https://doi.org/10.5194/egusphere-2025-3142
Preprint. Discussion started: 17 September 2025

Assimilating high-resolution satellite snow cover data in a permafrost model

Clarissa Willmes[1], Kristoffer Aalstad[1], and Sebastian Westermann[1,2]

[1]Department of Geosciences, University of Oslo, Norway.
[2]Centre for Biogeochemistry in the Anthropocene, University of Oslo, Norway.

Correspondence: Clarissa Willmes (clarissa.willmes@geo.uio.no)

Abstract.

In high-latitude and mountain areas, the seasonal snow cover features considerable variability over spatial scales of tens of meters which strongly impacts the ground thermal regime. In permafrost areas, spatial patterns of snow accumulation and melt strongly influence the landscape-scale response to atmospheric warming, creating a need to develop simulation tools that can represent such small-scale variability. In this study, we introduce a data assimilation scheme designed to integrate satellite-derived fractional snow covered area (FSCA) into the permafrost model CryoGrid. The idea is to reconstruct the winter snowpack from the satellite-derived timing of the meltout, which at the same time constrains the insulating effect of the snow cover and thereby improves simulations of the ground thermal regime. For this purpose, we employ an iterative ensemble-based data assimilation approach combing a Particle Batch Smoother with Adaptive Multiple Importance Sampling, which proves efficient to manage the high computational demands of the CryoGrid model. The outcomes of the study are assessed using spatially distributed measurements of ground surface temperature (GST) and snow water equivalent collected during four years in an area of approximately $0.5\ \mathrm{km}^2$ on Svalbard. The data assimilation scheme is evaluated using idealized, synthetic FSCA observations compiled from the GST measurements, as well as Sentinel-2 derived FSCA at a spatial resolution of $10\ \mathrm{m}$. For both data sets, the assimilation workflow markedly enhances the representation of the annual GST cycle in the model when simulating extremes in the snow distribution (wind-exposed ridges and snowdrifts) at the scale of individual pixels. However, due to frequent cloudiness and spatial mismatches between the point-scale GST and the $10\ \mathrm{m}$ resolution Sentinel-2 data, the enhancements are less consistent using Sentinel-2 derived FSCA, even resulting in poorer performance in certain years and locations. The data assimilation scheme is further adapted to simulate GST distributions over the entire study area using snowmelt statistics compiled from the $10\ \mathrm{m}$ Sentinel-2 FSCA pixels. The results show that the algorithm can not only reproduce the observed spatial variability of mean annual GST of up to $4.5\ ^\circ\mathrm{C}$, but also the temporal pattern of the spatial variability, with GST varying significantly more in space during snow-covered winter and the snowmelt period compared to the snow-free months. The study showcases that a close integration of observations and models can significantly enhance our ability to characterize the state of the terrestrial cryosphere. In particular, we demonstrate that the data assimilation can uncover the "hidden" variable GST when using a process-rich land surface model like CryoGrid.

https://doi.org/10.5194/egusphere-2025-3142
Preprint. Discussion started: 17 September 2025

1 Introduction

Increases in permafrost temperatures are observed globally (e.g. Biskaborn et al., 2019) and models predict this trend to continue with future climate change. However, model projections of the future permafrost state feature considerable uncertainties not only due to unrepresented processes (e.g. Nitzbon et al., 2019), but also because methods to determine the present-day permafrost state with high confidence at fine spatial resolutions are lacking. While most current permafrost maps, such as the $1\ \mathrm{km}$ resolution map by Obu et al. (2019), are created employing model-based approaches, their spatial resolution is limited by the availability of Earth observation data. Permafrost maps at $1\ \mathrm{km}$ resolution may provide more spatial detail than classic permafrost maps (e.g. Brown et al., 1997), but their resolution is still too coarse to be of value to assess local-scale climate impact and landscape development. The limited spatial resolution also complicates model validation exercises when using point-scale observations of ground temperatures from individual boreholes.

A critical factor influencing permafrost thermal dynamics is the insulating effect of winter snow cover, which can vary dramatically over distances of $10-1000\ \mathrm{m}$ due to wind-driven redistribution of snow (e.g. Zhang, 2005). Recent advances in generating moderate to high resolution satellite retrievals have provided valuable snow data, primarily by leveraging passive optical sensors operating in the shortwave electromagnetic spectrum, such as MODIS, Landsat, and Sentinel-2 (Aalstad et al., 2020; Bair et al., 2020). These passive optical satellite retrievals include the fractional snow-covered area (FSCA) which can offer detailed insight into snowmelt dynamics at a spatial resolutions of up to $10\ \mathrm{m}$ with frequent revisit times of $\leq 5\ \mathrm{days}$. An expanding field of research is also demonstrating how active optical sensor from space-borne lidar onboard ICESat-2 can be used to accurately retrieve snow depth at high spatial but relatively low temporal resolution (Fair et al., 2025). Furthermore, recent studies have shown that satellite-based radar data from Sentinel-1 can be leveraged to map both the wet-snow line (Cluzet et al., 2024) and snow depth (Dunmire et al., 2024) at $100\ \mathrm{m}$ resolution with a subweekly frequency in all weather conditions. Despite their potential, such high-resolution satellite-based snow datasets are currently underutilized in permafrost modeling.

Data assimilation (DA) provides a robust probabilistic framework for combining Earth observations with environmental models that improves model performance and adds value to observations, while accounting for uncertainty in both information sources (Evensen et al., 2022; Sanz-Alonso et al., 2023). The field of data assimilation emerged in operational meteorology where it is widely used to initialize numerical weather predictions and produce atmospheric reanalyses (Kalnay et al., 2024). In the last couple of decades the field has grown tremendously and has been widely adopted across geosciences as an inverse modeling tool to estimate model states and/or parameters using observations (Carrassi et al., 2018; Ismail-Zadeh et al., 2023). Some recent developments in applying data assimilation to new problems in Earth system science have focused on incorporating a wider range of Earth observations, including emerging satellite (Aalstad et al., 2018; Girotto et al., 2023; Mazzolini et al., 2024) and drone data (Alonso-González et al., 2022; Pirk et al., 2022; van Hove et al., 2025) to improve the representation of complex processes and interactions in environmental systems. For example, data assimilation is gaining considerable traction in cryospheric science where it is already widely used in seasonal snow modeling (Alonso-González et al., 2022; Girotto et al., 2023) and glacier modeling (Morlighem and Goldberg, 2023; Cao et al., 2025). Despite a handful of promising pioneering studies (e.g. Zwieback et al., 2019; Groenke et al., 2023), data assimilation has yet to make as broad an impact in permafrost modeling particularly when it comes to leveraging the largely untapped information contained in emerging Earth observations.

One reason for the relative scarcity of data assimilation studies in permafrost science is the fact that permafrost is a subsurface phenomenon that can generally not be observed directly via satellite remote sensing (Trofaier et al., 2017). Using data assimilation to improve permafrost models coupled to other components of the Earth system that strongly modulate the state of permafrost, such as the aforementioned seasonal snow cover, can thus be a promising step forward. In a proof of concept study, Aalstad et al. (2018) showed how assimilating satellite retrievals of FSCA from moderate (MODIS) and higher (Sentinel-2) resolution sensors could be used to constrain the evolution of the snowpack in a high-arctic permafrost environment. The next logical but considerable step to build on the promising findings of the stand alone snow data assimilation in Aalstad et al. (2018) is to perform data assimilation in a more complex modeling system that couples seasonal snow and permafrost modeling.

In this study, we explore different approaches to constrain a multiphysics terrestrial cryospheric model that couples a multilayer seasonal snowpack model to an underlying permafrost model with satellite retrievals of high-resolution fractional snow-covered area (FSCA) data. In particular, we use a state-of-the-art ensemble-based data assimilation method to assimilate FSCA time series into the CryoGrid community model which features a physically-based multilayer representation of snow processes (Westermann et al., 2023). The presented data assimilation workflows target GST on spatial scales from point to small catchments, but they also provide a sophisticated data assimilation scheme for snow variables.

2 Study area

The study area is a high-arctic permafrost site in North-Western Svalbard, Norway. It is located close to the Kongsfjorden about $2\ \mathrm{km}$ south-west of the village of Ny-Ålesund (Fig. 1) in the catchment of the Bayelva river. The area of interest (AOI) is characterized by low-relief ridges and hills with elevations between 10 and 50 m a.s.l. and the tundra surface exhibits sparse vegetation, exposed soil and rock fields (Boike et al., 2018). The area is underlain by continuous permafrost to a depth of about $100\ \mathrm{m}$ (Orvin, 1944) with an active layer thickness between 1 and 2 m (Boike et al., 2018). The permafrost in the area has been observed to be warming since the early 2000s (Grünberg et al., 2024).

Due to the influence of the North Atlantic Current, the areas has a high-arctic maritime climate with cool summers and comparatively mild winters (Maturilli et al., 2013). The mean annual air temperature at the Bayelva climate station located within the AOI was $-4.3°C$ (Boike et al., 2018) in 1998-2021, with an increase in air temperature of $0.091\ \mathrm{°C/year}$ observed between 1999 and 2023 (Grünberg et al., 2024). In summer, the area is mostly under the influence of moist atlantic air masses, whilst it is under the influence of both most atlantic and dry polar air masses in winter (Førland and Hanssen-Bauer, 2003). The annual precipitation in the area was reported to be around $400\ \mathrm{mm}$ that mostly falls as snow in the snow season which is roughly between September to May (Boike et al., 2018). The nearby Kongsfjorden fjord has been largely ice-free during winter since 2006 (Hanssen-Bauer et al., 2019).

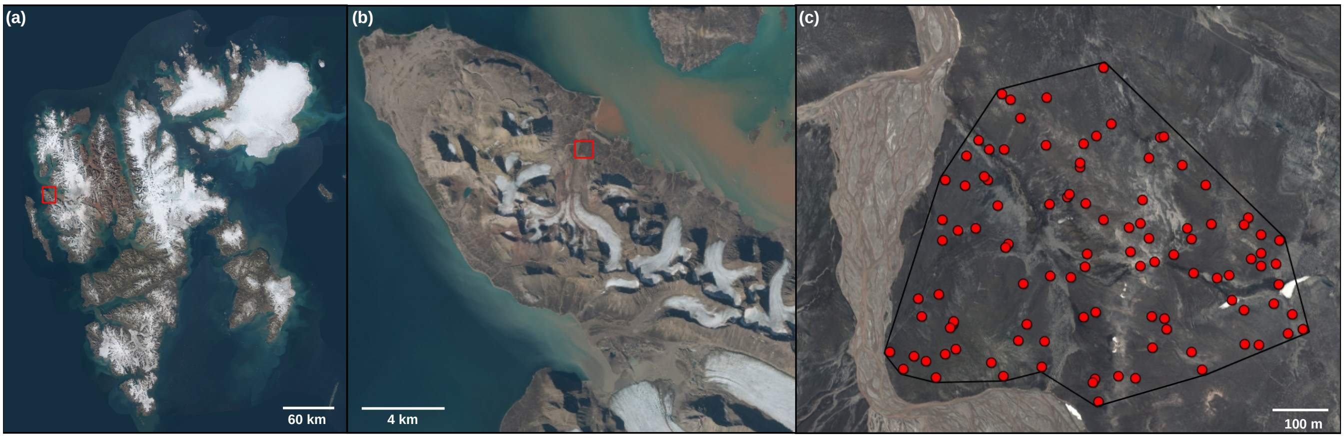

Figure 1. (a) Arial image of Svalbard with Brøgger peninsula and the location of the study site in Western Svalbard. (b) Brøgger peninsula with Ny-Ålesund and the study site. (c) The study area located east of the Bayelva riverbed and surrounding area. Red dots indicate locations of in-situ GST measurements. Black lines outline the AOI. All images are in UTM33X and retrieved and adapted from the Norwegian Polar Institute (https://toposvalbard.npolar.no on 26.06.25)

In the Bayelva area, the seasonal snow cover and its impact on the ground thermal regime have been extensively studied in the past decade. In particular, measurements of ground surface temperature (GST) are conducted at a 100 randomly selected points within an approximately 600 x 800 m^2 area (Fig. 1 c), with corresponding in-situ measurements of snow depth and snow water equivalent at each of the sites (Gisnås et al., 2014). Due to the hilly topography, the area features considerable differences in snow depths due to wind drift, with localized sites almost snow-free, while snow depths exceed 2 m in drifts (Zweigel et al., 2021). In addition, snow depths and snow water equivalent are strongly affected by wintertime melt and rain-on-snow events which also have a strong impact on ground surface and permafrost temperatures (Westermann et al., 2011). The availability of concurrent data sets of the spatial distributions of ground surface temperature and snow parameters makes the Bayelva area a well-suited test site for different permafrost modeling and mapping schemes.

3 Methods

We present a data assimilation framework that ingests satellite retrievals of fractional snow-covered area (FSCA) into the CryoGrid community model, a land-surface model focused on applications in the terrestrial cryosphere (Westermann et al., 2023). The goal is to better directly constrain the snowpack dynamics in the model and thereby indirectly improve the representation of the ground thermal regime from point scale to the entire AOI.

3.1 The CryoGrid community model

The CryoGrid community model is a modular simulation toolbox for applications in the terrestrial cryosphere, in particular permafrost, snow, and glaciers (Westermann et al., 2023). Users can select between different levels of process representation for the subsurface, the snow cover and other key factors (e.g. vegetation, Stuenzi et al., 2021; Zweigel et al., 2024). Due to its modularity, CryoGrid promotes the configuration of customized model setups depending on the characteristics of the study area and the available meteorological model forcing. In particular, the CryoGrid community model facilitates accounting for lateral fluxes of heat, water, and snow through so-called lateral interaction classes (Westermann et al., 2023) to represent the small-scale variability of cryospheric variables typical for many permafrost areas (e.g., Zweigel et al., 2021). In this study, we use a physically-based, one-dimensional (single column) model configuration of a layered subsurface and snowpack with lateral fluxes of water and snow between the model column and external reservoirs.

The ground or snow surface is coupled to the atmosphere by the surface energy balance as well as the mass balance for liquid and solid precipitation, driven by the meteorological forcing data (Sect. 3.1.1). At the lower boundary of the model domain (set to 100 m depth in this study), a geothermal heat flux of 50 $\mathrm{mW/m^2}$ is applied, as in Westermann et al. (2023).

