Advancing snow data assimilation with a dynamic observation uncertainty

Devon Dunmire[1], Michel Bechtold[1], Lucas Boeykens[1, 2], and Gabriëlle J. M. De Lannoy[1]

[1]Department of Earth and Environmental Sciences, KU Leuven, Leuven, Belgium
[2]Department of Environment, Ghent University, Ghent, Belgium

Correspondence: Devon Dunmire (devon.dunmire@kuleuven.be)

Abstract. Seasonal snow is a critical resource for society by providing water for billions, supporting agriculture, clean energy, and tourism, and is an important element within the climate system by influencing the global energy balance. However, accurately quantifying snow mass, particularly in mountainous regions, remains a challenge due to substantial observational and modeling limitations. As such, data assimilation (DA) offers a powerful solution by integrating observations with physically-based models to improve estimates of the snowpack. Previous snow DA studies have employed an Ensemble Kalman Filter (EnKF) to assimilate Sentinel-1 satellite-based snow depth retrievals, demonstrating improved accuracy in modeled snow depth, mass, and streamflow when evaluated against in-situ measurements. In those studies, the uncertainty of the assimilated retrievals was assumed to be static in time and space, likely leading to a suboptimal use of the observational information. Here, we present several advances in snow DA. Using an EnKF, we assimilate novel snow depth retrievals derived from a machine learning product that leverages Sentinel-1 backscatter observations, land cover, and topographic information over the European Alps. We also incorporate a spatiotemporally dynamic observation error, whereby the uncertainty of the assimilated snow depth retrieval varies in space and time with snow depth. The machine learning snow depth retrieval product is assimilated into the Noah-MP land surface model over the entire European Alps at 1 km resolution for the years 2015-2023 and snow depth, snow water equivalent, and snow cover are evaluated against independent in-situ data and satellite observations. This work demonstrates the benefits of machine learning based snow depth retrievals and dynamic observation errors in EnKF-based snow DA.

1 Introduction

Snow is a valuable natural resource, integral for societal needs and in the climate system. The runoff from seasonal snow serves as a water source for billions of people (Barnett et al., 2005; Mankin et al., 2015), supports clean hydroelectric energy generation (Wasti et al., 2022), and sustains irrigated agriculture (Qin et al., 2020). Snow is also necessary for the multi-billion dollar winter tourism industry (Outdoor Industry Association, 2017; Parthum and Christensen, 2022; Steiger et al., 2019). The total economic value of snow is estimated to be in the trillions of dollars (Sturm et al., 2017).

Snow has a high albedo and therefore plays an important role within the climate system by exerting a large-scale cooling effect. Variability in snow cover therefore impacts the Earth's surface energy balance and has been shown to potentially

25 affect Northern Hemisphere atmospheric circulation (Henderson et al., 2018). Significant changes including a decline in snow-covered area, particularly at low elevations (Bormann et al., 2018; Estilow et al., 2015), shifts in the timing of snow melt (Musselman et al., 2021; Vorkauf et al., 2021), and an increasing transition from snowfall to rainfall at lower elevations (Safeeq et al., 2016) have been observed in recent decades, with these changes projected to intensify throughout the 21st century (IPCC, 2021).

30 Despite the importance of snow within Earth's climate and as a natural resource, accurately quantifying snow mass (or snow water equivalent, SWE) in mountainous, complex terrain remains a challenge. In-situ observation stations provide point-based measurements with good temporal frequency, but fail to capture spatial snow variability, which can be great even in a small area (López-Moreno et al., 2015; Miller et al., 2022). Airborne surveys provide accurate snow depth maps at a fine spatial resolution (Deems et al., 2013), but their high costs and logistical constraints limit the frequency and spatial coverage of these

35 measurements. Snow depth has also been retrieved using satellite observations, which have the benefit of providing frequent, global coverage (Lievens et al., 2019). Passive microwave and synthetic aperture radar (SAR) observations are currently widely used to estimate snow depth (Kelly et al., 2019; Luojus et al., 2021; Lievens et al., 2022); however, these satellite-based estimates also have substantial limitations and uncertainties. For example, passive microwave imagery has a coarse spatial resolution (∼25 km) and saturates above 1 m snow depth (Tedesco and Narvekar, 2010; Vander Jagt et al., 2013), while

40 SAR observations are challenged by wet snow, shallow snow, and forest cover (Broxton et al., 2024; Hoppinen et al., 2024; Lievens et al., 2022). Although recent work has utilized machine learning (ML) techniques to enhance SAR-based snow depth retrievals (Daudt et al., 2023; Broxton et al., 2024; Dunmire et al., 2024), there is still some way to go for accurate global SWE estimation.

Ultimately, complex feedbacks between changes in snow and other components of the global climate system are currently

45 best studied using physics-based models (Girotto et al., 2020). Since in-situ SWE observations are far sparser than snow depth measurements (Dunmire et al., 2024), snow mass estimates also rely primarily on modeling approaches. However, these models are limited by uncertainties in mountain precipitation and low-quality forcing data (Günther et al., 2019; Raleigh et al., 2016; Terzago et al., 2020). In light of these observational and modeling challenges, data assimilation (DA) offers a way to overcome shortcomings of both the model and observations by integrating in-situ and remote satellite observations with physics-based

50 models to improve modeled snow variables (Helmert et al., 2018; Smyth et al., 2020, 2022).

The Ensemble Kalman Filter (EnKF; Reichle et al. (2002)) is a commonly utilized scheme for assimilating snow observations into land surface models via state updating. With an EnKF, the background-error covariance in the Kalman Filter is not explicitly computed, but instead estimated using an ensemble of model trajectories. While this ensemble approach is advantageous for high-dimensional, nonlinear systems where an exact computation of the background-error covariance is impractical,

55 the assumption of unbiased, normally distributed model-state errors is often violated for cumulative state variables like snow depth. Despite its reliance on Gaussian assumptions, the EnKF has been extensively used in previous snow data assimilation work (Slater and Clark, 2006; Durand and Margulis, 2006; De Lannoy et al., 2012; Huang et al., 2017; Pflug et al., 2024). In contrast, particle batch filters and smoothers provide as alternative capable of handling non-Gaussian noise and complex posterior distributions, though they can be more computationally expensive given the large number of particles required.

60   Recent studies have used both particle batch smoothers and the EnKF to assimilate SAR-based snow depth retrievals from Sentinel-1, thereby improving modeled snow depth, SWE and streamflow compared to in-situ measurements (De Lannoy et al., 2024; Brangers et al., 2024; Girotto et al., 2024). However, these previous snow DA studies make the simplifying assumption that the observation uncertainty is constant in space and time, meaning that a 10 cm snowpack is assumed to have the same absolute uncertainty as a 400 cm snowpack, contributing to a suboptimal use of the observational information.

65   Here, we present several advances in snow DA. First, we assimilate snow depth retrievals from an ML product that uses Sentinel-1 (S1) observations, land cover, and topographic information to estimate snow depth in the European Alps (Dunmire et al., 2024). These ML-based snow depth retrievals have a higher accuracy and lower bias compared to previous S1-based retrievals from a conceptual model (Lievens et al., 2022), when validated against in-situ observations and airborne snow depth maps. We assimilate these ML-based snow depth retrievals within a land surface model over the entire European Alps, a domain

70 much larger than most previous snow DA efforts which focus primarily on smaller, regional scales. Finally, we incorporate a dynamic observation error, whereby the uncertainty of the assimilated snow depth observation varies in space and time, reflecting the more realistic dynamics of uncertainty in snowpack observations. The primary goal of this work is to assess the utility of incorporating dynamic observation errors in EnKF-based snow DA.

2   Materials and methodology

75 In this work, we utilized the NASA Land Information System (LIS; Kumar et al. (2006); Peters-Lidard et al. (2007)) version 7.5.0 to assimilate snow depth retrievals in the Noah-MP land surface model (Niu et al., 2011; Yang et al., 2011) version 4.0.1. The snow depth retrievals, land surface model, DA experiments, and evaluation data and methods are further described below.

2.1   Noah-MP land surface model

To simulate snow processes over the European Alps (3.9945°E–17.0175°E, 42.9945°N–48.6195°N), we ran Noah-MP on a

80 regular latitude-longitude grid with a spatial resolution of 0.009°. The model was forced with atmospheric forcing from the ECMWF Reanalysis, version 5 (ERA5; Hersbach et al. (2020)). The ERA5 data were downscaled from their native resolution (31 km) to the domain grid through bilinear spatial interpolation and by applying a topographic lapse-rate correction. ERA5 has previously been used as atmospheric forcing in other snow DA studies (Pflug et al., 2024; De Lannoy et al., 2024), and Brangers et al. (2024) additionally demonstrated that ERA5 forcing leads to superior modeled snow depth, compared with

85 simulations forced with The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2; Gelaro et al. (2017)), and MERRA-2 gauge-corrected precipitation (M2CORR; Reichle et al. (2017)).

    In Noah-MP, snow is simulated in up to 3 layers, depending on the total snow depth. Snow processes and properties such as melt metamorphism, canopy interception, and snow cover fraction are represented by detailed physically-based parameterizations (Niu et al., 2011). For snow albedo, we used the Canadian Land Surface Scheme (CLASS; Verseghy (1991)). For other

90 parameterization options, we followed Brangers et al. (2024).

Before beginning our DA experiments, we performed a 15-year model spin-up (2000-2015). The experiments were conducted over the period spanning October 1, 2015 – April 30, 2023 (8 snow seasons). Noah-MP was run with a 15 minute model time step and daily averages of state variables were written to output.

2.2 Machine learning snow depth retrieval

Previous work has assimilated snow depths retrieved from the S1 satellite constellation (SD_{S1}; Brangers et al. (2024); De Lannoy et al. (2024)). Here, we assimilated snow depth estimates from Dunmire et al. (2024) (SD_{ML}), which uses machine learning to enhance S1-based snow depth retrievals. Dunmire et al. (2024) use an eXtreme Gradient Boosting (XGBoost) model that incorporates 12 input features (elevation, slope, aspect angle, topographical position index, snow class, forest cover fraction, day of snow season, snow cover fraction, cumulative snow cover fraction, local incidence angle of the S1 observation, S1 VV backscatter, and S1 cross-polarization ratio) to estimate snow depth across the European Alps at 100 m resolution. When compared to in-situ snow depth stations and airborne photogrammetry snow depth maps, SD_{ML} is shown to be more accurate than SD_{S1} (Dunmire et al., 2024; Lievens et al., 2022).

We spatially averaged the SD_{ML} retrievals to the 0.009° model resolution and masked pixels with a glacier fraction above 50%, according to version 7 of the Randolph Glacier Inventory (Pfeffer et al., 2014; RGI 7.0 Consortium, 2023). We also temporally averaged the SD_{ML} retrievals every 7 days and assimilated these estimates weekly, in the center of the 7-day averaging window. This step was taken to avoid assimilating outlier snow depths (the SD_{ML} retrievals can be noisy in time) and to avoid negative consequences (e.g. spurious temporal trends) associated with a changing assimilation frequency (Dee, 2005).

2.3 Data assimilation approach and experiments

We conducted 3 different experiments: (1) an open loop, model-only experiment (OL) which serves as a benchmark to evaluate the added value of assimilating SD_{ML} retrievals, (2) a DA experiment with an assumed constant observation error (DA$_{\text{const}}$), and (3) a DA experiment with a dynamic observation error that varies spatially and temporally (DA$_{\text{var}}$). For all experiments, we utilized 12 ensemble members, created by perturbing forcing variables (precipitation, 2 m air temperature, and incident longwave and shortwave radiation) and the total forecasted snow depth (with the total snow depth perturbations distributed over the snow layers). Although a larger ensemble size is more optimal, our choice of 12 ensembles is reasonable as the control vector used in the assimilation consists of just total snow depth (Pflug et al., 2024). The perturbation parameters are summarized in Table 1 and follow Modanesi et al. (2022), Bechtold et al. (2023), and Pflug et al. (2024).

