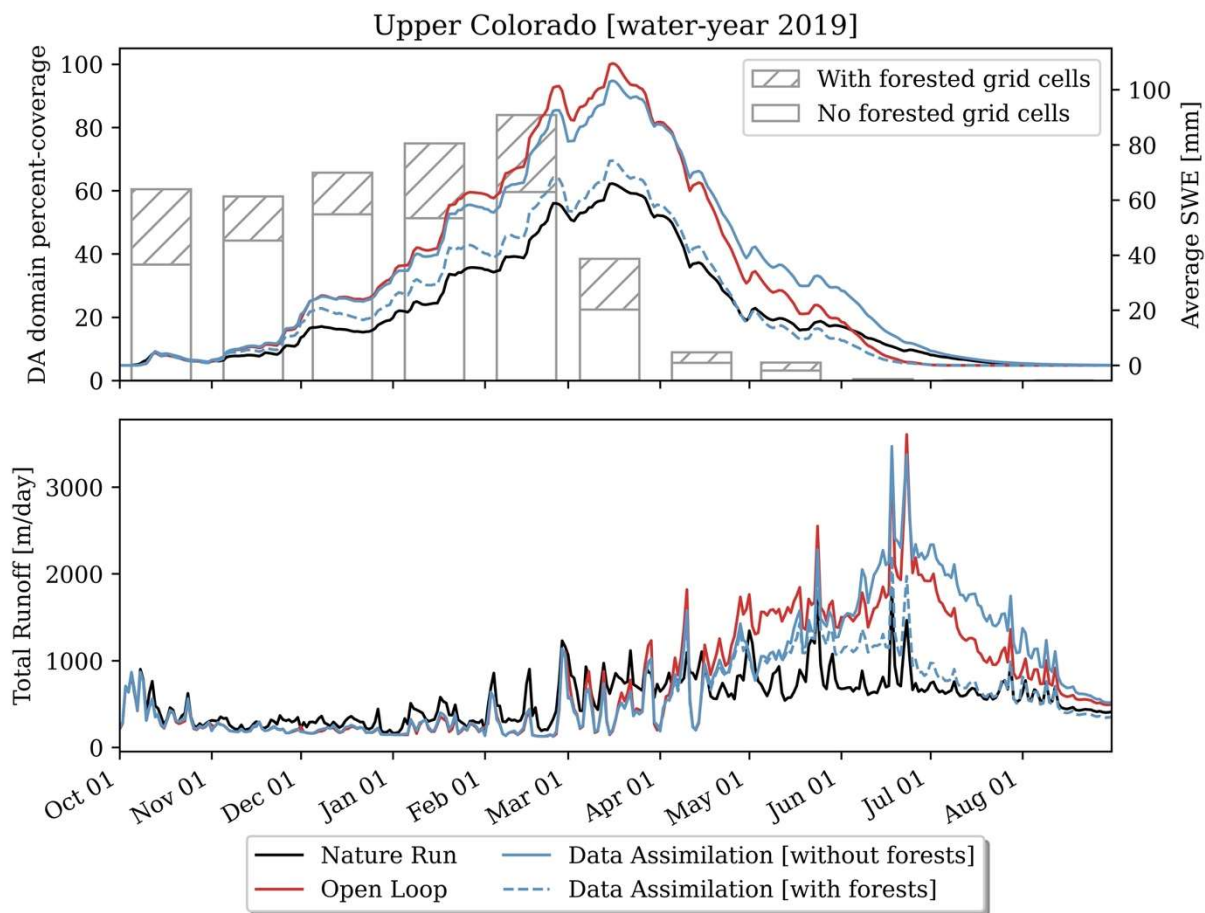



Figure 7. Time series comparison of ~~basin~~ mean SWE (top) and total runoff (bottom) between the open loop, nature run, and
simulations assimilating the synthetic observations, both with and without the forest strategy in the Upper Colorado ~~basin~~.
Dashed bars in the top plot represent the monthly percentage of the Upper Colorado grid cells with no snowmelt. Solid bars
also exclude grid cells with forest coverage.

385 Much like the Upper Colorado ~~Basin~~, SWE simulated by the open loop simulation in the Pacific Northwest (Fig. ~~S5~~S7)

386 was biased high for the entirety of the snow season. Both domains also had greater than 80% synthetic snow observation

~~19~~

387   coverage in March (including grid cells that filled snow estimates using the forest strategy), and as a result, the simulation

388   with data assimilation using the forest strategy closely matched SWE from the nature run. However, both of these domains

389   had a significant portion of the seasonal snowpack in forested landcover (Fig. 7 and Fig. ~~S5~~S7, difference between the

390   hatched and solid bars). These grid cells had winter SWE estimates from the open loop simulation that were predominately

391   high-biased (Fig. 3 and Fig. 5). Therefore, although data assimilation improved winter SWE in non-forested landcover types

392   (Fig. 6), the simulation without the forest strategy caused little-to-no improvement in the simulated domain mean

393   ~~peak~~maximum SWE (Table 2). This highlights the value of the forest strategy used here (Fig. 2), which drew information

394   from synthetic observations in relatively few nearby pixels to infer the mean snow volume in forested grid cells. Given the

395   four ~~basins~~hydrologic regions investigated in this study, a far smaller volume of snow existed in forested landcover for the

396   California (Fig. ~~S6~~S8) and Great ~~Basins~~Basin regions (Fig. ~~S4~~S6), resulting in DJF domain-mean SWE evolution that was

397   more similar between the simulations with and without the forest strategy. We expect results in these domains to be more

398   indicative of the value of winter SAR observations in less-vegetated snowy landscapes, such as Tundra and Prairie snow

399   regimes (Sturm and Liston, 2021).