The model domain is represented by vertically layered grid cells associated with different "stratigraphy classes" representing layers in the subsurface ground and the seasonal snowpack, for which a certain set of governing equations and process parametrizations is applied (Westermann et al., 2023). For the subsurface, the state variables are the volumetric contents of mineral, organic, air, water and ice, as well as the enthalpy related to the temperature and ice content of the grid cell. Energy is exchanged between the grid cells through heat conduction and advection through water fluxes. For the water balance, a bucket scheme representing gravity-driven vertical flow is applied. If the water content exceeds the field capacity of a grid cell, water flows to the grid cell below according to its hydraulic conductivity. Lateral overland water flow is accounted for according to the Gauckler-Manning equation (Westermann et al., 2023). Lateral subsurface drainage is realized by applying a seepage face as lateral boundary conditions which continuously drains water from the saturated zone above the permafrost, thus ensuring well-drained conditions for the conditions in the study area. Subsurface parameters are based on (Boike et al., 2018) and Westermann et al. (2023). Up to 5 m depth, we assume a volumetric mineral content of 60% and a porosity of 40 %, which is filled with ice in the permafrost, while water, ice and air contents vary in the active layer according to the applied model forcing. Below 5 m depth, bedrock is assumed. In a model sensitivity study, this setting was found to be an adequate representation for the conditions at the Bayelva site (Westermann et al., 2023). We apply the same stratigraphy and hydrological conditions to all locations in the AOI and we do not consider differences in exposition, which is a good approximation considering the gentle topography of the study area.

In the presence of a seasonal snow cover, a multilayered snow scheme is applied on top of the ground class, which follows the Crocus snow model by Vionnet et al. (2012) and is described in detail in Westermann et al. (2023). Snowfall is added with the temperature and wind-speed dependent properties density, grain size, dendricity, and sphericity (Vionnet et al., 2012). After deposition, snow metamorphism is parameterized in each grid cell based on snow age, temperature gradients, water content and overlying snow mass (Westermann et al., 2023). Liquid water content is again simulated with a bucket scheme which allows for infiltration and refreezing of water within the snowpack. Transmission of solar radiation into the snowpack and the resulting surface albedo are calculated according to the snow microphysical parameters for three spectral bands (0.3–0.8, 0.8–1.5 and 1.5–2.8 μm) following Vionnet et al. (2012). Lateral flow of liquid water within the snowpack is not taken into account, but overland flow on the surface of the snowpack is used to remove excess water when it overtops the snowpack. The

https://doi.org/10.5194/egusphere-2025-3142
Preprint. Discussion started: 17 September 2025

snow is parameterized as in Westermann et al. (2023), but with of a slightly higher maximum wind slab density of 600 $\mathrm{kg/m^3}$. We apply a surface roughness length of $z_0 = 0.0001\ \mathrm{m}$ for the snow.

Extending the work of Zweigel et al. (2021) to a one-dimensional model setup with an external snow reservoir, wind-driven snow removal and deposition is parameterized as a function of the snow microphysical parameters (dendricity, sphericity, grain size, density), building on the existing formulations for one-dimensional wind compaction of the snow in Crocus (Vionnet et al., 2012). Each grid cell is assigned a characteristic timescale for the change of the parameters under wind transport (τ_i), which assumes an exponential decrease of snow mobility with depth (Vionnet et al., 2012):

$$\tau_i = \frac{\tau}{\max\left(S_i \exp(-z_i/\lambda), 0\right)}, \tag{1}$$

Here, z_i is the depth below the snow surface, $\lambda = 0.1\,\mathrm{m}$ is an e-folding length scale, and τ is an empirical constant set by the user. The driftability index S_i of layer i is given by an empirical function of the horizontal wind speed and snow microphysical parameters (Vionnet et al., 2012). As in Zweigel et al. (2021), the characteristic timescale τ_i is used to parameterize lateral snow removal or deposition by wind. We assume the amount of mobilized snow $F_{i,\mathrm{snow}}$ per unit time to be proportional to the inverse of τ_i, with the exposure e determining how much snow is either removed or deposited, allowing to adapt the model to site-specific snow conditions, e.g. snow drifts are ridges with little snow. To constrain the range of e, we introduce an empirical drift factor N_{drift} (as in Zweigel et al., 2021), which is essentially a global scaling factor for the lateral snow fluxes:

$$F_{i,\mathrm{snow}} = \frac{e N_{\mathrm{drift}}}{\tau_i}. \tag{2}$$

For $e < 0$, the mobilized snow is deposited in the model column, while snow is eroded from the model column for $e > 0$, making it possible to represent areas with different snow depths without changing the snowfall in the model forcing (as e.g. in Martin et al., 2019). While the exposure in principle parametrizes the elevation of the model column relative to the surrounding local topography, the value of the exposure parameter e is inferred by the DA scheme (Sect. 3.3) based on observations of FSCA (Sect. 3.2) in this study.

3.1.1 Model forcing data

For meteorological forcing, we use downscaled ERA5 reanalysis (Hersbach et al., 2020) surface fields bias-corrected with in-situ observations from the nearby Bayelva climate station as described in Westermann et al. (2023). The forcing data include air temperature, incoming shortwave radiation, incoming longwave radiation, specific humidity, air pressure, wind speed, rainfall, and snowfall. When studying wind redistribution of snow, the near-surface wind speed is an especially important factor as it controls the occurrence and intensity of drifting snow through the driftability index S_i (Sect. 3.1). The near-surface wind speed is often misrepresented in atmospheric reanalysis such as ERA5, especially during high wind periods (e.g., Wilczak et al., 2024; Gualtieri, 2021) which strongly contribute to snow redistribution and metamorphism. Furthermore, the wind speed is influenced by the local topography, with elevated ridges experiencing increased wind speeds compared to sheltered locations, thus leading to a small-scale variability of wind speeds that can affect snow drift intensity and patterns (Liston and Elder, 2006). Therefore, we scale the surface wind speed using quantile mapping (e.g., Fiddes et al., 2022). We apply a linear scaling for quantiles between minimum wind speeds, which are scaled with factor 1, and maximum wind speeds, which are scaled with a wind scaling factor f. In this study, f is an uncertain parameter determined by the DA scheme which enables a better representation of the impact of surface wind on the snowpack.

3.2 Fractional Snow Covered Area

The fractional snow-covered area (FSCA) is a measure of the fraction of ground surface covered by snow, quantified by values between 0 for snow-free ground and 1 for fully snow-covered ground. Time series of FSCA retrievals during the snowmelt season have been widely used to reconstruct snow water equivalent (SWE) (e.g. Margulis et al., 2015; Aalstad et al., 2018) and snow depth (SD) (Fiddes et al., 2019; Mazzolini et al., 2024), but generally without a focus on the underlying ground thermal regime and often with a standalone snow model. Assimilating FSCA into a process-based coupled snow-land surface model like CryoGrid allows for bidirectional interaction between the snowpack and the underlying ground. In this study, we assimilate remotely sensed FSCA retrieved from Sentinel-2 observations, as well as in-situ derived FSCA compiled from ground surface temperature (GST) observations.

Deriving FSCA from in-situ GST measurements (see Sect. 3.4) offers binary snow / no snow point-scale observations, which are well-suited for testing the DA scheme. The point in time when the snow has fully melted above the measurement location becomes manifest when GSTs show a clear daily cycle in line with air temperatures, in particular positive GST values. We assume $\mathrm{FSCA} = 1$ before this date and $\mathrm{FSCA} = 0$ after.

For satellite retrievals of FSCA, we use VSWIR imagery from the Sentinel-2A and 2B twin satellites and the retrieval algorithm described in Aalstad et al. (2020). We employ the level 2 surface reflectance product in the blue ($\sim 0.49\,\mu\mathrm{m}$), green ($\sim 0.56\,\mu\mathrm{m}$), red ($\sim 0.66\,\mu\mathrm{m}$), NIR ($\sim 0.83\,\mu\mathrm{m}$), SWIR1 ($\sim 1.6\,\mu\mathrm{m}$) and SWIR2 ($\sim 2.2\,\mu\mathrm{m}$) bands with spatial resolutions between 10 and 20 m depending on the band. Bands with the coarser resolution of 20 m (SWIR1 and SWIR2) were resampled to 10 m resolution using nearest neighbor interpolation. A fully constrained linear spectral unmixing approach described in Aalstad et al. (2020) is then applied to retrieve FSCA from the surface reflectances. With a frequent revisit time of ≤ 5 days per satellite and a large swath of 290 km (Drusch et al., 2012), we obtain an acquisition approximately every $1-2$ days for our high-arctic AOI in the studied years. Cloud-free scenes were manually selected. Fig. 2 shows example scenes of FSCA retrieved from Sentinel-2 optical imagery during the 2017 snowmelt season over the AOI within the Bayelva catchment. The uncertainty of the FSCA observations is considered in the DA scheme by applying an observational variance of $\sigma^2_{\mathrm{FSCA}} = 0.05$ for both the in-situ derived FSCA and Sentinel-2 derived FSCA.

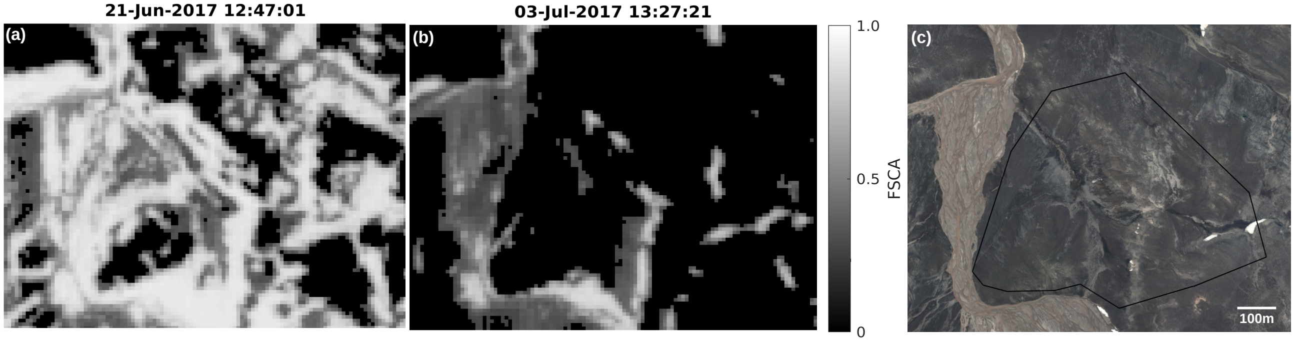

Figure 2. (a),(b): Sentinel-2 FSCA snapshots during the snow meltout period. (c): Aerial image of the study area (same extent as in a, b), and AOI outlined in black. Retrieved and adapted from the Norwegian Polar Institute (https://toposvalbard.npolar.no on 26.06.25). All images are in UTM 33X.

3.3 Data Assimilation

Data Assimilation (DA) is an uncertainty-aware approach based on Bayesian inference that can identify system states that are consistent with both the uncertainty of the model and the observation uncertainty. While commonly used in operational weather forecasting (Kalnay et al., 2024), DA has shown its usefulness in geosciences (Carrassi et al., 2018) including cryosphere applications (Girotto et al., 2023; Morlighem and Goldberg, 2023), but it has not yet been widely applied for permafrost modeling. In this study, it allows us to find the best representation of the true snowpack and associated ground surface temperatures, given the physical constraints of the system (i.e. the CryoGrid model) and the observations (i.e. remotely sensed FSCA). In general, we solve an inference problem for model parameters via Bayes' rule (MacKay, 2003; Evensen et al., 2022)

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{Z} \tag{3}$$

in which $p(\boldsymbol{\theta})$ is the prior probability distribution over parameters $\boldsymbol{\theta}$, $p(\mathbf{y}|\boldsymbol{\theta})$ is is a measure of how well the model fits the observations (i.e. the Gaussian likelihood, described in detail below), and Z is a normalizing constant. The target of inference is the posterior probability distribution over parameters $p(\boldsymbol{\theta}|\mathbf{y})$ that encodes updated knowledge about the parameters after considering the observations. A posterior distribution over dynamical states $\mathbf{x} = \mathcal{M}(\boldsymbol{\theta})$ and predicted observations $\widehat{\mathbf{y}} = \mathcal{G}(\boldsymbol{\theta})$ is obtained by running the forward model with some representation of the parameter posterior. The forward model generating synthetic observations $\mathcal{G} = \mathcal{H} \circ \mathcal{M}$ is a composition of the observation operator $\mathcal{H}$ that picks out the predicted observations from the dynamical model $\mathcal{M}$ (i.e. CryoGrid in this study).

3.3.1 Ensemble generation in CryoGrid

For each hydrological year, we simulate an ensemble of model realizations consisting of ensemble members (also referred to as particles). The state trajectory for each of these particles consists of running CryoGrid with a different combination of key parameters that strongly influence the snowpack evolution. In particular, we perturb the exposure e and the scaling factor for surface wind speeds f, which both modulate the lateral snow redistribution and thus the overall snow water equivalent in our one-dimensional (single column) model setup (Sect. 3.1). The perturbed parameters are assumed to be constant throughout each individual hydrological year (start of September to the end of August), and we assimilate FSCA observations independently between individual snowmelt seasons. Each particle is run for two consecutive hydrological years, of which the second year is the hydrological year of interest for which FSCA is assimilated. Although a common initial condition is used, this ensures a consistent spinup of the ground thermal regime and hydrology in the uppermost soil layers which is sufficient when targeting ground surface temperatures (Sects. 2, 3.4).

The ensemble consists of $N_e = 20$ ensemble members per iteration (Sect. 3.3.2). Initially, each ensemble member draws random values from probability distributions for the independent parameters e and f. We assume Gaussian prior distributions for both parameters with $\mu_e = 0$ and $\sigma_e = 1$, as well as $\mu_f = 1$ and $\sigma_f = 0.3$. The probability distributions are iteratively updated by the DA scheme (Sect. 3.3.2), based on the performance of each ensemble member compared to the assimilated FSCA observations. All other model parameters are considered constant across the ensemble.

3.3.2 Adaptive Particle Batch Smoother (AdaPBS)

In this study, we solve a strong constraint forcing formulation of the data assimilation problem (Evensen et al., 2022) by employing a batch smoothing variant of the Adaptive Multiple Importance Sampling (AMIS) scheme originally proposed by Cornuet et al. (2012) which we modify to fit the requirements and computational constraints of the CryoGrid model (Sect. 3.1). In particular, we use the AMIS algorithm at the core of an iterative extension of the traditional (non-iterative) particle batch smoother (PBS, Margulis et al., 2015), a widely used cryospheric data assimilation scheme (e.g. Fiddes et al., 2019; Alonso-González et al., 2022; Girotto et al., 2024; Cao et al., 2025), to devise a more robust adaptive sampling scheme (see Bugallo et al., 2017) that we refer to as the Adaptive Particle Batch Smoother (AdaPBS). Herein, the AdaPBS is used as an iterative DA scheme that combines a parallel (i.e. an ensemble of model realizations in each iteration) with a sequential component (i.e. repeated model ensembles drawn from distributions evolving with each iteration). This AdaPBS method is well-suited for a relatively "slow", computationally complex model like CryoGrid (in our case on the order of 15 minutes per model year), as the ensemble in each iteration can be run in parallel on a relatively small number of cores (in our case 20), while sequential iterations facilitate adaptively migrating to and zooming in on promising regions of the parameter space. This reduces unnecessary model runs in less promising parameter regions, thus providing the potential of convergence with fewer model runs compared to non-iterative schemes in which the entire possible parameter space must be sampled at high enough densities in order to determine a sufficient number of "successful" ensemble members. Moreover, unlike iterative ensemble Kalman methods with a fixed number of iterations (Aalstad et al., 2018), the adaptive nature of this AdaPBS scheme means that it can converge early and will only in the worst case run for a predetermined maximum number of iterations.

