Spatio-temporal snow data assimilation with the ICESat-2 laser altimeter

Marco Mazzolini[1], Kristoffer Aalstad[1], Esteban Alonso-González[2], Sebastian Westermann[1], and Désirée Treichler[1]
[1]Department of Geosciences, University of Oslo, Oslo, Norway
[2]Pyrenean Institute of Ecology, CSIC, Jaca, Spain.
Correspondence: Marco Mazzolini (marcomaz@uio.no)

Abstract. The satellite laser altimeter ICESat-2 provides accurate surface elevation observations across the globe. ~~Where~~ With a high-resolution DEM~~is available~~, we can use ~~these~~ such measurements to retrieve snow depth profiles even in remote areas where snow amounts are poorly constrained, despite being of great societal interest. However, the adoption of these retrievals remains low since they are very sparse in space~~(~~, as the satellite measures along profiles~~)~~ , and in time~~(~~, as the revisit is 3 months~~)~~. Data assimilation methods can exploit snow observations to constrain snow models and provide gap-free ~~snow map time series. Assimilation~~ distributed simulations. The assimilation of observations like snow cover is established, but there are currently no methods to assimilate sparse ICESat-2 snow depth profiles. We propose an approach that spatially propagates information using – instead of the classic geographical distance – an abstract distance measured in a feature space defined by ~~topographical parameters~~ a topographical index and the melt-out date climatology.

We demonstrate this framework for a small experimental catchment in the Spanish Pyrenees through three experiments. We assimilate different snow observations in an intermediate-complexity snow model: fractional snow cover retrievals from Sentinel-2, snow depth profiles from ICESat-2 located in proximity of the catchment, or both snow cover and depth in a joint assimilation experiment. Results show that assimilating ICESat-2 snow depth profiles successfully updates the ~~neighboring~~ neighbouring *unobserved* catchment, improving the simulated average snow depth compared to the prior run. ~~Moreover,~~ Another encouraging finding is that adding the snow depth profiles to fractional snow-covered area observations leads to an accurate reconstruction of the snow depth spatial distribution, improving the skill score by ~~22%~~19% in the accumulation season.

1 Introduction

Seasonal snow is ~~characterized by a~~ a crucial element for sustaining human life and an essential climate regulator (Sturm et al., 2017) . It is characterized by strong spatial and temporal variability ~~(Mott et al., 2018) arising~~ which arise from several processes such as preferential deposition, wind transport, differential radiation and turbulent heat fluxes, ~~metamorphism~~and metamorphism (Mott et al., 2018). This is a challenge for spatially-distributed modelling efforts as point measurements have a limited utility. ~~Manual snow measurements (of depth, snow water equivalent (SWE) or density) are present in populated areas, but this is not the case for remote mountain ranges (Orsolini et al., 2019), or above the treeline (Treichler and Kääb, 2017). Similarly,~~

field campaigns are expensive and limited in their spatio-temporal extent. Remote sensing is thus the most Remote sensing thus presents a more promising method to estimate the spatial distribution of snow (Clark et al., 2011). Moreover, Dozier et al. (2016) state that the study of snow water equivalent (SWE) distribution is the sector in hydrology that would benefit the most from remote sensing innovations.

Optical satelliteplatforms are routinely used to retrieve the fractional snow-covered area (fSCA) which is the key observed variable to reconstruct the seasonal snow evolution (Margulis et al., 2016). Several products are available, including operational ones such as long-term but coarse global ESA Snow_cci snow cover data from AVHRR (Naegeli et al., 2022), moderate resolution global snow cover data from MODIS and VIIRS (Riggs et al., 2017), and higher resolution snow cover retrievals from Sentinel-2 and Landsat available in limited regions (Gascoin et al., 2019). The accuracy of fSCA retrievals varies considerably depending on the retrieval algorithm and sensor employed (Aalstad et al., 2020). Both coarser and higher resolution fSCA retrievals have been shown to help reconstruct the evolution of the seasonal snowpack across entire mountain ranges (Alonso-González et al. .

Accurately measuring snow amounts Despite the many approaches involving satellite, airborne and drone sensors of various different types currently being used, accurately measuring the temporal and spatial variability in snow amount (i.e., mass or depth) from space is still a major scientific challenge (Dozier et al., 2016), but many approaches are currently used. Satellite measurements with passive microwave sensors have coarse spatial resolution (tens of kilometers) and saturate with a deep snowpack, limiting their applicability in complex terrain,(Foster et al., 2005) as in the state-of-the-art global passive microwave SWE product GlobSnow v3.0 (Luojus et al., 2021) where mountain regions are masked out. Active microwave sensors have also been experimented with, and Lievens et al. (2022) obtained snow depth estimates at medium spatial resolution (500 m) from Sentinel-1 Synthetic Aperture Radar (SAR) backscatter information and an empirical change detection algorithm.

In the last decades, unpiloted aerial vehicles (UAV) and airplanes have been widely used to map the spatial distributions of snow depth on a catchment to smaller scales. Photogrammetry (Eberhard et al., 2021), and light detection and ranging (lidar; Geissler et al., 2023; Harder et al., 2019) have been extensively used for this purpose. The most notable effort is the airborne snow observatory (ASO; Painter et al., 2016)), where snow depth has been extensively mapped over a large number of basins over the western part of the North American continent. However, the costs of performing such campaigns remain high and organizationally complex, making this approach prohibitively expensive for mapping seasonal snow globally.

Photogrammetry can be applied also with space-borne platforms acquiring high-resolution stereo imagery (Marti et al., 2016)), but the costs of acquiring imagery from private companies that operate very high-resolution optical satellites with the required specifications currently limit its use. Lidar technology is also used onboard satellites, for example snow depth has been estimated with the NASA satellite ICESat (Treichler and Kääb, 2017). However, the coarse footprint limits the applications of this technology in complex topography(Gascoin et al., 2024).

In Since October 2018, the ICESat-2 mission was launched. The Advanced Topographic Laser Altimeter System (ATLAS) mounted onboard has significantly better sensor characteristics for measuring snow depth. It emits measures the Earth's surface

elevation through a 532 nm (green) laser ~~pulses that illuminate~~ along six parallel profiles ~~along which it measures the earth's surface elevation (Markus et al., 2017). On average, for an area located at mid-latitudes extending around 10 km in longitude, ICESat-2 is expected, in the absence of clouds, to scan the surface with the 6 beams once or twice during the snow season.~~ (Markus et al., 2017). The geolocated photons have a centimetric vertical measurement error ~~(although on flat terrain Markus et al., 2017) and the geolocation~~ on flat terrain (Markus et al., 2017), while the horizontal accuracy is estimated ~~at 3 to~~ to be between 3 and 4 m (Magruder et al., 2021). ICESat-2 data products derived from the spatial aggregation of photons over tens to a hundred meters have been used to measure snow depth by differencing with a digital elevation model (DEM) acquired during snow-free conditions. From comparison to airborne lidar snow depth observations, Enderlin et al. (2022) found that the median absolute deviation (MAD), an index of accuracy, is around 0.2 m for slopes $< 5°$, while it increases to be > 1 m for slopes $> 20°$. Deschamps-Berger et al. (2023) found similar results: the random error (precision) was 0.5 m for $< 10°$ sloped terrain. Besso et al. (2024) improved the results in the same basins of the two ~~works~~ aforementioned studies by focusing on customized data products generated with the SlideRule Earth service (Shean et al., 2023).

~~All of the aforementioned observations~~ The ICESat-2 snow depth retrievals can quantify the snow spatial distribution at the acquisition time. However, they have limited value in directly estimating snow water resources, ~~because they~~ a trait they share with other commonly used remotely-sensed observations such as fractional Snow Covered Area (fSCA). This limitation arises because ICESat-2 observations are only able to capture a subset of the full state of the snowpack at one to several points in time and space ~~(because of their coverage, and spatial and temporal resolution)~~(Alonso-González et al., 2021; Margulis et al., 2016). In order to ~~capture the spatiotemporal~~ completely capture the spatio-temporal evolution of the snowpack, ~~these~~ satellite observations have been used to constrain ~~a plethora of~~ snow models with varying complexity~~, from simple empirical models with one to several parameters (Hock, 1999), to physically-based models that represent the snow energy and mass balance with many distinct layers (Lehning et al., 2002)~~. Herein, we focus on an intermediate complexity snow model, namely the Flexible Snow Model ~~(FSM2 Essery, 2015)~~(FSM2; Essery et al., 2024), which is a compromise between a detailed representation of physical processes that influence the key snow hydrological state variables and parsimonious parametrizations that allow for increased computational efficiency. ~~The~~ Observations such as snow depth or fSCA can mitigate the main limitation of all snow hydrology models ~~, independently of complexity,~~ which has been shown to be the accuracy of atmospheric forcing data (Raleigh et al., 2015)~~, especially in the~~ . This applies especially in a potentially globally-applicable spatially-distributed ~~case~~ setup when the forcing needs to be extracted from ~~larger~~ large scale atmospheric model outputs such as ~~coarse resolution~~ coarse-resolution (30 km) global atmospheric reanalyses (e.g. ERA5; Hersbach et al., 2020).

Data Assimilation (DA) ~~enables the fusion of~~ offers a wealth of algorithms to fuse noisy observations with uncertain models in a Bayesian ~~statistical~~ framework (Evensen et al., 2022), ~~obtaining (statistically) optimal estimates with an associated~~ to obtain statistically optimal estimates together with a quantification of their uncertainty. This technique ~~, especially ensemble-based (Monte Carlo) implementations,~~ has shown considerable promise in the snow science community to meet the reconstruction and forecasting requirements needed to more accurately map the water storage services that snow provides to downstream ecosystems and communities (Girotto et al., 2020). ~~In terms of time dynamics, these schemes can be employed either in~~

~~a strictly sequential forward manner as filters (e.g. for initializing short-term snow hydrological forecasts Mott et al., 2023), or instead as retrospective smoothers that allow information from observations to transfer backward in time, yielding a constrained and consistent reconstruction (ideal for snow reanalysis problems Margulis et al., 2016).~~

~~Ensemble DA methods can be further subdivided between particle and Kalman methods (Evensen et al., 2022). Particle filters (Leisenring and Moradkhani, 2011) and smoothers (Margulis et al., 2015) are particularly popular snow DA methods given their flexibility, ease of implementation, and relative lack of assumptions. However, particle methods are prone to undesirable ensemble collapse due to weight and path degeneracy (Murphy, 2023), especially in higher dimensional spatio-temporal problems (Cressie, 2011) that remains an active area at the frontier of particle DA research (Evensen et al., 2022).~~

~~Herein, we restrict our attention to ensemble Kalman methods (Evensen et al., 2022) which, despite (and thanks to) stronger Gaussian linear assumptions, have been shown to be robust also in very high-dimensional geophysical DA problems (Carrassi et al., 2018) .These ensemble Kalman techniques represent distributions through an ensemble of model realizations that are updated in state and/or parameter space using available observations. Moreover, the use of iterative ensemble Kalman updates that temper the likelihood such as the ensemble smoother with multiple data assimilation scheme (ES-MDA Emerick and Reynolds, 2013) , and the deterministic version (DES-MDA Alonso-González et al., 2023) has been shown to strongly mitigate the negative impact of a linear forward model assumption implicit of such methods (Aalstad et al., 2018; Evensen et al., 2022).~~

~~Ensemble Kalman methods have been widely applied to various snow DA problems and have been used to assimilate SWE data from stations(Magnusson et al., 2014) and passive microwave satellite retrievals (De Lannoy et al., 2012), snow depth data from stations (Stigter et al., 2017) and drones (Alonso-González et al., 2022), and fSCA satellite retrievals (De Lannoy et al., 2012; Stigter . Moreover, the ensemble smoother, a batch ensemble Kalman smoother (see Evensen et al., 2022; Alonso-González et al., 2022) , was suggested as a method for Bayesian snow reconstruction by Durand et al. (2008) and has been subsequently used to assimilate both moderate (Oaida et al., 2019) and higher resolution (Girotto et al., 2014) fSCA retrievals. More recently, iterative ensemble smoothers based on the ES-MDA have shown promise in the assimilation of satellite-based fSCA retrievals (Aalstad et al., 2018) or drone-based snow depth retrievals (Alonso-González et al., 2023).~~

~~Most~~ Girotto et al. (2020) noted that most snow DA research ~~, –~~ with a few exceptions ~~(e.g. Magnusson et al., 2014; De Lannoy et al., 2012~~ , (e.g. De Lannoy et al., 2012; Magnusson et al., 2014) – has focused on purely temporal DA where the snow in each model grid cell (or more generally spatial unit) is simulated and updated independently of its ~~neighboring~~ neighbouring cells and the observations therein. ~~However,~~ De Lannoy et al. (2022) recommend a greater adoption of spatio-temporal multivariate DA. Recent studies have demonstrated the added value of spatio-temporal DA in exploiting spatially sparse snow depth observations ~~propagating the~~ by propagating information through covariances based on geographical distance ~~(Cluzet et al., 2022). Alonso-González et al. (2023), in contrastwith the other aforementioned exceptions,~~ (Cluzet et al., 2022; Cho et al., 2023). In contrast, Alonso-González et al. (2023) have shown promising results with ~~the ES-MDA scheme and~~ a more general technique using a prior covariance matrix dependent on pixel proximity in a multi-dimensional space based on morphometric terrain features. However, this approach was shown only with high resolution and high accuracy UAV data, and has yet to be extended

to emerging yet sparse space-borne ~~observations (see comment 6 ?) of snow depth, such as~~ snow observations such as snow depth profiles derived from the laser altimeter ~~on~~ onboard ICESat-2 (Deschamps-Berger et al., 2023; Besso et al., 2024), as recommended by comment 6 in Anonymous (2023) and by Gascoin et al. (2024).

In this study, we aim to demonstrate for the first time the value of assimilating snow depth retrieved from the satellite laser altimeter ICESat-2 in a small experimental catchment. While other studies have focused on evaluating ~~spatially~~ aggregated data products from this satellite altimeter (Deschamps-Berger et al., 2023; Enderlin et al., 2022; Besso et al., 2024), we use the geolocated photon data product ATL03 (Neumann et al., 2019) with finer spatial resolution. The validation of this product is ~~undergoing work~~currently in progress, but preliminary results show good agreement in comparison to drone-based snow depth maps. Because of the temporal and spatial sparsity of these observations, their utility for direct mapping of snow depth along profiles via temporal DA is limited. Hence, we showcase the use of ICESat-2 data in a spatio-temporal DA scheme (Alonso-González et al., 2023) that enriches this data in that it can now be used to constrain distributed seasonal snow models for an entire catchment where the majority of grid cells are not observed by ICESat-2.

