https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

Machine learning for snow depth estimation over the European Alps, using Sentinel-1 observations, meteorological forcing data and physically-based model simulations

Lucas Boeykens[1,2], Devon Dunmire[2], Jonas-Frederik Jans[1], Willem Waegeman[3], Gabriëlle De Lannoy[2], Ezra Beernaert[1], Niko E.C. Verhoest[1], and Hans Lievens[1]

[1]Hydro-Climate Extremes Lab, Ghent University, Coupure links 653, 9000 Gent, Belgium
[2]Department of Earth and Environmental sciences, KU Leuven, Celestijnenlaan 200E, 3001 Leuven, Belgium
[3]Department of Data Analysis and Mathematical Modelling, Ghent University, Coupure links 653, 9000 Gent, Belgium

Correspondence: Lucas Boeykens (lucas.boeykens@ugent.be, lucas.boeykens@kuleuven.be)

Abstract. Seasonal mountain snow is an indispensable resource, providing drinking water to more than a billion people worldwide, supporting agriculture, industry and hydropower generation, and sustaining river discharge, soil moisture and groundwater recharge. However, accurate estimates of this seasonal water storage remain limited, even in the European Alps, where there is a dense network of in situ monitoring stations. In this study, we address this issue by estimating Alpine snow depth at a 100 m spatial and sub-weekly temporal resolution with an extreme gradient boosting model (XGBoost) for the time period 2015-2024. We explore the potential for using Sentinel-1 C-band dual-polarized synthetic aperture radar polarimetry (PolSAR) observations to improve upon backscatter-based approaches, and include regionally downscaled meteorological forcing data and modeled snow depth inputs to further explain interannual and spatial variability. To account for the spatio-temporal dependencies present in the snow depth data, we conduct a threefold nested cross-validation, and incorporate spatial training data to better represent topographical patterns in snow depth variability. Finally, we utilize XGBoost's booster and Shapley additive explanation values to understand the relationship between the input features and predicted snow depths during both dry and wet snow conditions. Our results demonstrate that incorporating Sentinel-1 PolSAR observations leads to more accurate snow depth retrievals compared to using backscatter alone. In addition, our analyses indicate that including either meteorological forcing data or modeled snow depth estimates substantially improves the XGBoost snow depth estimates, both of which yield comparable accuracy. Finally, we demonstrate that the inclusion of spatial training data is essential for capturing the topographic influence on snow depth estimates, and to obtain good spatial prediction accuracy. Overall, this work contributes to an improved large-scale monitoring of water stored in seasonal mountain snow.

1 Introduction

Seasonal snow plays a critical role in nature, society and climate. First, snow serves as a natural water reservoir that contributes to river discharge, soil moisture and groundwater recharge (Gascoin et al., 2024; Dozier et al., 2016). In addition, snow provides drinking water to over 1.2 billion people worldwide, supports agricultural irrigation via snow melt runoff, and meets other

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

essential human needs (Bales et al., 2006; Barnett et al., 2005; Luce et al., 2014; Bormann et al., 2018; Qin et al., 2020). Snow melt runoff is also indispensable for hydropower generation, can contribute to traffic delays and accidents, and sustains a multi-billion dollar winter recreation industry that is the dominant revenue source for many Alpine communities (Burakowski and Magnusson, 2012; Seeherman and Liu, 2015; Sturm et al., 2017; Parthum and Christensen, 2022). The climatic significance of snow is arguably even greater, albeit harder to monetize (Sturm et al., 2017). The combination of the high albedo of snow, which reflects of a large portion of the incoming solar radiation, with the vast area of the world covered yearly by snow, enhances global cooling, and thus influences the Earth's surface energy balance (Warren, 1982; Groisman et al., 1994; Marks and Dozier, 1992).

Globally, the spatial and temporal dynamics of seasonal snow have been changing over the last decades. Notarnicola (2022) and Bormann et al. (2018) report a negative trend in global annual snow cover, while Matiu et al. (2021) observed decreasing trends in monthly mean snow depth (SD) for the months of November through May, at most of their surveyed sites across the European Alps. Given the importance of seasonal snow, key snowpack properties such as SD and snow water equivalent (SWE), the latter relating to SD through snow bulk density, have been designated as essential climate variables, and various scientific institutions and international organizations have prioritized their enhanced observation and monitoring (Gascoin et al., 2024; Dozier et al., 2016; World Meteorological Organization, 2023).