In each iteration ℓ, we sample $i = 1, \ldots, N_e$ particles in the form of parameter vectors $\boldsymbol{\theta}_i^{(\ell)}$ containing the exposure e (Sect. 3.1) and the wind scaling factor f (Sect. 3.1.1) from the sampling proposal $q^{(\ell)}(\boldsymbol{\theta})$ which evolves via an adaptation strategy across in each iteration (see below). For the first iteration, the sampling proposal $q^{(1)}(\boldsymbol{\theta})$ is set to the prior

$p(\boldsymbol{\theta}) = \mathrm{N}(\boldsymbol{\theta}|\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$, with samples drawn from independent Gaussian distributions for e and f (Sect. 3.3.1), centered around $\mu_f^{(0)} = 1$ and $\mu_e^{(0)} = 0$. We note that a simulation with $\mu_f^{(0)} = 1$ and $\mu_e^{(0)} = 0$ corresponds to the regular forward run with CryoGrid (denoted "reference run" in the following), i.e. using the meteorological forcing as described in Sect. 3.1 and snow deposition solely by snowfall, not accounting for snow redistribution. In each iteration, nonlinear forward simulations $\mathcal{G}(\cdot)$ with CryoGrid are performed for all particles with parameter vectors $\boldsymbol{\theta}_i^{(\ell)}$ sampled from the current proposal $\boldsymbol{\theta}_i^{(\ell)} \sim q^{(\ell)}(\boldsymbol{\theta})$, resulting in predicted (i.e. modeled) observations of FSCA, denoted $\widehat{\mathbf{y}}_i^{(\ell)} = \mathcal{G}(\boldsymbol{\theta}_i^{(\ell)})$. By comparing these predicted observations $\widehat{\mathbf{y}}_i^{(\ell)}$ against the actual yet uncertain FSCA observations $\mathbf{y}$ their 'goodness of fit' is evaluated through the likelihood function $p(\mathbf{y}|\boldsymbol{\theta}_i^{(\ell)})$. Assuming typical additive Gaussian observation error (Carrassi et al., 2018) results in a nonlinear Gaussian likelihood $p(\mathbf{y}|\boldsymbol{\theta}) = \mathcal{N}(\mathbf{y}|\widehat{\mathbf{y}}, \mathbf{R})$ of the form

$$p(\mathbf{y}|\boldsymbol{\theta}) = \frac{1}{c_y} \exp\left(-\frac{1}{2} \left[\mathbf{y} - \widehat{\mathbf{y}}\right]^{\mathrm{T}} \mathbf{R}^{-1} \left[\mathbf{y} - \widehat{\mathbf{y}}\right] \right), \tag{4}$$

where $\mathbf{R} = \sigma_y^2 \mathbf{I}$ is a diagonal observation error covariance matrix with observation error variance σ_y^2, $\widehat{\mathbf{y}} = \mathcal{G}(\boldsymbol{\theta})$ are the predicted observations from the nonlinear forward model $\mathcal{G}$ with parameters set to $\boldsymbol{\theta}$, and $c_y = \det(2\pi\mathbf{R})^{1/2}$ is a normalizing constant. We then perform the DA step in which the self-normalized importance weights $w_i^{(j)}$ are (re)calculated for all generations of particles i across the full history of currently available iterations $j = 1, \ldots, \ell$ as follows

$$w_i^{(j)} = \frac{1}{c^{(\ell)}} \frac{p(\mathbf{y}|\boldsymbol{\theta}_i^{(j)}) p(\boldsymbol{\theta}_i^{(j)})}{m^{(\ell)}(\boldsymbol{\theta}_i^{(j)})}, \tag{5}$$

which is obtained by applying importance sampling to Bayes' rule in Eq. (3). In practice, the weights in Eq. (5) are computed in log-space using the so-called log-sum-exp trick (Murphy, 2023) to ensure numerical stability. Furthermore, the normalizing constant in Eq. 5 is given by

$$c^{(\ell)} = \sum_{j=1}^{\ell} \sum_{i=1}^{N_e} \frac{p(\mathbf{y}|\boldsymbol{\theta}_i^{(j)}) p(\boldsymbol{\theta}_i^{(j)})}{m^{(\ell)}(\boldsymbol{\theta}_i^{(j)})}, \tag{6}$$

which ensures self-normalization so that the weights over the entire particle history sum up to one $\sum_{j=1}^{\ell} \sum_{i=1}^{N_e} w_i^{(j)} = 1$. Recall that in Eq. (5) $p(\mathbf{y}|\boldsymbol{\theta}_i^{(j)})$ and $p(\boldsymbol{\theta}_i^{(j)})$ are the likelihood and prior, respectively, of $\boldsymbol{\theta}_i^{(j)}$ whereas the new term $m(\boldsymbol{\theta}_i^{(j)})$ is the so-called deterministic mixture proposal (Owen and Zhou, 2000) defined as

$$m^{(\ell)}(\boldsymbol{\theta}_i^{(j)}) = \sum_{k=1}^{\ell} \frac{1}{\ell} q^{(k)}(\boldsymbol{\theta}_i^{(j)}), \tag{7}$$

with $1/\ell$ serving as a constant (equal) mixture weight. This deterministic mixture proposal is a mixture of the entire history of proposals $q^{(k)}$ for $k = 1, \ldots, \ell$ that have been computed thus far. In this way, we can reuse all the previously sampled particles $\boldsymbol{\theta}_i^{(j)}$ and their associated (costly) likelihood evaluations to obtain an extended and more diverse set of particles that is less likely to be degenerate than those from a single (let alone a non-adaptive) proposal. Crucially, we can directly reuse promising particles from past iterations without having to repeat computationally expensive CryoGrid simulations. From the weights, the effective ensemble size $N_{\text{eff}} = \sum_{k=1}^{\ell} \sum_{i=1}^{N_e} (w_i^k)^{-2}$ (Elvira et al., 2022) is employed to measure particle diversity (non-degeneracy) and thus for gauging the success of the DA step. If the particle diversity $D = N_{\text{eff}}/N_e$ is below a user-defined threshold $D_{\min}$, another iteration is initiated by constructing a new sampling proposal $q^{(\ell+1)}$ from which a new particle ensemble (as defined by model parameters e and f) is drawn for the next iteration $\ell \leftarrow \ell+1$.

As a first step to construct the new proposal $q^{\ell+1}$, we resample N_e particles with replacement from the set of $N_e \times \ell$ history of particles $\boldsymbol{\theta}_i^{(j)}$ according to their weights $w_i^{(j)}$. Ideally, the combined historical ensemble of particles contains a sufficient number of unique particles carrying non-negligible weights, so that the ensemble mean vector $\boldsymbol{\mu}^{(\ell+1)}$ and covariance matrix $\boldsymbol{\Sigma}^{(\ell+1)}$ of the resampled particles can be used to define a non-degenerate normal distribution as the proposal $\widetilde{q}_r^{(\ell+1)}(\boldsymbol{\theta}) = \mathrm{N}(\boldsymbol{\theta}|\boldsymbol{\mu}^{(\ell+1)},\boldsymbol{\Sigma}^{(\ell+1)})$, from which a new ensemble $\boldsymbol{\theta}_i^{(\ell+1)}$ $(i = 1,\ldots,N_e)$ can be drawn. However, in this study, the relatively small number of ensemble members per iteration often leads to a degenerate ensemble at least for the first iterations, with weights concentrated on a single or a few ensemble members. In this case, we instead sample from a (typically) broader Gaussian proposal distribution $\widetilde{q}_b^{(\ell+1)}(\boldsymbol{\theta}) = \mathrm{N}(\boldsymbol{\theta}|\boldsymbol{\mu}^{(\ell+1)},\boldsymbol{\Sigma}_b^{(\ell+1)})$ with the same adapted mean $\boldsymbol{\mu}^{(\ell+1)}$ from the resampled particles as used in $\widetilde{q}_r^{(\ell+1)}$ but with a larger spread set by a covariance $\boldsymbol{\Sigma}_b^{(\ell+1)}$ based on the prior covariance $\boldsymbol{\Sigma}_0$. To determine the covariance $\boldsymbol{\Sigma}_b$ of this broader Gaussian distribution we employed two strategies which both can be promising depending on the situation. On the one hand, retaining the large prior covariance makes it possible that the ensemble sampled from the proposal can more effectively explore by migrating within the parameter space from iteration to iteration, even outside the original plausible range covered by the prior. On the other hand, iteratively reducing the proposal covariance allows the ensemble to exploit by obtaining a denser coverage in promising parameter regions, which may increase the chance of obtaining a non-degenerate ensemble from which $\widetilde{q}_r$ can be constructed for the next iteration. In this study, we alternate between both strategies, with $\boldsymbol{\Sigma}_b = \alpha^{(\ell+1)}\boldsymbol{\Sigma}_0 + (1-\alpha^{(\ell+1)})R_D^{\ell}\boldsymbol{\Sigma}_0$ where $R_D = \min(D/D_{\min},1)$ is the diversity ratio, and $\alpha^{(\ell+1)}$ is a realization of a uniform random variable between 0 and 1 sampled in each iteration. While the second term leads to a continuous reduction of the width as the number of iterations ℓ increases, the first term retains the width of the prior distributions, and the varying α-parameter ensures that both cases occur throughout the iterations. Finally, $q^{(\ell+1)}$ is obtained through

$$q^{(\ell+1)}(\boldsymbol{\theta}) = R_D\,\widetilde{q}_r^{(\ell+1)}(\boldsymbol{\theta}) + (1-R_D)\,\widetilde{q}_b^{(\ell+1)}(\boldsymbol{\theta})\,. \tag{8}$$

For a degenerate ensemble (i.e. $R_D \ll 1$), the first term collapses to a single unique particle or at least a much too a very narrow distribution, so the second term ensures that a sufficiently diverse ensemble can still be sampled. On the other hand, if D approaches the termination threshold $D_{\min}$ so that $R_D \simeq 1$, the first term is prioritized so that the new ensemble is to a large extent drawn from the distribution spanned by the "successful" particles from the previous iterations. In this study, the termination threshold is set to $D_{\min} = 6$.

3.4 In-situ observations

To evaluate the performance of the CryoGrid snow-permafrost DA system, we compare the modeling results to field observations of ground surface temperature (GST) and in-situ snowpack observations. Furthermore, the GST observations are used to create in-situ derived FSCA time series for individual measurement sites (Sect. 3.2) which are assimilated in the model to benchmark the performance without the additional complexities introduced by employing remotely sensed FSCA observations retrieved from space-borne satellite sensors (Sect. 3.5).

GST measurements from about 100 iButton loggers of the hydrological years 2016-2019 are used, which are randomly distributed in space within the study area in order to capture small-scale spatial variability of the thermal regime (Gisnås et al., 2014). The GST is observed every 4 hours at 2–3 cm below the surface with an accuracy of approximately 0.2 °C (Zweigel et al., 2021). More details on these observations can be found in Gisnås et al. (2014).

In-situ snow observations from the end of the snow accumulation season of the hydrological years 2016-2019 are used for additional validation. As described in Zweigel et al. (2021), the observations were conducted around the time of peak snow water equivalent (SWE) and consist of snow depth (SD) observations at each GST logger site and SWE estimates for each logger site which were reconstructed with snow densities measured each year at a handful of representative, detailed snow pits at selected GST logger sites.

3.5 DA Experiments

In this study, we apply the aforementioned DA in different configurations:

We conduct a *Point-Scale DA with in-situ derived FSCA* experiment, in which we assimilate binary FSCA observations derived from GST measurements as described in Sect. 3.2. The resulting FSCA timeseries correspond to a hypothetical satellite system with effectively infinitely high resolution in space and time and serves as a proof of concept for the DA scheme.

We then extend this point-scale study to the use of real-world satellite observations, i.e. FSCA derived from Sentinel-2, in a *Point-Scale DA with Sentinel-2 derived FSCA* experiment. This has the purpose of obtaining high-resolution (10 m x 10 m Sentinel-2 resolution) simulations for individual sites, approaching the spatial scale of individual GST measurement sites. The Sentinel-2 FSCA observations can contain relatively long (a week or more) data gaps due to cloudiness.

Furthermore, we employ the DA in a spatial manner in a *Spatial DA with Sentinel-2 derived FSCA* experiment. This allows us to model the spatial distributions of the ground thermal regime within a medium-scale (500 m-1 km) area of interest (AOI). We use the cumulative FSCA retrieved from Sentinel-2 over the AOI, assume a constant melt-rate between individual retrievals and obtain an area-weighted histogram of meltout dates for each of the studied years. For each of the actual Sentinel-2 FSCA observation dates, we then run the point-scale DA scheme described above by assimilating the corresponding synthetic binary FSCA in the same manner as in *Point-Scale DA with in-situ derived FSCA*. The resulting ensemble means (of e.g. GST or SWE) from these DA runs are then area-weighted in accordance with the Sentinel-2-derived meltout histogram, yielding distributions of ground and snow parameters, in particular mean monthly and mean annual ground surface temperature and snow water equivalent (SWE).

We further produce 10 m resolution GST maps in a *GST Backmapping* experiment which infers the spatial patterns of GST by clustering the pixels based on their initial observation of meltout in the Sentinel-2 FSCA data. For each pixel, the GST is calculated as the average between the GSTs obtained from the DA simulations for the last snow-covered and first snow-free Sentinel-2 scenes, i.e. for the two scenes between which the actual meltout of the pixel occurred. This provides insights into spatial GST patterns in the AOI without incurring additional computational costs, as it utilizes the simulations of the *Spatial DA with Sentinel-2 derived FSCA* experiment.

https://doi.org/10.5194/egusphere-2025-3142
Preprint. Discussion started: 17 September 2025

4 Results

4.1 Point-Scale DA with in-situ derived FSCA

We present the results of the DA assimilating in-situ derived FSCA on the point-scale at two benchmark sites representing endmembers in the snow distribution, namely a snow drift with deep snow and a wind-blown ridge with generally little snow.