For the DA experiments, we used a one-dimensional EnKF to assimilate the SD_{ML} retrievals into Noah-MP. The Kalman gain matrix determines the strength of the model corrections at each location (x) and timestep (t), and is given by Equation 1 below:

$$K(x,t) = \frac{\sigma_f^2(x,t)}{\sigma_f^2(x,t) + \sigma_{obs}^2} \tag{1}$$

Table 1. Perturbation parameters applied for the OL and DA runs. * We perturb the total snow depth and propagate these perturbations into the different snow layers.

Variable	Perturbation type	Standard deviation	Cross-correlation			
Forcing variables			SW	LW	P	T
SW: Incident shortwave ($\mathrm{W\,m^{-2}}$)	multiplicative	0.6	1	-0.5	-0.5	0.3
LW: Incident longwave ($\mathrm{W\,m^{-2}}$)	additive	50.0	0.5	1	0.5	0.6
P: Precipitation ($\mathrm{kg\,m^{-2}\,s^{-2}}$)	multiplicative	0.5	-0.5	0.5	1	-0.1
T: 2 m air temperature (K)	additive	1.0	0.3	0.6	-0.1	1
Forecast variable*						
Snow depth (m)	multiplicative	0.0005				

where σ_f is the standard deviation of the forecast error and represents the uncertainty in the forecast's total snow, and σ_{obs} is the standard deviation of the observation error and represents the uncertainty in the observations. The EnKF extends a traditional Kalman Filter by estimating σ_f using forecast ensembles, while σ_{obs} is a user-defined parameter. Here, we tested
125 two different approaches for σ_{obs}, one that is constant ($\mathrm{DA_{const}}$) and one that varies in space and time ($\mathrm{DA_{var}}$).

As per De Lannoy et al. (2024), the $\mathrm{DA_{const}}$ experiment assumes a constant value of $\sigma^{obs} = 0.3$ m. The multiplicative factor for the snow depth state perturbations (Table 1) was determined experimentally through trial and error, with the optimal value selected based on its performance compared to in-situ snow depth observations over a subset region (Brangers et al. (2024), personal communication, Isis Brangers).

130 The $\mathrm{DA_{var}}$ experiment expands upon $\mathrm{DA_{const}}$ by varying σ_{obs} throughout space (x) and time (t) following Equation 2 below:

$$\sigma^{obs}(x,t) = \begin{cases} 0.05, & SD_{ML}(x,t) \leq 0.167, \\ m * SD_{ML}(x,t), & 0.167 < SD_{ML}(x,t) < 3.5, \\ 1.05, & SD_{ML}(x,t) \geq 3.5. \end{cases} \qquad (2)$$

where $SD_{ML}(x,t)$ is the assimilated observation at location x and time t and m is a user-defined multiplier. We defined m experimentally by selecting the optimal value when comparing modeled snow depth with in-situ observations in a subset
135 region (6-8°E, 45-46°N). Here, we used $m = 0.3$. To avoid issues when $SD_{ML}(i,t) = 0$ m, we defined a minimum value of 0.05 m for σ_{obs}. We also defined a maximum value for σ_{obs} of 1.05 m. Figure 1 compares σ_{obs} from $\mathrm{DA_{const}}$ and $\mathrm{DA_{var}}$ as a function of the assimilated snow depth observation (SD_{ML}).

For both DA experiments, the snow updates were applied following the methodology of Brangers et al. (2024), whereby the increments applied to the total forecasted snow depth and SWE are divided over the different snow layers, proportionate to each
140 layer's forecasted share of the total snowpack, and SWE is updated accordingly assuming snow density remains unchanged
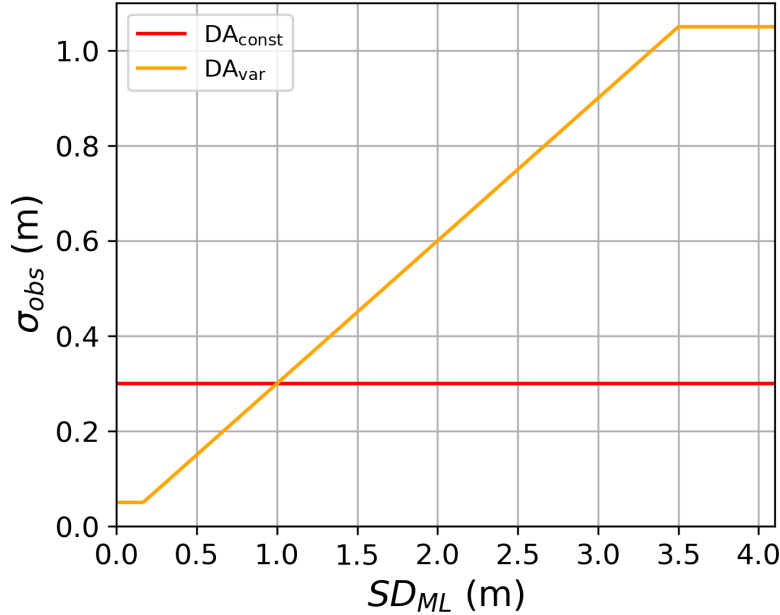
Figure 1. Assumed observation error standard deviation (σ_{obs}) as a function of the assimilated snow depth (SD_{ML}) for the two DA experiments.

during each update. We assimilated SD_{ML} estimates weekly each year from September 1 through March 31, except over forested areas and glaciers, or when the soil or vegetation temperature is above 5°C (De Lannoy et al., 2024).

2.4 Evaluation

To evaluate the snow depth estimates from each experiment (OL, DA$_{\text{const}}$, DA$_{\text{var}}$), we first compared with in-situ snow depth
145   observations from a variety of snow monitoring networks. For comparing the performance of our DA experiments against the
OL experiment, we utilized independent in-situ observations that were not included in the training for the ML model from
Dunmire et al. (2024), and sites located in places where the SD_{ML} retrievals were assimilated (i.e. not in dense forest, over
glaciers). In total, we utilized snow depth data from 588 measurement sites, which report for varying parts of the 8 year study
period. We obtained these point-scale snow depth measurements from the WSL – Institute for Snow and Avalanche Research
150   SLF (Switzerland, 220 sites), Météo-France (France, 57 sites), GeoSphere Austria (Austria, 108 sites), the International Center
for Environmental Monitoring CIMA Research Foundation (Italy, 10 sites), Provincia autonoma di Trento (Italy, 48 sites),
Provincia autonoma di Bolzano - Alto Adige (Italy, 19 sites), Valle d'Aosta (Italy, 27 sites), the Agenzia Regionale per la
Protezione Ambientale - Piemonte (Italy, 28 sites), the European Centre for Medium-Range Weather Forecasts' SYNOP snow
depth measurement network (Global, 35 sites; de Rosnay et al. (2015)), and Global Historical Climatology Network (Global,
155   36 sites). For each experiment, we computed the mean absolute error (MAE), bias, and Pearson correlation coefficient (R)
of the modeled snow depth compared with the in-situ observations obtained at these sites. To investigate how well the model

captures spatial and temporal anomalies in snow depth patterns, we also computed spatial and temporal anomaly correlation coefficients (ACC). The spatial ACC was computed for each day throughout the snow season with more than 10 in-situ snow depth measurements available. Spatial anomalies were computed for each site by subtracting the spatial mean snow depth

160 recorded across all measurement sites on that day. The temporal ACC was computed for each measurement site with 5 or more years of in-situ observations. Temporal anomalies were calculated at each site by subtracting the site's multi-year climatology (2015-2023) with a 10-day moving mean smoothing function applied. In order to utilize more sites with a longer time series of observations, we also included sites that were used in the ML training. Thus, for this metric, we only compared the two DA experiments, which both assimilated the same SD_{ML} retrievals.

165 Next, we evaluated modeled SWE, with in-situ measurements of SWE located (1) in places where DA was applied, and (2) not on a glacier, according to the Noah-MP glacier land cover class and the Randolph Glacier Inventory (Pfeffer et al., 2014; RGI 7.0 Consortium, 2023). We consolidated 8211 manual SWE measurements from the Bundesministerium für Land- und Forstwirtschaft, Regionen und Wasserwirtschaft (Austria, 676 measurements), The Climate Data Center of the German Weather Service (Germany, 2311 measurements), the WSL – Institute for Snow and Avalanche Research SLF (Switzerland,

170 1546 measurements), Provincia autonoma di Trento (Italy, 944 measurements), and Valle d'Aosta (Italy, 2793 measurements). As with snow depth, we compared MAE, bias, and R for the different experiments.

Finally, we evaluated the impact of the DA on the timing of snow disappearance and modeled snow cover fraction (SCF). We first compared the snow disappearance date (SDD) of the model experiments at the in-situ snow measurement sites. We computed the SDD as the first day of five consecutive days with less than 0.1 mm snow depth, following the date of peak snow.

175 For in-situ SDD, the day of peak snow was computed using the in-situ snow depth and for model SDD, the day of peak snow was computed using snow depth output from the appropriate model experiment. We also, in the same manner, computed SDD using the Interactive Multisensor Snow and Ice Mapping System (IMS) product. IMS is a 1 km horizontal resolution binary snow cover dataset that is derived from a variety of satellite and in-situ data.

We also compared SCF and total snow covered area from our three model experiments with both the IMS product and the

180 Copernicus Fractional Snow Cover product. The Copernicus product is available at a 20 m spatial resolution and is computed from Sentinel-2 Level1-C imagery. The product is not gap-filled, thus data gaps exist when clouds are present. We regridded both snow cover products to our model domain grid using nearest neighbor interpolation for IMS, and averaging for the Copernicus product. For comparison with the IMS product, we converted modeled SCF to a binary value: $SCF < 50\% = 0$, $SCF \geq 50\% = 1$. For comparison with the Copernicus product, we ignored areas with data gaps.