To date, SD and SWE have been observed using a wide range of measurement techniques, including manual and automated point measurements; airborne passive and active optical sensors; airborne and space-borne passive microwave radiometers; and airborne and space-borne active microwave sensors, such as synthetic aperture radar (SAR). Manual and automated point measurements offer frequent data at many locations globally, but fall short in capturing the snowpack's spatial variability due to their relatively sparse and uneven spatial distribution within alpine regions (Miller et al., 2022; López-Moreno et al., 2015). Moreover, this type of monitoring is often conducted at relatively flat and open sites, rarely covers the highest elevations, and can disturb the snowpack (Dozier et al., 2016; Deems et al., 2013; Miller et al., 2022). Snow measurements derived from airborne passive and active optical sensors, such as photogrammetry-based SD maps (Bührle et al., 2023; Marty et al., 2019) and lidar altimetry-derived SD and SWE products (Painter et al., 2016; Deems et al., 2013), provide high-resolution, spatially distributed measurements, yet their use is constrained by limited spatial coverage and sparse temporal sampling (Tsang et al., 2022; Dozier et al., 2016; Hill et al., 2019). In contrast, space-borne passive microwave radiometers provide broad spatial coverage, but their effectiveness is reduced by a coarse spatial resolution ($>$ 25 km) and signal saturation in deep ($\geq$ 0.8 m) snowpacks (Tedesco and Narvekar, 2010). Alternatively, space-borne SAR has demonstrated potential for monitoring SD or SWE at finer spatial scales, particularly with Ku- and X-band observations (Tsang et al., 2022). Unfortunately, X-band observations are not publicly available, and no Ku-band satellite mission currently exists (Tsang et al., 2022).

Recent work (Lievens et al., 2019; Brangers et al., 2024b) has demonstrated the sensitivity of active microwave observations to snow at C-band ($\sim$ 5.4 GHz). This work focuses on C-band backscatter in co- (VV) and cross- (VH) polarization, demonstrating that under dry snow conditions, an increasing snowpack leads to enhanced signal depolarization. This depolarization is thought to be related to snow volume scattering; (multiple) scattering on anisotropic snow crystals or clusters of crystals; and scattering at snow (or snow-ground) layer interfaces, etc. As a result, mainly cross-polarized backscatter, or the ratio of cross-

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

to co-polarized backscatter, is sensitive to an increasing SD (Lievens et al., 2019; Tsang et al., 2022). Jans et al. (2025) further explored the sensitivity of Sentinel-1 (S1) C-band backscatter satellite observations to seasonal patterns of SD and SWE. This work emphasizes the potential of S1 C-band dual-polarized SAR polarimetry (PolSAR) for SD retrieval, through the use of derived PolSAR variables that inform about the dominant snow scattering mechanism (i.e., a polarimetric scattering angle that increases with a growing snowpack), or about the intensity of the total received backscatter (i.e., the first Stokes parameter). Previous work has explored PolSAR for wet snow or snow cover area detection (Tsai et al., 2019; Varade et al., 2020). However, no operational retrieval algorithms currently exist that utilize dual-polarized PolSAR observations for estimating mountain SD or SWE.

In contrast, SD retrieval algorithms that incorporate S1 C-band backscatter observations have been developed previously. For instance, Lievens et al. (2022) use C-band co- and cross-polarized backscatter observations in a conceptual change detection algorithm to estimate SD across the European Alps at sub-kilometer spatial resolution. However, the algorithm's performance is limited during periods of wet and shallow snow, in forested areas, or after frequent freeze-thaw events (Lievens et al., 2022; Brangers et al., 2024b). To address some of these limitations, Dunmire et al. (2024) implemented machine learning (ML) to enhance S1-based SD estimates at a high spatial resolution (100 m). They utilized an extreme gradient boosting model (XGBoost; Chen and Guestrin (2016)), incorporating input features related to topography, land and snow cover, and S1 C-band backscatter. For the latter, they focused on the same satellite observations used by Lievens et al. (2022), namely co- (VV) polarized backscatter, and the cross-polarization ratio, defined as the ratio of cross (VH)- to co-polarized backscatter. Other studies have similarly explored the application of ML for estimating SD and SWE, both with and without integrating S1 C-band satellite observations (Xiong et al., 2022; Daudt et al., 2023; Broxton et al., 2024; Duan et al., 2024). However, no work currently exists that incorporates dual-polarized PolSAR variables for snow retrieval; not in a conceptual algorithm, nor within an ML model.