400   Finally, the improvements to the spatial and temporal estimates of SWE discussed above had trickle-down improvements

401   on simulated runoff. For example, in the Upper Colorado (Fig. 7), total annual runoff from the open loop simulation was

402   biased high by approximately 35%, relative to the nature run. This error was driven most by high-biased winter snow

403   accumulation, which nearly doubled the melt season (March – July) runoff estimated by the nature run simulation. Here, by

404   assimilating the synthetic SWE observations, and estimating forest snowpack from the relationship between modeled and

405   observed SWE from the nearest canopy-free pixels, total annual streamflow in this domain was improved to within 1%. Not

406   only was domain total runoff improved, but the seasonal evolution of high and low-flows vital for water management and

407   planning was also improved. This improved the Nash-Sutcliffe Efficiency (NSE) from –2.59 to 0.22 between the open loop

408   simulation and simulation with data assimilation employing the forest strategy (Table 2). These results were similar for the

409   Pacific Northwest, which had an NSE that improved from –0.17 to 0.39. ~~Here~~However, due to the smaller changes to SWE

410   and more-rapid snowmelt simulated by Noah-MP, changes to runoff from data assimilation in California and the Great Basin

411   were small (Table 2~~).~~), with improvements that were largely outweighed by the difference in snowmelt timing and rates

412   between Noah-MP and SnowModel.

413   **4. Discussion**

414   The differences between the open loop simulation and nature run in this study were representative of ~~typical~~ snow modeling

415   errors ~~reported in the literature~~common for continental and global-scale models used for seasonal to long-term future snow

416   predictions (e.g., Franz et al., 2010; Garousi-Nejad and Tarboton, 2022; Kim et al, 2021; Liu et al, 2022). The greatest source

417   of these snow modeling errors is commonly errors in meteorological forcing data, and in particular, biases in precipitation

(Garousi-Nejad and Tarboton, 2022; Henn et al., 2018; Pflug et al., 2021; Raleigh and Lundquist, 2012; Wayand et al., 2013). These biases are especially prevalent in the portions of the earth's surface with the greatest volumes of snow, such as the tundra and montane regions (Kim et al., 2021), where ground observations and observation station maintenance are hindered by harsh winter conditions and inaccessibility. This suggests that the greatest need for improving global estimates of snow is improved estimates of snow accumulation in remote, under sampled landscapes. Here, we expect that the SAR observations evaluated in this study could address these needs., thus providing a path forward for pairing common snow models with observations as a basis for determining global snow mass. For example, assimilating SAR observations at 10 – 14 day intervals with the observational error characteristics reported in Sect. 2.2, improved midlatitude winter SWE volume by approximately 22%, on average (Table 2). In unforested landscapes, which account for a majority of the Earth's snow water storage (Kim et al., 2021), assimilation improved the mean SWE bias at peak maximum SWE timing to within 1%, on average, and reduced the standard deviation of errors by approximately 45 mm (~85%) (Fig. 6).

Despite the benefits discussed above, SAR observations have known limitations in forested landscapes where the canopy overstory obstructs retrievals from the underlying snowpack (Huang et al., 2019; Ruiz et al., 2022; Tsang et al., 2022). Therefore, this study was designed to investigate a forest strategy that uses the relationship between modeled SWE estimates and synthetic SWE observations from neighboring grid cells as the basis for inferring snow distribution in regions with forested landcover (Fig. 2). To focus on the benefits of this approach, we chose a domain (Fig. 1) that included both significant forest spatial coverage (22%) with disproportionate amounts of winter snow (34%) within the forested pixels (Fig. 6). Relative to the open loop simulation, the simulation with data assimilation and the forest strategy dramatically improved the spatial distribution of SWE (e.g., Fig. 3 and Fig. 4) and the resulting SWE biases at peak domain maximum snowpack timing (Fig. 5). In fact, in forested grid cells, SWE on 13 March was only biased by 14 mm (mean absolute error of 27 mm), on average, for the simulation with data assimilation and the forest strategy., relative to the nature run. This was opposed to the open loop simulation, which was biased by 87 mm (mean absolute error of 111 mm) over the same regions and date. Despite the fact that the two simulations with data assimilation agreed in all grid cells except forested grid cells, the simulation employing the forest strategy had a mean absolute error (17 mm) across the full modeling region that was approximately 51% smaller than the simulation without the forest strategy. Here, we recognize that this study used a single date (13 March) to represent snow water resources at maximum SWE timing. However, the date of maximum SWE volume from the nature run varied by less than a week across the four hydrologic regions (11 - 16 March; Fig. 7, Fig. S6 - S8). Therefore, this was a relevant date for model comparisons, especially given that water resource and allocation decisions in the Western US are often based on the volume of snow at maximum snow timing.

This research shows how a modeling framework and relatively few observations can be used to gap-fill estimates of snow in regions where remote sensing observations from a future platform may be most challenged. Despite the fact that snowpack with properties able to be retrieved by SAR instrumentation (i.e., canopy-free landcover and no snowmelt) sometimes only

450 accounted for only small portions of a modeling domain (e.g., Fig. 7), SWE from the model and SAR observations in nearby

451 canopy-free grid cells were predictive of the snow in forested grid cells. We hypothesize that this could have partly been driven

452 by the 250 m resolution of synthetic observations and simulations. At this length scale, snow distribution is typically driven

453 by processes like ~~synoptic~~mesoscale weather patterns and their interaction (e.g., orographic lapse rates, wind

454 loading/sheltering, terrain-shading, etc.) with static topographical features like elevation, slope, and aspect (e.g., Clark et al.,

455 2011; Lehning et al., 2011; McGrath et al., 2018; Minder et al., 2008; Trujillo et al., 2007). However, we acknowledge that

456 snow in forested and open grid cells is subject to different snow processes. In fact, the nature run simulation used here attempts

457 to simulate snow-canopy interactions, such as snow interception and solar shading from the canopy overstory (Liston and

458 Elder, 2006). Here, since we focus predominantly on model improvements from data assimilation in the SWE accumulation