185 ## 3 Results

3.1 Snow depth evaluation

The practical impact of the DA_{var} and DA_{const} experiments on snow depth estimates is illustrated in Figure 2. When the assimilated snow depth retrieval is 1 m, the observation uncertainty is equivalent for both experiments (Fig. 1). The variable observation uncertainty approach in DA_{var} dynamically adapts to assimilated snow depth, resulting in stronger corrections for
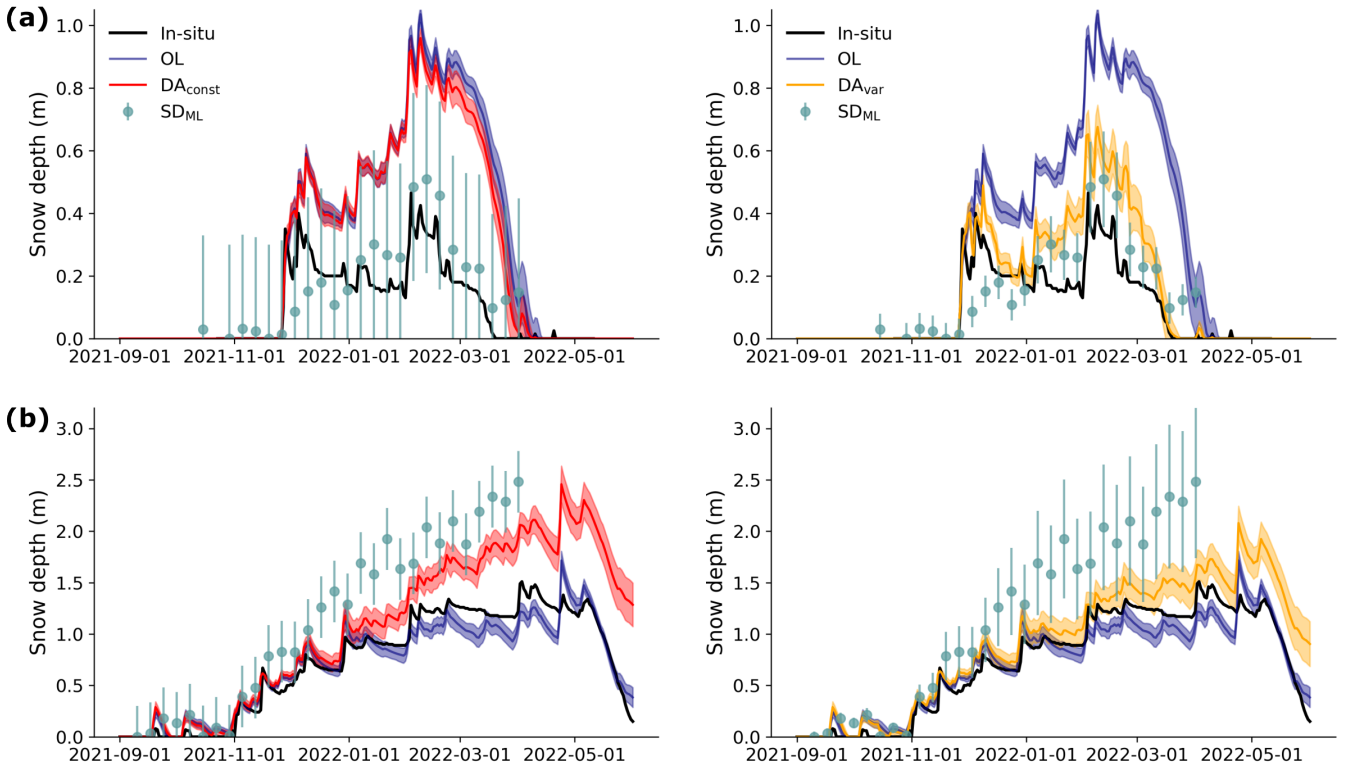
Figure 2. Snow depth estimates and in-situ measurements at two example sites. (a) Snow depth from DA_{const} (red, left) and DA_{var} (orange, right) compared with the OL (navy) from a measurement station in Austria ($\sim$13.6228 $^\circ E$, 47.0944 $^\circ N$, 1050 m elevation). The shading represents $\pm$1 standard deviation in the model ensembles. The sage green dots represent the assimilated SD_{ML} retrievals, with error bars for the assumed observation error standard deviation (σ_{obs}). (b) Same as (a), but for a different measurement station in Switzerland ($\sim$7.7836 $^\circ E$, 45.9872 $^\circ N$, 2948 m elevation)

shallow depths while DA_{const} provides stronger corrections at higher depths (Fig. 2). For assimilated snow depths below 1 m, the observation uncertainty is smaller in DA_{var} than in DA_{const}, resulting in a lower observation error covariance (σ_{obs}) in the EnKF (Equation 1) and stronger corrections of the posterior state toward the observations in DA_{var} (Fig. 2a). In contrast, the constant σ_{obs} of 0.3 m in DA_{const} is relatively large for shallow snow depths and results in minimal corrections of the posterior state. A measurement site with assimilated snow depths substantially greater than 1 m is demonstrated in Figure 2b. In this case, the observation uncertainty is smaller for DA_{const} than for DA_{var}, resulting in stronger posterior state adjustments in DA_{const}.

Across the 588 in-situ snow depth measurement sites used for evaluation, the corrections applied in DA_{var} result in snow depth estimates that align more closely with in-situ observations (Fig. 3). The OL experiment yields a site-average MAE of 0.244 m, a bias of 0.113 m and a Pearson correlation coefficient of 0.75. Both the DA_{const} and DA_{var} experiments show improved performance, with site-average MAE values of 0.237 m and 0.215 m, and biases of 0.106 m and 0.055 m, respectively.

These improvements are illustrated in Figure 3, which compares MAE from the DA_{var} experiment with the OL experiment (Fig. 3a) and DA_{const} (Fig. 3b). Relative to the OL, MAE is reduced in DA_{var} by more than 25 mm at 245 sites (42%), while 92 sites (16%) have an MAE increase exceeding 25 mm. Comparing DA_{var} to DA_{const}, we find that MAE is reduced in DA_{var} by more than 15 mm at 297 sites (51%), while 71 sites (12%) experience a deterioration greater than 15 mm. While improvement in MAE from the OL experiment is not significant for DA_{const}, the MAE improvement is significant for DA_{var} (p=0.001). The site-average Pearson correlation coefficient slightly deteriorated for DA_{const} and improved for DA_{var} to 0.75 and 0.76, respectively.

Figure 3c highlights that, for most of the snow season, the DA_{var} experiment provides the best representation of spatial snow depth patterns. Averaged across the entire year, the spatial ACC is 0.71 for the OL experiment, 0.72 for DA_{const}, and 0.73 for DA_{var}. The greatest improvement in spatial ACC for DA_{var} occurs during the early snow season (November), with values exceeding those of the OL and DA_{const} experiments by 0.058 and 0.047, respectively. From December through April, the spatial ACC for DA_{var} remains approximately 0.021 greater than that of the OL experiment. By mid-April, all three experiments exhibit similar performance in capturing spatial snow depth patterns. Additionally, DA_{var} significantly outperforms DA_{const} (p < 0.001) in capturing temporal snow depth patterns (Fig. 3d). Across the 948 sites evaluated, 491 sites (52%) have an improved temporal ACC in DA_{var} (>+0.02 compared to DA_{const}), while only 103 sites (11%) experience a deterioration in temporal ACC (<-0.02 compared to DA_{const}).

The OL experiment has an elevation-dependent snow depth bias, characterized by an overestimation of snow depth at lower elevations and early in the snow season, and an underestimation at higher elevations during peak snow accumulation (Fig. 4a). Both of these issues are mitigated in the DA_{var} experiment, which brings seasonal biases closer to zero across all elevation bands (Fig. 4c). In contrast, the DA_{const} experiment minimally corrects snow depth overestimation in the early season and at low elevations, due to the relatively higher assumed observation uncertainty for shallow snow (e.g., Fig. 2a). From September 1 through January 31, the DA_{var} experiment reduces the average bias across all sites by 46%, while the DA_{const} experiment achieves only a 10% reduction over the same period. These improvements are particularly notable at mid-elevations (1000–2000 m), where DA_{var} reduces model bias by 54% throughout the season, compared to a 13% reduction in model bias at these same sites in DA_{const}.

The MAE is also reduced by DA_{var} across most elevation bands and throughout much of the season. The difference in MAE between the OL and DA_{var} experiments (Fig. 4d) indicate that the largest MAE improvements occur from early winter through peak accumulation. However, an increase in MAE at high elevations during the melt season (March onwards) suggests a tendency for the DA experiments to retain snow for too long, which could be due to limitations in the modeled melt processes or biases introduced by the assimilated observations at higher elevations.

3.2 SWE evaluation

Compared with 8,211 manual SWE measurements from 231 different measurement sites across the Alps, the DA_{var} experiment also significantly improves the MAE of SWE estimates compared to both the OL and DA_{const} experiments (p<<0.001). Relative to the OL, DA_{var} reduces SWE MAE by at least 15 mm at a majority of these sites (57%), while only 23% of sites
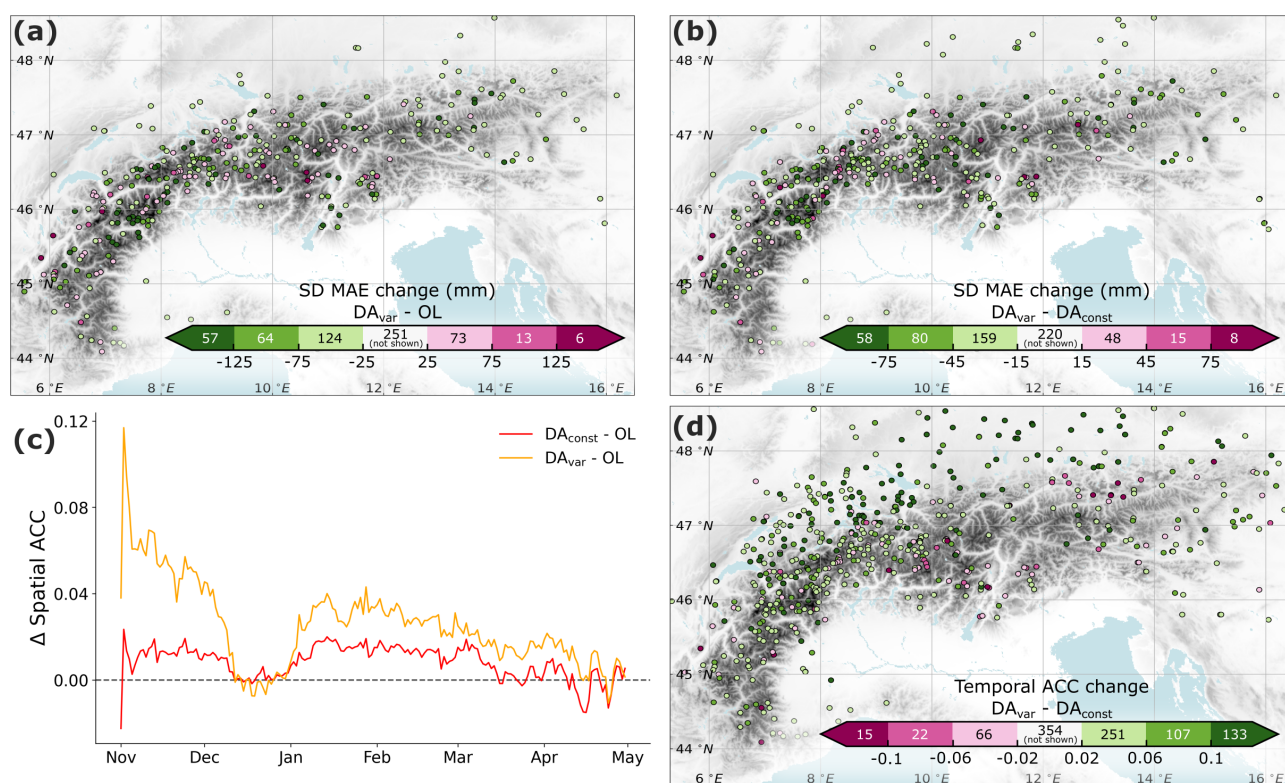
Figure 3. Experiment evaluation at in-situ snow depth measurement sites. (a) Change in MAE at each measurement site from the OL experiment to DA_{var}. Green colors indicate an improvement in MAE in the DA_{var} experiment. On the color bar, the number of sites that fall within each color range is indicated and points within the white color are not plotted on the map. (b) Change in MAE at each measurement site from the DA_{const} experiment to DA_{var}. (c) Change in the spatial anomaly correlation coefficient (ACC) for each DA experiment from the OL experiment. The spatial ACC is averaged over all snow seasons (2015/16 - 2022/23). (d) Change in the temporal ACC from the DA_{const} experiment to DA_{var}. Green colors indicate an improvement in temporal ACC in the DA_{var} experiment.
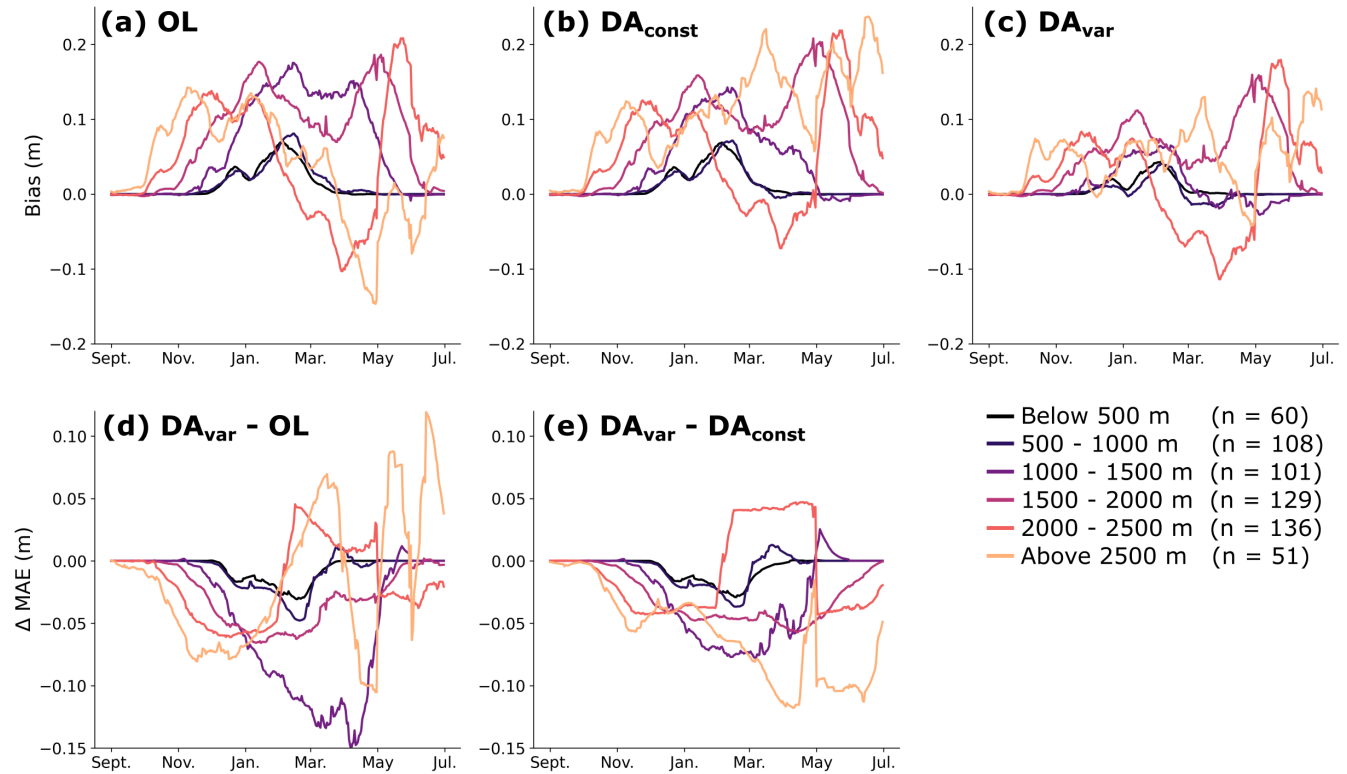
Figure 4. Seasonal evolution of bias and mean absolute error (MAE) stratified by elevation. Panels (a)-(c) snow the seasonal snow depth bias for the (a) OL, (b) DA_{const}, and (c) DA_{var} experiments. Bias is computed relative to in-situ snow depth measurements and is grouped by elevation bands (indicated by different colors). Panels (d)–(e) show the change in MAE between the OL and DA_{var} experiments (d) and between the DA_{const} and DA_{var} experiments (e). Negative values in (d)–(e) indicate improved performance (decreased MAE). Statistics are computed for each day, averaged over the entire 8 year period (2015–2023). A 14-day smoothing is applied to each timeseries and the number of in-situ measurement sites within each elevation band is provided in the legend.