We also propose the joint assimilation of ICESat-2 snow depth and fSCA from Sentinel-2. This is compared to assimilating only ICESat-2 data or fSCA data alone, ~~whereby~~ where the latter is widely available and routinely used ~~in~~ by the snow DA community (Largeron et al., 2020). The two datasets have complementary features: ICESat-2 retrieves snow depth directly, but only along profiles; while fSCA has an indirect ~~correlation~~ relationship with snow depth, but ~~this dataset~~ is spatially distributed. ~~Our hypothesis is that the joint assimilation will be able to exploit these to better infer the seasonal snowevolution~~The novel scientific questions we aim to answer are:

a) Can information from snow depth profiles retrieved with ICESat-2 be used to provide information about average catchment-scale snow depth and its complete spatial distribution?

b) Is assimilating sparse ICESat-2 snow depth retrievals better than more commonly used fSCA observations derived from optical satellites?

c) Is ensemble-based DA able to leverage information from both observation types when jointly assimilating both fSCA and sparse snow depth observations?

The underlying hypothesis of this work, shared with other high-resolution DA assimilation studies, can be outlined as follows. The snow model does not represent some of the lateral processes influencing the snowpack such as wind and gravitational redistribution: the atmospheric forcing information that is available at scale is not detailed enough to accurately capture such processes in a distributed manner. We aim at showing that assimilating information from higher resolution satellite observations can constrain the spatial distribution of snow, mitigating the effects resulting from missing the aforementioned processes that are not represented in the snow model.

2 Study Area and Data

The study area is the Izas experimental catchment~~(Revuelto et al., 2017)~~, located in the Central Spanish Pyrenees. The elevation ranges between 2075 and 2325 m a.s.l. and the total annual precipitation typically sums up to 2000 mm, with around half of this in the form of snowfall as is typical for such a sub-alpine environment. Snow usually covers a large fraction of the catchment from late November to the end of May. The vegetation type is a high mountain steppe ~~:~~ where most of the surface is covered by bunch grass ~~.~~(Revuelto et al., 2017).

We chose this site because of the availability of several spatially distributed drone-based snow depth measurements (Alonso-González, 2022) in the 55 ha ~~area~~ sized experimental catchment highlighted with black in Figure 1~~(Revuelto et al., 2021)~~. The drone surveys were conducted in 2020 on 14/01, 03/02, 24/02, 11/03, 29/04, 3/05, 12/05, 19/05, 26/05, 02/06, 10/06, 21/07 and were already used in various DA experiments ~~(e.g. Alonso-González et al., 2022, 2023)~~(Alonso-González et al., 2022, 2023). The original spatial resolution is 1 m, and the measurement error is assumed to have a standard deviation equal to 20 cm. ~~The right panel~~ Panel c) of Figure 1 shows the spatial distribution of snow depth close to the seasonal peak at the native resolution. Very high snow depths (> 300 cm) are accumulated under the ~~ridgeline on the west side of the area~~western ridgeline, in the deep valleys throughout the catchment, and at the foot of the ~~slope on the east side~~eastern slope. Very low snow depths (< 50 cm) ~~can be seen in the south-southeast facing~~ areas are located on the southeast-facing aspects located on the ~~north~~ northern side of the catchment as well as on the wind-blown ridges located on the southern side of the catchment.

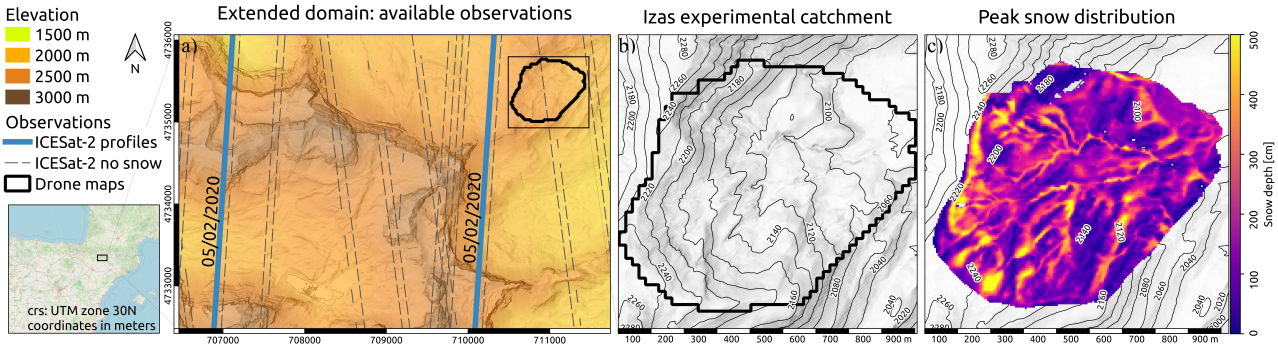

Figure 1. Panel a): Topography of the ~~Izas experimental basin~~extended domain, located in the Pyrenees. The raw DEM data is obtained from the Centro Nacional de Información Geográfica. Blue lines: ICESat-2 snow depth profiles in 2020 used in the experiments; gray dashed lines: ICESat-2 snow-free profiles that are used for coregistration purposes; black area: experimental catchment where 12 snow depth maps were acquired during the 2020 snow season. ~~The~~ Panel b): the small-scale topography ~~, shown in~~ of the ~~middle panel~~Izas experimental catchment, ~~dictates~~ dictating the ~~peak~~ snow depth patterns~~shown in the right panel, which was~~ . Panel c): drone-derived snow depth map acquired on 11th of March shown at the original 1 m resolution~~and is~~ , obtained from Alonso-González (2022).

We ~~obtain~~ retrieve snow depth by subtracting the snow-off elevation obtained with a reference DEM to the snow surface elevation measured ~~by the ICESat-2 satellite altimeter to a snow-off DEM. The~~ with the ATLAS instrument onboard the National Aeronautics and Space Administration (NASA) satellite ICESat-2~~provides very accurate surface height measurements along six beams (Markus et al., 2017)~~. We use the level 2A product ATL03 ~~(Neumann et al., 2023a)~~(Neumann et al., 2023b), which we downloaded through the SlideRule python-client (Shean et al., 2023). This low-level product combines the time of flight data with the location of the satellite to obtain the geolocation of single photons' reflection events off the snow surface. The ATLAS transmitter generates six beams ~~(across-track direction). The beams are~~ arranged in three pairs, with a strong and weak beam located at a distance of 90 m, while the pairs are spaced 3.3 km (Neumann et al., 2019). The orbit has a 92° inclination. The signal ratio between strong and weak beams is 4 to 1 ~~. One~~ and a pulse of the 532 nm laser illuminates a region (footprint) ca. 14 m in diameter. Pulses are transmitted every 1.5 ns, hence the footprints are spaced 0.7 m in the ~~along-track~~ along-profile direction and have a large overlap. The expected number of photons backscattered to the ATLAS receiver for a single strong beam shot from a reflective surface such as snow ranges from 7 to 3 depending on the slope and roughness of the surface (Neumann et al., 2019). We select two of the strong beam profiles acquired ~~on~~ in the night of the 5th of February 2020, as they sample snow depth during the accumulation phase. We tested the weak beams but they were not used as those provided an observation set with fewer returns in the 20 m cell we used as unit, and a consistent vertical shift was detected. We discard the third strong beam as it is located further away at lower elevations and in forested terrain~~, and the weak beams as preliminary results showed their signal-to-noise ratio is too weak for this application~~. ICESat-2's next acquisition during the 2020 snow season in the study area is in May, but the cloudy conditions weakened the surface signal and made these profiles unreliable. ~~ICESat-2 measures the surface elevation through the geolocation of photon reflection events, and we access this information through the low-level ATL03 data product (Luthcke, 2021).~~ Most of these events happen on the highly reflective snow surface during the ~~cold~~ snow-covered season, but there is a ~~substantial~~ relevant amount of noise due to ~~solar radiation~~ atmospheric scattering or double bounces (Neumann et al., 2023a). ~~Photon events are divided into several classes depending on the identified location of the event (e.g. ground, canopy, top of the canopy) in the algorithm vegetation product ATL08 (Shean et al., 2023).~~

The DEM used as the snow-off reference surface is available thanks to the Spanish government's Plan Nacional de Ortofotografía Aérea (PNOA) airborne lidar initiative and its spatial resolution is 2 m. The accuracy in terms of RMSE for this lidar-based model is declared to be 20 cm in the vertical direction and 30 cm in the horizontal plane (Centro Nacional de Información Geográfica). ~~Moreover, as we will detail in the following section 3, we~~ We use snow-off ICESat-2 profiles to evaluate the vertical offset between the ~~mentioned~~ aforementioned DEM and the ICESat-2 acquisitions, as detailed in Section 3. We employ 18 profiles depicted with a gray dashed line in Figure 1.

In addition to the ICESat-2's snow depth data, we also employ snow cover information from high-resolution ($\sim$ 10 m) multispectral satellite imagery. We used surface reflectances (the Level 2A product) obtained from the MultiSpectral Instrument (MSI) onboard the Sentinel-2A and Sentinel-2B twin satellites (Drusch et al., 2012) operated by the European Space Agency as part of the Copernicus Programme. The Sentinel-2 imagery was downloaded from Google Earth Engine (Gorelick et al.,

2017), which is a cloud-based platform that ~~harvests~~ hosts open Earth observation data from its original source, in this case Copernicus. By manually selecting cloud-free imagery, we obtained a total of 19 scenes covering the entire study area and snowmelt season. The acquisition dates of the used scenes are shown with blue stars in panel a) of Figure ??5, and are irregularly spaced between the 5th of February (before peak snow) and the 17th of July 2022 (complete melt-out), with a median (maximum) spacing of 5 (33) days.

3 Methods

3.1 ~~ICESat-2 snow depth retrieval~~Modelling

We simulate the snowpack at 20 m spatial resolution with the Flexible Snow Model (FSM2; Essery, 2015, 2023) . This intermediate-complexity model represents the snowpack with up to three different layers and solves the coupled mass and energy balance equations to simulate the seasonal evolution of snow. In FSM2, seven physical processes occurring in the snowpack are represented with multiple available parameterizations. We choose the most complex representation for all the processes to obtain a more comprehensive snowpack ensemble simulation. These parametrizations are: albedo decay with elapsed time since the last significant snowfall, thermal conductivity as a function of snow density, density as a function of overburden and metamorphism, turbulent fluxes diagnosed using the Monin-Obukhov similarity theory, and melt-water percolation as a function of gravitational drainage, fractional snow cover as an asymptotic function of snow depth. No lateral processes are represented in FSM2.

We drive the simulations with meteorological forcing derived from the globally-available ERA5 reanalysis (Hersbach et al., 2020) . Because the coarse resolution of 30 km misses the subgrid topography-driven heterogeneity of the atmospheric variables, we downscale the reanalysis with the widely used hillslope scale topograpy-based downscaling tool TopoSCALE (see Filhol et al., 2023, and references therein). This process uses the pressure level data to interpolate the forcing variables to the grid cell elevation, and radiative fluxes are adjusted by taking into account local and surrounding topography as well as the position of the sun. The TopoSUB routine included in the downscaling tool allows for efficient semi-distributed spatial downscaling through topography-based clustering (Fiddes et al., 2019). We select 400 as an appropriate number of clusters to run the semi-distributed downscaling for the extended domain which has a rather small area of about 5 by 3 km (equal to the extent of panel a) of Figure 1). The obtained semi-distributed forcing is then mapped back to a 20 m fully distributed grid covering the whole extended domain. This globally applicable TopoSCALE metod has for example been used in similar settings to successfully topographically-downscale ERA5 reanalysis data to drive snow models both in hyper-resolution snow data assimilation experiments Fiddes et al. (2019) and in nationwide hillslope-scale snow simulations with FSM (Fiddes et al., 2022) . The application of topographically downscaled ERA5 data to force the snow model FSM2 has already been conducted in this area in Alonso-González et al. (2022, 2023) – although another downscaling model was used. We now select the cells covering the drone maps in the Izas experimental catchment (solid black line in Figure 1) and the grid cells in the extended domain that

are intersected by the ICESat-2 tracks (blue lines in panel a), Figure 1), summing up to a total of ~ 1900 cells where the FSM2 model is run.

3.2 ICESat-2 snow depth retrieval

~~As mentioned in the previous section, the laser altimeter ATLAS on-board of the satellite~~ The retrieval of snow depth observations is based on ICESat-2~~geo-locates photon reflection events to retrieve the surface elevation~~'s accurate surface elevation measurements from individual photon reflection events. We filter such events by selecting only the ones classified as ground, as results from Besso et al. (2024) indicate such filtering improves the median absolute error. Moreover, we assign the photon events a weight based on the local ~~neighborhood~~ neighbourhood density, using the Yet Another Photon Classifier (YAPC) algorithm~~(Sutterley and Gibbons, 2021)~~. This determines the significance of individual photon events with a customized inverse-distance weighted kNN algorithm ~~. The neighborhood~~ (Sutterley and Gibbons, 2021). The neighbourhood is defined with a window length (parallel to the line of flight) of 5 m and a 3 m height. The rationale behind this is that photons returning from the ground have a large number of ~~neighbors~~ neighbours as it is less likely for photons reflected from atmospheric particles or objects above the ground to be clustered together. For each ~ 4 km profile we select 60% of the photons with the largest significance according to YAPC. In Figure 2, the orange photons' size is proportional to their YAPC score, while the filtered-out photons are ~~grey~~gray. Before comparing the ATL03 photon events to the snow-free reference surface elevation it is necessary to co-register this dataset with the snow-off DEM. Every beam is independently co-registered with a horizontal displacement~~, and a~~ . A vertical offset common to all the acquisitions and beams is obtained by computing the median of all the snow-off acquisitions vertical offsets, as other studies have done (Enderlin et al., 2022; Besso et al., 2024). We employ the Nuth-Kääb algorithm to obtain the horizontal shifts (Nuth and Kääb, 2011), implemented in the xdem python library (Dehecq et al., 2021).

Snow depth is computed for each selected photon event by subtracting the elevation from the co-registered DEM. We linearly interpolate the DEM to obtain the snow-off elevation at the location of the photon event. Subsequently, we divide the snow depth observations into cells with a 20 m spatial resolution, in order to match the spatial resolution of the simulation (see ~~3.6). Since the ICESat-2 orbit is 92°, each cell has around 29 footprints summing up to~~ Section 3.1). On average, 45 ± 18 ~~photons after the YAPC filtering. As~~ individual snow depth measurements per cell are available. Due to ICESat-2's inclined ground track (see panel a of Figure 1) some of the cells defined by modelling grid ~~might~~ have very few measurements, we hence filter out cells ~~where~~ with less than 10 photon reflection events. In addition, also the cells with an average slope ~~larger~~ greater than $40°$ are filtered out, as the horizontal positioning uncertainty makes snow depth retrievals ~~not~~ less reliable for steep terrain. In Figure 2, a 400 m transect with a comparison between the co-registered photons and the high-resolution DEM is shown (~~upper panel~~ panel a)) as well as the snow depth sample distribution available along the transect (~~lower panel~~ panel b)). We sample the obtained snow depth distribution to retrieve an observation in each cell with the median operator. To estimate a domain-consistent statistic for the spread of the snow depth observation error~~(σ_y)~~, we compute the standard deviation of snow depth samples in each cell and average it over the profiles, obtaining $\sigma_y = 0.92$ m.