459 season, we hypothesize that the primary difference between SWE accumulation in forested pixels, and the nearest canopy-free

460 grid cells could be driven by canopy interception, or the lack thereof. In other words, inferring forested snowpack using the

461 nearest canopy-free grid cells could bias snow in forested regions where snow processes differ slightly. While the forest

462 strategy improved SWE simulated in forested grid cells at the date of ~~peak~~ maximum SWE volume, SWE was still biased high

463 relative to the nature run (Fig. 6). We hypothesize that a correction factor, based on variables like forest canopy type, vegetation

464 density, wind speed, and temperature during snowfall, all of which influence snow interception, (Lundquist et al., 2021), could

465 be used to facilitate the difference in snow accumulation expected between a forest pixel and SWE observations from nearby

466 canopy-free grid cells. This approach will be a topic of future research. However, since errors with precipitation are often the

467 overwhelming source of model errors, we hypothesize that the forest strategy (Fig. 2), which corrected modeled SWE in

468 forested areas using the ratio between modeled and observed SWE in nearby open areas, was well-suited to correct precipitation

469 biases.

470   The results presented here are subject to a number of assumptions. These assumptions were intended to apply regionally-

471 consistent and conservative rules about how 1) synthetic SAR observations were generated, and 2) the grid cells and time

472 periods that SAR observations occurred in. For example, we used a 20% and zero-mean random distribution of errors to

473 generate observations from the nature run. We expect the error from a future satellite mission to be less than 20% over the

474 majority of snow covered regions (Sect. 2.2). However, observational biases may be more common in certain locations and

475 periods based on snow depth, particularly in very shallow or very deep snowpacks, terrain characteristics and vegetation

476 characteristics. Additionally, the landcover classification used in this study (Fig. 1) was based on the dominant landcover type

477 within each model grid cell, as defined from the North American Land Change Monitoring System (Latifovic et al., 2017). For

478 forested grid cells, this included needleleaf, broadleaf, and mixed forest types. To be conservative, this study completely

479 masked synthetic observations in 250 m grid cells classified as forest, thereby assuming 1) no observation capabilities in

480 predominantly forested areas, and 2) full observation capabilities in grid cells where forests were not the dominant landcover

481 type. In reality, SAR may be able to achieve accurate snow retrievals in some forested-dominated regions based on the forest

482 type, forest distribution, and canopy density (Tsang et al., 2022). Conversely, some regions with sparser or no forest cover

483 may still have observation limitations based on the domain and snow characteristics mentioned above. The large domain used

484 in this study also made tests over multiple years computationally challenging. Here, the intent of this study was to investigate

485 a strategy for deriving SWE corrections in difficult to observe forest landscapes, and we hypothesize that precipitation biases

486 and the resulting modeled SWE accumulation could be improved to a similar degree in years with both larger and smaller

487 snow volumes. Finally, while strategies for identifying and correcting systematic SAR observation errors are a topic of

488 continued research (e.g., Durand et al., 2023; Singh et al., 2023), OSSEs are an inherently flexible framework for evaluating

489 sensor utility, so future research could use the simulations performed here to test a wider array of sensor configurations and

490 non-normal retrieval errors. Future work could build upon these results to investigate multiple years, perhaps considering

491 warmer and/or drier snow years, when the role of snowpack for water supply and midwinter snowmelt and rain-on-snow

492 frequency may be more likely to increase snowpack liquid water content, or years with late-season spring snow accumulation.

493 Future research should also investigate other gap-filling approaches, like methods to infer SWE in grid cells where snowmelt

494 is occurring and liquid water may prevent SAR retrievals, and gap-filling approaches using different window sizes and/or

495 searching windows that more heavily weight unforested grid cells with similar characteristics (elevation, aspect, etc.).

496 This study tested a simple model setup using a popular land surface model (Noah-MP) and Kalman-based data assimilation

497 procedure. This data assimilation procedure updated modeled snow states, like snow depth and SWE, based only on synthetic

498 SWE observations at 10 – 14 day temporal frequencies where/when snowmelt was not occurring. WeDespite the limitations

499 and assumptions discussed above, we expect that the results presented here could represent the lower-bound of performance

500 that could be achieved from a real-time modeling framework that could accompany a space-borne SAR remote sensing

501 platform. For example, many studies have demonstrated repeatable patterns of snow accumulation in years with similar winter

502 meteorological characteristics (e.g., Deems et al., 2008; Pflug et al, 2022; Schirmer et al., 2011; Sturm and Wagner, 2010;

503 Woodruff and Qualls, 2019). This suggests that retrospective information about snow distribution patterns in previous years,

504 could be used as the basis for extrapolating and updating snow model states in grid cells not covered by SAR observations on

505 a given date. From the modeling perspective, only 5 ensemble members were used in the Ensemble-Kalman data assimilation

506 (Sec 2.3), when a larger ensemble of simulations may have improved uncertainty characterization of simulated snow and

507 hydrological states even more. This study also assumed that synthetic SAR observations were unable to observe snow in all

508 forested landscapes, when retrievals of snow in forested stands could be achievable for some forested regions with smaller tree

509 cover fractions and biomass (Montomoli et al., 2015; Tsang et al., 2022). Finally, the SAR configuration tested here had 10 –

510 14 day repeat times, but future satellite configurations with more-frequent observational repeats could be possible.are possible

511 and have been recommended by the 2018 Decadal Survey (NASEM 2018). Despite all of these conservative assumptions, the

512 difference between the open loop simulation (representative of current modeling accuracies), and the simulation with synthetic

513 observation data assimilation using the forest strategy, demonstrated large-magnitude and widespread improvements to real-

514 time estimates of winter SWE and the associated improvement to spring SWE and runoff. Therefore, we expect the findings

515 of this study, particularly the strategy to extend the observational utility to forested areas, to significantly aid in the full

516 exploitation of the information from a future SAR-based snow satellite mission.