In addition to the previously mentioned conceptual, empirical, and ML approaches, physically-based models are also used to estimate snow properties, including SD and SWE (e.g., Lute et al. (2022); Bartelt and Lehning (2002); Niu et al. (2011)). Such models use meteorological forcings — often reanalysis data — to simulate the snowpack as one or multiple layers; solving mass and energy balance equations and mathematically representing the physical processes occurring within the snowpack. However, due to uncertainties in the forcing data and/or limitations in model representation, modeled SD may differ substantially compared to in situ measurements. In an attempt to further improve the estimates, multiple authors have explored the potential of ML in a hybrid physical/data-driven approach (e.g., Wang and Tian (2024); Pomarol Moya et al. (2025)). Thereby, additional information relevant to the snowpack can be incorporated, e.g., S1 C-band observations. A similar approach has been used by Brangers et al. (2024a) and De Lannoy et al. (2024) within a data-assimilation setup, who used S1 C-band snow retrievals to enhance modeled SD. Additionally, ML models can also be used as emulators, replacing physical models by directly estimating SD from meteorological forcings.

In this manuscript, we further investigate the potential of ML to improve SD estimates across the European Alps. To this end, we conduct various experiments comparing the performance of S1 C-band backscatter with PolSAR observations. Additionally, we use feature importance (FI) analysis to examine when and where satellite observations contribute to SD predictions, under

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

Table 1. Overview of the datasets and their associated targeted resolution marked by a cross(×). The 100 m and 500 m resolutions refer to the spatial resolution of the WGS84 grid.

Resolution WGS84	In situ SD photogrammetry maps	S1 C-band variables	Snow cover data	Wet snow mask	Meteorological forcing data	*Snowclim* SD	Auxiliary data
100 m	×			×			×
500 m		×	×		×	×	

both dry and wet snow conditions. We also evaluate the added value of incorporating meteorological forcing data or physically-based snow model SD estimates as features in our ML model, to improve the representation of interannual and site-specific variability. To validate our approach, we implement a threefold nested cross-validation (nested CV) framework, which masks subsets of the data during training and predicting (testing). This framework accounts for the spatial, temporal and spatio-temporal dependencies in the data, an essential consideration when validating ML models for spatio-temporal purposes (Meyer et al., 2018). Finally, to address the limitations of relying solely on point-based training data, we compare input configurations trained with and without airborne snow survey data, and validate the predictions against airborne photogrammetry snow surveys conducted across the Dischma Valley, Switzerland.

2 Data

Various datasets were collected and reprojected (and/or resampled) to two geographic grids used in this study: the ∼100 m (1/1008°) resolution World Geodetic System 1984 (100 m WGS84) grid, for fine-resolution datasets or to preserve more details in the target (SD) and auxiliary datasets, and the ∼500 m (5/1008°) WGS84 grid (500 m WGS84), used for the course-scale and satellite datasets. Table 1 provides an overview of the data products and the grid to which each was aligned to be used within this study.

2.1 In situ SD measurements

We compiled in situ SD data over the European Alps by combining stationary and point-based measurements and airborne photogrammetry snow surveys, sourced from multiple regional providers across the Alpine countries (Fig. 1). Point-based measurements were collected for the time period 2015-2024, and include data from Austria (provided by Geosphere Austria, formerly ZAMG, and TIWAG), France (from measurement stations operated by partners under agreement of MetéoFrance), Italy (collected in the autonomous regions of Valle d'Aosta, Trento and Alto-Adige, Piemonte, Lombardy and data provided by the International Center for Environmental Monitoring CIMA Research Foundation), and Switzerland (manual and automated point-based SD measurements collected and managed by the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL) - Institute for Snow and Avalanche Research (SLF)). In addition, we supplemented our SD dataset with

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

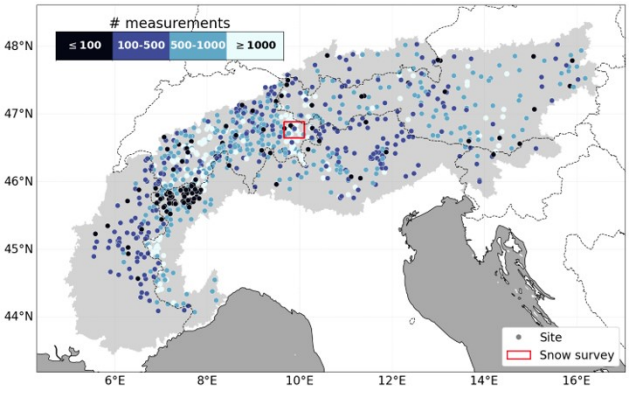

Figure 1. Spatial distribution of (dots) static point measurement sites colored by the number of available data in time, and (red box) airborne photogrammetry snow surveys conducted across the Dischma valley, used to train XGBoost and evaluate the SD predictions. The light gray shaded area delineates the Alpine region as defined by the Alpine Convention (Aureliano, 2020).

measurement sites originating from the Synoptic Data platform and the National Oceanic and Atmospheric Administration's Global Historical Climatology (GHC) network, extending our coverage to locations in Germany and Slovenia.