235   experience a deterioration in SWE MAE of more than 15 mm (Fig. 5a). Similar improvements are observed when comparing DA_{var} to DA_{const}, with DA_{var} outperforming DA_{const} at 56% of measurement sites (Fig. 5b).

     In the OL experiment, we observe a positive bias for low observed SWE and a negative bias for high observed SWE (Fig. 5c), similar to the bias patterns seen for snow depth. The DA_{var} experiment effectively reduces both biases, with the most substantial improvement occurring for low observed SWE values. As a result, the overall average SWE bias decreases from

240   +81 mm in the OL to +18 mm in DA_{var}. This bias reduction is significantly greater than that for DA_{const}, which only marginally corrects the high bias for low observed SWE, due to minimal model adjustments for shallow assimilated snow depths (e.g., Fig. 2a). Both DA_{const} and DA_{var} also substantially improve the Pearson correlation coefficient, indicating a stronger correlation with measured SWE.
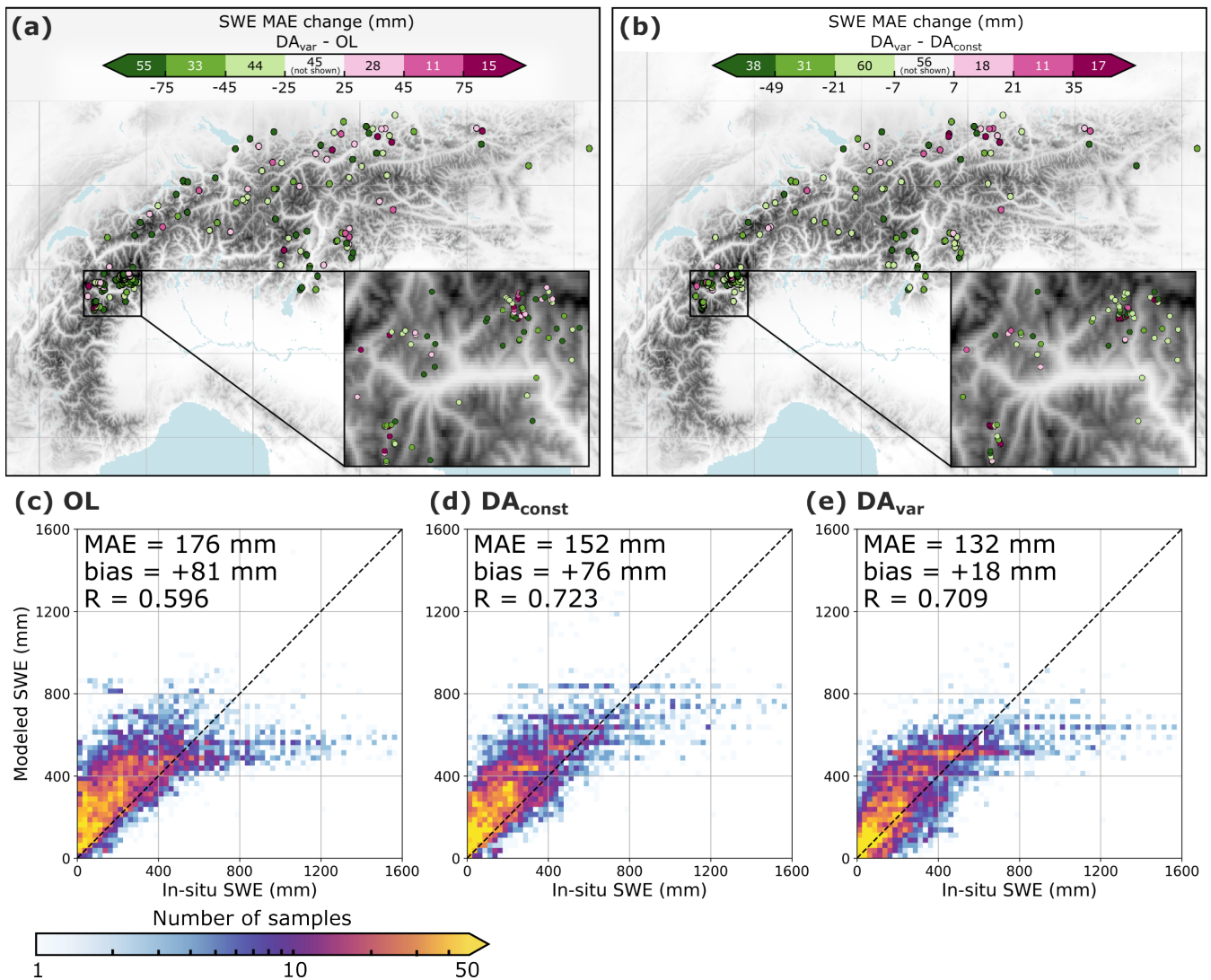
Figure 5. Evaluation of SWE from the OL and DA experiments against in-situ measurements. (a-b) Change in SWE MAE (a) DA_{var} relative to the OL experiment, and (b) DA_{var} relative to DA_{const}, where green indicates error reduction and magenta indicates a deterioration in performance. Measurements from within the same 1 km model grid cell are averaged for visualization purposes. On the color bar, the number of sites that fall within each color range is indicated and points within the white color are not shown on the map. (c-e) 2D histograms comparing modeled SWE to in-situ SWE observations for (c) OL, (d) DA_{const}, and (e) DA_{var}. All non-zero SWE measurements are included and the spatiotemporal MAE, bias, and R are provided for each approach.

Across all experiments, SWE typically peaks during the first week of March (March 1–7). The 2016/2017 snow season
recorded the lowest modeled SWE in our OL experiment, and correspondingly saw the largest SWE increases in DA_{var} prior
to early March, particularly in the Central Alps and Austrian Alps (Fig. 6a). However, DA_{var} SWE improvements were mixed
during this year. Of the 41 manual measurements taken between March 1 and March 7, 2017, only 24% demonstrated improved
SWE of more than 15 mm in DA_{var}. While the data assimilation led to more accurately estimated SWE at some sites (e.g.,
Supplemental Fig. S1b,d), it resulted in an overestimation of SWE at others (e.g., Supplemental Fig. S1c,e,f).

The most significant reductions of the positive SWE bias from the OL to the DA_{var} experiment occurred during the
2017/2018 snow season, particularly in the Bavarian Alps, Swiss Alps, and French Alps (Fig. 6b). In general, the reduced
SWE in DA_{var} aligns more closely with in-situ observations (e.g., Supplemental Fig. S2), improving SWE error by more than
15 mm in 59% of the 68 manual measurements taken between March 1 and March 7, 2018.

The 2020/2021 snow season also experienced substantial SWE reductions, especially in the Swiss Alps and Eastern Dolomites.
A lack of in-situ observations in the Dolomites region makes it difficult to assess whether these reductions are realistic; however,
limited observation sites along the Italy-Austria border suggest that the SWE reductions may be too strong (e.g., Supplemental
Fig. S3d).

3.3 Snow cover evaluation

The DA also affects snow cover estimates, contributing to a decrease in total snow-covered area leading up to peak snow
accumulation in early March, and a slight increase in snow-covered area later in the season (April-May), compared with the
OL experiment (Fig. 7a). During peak snow accumulation in early March (March 1–7), the DA_{var} experiment reduces total
snow-covered area by 6,077 km^2 compared to the OL, averaged across the 2016-2023 period. Total snow-covered area during
this same period in the DA_{const} experiment is comparatively reduced by only 1,409 km^2. The reduction in snow cover primarily
occurs in low-elevation areas along the northern Alps (Fig. 7b), and aligns more closely with observed snow cover estimates
from the IMS and Copernicus snow cover products, both of which indicate substantially less snow-covered area than any of
our model simulations. For example, on March 1, 2021, the OL and DA_{var} experiments have, respectively, 79,345 km^2 and
58,091 km^2 more snow-covered area than the Copernicus fractional snow cover product, and 55,578 km^2 and 32,526 km^2
more than the IMS snow cover product (Supplemental Fig. S4). These discrepancies will be discussed further in Section 4.