517 5. **Conclusion**

518 Global estimates of snow volume and distribution have uncertainties stemming from limited snow observations and

519 biases in meteorological forcing data. These uncertainties stress the need for a global snow-focused satellite remote

520 sensing platform. Here, we investigate the degree to which synthetic observations of ~~snow water equivalent (~~SWE

521 representative of a Synthetic Aperture Radar (SAR) remote sensing platform, could correct common snow modeling

522 errors and provide spatiotemporally continuous SWE estimates. We investigate this using an ~~Observation~~Observing

523 System Simulation Experiment, specifically investigating how much snow simulated using a ~~popular~~widely used land

524 surface model and meteorological forcing dataset, could be improved by assimilating synthetic SAR observations of

525 SWE.

526 The difference between the open loop simulation and the nature run ~~were~~was representative of common modeling

527 errors. Snow simulated by the open loop simulation had larger winter snow extents, and total snow volume that was

528 biased high by approximately 35%. The open loop simulation also simulated snow that was more spatially homogeneous,

529 underestimating the variability across variations in topography and underestimating lower-elevation snowmelt from the

530 nature run. Assimilating the synthetic SWE observations improved SWE simulated in the shrub, grass, crop, bare-ground,

531 and wetland land cover types. In fact, SWE biases on the date of domain ~~peak~~maximum SWE volume (13 March 2019)

532 in these landcover types improved from 14% for the open loop simulation to within 1% after data assimilation. However,

533 despite only covering 22% of the study area, forested grid cells contained just over 34% of the domain SWE on 13

534 March. The open loop simulation and the simulation with data assimilation without the forest strategy had SWE that was

535 high biased by 150% (87 mm), on average, in these forested grid cells. The relationship between modeled SWE and

536 synthetic SWE observations in forested grid cells exhibited similarities with the nearest canopy-free grid cells. Therefore,

537 SWE in forested regions was able to be inferred using the simple modeling framework and synthetic SAR observations

538 from nearby canopy-free grid cells. In fact, the simulation with data assimilation using this forest gap-filling strategy

539 substantially improved SWE biases to 4% (~22% improvement) at ~~peak~~ maximum SWE timing, with a SWE mean

540 absolute error of 17 mm (24 mm improvement) and spatial correlation of 0.91 (0.17 improvement) across the Western

541 US

542 Improvements in winter SWE accumulation also improved estimates of melt-season SWE evolution and total runoff

543 in four major Western ~~United States hydrological basins~~US hydrologic regions, even in periods when winter snowmelt

greatly reduced the number of grid cells that could be observed by the synthetic SWE observations. In fact, in the Upper Colorado River basin, melt season SWE biases improved from 63% to less than 1% after assimilation, and the runoff Nash Sutcliffe Efficiency improved from -2.59 to 0.22. These results demonstrate the value of SAR observations and simple spatial-filling strategies in grid cells where SAR retrievals could be obstructed by the canopy. Here, we expect our results to represent a lower-boundary of model performance which could be improved further by more robust assimilation approaches, more-frequent SAR observations, further developments to SAR retrieval algorithms in forested landscapes, and adaptations to the forest gap-filling strategy developed here. However, our results also show that widespread improvements to global SWE could be available in near real-time provided data assimilation approaches and a SAR remote sensing platform.

Code availability: The Land Information System (LIS; lis.gsfc.nasa.gov) framework used to perform the nature run, open loop, and data assimilation simulations from this study can be accessed from a GitHub public repository (https://github.com/NASA-LIS/LISF). Model documentation and LIS tutorials can also be accessed from this repository. Users are encouraged to reference Kumar et al. (2006) for more information on LIS. The Trade-space Analysis Tool for designing Constellations (TAT-C) tool is currently available on-request for federal employees and contractors (https://software.nasa.gov/software/GSC-18399-1).

Data availability: The model outputs and data necessary to reproduce the figures and statistics reported in this study can be found at https://www.hydroshare.org/resource/e0ad80f818bf4062a335e9e0d7362834/. This repository includes domain static variables, such as land cover, elevation, and spatial coordinates, in addition to model outputs of winter-average SWE, SWE at the date of peak maximum SWE volume (13 March 2019), and basin aggregated SWE and runoff aggregated across each region. MERRA-2 forcing data can be accessed from the Goddard Earth Sciences Data and Information Services Center (GES DISC, https://disc.gsfc.nasa.gov/), and ERA5 data can be accessed from the European Centre for Medium-Range Weather Forecasts climate data store (https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5).

Author contributions: CV and SK coordinated the manuscript question and research methodology. SK adapted LIS model source code to implement the forest gap-filling strategy in Sect. 2.4. MW set up the model domains and model configurations, and performed the open loop simulation. EC assisted with generating the synthetic SWE observations. With assistance from KA, JP performed code developments for the simulations using both the Noah-MP and coupled SnowModel models. KA implemented the SnowModel code and parameters into LIS and LDT. JP also performed the nature run and both data assimilation simulations. The manuscript was written provided text, figures, and feedback from all coauthors.

Competing interests: The authors declare that they have no conflict of interest.

573 *Acknowledgements:* This work was supported by the funding from the NASA Terrestrial Hydrology program. Computing was

574 supported by the NASA Center for Climate Simulation (NCCS).