The resulting modeled GSTs without DA and with DA as well as the observed GSTs are shown in Fig. 3. The reference run without DA (open loop) is the same for both locations (red solid line) as it does not account for lateral snow redistribution. We applied the same prior distributions of the perturbed parameters $\boldsymbol{\theta} = [e, f]^{\mathrm{T}}$ for all sites and years (see Supplement Fig. S1). The prior ensemble GST range (gray shaded area) covers the entire range of GSTs between a multi-annually persistent snow patch and fully bare-blown ground. The posterior mean RMSE of the monthly mean GST (RMSEmm) is reduced by 1.5 °C at the snowdrift location and 0.7 °C at the windblown ridge location for the hydrological year of 2016 compared to the reference run. For the hydrological year of 2017, the RMSEmm is reduced by 2.7 °C at the snowdrift location and by 3.0 °C at the windblown ridge location compared to the reference run. The DA thus markedly improves the simulated GSTs in comparison to the reference run. However, the posterior GST ensembles generally display a small spread that often does not overlap with the observations, which indicates overconfidence by placing most of the weight on just a few parameter combinations. This is a tradeoff from using a small initial ensemble size for numerical cost reasons which makes well calibrated uncertainty quantification using this DA scheme in our experiments difficult, but it does not compromise the main goal of the scheme, which is improving the local GST representation as obtained by the ensemble mean as an optimal point estimate. By reproducing local binary FSCA the scheme finds better parameter combinations of e and f that directly allow for an overall improved representation of the local snow cover, which indirectly results in an improved model estimates of GST.
In case of the snow drift (Fig. 3, left panels), the reference run clearly underestimates the snow depth, leading to too low winter temperatures and a too early meltout of the snowpack. This is greatly improved by the DA scheme, which results in a more accurate later meltout date, a deeper snowpack, and higher winter GST. For the windblown ridge (Fig. 3, right panels), the reference run overestimates the snow depth, leading to overestimated winter GSTs and underestimated spring GSTs due to an unrealistically late meltout. Application of the DA scheme results in earlier meltout, less snow, and colder winter ground temperatures.

As the effect of subgrid snow dynamics on GST variability is most relevant in years with relatively little snow, the improvement of GST representation by applying the DA scheme is largest in those years. We show results for the hydrological years of 2016 and 2017, but found comparable results for the other studied years of 2018 and 2019 (see Fig. S1). The improvement of GST representation by the model is seen for all years at both benchmark sites and showcases the functionality of the DA scheme for close to ideal assimilated observations with high spatio-temporal resolution.

For the wind-blown ridge in the hydrological year of 2017 (Fig. 3), the modeled GSTs show a too late meltout for both the reference run and the DA. Whilst the DA improves the melt-out date and GSTs, the scheme fails to reproduce the early meltout of the windblown ridge in this particular year. This is likely due to the occurrence of snowfall in the atmospheric forcing which

https://doi.org/10.5194/egusphere-2025-3142
Preprint. Discussion started: 17 September 2025

385 did not occur in reality, which is a common limitation occurring when using coarse scale forcing in physics-based modeling on such small scales (Fang et al., 2023).

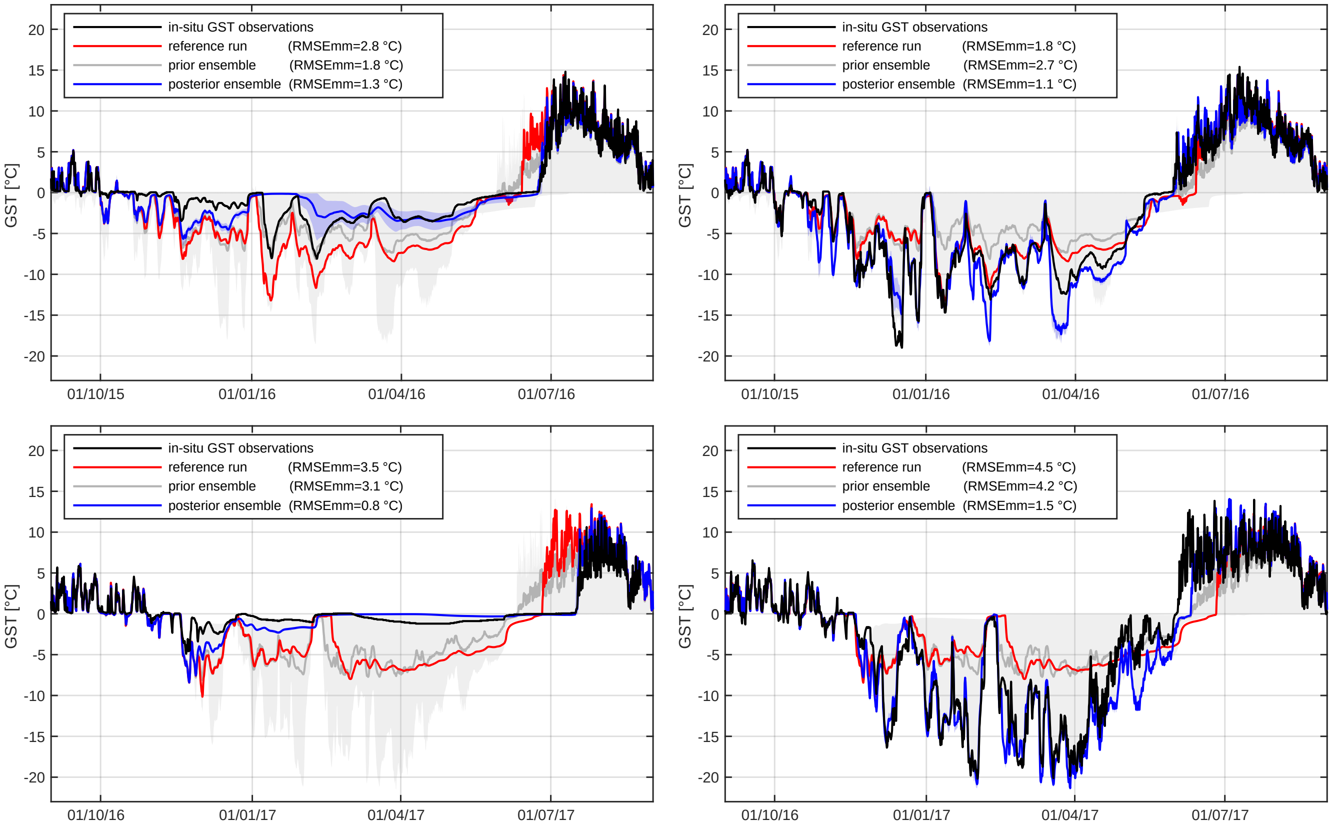

Figure 3. Observed and modeled GSTs during the hydrological years of 2016 and 2017 for a snow drift location (left panels) and a wind-exposed ridge (right panels). The modeled GST of the reference run without DA are shown in red. The prior ensemble (before DA) is shown in gray, with the prior ensemble mean displayed as the gray solid line and the prior ensemble range as the gray shaded area. The posterior results of the DA scheme assimilating in-situ derived FSCA are shown in blue, with the blue solid line being the posterior ensemble mean and the blue shaded area the posterior ensemble range. The RMSE of the monthly mean GSTs relative to the in-situ measurements is given for the reference run and the ensemble means of the prior and posterior ensembles each.

https://doi.org/10.5194/egusphere-2025-3142
Preprint. Discussion started: 17 September 2025

4.2 Point-Scale DA with Sentinel-2 derived FSCA

We present the results of the DA scheme assimilating Sentinel-2 derived FSCA at the point-scale for the benchmark sites described in Sect. 4.1. The same model and DA setup as in the previous experiment is applied.

Fig. 4 displays the Sentinel-2 retrieved FSCA observations as well as the modeled FSCA of the reference run and of the DA scheme during the two exemplary years of 2016 and 2017 (additional years in Fig. S2). As discussed in Sect. 3.2, the satellite-derived FSCA observations are less frequent than the in-situ derived FSCA. The Sentinel-2 derived FSCA observations of the two benchmark sites show a gradual meltout over extended periods (e.g. approximately one month, Fig. 4), indicating substantial sub-pixel variability. As the reference run is a single point-scale simulation, the modeled FSCA timeseries of the reference run is binary. The DA results are displayed as the ensemble mean and range. The prior ensemble range containing FSCA values from 0 to 1 in mid-winter, as present for both benchmark sites in 2017, indicates that the prior ensemble contains members with low SWE which become snow-free during individual wintertime melt events or rain-on-snow (ROS) events. In case of the snow-drift (left panel in Fig. 4), the posterior ensemble does no longer contain members which become snow-free during winter, whilst it contains members which are periodically snow-free during mid-winter at the wind-exposed benchmark site (right panel in Fig. 4).

The observed and modeled GSTs are presented in Fig. 5. Compared to the reference run, the DA scheme using Sentinel-2 derived FSCA improves the representation of the local GSTs in most cases, for example in 2017 for which the RMSEmm is reduced by 2.6 °C and by 2.8 °C for the snow drift and wind-blown ridge, respectively. However, in some cases the DA slightly worsens the GST representation in comparison to the reference run, for example in 2016 for the wind-blown ridge where the RMSEmm is increased by 0.2 °C. This likely occurs due to spatial variability within Sentinel-2 pixels in the AOI which causes a resolution mismatch, so-called representation error, when comparing to point-scale observations. In some cases, this uncertainty is captured by the posterior ensemble spread, as e.g. in 2017, whilst in other cases the DA fails to capture this uncertainty and is overconfident. Whether the DA reliably offers an improvement in comparison to the reference run thus depends on the representativeness of the 10x10 m pixel for the point scale, i.e. the spatial variability of the snow cover within each pixel. Comparable results have been found for other years (Fig. S3).

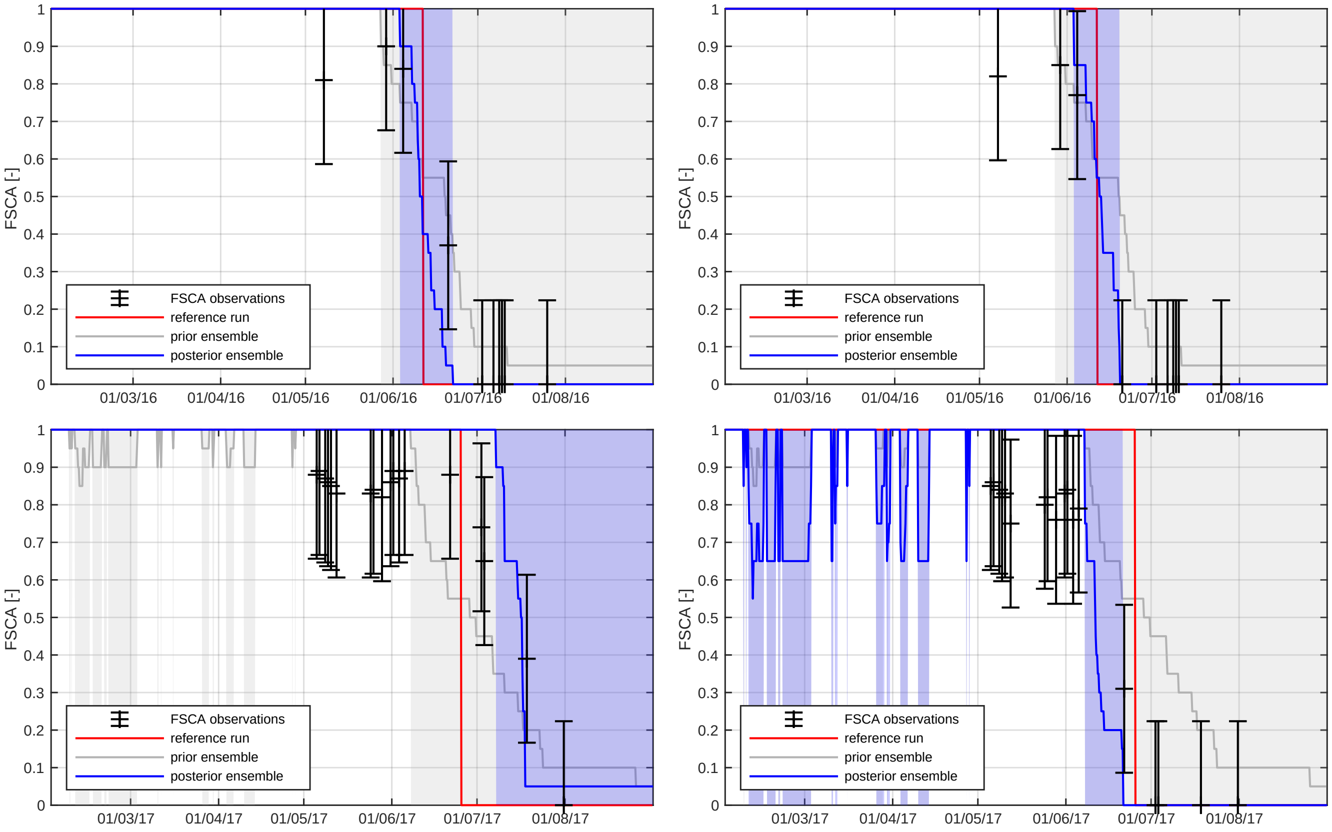

Figure 4. Sentinel-2 derived FSCA and modeled FSCA during the hydrological years of 2016 and 2017 for a snow drift location (left panels) and a wind-exposed ridge (right panels). The modeled FSCA of the reference run without DA is shown in red. The prior ensemble (before DA) is shown in gray, with the prior ensemble mean displayed as the gray solid line and the prior ensemble range as the gray shaded area. The posterior results of the DA scheme assimilating Sentinel-2 derived FSCA are shown in blue, with the blue solid line being the posterior ensemble mean and the blue shaded area the posterior ensemble range.

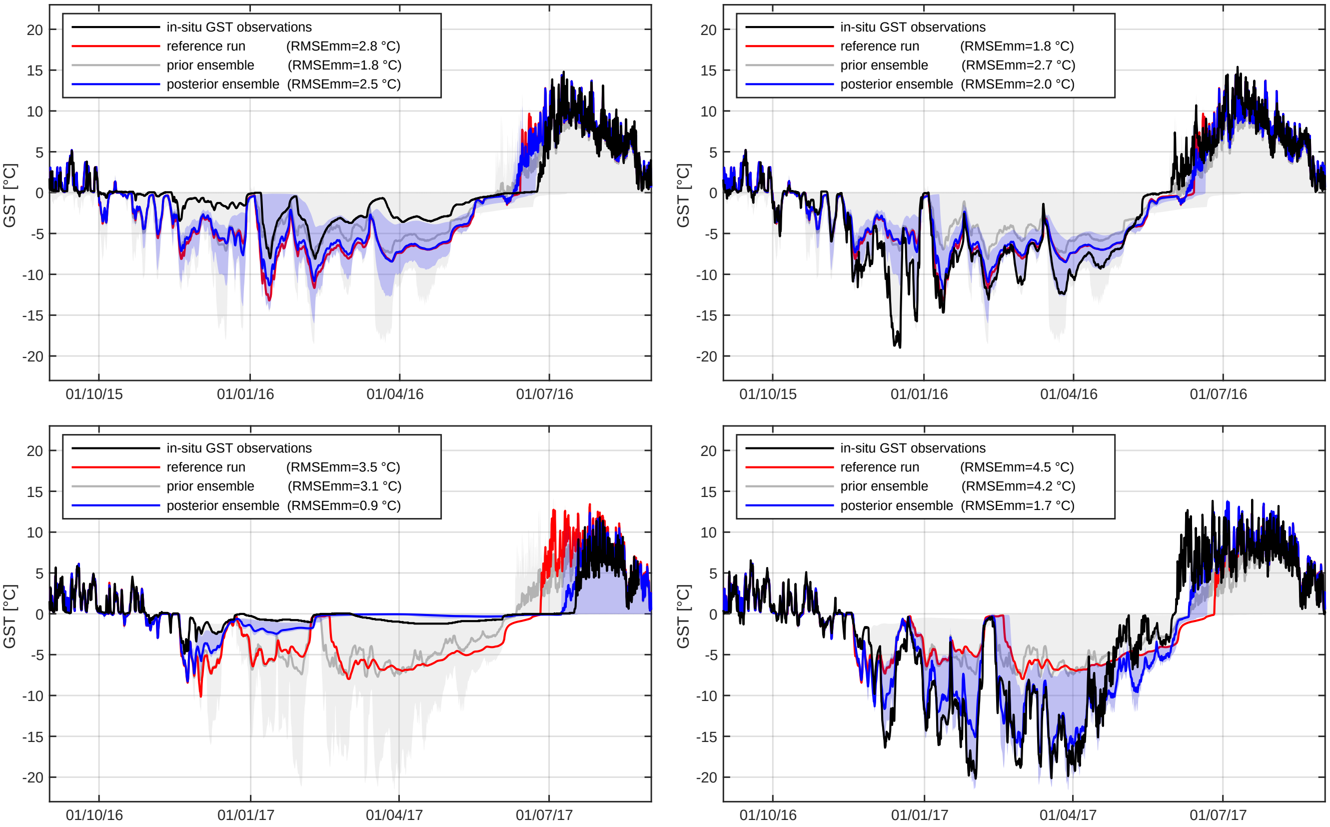

Figure 5. Observed and modeled GSTs during the hydrological years of 2016 and 2017 for a snow drift location (left panels) and a wind-exposed ridge (right panels). The modeled GST of the reference run without DA are shown in red. The prior ensemble (before DA) is shown in gray, with the prior ensemble mean displayed as the gray solid line and the prior ensemble range as the gray shaded area. The posterior results of the DA scheme assimilating Sentinel-2 derived FSCA are shown in blue, with the blue solid line being the posterior ensemble mean and the blue shaded area the posterior ensemble range. The RMSE of the monthly mean GSTs relative to the in-situ measurements is given for the reference run and the ensemble means of the prior and posterior ensembles each.