At the majority of in-situ snow depth measurement sites, the estimated snow persists for too long compared to in-situ
observations. Figure 8 presents cumulative distribution functions (CDFs), which show the cumulative number of sites with
snow-free conditions after peak snow, stratified by elevation band. In all three model experiments, the snow disappearance date
(SDD) occurs later than observed, indicating an overestimation of snow persistence across all elevation bands.

In DA_{var}, the SDD timing is improved at a majority of the observation sites located below 2000 m, with 51% of sites expe-
riencing a SDD closer to in-situ observations, 22% experiencing a SDD farther from in-situ observations, and 27% remaining
unchanged. The improvement is less pronounced for DA_{const}, in which 40% of sites show better agreement with observations,
24% show worse agreement, and 36% remain unchanged. The reduced SWE at lower elevations in DA_{var} (see Section 3.2)
likely results in more realistic timing for snow-free conditions at these sites. As we only assimilate observations through March,
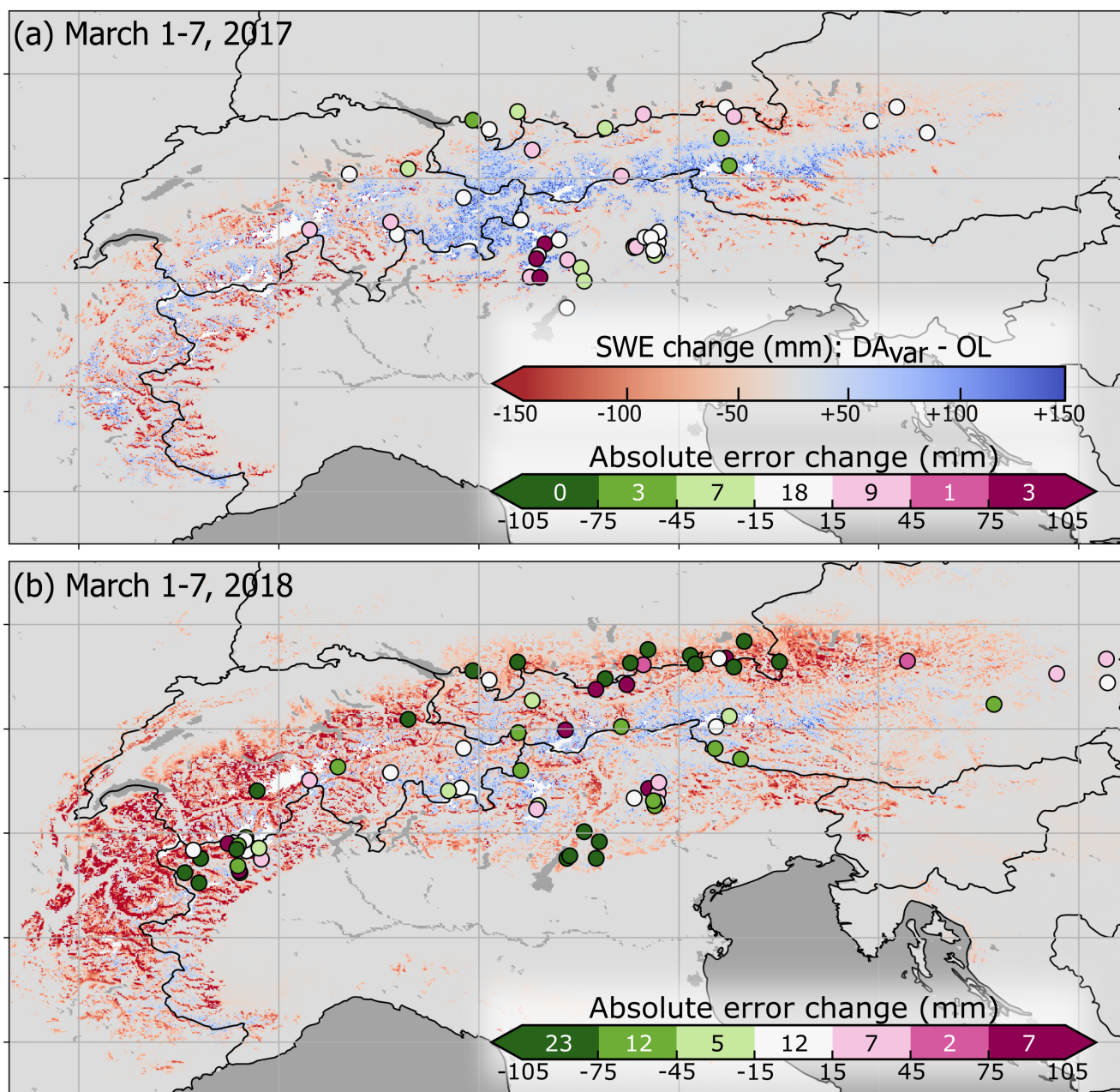
Figure 6. Change in SWE during the period March 1-7, between the DA_{var} and OL experiments for (a) 2017 and (b) 2018. Manual SWE measurements taken during this period are plotted as dots, colored according to the change in absolute error between the DA_{var} and OL experiments. On the error change color bar, the number of sites that fall within each color range is indicated.
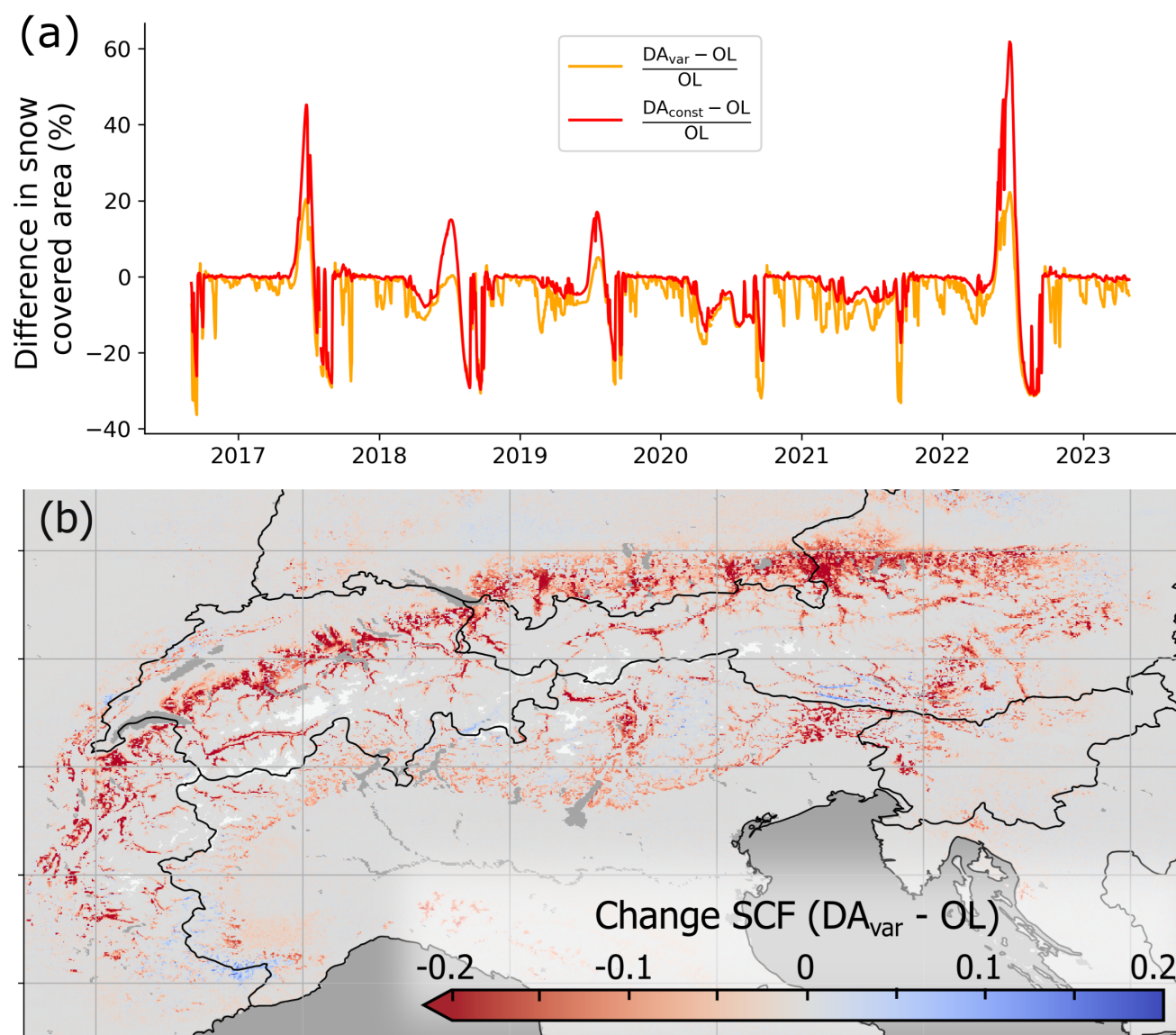
Figure 7. Difference in estimated snow covered area. (a) Timeseries of the percent difference in total snow covered area between DA_{var} and OL (orange), and DA_{const} and OL (red). (b) Average change in snow cover fraction (SCF) between the DA_{var} and OL experiments during the period March 1-7 (all years). White indicates glaciers.
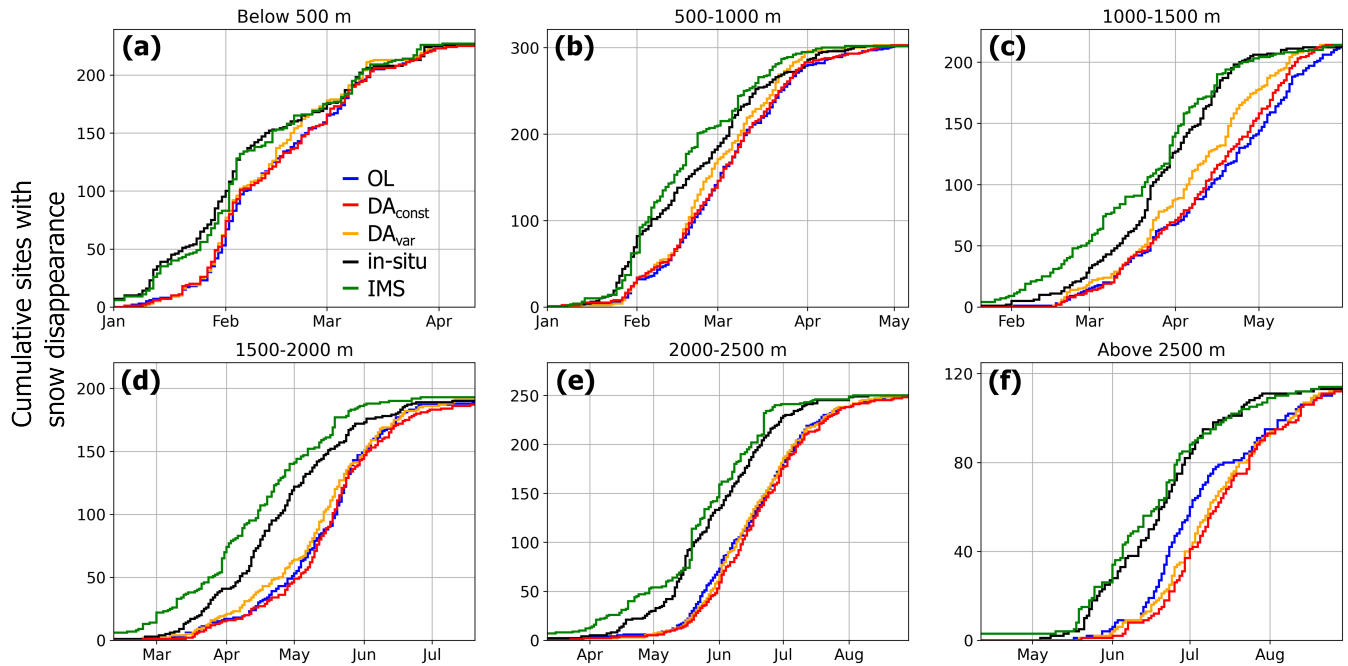
Figure 8. Cumulative number of measurement sites with snow disappearance (following peak snow depth) at in-situ measurement sites, stratified by elevation: (a) below 500 m, (b) 500-1000 m, (c) 1000-1500 m, (d) 1500-2000 m, (e) 2000-2500 m, (f) above 2500 m.

thus limiting assimilation during times of ablation, changes in SDD are mainly a result in changes of peak SWE. In general, the IMS observations underestimate snow persistence (Fig. 8), leading to an earlier SDD compared to in-situ observations, which
280 may result from the binary (as opposed to fractional) nature of the IMS observations.