Next, we collected spatial SD measurements, derived from airborne photogrammetry (Marty et al., 2019; Bührle et al., 2023) over the Dischma valley, Switzerland (Fig. 1, red box). A total of nine photogrammetry SD surveys (snow surveys) were added to our dataset, with an original resolution of 0.5 m (6 maps, measured yearly near peak SD between 2017 and 2022) or 2 m (3 maps, collected on 3 distinct days in 2016). For both resolutions, we applied a linear averaging and reprojected all surveys to the target grid of our SD estimates, the 100 m WGS84 grid.

2.2 S1 C-band PolSAR and backscatter observations

Within our research, we collected data from the European Space Agency (ESA) and Copernicus S1A and -B satellites, which operate in the C-band (5.4 GHz) and share the same orbital plane with a 180° orbital phasing difference. Both have a 12 day repeat cycle. Combined, they offer a repeat cycle of 6 days over the study area. S1B, however, has been out of operation since 24 December 2021, and thus does not cover the complete study period (2015-2024).

S1 dual-polarized PolSAR observations were acquired from interferometric wide swath (IW) single look complex (SLC) data in dual-polarization (VV and VH) mode. The data were processed into two dual-polarized PolSAR variables: the polarimetric scattering angle (α) and the first Stokes parameter (S_0), using the Sentinel application platform (SNAP) toolbox, version 9, with processing steps described in Jans et al. (2025). These processing steps included applying precise orbit file corrections, radiometric calibration, debursting TOPSAR bursts, and merging sub-swaths into a single image. This was followed by the generation of the 2×2 covariance matrix (C2-matrix). The C2-matrix was then used as input for two distinct processing chains. The first one involved an eigenvector/value decomposition of the C2-matrix, followed by the H/α decomposition (Cloude and Pottier, 1997), from which α was derived. The second processing chain utilized the compact-polarimetry Stokes parameter

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

generation to derive S_0 from the C2-matrix. Next, both variables were multilooked to $\sim$ 25 m to reduce speckle and finally, a Range Doppler terrain correction was applied to both variables, reducing (geometric) distortions due to topographical variations in the scene. For the latter, the Copernicus 30 m global digital elevation model (GLO-30 DEM) was used.

Next, co- (γ^0_{VV}) and cross- (γ^0_{VH}) polarized backscatter variables were retrieved from ground range detected (GRD) amplitude data. Processing steps, described in Lievens et al. (2022) and Jans et al. (2025), were applied, including precise orbit file application, GRD border and thermal noise removal, radiometric calibration, terrain flattening to backscatter as γ^0 (Small, 2011) and finally a Range Doppler terrain correction. As for the PolSAR variables, the Copernicus GLO-30 DEM was used during the processing steps. Additionally, local incidence angle (LIA) information was preserved and the cross-polarization ratio (γ^0_{CR}; calculated as the ratio between γ^0_{VH} and γ^0_{VV} in linear scale) was calculated.

Finally, both the PolSAR and backscatter observations were rescaled to reduce inter-orbital and interannual start-of-season differences. To this end, we first reprojected the satellite observations onto the 100 m WGS84 grid using linear averaging, and resampled to the 500 m WGS84 grid by taking the mean. Then, we performed a first and second-order moment scaling (to correct for differences in the mean and variance) between orbits, followed by subtracting the mean summer value (July-September) from the satellite observations for each snow season (September-June). To indicate that these variables have been rescaled, we further append a superscript "s" ($^{\mathrm{s}}$) to the satellite variable notation.

2.3 Snow cover data and wet snow mask

Snow cover data were obtained from two sources: binary snow cover information from the interactive multisensor snow and ice mapping system (IMS; U.S. National Ice Center (2008)) at 1 km resolution, and fractional snow cover (FSC) information from the moderate resolution imaging spectroradiometer (MODIS; Hall and Riggs (2020a, b)) at 500 m resolution. For the latter, we collected cloud gap-filled FSC data separately from the Terra and Aqua satellites, which we both reprojected onto the 500 m WGS84 grid using linear averaging. Subsequently, we combined the individual products using a weighted average, which accounted for the quality flags of each product, the time lag between the last cloud-free FSC observation and the current date, and the time lag of the individual satellite observations relative to each other (Appendix A). Next, a cutoff was set on the combined product, only allowing for FSC information less than 7 days before the current date. Finally, remaining gaps were filled using IMS' binary snow cover information, reprojected to the same 500 m WGS84 grid using majority resampling. Hereby, a no-snow indication of IMS was set to an FSC value of 0%, while an IMS snow covered pixel received a value of 100%.