4.3 Spatial DA with Sentinel-2 derived FSCA

To assess the ability of the scheme to reproduce the spatial variability of the snowpack and GST within the AOI, we apply the DA scheme in a spatially distributed fashion as described in Sect. 3.5.

The observed and modeled SWE histograms at the measurement date around peak SWE in April/May are displayed in Fig. 6. In the studied years, the range of peak SWE typically extends from close to 0 m w.e. at wind-exposed locations to 0.315 m w.e. to 0.875 m w.e. at snow drift locations. The observed peak SWE histogram is right-skewed in most years with both the mode and median peak SWE occurring in the 0.125 to 0.25 m w.e. range for all studied years. The modeled peak SWE distributions reproduce the range and shape of the observations for most years, with the median and mode of the distribution being represented by the model scheme within an error of ± 0.125 m w.e. compared to the in-situ measurements. For most years, the model scheme produces slightly higher maximum peak SWEs than those observed, which is especially pronounced in 2017. At the same time, this tail of high peak SWEs only represents a small areal fraction of the AOI.

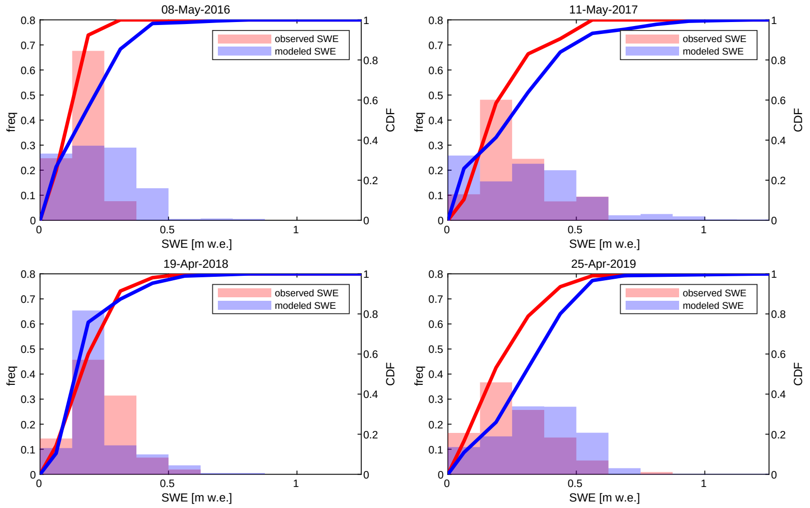

Figure 6. Observed and modeled SWE in the AOI, showing histograms of the actual values (bars) and the cumulative distribution function (CDF, lines). Observations were taken around peak SWE in April/May and the results of the spatial DA scheme with Sentinel-2 derived FSCA are presented for the same date as the observations.

The observed and modeled monthly mean GSTs (mmGST) are displayed in Fig. 7. The observed spatial variability of mmGSTs is largest during snow-covered months (typically December-April) as well as during snow-melt (typically May-June/July). During snow-free months (August and September), only a small spatial variability in mmGSTs is observed. The range and variability of winter GSTs vary considerably between years, with snow-rich years like 2017 displaying higher median mmGSTs, a distribution shifted towards higher temperatures, as well as maximum mmGSTs very close to 0 °C during the snow-covered winter months. Observed winter mmGST reach as low as -15 °C in March 2017, whilst maximum mmGST in the same month are close to 0 °C, emphasizing the variability during the snow-covered winter months. The range of observed mmGSTs is especially large in snow-rich years and the regime shift to positive mmGSTs after snow-melt occurs later. This causes wider spreads in the mmGST distributions and generally lower temperatures during the snowmelt season as seen in June 2017 compared to June 2016. The posterior mean modeled mmGST distributions align well with the observations in range, shape and median, in most cases deviating by less than 2 °C in their median mmGST from the observed median mmGSTs. Larger deviations in median and spread are present during the shoulder seasons in the beginning of the snow-covered season (typically November/December) and snow-melt (typically May-June).

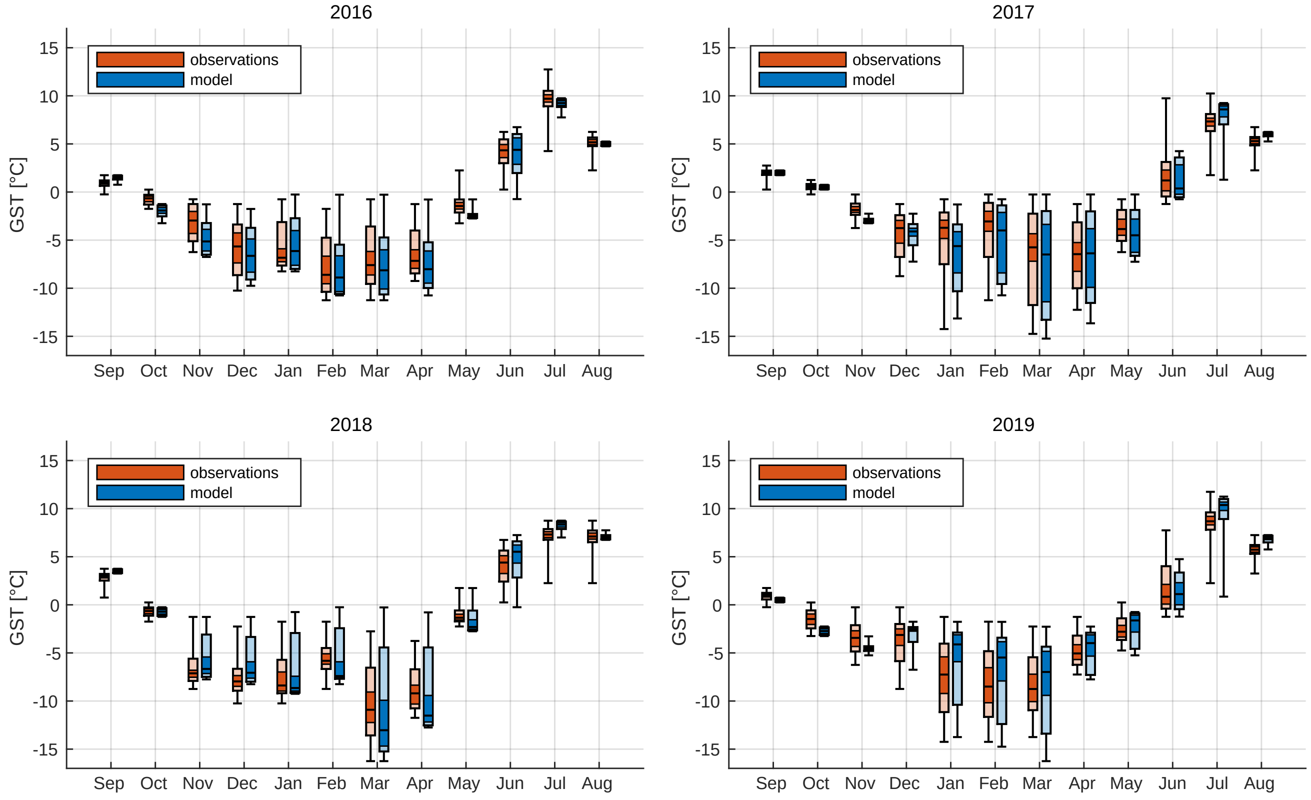

Figure 7. Monthly mean GSTs of the hydrological years 2016 - 2019. In-situ observations are shown in red, modeled GSTs in blue. The boxplots mark the 10th-90th percentile range, 25th-75th percentile range as well as the median. The whiskers mark maxima and minima.

435 The observed and modeled mean annual GSTs (MAGST) are displayed in Fig. 8. In the studied years of 2016-2019, the observations show median MAGSTs between -1.0 °C and -2.6 °C, which are well represented by the model scheme with deviations of 1 °C from the observed median. The modeled MAGST distributions generally show a slightly larger spread and are less centered than the observed distributions, but represent the observed range in MAGST between approximately 1.2 °C and -4.2 °C in the studied years well. The observed variability between extremes in MAGST lies between 4 °C and 4.5 °C, depending on the year. The observed range in MAGST generally is well reproduced by the model scheme, with larger deviations

https://doi.org/10.5194/egusphere-2025-3142
Preprint. Discussion started: 17 September 2025

for the range but not for the center of distribution for the hydrological year of 2018 only. In all years, the observations show the 90th percentile below the 0 °C MAGST marking, which is reproduced by the model.

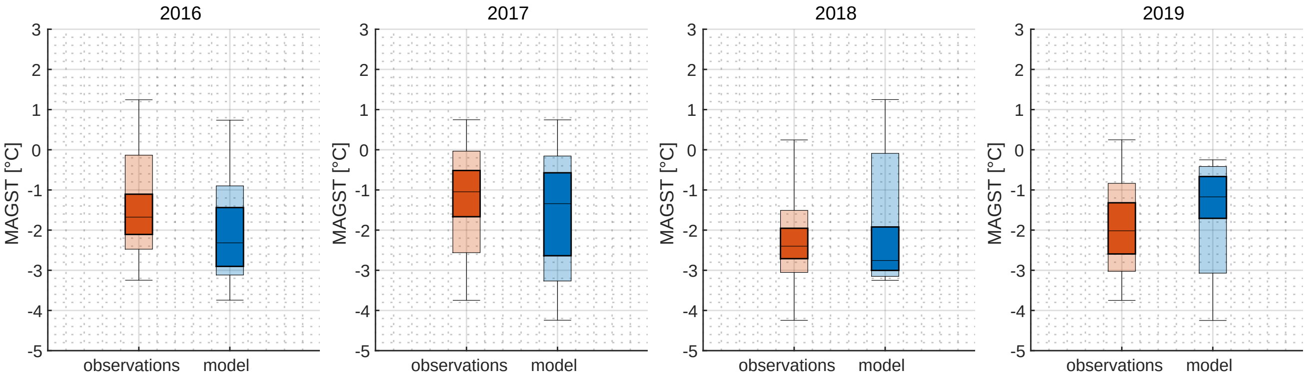

Figure 8. Mean annual GSTs of the hydrological years 2016 - 2019. In-situ observations are shown in red, modeled GSTs in blue. The boxplots mark the 10th-90th percentile range, 25th-75th percentile range and the median. The whiskers mark maxima and minima.

Figs. 9 and 10 show maps of the mean annual GST (MAGST) and the mean GST during the snow-covered winter months (December-April) over the AOI (Fig. 2) for the hydrological years of 2016-2019, as resulting from the backmapping experiment described in Sect. 3.5. In all maps, the patterns of the meltout dates derived from the Sentinel-2 scenes (see Fig. S4) directly translate to spatial differences in GST. Prominent examples are snowdrifts with late meltout (compare Fig. 2 b for the year 2017) where the algorithm yields the highest MAGST both during winter and for the annual mean. Likewise, areas with early meltout (compare Fig. 2 a for the year 2017) feature the highest GSTs. The spatial patterns resulting from the DA backmapping generally correspond well to the in-situ GST observations, typically deviating by at most a few degrees Celsius during winter (Fig. 10) and less than 1.5 °C for the annnual mean (Fig. 9). The most pronounced mismatches with observed GST commonly occur around the edges of spatial features, which may be attributed to georeferencing challenges in the Sentinel-2 scenes, sub-pixel variability of the snow meltout, as well as the limited accuracy of the logger locations (Sect. 4.2). Furthermore, Figs. 9 and Fig. 10 are evidence that the spatial detail of the obtained maps strongly varies between years, which is a result of the varying number of available Sentinel-2 FSCA observations and thus recorded differences in meltout dates (Fig. S4). For example, we obtain comparatively little spatial detail in 2016, where only extremes in GST (wind-blown ridge with early meltout vs. snowdrift with late meltout) can be distinguished. In 2018, more spatial detail in the meltout dates is obtained from the available Sentinel-2 scenes, resulting in a more detailed representation of the GST variability.