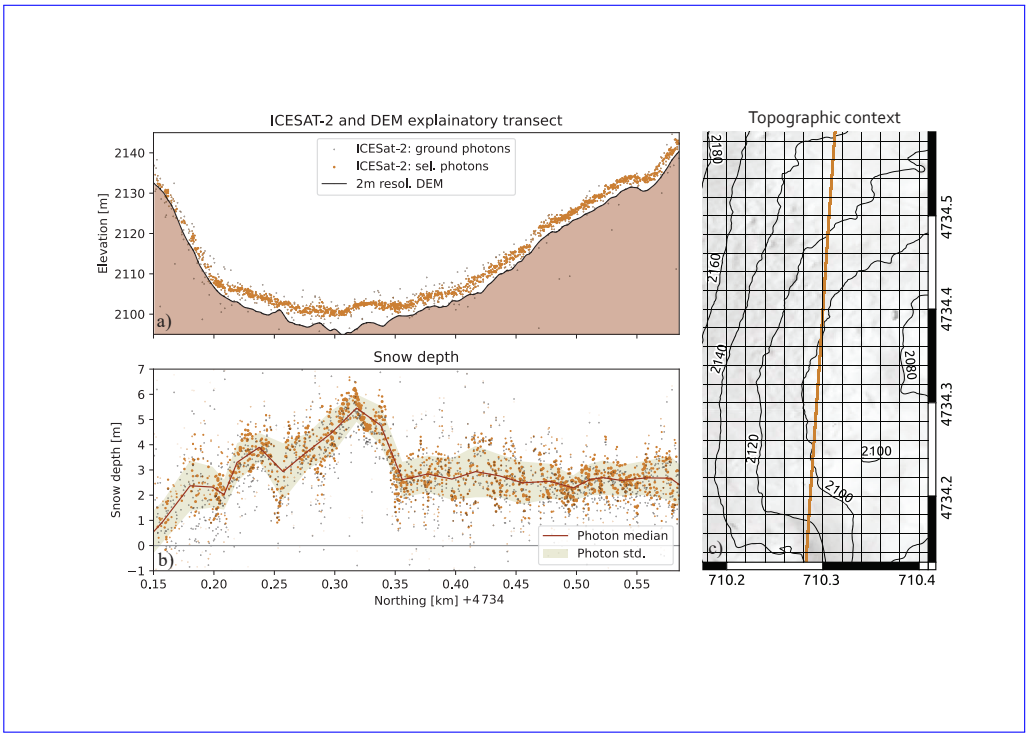

Figure 2. ~~Left panel: the topography of the ICESat-2 profile. Upper panel~~Panel a): A 400 m long segment where the photon events selected are compared with the snow-off information obtained from the 2 m resolution DEM. ~~Lower panel~~Panel b): the median operator is applied to snow depth observations from the selected photon on a 20 m moving window and the shaded area represents the 2 standard deviation range around the median. Panel c): the general topography of the ICESat-2 profile shown.

3.3 Sentinel-2 fSCA retrieval

The fSCA retrieval algorithm is based on surface reflectances estimated from MSI onboard the Sentinel-2 satellites in 6 bands in the visible, near-infrared, and shortwave infrared. By comparing the reflectances measured in these bands to modeled spectra for snow and snow-free endmembers we infer the fSCA using spectral unmixing via a fully constrained least squares algorithm. As shown in Aalstad et al. (2020), this approach can outperform simpler linear regression and thresholding-based approaches to fSCA retrieval, albeit at a higher computational cost. This unmixing approach to retrieving fSCA from Sentinel-2 imagery has been successfully applied both in the high-Arctic (Aalstad et al., 2020), as well as at Alpine sites (Pirk et al., 2023) in Norway. The retrievals were performed in the native 10 m grid of the visible and near-infrared reflectances from Sentinel-2, and were subsequently ~~regridded~~ resampled to the 20 m spatial resolution of the simulations ~~3.6~~ through averaging and subsequent inverse distance weighting. We estimate the observation error for the fSCA retrievals at 20 m resolution to be $\sigma = 0.34$. ~~As independent~~ Independent validation estimated the observation error σ_N at 100 m resolution to be equal to $\sigma_N = 0.07$ (see Table 2 Aalstad et al., 2020)~~, we expect the error at coarser resolution to improve~~. We obtained our 20 m σ estimate using that the

error from coarser resolutions should increase at higher resolutions according to the central limit theorem ~~$\sigma_N = \frac{\sigma}{\sqrt{N}}$, where N~~ through $\sigma = \sigma_N\sqrt{N}$, where $N = 25$ is the number of independent 20 m cells ~~being aggregated~~ contained in the coarser ~~validation ($N = 25$ in this case). Thus, the disaggregated observation error at 20 m resolution for these fSCA retrievals should be on the order $\sigma = \sigma_N\sqrt{N} \simeq 0.34$ from which we obtained our estimate.~~

3.4 ~~Modelling~~

~~We simulate the snowpack at 20 m spatial resolution with the Flexible Snow Model (FSM2 Essery, 2015, 2023). This intermediate-complex model represents the snowpack with up to three different layers and solves the coupled mass and energy balance equations to simulate the seasonal evolution of snow. To obtain a more comprehensive snowpack representation, 7 physical processes are parameterized with the most detailed process representation among those available in the snow model. These parametrizations are: albedo decay with elapsed time since the last significant snowfall, thermal conductivity depending on snow density, density influenced by overburden and metamorphism, turbulent fluxes diagnosed using the Monin-Obukhov similarity theory, and melt-water percolation depending on gravitational drainage, fractional snow cover asymptotic to snow depth.~~

~~We drive the simulations with meteorological forcing derived from the ERA5 reanalysis (Hersbach et al., 2020). Because the coarse resolution of 30 km misses the subgrid topography-driven heterogeneity of the atmospheric variables, we downscale the reanalysis with TopoSCALE (Filhol et al., 2023). This process uses the pressure level data to interpolate the forcing variables at the cell elevation, and radiation components are scaled depending on the topography. The TopoSUB routine in TopoSCALE allows for efficient semi-distributed spatial downscaling through topography-based clustering (Fiddes et al., 2019). We select 400 as an appropriate number of clusters to run the semi-distributed downscaling for this rather small area (equal to the extent of the map in the left panel of Figure 1). The obtained semi-distributed forcing is then mapped back to the 20 mfully distributed grid. Such a combination of topographic downscaling (although another downscaling model was used) and the snow model FSM2 has already been used in this area in Alonso-González et al. (2022, 2023). Note that to save computational resources we simulate only the cells in the domain where an ICESat-2 observation or the drone validationis available (~ 1900 cells).~~ (100 m) validation.

3.4 Data assimilation

The Multiple Snow DA System ~~(MuSA Alonso-González et al., 2022)~~ (MuSA; Alonso-González et al., 2022) allows the execution of various forms of ensemble-based snow DA. Therein, the prior ~~uncertainty~~ distribution — a probabilistic distribution representing uncertainty over the system's state and parameter space before observations are taken into account — is represented by the spread of ~~the~~ a finite collection of samples known as ensemble members. ~~Each~~ This spread in terms of basin-average snow depth can be seen in the gray trajectories of panels a), b) and c) of Figure 5. Each prior ensemble member is an FSM2 simulation obtained by perturbing a selection of forcing variables. In the presented experiments, the perturbed forcing variables are air temperature, precipitation and downwelling longwave radiation. The perturbation parameters are time-invariant throughout the water year, and ~~the prior perturbation parameters are extracted via transformations~~ are extracted from a logit-

Table 1. Hyperparameters for the extraction of the logit-normally distributed prior perturbation parameters. Numerical entries without units are implicitly dimensionless. Note that the hyperprameters μ and σ are the mean and standard deviation, respectively, of the associated Gaussian distributions that the logit-transformed prior perturbation parameters follow.

Variable	Type	μ	σ	Lower bound	Upper bound
Precipitation	Multiplicative	-0.9	0.7	0.1	5
Temperature	Additive	0	0.5	$-8\,\mathrm{K}$	$8\,\mathrm{K}$
Longwave radiation	Additive	0	0.5	$-8\,\mathrm{W m^{-2}}$	$8\,\mathrm{W m^{-2}}$

normal distribution ~~rather than Gaussian , to restrict~~ whose prior hyperparameters μ and σ can be seen in Table 1. We choose this distribution over a log-normal or a Gaussian distribution as the logit-normal restricts the perturbation within defined ~~bounds (Aalstad et al., 2018; Guidicelli et al., 2023)~~upper and lower bounds (shown in Table 1), in contrast to other distributions which would have respectively only one or no bounds (Aitchison and Shen, 1980). The nature of the perturbation is multiplicative for the precipitation (in part to prevent non-physical negative values) and additive for the other variables.

The ensemble members representing the prior perturbation parameter distribution are updated with a ~~DES-MDA with four iterations, as the mapping from perturbation parameters to observations – the snow model~~ deterministic version of the ensemble smoother with multiple data assimilation scheme (DS-MDA; Emerick, 2018). Such iterative 'multiple data assimilation' algorithms mitigate the negative impact of a linear forward model assumption implicit in Kalman methods (Aalstad et al., 2018; Alonso-González et al. , Since the FSM2 ~~– is clearly~~ model is non-linear in the mapping between atmospheric forcing and snow states, we choose four iterations, in line with the aforementioned literature. We select 40 as an appropriate number of ensemble members in order to adequately represent the prior distributions while maintaining a reasonable computational cost.

3.5 Spatial propagation of information

The key to spatially propagate information from local observations in space — that is to other simulated cells — is in the construction of the prior covariance matrix with spatial dependence (Cressie, 2011). ~~Details on the practical implementation (and the theoretical background) of the system can be found in Alonso-González et al. (2023).~~ As the spatial distribution of snow depth is strongly governed by topography, and as the relative patterns are often repeated year after year, we ~~follow a concept introduced in experiment III of Alonso-González et al. (2023). We employ~~ adopt a spatio-temporal snow DA approach based on generalized non-dimensional distances (Alonso-González et al., 2023). Details on the practical implementation and the theoretical background of this method can be found in Alonso-González et al. (2023). The key innovation of this method was to define a generalized prior correlation function dependent on the proximity of cells in a multidimensional feature space. ~~We selected the following~~ Herein, we adapt the method to our novel ICESat-2 snow depth DA problem. After a careful assessment of various predictors, we select the following two features as the dimensions for the ~~generalized~~ feature space:

1. **CSMD**: the Climatology of Snow Melt-out Date is obtained by averaging the date (day-of-water-year) when the snow melted out in a selected pixel, extracted from Sentinel-2 fSCA time series provided by the Theia land data center

(Gascoin et al., 2019). A climatology consisting of five water years was used (2017-2021). This is a new and potentially globally-available feature that was not used in Alonso-González et al. (2023).

2. **TPI**: the topographical position index (Weiss, 2001) is computed as the difference of the cell's elevation compared with its ~~neighborhood~~ neighbourhood defined as a 24 m radius — selected based on the results of regression experiments targeting snow depth in the Izas experimental catchment (Revuelto et al., 2014). It represents the exposure of the terrain at the mentioned scale, so that cells with negative TPI are located in a concavity or a valley and positive TPI are cells in convex terrain such as a ridge relative to its surroundings~~;~~

3. ~~**Sx**: the maximum upwind slope parameter of Winstral et al. (2002) provides information about topographical sheltering of individual cells. This~~ . The index is computed ~~in the prevailing wind direction: northwest for the Izas study site. It corresponds to the maximum elevation gradient between the selected cell and all the cells upwind that lie within a maximum distance of 200 m ;~~

4. ~~**CSMD**: the Climatology of Snow Melt-out Date is obtained by averaging the date (day-of-water-year) when the snow melted out in a selected pixel, extracted from Sentinel-2 fSCA time series provided by the Theia land data center (Gascoin et al., 2019). Five water years were used (2017-2021).~~

   ~~The topographic features (points 2 and **??** of the list) are computed~~ with the high-resolution snow-off 2 m DEM to capture small-scale topographic effects. To match the simulation resolution of 20 m, we use the average of all high-resolution ~~values~~ TPIs within each model grid cell. ~~The terrain indexes, their size and direction were selected based on the results of regression experiments targeting snow depth in the Izas experimental catchment (Revuelto et al., 2014).~~

Figure 3 shows the spatial distribution of the aforementioned features at the modelled resolution (20 m) in the experimental catchment as well as ~~a scatter-plot showing~~ the relation between snow depth and the selected coordinates~~and snow depth~~.