Additionally, a wet snow mask was acquired to differentiate between dry and wet snow periods. This dataset, based on S1 C-band satellite observations and obtained from the Copernicus Land Monitoring Service (CLMS) High Resolution Snow and Ice Monitoring (HRSIM) SAR Wet Snow in high mountains (SWS) product (Copernicus Land Monitoring Service, 2024), provides a binary classification between wet snow conditions, and dry, patchy or snow-free conditions. The data, available at 60 m resolution, were first reprojected onto the 100 m WGS84 grid using majority resampling, meanwhile excluding no-data values. Subsequently, all non-wet pixels were classified as dry, while quality flags were retained to identify locations and periods with no (reliable) classification.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

2.4 Meteorological forcing data

We downscaled three-hourly meteorological forcing data over the European Alps to a 500 m spatial resolution, to serve as additional input to the ML model. To this end, we first collected coarse three-hourly meteorological forcing data with a 0.1° spatial resolution. Precipitation data were sourced from the multi-source weighted-ensemble precipitation (MSWEP, Beck et al. (2019)) dataset, which combines various state-of-the-art reanalysis products with gauge and satellite observations (Beck et al., 2019). Other variables, including downward shortwave and longwave radiation, 2-m air temperature, wind speed, and 2-m relative humidity, were obtained from the multi-source weather (MSWX, Beck et al. (2022)) product. This dataset is derived from coarse-scale European centre for medium-range weather forecasts reanalysis v5 (ERA5) data refined through bias correction and downscaling using monthly 0.1° reference climatologies (Beck et al., 2022).

Subsequently, we applied bilinear interpolation in combination with other downscaling techniques to account for local terrain features. To this end, the Copernicus GLO-30 DEM was employed, linearly averaged to the 500 m WGS84 grid to match the target resolution for the meteorological forcings. For each location x within the 500 m grid and time step t, the coarse scale precipitation ($P_{\mathrm{coarse}_{x,t}}$; [mm $3\mathrm{h}^{-1}$]) data were downscaled using the following method:

$$P_{x,t} = (1-D)\cdot P_{\mathrm{coarse}_{x,t}} + D\cdot P_{\mathrm{coarse}_{x,t}}\cdot\left[0.75+0.5\frac{z_x - z_{\min}}{z_{\max}-z_{\min}}\right] \quad \text{with } D=\frac{z_{\max}-z_{\min}}{D_{\mathrm{dif}}} \quad 0\le D\le 1 \tag{1}$$

with z_x the elevation [m] of the 500 m Copernicus GLO-30 DEM, $z_{\min}$ and $z_{\max}$ the minimum and maximum elevation [m] within an interpolation window — centered on the location x and spanning an area roughly matching the original 0.1°grid size — and D_{dif} a user defined difference in elevation, set to 250 m. The user defined difference was introduced to focus the corrections on the study area, with minor adjustments for areas with small elevation differences.

The other forcings were adjusted as follows. The temperature was corrected using a seasonally varying lapse rate, following Liston and Elder (2006), to account for elevation-dependent temperature variations. For relative humidity, we first converted to dewpoint temperature, then applied a lapse rate correction — similar as for the temperature — and finally converted the corrected dewpoint temperature back to relative humidity (Liston and Elder, 2006). Downward shortwave radiation was adjusted following the method of Fiddes and Gruber (2014), which involved partitioning the radiation data into direct and diffuse components, accounting for the elevation effect on the direct component, and applying a topographic correction to both components. In contrast, downward longwave radiation and wind speed were not adjusted after bilinear interpolation.

2.5 Physically-based snow model *Snowclim*

Modeled SD estimates were obtained using *Snowclim* (Lute et al., 2022), a physically-based enhanced single layer snow model utilizing a fully distributed energy and mass balance. We ran the model over the entire European Alps with the downscaled meteorological forcing data at a 500 m spatial and 3-hourly temporal resolution. To obtain daily values, the 3-hourly SD estimates were averaged to a daily time step. A description of the model's parameter settings, how they have been calibrated, and the model's performance when validated against in situ measurements, is provided in Appendix B.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

2.6 Time-independent auxiliary data

In addition to time-dependent data, we employed time-independent auxiliary data, sourced from various providers. As with the other datasets (except the meteorological forcing and *Snowclim* data), these data were first projected onto the 100 m WGS84 grid, using linear averaging for continuous variables and majority resampling for categorical data. Forest cover fraction (FCF) data and land cover information, including a water mask, were acquired from the 2018 epoch of the 100 m Copernicus PROBA-V global land cover dataset (Buchhorn et al., 2020). Next, we used the Copernicus GLO-30 DEM to extract elevation and three topographic features: slope, aspect, and topographic position index (TPI), the latter quantifying the relative elevation of a grid cell concerning its surroundings within a predefined diameter. Slope and aspect were computed in MATLAB using the Geodetic (Craymer, 2022) and TopoToolbox (Schwanghart and Scherler, 2014) toolboxes, while TPI was calculated following the methodologies of Weiss (2001) and Lundblad et al. (2006). For this study, we computed the TPI using only the neighboring pixels of a specific grid cell, corresponding to a 3×3 grid window. Finally, we included a glacier mask, retrieved from the Randolph Glacier Inventory version 7.0 (RGI 7.0 Consortium, 2023).