References

576 Arsenault, K.R., Wrzesien, M., Gutmann, E.D., Vuyovich, C., Liston, G.E., Mower, R., Reinking, A., Newman, A.J., Kumar,
577   S.V., Wang, S., Navari, M., Forman, B.A., Jessica, L.: Implementing SnowModel into the Land Information System
578   Framework to Support High Resolution Modeling of Snow Heterogeneity. Presented at the AGU Fall Meeting 2021,
579   AGU, 2021.
580 ~~Arsenault, K.R., Kumar, S.V., Geiger, J.V., Wang, S., Kemp, E., Mocko, D.M., Beaudoing, H.K., Getirana, A., Navari, M.,~~
581   ~~Li, B., Jacob, J., Wegiel, J., Peters Lidard, C.D.: The Land surface Data Toolkit (LDT v7.2) a data fusion~~
582   ~~environment for land data assimilation systems. Geoscientific Model Development 11, 3605 3621.~~
583   ~~https://doi.org/10.5194/gmd 11 3605 2018, 2018.~~
584 Barnett, T.P., Adam, J.C., Lettenmaier, D.P.:Potential impacts of a warming climate on water availability in snow-dominated
585   regions. Nature 438, 303. https://doi.org/10.1038/nature04141, 2005.
586 Barry, R.G.: The Role of Snow and Ice in the Global Climate System: A Review. Polar Geography 26, 235–246.
587   https://doi.org/10.1080/789610195, 2002.
588 Best, M.J., Pryor, M., Clark, D.B., Rooney, G.G., Essery, R.L.H., Ménard, C.B., Edwards, J.M., Hendry, M.A., Porson, A.,
589   Gedney, N., Mercado, L.M., Sitch, S., Blyth, E., Boucher, O., Cox, P.M., Grimmond, C.S.B., Harding, R.J.: The Joint
590   UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes. Geoscientific Model
591   Development 4, 677–699. https://doi.org/10.5194/gmd-4-677-2011, 2011.
592 Beven, K.J., Kirby, M.J., Freer, J.E., Lamb, R.: A history of TOPMODEL. Hydrology and Earth System Sciences 25, 527 549.
593   https://doi.org/10.5194/hess-25-527-2021, 2021.
594 Cho, E., Vuyovich, C. M., Kumar, S. V., Wrzesien, M. L., & Kim, R. S.~~.: (2023).~~ Evaluating the ~~Utility of Active Microwave~~
595   ~~Observations~~utility of active microwave observations as a ~~Snow Mission Concept Using Observing System~~
596   ~~Simulation Experiments.~~snow mission concept using observing system simulation experiments. The Cryosphere
597   ~~Discussions 1 24. https://doi.org/10.5194/tc-2022-220, 2022,~~ 17(9), 3915-3931.
598 Clark, M.P., Hendrikx, J., Slater, A.G., Kavetski, D., Anderson, B., Cullen, N.J., Kerr, T., Hreinsson, E.Ö., Woods, R.A.:
599   Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review. Water
600   Resources Research 47. https://doi.org/10.1029/2011WR010745, 2011.
601 Cosgrove, B.A., Lohmann, D., Mitchell, K.E., Houser, P.R., Wood, E.F., Schaake, J.C., Robock, A., Marshall, C., Sheffield,
602   J., Duan, Q., Luo, L.: Real-time and retrospective forcing in the North American Land Data Assimilation System
603   (NLDAS) project. Journal of Geophysical Research: Atmospheres, 108.D22, 2003.
604 De Lannoy, G.J.M., Reichle, R.H., Arsenault, K.R., Houser, P.R., Kumar, S., Verhoest, N.E.C., Pauwels, V.R.N.: Multiscale
605   assimilation of Advanced Microwave Scanning Radiometer–EOS snow water equivalent and Moderate Resolution
606   Imaging Spectroradiometer snow cover fraction observations in northern Colorado. Water Resources Research 48.
607   https://doi.org/10.1029/2011WR010588, 2012.
608 Deems, J.S., Fassnacht, S.R., Elder, K.J.: Interannual Consistency in Fractal Snow Depth Patterns at Two Colorado Mountain
609   Sites. J. Hydrometeor 9, 977–988. https://doi.org/10.1175/2008JHM901.1, 2008.
610 Derksen, C., Lemmetyinen, J., Toose, P., Silis, A., Pulliainen, J., Sturm, M.: Physical properties of arctic versus subarctic
611   snow: Implications for high latitude passive microwave snow water equivalent retrievals. Journal of Geophysical
612   Research: Atmospheres 119, 7254 – 7270. https://doi.org/10.1002/2013JD021264, 2014.
613 Durand, M., Johnson, J.T., Dechow, J., Tsang, L., Borah, F., Kim, E.J.: Retrieval of SWE from dual-frequency radar
614   measurements: Using time series to overcome the need for accurate a priori information. EGUsphere 1 – 23.
615   https://doi.org/10.5194/egusphere-2023-1653, 2023.
616 Ek, M.B., Mitchell, K.E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G., Tarpley, J.D.: Implementation of Noah
617   land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model.
618   Journal of Geophysical Research: Atmospheres 108. https://doi.org/10.1029/2002JD003296, 2003.

Errico, R.M., Yang, R., Masutani, M., Woollen, J.S.: The estimation of analysis error characteristics using an observation systems simulation experiment. Meteorologische Zeitschrift 16, 695–708, 2007.

Fang, Y., Liu, Y., Margulis, S.A.: A western United States snow reanalysis dataset over the Landsat era from water eyars 1985 to 2021. Sci Data 9, 677. https://doi.org/10.1038/s41597-022-01768-7, 2022.

Foster, J.L., Sun, C., Walker, J.P., Kelly, R., Chang, A., Dong, J., Powell, H.: Quantifying the uncertainty in passive microwave snow water equivalent observations. Remote Sensing of the Environment 94, 187 – 203. https://doi.org/10.1016/j.rse.2004.09.012, 2005.

Franz, K.J., Butcher, P., Ajami, N.K.: Addressing snow model uncertainty for hydrologic prediction. Advances in Water Resources 33, 820–832. https://doi.org/10.1016/j.advwatres.2010.05.004, 2010.

Garnaud, C., Bélair, S., Carrera, M.L., Derksen, C., Bilodeau, B., Abrahamowicz, M., Gauthier, N., Vionnet, V.: Quantifying Snow Mass Mission Concept Trade-Offs Using an Observing System Simulation Experiment. Journal of Hydrometeorology 20, 155–173. https://doi.org/10.1175/JHM-D-17-0241.1, 2019.