As the dimensions of the feature space have different units, each is standardized and made non-dimensional by applying a z-score, so that ~~the~~ each set of coordinates has a ~~null mean and~~ zero sample mean and sample standard deviation equal to 1. Secondly, inspired by the concept of automatic relevance determination (Murphy, 2023), CSMD coordinates are increased (multiplied by ~~1.5~~3) to make their relative weight larger as this feature correlates the most with snow depth. ~~This effectively creates a space where we can measure~~ These operations effectively create an abstract bidimensional feature space that we use to measure the similarity between cells ~~. Covariance localization is~~ using generalized distances. Figure 4 depicts both the cells in the drone domain (panel a)) as well as the cells where ICESat-2 (panel b)) are available in the feature space, where the coordinates are not yet standardized. The subsequent step is localisation, i.e., the practice of limiting the effect of long-range spurious correlation ~~(Evensen et al., 2022). We employ the~~ (Sakov and Bertino, 2011). Both to define the prior correlations matrix and to perform localisation we employ a widely used damping operator in the form of the Gaspari and Cohn correlation function (GC; Gaspari and Cohn, 1999), which is defined by ~~a single hyper-parameter~~distance in feature space and a single hyperparameter: the correlation length scale ~~. In practice, we define the neighborhood of each cell as the set of cells that are located~~ that we set to $c = 1.5$. This operator truncates the observations located further than a radius of two times the correlation length, while including in the

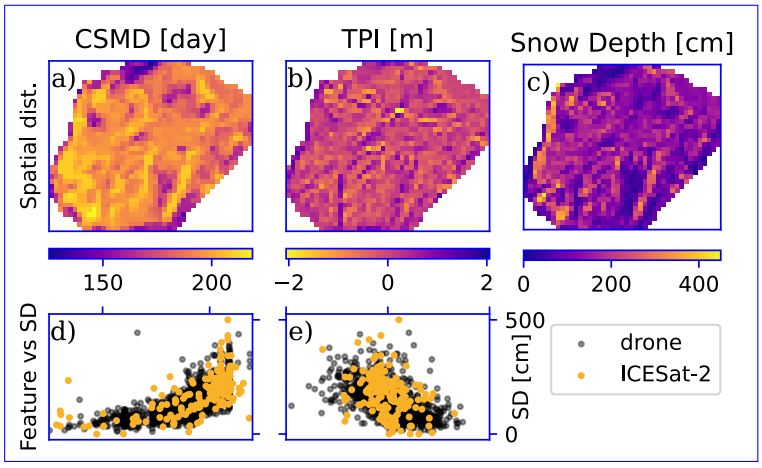

Figure 3. ~~Upper panels~~Panels a) and b): spatial distribution of the ~~three~~ feature dimensions ~~(~~Climatology of Snow Melt-out Date ~~:~~ (CSMD~~,~~) and Topographical Position index ~~:~~ (TPI~~, Winstral index: Sx~~) in the Izas experimental catchment~~, outlined with a green line~~. ~~Lower panels~~Panel c): spatial distribution of snow depth derived from the drone map of the 3rd of February. Panels d) and e): the dimension's relation with snow depth from the drone map acquired on the 3rd of February (black) and with ICESat-2 for the observed profiles on the 5th of February (yellow).

neighbourhood all observations located within this radius. Each cell in the neighbourhood is assigned a non-zero correlation value $\rho = \mathrm{GC}(d)$, computed as the GC function of the generalized distance in the feature space ~~within two times the correlation length scale. The closer the cells~~ (as defined in Alonso-González et al., 2023). This definition of neighbourhood also leads to a reduction in computational cost through domain localisation (Sakov and Bertino, 2011), whereby fewer observations are included when comparing state and observations in the update computation step of the assimilation algorithm.

In panel b) of Figure 4, we show one example for the definition of a neighbourhood and the assignment of the correlation value ρ to the included cells. We exemplify a situation where a cell in the catchment with drone data — depicted therein with a cross — has to be updated. The solid points in the scatterplot are included in the neighbourhood, hence information is transferred from those cells to the target cell (the cross) by means of the Kalman update (see step 11 in Algorithm 1 of Alonso-González et al., 2023), used to constrain the local forcing perturbation ensemble of the target grid cell. As cells closer in feature space to the ~~more similar the snowpack the system will constrain the cells to be, and the further away the less strict the system will enforce similarity. Cells outside each other's neighborhood will not influence the respective simulation. The correlation length scale was set to 1.5 after testing several values and making sure the~~ target cell should have a larger influence, their correlation ρ is larger as can be appreciated by looking at the size of the ~~resulting neighborhoods was acceptable.~~ scatter points. In simpler words, more information is transferred to the target cell from cells with large ρ compared to cells with low ρ, which are less similar in terms of TPI and CSMD. The shape of the neighbourhood is elliptic due to the differential weighting of the features, as we designed the CSMD to have a larger effect. Note that in Figure 4 the coordinates are shown with their (non-weighted) physical dimensions.

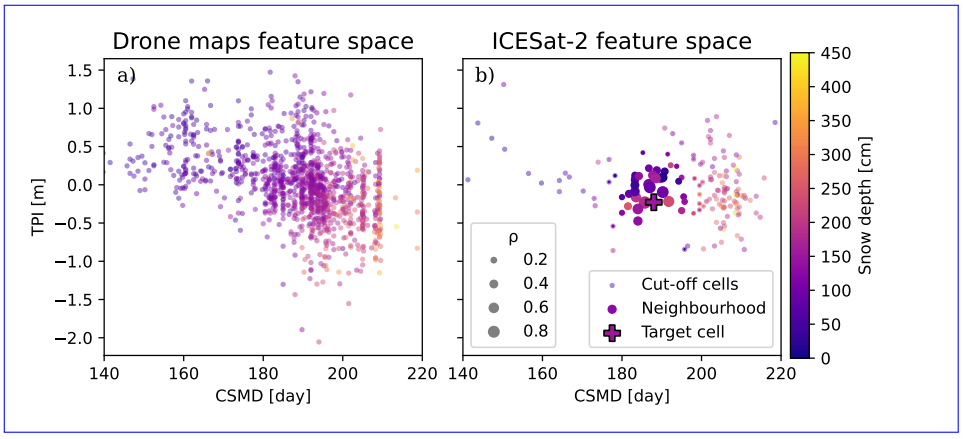

Figure 4. Panel a): scatterplot depicting the position in the feature space of the cells from the drone maps. This feature space — created with TPI and CSMD — is adopted to define the similarity between cells. The points are coloured according to the snow depth observed with the drone on the 5th of February. Panel b): ICESat-2 snow depth observations in the extended catchment, displayed in feature space, with snow depth-based colouring. The cross represents one cell from the drone domain where a snow depth of 150 cm was measured. The solid points are ICESat-2 data points included in the neighbourhood for this cell, with their size proportional to the correlation ρ.

3.6 Experiments

Three experiments are carried out where different observations are assimilated for the water year 2020. ~~Only information from the ICESat-2 snow depth retrievals – located outside the measured drone field as visible in Figure 1 – is spatially propagated, as fSCA maps are available for all cells in the model domain.~~ To allow a comparison where we change exclusively the assimilated variables, we use the same spatially correlated prior defined in Table 1 for all the ~~experiments~~simulations. The three experiments are designed ~~as follows~~to answer the scientific questions we presented in Section 1 and contextualize the value of ICESat-2 to the current standard snow DA practices:

- Snow **e**Cover experiment (**C**): Temporal assimilation of local fSCA retrievals from Sentinel-2. For each grid cell, all the available local (not ~~neighboring~~neighbouring) fSCA observations are assimilated. We use this as a baseline simulation.
- Snow **d**Depth experiment (**D**): Spatio-temporal assimilation of neighbouring snow depth retrievals from the ICESat-2 satellite altimeter. The ICESat-2 observations of the snowpack on the 5$^{\text{th}}$ of February are assimilated. Since the profiles are located outside the experimental catchment, the information is ~~spatially propagated using the dimensions we introduced above and depicted in Figure 3~~transferred using the methods explained in Section 3.5 and exemplified with Figure 4.
- **J**oint assimilation experiment (**J**): Spatio-temporal assimilation of local fSCA and ~~neighboring~~ neighbouring snow depth. All the observations of experiments (**C**) and (**D**) are assimilated, and their respective observation ~~error standard deviations~~

errors are the same as for these experiments. Note that only the sparse snow depth observations are spatially propagated, while the spatially-complete fSCA observations are not.

All the experiments are set up with the MuSA system (Alonso-González et al., 2023). The system was already ~~set up for~~ capable of jointly assimilating different observed variables at the same location, although joint assimilation ~~was tested only~~ has previously only been tested in a temporal DA setting jointly assimilating fSCA and skin temperature for a single cell experiment (Alonso-González et al., 2022). Here, for the first time, we perform a hyper-resolution spatio-temporal joint assimilation of snow depth and fSCA. Each cell is updated with the local observation (if any), located in the same cell position, together with the observations in its ~~neighborhood~~ neighbourhood defined as the set of cells located inside a search radius ~~in the afore-mentioned~~ ($d < 2c$) in the aforementioned feature space with dimensions TPI, ~~Sx~~ and CSMD (Section 3.5). This domain ~~localization~~ localisation step greatly reduces the computational effort, as it avoids searching through all the simulated cells. ~~For thisstudy~~Only information from the ICESat-2 snow depth retrievals — the blue lines in panel a) of Figure 1 located in the extended catchment — is spatially propagated, as fSCA maps are available for all cells in the model domain. To allow this, we updated the MuSA system ~~in order~~ which is now able to select a subset of the available observation types to spatially propagate ~~Alonso-González et al. (2024)~~information. This allowed us to exclude ~~fSCA observations (which are spatially complete)~~ spatially complete fSCA observations from the spatial propagation, which leads to ~~an additional~~ a marked decrease in the computational cost. In practice, that implied modifying the criteria for the selection of the observations to be included in the neighbourhood. The code for the updated version of MuSA developed in this study is available online (MuSA: v2.1, Alonso-González et al., 2024).

All the experiments were executed on a local server equipped with a 1 TB Memory and using about 40 processors. The computational time varied widely from 8 to 60 hours, depending mainly on the server load and on the number of observations in a given experiment, as that number influences the amount I/O operations necessary for each update.

3.7 Evaluation

We use the twelve drone-based snow depth maps of the Izas experimental catchment (Revuelto et al., 2021) as ground truth to independently evaluate the relative performance of the DA in the three experiments. ~~Their measurement error is typically one order of magnitude lower than the uncertainty in the snow-pack reconstruction, as we resample~~ A lower bound (assuming independent errors) on the measurement error can be estimated at about 1 cm as we have resampled the snow depth maps to the modelling resolution (from 1 m to 20 m) with the averaging operator. This is typically one order of magnitude lower than the uncertainty in the snowpack reconstruction. We aim at evaluating the sampled posterior distribution in terms of the resulting simulated snow depth goodness of fit to these independent observations. We evaluate our results in three ways: i) in time, ii) in space and in terms of snow depth distribution, and iii) for the entire spatio-temporal ensemble distribution. For i), all the ensemble member snow depth simulations are spatially averaged over the measured experimental catchment for each day in the simulated water year, and compared to the corresponding spatially averaged drone observations. ~~This~~ In this way, we can also evaluate how well the experiments perform in estimating the total snow volume in the catchment. For ii), we visually compare

the spatial distribution of the simulations against the drone-based map measured on the 11th of March, the closest acquisition to the seasonal SWE maximum. Since the result of the DA problem is a spatially correlated ensemble representing a statistical distribution, we show one single ensemble member simulation in order to appreciate the spatial structure embedded in the simulation. ~~We~~ To choose the representative member, we first select the simulation state on 11th of March as this is the closest drone acquisition to the peak-SWE. Then we spatially average the ~~ensemble members~~ ensembles and pick the ~~member whose average snow depth state is selected by the median operator for the 11th of March.~~ median member of those spatial averages. Note that the selection of one ensemble member introduces an objective but casual element, as different ensemble members exhibit different spatial patterns. In the supplement, the interested reader will find maps of the pixel-wise ensemble medians, showing a smoother spatial distribution.

To evaluate the probabilistic inference results of the three experiments (iii), we employ the Continuous Ranked Probability Score ~~(CRPS Hersbach, 2000)~~(CRPS; Hersbach, 2000), as this metric evaluates the performance of the inferred distribution represented by all the ensemble members (rather than a single point-estimate such as the median of the distribution) of each simulated cell in terms of snow depth compared to the observed reference. In particular, the CRPS quantifies both the precision (certainty or confidence) and the accuracy (ability to match the observations) of the ensemble as a whole. This is a strictly positive score, where a perfect match between the compared distributions would result in a score of 0, while the larger the CRPS score the worse the result. We compute the CRPS metric for each experiment for all the available drone-based snow depth maps. To evaluate the spatial distribution of the errors we present two maps per experiment with the average score for accumulation and melting season. To evaluate how the experiment's performance varies in time we average the experiment's score throughout the measured catchment~~; while to evaluate the spatial distribution of the errors we present two maps per experiment with the average score for accumulation and melting season~~, thus receiving a single score for each drone-based snow depth map.

4 Results

~~Figure??~~

4.1 Comparison with drone maps

Figure 5 summarizes the results for ~~Experiment~~ all the experiments and shows a comparison with the drone-based validation snow maps. We first focus on experiment **(C)** assimilating fSCA as a benchmark representing the current best-practice snow DA. The time series in panel a) shows the prior and posterior ensemble members averaged over the Izas ~~catchment~~experimental catchment for experiment **(C)**. The ensemble of simulations is precise (low ensemble spread) and accurate (good match to validation data) towards the end of the snow season. However, the ensemble spread is ~~larger~~ wider in the snow accumulation months and the ensemble median overestimates the catchment-average snow depth. From the map in panel ~~b~~d), it is discernible that the simulation correctly reconstructs the ~~observed~~ drone-observed snow depth patterns in a relative sense (~~panel c~~comparison with panel g)): the areas with larger-than-average snow depth are correctly recognized, as well as the ones with lower-than-

average snow depth, despite the absolute values not being correct. ~~The histograms in panels d) and e) highlight~~ Comparing the histogram in panel h) with the reference in l) highlights that the simulation average overestimates snow depth at peak SWE by ~~47%~~51%. In the simulation, only ~~10%~~ 24% of the simulated cells are inferred to have less than 150 cm of snow, while in the drone validation ~~about 50%~~ 46% of them are measured in this part of the range.