3 Methods

3.1 Dataset preparation

We trained, validated, and tested XGBoost with various configurations on a dataset containing input features associated with each SD measurement (predictor variable) from the in situ measurement sites and snow surveys. To construct this dataset, we first excluded those stationary measurement sites located within glaciated areas, as well as those with less than 5% non-zero SD observations. Subsequently, we identified the unique locations of the stationary sites within the 100 m WGS84 grid, and averaged time series from sites within the same grid cell. Furthermore, we excluded SD measurements during the summer months July and August, and resampled the data with a 500 m spatial resolution to the 100 m WGS84 grid using value replication.

The S1 data spatially and temporally coinciding with available SD measurements on the 100 m WGS84 grid include both ascending and descending orbits. This allows a single location to have two S1 observation values for the same SD measurement on a given date. Importantly, we chose not to exclude satellite observations with high or low LIA, as we assumed that XGBoost could effectively learn the relationships between these satellite observations and the corresponding SD. For each location and date with an available satellite observation, we also selected the corresponding (nearest) time-dependent non-satellite data and time-independent auxiliary data.

A similar procedure was applied for the photogrammetry snow survey data: for each SD measurement, the nearest input feature values within the 100 m WGS84 grid were selected and compiled into a dataset, linking each SD observation with its corresponding input features. Since satellite acquisitions did not always coincide with the exact date of the snow surveys, we used the most recent S1 observation prior to the survey, with a maximum offset of two days. Lastly, to represent the temporal component of our datasets, we computed the day of the year (DOY) for each instance. To account for the cyclic nature of

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

the calendar year and prevent discontinuity between 31 December and 1 January, we transformed the DOY into two separate features using sine ($\mathrm{DOY}_{\sin}$) and cosine ($\mathrm{DOY}_{\cos}$) functions:

$$\mathrm{DOY}_{\sin} = \sin\left(\frac{2\pi \mathrm{DOY}}{n}\right) \tag{2}$$

$$\mathrm{DOY}_{\cos} = \cos\left(\frac{2\pi \mathrm{DOY}}{n}\right) \tag{3}$$

with n the number of days in the corresponding year.

3.2 Feature configurations and SD prediction

3.2.1 S1 PolSAR vs. backscatter observations

XGBoost was trained and validated with three different configurations to compare the performance of ML setups with PolSAR and backscatter observations (Table 2). For each configuration, XGBoost incorporated the same time-independent auxiliary features: elevation, slope, aspect, TPI, and FCF. Next, FSC data and LIA information were included, the latter to indirectly account for orbital differences. The first configuration, referred to as $\mathrm{PolSAR}_{\mathrm{ML}}$, incorporated α^{s} and S_0^{s} next to the shared input features. Conversely, the configuration focusing on backscatter observations ($\mathrm{Backscatter}_{\mathrm{ML}}$ configuration) included $\gamma^{0,\mathrm{s}}_{\mathrm{VV}}$ and $\gamma^{0,\mathrm{s}}_{\mathrm{CR}}$ as satellite input features. In addition, we combined the PolSAR and backscatter satellite input features ($\mathrm{Combination}_{\mathrm{ML}}$ configuration), to identify the satellite variables most effectively used within XGBoost.

3.2.2 S1 observations plus meteorological or *Snowclim* SD estimates

The satellite observations were further complemented with either meteorological forcing data ($\mathrm{Weather}_{\mathrm{ML}}$ configuration) or *Snowclim* SD estimates ($\mathit{Snowclim}_{\mathrm{ML}}$ configuration) to better resolve interannual and in-between-site variability (Table 2). The $\mathrm{Weather}_{\mathrm{ML}}$ configuration was an extension of the $\mathrm{Combination}_{\mathrm{ML}}$ setup, now including cumulative snowfall, cumulative wind speed, cumulative shortwave radiation, a cumulative number of melt days, and the mean daily temperature (T_{a}). Cumulative snowfall ($P_{\mathrm{s,c}}$) was estimated by first calculating the daily fraction of precipitation falling as snow, based on temperature, precipitation and relative humidity using the bivariate logistic regression model described by Jennings et al. (2018), and subsequently summing these fractions from the 1st of September of the corresponding snow year up to the prediction date. Similarly, the cumulative wind speed ($\mathrm{Ua_c}$) was derived as the sum of the daily wind speed from the 1st of September of the corresponding snow year up to the prediction date. Cumulative shortwave radiation (SWd_7), on the other hand, was summed over the seven days preceding the prediction date. Similarly, the cumulative number of melt days (MD_7) was defined as the number of days with a maximum temperature above 0°C within the 7 days preceding the prediction date. The $\mathit{Snowclim}_{\mathrm{ML}}$ configuration did not include meteorological forcing data as direct inputs, but instead makes use of modeled SD estimates from *Snowclim* ($\mathrm{SD}_{\mathrm{SC}}$), which are driven by meteorological conditions.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