Garousi-Nejad, I., Tarboton, D.G.: A comparison of National Water Model retrospective analysis snow outputs at snow telemetry sites across the Western United States. Hydrological Processes 36, e14469. https://doi.org/10.1002/hyp.14469, 2022.

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., Silva, A.M. da, Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J.E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.D., Sienkiewicz, M., Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). Journal of Climate 30, 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

Henn, B., Newman, A.J., Livneh, B., Daly, C., Lundquist, J.D.: An assessment of differences in gridded precipitation datasets in complex terrain. Journal of Hydrology 556, 1205–1219. https://doi.org/10.1016/j.jhydrol.2017.03.008, 2018.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.-N.: The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society 146, 1999–2049. https://doi.org/10.1002/qj.3803, 2020.

Hiemstra, C.A., Liston, G.E., Reiners, W.A.: Snow Redistribution by Wind and Interactions with Vegetation at Upper Treeline in the Medicine Bow Mountains, Wyoming, U.S.A. Arctic, Antarctic, and Alpine Research 34, 262–273. https://doi.org/10.1080/15230430.2002.12003493, 2002.

Huang, H., Tsang, L., Colliander, A., Yueh, S.H.: Propagation of Waves in Randomly Distributed Cylinders Using Three-Dimensional Vector Cylindrical Wave Expansions in Foldy–Lax Equations. IEEE Journal on Multiscale and Multiphysics Computational Techniques 4, 214–226. https://doi.org/10.1109/JMMCT.2019.2948022, 2019.

Kim, R.S., Kumar, S., Vuyovich, C., Houser, P., Lundquist, J., Mudryk, L., Durand, M., Barros, A., Kim, E.J., Forman, B.A., Gutmann, E.D., Wrzesien, M.L., Garnaud, C., Sandells, M., Marshall, H.-P., Cristea, N., Pflug, J.M., Johnston, J., Cao, Y., Mocko, D., Wang, S.: Snow Ensemble Uncertainty Project (SEUP): quantification of snow water equivalent uncertainty across North America via ensemble land surface modeling. The Cryosphere 15, 771–791. https://doi.org/10.5194/tc-15-771-2021, 2021.

Koster, R.D., Mahanama, S.P.P., Livneh, B., Lettenmaier, D.P., Reichle, R.H.: Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow. Nature Geosci 3, 613–616. https://doi.org/10.1038/ngeo944, 2010.

Koster, R.D., Suarez, M.J., Ducharne, A., Stieglitz, M., Kumar, P.: A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. Journal of Geophysical Research: Atmospheres 105, 24809–24822. https://doi.org/10.1029/2000JD900327, 2000.

Kumar, S.V., Kolassa, J., Reichle, R., Crow, W., de Lannoy, G., de Rosnay, P., MacBean, N., Girotto, M., Fox, A., Quaife, T., Draper, C., Forman, B., Balsamo, G., Steele-Dunne, S., Albergel, C., Bonan, B., Calvet, J.-C., Dong, J., Liddy, H., Ruston, B.: An Agenda for Land Data Assimilation Priorities: Realizing the Promise of Terrestrial Water, Energy, and Vegetation Observations From Space. Journal of Advances in Modeling Earth Systems 14, e2022MS003259. https://doi.org/10.1029/2022MS003259, 2022.

Kumar, S.V., Peters-Lidard, C.D., Tian, Y., Houser, P.R., Geiger, J., Olden, S., Lighty, L., Eastman, J.L., Doty, B., Dirmeyer, P., Adams, J., Mitchell, K., Wood, E.F., Sheffield, J.: Land information system: An interoperable framework for high resolution land surface modeling. Environmental Modelling & Software 21, 1402–1415. https://doi.org/10.1016/j.envsoft.2005.07.004, 2006.

Kumar, S.V., Peters-Lidard, C.D., Mocko, D., Tian, Y.: Multiscale Evaluation of the Improvements in Surface Snow Simulation through Terrain Adjustments to Radiation. Journal of Hydrometeorology 14, 220–232. https://doi.org/10.1175/JHM-D-12-046.1, 2013.

Kwon, Y., Yoon, Y., Forman, B.A., Kumar, S.V., Wang, L.: Quantifying the observational requirements of a space-borne LiDAR snow mission. Journal of Hydrology 601, 126709. https://doi.org/10.1016/j.jhydrol.2021.126709, 2021.

Lahmers, T.M., Kumar, S.V., Rosen, D., Dugger, A., Gochis, D.J., Santanello, J.A., Gangodagamage, C., Dunlap, R.: Assimilation of NASA's Airborne Snow Observatory Snow Measurements for Improved Hydrological Modeling: A Case Study Enabled by the Coupled LIS/WRF-Hydro System. Water Resources Research 58, e2021WR029867. https://doi.org/10.1029/2021WR029867, 2022.

Latifovic, R., Pouliot, D., Olthof, I.: Circa 2010 Land Cover of Canada: Local optimization methodology and product development. Remote Sensing 9, 11. https://doi.org/10.3390/rs9111098, 2017.

Le Moigne, J., Dabney, P., de Weck, O., Foreman, V., Grogan, P., Holland, M., Hughes, S., Nag, S.: Tradespace analysis tool for designing constellations (TAT-C), in: 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, pp. 1181–1184, 2017.

Lehning, M., Grünewald, T., Schirmer, M.: Mountain snow distribution governed by an altitudinal gradient and terrain roughness. Geophysical Research Letters 38. https://doi.org/10.1029/2011GL048927, 2011.

Li, D., Wrzesien, M.L., Durand, M., Adam, J., Lettenmaier, D.P.: How much runoff originates as snow in the western United States, and how will that change in the future? Geophysical Research Letters 44, 6163–6172. https://doi.org/10.1002/2017GL073551, 2017.

Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J.: A simple hydrologically based model of land surface water and energy fluxes for general circulation models. Journal of Geophysical Research: Atmospheres 99, 14415–14428. https://doi.org/10.1029/94JD00483, 1994.

Lievens, H., Demuzere, M., Marshall, H.P., Reichle, R.H., Brucker, L., de Rosnay, P., Dumont, M., Girotto, M., Immerzeel, W.W., Jonas, T., Kim, E.J., Koch, I., Marty, C., Saloranta, T., Schober, J., De Lannoy, G.J.M.: Snow depth variability in the Northern Hemisphere mountains observed from space. Nature Communications, 2019.

Liston, G.E., Elder, K.: A Distributed Snow-Evolution Modeling System (SnowModel). Journal of Hydrometeorology 7, 1259–1276. https://doi.org/10.1175/JHM548.1, 2006.

Liston, G.E., Perham, C.J., Shideler, R.T., Cheuvront, A.N.: Modeling snowdrift habitat for polar bear dens. Ecological Modelling 320, 114–134. https://doi.org/10.1016/j.ecolmodel.2015.09.010, 2016.

Liu, Y., Fang, Y., Li, D., Margulis, S.A.: How well do global snow products characterize snow storage in high mountain Asia? Geophysical Research Letters 49, e2022GL100082. https://doi.org/10.1029/2022GL100082, 2022.

Livneh, B., Badger, A.M.: Drought less predictable under declining future snowpack. Nat. Clim. Chang. 10, 452–458. https://doi.org/10.1038/s41558-020-0754-8, 2020.

Lundquist, J.D., Dickerson-Lange, S., Gutmann, E., Jonas, T., Lumbrazo, C., Reynolds, D.: Snow interceptions modelling: Isolated observations have led to many land surface models lacking appropriate temperature sensitivities. Hydrological Processes 35, 7. https://doi.org/10.1002/hyp.14274, 2021.

Mahoney, P.J., Liston, G.E., LaPoint, S., Gurarie, E., Mangipane, B., Wells, A.G., Brinkman, T.J., Eitel, J.U.H., Hebblewhite, M., Nolin, A.W., Boelman, N., Prugh, L.R.: Navigating snowscapes: scale-dependent responses of mountain sheep to snowpack properties. Ecological Applications 28, 1715–1729. https://doi.org/10.1002/eap.1773, 2018.

McGrath, D., Sass, L., O'Neel, S., McNeil, C., Candela, S.G., Baker, E.H., Marshall, H.P.: Interannual snow accumulation variability on glaciers derived from repeat, spatially extensive ground-penetrating radar surveys. The Cryosphere 12, 3617–3633. https://doi.org/10.5194/tc-12-3617-2018, 2018.

Mernild, S.H., Liston, G.E., Hiemstra, C., Wilson, R.: The Andes Cordillera. Part III: glacier surface mass balance and contribution to sea level rise (1979–2014). International Journal of Climatology 37, 3154–3174. https://doi.org/10.1002/joc.4907, 2017.

Minder, J.R., Durran, D.R., Roe, G.H., Anders, A.M.: The climatology of small-scale orographic precipitation over the Olympic Mountains: Patterns and processes. Quarterly Journal of the Royal Meteorological Society 134, 817–839. https://doi.org/10.1002/qj.258, 2008.

Montomoli, F., Macelloni, G., Brogioni, M., Lemmetyinen, J., Cohen, J., Rott, H.: Observations and simulation of multifrequency SAR data over a snow-covered boreal forest. IEEE journal of selected topics in applied earth observations and remote sensing 9, 1216–1228, 2015.

NASM: National Academies of Sciences, Engineering, and Medicine: Thriving on our changing plaent: A decadal strategy for Earth observation from space. Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/24938, 2018.

Niu, G., Yang, Z., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., Xia, Y.: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. Journal of Geophysical Research: Atmospheres 116. https://doi.org/10.1029/2010JD015139, 2011.

Niu, G.-Y., Yang, Z.-L.: Effects of vegetation canopy processes on snow surface energy and mass balances. Journal of Geophysical Research: Atmospheres 109. https://doi.org/10.1029/2004JD004884, 2004.

Painter, T.H., Berisford, D.F., Boardman, J.W., Bormann, K.J., Deems, J.S., Gehrke, F., Hedrick, A., Joyce, M., Laidlaw, R., Marks, D., Mattmann, C., McGurk, B., Ramirez, P., Richardson, M., Skiles, S.M., Seidel, F.C., Winstral, A.: 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, 2016.

Pflug, J.M., Hughes, M., Lundquist, J.D.: Downscaling snow deposition using historic snow depth patterns: Diagnosing limitations from snowfall biases, winter snow losses, and interannual snow pattern repeatability. Water Resources Research e2021WR029999. https://doi.org/10.1029/2021WR029999, 2021.

Pflug, J.M., Margulis, S.A., Lundquist, J.D.: Inferring watershed-scale mean snowfall magnitude and distribution using multidecadal snow reanalysis patterns and snow pillow observations. Hydrological Processes 36, e14581, 2022.

Raleigh, M.S., Lundquist, J.D.: Comparing and combining SWE estimates from the SNOW-17 model using PRISM and SWE reconstruction. Water Resources Research 48. https://doi.org/10.1029/2011WR010542, 2012.

Reichle, R.H., McLaughlin, D.B., Entekhabi, D.: Hydrologic Data Assimilation with the Ensemble Kalman Filter. Monthly Weather Review 130, 103–114. https://doi.org/10.1175/1520-0493(2002)130<0103:HDAWTE>2.0.CO;2, 2002.

Rott, H., Yueh, S. H., Cline, D. W., Duguay, C., Essery, R., Haas, C., Hélière, F., Kern, M., Macelloni, G., Malnes, E., Thompson, A.: Cold regions hydrology high-resolution observatory for snow and cold land processes. Proceedings of the IEEE, 98(5), 752-765, 2010.