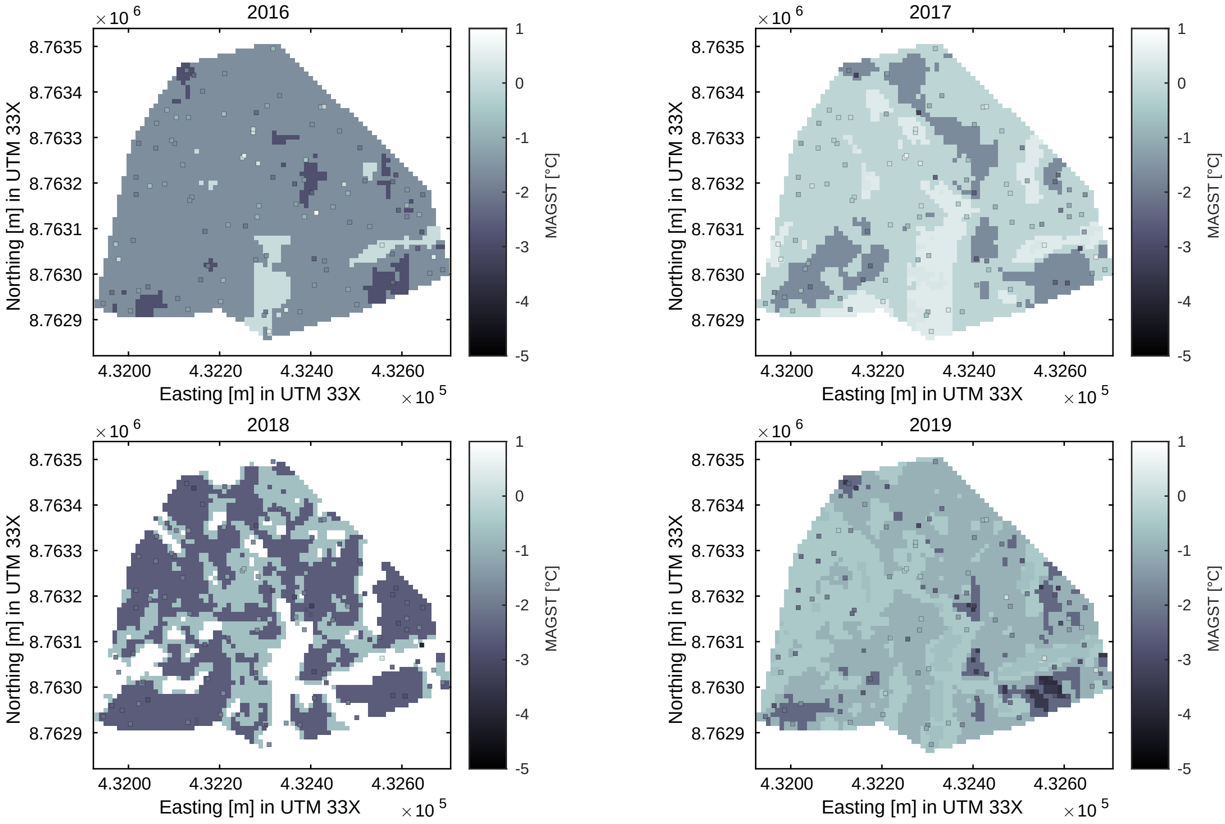

Figure 9. Maps of the mean annual GSTs (MAGST) in the AOI (see Fig. 1 c) for the hydrological years 2016-2019, obtained from the backmapping experiment (Sect. 3.5). The symbols display MAGST from in-situ observations (Sect. 3.4).

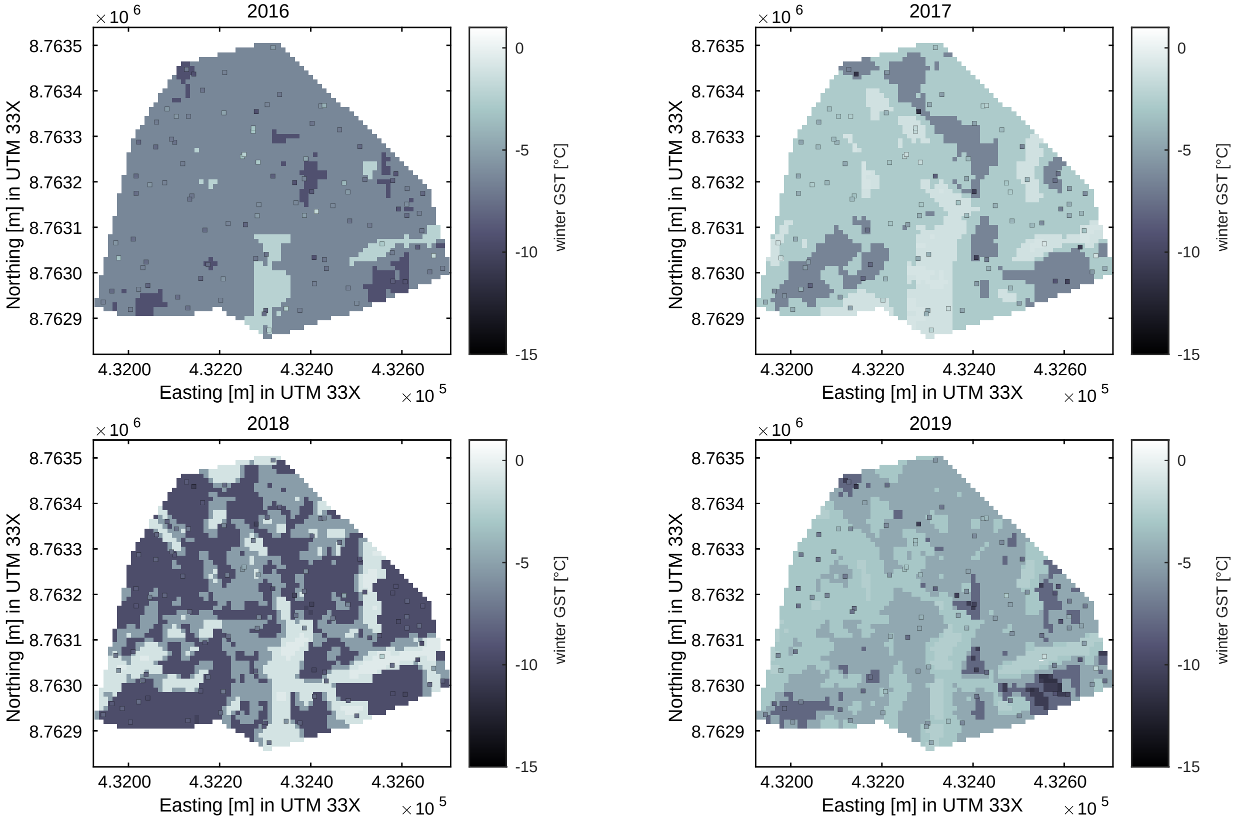

Figure 10. Maps of mean winter GST (snow-covered winter months December-April) in the AOI (see Fig. 1 c) for the hydrological years 2016-2019, obtained from the backmapping experiment (Sect. 3.5). The symbols display mean winter GST from in-situ observations (Sect. 3.4).

5 Discussion

5.1 Challenges and Limitations

A key limiting factor in our study is the spatial and temporal resolution of the assimilated observational data. The availability of satellite-derived FSCA observations is limited by cloudiness which causes the temporal resolution and thus the performance of the DA scheme to vary between individual years. While Svalbard is a particularly cloudy region, this limitation can be expected to persist in most parts of the Arctic. However, our study suggests that even in years with few cloud-free Sentinel-2 scenes (e.g. 2016), it is possible to distinguish the extremes of the snow distribution with the presented DA scheme, thus providing an advantage over the unconstrained reference run. A further limitation arises from biases in the geolocation (Pandžic et al., 2016) and co-registration (Stumpf et al., 2018) of the Sentinel-2 scenes, which mainly affects the point-scale DA for individual Sentinel-2 pixels. In addition, we found that subpixel variability of snowmelt poses challenges even for high-resolution

https://doi.org/10.5194/egusphere-2025-3142
Preprint. Discussion started: 17 September 2025

Sentinel-2 pixels (10x10 m), which could also be a problem when applying the method to important permafrost landforms with variable snow cover, such as tundra polygons (Nitzbon et al., 2019) or peat plateaus (Martin et al., 2019). These challenges posed by the observations are a major limiting factors for satellite-based DA at pixel resolution in our study, as described in Sect.4.2. The method can be expected to be feasible in different regions, though, if the topographic complexity occurs on scales larger than the observational footprint, and more frequent and correctly registered observations are available.

The comparison of in-situ GST measurements with modeled GST data in space (Sect. 4.3, Figs. 9, 10) suggests that snowdrift areas with late snowmelt and consequently high MAGST may be insufficiently represented in the GST validation dataset. Although the temperature sensors were installed in a randomized fashion (Gisnås et al., 2014), they may not fully capture the true areal distributions of the various snow regimes, especially for snowdrift areas. This could explain the apparent over-representation of high SWE in the spatially distributed DA model results (Fig. 6), indicating that the high SWE values obtained in the DA scheme may indeed have occurred in the AOI, but were not captured by the point-based validation data set.

Another challenge is constraining unknown model parameters. This is always a challenge in process-rich, complex models, and requires finding a balance between complexity and improved results (Sect. 5.3). In our study, the results are sensitive to the description of albedo development and surface roughness of the snowpack, both of which can be temporarily (event-based) as well as spatially variable within the AOI. Such dynamics cannot be constrained within the DA and the corresponding model parameters were thus set to be constant in both space and time. However, we emphasize that the same limitations apply to the reference run, in which albedo and roughness parameters also must be assumed constant. Furthermore, the performance of the DA scheme is influenced by the accuracy of the meteorological data employed to force the CryoGrid model. Individual weather events like winter-time melt or rain-on-snow events are challenging to represent accurately, while strongly impacting snow properties not only during the event itself, but also in the following period (e.g. Westermann et al., 2011). This is an unavoidable issue commonly faced in land surface modeling, which is partially mitigated by applying a bias-correction using local weather station data to the downscaled ERA5 reanalysis.

Another major limitation of our study were computational costs. While a forward model run required $N_{\text{CPUhours/yr}} = 0.2 - 4$ CPU hours per modeled year, depending on the meteorological forcing and snow conditions, a DA run requires $N_{\text{ensemble}} \cdot N_{\text{iteration}} \cdot N_{\text{CPU hours/yr}}$, thus increasing the computational costs by up to two orders of magnitude compared to the reference run. This limited the manageable initial ensemble size and impacted the convergence of the DA scheme which was moderated by the choice of an iterative DA scheme (Sect. 3.3). Due to the computational costs, it is not feasible to upscale the pixel-based assimilation of the satellite observations (Sect. 4.2) at this resolution over larger areas. To obtain pixel-resolved results, the method has to be combined with clustering approaches (see Sect. 5.5) or a simpler, computationally cheaper forward model.

5.2 Spatial variability of snow cover and ground thermal regime

As shown by Gisnås et al. (2014) and Zweigel et al. (2021), the seasonal snow cover is the largest source of spatial variability of MAGST in the AOI. In the studied years, observations in the AOI reveal a spread in MAGST of around 4.5 °C, which highlights the importance of small-scale spatial variability of GST in the area. Most of this variability accumulates during the snow-covered winter months, with spatial differences in mmGSTs of up to 15 °C (e.g. in March 2017). Other factors contributing to GST variability in the AOI are variability in soil moisture conditions, ground stratigraphy and vegetation cover (Boike et al., 2018), which especially contribute to the variability of summer GST. Whilst the MAGST is a crucial variable for assessing the occurrence of permafrost, the seasonality of GST is important for the thermal regime in the uppermost soil layers, as well as the active layer thickness. Despite substantial increases in summer air temperature, observations at the Bayelva permafrost borehole located within the AOI reveal only little permafrost warming in recent years (Grünberg et al., 2024). This compensation of increased summer GST is attributed to changes in snow cover causing less insulation during cold winter months, i.e. a reduction the of snow depth and snow duration, which underlines the complex interaction of various climatic factors (Grünberg et al., 2024). The compensation of summer warming can be expected to most strongly affect locations with intermediate snow cover, such as the Bayelva borehole site. Locations with deeper snow cover remain thermally insulated from cold winter air temperatures even with decreasing snow cover, potentially making them more prone to permafrost degradation. This highlights the limited representativeness of individual permafrost boreholes in heterogeneous permafrost landscapes and emphasizes the importance of considering small-scale spatial variability when assessing climatic trends in the ground thermal regime, both for in-situ measurements and in numerical models.

Snow and GST variability additionally impact local hydrology, such as the timing and magnitude of snowmelt runoff, as well as the carbon cycle and human settlements (e.g. Anderton et al., 2002; Pirk et al., 2023; Aga et al., 2025). Snow properties such as internal and basal ice layers from ROS events exhibit large spatial variability and impact underlying vegetation as well as the feeding behavior of herbivores (Hansen et al., 2011). Whilst the importance of the snowpack on the surface energy balance and hydrology is well studied (e.g. McCabe and Clark, 2005; Westermann et al., 2009; Langer et al., 2011), the representation of its spatial variability in complex terrain remains a significant challenge in models (e.g. Freudiger et al., 2017).

5.3 Approaches for representation of snow spatial variability

Key processes influencing the spatial heterogeneity of the snowpack include variations in snowfall and precipitation rates (typically on scales of kilometers), variations in the surface energy balance and thus snow sublimation and melt (tens of meters to kilometers), avalanches (hundreds of meters), as well as snow drifting due to wind (tens of meters to kilometers). Explicit representation of wind-driven redistribution of snow in complex terrain requires multidimensional models and comes at high computational costs, limiting their application to case studies, intermediate complexity snow-models, or small areas. Examples of such studies are Marsh et al. (2024) using an intermediate complexity snow-model on resolutions down to 50 m as well as Zweigel et al. (2021) using the CryoGrid model as a laterally coupled land surface model with explicit representation of wind redistribution of snow.

A commonly used alternative to represent spatial variability in the snowpack is assimilating satellite-retrieved snow observations into physical snow models. This facilitates accounting for small-scale variability on the scale of the satellite resolution using computationally less expensive, one-dimensional model schemes which do not explicitly exchange snow with neighboring pixels. Such models can be applied for individual pixels (e.g. Aalstad et al., 2018), in a spatially distributed manner (e.g. Margulis et al., 2016) or for a number of representative sites in combination with clustering (e.g. Fiddes et al., 2019).

All of these approaches primarily focus on the correct representation of snow depth and SWE. To study the impact of the snowpack on the ground thermal regime, the insulating properties of the snowpack must be represented, which requires a more sophisticated representation of the snowpack, e.g. through multiple layers, transient evolution of snow microphysics and density, as well as albedo. Therefore, the assimilation of satellite-retrieved snow observations into land-surface models like CryoGrid presents significant potential (Reynolds et al., 2024). In particular, this study shows that data assimilation is indeed capable of inferring "hidden" variables related to the ground thermal regime, such as GST, which are inaccessible through direct satellite observations.

5.4 Data Assimilation in complex Land Surface Models

The practical implementation of DA techniques in complex models introduces significant challenges primarily centered around the computational constraints with a potentially large number of model parameters. In reality, it is impractical to constrain all model parameters through the data assimilation exercise. Instead, a limited number of parameters which are sensitive to the assimilated observations needs to be selected, whilst the other parameters are held fixed. In our study, we chose to perturb and thereby constrain two critical parameters within the DA: the exposure to snow drifting linked to topography and a wind scaling factor adjusting high near-surface wind speeds. The majority of model parameters in CryoGrid and the meteorological forcing were set to be constant across the ensemble and are calibrated using traditional methods, i.e. by adopting literature values or manually adjusting parameters to fit in-situ measurements. Introducing a large number of degrees of freedom (i.e. many unconstrained parameters) in the DA is generally not feasible, as the assimilated FSCA is insensitive to many model parameters and the DA scheme therefore cannot constrain them. Furthermore, this would require a significantly larger ensemble and computational resources, which in practice is the main limiting factor.