Table 2. Features included in the different ML configurations tested within this manuscript. A cross (×) denotes the presence of a feature in a given configuration. Additionally, all configurations include the following baseline input features, which are not listed in the table: elevation, slope, aspect, TPI, FCF, FSC, $\mathrm{DOY}_{\sin}$, $\mathrm{DOY}_{\cos}$ and LIA.

ML configuration	α^{s}	S_0^{s}	$\gamma_{\mathrm{CR}}^{0,\mathrm{s}}$	$\gamma_{\mathrm{VV}}^{0,\mathrm{s}}$	$P_{\mathrm{s,c}}$	SWd_7	Ua_{c}	T_{a}	MD_7	$\mathrm{SD}_{\mathrm{SC}}$
$\mathrm{PolSAR}_{\mathrm{ML}}$	×	×								
$\mathrm{Backscatter}_{\mathrm{ML}}$			×	×						
$\mathrm{Combination}_{\mathrm{ML}}$	×	×	×	×						
$\mathrm{Weather}_{\mathrm{ML}}$	×	×	×	×	×	×	×	×	×	
Snowclim$_{\mathrm{ML}}$	×	×	×	×						×

3.2.3 Snow depth prediction

Prior to being used in XGBoost for SD prediction, each feature was standardized individually, using the mean and standard deviation calculated from the training set values. This standardization was consistently applied to the validation and predicting (test) datasets as well, ensuring compatibility across all data splits. Input features from ascending and descending orbits were fed to XGBoost separately. In case a location had both an ascending and descending satellite observation on a given date, the mean of both XGBoost SD outputs was taken.

3.3 Spatio-temporal nested cross-validation and hyperparameter tuning

The concept of nested cross-validation (nested CV) has been extensively utilized in other studies (Dora et al., 2018; Abdulaal et al., 2018; Parvandeh et al., 2020), since nested CV allows to characterize the generalization error of an ML model correctly (Blancas, 2021). Figure 2 illustrates that nested CV involves both inner and outer resampling to, respectively, train and tune the ML model (and its hyperparameters), and evaluate predictions on unseen and independent test data. Considering the susceptibility of SD measurements to spatial and temporal autocorrelation, we implemented the cross-validation strategy described in Landschoot et al. (2012) within a threefold nested CV framework for the site data (excluding the snow surveys), in which subsets of the data are masked during training, validation and predicting (testing). To this end, the data of the site measurements were partitioned into spatial, temporal and spatio-temporal folds.

The spatial folds were constructed using a two-step approach: first, the sites located within five km of one another were grouped into clusters, which were then randomly assigned to five unique folds, ensuring that each fold contained a comparable amount of data and maintained a similar SD distribution. For the temporal folds, measurement sites were kept separate, but data from all sites within a single snow season (September–June) were grouped and then divided into five folds, again ensuring comparable fold sizes and SD distribution. Finally, by combining spatial and temporal folds, we constructed 25 unique spatio-temporal folds.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