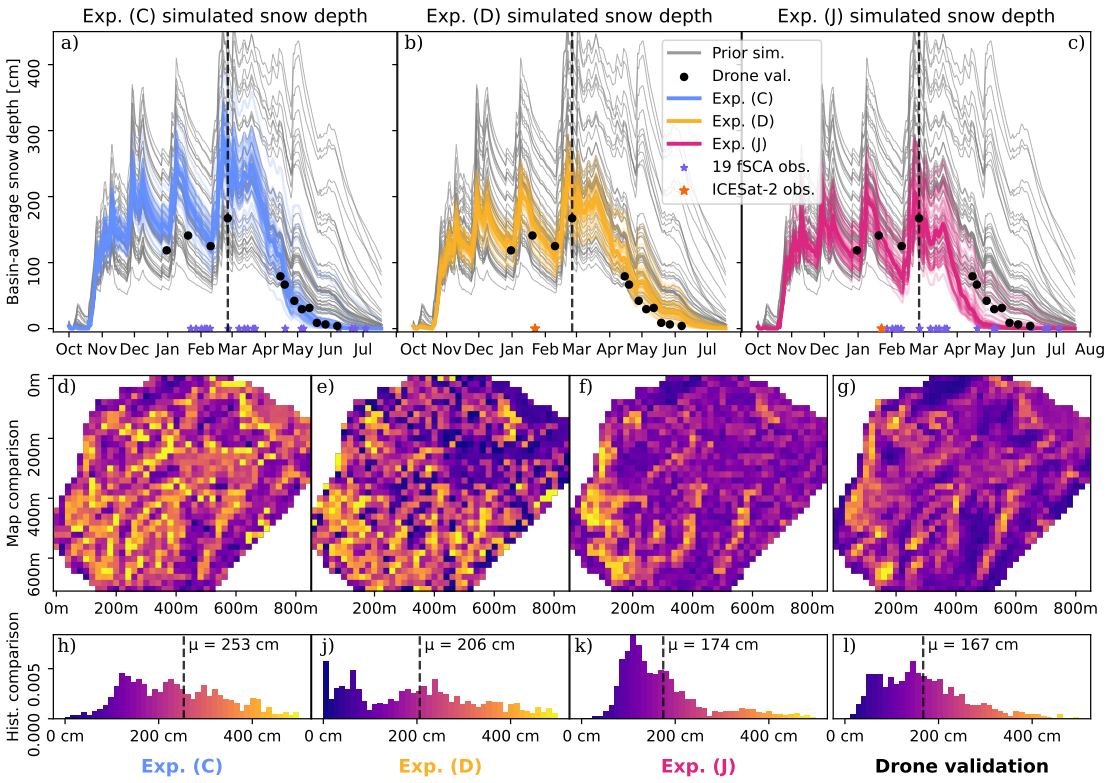

Figure 5. ~~Results from experiment (C). Panel~~ Panels a)~~:~~, b) and c) show the prior (gray) and posterior ensemble simulations (~~blue~~coloured) as catchment-average snow depths over the whole water year for experiment (C), (D), (J), respectively. The black points are the drone-based snow map averages serving as validation data. ~~Panel b~~The blue and orange stars show the timing of the fSCA and snow depth observations that are assimilated in each experiment. Panels d), e) and f): simulated snow ~~depth map~~ depths are shown as maps for ~~11th March 2020~~ the 11.03.2020 (date shown with a vertical dashed line in ~~panel a~~panels above) ~~: the median~~ for a representative ensemble ~~with respect~~ member that is nearest to the ~~catchment-average peak SWE is selected~~ensemble median catchment average snow depth for experiment (C), ~~color map explained in panels d)/e)~~(D), (J), respectively. Panel ~~e~~g): the corresponding 11.03.2020 drone-based snow depth map at the model's spatial resolution (20 m). ~~Panels d~~Bottom panels h), j), k) and ~~e~~l): snow depth histograms ~~of~~ corresponding to the maps in the panels above, using the same colour scale. μ indicates the experimental catchment-average snow depth.

~~Figure ?? summarizes the results of Experiment~~ For experiment **(D)**~~. Panel a)~~ , panel b) in Figure 5 shows that the assimilation of the sparse snow depth ~~profile data~~ profiles from a single date substantially narrows and improves the posterior distribution compared to the prior. Notably, during the winter season, the accuracy of the posterior is improved~~(~~: the ensemble median and the black validation points are closer ~~) compared to the first experiment~~ compared to **(C)**~~, and also~~ . Also the precision is improved~~(~~: the ensemble spread is ~~lower)~~narrower compared to **(C)**. Towards the other end of the season, both accuracy and precision ~~largely~~ slightly decrease compared to the first experiment, despite still improving ~~from~~ that of the prior simulation. Panel ~~b~~e) shows that the posterior simulation is able to match ~~only some~~ many of the spatial patterns that the drone map shows in panel ~~c~~g), such as in the ~~northeast~~ high-elevation accumulations under the ridge on the western side of the catchment. Also, ~~the valley~~ patterns in the north-east part of the catchment are well reproduced. However, the valleys affected by wind drift and the corresponding wind-blown ridges in the ~~southeast~~ south-east portion of the catchment are ~~partially recognizable. However, patterns in the high-elevation part of the catchment (west side) are not~~ only partially reproduced. The histograms in panels ~~d) and e) show that the simulated snow depths at peak-SWE have similar ranges, despite the mean being~~ j) and l) show the snow depth distributions' comparison. The measured catchment-average snow depth is only over-estimated by ~~33%~~23%. In contrast with ~~Experiment~~ experiment **(C)**, low snow depth areas are ~~represented as in~~ well represented and better fit the histogram of the measured map. However ~~, also in this experiment, 32%~~ 25% are simulated with very high snow depth (> 300 cm)~~despite the fact that~~ , while in the drone validation only 9% of the cells are measured with such snow depths. ~~Results from Experiments (D), presented in the same way as in Figure ??.~~

~~Figure ?? summarizes the results of Experiment~~ For experiment **(J)** ~~, where we use all the available observations. The~~ the time series in ~~panel a~~Figure 5c) shows that the inference ~~produced a precise result~~ produces a precise posterior ensemble (low spread) throughout the water year. In the accumulation period and up to the peak-SWE, the catchment-average snow depth is accurately reconstructed. Notably, for the ~~11th of March~~ 11.03.2020 acquisition the measured and simulated ~~average snow depth differs only 2~~ catchment-average snow depth differ by only 7 cm. However, there is a negative bias in spring, which we do not see in experiments **(C)** (unbiased) or **(D)** (positive bias). The comparison of simulated and measured maps in panels ~~b) and c~~f) and g) shows that the posterior simulation is able to match ~~only some~~ well most of the relative spatial patterns. For example, the accumulation below the ridge on the western border of the catchment is simulated with higher-than-average snow depth~~only in its center and south portion of the feature, and not in its northern part. Only some of~~ . Moreover, the absolute snow depth values are very similar in the western edge of the catchment. The cells located in the deep valleys and ~~depressions~~ on the corresponding ridgelines in the south~~and some of the flat~~ , and most of the flatter accumulation areas in the north-east have the correct relative snow depth patterns. The nearly snow-free ~~south-facing slope in the north is~~ south-east-facing slope located in the northern part of the catchment is also correctly simulated. In terms of distribution, the histograms in panels ~~d) and e~~k) and l) show clearly that the simulation reproduces ~~the mean and the frequencies of the~~ not only the mean but also the general shape of the distribution, including the tails. Low snow depths (<150 cm), simulated for ~~47%~~ 51% of the cells, ~~match~~ are only slightly more than the number of cells measured ~~with such~~ in this range (46%); as well as for very high snow accumulation values (> 300 cm): ~~5%~~ 12% of the cells are simulated and 9% measured.

~~Results from Experiments (J), presented in the same way as in Figure ??. Scoring the three experiments' whole ensemble~~

4.2 Ensemble quantitative evaluation

Scoring the entire spatio-temporal ensemble of the three experiments allows for a more quantitative comparison of the inference performance, as panels ~~b) and d)~~d)-f) of the previous figures only show one representative ensemble member's simulation. Figure 6 shows the CRPS computed ~~against the drone acquisitions for~~ using the drone-based snow depth retrievals as a reference for all the available snow depth maps, ~~averaged in space (left panel~~ focusing on the seasonal (panel a)), and ~~time (right panels)~~spatial (panels d)-i)) distribution. Focusing on the ~~right panel, the right panel shows that for all experiments~~time evolution of catchment-average CRPS, the largest ~~errors (largest CRPS)~~ CRPS values are found during the accumulation season, while the errors decrease in the melting season, together with a reduction in the absolute values of snow depth.

~~Averaging the CRPSs over~~ In the accumulation season, the average of the CRPS values show that the ICESat-2 assimilation **(D)** ~~shows~~ has a similar performance (~~48 ± 12~~44 ± 18 cm) to the fSCA assimilation **(C)**, whose CRPS is ~~45 ± 18~~44 ± 10 cm. Both these experiments are substantially worse than the joint assimilation **(J)**, scoring 35 ± 7 cm. Thus, adding the ICESat-2 snow depth profiles to fSCA in the set of assimilated observations ~~lowers (i.e., improves) the~~ improves the overall error score by ~~22%.~~

19% for the accumulation season. In the melting season, ~~the fSCA~~ experiment **(D)** scores worst but still improves the prior CRPS by 25%. However, experiment **(C)** ~~has very similar results to~~ simulations are not outperformed by the joint assimilation **(J)**~~, even being the best for two snow depth maps. The experiment with the ICESat-2 assimilation (D) shows the worst performance, with the highest CRPS~~: their results are respectively 14 ± 11 cm and 16 ± 11 cm.

~~Looking at~~ Panels d) to i) in Figure 6 show the spatial distribution of the ~~errors (right panels in Figure 6), experiment (C) shows that~~ CRPS. For experiment **(C)** (panels d) and e)), the CRPS is generally ~~low for both accumulation and melt~~ lower for the melting season. Despite ~~slightly~~ the higher values in the accumulation season, only ~~2% of the~~ 7% of the evaluated cells have a ~~very~~ high CRPS (> 100 cm). ~~For the accumulation season, the~~ The location where large errors cluster are the windblown ridges in the ~~center-south~~ south-east and on the west of the experimental catchment where a close look at panels ~~b) and c~~d) and g) of Figure~~?? shows that low snow depth is~~ 5 shows that medium snow depths are inferred instead of ~~snow absence~~low snow depth or no snow. Experiment **(D)** performs better during the accumulation season: ~~10%~~ only 1% of the cells have a very high CRPS (> ~~100~~150 cm) in that season, compared to 4% in the melting season~~, and these~~. Such large errors are clustered in the south-west of the catchment. The low elevations on the eastern side of the catchment have ~~lower~~ better scores (compared to higher elevations) in both the accumulation and melt season. Experiment **(J)** shows a very similar error pattern distribution in both the accumulation and ablation season (in contrast with the other experiments). Less than 4% of the simulated cells have a ~~very~~ high CRPS (>100 cm) for both the accumulation and melting season. ~~An inspection of~~ By comparing panels h) and i) with the absolute snow depths (panels b) and c)~~of Figure ??, shows that very high~~), it is clear that the largest errors are located where very high snow ~~depth~~ accumulation are located (> 300 cm)~~are measured:~~ , such as the accumulation areas under

the ridge on the ~~west~~ western side of the ~~border~~ catchment and some of the areas in the valleys in the south of the catchment. Supplemental Figure 1 clarifies that the snow depth in the mentioned areas is underestimated.

Focusing on the use of the novel ICESat-2 snow depth retrievals, the experiments show that adding these observations for constraining the inference of the seasonal snow evolution has ~~mostly~~ a positive impact on the simulation results. Comparing the CRPS over the experimental catchment of experiment **(D)** to the prior (no observation assimilated) shows an improvement of ~~7.5 cm (3.5 cm) for the accumulation (melting) phase of the season~~20 cm for the accumulation season and 10 cm for the melting. When adding snow depth to fSCA in the set of ~~observations – hence comparing experiment~~ assimilated observations – hence comparing experiments **(J)** ~~to~~ and **(C)** ~~–~~ – the improvement is clear with a 10 cm ~~clear~~ CRPS reduction for the accumulation part of the season. However, there is a slight ~~decrease of 4~~increase of 2 cm in the ~~score~~ CRPS for the melting season, despite the left panel of Figure 6 showing experiment **(J)**'s score being worse only for two of the drone acquisitions.

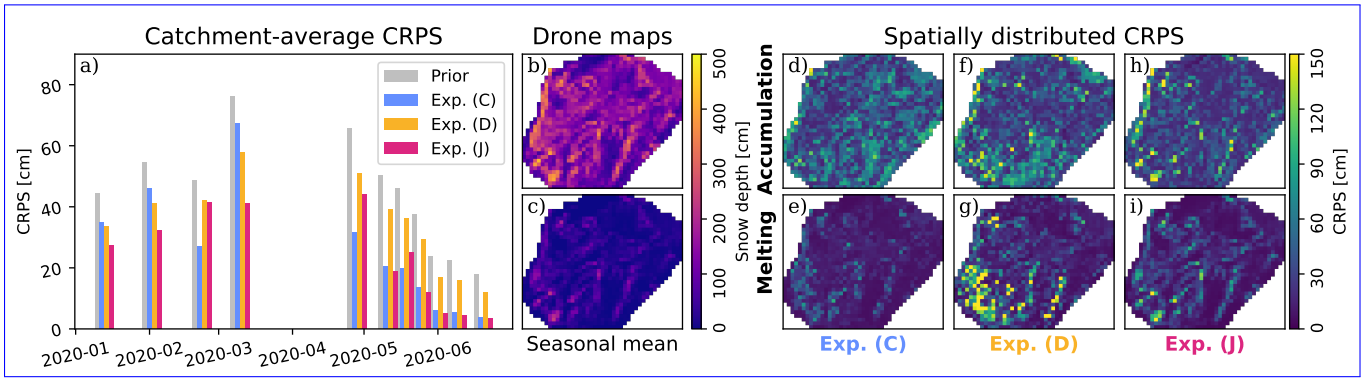

Figure 6. The performance of the three assimilation experiments is presented through the Continuous Ranked Probability Score (CRPS), where a perfect match of the compared distribution would score 0, while the larger the score the worse the result. In ~~the left~~ panel a), the temporal evolution of the experiments' results is shown by spatially averaging throughout the experimental catchment. ~~In~~ Panels b) and c), show the ~~upper right~~ drone maps averaged for the accumulation (b)) and melting (c)) season. In panels d) – i), the spatial distribution of the errors ~~is shown~~ obtained by averaging the CRPS computed for the validation maps acquired during the accumulation season (up to and including the 11^{th} of March) ~~; while~~ is shown in the ~~lower right~~ upper panels~~show the same metric averaged over the validation maps acquired in~~ , and for the melting season in the lower panels.

5 Discussion

In this paper, three DA experiments are carried out. ~~For all of them, the assimilation algorithm as well as the prior~~ All the experiments employ the same assimilation algorithm (DES-MDA), and the spatially correlated prior distributions (Table 1) are the same. ~~We use observations of different~~ In each experiment we assimilate remotely sensed observations using a different mix of snow variables, and all ~~of them~~ experiments successfully improve the prior simulation by constraining the simulations with information coming from the ~~observations, obtaining~~ retrievals. However, we obtain very different results depending on

the assimilated ~~variable~~variables. The observations in the three experiments are fSCA retrieved from Sentinel-2 **(C)**~~as well as~~ , the novel snow depth retrievals from the laser altimeter ICESat-2 **(D)**, and finally a joint assimilation experiment using both these retrievals **(J)**. To the best of our knowledge, this is the first example of ~~high-resolution (almost hyper)~~ hyper-resolution multivariate snow DA. Moreover, this is the first study not only to assimilate the ICESat-2 snow depth observations but also to show how these observations acquired along profiles located outside the area of interest can be harnessed directly through spatial propagation of information using spatio-temporal DA. This system successfully updates the snowpack states in the Izas experimental catchment ~~with observations~~ both with local Sentinel-2 fSCA observations as well as ICESat-2 observations from profiles located in its proximity~~, by exploiting~~ . The experiments with the ICESat-2 profiles exploit topographical and snow climatological ~~similarities~~ indexes to define a similarity measure used in the correlation function.