3.4 DA increments and spread

In DA_{const}, model updates predominantly occur later in the accumulation season, with positive average increments above 2500 m and negative average increments below 1500 m (Fig. 9a). In contrast, DA_{var} exhibits stronger negative increments earlier in the snow season, and at lower elevations (Fig. 9b), suggesting that assimilated observations influence the entire
285 accumulation period rather than just times near peak SWE. Additionally, the magnitude of positive increments in DA_{var} is reduced, meaning that less snow is added at higher elevations in DA_{var}. While the OL has a negative snow bias in these higher elevation areas, the weaker positive increments in DA_{var} may be more realistic, given that Figure 4b indicates a strong positive snow depth bias for sites about 2500 m in DA_{const}, and a reduced positive bias at these same sites in DA_{var}.

The change in observation uncertainty also has an impact on the analysis ensemble spread, with primarily decreased ensem-
290 ble spread in DA_{var}, compared to DA_{const}, especially in lower elevation regions (Supplemental Fig. S5a). Changes in analysis spread are related to changes in the observation uncertainty, with decreases in spread corresponding to decreases in average observation uncertainty (Supplemental Fig. S5b). For example, for all model grid cells where σ_{obs} decreases, on average, from
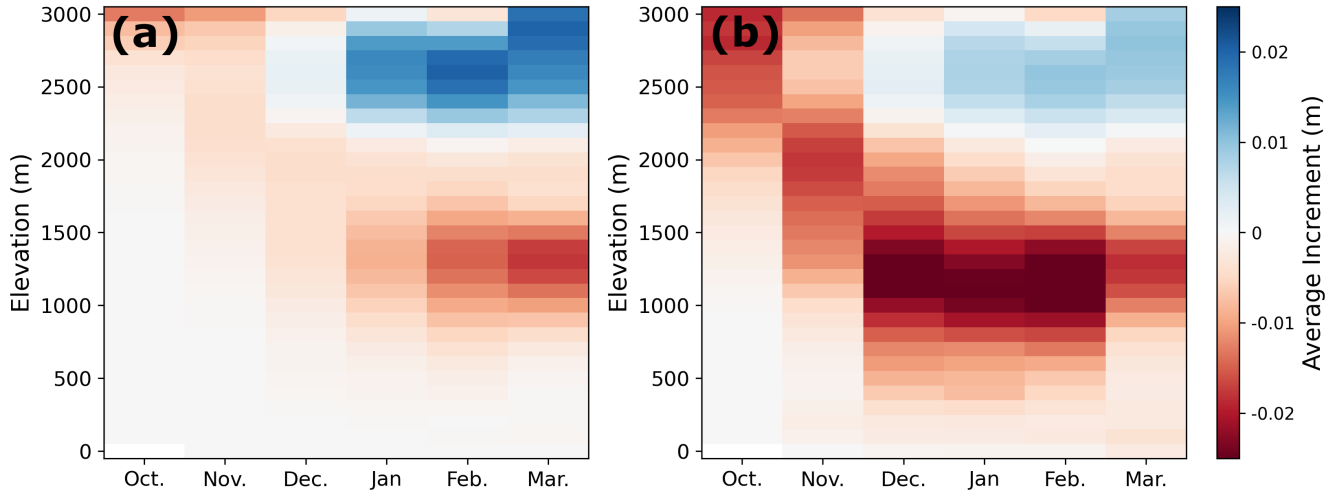
Figure 9. Average total snow depth increments (m) over the accumulation season (x-axis), stratified by elevation (y-axis) for (a) DA_const and (b) DA_var. Increments are averaged over all years (2015-2023)

DA_const to DA_var, 83% indicate a corresponding decrease in the snow depth analysis ensemble spread. In contrast, for all grid cells where σ_{obs} increases, 65% have a corresponding increase in analysis ensemble spread. The reason for this decrease in ensemble spread is likely two-fold. This overall decrease in ensemble spread is likely driven by two factors: (1) lower observation uncertainty in many regions, and (2) reduced snow depth, which results in smaller multiplicative perturbations to the forecast state.

4 Discussion

This work enhances snow DA by incorporating an ML-based snow depth retrieval product using spatio-temporally dynamic error estimates into the assimilation scheme. The ML snow depth model integrates multiple sources of information, including S1 backscatter observations, fractional snow cover from optical imagery, and land cover information to estimate snow depth. The snow depth estimated from this ML model has been shown to possess superior accuracy compared to prior S1 snow depth retrieval work by Lievens et al. (2022) (Dunmire et al., 2024), which has previously been assimilated into the Noah-MP land surface model using an Ensemble Kalman Filter (De Lannoy et al., 2024; Brangers et al., 2024). Future work could experiment with integrating additional satellite-based information into the assimilated ML product (e.g., passive microwave, X-band, ICESat-2 data).

To compare with previous work that assimilates snow depth retrievals from the S1 change detection algorithm (SD_{S1}; Lievens et al. (2022)), we compared output from our two DA experiments with DA output from De Lannoy et al. (2024) (DA_{S1}). The DA_{S1} experiment utilizes the same DA setup as DA_const, with a static observation uncertainty ($\sigma obs = 0.3\ m$), but assimilates SD_{S1} retrievals instead of SD_{ML}. We find that SWE MAE is significantly lower in DA_var compared to DA_{S1}

($p \ll 0.001$), whereas DA_{const} does not demonstrate a meaningful improvement (Supplemental Fig. S6). Using 4548 manual SWE measurements collected within the Po River basin (the study domain of De Lannoy et al. (2024)), we find an MAE of 225 mm from the OL experiment, while the MAE for DA_{S1}, DA_{const}, and DA_{var} is 193, 195, and 177 mm, respectively. The SD_{ML} retrievals are more accurate than SD_{S1} for in-situ snow depths below 2.5 m, while for snow depth exceeding
315   3 m, SD_{S1} performs better (see Figure 3a from Dunmire et al. (2024)). This suggests that assimilating SD_{ML} should provide improvements particularly for shallower snow. However, in the DA_{const} experiment, the use of a static observation uncertainty, where relatively large errors are assumed for shallow snow observations, limits these potential improvements (e.g. Fig. 2a) and results in an overall performance of DA_{const} that is similar to DA_{S1}. This analysis highlights that the treatment of the observation uncertainty is as critical as the observations themselves. A poorly parameterized observation uncertainty can restrict
320   the benefits of DA, underscoring the need for options in DA systems to dynamically vary the observation error.

The EnKF depends on accurate uncertainty estimates for both the model and observations, using these to weigh the information and obtain an optimal state. With this in mind, Dee (1995) argues that proper characterization of both model and observation uncertainties is necessary for successful implementation of the EnKF. While the specification of observation uncertainty substantially influences DA performance, in snow DA systems, this uncertainty is often prescribed as a constant value
325   (Helmert et al., 2018), and the use of dynamic observation error estimation in snow DA had not yet been explored prior to this work. Moreover, most operational land data assimilation systems (e.g., NASA Land Data Assimilation Systems, ECMWF Land Data Assimilation System) and recent studies that have applied the EnKF to assimilate SAR-based snow depth retrievals assume a static observation error. For instance, Brangers et al. (2024) assumed $\sigma obs = 0.36\ m$, and De Lannoy et al. (2024) assume $\sigma obs = 0.30\ m$ (applied here in DA_{const}).

330   Here, we highlight the implications of accounting for dynamical estimates of observation uncertainty, and demonstrate that this system results in a more realistic modeled snow state. In DA_{var}, Equation 2 ($\sigma^{obs} = m * SD_{ML}$, $m = 0.3$) is used to adapt the standard deviation of the observation error based on the assimilated snow depth. This relationship is a first-order approximation that assumes that the observation error increases linearly with the observation magnitude; however, σ_{obs} could be defined to vary in more complex ways. Future work could explore applying relationships where σ_{obs} varies non-linearly
335   with the assimilated snow depth observation, or statistical parameterizations of σ_{obs} depending on other conditions such as elevation, or forest cover. The effectiveness of a variable observation error also depends on the magnitude of the forecast error, as the Kalman gain matrix, which determines the strength of the corrections, depends on both forecast and observation error. To maximize benefits, the observation error, whether static or dynamic, should be properly tuned in relation to forecast error.

The EnKF is widely used in snow DA systems due to its efficiency; however, a key assumption is that both the observations
340   and model are unbiased. Here, we implement a bias-blind system by not bias-correcting either the observations or the model, thereby violating this assumption. Bias-aware systems which a priori correct the model bias to align with the observation climatology assume that the assimilated observations are more realistic than the model. While this assumption may be realistic in many situations, satellite-based snow retrievals also exhibit substantial bias. Since snow is a cumulative variable, biases in either the observations or the model typically persist throughout the snow season. While in-situ measurement stations can

345   help quantify these biases, they are often inconsistent spatially and on an interannual basis (i.e. Supplemental Fig. S7), which
provides a challenge for correcting them a priori.

Two major issues exist with bias-blind systems: (1) model drift towards its original state, leading to a sawtooth-like pattern
that can result in unrealistic fluxes in other variables, and (2) unrealistic model trends in DA output due to changes in assim-
ilated observation frequency (Dee, 2005). For snow, model biases primarily stem from errors in precipitation forcing data.
350   Consequently, we do not expect model drift to occur as observed in De Lannoy et al. (2007); Mocko et al. (2021); Scherrer
et al. (2023), unless there is an instantaneous precipitation forcing error. We also assimilate observations weekly throughout the
study period, thereby mitigating the potential effects of assimilation frequency in bias-blind DA. Scherrer et al. (2023) further
compare bias-blind and bias-aware assimilation of leaf area index - a cumulative variable - using the EnKF. Their results show
that the bias-blind DA more effectively updates the model state variable, and leads to larger improvements in water balance
355   components such as evapotranspiration and runoff. In contrast, while the bias-aware approach yields smaller improvements in
state variables, it improves temporal anomalies and internal DA diagnostics indicate a more optimal DA system performance.
Given our focus on improving the modeled snow state rather than snow anomalies, along with the inherent challenges of a pri-
ori bias correcting the observations and model, we opt for a bias-blind approach, recognizing that this may lead to suboptimal
DA performance (i.e. temporally correlated residuals).