Rott, H., Duguay, C., Etchevers, P., Essery, R., Hajnsek, I., Macelloni, G., Malnes, E., and Pulliainen, J.: CoReH2O Report for mission selection: An Earth Explorer to observe snow and ice, Tech. rep., European Space Agency, https://earth.esa.int/eogateway/documents/20142/37627/CoReH2O-Report-for-Mission-Selection-An-Earth-Explorer-to-observe-snow-and-ice.pdf, 2012.

Ruiz, J.J., Lemmetyinen, J., Kontu, A., Tarvainen, R., Vehmas, R., Pulliainen, J., Praks, J.: Investigation of Environmental Effects on Coherence Loss in SAR Interferometry for Snow Water Equivalent Retrieval. IEEE Transactions on Geoscience and Remote Sensing 60, 1–15. https://doi.org/10.1109/TGRS.2022.3223760, 2022.

Schirmer, M., Wirz, V., Clifton, A., Lehning, M.: Persistence in intra-annual snow depth distribution: 1. Measurements and topographic control: PERSISTENT SNOW DEPTH DEVELOPMENT, 1. Water Resources Research 47. https://doi.org/10.1029/2010WR009426, 2011.

Singh, S., Durand, M., Kim, E., Barros, A.P: Bayesian physical-statistical retrieval of snow water equivalent and snow depth from X- and Ku-band synthetic-aperture-radar demonstration using airborne SnowSAR in SnowEx17. EGUsphere 1 – 35. https://doi.org/10.5194/egusphere-2023-1987, 2023.

Sturm, M., Wagner, A.M.: Using repeated patterns in snow distribution modeling: An Arctic example. Water Resources Research 46, 2010.

Terzago, S., Bongiovanni, G., von Hardenberg, J.: Seasonal forecasting of snow resources at Alpine sites. Hydrology and Earth System Sciences 27, 519–542. https://doi.org/10.5194/hess-27-519-2023, 2023.

767 Trujillo, E., Ramírez, J.A., Elder, K.J.: Topographic, meteorologic, and canopy controls on the scaling characteristics of the
768   spatial distribution of snow depth fields. Water Resources Research 43, 2007.
769 Tsang, L., Durand, M., Derksen, C., Barros, A.P., Kang, D.-H., Lievens, H., Marshall, H.-P., Zhu, J., Johnson, J., King, J.,
770   Lemmetyinen, J., Sandells, M., Rutter, N., Siqueira, P., Nolin, A., Osmanoglu, B., Vuyovich, C., Kim, E., Taylor, D.,
771   Merkouriadi, I., Brucker, L., Navari, M., Dumont, M., Kelly, R., Kim, R.S., Liao, T.-H., Borah, F., Xu, X.: Review
772   article: Global monitoring of snow water equivalent using high-frequency radar remote sensing. The Cryosphere 16,
773   3531–3573. https://doi.org/10.5194/tc-16-3531-2022, 2022.
774 Wayand, N.E., Hamlet, A.F., Hughes, M., Feld, S.I., Lundquist, J.D.: Intercomparison of Meteorological Forcing Data from
775   Empirical and Mesoscale Model Sources in the North Fork American River Basin in Northern Sierra Nevada,
776   California. Journal of Hydrometeorology 14, 677–699. https://doi.org/10.1175/JHM-D-12-0102.1, 2013.
777 Woodruff, C.D., Qualls, R.J.: Recurrent Snowmelt Pattern Synthesis using Principal Component Analysis of Multi-Year
778   Remotely Sensed Snow Cover. Water Resources Research 55, 6869–6885. https://doi.org/10.1029/2018WR024546,
779   2019.
780 Wrzesien, M., Kumar, S.V., Vuyovich, C., Kim, R.S., Cho, E., Pflug, J.M., Konapala, G., Arsenault, K.R.: Merging remote
781   sensing and models to improve performance and accessibility of snow information, AGU Fall Meeting, Conference
782   on Hydrology, 2022.
783 Wrzesien, M.L., Kumar, S., Vuyovich, C., Gutmann, E.D., Kim, R.S., Forman, B.A., Durand, M., Raleigh, M.S., Webb, R.,
784   Houser, P.: Development of a "Nature Run" for Observing System Simulation Experiments (OSSEs) for Snow
785   Mission Development. Journal of Hydrometeorology 23, 351–375. https://doi.org/10.1175/JHM-D-21-0071.1, 2022.
786 Ying, Y.: Assimilating Observations with Spatially Correlated Errors Using a Serial Ensemble Filter with a Multiscale
787   Approach. Monthly Weather Review 148, 3397–3412. https://doi.org/10.1175/MWR-D-19-0387.1, 2020.
788 Yu, L., Fennel, K., Wang, B., Laurent, A., Thompson, K.R., Shay, L.K.: Evaluation of nonidentical versus identical twin
789   approaches for observation impact assessments: an ensemble-Kalman-filter-based ocean assimilation application for
790   the Gulf of Mexico. Ocean Science 15, 1801–1814. https://doi.org/10.5194/os-15-1801-2019, 2019.
791 Yueh, S. H., Dinardo, S. J., Akgiray, A., West, R., Cline, D. W., Elder, K.:. Airborne Ku-band polarimetric radar remote
792   sensing of terrestrial snow cover. IEEE Transactions on Geoscience and Remote Sensing, 47(10), 3347-3364. 2009.
793 Zhu, J., Tan, S., King, J., Derksen, C., Lemmetyinen, J., and Tsang, L.: Forward and inverse radar modeling of terrestrial snow
794   using SnowSAR Data, IEEE Transactions on Geoscience and Remote Sensing, 56(12), 7122-7132,
795   doi:10.1109/TGRS.2018.2848642, 2018.
796 Zhu, J., Tan, S., Tsang, L., Kang, D.K., and Kim, E.: Snow water equivalent retrieval using active and passive microwave
797   observations. Water Resources Research 57, no. 7, 2021.