~~In general, DA consists in~~ DA consists of combining uncertain information coming from models and observations. In our case, the unconstrained model (i.e., the prior) shows an overestimation of the snow depth, especially during the melting season (see gray simulations in panel a) of ~~Figures ??, ??, ??~~Figure 5). The errors in the prior can be attributed to two different sources: the forcing data and the model itself (Raleigh et al., 2015). ~~Errors~~ Considerable errors in the forcing data are ~~to be expected~~ expected to dominate and thus, by design, the forcing formulation of DA allows for the correction of such errors (Evensen et al., 2022). We emphasize that when carrying out high-resolution, spatially-distributed modelling ~~and using coarse reanalysis such as ERA5 as forcing information (despite the topographic downscaling),~~ one has to expect large ~~biases~~ errors despite carefully chosen climatic forcing and topographic downscaling, given that the high spatial heterogeneity of the drivers of snow accumulation and redistribution processes cannot be explicitly represented in the forward modelling. On the modelling side, one can expect that some of the ~~parameterized~~ parametrised processes (e.g. albedo decay or precipitation partitioning) in FSM2 suffer from errors. Here, some of the snow processes can be substantially different and not well captured by FSM2's ~~parametrizations~~parametrisations, which were originally implemented ~~on~~ for a field site in the European Alps (Essery, 2015). With the ~~given prior~~ prior employed in our study ~~, there may be a greater~~ (Table 1), there is just as great a need for information about the melting season ~~rather than~~ as well as about the accumulation season. The ~~correction of the main errors in the prior is accomplished through the assimilation of fSCA, and the DA formulation is able to compensate for forcing errors~~fSCA observations contains a cumulative signal of both accumulation and melting processes. On the other hand, ICESat-2 snow depth observations from the 5th of February mostly contain accumulation processes-related information. This could ~~also~~ explain the better performance for the melting season of experiments **(C)** and **(J)** compared to **(D)** ~~,~~ in which only observations from the accumulation season were assimilated. ~~However, if snow depth profiles were acquired by ICESat-2 later in the season, the performance during the melting phase might also improve when assimilating only snow depth information.~~

5.1 Experiments' results

In experiment **(C)**, we simulate the snowpack and assimilated fSCA to create a baseline. It has been shown that, despite fSCA exhibiting a lower instantaneous correlation with early season snow depth in a deep snowpack compared to the end of the season with a melting snowpack (Girotto et al., 2020), the assimilation of fSCA allows for an accurate reconstruction of peak

SWE (Girotto et al., 2014). Indeed, experiment **(C)** shows high accuracy and precision towards the end of the season. As the experiments adopt a smoother approach, such information is also propagated backward in time and the posterior simulation offers ~~an~~ a relatively accurate reconstruction for the peak-SWE: the validation is clearly close to the median ensemble spread, but the reconstruction for this part of the season is less precise compared to experiment **(D)**, and both less precise and accurate ~~for~~ compared to experiment **(J)**. In terms of relative spatial patterns, experiment ~~(C~~**(J)** shows the best visual agreement with the validation of drone-based maps for peak-SWE, despite overestimating the absolute values ~~.~~ in the melting season (Figure 5).

~~In~~ Research question a) asked about the value of ICESat-2 observations for inferring catchment-scale snow depth and its distribution. We can answer this question with the results of experiment **(D)**, where we assimilate snow depth observations from ICESat-2 on two profiles outside the catchment, spatially propagating this information to the experimental catchment. The snow depth was acquired on the 5th of February, about a month before peak SWE. As ~~Figure **??** panel~~ a) in Figure 5 shows, this information leads to a more precise reconstruction of the catchment-average peak-SWE compared to experiment **(C)**, which adopted state-of-the-art methods for seasonal snow reconstruction with DA. This demonstrates that the ~~spatio-temporal DA is successful, as the information propagated from observations outside the Izas catchment carries more or at least a similar amount of information compared to the temporal-only information propagation that happens in experiment **(C)**. Compared to experiment **(C)**, this simulation has a better agreement with the observed snow depth histogram distribution, as the range of the snow depth histograms has a better match (panels d) and e), Figure **??**)~~spatial transfer of information method successfully relates snow depth and the features when averaging over the whole basin. However, not all the relative spatial patterns of the simulation ~~only partially~~ match those of the validation maps (panels ~~b) and c~~e) and g), Figure ~~**??**). Since the~~ 5). The observations we use in this experiment are not direct measurements in the ~~catchment, this result is in the end not surprising: the similarity measure we define is only partially able to propagate snow depth information properly. Nevertheless, single pixels~~ experimental catchment, so the fact that the spatial distribution of the simulation is not entirely reproduced is not surprising. While the entire area experiences similar snow conditions there are local differences which can be only partially captured with a low dimensional space, as TPI and CSMD do not fully characterize the snow depth distribution. For example, single cells with extreme values located in the ~~basin might not be similar (~~experimental catchment might be similar in terms of ~~topography and meltout date) to the ones which are~~ TPI and CSMD to cells observed by ICESat-2 ~~.~~ with very high snow depth, but also to some medium snow depth cells, and thus will not receive an optimal update.

Towards the end of the season, both the precision and accuracy of the ~~simulations~~ experiment **(D)** are degraded when compared to experiment **(C)**, and the CRPS score is almost the same as for the prior for this point of the season. Here, the timing of the single assimilated observation is an important factor: most of the information of early-season snow depth observations is related to accumulation processes (precipitation perturbation) rather than melting processes, ~~as found by Guidicelli et al. (2023)~~in line with previous findings by Guidicelli et al. (2024). The very late melt-out date we obtain with this experiment can mostly be attributed to the fact that also the prior simulates the snowpack with such a late melt-out date. Moreover, Margulis et al. (2019) showed that assimilating snow depth observations later in the season, when ~~more consistent~~ some ablation processes have taken

place, ~~could improve the melting season estimates . If the previous suggested improvements would improve the results of the experiments for this setting, this could be used, in principle, in a forecasting system, as~~ improves the melt season estimates compared to assimilating earlier snow depth observations~~have instantaneous value; while fSCA are more useful in a reanalysis setting. There would be the need to speed-up the~~ .

To research question b), asking whether ICESat-2 ~~processing for the low-level product, as it is now usually three months~~snow depths or fSCA observations perform better for snow data assimilation, we cannot give a clear answer for the entirety of the snow season. We showed that assimilating ICESat-2 provides a better estimate of average snow depths during accumulation season, which is very valuable for water managers, while assimilating fSCA results in a better relative spatial distribution and melting season reconstruction.

In experiment **(J)**, we assimilate all the observations used for the previous two experiments. ~~Note that these~~ These two sets of observations complement each other's coverage in time and space. ICESat-2 observations occur in February, when a deep ~~snow pack~~ snowpack causes fSCA relation with snow depth to saturate, and hence provides little or no ~~information~~instantaneous information (Margulis et al., 2015). However, fSCA is spatially complete and ~~thus nicely~~ temporally denser which thus complements the sparse, but more direct, snow depth observations of ICESat-2, which are located along two profiles outside the experimental catchment. We show that ~~executing~~ the joint assimilation ~~leads to higher precision (smaller ensemble spread) throughout the water year. This simulation~~ experiment **(J)** performs better compared to the experiments **(C)** and **(D)** in terms of the CRPS in the accumulation season, showing a better representation of the spatial distribution of snow depth. For the melting season, experiment **(J)** has a comparable score to experiment **(C)**, in which the melting season seems already accurately modelled. The time series in ~~panel aof Figures **??** and **??**~~ panels a) and c) of Figure 5 show that in terms of catchment average snow depth, experiment **(C)** performs slightly better. However, Figure 6 explains the spatial distribution of the errors and it ~~'s~~ is clear that for most of the cells both the simulations have a similar score to **(C)**. It is only in the large snow drifts that experiment **(J)** is not able to simulate large snow ~~depth.~~

~~For both experiments **(D)** and~~ depths. With experiment **(J)**~~, improvements in~~ 's results in mind, it is possible to answer research question c) on whether DA is able to use information from both snow depth and fSCA combined. We found that the DES-MDA algorithm is able to jointly leverage observations of different snow variables in the same experiment, with better results compared to the independent assimilation of either observation.

5.2 Outlook and future developements

Simulating the snowpack with a physically-based model (i.e. FSM2) allows to infer unobserved but socially relevant variables (such as SWE) or fluxes (such as snowmelt). The assimilation of fSCA has already been shown to accurately reconstruct the peak SWE (Girotto et al., 2014), and here we show that adding snow depth observations from ICESat-2 in the pool of assimilated variables improves the peak basin-average snow depth. Despite not being able to directly demonstrate it because of missing large-scale SWE validation measurements for Izas, we can postulate (based on results on SWE reconstruction experiments using complete snow depth maps Margulis et al., 2019; Ma et al., 2023)

that ICESat-2-retrieved snow depth can improve the total water resources estimation for basins in complex terrain. The satellite ICESat-2 will potentially collect data until the mid-2030s, and it is necessary to further refine the methods to inform seasonal snow models and exploit this dataset for mountain water resources management. Our results show that this dataset has the potential to become a functional tool for water managers to estimate the maximum seasonal snow accumulation. However, especially within an operational snow hydrological forecasting context (e.g. Mott et al., 2023), there is a clear need to reduce the processing time of the ~~spatial distribution are expected in the case of~~ ICESat-2 ~~profiles observing snow in a terrainwith more similar characteristics. Moreover, herein the process of~~ data products which currently takes months, and ICESat-2 snow depth measurements from the accumulation season may thus not be available before the end of the melt season.

In this paper, the spatial information propagation depends on a feature space defined only with ~~few characteristics~~two dimensions, that has not been tuned systematically to improve the results~~, but that was based on previous studies (Revuelto et al., 2014). Optimizing or tuning~~ . Optimizing the selection of the ~~dimensions in the feature space~~ feature space dimensions as well as tuning their relative weight and the correlation length scale ~~, despite being computationally expensive ,~~ (the hyperparameters of the spatial dependence) would be computationally expensive but might lead to better results, as ~~Experiments~~ experiments II and III in Alonso-González et al. (2023) show. For example, we found that the exclusion of the Winstral index (Winstral et al., 2002) from the definition of the feature space improved the inferred spatial patterns for all the experiments. We envision future work in which the ~~an~~ optimization process could be fruitful for data-rich and vast basins such as the repeated measurements in the western ~~U.S by the ASO (Painter et al., 2016)~~US by the Airborne Snow Observatory (ASO, Painter et al., 2016). Despite the optimization or inference of the ~~aforementioned hyper-parameters~~ hyperparameters being foreseen to be very expensive in terms of computational resources, we acknowledge that it could be performed off-line for a single season, exploiting the repeated patterns of the seasonal snow evolution ~~(Revuelto et al., 2014)~~(Sturm and Wagner, 2010; Revuelto et al., 2014). In terms of methods, we envision the use of hierarchical data assimilation ~~(Katzfuss et al., 2020)~~ for the statistically-optimal inference of the aforementioned ~~hyper-parameters~~hyperparameters (Katzfuss et al., 2020). In simple terms, this would increase the complexity of the DA system by defining a hyperprior on ~~these~~ the hyperparameters that govern the spatial propagation of information. The assimilation would lead to the optimal inference of the seasonal snow evolution for the current water year after having inferred the spatial hyperparameters. Moreover, the learned hyperparameters could conceivably be transferred to other water years without the need for a second round of hierarchical inference. To alleviate computational bottlenecks, hyperparameter inference could be achieved using simpler snow models that are ~~now~~ available in MuSA. Moreover, splitting the spatial propagation problem from the temporal DA — provided that several year's snow depth maps are available for training and testing independently the hyperparameters — could also alleviate the computational cost by avoiding the need to run a snow model is another option. A feature space including various terrain parameters and geographical, elevation, land-use-related coordinates to define the similarity in terms of snow depth or SWE has been adopted in several recent statistical snow modelling studies based on satellite or airborne data covering larger domains (Guidicelli et al., 2024; Hultstrand et al., 2022; Liu et al., 2024) .

A final note is warranted concerning the ~~very high-resolution (almost hyper)~~ spatial resolution of the experiments (20 m). We selected this cell size to test the ability to ~~measure~~ retrieve snow depth at a hill-slope scale with the ~~photon-counting~~ photon-counting technology that ICESat-2's ATLAS is equipped with. ~~The results are not clear yet in this sense, as validation is not available for the measured track. However, some of the larger spatial patterns are reproduced, even with this very high resolution, in experiment (D). Lower resolutions are also~~ In experiment **(D)**, we partially reproduced the spatial patterns observed through the drone maps. Lowering the spatial resolutions of the simulation would still be adequate for water resources mapping, and would make the assimilation exercise much easier: many of the accumulation features are averaged out already at 100 m resolution. Assimilation of snow depth data at lower spatial resolution has been found to give better results in previous synthetic ICESat-2 observations propagation with neural networks ~~(Guidicelli et al., 2023)~~(Guidicelli et al., 2024).

~~Simulating the snowpack with a physical-based model as we do with FSM2 provides inference on unobserved but socially relevant variables (such as SWE) or fluxes (such as the snowmelt flux). The assimilation of fSCA has already been shown to accurately reconstruct the peak SWE (Girotto et al., 2014), and here we show that adding snow depth observations from ICESat-2 in the pool of assimilated snow data improves the peak basin-average snow depth. Despite not being able to demonstrate it because of missing large-scale SWE validation measurements, we can postulate (based on results on SWE reconstruction exp that ICESat-2-retrieved snow depth can improve the total water resources estimation for basins in complex terrain.~~

6 Conclusions~~and outlook~~

In this study, we present a novel approach where ICESat-2 snow depth observations along profiles ~~were used~~ are assimilated for the first time to ~~update~~ constrain high-resolution snowpack simulations ~~. We exploited~~ across an entire unobserved catchment. We worked with the data assimilation system MuSA, which ~~has~~ had capabilities to propagate information in space and time developed in Alonso-González et al. (2023)~~, to bridge the observations'sparsity. We perform~~ . However, we modified it to exploit ICESat-2's sparse observations and jointly assimilate a spatially-complete variable. We performed a set of three experiments where ~~fSCA,~~ only Sentinel-2 fSCA, only ICESat-2 snow depth, and both these observations together, were assimilated while keeping the assimilation algorithm and the prior information the same, in order to evaluate the potential of the satellite laser altimeter ICESat-2 for updating seasonal snow models.