360   Snow cover fraction affects the energy balance, and consequently, has implications for numerical weather prediction. While
the DA experiments generally reduce the snow-covered area by largely removing snow at lower elevation regions, all three
experiments still exhibit a substantial overestimation of total snow-covered area compared with both Copernicus and IMS snow
cover products. Several factors may contribute to this discrepancy. First, a high bias in snowfall forcing data at low elevations
will result in unrealistically large snow-covered area. Second, the higher-resolution Copernicus product (20 m) inherently
365   captures finer-scale variation between snow-covered and snow-free conditions, often resulting in lower overall snow cover
estimates compared to coarser-resolution products. Third, inaccuracies in the parameterization of snow cover fraction within
Noah-MP may also play a role. In Noah-MP, the snow cover fraction is parameterized as a function of snow depth, density, and
ground roughness length. (Niu et al., 2011; Lee et al., 2024). It should be investigated whether the current parameterizations
in Noah-MP remain appropriate for regions with complex terrain, where subgrid variation in topography can significantly
370   influence fractional snow cover. Finally, uncertainty in the Copernicus and IMS snow cover, for example due to cloud and
forest cover, contribute to errors in these validation data sets and potentially influence the perceived model biases.

5   Conclusions

In this manuscript, we demonstrate the utility of incorporating a dynamic observation uncertainty into a snow depth data
assimilation scheme. For the first time, we assimilate satellite-based snow depth estimates from a novel machine learning
375   model into the Noah-MP land surface model using the EnKF to update snow depth and SWE. We compare two data assimilation
experiments: one with a static observation error (DA_{const}), and one with an observation error that is dynamic in space and time
(DA_{var}).The performance of these DA experiments is evaluated against the open-loop experiment (OL, model-only) using in-

situ snow depth observations, manual SWE measurements, and two different snow cover products. We show that the dynamic observation error makes better use of the assimilated observations, thereby leading to stronger model corrections, particularly at times when the assimilated snow depth observation is much shallower than the model forecast (e.g., early in the accumulation period or at lower elevations). The DA experiment that incorporates this dynamic observation error more effectively corrects biases introduced by errors in the forcing data, and improves SWE estimates by 25% and 13% compared to the OL and DA_{const} experiments, respectively. While snow cover is overestimated in all three model experiments, DA_{var} also leads to stronger reductions in snow cover than DA_{const}, better aligning with existing snow cover products. As most snow DA work and operational snow DA systems assume that the observational uncertainty is constant in space and time, this work highlights the impact of a better constrained observational error and the importance of these considerations when designing a DA system.

Code and data availability. The ML-based snow depth retrieval product is publicly available at (https://doi.org/10.5281/zenodo.13342108). The NASA LIS software is available at https://github.com/NASA-LIS/LISF. Publicly available in-situ snow depth and SWE data used for evaluation can be accessed at:

– https://www.doi.org/10.16904/15 (Switzerland)

– https://www.doi.org/10.16904/envidat.380 (Switzerland)

– https://www.doi.org/10.16904/envidat.590 (Switzerland)

– https://www.doi.org/10.16904/envidat.406 (Switzerland)

– https://www.arpa.piemonte.it/rischi_naturali/snippets_arpa_graphs/map_meteoweb/?rete=stazione_meteorologica (Italy)

– https://www.meteotrentino.it/index.html#!/home (Italy)

– https://data.civis.bz.it/de/dataset/p-bz-southtyrolean-weatherservice-weatherstations/resource/ef2f6f24-cffd-4993-8699-5023696a49b5 (Italy)

– https://dataset.api.hub.geosphere.at/app/frontend/station/historical/klima-v2-1d (Austria)

– https://donneespubliques.meteofrance.fr/?fond=recherche (France)

– https://cdc.dwd.de/portal/ (Germany)

Additional snow depth and SWE data were obtained from the Italian Department of Civil Protection and processed by the Centro Internazionale in Monitoraggio Ambientale (CIMA).

The configuration files used for the modeling experiments and code used for the analysis and creation of figures will be made freely available after review.

Author contributions. DD and GDL conceived and designed the study. DD conducted the model experiments, with help from MB. Material preparation and data collection were performed by DD and LB. Formal analysis, visualization, and original draft preparation was done by DD and all authors contributed to the manuscript reviewing and editing.

Competing interests. The authors declare that there are no competing interests.

Acknowledgements. This work was funded by the project C14/21/057 of KU Leuven and SNOWTRANE (SR/00/407) of the Belgian Science
410 Policy (Belspo). The computer resources and services were provided by the High Performance Computing system of the Vlaams Supercomputer Center, funded by FWO and the Flemish Government (incl. Storage4Climate collaborative grant). The authors acknowledge CIMA for providing reference data. During the preparation of this work the authors used ChatGPT in order to improve readability, flow, and language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

415 Barnett, T. P., Adam, J. C., and Lettenmaier, D. P.: Potential impacts of a warming climate on water availability in snow-dominated regions, Nature, 438, 303–309, https://doi.org/10.1038/nature04141, 2005.

Bechtold, M., Modanesi, S., Lievens, H., Baguis, P., Brangers, I., Carrassi, A., Getirana, A., Gruber, A., Heyvaert, Z., Massari, C., Scherrer, S., Vannitsem, S., and De Lannoy, G.: Assimilation of Sentinel-1 Backscatter into a Land Surface Model with River Routing and Its Impact on Streamflow Simulations in Two Belgian Catchments, Journal of Hydrometeorology, 24, 2389–2408, https://doi.org/10.1175/JHM-D-
420 22-0198.1, 2023.

Bormann, K. J., Brown, R. D., Derksen, C., and Painter, T. H.: Estimating snow-cover trends from space, Nature Climate Change, 8, 924–928, https://doi.org/10.1038/s41558-018-0318-3, 2018.

Brangers, I., Lievens, H., Getirana, A., and De Lannoy, G. J. M.: Sentinel-1 Snow Depth Assimilation to Improve River Discharge Estimates in the Western European Alps, Water Resources Research, 60, https://doi.org/10.1029/2023WR035019, 2024.

425 Broxton, P., Ehsani, M. R., and Behrangi, A.: Improving Mountain Snowpack Estimation Using Machine Learning With Sentinel-1, the Airborne Snow Observatory, and University of Arizona Snowpack Data, Earth and Space Science, 11, https://doi.org/10.1029/2023EA002964, 2024.

Daudt, R. C., Wulf, H., Hafner, E. D., Bühler, Y., Schindler, K., and Wegner, J. D.: Snow depth estimation at country-scale with high spatial and temporal resolution, ISPRS Journal of Photogrammetry and Remote Sensing, 197, 105–121,
430 https://doi.org/10.1016/j.isprsjprs.2023.01.017, 2023.

De Lannoy, G. J. M., Reichle, R. H., Houser, P. R., Pauwels, V. R. N., and Verhoest, N. E. C.: Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter, Water Resources Research, 43, https://doi.org/10.1029/2006WR005449, 2007.

De Lannoy, G. J. M., Reichle, R. H., Arsenault, K. R., Houser, P. R., Kumar, S., Verhoest, N. E. C., and Pauwels, V. R. N.: Multiscale assimilation of Advanced Microwave Scanning Radiometer–EOS snow water equivalent and Moderate Resolution Imaging Spectroradiometer
435 snow cover fraction observations in northern Colorado, Water Resources Research, 48, https://doi.org/10.1029/2011WR010588, 2012.

De Lannoy, G. J. M., Bechtold, M., Busschaert, L., Heyvaert, Z., Modanesi, S., Dunmire, D., Lievens, H., Getirana, A., and Massari, C.: Contributions of Irrigation Modeling, Soil Moisture and Snow Data Assimilation to High-Resolution Water Budget Estimates Over the Po Basin: Progress Towards Digital Replicas, Journal of Advances in Modeling Earth Systems, 16, https://doi.org/10.1029/2024MS004433, 2024.

440 de Rosnay, P., Isaksen, L., and Dahoui, M.: Snow data assimilation at ECMWF, Tech. rep., ECMWF Newsletter, 2015.

Dee, D. P.: On-line Estimation of Error Covariance Parameters for Atmospheric Data Assimilation, Monthly Weather Review, 123, 1128–1145, https://doi.org/10.1175/1520-0493(1995)123<1128:OLEOEC>2.0.CO;2, 1995.

Dee, D. P.: Bias and data assimilation, Quarterly Journal of the Royal Meteorological Society, 131, 3323–3343, https://doi.org/10.1256/qj.05.137, 2005.

445 Deems, J. S., Painter, T. H., and Finnegan, D. C.: Lidar measurement of snow depth: a review, Journal of Glaciology, 59, 467–479, https://doi.org/10.3189/2013JoG12J154, 2013.

Dunmire, D., Lievens, H., Boeykens, L., and De Lannoy, G. J.: A machine learning approach for estimating snow depth across the European Alps from Sentinel-1 imagery, Remote Sensing of Environment, 314, 114 369, https://doi.org/10.1016/j.rse.2024.114369, 2024.

Durand, M. and Margulis, S. A.: Feasibility Test of Multifrequency Radiometric Data Assimilation to Estimate Snow Water Equivalent,
450 Journal of Hydrometeorology, 7, 443–457, https://doi.org/10.1175/JHM502.1, 2006.

Estilow, T. W., Young, A. H., and Robinson, D. A.: A long-term Northern Hemisphere snow cover extent data record for climate studies and monitoring, Earth System Science Data, 7, 137–142, https://doi.org/10.5194/essd-7-137-2015, 2015.

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., da Silva, A. M., Gu, W., Kim, G.-K., Koster, R.,
455   Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and 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.

Girotto, M., Musselman, K. N., and Essery, R. L. H.: Data Assimilation Improves Estimates of Climate-Sensitive Seasonal Snow, Current Climate Change Reports, 6, 81–94, https://doi.org/10.1007/s40641-020-00159-7, 2020.

460 Girotto, M., Formetta, G., Azimi, S., Bachand, C., Cowherd, M., De Lannoy, G., Lievens, H., Modanesi, S., Raleigh, M. S., Rigon, R., and Massari, C.: Identifying snowfall elevation patterns by assimilating satellite-based snow depth retrievals, Science of The Total Environment, 906, 167 312, https://doi.org/10.1016/j.scitotenv.2023.167312, 2024.

Günther, D., Marke, T., Essery, R., and Strasser, U.: Uncertainties in Snowpack Simulations—Assessing the Impact of Model Structure, Parameter Choice, and Forcing Data Error on Point-Scale Energy Balance Snow Model Performance, Water Resources Research, 55,
465   2779–2800, https://doi.org/10.1029/2018WR023403, 2019.

Helmert, J., Şensoy Şorman, A., Alvarado Montero, R., De Michele, C., de Rosnay, P., Dumont, M., Finger, D., Lange, M., Picard, G., Potopová, V., Pullen, S., Vikhamar-Schuler, D., and Arslan, A.: Review of Snow Data Assimilation Methods for Hydrological, Land Surface, Meteorological and Climate Models: Results from a COST HarmoSnow Survey, Geosciences, 8, 489, https://doi.org/10.3390/geosciences8120489, 2018.

470 Henderson, G. R., Peings, Y., Furtado, J. C., and Kushner, P. J.: Snow–atmosphere coupling in the Northern Hemisphere, Nature Climate Change, 8, 954–963, https://doi.org/10.1038/s41558-018-0295-6, 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.,
475   Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and 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.