Particularly in non-iterative particle DA schemes (van Leeuwen et al., 2019), such as the Particle Batch Smoother (Margulis et al., 2015), the prior ensemble must sample the entire space of possible parameter combinations. This causes the number of required particles to avoid degeneracy, and thus total computation time, to scale exponentially with the number of perturbed parameters due to the so-called curse of dimensionality (Murphy, 2023). While not resolving this general problem, using iterative DA schemes such as the presented AdaPBS algorithm can considerably reduce the required computation time by adapting the particle ensemble (Bugallo et al., 2017). The main idea is to iteratively zoom into the promising regions of the parameter space, thus eliminating the need to densely sample less promising regions. In particular, AdaPBS takes the results of previous iterations into account using adaptive multiple importance sampling (Cornuet et al., 2012). As such, the total number of ensemble members grows with each iteration, until a stopping criterion is reached. Therefore, AdaPBS efficiently leverages an adaptive ensemble of particles to better fit the requirements of the specific application case and thereby reduces the computation time compared to costly non-iterative schemes.

To further reduce computation time, machine learning-based emulators are a promising way forward to improve data assimilation (Keetz et al., 2025). Such emulators serve as approximations of the primary model, tailored for a specific range of model parameters and model forcing conditions. They can accelerate computation by generating the synthetic observations required for DA, or even directly delivering key DA terms, as a function of the model parameters included in the DA. For the case of this study, an emulator would have to deliver time series of FSCA, or a summary of how well they match the true observations (such as the nonlinear Gaussian likelihood in Eq. 4), as a function of exposure and wind scaling factor. As the emulator is significantly faster than a CryoGrid forward simulation, a much denser ensemble of model parameters could be obtained, thus making it easier to identify the parameter regions that are in agreement with the observations. However, establishing reliable emulators is not trivial, as they must be trained to accurately reflect the model output across the possible range of multiple model parameters. Consequently, if the DA scheme incorporates many target parameters and/or observables, training these emulators could prove to be as time-consuming as the original DA process due to the large number of ensemble runs with the full model that would be required to generate the training data. Therefore, while emulators may partially alleviate computational constraints, they are unlikely to fully resolve the computational challenges associated with scaling DA in complex land-surface models. A promising way forward could be the use of active learning (e.g. Alsing et al., 2019) where uncertainty-aware emulators are used to design new full model runs which are then in turn used to train the emulator in a virtuous cycle to improve inference.

5.5 Outlook and future work

In this study, we demonstrate the assimilation of FSCA in the ground thermal model CryoGrid for individual hydrological years, which is sufficient for reproducing GST. However, the ground thermal regime in greater depths is the result of thermal forcing over multiple years. For point-scale simulations with single Sentinel-2 pixels (Sect. 4.3), it is straight-forward to extend the presented scheme to multiple years by applying the DA sequentially year after year. Model states (e.g. ground temperatures) resulting from the successful application of the DA in one year can then be used as the initial states for the ensemble simulations of the following year. In such multi-annual (i.e. sequential) experiments, the ensemble parameters can either be considered independent between individual years, sampling from the same prior distributions every year, or continuously adjusted to learn from the posterior parameter distributions of the previous years. The latter has the potential to direct the DA to the most promising regions of the parameter space already in the first iteration, thus reducing the number of iterations required to achieve a sufficiently diverse posterior ensemble. In the case of snow DA, such a learning strategy seems promising as the spatial snow patterns are often related to terrain properties and thus are similar between years (e.g. Kępski et al., 2017), which is also evident in the AOI (Sect. 4.3).

For spatially distributed applications as presented in Sect. 4.3, assimilating FSCA observations over consecutive years is more challenging. As it is too computationally expensive to apply the DA scheme for all individual pixels in the AOI, a clustering approach (e.g. Fiddes and Gruber, 2012) based on entire FSCA time series of the target period seems most promising, capitalizing on the inter-annual consistency of snow distribution patterns. The single-pixel DA scheme would be applied to individual locations representative for clusters in the variability of the meltout. The FSCA variability within individual clusters could be employed as observation uncertainty into the DA, which likely results in a larger spread of the posterior ensemble representing the spatial variability of the target variables within the cluster. Note that such procedures are compatible with postprocessing of ground temperature distributions through interpolation between individual meltout dates, as demonstrated in this study for individual years (Figs. 7, 8).

https://doi.org/10.5194/egusphere-2025-3142
Preprint. Discussion started: 17 September 2025

Within our largely flat, high-latitude study area, it was not necessary to account for topography- or land cover-induced variability of the surface energy balance and thus snowmelt rates. However, this can be a critical process governing the melt patterns in other regions in addition to snow redistribution targeted in the DA procedure. For individual pixels, it is straight-forward to adjust e.g. the incoming radiation for the exposition (e.g. Fiddes and Gruber, 2014). For spatially distributed applications, it could be possible to include topographic factors as additional variables in an FSCA clustering procedure, as described above. In this case, pixels with similar meltout dates would be represented by different clusters if they feature different expositions.

A major limitation of the presented DA scheme is the lack of a dense time series of satellite-derived FSCA due to frequent cloudiness. This introduces uncertainty in the DA results (e.g. GST) and for spatially distributed applications effectively limits the spatial detail of the obtained GST maps (Sect. 4.3). A way to moderate this problem and/or constrain the ensemble further is by assimilating complementary observations in addition. Promising additional observations are for example remotely sensed snow depth, shortwave reflectance or microwave backscatter from active microwave remote sensors. Snow depth observations as input for snow DA could be obtained from the ICESat-2 laser altimeter (Mazzolini et al., 2024), but the spatial coverage is limited to narrow tracks, while the measurements are similarly affected by cloudiness. Shortwave reflectance in different spectral bands is provided by the Sentinel-2 observations used in this study, and whilst it suffers from the same cloudiness limitations, it could constrain e.g. snow grain properties, the albedo evolution and the moisture content in the snow. Microwave backscatter is available through e.g. Sentinel-1 at a similarly high resolution as the Sentinel-2 observations used in this study, and its availability is independent of cloudiness. Under ideal conditions, the Sentinel-1 constellation can provide a revisit time of 6 days, which at least during prolonged periods of cloudiness during the snowmelt (e.g. in 2016) could provide additional information. However, the modification of the radar backscatter signal between melting snow and snow-free ground is much less defined than for the reflectivity in the visible electromagnetic spectrum, being strongly influenced by the moisture conditions of the ground after snowmelt. Therefore, assimilating microwave backscatter data is highly challenging, although it may be possible to ingest snow parameters computed by CryoGrid (such as grain size, density, water content) into microwave transfer models (e.g. Picard et al., 2018) to obtain synthetic observations of the actual radar backscatter signal, which could subsequently be ingested in the data assimilation procedure. While it remains unclear whether microwave backscatter can indeed add information on the snowmelt progression, it might be useful to constrain other parts of the land-surface model, especially changes in soil moisture conditions during the snow-free period (Kornelsen and Coulibaly, 2013), or the occurrence of wintertime rain and melt events (Bartsch et al., 2023). Furthermore, model parameters related to subsurface properties could be constrained by assimilating InSAR (Interferometric Synthetic Aperture Radar) retrievals of vertical surface heave and subsidence from active microwave sensors such as Sentinel-1, which contain information on the ground ice content (e.g. Wendt et al., 2024). This could contribute to a more comprehensive assessment of key permafrost variables, such as the active layer thickness. The presented DA workflow is in principle designed to assimilate a range of different observations simultaneously. In practice, such multi-sensor data fusion would make it necessary to include additional parameters in the DA procedure, potentially resulting in a significant increase of the computational costs (Sect. 5.4), thus compromising the practical feasibility. However, at least in some applications, the cross-sensitivity of different observational data sets may be relatively low, so that

https://doi.org/10.5194/egusphere-2025-3142
Preprint. Discussion started: 17 September 2025

each data set could constrain its own set of model parameters without strongly affecting the other observations and associated model parameters.

6 Conclusions

In this study, we present a data assimilation scheme which ingests satellite-derived fractional snow cover at 10 m resolution into the process-rich permafrost model CryoGrid, resulting in a much improved characterization of the spatial variability of snow cover and ground surface temperatures. For this purpose, we employ an iterative ensemble-based data assimilation scheme (Adaptive Particle Batch Smoother - AdaPBS) which is efficient enough to handle the significant computational costs of the CryoGrid model. The results are evaluated against spatially distributed measurements of ground surface temperature and snow water equivalent conducted for four years in a study area of about 0.5 km^2. The scheme is tested at point-scale using evaluation sites representing extremes in the snow distribution, i.e. an exposed ridge site with little snow and a snowdrift. We demonstrate the scheme's performance by assimilating idealized, synthetic observations of FSCA derived from the in-situ GST measurements. Furthermore, we apply the scheme with real FSCA observations derived from Sentinel-2, both at the scale of individual pixels and in a spatially distributed fashion, with the goal to simulate the distributions of ground surface temperature and snow water equivalent in the entire study area. The capability of the data assimilation experiments to reproduce the in-situ observations is compared against a reference simulation, i.e. a standard forward run of the model, which does not account for wind redistribution of snow. From this study, the following conclusions can be drawn:

- For point-scale simulations at sites representing extremes in snow cover (snowdrift, exposed ridge), assimilating time series of fractional snow-covered area significantly improves the representation of the annual GST cycle compared to the reference run, with GST especially improved during winter and during the snowmelt period when GST is isothermal around 0 °C.

- At both sites, the RMSE for monthly average ground surface temperatures (evaluated against the in-situ measurements) improved by up to 3 °C compared to the reference simulation. While this improvement occurred for both the assimilation of in-situ and Sentinel-2 derived FSCA observations, the improvement is less consistent for the latter, with the data assimilation procedure even leading to a worse performance in some of the years and locations.

- The spatially distributed setup successfully reproduces the annual cycle of the GST variability within the study area, with a high spread in the winter months ($>$ 10 °C differences in mean monthly GST) and during snowmelt ($>$ 5 °C differences), while the spatial GST variability was lower ($<$ 3 °C differences) during the snow-free periods.

- High-resolution maps of ground surface temperatures for the study area can be obtained by backmapping on individual Sentinel-2 pixels, using the relationship between the timing of the meltout and GST obtained from the DA scheme. The resulting maps can represent the thermal impact of key features in the snow distribution, such as snowdrifts and ridges with shallow snow cover, but the spatial detail is limited by the number of cloud-free satellite scenes within the snowmelt period.

Our study demonstrates the potential of modern data assimilation techniques for integrating satellite observations into numerical models for land surface processes. In particular, we have been able to recover ground surface temperatures, which is a key environmental variable in permafrost areas, from a priori unrelated satellite observations of fractional snow-covered area. The study highlights how a tight integration of observations and models can advance our capacity to understand, monitor, and project the state of complex Arctic systems in a changing climate.

Code availability. The CryoGrid model code is available at https://github.com/CryoGrid/CryoGridCommunity_run (last access: 1 June 2025).

Author contributions. CW, KA and SW conceptualized the research. KA developed the AdaPBS scheme, with input from CW and SW. CW and SW implemented and tested the AdaPBS scheme in CryoGrid. CW conducted and analyzed the data assimilation experiments. CW prepared the manuscript, and all co-authors revised and edited the final version.

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

Financial support. This research has been supported by the European Space Agency Permafrost_cci (grant no. 4000123681/18/I-NB), as well as the European Union through grant no. 101184962 (LIQUIDICE). We acknowledge financial support by the Faculty for Mathematics and Natural Sciences at the University of Oslo, Norway, through the ACT-Pilot project.

Acknowledgements. The study uses modified Copernicus Sentinel data (years 2016-2019) obtained from the Google Earth Engine. The simulations were performed on resources provided by Sigma2 - the National Infrastructure for High-Performance Computing and Data Storage in Norway through projects NN8107K and NS9607K. The results contain modified Copernicus Climate Change Service information 2020. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains.

https://doi.org/10.5194/egusphere-2025-3142
Preprint. Discussion started: 17 September 2025

References

Aalstad, K., Westermann, S., Schuler, T. V., 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.

Aga, J., Willmes, C., Sinitsyn, A. O., Arlov, T. B., Boike, J., and Westermann, S.: Impact of snow and building management on ground surface temperatures in permafrost environments-A case study from the historical mining town Ny-Ålesund, Svalbard, Cold Regions Science and Technology, p. 104516, 2025.

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.

Alsing, J., Charnock, T., Feeney, S., and Wandelt, B.: Fast likelihood-free cosmology with neural density estimators and active learning, Monthly Notices of the Royal Astronomical Society, 488, 4440–4458, https://doi.org/10.1093/mnras/stz1960, 2019.

Anderton, S., White, S., and Alvera, B.: Micro-scale spatial variability and the timing of snow melt runoff in a high mountain catchment, Journal of Hydrology, 268, 158–176, 2002.

Bair, E., Stillinger, T., and Dozier, J.: A Generalized Multispectral Unmixing Approach With Examples From MODIS and Landsat 8 OLI, IEEE Transactions on Geoscience and Remote Sensing, 59, 7270–7284, https://doi.org/10.1109/TGRS.2020.3040328, 2020.

Bartsch, A., Bergstedt, H., Pointner, G., Muri, X., Rautiainen, K., Leppänen, L., Joly, K., Sokolov, A., Orekhov, P., Ehrich, D., et al.: Towards long-term records of rain-on-snow events across the Arctic from satellite data, The Cryosphere, 17, 889–915, 2023.

Biskaborn, B. K., Smith, S. L., Noetzli, J., Matthes, H., Vieira, G., Streletskiy, D. A., Schoeneich, P., Romanovsky, V. E., Lewkowicz, A. G., Abramov, A., et al.: Permafrost is warming at a global scale, Nature communications, 10, 264, 2019.

Boike, J., Juszak, I., Lange, S., Chadburn, S., Burke, E., Overduin, P. P., Roth, K., Ippisch, O., Bornemann, N., Stern, L., et al.: A 20-year record (1998–2017) of permafrost, active layer and meteorological conditions at a high Arctic permafrost research site (Bayelva, Spitsbergen), Earth System Science Data, 10, 355–390, 2018.

Brown, J., Jr., O. J. F., Heginbottom, J. A., and Melnikov, E. S.: Circum-Arctic map of permafrost and ground-ice conditions, https://pubs.usgs.gov/publication/cp45, report number 45, 1997.

Bugallo, M. F., Elvira, V., Martino, L., Luengo, D., Miguez, J., and Djuric, P. M.: Adaptive Importance Sampling: The past, the present, and the future, IEEE Signal Processing Magazine, 34, 60–79, https://doi.org/10.1109/MSP.2017.2699226, 2017.

Cao, W., Aalstad, K., Schmidt, L. S., Westermann, S., and Schuler, T. V.: Glacier data assimilation on an Arctic glacier: Learning from large ensemble twin experiments, arXiv, https://doi.org/10.48550/arXiv.2502.09314, preprint submitted to Journal of Glaciology (in review), 2025.

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, 2018.