We find that including two snow depth profiles in the set of assimilated observations improves the snowpack simulation in terms of average snow depth, especially during the accumulation phase of the season ~~–~~ even though the snow depth profiles were located completely outside the experimental catchment of Izas (55 ha), where we validate the experiments using drone-based snow depth maps. Results show that the relative spatial patterns can ~~only~~ partially match the validation drone maps when assimilating exclusively the snow depth profiles~~; as~~ , since the snow depth pattern is very sensitive to the design of the spatially correlated prior ~~covariance which governs~~ which governs the spatial propagation of information. Nevertheless, the joint assimilation of Sentinel-2 fSCA and snow depth from ICESat-2 bridged such limitations and performed best in terms of average snow depth as well as capturing the spatial distribution.

These findings indicate that the ~~satellite~~ ICESat-2 satellite can be exploited to improve the current state-of-the-art ~~reanalysis~~ snow reanalyses generated by assimilating fSCA. Notably, the proposed workflow exploits globally-available datasets. Provided a high-resolution DEM ~~–~~ which the geosciences community generally advocates the need for ~~–~~ ICESat-2's surface elevation measurements can be used to observe snow depth along profiles in inaccessible regions where snow amounts are still very hard to quantify. As multivariate DA techniques mature, the snow community can begin to exploit the plethora of snow observations (e.g. snow depth, fSCA, land surface temperature etc.) to constrain ~~the~~ snow models and to shed light on the still unsolved snow hydrology problem of inferring the spatial and temporal distribution of SWE.

Code and data availability. The MuSA code used for the experiments is version 2.1 and can be found at https://zenodo.org/records/11147258. The ICESat-2 data was downloaded with SlideRule (Shean et al., 2023), the Sentinel-2 data with GEE (Gorelick et al., 2017). The complete input data for the experiments in the Izas basin can be found in this data repository: https://zenodo.org/records/13860511. Validation maps are available at https://doi.org/10.5281/zenodo.7248635 (see the Obs folder).

Author contributions. Conceptualization was by MM, with key contributions from DT and KA. Data curation: ICESat-2 data was by MM and DT, Sentinel-2 data and ERA5 downscaling was by KA. Formal analysis was by MM, KA, DT, and EAG. Funding acquisition was by DT. Investigation was by MM, KA, DT, SW, and EAG. Methodology was developed by EAG and KA, with key contributions from MM. Project administration was by DT. Software was by EAG, MM and KA. Supervision was by DT, KA, EAG and SW. Validation was by MM. Visualization was by MM and DT. Writing — original draft preparation was led by MM with key contributions by all co-authors.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Acknowledgements. Marco Mazzolini and Désirée Treichler acknowledge funding from the Research Council of Norway (SNOWDEPTH project, contract 325519), DT additionally from the ESA project Glaciers_cci+ (#4000127593/19/I-NB).
Kristoffer Aalstad acknowledges funding from the Research Council of Norway (Spot-On project, contract 301552), the ERC-2022-ADG under grant agreement No 01096057 GLACMASS, and an ESA CCI Research Fellowship (PATCHES project).
Esteban Alonso-González acknowledges funding from an ESA CCI Research Fellowship (SnowHotspots project)
We are grateful to NASA and USGS for the free provision of the ICESat-2 data, to Copernicus (and ECMWF) for the free provision of the Sentinel-2 (and ERA5) data. We are also grateful to all the developers of free software, mostly based on the Python programming language (Python Software Foundation, https://www.python.org/), used in this research such as: MuSA (Alonso-González et al., 2024) , xdem (Dehecq et al., 2021), SlideRule (Shean et al., 2023), xarray, etc. ~~.~~
~~Value-added~~ We acknowledge the use of value-added data processed by CNES for the Theia data centre www.theia-land.~~frusing~~ fr using Copernicus products. The processing uses algorithms developed by Theia's Scientific Expertise Centres.

References

Aalstad, K., Westermann, S., Schuler, T., Boike, J., and Bertino, L.: Ensemble-based assimilation of fractional snow-covered area satellite retrievals to estimate the snow distribution at Arctic sites, The Cryosphere, 12, 247–270, https://doi.org/10.5194/tc-12-247-2018, 2018.

Aalstad, K., Westermann, S., and Bertino, L.: Evaluating satellite retrieved fractional snow-covered area at a high-Arctic site using terrestrial photography, Remote Sensing of Environment, 239, 111 618, https://doi.org/10.1016/j.rse.2019.111618, 2020.

Aitchison, J. and Shen, S. M.: Logistic-Normal Distributions: Some Properties and Uses, Biometrika, 67, 261–272, https://doi.org/10.2307/2335470, publisher: [Oxford University Press, Biometrika Trust], 1980.

Alonso-González, E.: Inputs (forcing and observations) ready for use by 'MuSA: The Multiscale Snow Data Assimilation System (v1.0)', https://doi.org/10.5281/zenodo.7248635, 2022.

Alonso-González, E., Gutmann, E., Aalstad, K., Fayad, A., Bouchet, M., and Gascoin, S.: Snowpack dynamics in the Lebanese mountains from quasi-dynamically downscaled ERA5 reanalysis updated by assimilating remotely sensed fractional snow-covered area, Hydrology and Earth System Sciences, 25, 4455–4471, https://doi.org/10.5194/hess-25-4455-2021, publisher: Copernicus GmbH, 2021.

Alonso-González, E., Aalstad, K., Baba, M. W., Revuelto, J., López-Moreno, J. I., Fiddes, J., Essery, R., and Gascoin, S.: The Multiple Snow Data Assimilation System (MuSA v1.0), Geoscientific Model Development, 15, 9127–9155, https://doi.org/10.5194/gmd-15-9127-2022, 2022.

Alonso-González, E., Aalstad, K., Pirk, N., Mazzolini, M., Treichler, D., Leclercq, P., Westermann, S., López-Moreno, J. I., and Gascoin, S.: Spatio-temporal information propagation using sparse observations in hyper-resolution ensemble-based snow data assimilation, Hydrology and Earth System Sciences, 27, 4637–4659, https://doi.org/10.5194/hess-27-4637-2023, 2023.

Alonso-González, E., Mazzolini, M., and Aalstad, K.: ealonsogzl/MuSA: v2.1 software for TC submission, https://doi.org/10.5281/zenodo.11147258, 2024.

Anonymous, R.: Review RC2 to the submission of the paper: Spatio-temporal information propagation using sparse observations in hyper-resolution ensemble-based snow data assimilation, https://doi.org/https://doi.org/10.5194/egusphere-2023-954-RC2, 2023.

Besso, H., Shean, D., and Lundquist, J. D.: Mountain snow depth retrievals from customized processing of ICESat-2 satellite laser altimetry, Remote Sensing of Environment, 300, 113 843, https://doi.org/10.1016/j.rse.2023.113843, 2024.

Carrassi, A., Bocquet, M., Bertino, L., and Evensen, G.: Data assimilation in the geosciences: An overview of methods, issues, and perspectives, WIREs Climate Change, 9, e535, https://doi.org/10.1002/wcc.535, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcc.535, 2018.

Centro Nacional de Información Geográfica, I. G. N.: Plan Nacional de Ortofotografía Aérea, https://pnoa.ign.es/.

Cho, E., Kwon, Y., Kumar, S. V., and Vuyovich, C. M.: Assimilation of airborne gamma observations provides utility for snow estimation in forested environments, Hydrology and Earth System Sciences, 27, 4039–4056, https://doi.org/10.5194/hess-27-4039-2023, 2023.

Clark, M. P., Hendrikx, J., Slater, A. G., Kavetski, D., Anderson, B., Cullen, N. J., Kerr, T., Örn Hreinsson, E., and Woods, R. A.: Representing spatial variability of snow water equivalent in hydrologic and land-surface models: A review, Water Resources Research, 47, https://doi.org/10.1029/2011WR010745, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2011WR010745, 2011.

Cluzet, B., Lafaysse, M., Deschamps-Berger, C., Vernay, M., and Dumont, M.: Propagating information from snow observations with CrocO ensemble data assimilation system: a 10-years case study over a snow depth observation network, The Cryosphere, 16, 1281–1298, https://doi.org/10.5194/tc-16-1281-2022, 2022.

Cressie, N. A. C.: Statistics for spatio-temporal data, iSBN: 1-119-24304-1 Place: Hoboken, New Jersey Series: Wiley series in probability and statistics, 2011.

De Lannoy, G., Reichle, R., Arsenault, K., Houser, P., Kumar, S., Verhoest, N., and Pauwels, V.: Multiscale assimilation of Advanced Microwave Scanning Radiometer–EOS snow water equivalent and Moderate Resolution Imaging Spectroradiometer 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., Albergel, C., Brocca, L., Calvet, J.-C., Carrassi, A., Crow, W. T., de Rosnay, P., Durand, M., Forman, B., Geppert, G., Girotto, M., Hendricks Franssen, H.-J., Jonas, T., Kumar, S., Lievens, H., Lu, Y., Massari, C., Pauwels, V. R. N., Reichle, R. H., and Steele-Dunne, S.: Perspective on satellite-based land data assimilation to estimate water cycle components in an era of advanced data availability and model sophistication, Frontiers in Water, 4, https://doi.org/10.3389/frwa.2022.981745, publisher: Frontiers, 2022.

Dehecq, A., Mannerfelt, E. S., Hugonnet, R., Knuth, F., and Tedstone, A.: xdem, https://doi.org/10.5281/zenodo.4809698, 2021.

Deschamps-Berger, C., Gascoin, S., Shean, D., Besso, H., Guiot, A., and López-Moreno, J. I.: Evaluation of snow depth retrievals from ICESat-2 using airborne laser-scanning data, The Cryosphere, 17, 2779–2792, https://doi.org/10.5194/tc-17-2779-2023, 2023.

Dozier, J., Bair, E. H., and Davis, R. E.: Estimating the spatial distribution of snow water equivalent in the world's mountains, WIREs Water, 3, 461–474, https://doi.org/10.1002/wat2.1140, 2016.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., and Bargellini, P.: Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services, Remote Sensing of Environment, 120, 25–36, https://doi.org/10.1016/j.rse.2011.11.026, 2012.

Durand, M., Molotch, N. P., and Margulis, S. A.: A Bayesian approach to snow water equivalent reconstruction, Journal of Geophysical Research: Atmospheres, 113, https://doi.org/10.1029/2008JD009894, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2008JD009894, 2008.

Eberhard, L. A., Sirguey, P., Miller, A., Marty, M., Schindler, K., Stoffel, A., and Bühler, Y.: Intercomparison of photogrammetric platforms for spatially continuous snow depth mapping, The Cryosphere, 15, 69–94, https://doi.org/10.5194/tc-15-69-2021, 2021.

Emerick, A. A.: Deterministic ensemble smoother with multiple data assimilation as an alternative for history-matching seismic data, Computational Geosciences, 22, 1175–1186, https://doi.org/10.1007/s10596-018-9745-5, 2018.

Emerick, A. A. and Reynolds, A. C.: Ensemble smoother with multiple data assimilation, Computers & Geosciences, 55, 3–15, https://doi.org/10.1016/j.cageo.2012.03.011, 2013.

Enderlin, E. M., Elkin, C. M., Gendreau, M., Marshall, H. P., O'Neel, S., McNeil, C., Florentine, C., and Sass, L.: Uncertainty of ICESat-2 ATL06- and ATL08-derived snow depths for glacierized and vegetated mountain regions, Remote Sensing of Environment, 283, 113 307, https://doi.org/10.1016/j.rse.2022.113307, 2022.

Essery, R.: A factorial snowpack model (FSM 1.0), Geoscientific Model Development, 8, 3867–3876, https://doi.org/10.5194/gmd-8-3867-2015, 2015.

Essery, R.: FSM2 quickstart guide, https://github.com/RichardEssery/FSM2, original-date: 2017-08-14T19:35:47Z, 2023.

Essery, R., Mazzotti, G., Barr, S., Jonas, T., Quaife, T., and Rutter, N.: A Flexible Snow Model (FSM 2.1.0) including a forest canopy, EGUsphere, pp. 1–37, https://doi.org/10.5194/egusphere-2024-2546, publisher: Copernicus GmbH, 2024.

Evensen, G., Vossepoel, F. C., and van Leeuwen, P. J.: Data Assimilation Fundamentals: A Unified Formulation of the State and Parameter Estimation Problem, Springer Textbooks in Earth Sciences, Geography and Environment, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-030-96709-3, 2022.

Fiddes, J., Aalstad, K., and Westermann, S.: Hyper-resolution ensemble-based snow reanalysis in mountain regions using clustering, Hydrology and Earth System Sciences, 23, 4717–4736, https://doi.org/10.5194/hess-23-4717-2019, 2019.

Fiddes, J., Aalstad, K., and Lehning, M.: TopoCLIM: rapid topography-based downscaling of regional climate model output in complex terrain v1.1, Geoscientific Model Development, 15, 1753–1768, https://doi.org/10.5194/gmd-15-1753-2022, publisher: Copernicus GmbH, 2022.

Filhol, S., Fiddes, J., and Aalstad, K.: TopoPyScale: A Python Package for Hillslope Climate Downscaling, Journal of Open Source Software, 8, 5059, https://doi.org/10.21105/joss.05059, 2023.

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

Gascoin, S., Grizonnet, M., Bouchet, M., Salgues, G., and Hagolle, O.: Theia Snow collection: high-resolution operational snow cover maps from Sentinel-2 and Landsat-8 data, Earth System Science Data, 11, 493–514, https://doi.org/10.5194/essd-11-493-2019, publisher: Copernicus GmbH, 2019.

Gascoin, S., Luojus, K., Nagler, T., Lievens, H., Masiokas, M., Jonas, T., Zheng, Z., and De Rosnay, P.: Remote sensing of mountain snow from space: status and recommendations, Frontiers in Earth Science, 12, https://doi.org/10.3389/feart.2024.1381323, publisher: Frontiers, 2024.

Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, Quarterly Journal of the Royal Meteorological Society, 125, 723–757, https://doi.org/10.1002/qj.49712555417, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.49712555417, 1999.

Geissler, J., Rathmann, L., and Weiler, M.: Spatio-temporal Snow Variability in a Sub-Alpine Forest predicted by Machine Learning and UAV-based LiDAR Snow Depth Maps, https://doi.org/10.22541/essoar.167458059.97519903/v1, 2023.

Girotto, M., Margulis, S. A., and Durand, M.: Probabilistic SWE reanalysis as a generalization of deterministic SWE reconstruction techniques, Hydrological Processes, 28, 3875–3895, https://doi.org/10.1002/hyp.9887, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hyp.9887, 2014.

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.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R.: Google Earth Engine: Planetary-scale geospatial analysis for everyone, Remote Sensing of Environment, 202, 18–27, https://doi.org/10.1016/j.rse.2017.06.031, 2017.

Guidicelli, M., Aalstad, K., Treichler, D., and Salzmann, N.: A Combined Data Assimilation and Deep Learning Approach for Continuous Spatio-Temporal SWE Reconstruction from Sparse Ground Tracks, https://doi.org/10.2139/ssrn.4489553, 2023.