Hoppinen, Z., Palomaki, R. T., Brencher, G., Dunmire, D., Gagliano, E., Marziliano, A., Tarricone, J., and Marshall, H. P.: Evaluating Snow Depth Retrievals from Sentinel-1 Volume Scattering over NASA SnowEx Sites, EGUsphere [preprint], 2024.

480 Huang, C., Newman, A. J., Clark, M. P., Wood, A. W., and Zheng, X.: Evaluation of snow data assimilation using the ensemble Kalman filter for seasonal streamflow prediction in the western United States, Hydrology and Earth System Sciences, 21, 635–650, https://doi.org/10.5194/hess-21-635-2017, 2017.

IPCC: IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Tech. rep., Cambridge University Press, Cambridge, United Kingdom and New York,
485   NY, USA, 2021.

Kelly, R., Li, Q., and Saberi, N.: 'The AMSR2 Satellite-Based Microwave Snow Algorithm (SMSA): A New Algorithm for Estimating Global Snow Accumulation, in: IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, pp. 5606–5609, IEEE, ISBN 978-1-5386-9154-0, https://doi.org/10.1109/IGARSS.2019.8898525, 2019.

Kumar, S., Peterslidard, C., Tian, Y., Houser, P., Geiger, J., Olden, S., Lighty, L., Eastman, J., Doty, B., and Dirmeyer, P.: 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.

Lee, W. Y., Gim, H.-J., and Park, S. K.: Parameterizations of Snow Cover, Snow Albedo and Snow Density in Land Surface Models: A Comparative Review, Asia-Pacific Journal of Atmospheric Sciences, 60, 185–210, https://doi.org/10.1007/s13143-023-00344-2, 2024.

Lievens, H., Demuzere, M., Marshall, H.-P., Reichle, R. H., Brucker, L., Brangers, I., de Rosnay, P., Dumont, M., Girotto, M., Immerzeel, W. W., Jonas, T., Kim, E. J., Koch, I., Marty, C., Saloranta, T., Schöber, J., and De Lannoy, G. J. M.: Snow depth variability in the Northern Hemisphere mountains observed from space, Nature Communications, 10, 4629, https://doi.org/10.1038/s41467-019-12566-y, 2019.

Lievens, H., Brangers, I., Marshall, H.-P., Jonas, T., Olefs, M., and De Lannoy, G.: Sentinel-1 snow depth retrieval at sub-kilometer resolution over the European Alps, The Cryosphere, 16, 159–177, https://doi.org/10.5194/tc-16-159-2022, 2022.

López-Moreno, J. I., Revuelto, J., Fassnacht, S. R., Azorín-Molina, C., Vicente-Serrano, S. M., Morán-Tejeda, E., and Sexstone, G. A.: Snow-pack variability across various spatio-temporal resolutions, Hydrological Processes, 29, 1213–1224, https://doi.org/10.1002/hyp.10245, 2015.

Luojus, K., Pulliainen, J., Takala, M., Lemmetyinen, J., Mortimer, C., Derksen, C., Mudryk, L., Moisander, M., Hiltunen, M., Smolander, T., Ikonen, J., Cohen, J., Salminen, M., Norberg, J., Veijola, K., and Venäläinen, P.: GlobSnow v3.0 Northern Hemisphere snow water equivalent dataset, Scientific Data, 8, 163, https://doi.org/10.1038/s41597-021-00939-2, 2021.

Mankin, J. S., Viviroli, D., Singh, D., Hoekstra, A. Y., and Diffenbaugh, N. S.: The potential for snow to supply human water demand in the present and future, Environmental Research Letters, 10, 114 016, https://doi.org/10.1088/1748-9326/10/11/114016, 2015.

Miller, Z. S., Peitzsch, E. H., Sproles, E. A., Birkeland, K. W., and Palomaki, R. T.: Assessing the seasonal evolution of snow depth spatial variability and scaling in complex mountain terrain, The Cryosphere, 16, 4907–4930, https://doi.org/10.5194/tc-16-4907-2022, 2022.

Mocko, D. M., Kumar, S. V., Peters-Lidard, C. D., and Wang, S.: Assimilation of Vegetation Conditions Improves the Representation of Drought over Agricultural Areas, Journal of Hydrometeorology, 22, 1085–1098, https://doi.org/10.1175/JHM-D-20-0065.1, 2021.

Modanesi, S., Massari, C., Bechtold, M., Lievens, H., Tarpanelli, A., Brocca, L., Zappa, L., and De Lannoy, G. J. M.: Challenges and benefits of quantifying irrigation through the assimilation of Sentinel-1 backscatter observations into Noah-MP, Hydrology and Earth System Sciences, 26, 4685–4706, https://doi.org/10.5194/hess-26-4685-2022, 2022.

Musselman, K. N., Addor, N., Vano, J. A., and Molotch, N. P.: Winter melt trends portend widespread declines in snow water resources, Nature Climate Change, 11, 418–424, https://doi.org/10.1038/s41558-021-01014-9, 2021.

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., and 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, 116, D12 109, https://doi.org/10.1029/2010JD015139, 2011.

Outdoor Industry Association: The outdoor recreation economy, Tech. rep., Boulder, CO, USA, 2017.

Parthum, B. and Christensen, P.: A market for snow: Modeling winter recreation patterns under current and future climate, Journal of Environmental Economics and Management, 113, 102 637, https://doi.org/10.1016/j.jeem.2022.102637, 2022.

Peters-Lidard, C. D., Houser, P. R., Tian, Y., Kumar, S. V., Geiger, J., Olden, S., Lighty, L., Doty, B., Dirmeyer, P., Adams, J., Mitchell, K., Wood, E. F., and Sheffield, J.: High-performance Earth system modeling with NASA/GSFC's Land Information System, Innovations in Systems and Software Engineering, 3, 157–165, https://doi.org/10.1007/s11334-007-0028-x, 2007.

525 Pfeffer, W. T., Arendt, A. A., Bliss, A., Bolch, T., Cogley, J. G., Gardner, A. S., Hagen, J.-O., Hock, R., Kaser, G., Kienholz, C., Miles, E. S., Moholdt, G., Mölg, N., Paul, F., Radić, V., Rastner, P., Raup, B. H., Rich, J., and Sharp, M. J.: The Randolph Glacier Inventory: a globally complete inventory of glaciers, Journal of Glaciology, 60, 537–552, https://doi.org/10.3189/2014JoG13J176, 2014.

Pflug, J. M., Wrzesien, M. L., Kumar, S. V., Cho, E., Arsenault, K. R., Houser, P. R., and Vuyovich, C. M.: Extending the utility of space-borne snow water equivalent observations over vegetated areas with data assimilation, Hydrology and Earth System Sciences, 28, 631–648,

530 https://doi.org/10.5194/hess-28-631-2024, 2024.

Qin, Y., Abatzoglou, J. T., Siebert, S., Huning, L. S., AghaKouchak, A., Mankin, J. S., Hong, C., Tong, D., Davis, S. J., and Mueller, N. D.: Agricultural risks from changing snowmelt, Nature Climate Change, 10, 459–465, https://doi.org/10.1038/s41558-020-0746-8, 2020.

Raleigh, M. S., Livneh, B., Lapo, K., and Lundquist, J. D.: How Does Availability of Meteorological Forcing Data Impact Physically Based Snowpack Simulations?*, Journal of Hydrometeorology, 17, 99–120, https://doi.org/10.1175/JHM-D-14-0235.1, 2016.

535 Reichle, R. H., McLaughlin, D. B., and 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.

Reichle, R. H., Liu, Q., Koster, R. D., Draper, C. S., Mahanama, S. P. P., and Partyka, G. S.: Land Surface Precipitation in MERRA-2, Journal of Climate, 30, 1643–1664, https://doi.org/10.1175/JCLI-D-16-0570.1, 2017.

RGI 7.0 Consortium: Randolph Glacier Inventory - A Dataset of Global Glacier Outlines, Version 7.0, 2023.

540 Safeeq, M., Shukla, S., Arismendi, I, Grant, G. E., Lewis, S. L., and Nolin, A.: Influence of winter season climate variability on snow–precipitation ratio in the western United States, International Journal of Climatology, 36, 3175–3190, https://doi.org/10.1002/joc.4545, 2016.

Scherrer, S., De Lannoy, G., Heyvaert, Z., Bechtold, M., Albergel, C., El-Madany, T. S., and Dorigo, W.: Bias-blind and bias-aware as-similation of leaf area index into the Noah-MP land surface model over Europe, Hydrology and Earth System Sciences, 27, 4087–4114,

545 https://doi.org/10.5194/hess-27-4087-2023, 2023.

Slater, A. G. and Clark, M. P.: Snow Data Assimilation via an Ensemble Kalman Filter, Journal of Hydrometeorology, 7, 478–493, https://doi.org/10.1175/JHM505.1, 2006.

Smyth, E. J., Raleigh, M. S., and Small, E. E.: Improving SWE Estimation With Data Assimilation: The Influence of Snow Depth Observation Timing and Uncertainty, Water Resources Research, 56, https://doi.org/10.1029/2019WR026853, 2020.

550 Smyth, E. J., Raleigh, M. S., and Small, E. E.: The Challenges of Simulating SWE Beneath Forest Canopies are Reduced by Data Assimilation of Snow Depth, Water Resources Research, 58, https://doi.org/10.1029/2021WR030563, 2022.

Steiger, R., Scott, D., Abegg, B., Pons, M., and Aall, C.: A critical review of climate change risk for ski tourism, Current Issues in Tourism, 22, 1343–1379, https://doi.org/10.1080/13683500.2017.1410110, 2019.

Sturm, M., Goldstein, M. A., and Parr, C.: Water and life from snow: A trillion dollar science question, Water Resources Research, 53,

555 3534–3544, https://doi.org/10.1002/2017WR020840, 2017.

Tedesco, M. and Narvekar, P. S.: Assessment of the NASA AMSR-E SWE Product, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 3, 141–159, https://doi.org/10.1109/JSTARS.2010.2040462, 2010.

Terzago, S., Andreoli, V., Arduini, G., Balsamo, G., Campo, L., Cassardo, C., Cremonese, E., Dolia, D., Gabellani, S., von Hardenberg, J., Morra di Cella, U., Palazzi, E., Piazzi, G., Pogliotti, P., and Provenzale, A.: Sensitivity of snow models to the accuracy of meteorological

560 forcings in mountain environments, Hydrology and Earth System Sciences, 24, 4061–4090, https://doi.org/10.5194/hess-24-4061-2020, 2020.

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

565 Verseghy, D. L.: Class—A Canadian land surface scheme for GCMS. I. Soil model, International Journal of Climatology, 11, 111–133, https://doi.org/10.1002/joc.3370110202, 1991.

Vorkauf, M., Marty, C., Kahmen, A., and Hiltbrunner, E.: Past and future snowmelt trends in the Swiss Alps: the role of temperature and snowpack, Climatic Change, 165, 44, https://doi.org/10.1007/s10584-021-03027-x, 2021.

Wasti, A., Ray, P., Wi, S., Folch, C., Ubierna, M., and Karki, P.: Climate change and the hydropower sector: A global review, WIREs Climate

570 Change, 13, https://doi.org/10.1002/wcc.757, 2022.

Yang, Z.-L., Niu, G.-Y., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Longuevergne, L., Manning, K., Niyogi, D., Tewari, M., and Xia, Y.: The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins, Journal of Geophysical Research, 116, D12 110, https://doi.org/10.1029/2010JD015140, 2011.