Cluzet, B., Magnusson, J., Quéno, L., Mazzotti, G., Mott, R., and Jonas, T.: Exploring how Sentinel-1 wet-snow maps can inform fully distributed physically based snowpack models, The Cryosphere, 18, 5753–5767, https://doi.org/10.5194/tc-18-5753-2024, 2024.

Cornuet, J.-M., Marin, J.-M., Mira, A., and Robert, C. P.: Adaptive Multiple Importance Sampling, Scandinavian Journal of Statistics, 39, 798–812, https://doi.org/10.1111/j.1467-9469.2011.00756.x, 2012.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., et al.: Sentinel-2: ESA's optical high-resolution mission for GMES operational services, Remote sensing of Environment, 120, 25–36, 2012.

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

Elvira, V., Martino, L., and Robert, C. P.: Rethinking the Effective Sample Size, International Statistical Review, 90, 525–550, https://doi.org/10.1111/insr.12500, 2022.

Evensen, G., Vossepeol, F. C., and van Leeuwen, P. J.: Data Assimilation Fundamentals, Springer, https://doi.org/10.1007/978-3-030-96709-3, 2022.

Fair, Z., Vuyovich, C., Neumann, T., Pflug, J., Shean, D., Enderlin, E. M., Zikan, K., Besso, H., Lundquist, J., Deschamps-Berger, C., and Treichler, D.: Review article: Using spaceborne lidar for snow depth retrievals: Recent findings and utility for global hydrologic applications, EGUsphere, 2025, 1–35, https://doi.org/10.5194/egusphere-2024-3992, 2025.

Fang, Y., Liu, Y., Li, D., Sun, H., and Margulis, S. A.: Spatiotemporal snow water storage uncertainty in the midlatitude American Cordillera, The Cryosphere, 17, 5175–5195, https://doi.org/10.5194/tc-17-5175-2023, 2023.

Fiddes, J. and Gruber, S.: TopoSUB: a tool for efficient large area numerical modelling in complex topography at sub-grid scales, Geoscientific Model Development, 5, 1245–1257, 2012.

Fiddes, J. and Gruber, S.: TopoSCALE v. 1.0: downscaling gridded climate data in complex terrain, Geoscientific Model Development, 7, 387–405, 2014.

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, 2022.

Førland, E. J. and Hanssen-Bauer, I.: Past and future climate variations in the Norwegian Arctic: overview and novel analyses, Polar research, 22, 113–124, 2003.

Freudiger, D., Kohn, I., Seibert, J., Stahl, K., and Weiler, M.: Snow redistribution for the hydrological modeling of alpine catchments, Wiley Interdisciplinary Reviews: Water, 4, e1232, 2017.

Girotto, M., Musselman, K. N., and Essery, R. L. H.: Data Assimilation of Seasonal Snow, in: Applications of Data Assimilation and Inverse Problems in the Earth Sciences, edited by Ismail-Zadeh, A., Castelli, F., Jones, D., and Sanchez, S., Special Publications of the International Union of Geodesy and Geophysics, p. 79–92, Cambridge University Press, https://doi.org/10.1017/9781009180412.006, 2023.

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

Gisnås, K., Westermann, S., Schuler, T. V., Litherland, T., Isaksen, K., Boike, J., and Etzelmüller, B.: A statistical approach to represent small-scale variability of permafrost temperatures due to snow cover, The Cryosphere, 8, 2063–2074, 2014.

Groenke, B., Langer, M., Nitzbon, J., Westermann, S., Gallego, G., and Boike, J.: Investigating the thermal state of permafrost with Bayesian inverse modeling of heat transfer, The Cryosphere, 17, 3505–3533, https://doi.org/10.5194/tc-17-3505-2023, 2023.

Grünberg, I., Groenke, B., Westermann, S., and Boike, J.: Permafrost and active layer temperature and freeze/thaw timing reflect climatic trends at Bayelva, Svalbard, Journal of Geophysical Research: Earth Surface, 129, e2024JF007 648, 2024.

Gualtieri, G.: Reliability of ERA5 reanalysis data for wind resource assessment: A comparison against tall towers, Energies, 14, 4169, 2021.

Hansen, B. B., Aanes, R., Herfindal, I., Kohler, J., and Sæther, B.-E.: Climate, icing, and wild arctic reindeer: past relationships and future prospects, Ecology, 92, 1917–1923, 2011.

Hanssen-Bauer, I., Førland, E., Hisdal, H., Mayer, S., Sandø, A. B., Sorteberg, A., et al.: Climate in svalbard 2100, A knowledge base for climate adaptation, 470, 2019.

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, 2020.

Ismail-Zadeh, A., Castelli, F., Jones, D., and Sanchez, S.: Inverse Problems and Data Assimilation in Earth Sciences, in: Applications of Data Assimilation and Inverse Problems in the Earth Sciences, edited by Ismail-Zadeh, A., Castelli, F., Jones, D., and Sanchez, S., Special Publications of the International Union of Geodesy and Geophysics, p. 3–8, Cambridge University Press, https://doi.org/10.1017/9781009180412.002, 2023.

Kalnay, E., Mote, S., and Da, C.: Earth System Modeling, Data Assimilation and Predictability: Atmosphere, Oceans, Land and Human Systems, Cambridge University Press, 2 edn., https://doi.org/10.1017/9780511920608, 2024.

Keetz, L. T., Aalstad, K., Fisher, R. A., Poppe Terán, C., Naz, B., Pirk, N., Yilmaz, Y. A., and Skarpaas, O.: Inferring Parameters in a Complex Land Surface Model by Combining Data Assimilation and Machine Learning, Journal of Advances in Modeling Earth Systems, 17, e2024MS004 542, https://doi.org/10.1029/2024MS004542, 2025.

Kępski, D., Luks, B., Migała, K., Wawrzyniak, T., Westermann, S., and Wojtuń, B.: Terrestrial remote sensing of snowmelt in a diverse high-Arctic tundra environment using time-lapse imagery, Remote Sensing, 9, 733, 2017.

Kornelsen, K. C. and Coulibaly, P.: Advances in soil moisture retrieval from synthetic aperture radar and hydrological applications, Journal of Hydrology, 476, 460–489, 2013.

Langer, M., Westermann, S., Muster, S., Piel, K., and Boike, J.: The surface energy balance of a polygonal tundra site in northern Siberia–Part 1: Spring to fall, The Cryosphere, 5, 151–171, 2011.

Liston, G. E. and Elder, K.: A distributed snow-evolution modeling system (SnowModel), Journal of Hydrometeorology, 7, 1259–1276, 2006.

MacKay, D. J. C.: Information Theory, Inference, and Learning Algorithms, Cambridge University Press, 2003.

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, 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, 2016.

Marsh, C. B., Lv, Z., Vionnet, V., Harder, P., Spiteri, R. J., and Pomeroy, J. W.: Snowdrift-permitting simulations of seasonal snowpack processes over large mountain extents, Water Resources Research, 60, e2023WR036 948, 2024.

Martin, L. C. P., Nitzbon, J., Aas, K. S., Etzelmüller, B., Kristiansen, H., and Westermann, S.: Stability conditions of peat plateaus and palsas in northern Norway, Journal of Geophysical Research: Earth Surface, 124, 705–719, 2019.

Maturilli, M., Herber, A., and König-Langlo, G.: Climatology and time series of surface meteorology in Ny-Ålesund, Svalbard, Earth System Science Data, 5, 155–163, 2013.

Mazzolini, M., Aalstad, K., Alonso-González, E., Westermann, S., and Treichler, D.: Spatio-temporal snow data assimilation with the ICESat-2 laser altimeter, EGUsphere, 2024, 1–29, https://doi.org/10.5194/egusphere-2024-1404, 2024.

McCabe, G. J. and Clark, M. P.: Trends and variability in snowmelt runoff in the western United States, Journal of Hydrometeorology, 6, 476–482, 2005.

Morlighem, M. and Goldberg, D.: Data Assimilation in Glaciology, in: Applications of Data Assimilation and Inverse Problems in the Earth Sciences, edited by Ismail-Zadeh, A., Castelli, F., Jones, D., and Sanchez, S., Special Publications of the International Union of Geodesy and Geophysics, p. 93–111, Cambridge University Press, https://doi.org/10.1017/9781009180412.007, 2023.

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

Nitzbon, J., Langer, M., Westermann, S., Martin, L., Aas, K. S., and Boike, J.: Pathways of ice-wedge degradation in polygonal tundra under different hydrological conditions, The Cryosphere, 13, 1089–1123, 2019.

Obu, J., Westermann, S., Bartsch, A., Berdnikov, N., Christiansen, H. H., Dashtseren, A., Delaloye, R., Elberling, B., Etzelmüller, B., Kholodov, A., et al.: Northern Hemisphere permafrost map based on TTOP modelling for 2000–2016 at 1 km2 scale, Earth-Science Reviews, 193, 299–316, 2019.

Orvin, A. K.: Litt om kilder på Svalbard, 57, 16–38, 1944.

Owen, A. and Zhou, Y.: Safe and Effective Importance Sampling, Journal of the American Statistical Association, 95, 135–143, https://doi.org/10.1080/01621459.2000.10473909, 2000.

Pandžic, M., Mihajlovic, D., Pandžic, J., and Pfeifer, N.: Assessment of the geometric quality of sentinel-2 data, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 41, 489–494, 2016.

Picard, G., Sandells, M., and Löwe, H.: SMRT: An active–passive microwave radiative transfer model for snow with multiple microstructure and scattering formulations (v1. 0), Geoscientific Model Development, 11, 2763–2788, 2018.

Pirk, N., Aalstad, K., Westermann, S., Vatne, A., van Hove, A., Tallaksen, L. M., Cassiani, M., and Katul, G.: Inferring surface energy fluxes using drone data assimilation in large eddy simulations, Atmospheric Measurement Techniques, 15, 7293–7314, https://doi.org/10.5194/amt-15-7293-2022, 2022.

Pirk, N., Aalstad, K., Yilmaz, Y. A., Vatne, A., Popp, A. L., Horvath, P., Bryn, A., Vollsnes, A. V., Westermann, S., Berntsen, T. K., et al.: Snow-vegetation-atmosphere interactions in alpine tundra, Biogeosciences Discussions, 2023, 1–24, 2023.

Reynolds, D., Quéno, L., Lehning, M., Jafari, M., Berg, J., Jonas, T., Haugeneder, M., and Mott, R.: Seasonal snow–atmosphere modeling: let's do it, The Cryosphere, 18, 4315–4333, https://doi.org/10.5194/tc-18-4315-2024, 2024.

Sanz-Alonso, D., Stuart, A., and Taeb, A.: Inverse Problems and Data Assimilation, Cambridge University Press, https://doi.org/10.1017/9781009414319, 2023.

Stuenzi, S. M., Boike, J., Gädeke, A., Herzschuh, U., Kruse, S., Pestryakova, L. A., Westermann, S., and Langer, M.: Sensitivity of ecosystem-protected permafrost under changing boreal forest structures, Environmental Research Letters, 16, 084 045, 2021.

Stumpf, A., Michéa, D., and Malet, J.-P.: Improved co-registration of Sentinel-2 and Landsat-8 imagery for Earth surface motion measurements, Remote Sensing, 10, 160, 2018.

Trofaier, A. M., Westermann, S., and Bartsch, A.: Progress in space-borne studies of permafrost for climate science: Towards a multi-ECV approach, Remote Sensing of Environment, 203, 55–70, https://doi.org/10.1016/j.rse.2017.05.021, earth Observation of Essential Climate Variables, 2017.

van Hove, A., Aalstad, K., Lind, V., Arndt, C., Odongo, V., Ceriani, R., Fava, F., Hulth, J., and Pirk, N.: Inferring methane emissions from African livestock by fusing drone, tower, and satellite data, EGUsphere, 2025, 1–36, https://doi.org/10.5194/egusphere-2024-3994, 2025.

van Leeuwen, P. J., Künsch, H. R., Nerger, L., Potthast, R., and Reich, S.: Particle filters for high-dimensional geoscience applications: A review, Quarterly Journal of the Royal Meteorological Society, 145, 2335–2365, https://doi.org/10.1002/qj.3551, 2019.

Vionnet, V., Brun, E., Morin, S., Boone, A., Faroux, S., Le Moigne, P., Martin, E., and Willemet, J.-M.: The detailed snowpack scheme Crocus and its implementation in SURFEX v7. 2, Geoscientific model development, 5, 773–791, 2012.

Wendt, L., Rouyet, L., Christiansen, H. H., Lauknes, T. R., and Westermann, S.: InSAR sensitivity to active layer ground ice content in Adventdalen, Svalbard, EGUsphere, 2024, 1–26, 2024.

Westermann, S., Lüers, J., Langer, M., Piel, K., and Boike, J.: The annual surface energy budget of a high-arctic permafrost site on Svalbard, Norway, The Cryosphere, 3, 245–263, 2009.

Westermann, S., Boike, J., Langer, M., Schuler, T., and Etzelmüller, B.: Modeling the impact of wintertime rain events on the thermal regime of permafrost, The Cryosphere, 5, 945–959, 2011.

Westermann, S., Ingeman-Nielsen, T., Scheer, J., Aalstad, K., Aga, J., Chaudhary, N., Etzelmüller, B., Filhol, S., Kääb, A., Renette, C., Schmidt, L. S., Schuler, T. V., Zweigel, R. B., Martin, L., Morard, S., Ben-Asher, M., Angelopoulos, M., Boike, J., Groenke, B., Miesner, F., Nitzbon, J., Overduin, P., Stuenzi, S. M., and Langer, M.: The CryoGrid community model (version 1.0) – a multi-physics toolbox for climate-driven simulations in the terrestrial cryosphere, Geoscientific Model Development, 16, 2607–2647, https://doi.org/10.5194/gmd-16-2607-2023, 2023.

Wilczak, J. M., Akish, E., Capotondi, A., and Compo, G. P.: Evaluation and bias correction of the ERA5 reanalysis over the United States for wind and solar energy applications, Energies, 17, 1667, 2024.

Zhang, T.: Influence of the seasonal snow cover on the ground thermal regime: An overview, Reviews of Geophysics, 43, 2005.

Zweigel, R. B., Westermann, S., Nitzbon, J., Langer, M., Boike, J., Etzelmüller, B., and Vikhamar Schuler, T.: Simulating snow redistribution and its effect on ground surface temperature at a High-Arctic site on Svalbard, Journal of Geophysical Research: Earth Surface, 126, e2020JF005 673, 2021.

Zweigel, R. B., Dashtseren, A., Temuujin, K., Aalstad, K., Webster, C., Stuenzi, S. M., Aas, K. S., Lee, H., and Westermann, S.: Simulating the thermal regime and surface energy balance of a permafrost-underlain forest in Mongolia, Journal of Geophysical Research: Earth Surface, 129, e2023JF007 609, 2024.

Zwieback, S., Westermann, S., Langer, M., Boike, J., Marsh, P., and Berg, A.: Improving Permafrost Modeling by Assimilating Remotely Sensed Soil Moisture, Water Resources Research, 55, 1814–1832, https://doi.org/10.1029/2018WR023247, 2019.