Guidicelli, M., Aalstad, K., Treichler, D., and Salzmann, N.: A combined data assimilation and deep learning approach for continuous spatio-temporal SWE reconstruction from sparse ground tracks, Journal of Hydrology X, 25, 100 190, https://doi.org/10.1016/j.hydroa.2024.100190, 2024.

Harder, P., Pomeroy, J. W., and Helgason, W. D.: Advances in mapping sub-canopy snow depth with unmanned aerial vehicles using structure from motion and lidar techniques, preprint, Snow/Remote Sensing, https://doi.org/10.5194/tc-2019-284, 2019.

Hersbach, H.: Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems, Weather and Forecasting, 15, 559–570, https://doi.org/10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2, publisher: American Meteorological Society Section: Weather and Forecasting, 2000.

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., 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, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3803, 2020.

Hock, R.: A distributed temperature-index ice- and snowmelt model including potential direct solar radiation, Journal of Glaciology, 45, 101–111, https://doi.org/10.3189/S0022143000003087, publisher: Cambridge University Press, 1999.

Hultstrand, D. M., Fassnacht, S. R., Stednick, J. D., and Hiemstra, C. A.: Snowpack Distribution Using Topographical, Climatological and Winter Season Index Inputs, Atmosphere, 13, 3, https://doi.org/10.3390/atmos13010003, number: 1 Publisher: Multidisciplinary Digital Publishing Institute, 2022.

Katzfuss, M., Stroud, J. R., and Wikle, C. K.: Ensemble Kalman methods for high-dimensional hierarchical dynamic space-time models, Journal of the American Statistical Association, 115, 866–885, https://doi.org/10.1080/01621459.2019.1592753, arXiv:1704.06988 [stat], 2020.

Largeron, C., Dumont, M., Morin, S., Boone, A., Lafaysse, M., Metref, S., Cosme, E., Jonas, T., Winstral, A., and Margulis, S. A.: Toward Snow Cover Estimation in Mountainous Areas Using Modern Data Assimilation Methods: A Review, Frontiers in Earth Science, 8, https://doi.org/10.3389/feart.2020.00325, 2020.

Lehning, M., Bartelt, P., Brown, B., and Fierz, C.: A physical SNOWPACK model for the Swiss avalanche warning: Part III: meteorological forcing, thin layer formation and evaluation, Cold Regions Science and Technology, 35, 169–184, https://doi.org/10.1016/S0165-232X(02)00072-1, 2002.

Leisenring, M. and Moradkhani, H.: Snow water equivalent prediction using Bayesian data assimilation methods, Stochastic Environmental Research and Risk Assessment, 25, 253–270, https://doi.org/10.1007/s00477-010-0445-5, 2011.

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.

Liu, Z., Filhol, S., and Treichler, D.: Retrieving snow depth distribution by downscaling ERA5 Reanalysis with ICESat-2 laser altimetry, https://arxiv.org/abs/2410.17934v1, 2024.

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, number: 1 Publisher: Nature Publishing Group, 2021.

Luthcke, S.: ATBD for ICESat-2 Received Photon Geolocation, version 2, https://doi.org/10.5067/QN0KOSRLV45J, publisher: NASA National Snow and Ice Data Center Distributed Active Archive Center, 2021.

Ma, X., Li, D., Fang, Y., Margulis, S. A., and Lettenmaier, D. P.: Estimating spatiotemporally continuous snow water equivalent from intermittent satellite observations: an evaluation using synthetic data, Hydrology and Earth System Sciences, 27, 21–38, https://doi.org/10.5194/hess-27-21-2023, publisher: Copernicus GmbH, 2023.

Magnusson, J., Gustafsson, D., Hüsler, F., and Jonas, T.: Assimilation of point SWE data into a distributed snow cover model comparing two contrasting methods, Water Resources Research, 50, 7816–7835, https://doi.org/10.1002/2014WR015302, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2014WR015302, 2014.

Magruder, L., Brunt, K., Neumann, T., Klotz, B., and Alonzo, M.: Passive Ground-Based Optical Techniques for Monitoring the On-Orbit ICESat-2 Altimeter Geolocation and Footprint Diameter, Earth and Space Science, 8, e2020EA001 414, https://doi.org/10.1029/2020EA001414, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2020EA001414, 2021.

Margulis, S. A., Girotto, M., Cortés, G., and Durand, M.: A Particle Batch Smoother Approach to Snow Water Equivalent Estimation, Journal of Hydrometeorology, 16, 1752–1772, https://doi.org/10.1175/JHM-D-14-0177.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology, 2015.

Margulis, S. A., Cortés, G., Girotto, M., and Durand, M.: A Landsat-Era Sierra Nevada Snow Reanalysis (1985–2015), Journal of Hydrometeorology, 17, 1203–1221, https://doi.org/10.1175/JHM-D-15-0177.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology, 2016.

Margulis, S. A., Fang, Y., Li, D., Lettenmaier, D. P., and Andreadis, K.: The Utility of Infrequent Snow Depth Images for Deriving Continuous Space-Time Estimates of Seasonal Snow Water Equivalent, Geophysical Research Letters, 46, 5331–5340, https://doi.org/10.1029/2019GL082507, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019GL082507, 2019.

Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., Farrell, S., Fricker, H., Gardner, A., Harding, D., Jasinski, M., Kwok, R., Magruder, L., Lubin, D., Luthcke, S., Morison, J., Nelson, R., Neuenschwander, A., Palm, S., Popescu, S., Shum, C. K., Schutz, B. E., Smith, B., Yang, Y., and Zwally, J.: The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): Science requirements, concept, and implementation, Remote Sensing of Environment, 190, 260–273, https://doi.org/https://doi.org/10.1016/j.rse.2016.12.029, 2017.

Marti, R., Gascoin, S., Berthier, E., de Pinel, M., Houet, T., and Laffly, D.: Mapping snow depth in open alpine terrain from stereo satellite imagery, The Cryosphere, 10, 1361–1380, https://doi.org/10.5194/tc-10-1361-2016, publisher: Copernicus GmbH, 2016.

Mott, R., Vionnet, V., and Grünewald, T.: The Seasonal Snow Cover Dynamics: Review on Wind-Driven Coupling Processes, Frontiers in Earth Science, 6, https://www.frontiersin.org/articles/10.3389/feart.2018.00197, 2018.

Mott, R., Winstral, A., Cluzet, B., Helbig, N., Magnusson, J., Mazzotti, G., Quéno, L., Schirmer, M., Webster, C., and Jonas, T.: Operational snow-hydrological modeling for Switzerland, Frontiers in Earth Science, 11, https://www.frontiersin.org/articles/10.3389/feart.2023.1228158, 2023.

Murphy, K.: Probabilistic Machine Learning: Advanced Topics, MIT Press, http://probml.github.io/book2, 2023.

Naegeli, K., Neuhaus, C., Salberg, A.-B., Schwaizer, G., Weber, H., Wiesmann, A., Wunderle, S., and Nagler, T.: ESA Snow Climate Change Initiative (Snow_cci): Daily global Snow Cover Fraction - snow on ground (SCFG) from AVHRR (1982 - 2018), version 2.0, https://doi.org/10.5285/3F034F4A08854EB59D58E1FA92D207B6, 2022.

Neumann, T., Brenner, A., Hancock, D., Robins, J., Saba, J., Harbeck, K., Gibbons, A., Lee, J., Luthcke, S., and Rebold, T.: Ice, Cloud, and Land Elevation Satellite (ICESat-2) Project Algorithm Theoretical Basis Document (ATBD) for Global Geolocated Photons ATL03, version 6, https://doi.org/10.5067/GA5KCLJT7LOT, publisher: NASA National Snow and Ice Data Center Distributed Active Archive Center, 2023a.

Neumann, T. A., Martino, A. J., Markus, T., Bae, S., Bock, M. R., Brenner, A. C., Brunt, K. M., Cavanaugh, J., Fernandes, S. T., Hancock, D. W., Harbeck, K., Lee, J., Kurtz, N. T., Luers, P. J., Luthcke, S. B., Magruder, L., Pennington, T. A., Ramos-Izquierdo, L., Rebold, T., Skoog, J., and Thomas, T. C.: The Ice, Cloud, and Land Elevation Satellite – 2 mission: A global geolocated photon product derived from the Advanced Topographic Laser Altimeter System, Remote Sensing of Environment, 233, 111 325, https://doi.org/10.1016/j.rse.2019.111325, 2019.

Neumann, T. A., Brenner A., Hancock D., Robbins J., Gibbons A., Lee J., Harbeck K., Saba J., Luthcke S. B., and Rebold T.: ATLAS/ICESat-2 L2A Global Geolocated Photon Data, version 6, https://doi.org/https://doi.org/10.5067/ATLAS/ATL03.006, 2023b.

Nuth, C. and Kääb, A.: Co-registration and bias corrections of satellite elevation data sets for quantifying glacier thickness change, The Cryosphere, 5, 271–290, https://doi.org/10.5194/tc-5-271-2011, 2011.

Oaida, C. M., Reager, J. T., Andreadis, K. M., David, C. H., Levoe, S. R., Painter, T. H., Bormann, K. J., Trangsrud, A. R., Girotto, M., and Famiglietti, J. S.: A High-Resolution Data Assimilation Framework for Snow Water Equivalent Estimation across the Western United States and Validation with the Airborne Snow Observatory, Journal of Hydrometeorology, 20, 357–378, https://doi.org/10.1175/JHM-D-18-0009.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology, 2019.

Orsolini, Y., Wegmann, M., Dutra, E., Liu, B., Balsamo, G., Yang, K., de Rosnay, P., Zhu, C., Wang, W., Senan, R., and Arduini, G.: Evaluation of snow depth and snow cover over the Tibetan Plateau in global reanalyses using in situ and satellite remote sensing observations, The Cryosphere, 13, 2221–2239, https://doi.org/10.5194/tc-13-2221-2019, publisher: Copernicus GmbH, 2019.

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., and 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.

Pirk, N., Aalstad, K., Yilmaz, Y. A., Vatne, A., Popp, A. L., Horvath, P., Bryn, A., Vollsnes, A. V., Westermann, S., Berntsen, T. K., Stordal, F., and Tallaksen, L. M.: Snow-vegetation-atmosphere interactions in alpine tundra, Biogeosciences Discussions, pp. 1–24, https://doi.org/10.5194/bg-2023-21, publisher: Copernicus GmbH, 2023.

Raleigh, M. S., Lundquist, J. D., and Clark, M. P.: Exploring the impact of forcing error characteristics on physically based snow simulations within a global sensitivity analysis framework, Hydrology and Earth System Sciences, 19, 3153–3179, https://doi.org/10.5194/hess-19-3153-2015, publisher: Copernicus GmbH, 2015.

Revuelto, J., López-Moreno, J. I., Azorin-Molina, C., and Vicente-Serrano, S. M.: Topographic control of snowpack distribution in a small catchment in the central Spanish Pyrenees: intra- and inter-annual persistence, preprint, Snow Hydrology, https://doi.org/10.5194/tcd-8-1937-2014, 2014.

Revuelto, J., Azorin-Molina, C., Alonso-González, E., Sanmiguel-Vallelado, A., Navarro-Serrano, F., Rico, I., and López-Moreno, J. I.: Meteorological and snow distribution data in the Izas Experimental Catchment (Spanish Pyrenees) from 2011 to 2017, Earth System Science Data, 9, 993–1005, https://doi.org/10.5194/essd-9-993-2017, publisher: Copernicus GmbH, 2017.

Revuelto, J., Alonso-Gonzalez, E., Vidaller-Gayan, I., Lacroix, E., Izagirre, E., Rodríguez-López, G., and López-Moreno, J. I.: Intercomparison of UAV platforms for mapping snow depth distribution in complex alpine terrain, Cold Regions Science and Technology, 190, 103 344, https://doi.org/10.1016/j.coldregions.2021.103344, 2021.

Riggs, G. A., Hall, D. K., and Román, M. O.: Overview of NASA's MODIS and Visible Infrared Imaging Radiometer Suite (VIIRS) snow-cover Earth System Data Records, Earth System Science Data, 9, 765–777, https://doi.org/10.5194/essd-9-765-2017, publisher: Copernicus GmbH, 2017.

Sakov, P. and Bertino, L.: Relation between two common localisation methods for the EnKF, Computational Geosciences, 15, 225–237, https://doi.org/10.1007/s10596-010-9202-6, 2011.

Shean, D., Swinski, J. p., Smith, B., Sutterley, T., Henderson, S., Ugarte, C., Lidwa, E., and Neumann, T.: SlideRule: Enabling rapid, scalable, open science for the NASA ICESat-2 mission and beyond, Journal of Open Source Software, 8, 4982, https://doi.org/10.21105/joss.04982, 2023.

Stigter, E. E., Wanders, N., Saloranta, T. M., Shea, J. M., Bierkens, M. F. P., and Immerzeel, W. W.: Assimilation of snow cover and snow depth into a snow model to estimate snow water equivalent and snowmelt runoff in a Himalayan catchment, The Cryosphere, 11, 1647–1664, https://doi.org/10.5194/tc-11-1647-2017, 2017.

Sturm, M. and Wagner, A. M.: Using repeated patterns in snow distribution modeling: An Arctic example, Water Resources Research, 46, https://doi.org/10.1029/2010WR009434, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2010WR009434, 2010.

Sturm, M., Goldstein, M. A., and Parr, C.: Water and life from snow: A trillion dollar science question, Water Resources Research, 53, 3534–3544, https://doi.org/10.1002/2017WR020840, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2017WR020840, 2017.

Sutterley, T. and Gibbons, A.: pyYAPC: Python interpretation of the NASA Goddard Space Flight Center YAPC ("Yet Another Photon Classifier") algorithm, https://doi.org/doi: 10.5281/zenodo.6717591, 2021.

Treichler, D. and Kääb, A.: Snow depth from ICESat laser altimetry — A test study in southern Norway, Remote Sensing of Environment, 191, 389–401, https://doi.org/https://doi.org/10.1016/j.rse.2017.01.022, 2017.

Weiss, A.: Topographic position and landforms analysis, in: Poster presentation, ESRI user conference, San Diego, CA, vol. 200, 2001.

Winstral, A., Elder, K., and Davis, R. E.: Spatial Snow Modeling of Wind-Redistributed Snow Using Terrain-Based Parameters, Journal of Hydrometeorology, 3, 524–538, https://doi.org/10.1175/1525-7541(2002)003<0524:SSMOWR>2.0.CO;2, publisher: American Meteorological Society Section: Journal of Hydrometeorology, 2002.