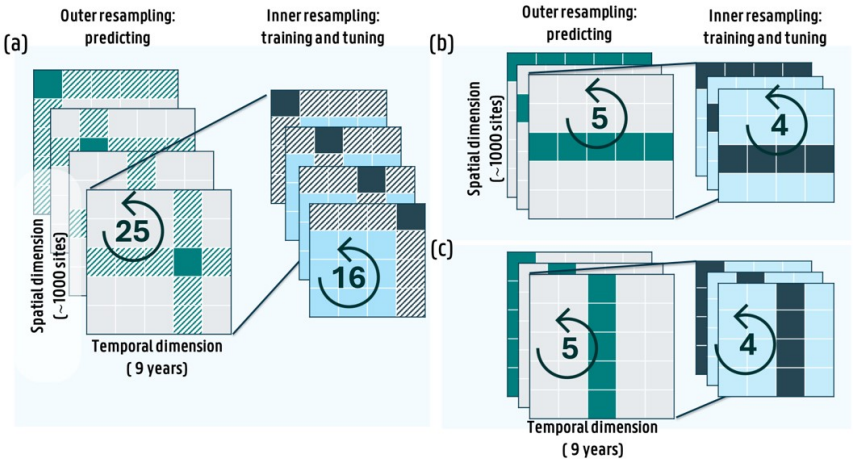

Figure 2. Principle of nested CV for (a) the spatio-temporal setup, (b) the spatial setup and (c) the temporal setup. The predicting (test) and training folds during outer resampling are colored in turquoise and light gray respectively. The validation and training folds during inner resampling are colored in dark and light blue. The striped boxes in (a) represent the excluded folds during training and testing within the spatio-temporal setup. During outer resampling, predictions are made with the trained ML model, with hyperparameters that are tuned during inner resampling. The numbers indicated in the arrows represent the amount of times outer and inner resampling are performed, respectively. For (a), four iterations are shown, while for (b) and (c), three iterations are displayed.

For each unique fold (five for spatial and temporal nested CV; 25 for spatio-temporal nested CV), we iteratively designated one fold as the test set during outer resampling, using the remaining folds for training and validation. To ensure independence of the test data, all data from the sites and/or years present in the test fold were excluded from the training and validation sets (Fig. 2).

The same approach was applied during inner resampling, to split the training and validation data used for tuning XGBoost's hyperparameters. To tune the hyperparameters, we utilized Scikit-learn's (Pedregosa et al., 2011) *RandomizedSearchCV* algorithm, employing the mean squared error (MSE) as loss function. A different random seed was set for each outer resampling loop, introducing variability in the selection of optimal hyperparameter combinations. During each loop, 150 hyperparameter combinations were chosen from a predefined tuning grid, based on the one provided in PyCaret (PyCaret, 2020). The hyperparameters yielding the best mean MSE score across the validation folds were selected. This approach produced independent predictions for every instance in the dataset, enabling the calculation of performance metrics and the construction of FI scores.

3.4 Feature importance scores

We used both XGBoost's booster and Shapley additive explanations (SHAP; Lundberg and Lee (2017)) values to assess the feature's impact on the SD estimates. For the booster, we selected the *gain* FI, which informs about the contribution of each

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

input feature in minimizing the loss function during model training. As such, for every outer resampling loop within the nested CV framework, we computed the gain FI of the individual features during model training.

SHAP values, on the other hand, quantify the contribution of each input feature to an individual (SD) prediction, measured relative to an expected prediction (SD) value. Within the nested CV framework, SHAP values thus indicate how each feature influences the SD predictions made for the outer resampling predicting sets, relative to the average (SD) prediction computed from the training set. Consequently, SHAP values can be both positive and negative. To assess the overall global SHAP value FI during SD prediction, mean absolute SHAP values were computed for each input feature during every outer resampling loop. The latter were converted into relative contributions by normalizing them with the sum across all features.

We further used SHAP values to assess the contribution of the PolSAR and backscatter satellite observations to the SD predictions throughout the seasonal evolution of the snowpack. To this end, for each site and snow season in the test sets, we first identified the periods during which a snowpack was present, based on the measured SD. The season start date was identified as the first date with SD $\geq$ 10 cm, that was followed by nine consecutive days with SD $\geq$ 1 cm. For low elevation sites with (very) shallow snowpacks, where this condition was not met for an entire snow season, the start of the snow season was marked by the first date with SD $\geq$ 10 cm. Conversely, we defined the end of a site's snow season as the first day of a 10-day period with 0 cm SD, near the end of snowpack presence. In addition, we used the CLMS SWS product to distinguish, where possible, between dry and wet snow conditions.

3.5 Performance metrics

In addition to the FI and SHAP value analysis, several performance metrics were used to validate XGBoost SD predictions with measured SD. For each configuration and across all outer resampling predicting sets (thus the total dataset), we computed the spatiotemporal Pearson correlation coefficient (R), mean absolute error (MAE), root mean squared error (RMSE) and bias between predicted and measured SD (all averaged over time and space). Since cross-validation is conducted in both space and time, the performance metrics reflect the ML model's ability to generalize across purely spatial, purely temporal, and combined spatio-temporal domains. Additionally, certain metrics were calculated on a per-site basis, such as the MAE.

Next, we evaluated the XGBoost's capability to generate spatial SD predictions by comparing them with the photogrammetry snow surveys conducted in the Dischma Valley, Switzerland. For this case, we trained XGBoost exclusively on measurement site data, excluding spatial data from the snow surveys. Model training and hyperparameter tuning were performed using fivefold spatial cross-validation, as this approach best reflects the intended application of the ML model: predicting SD on unseen locations across the European Alps within the time period of our S1 data collection. After training, XGBoost was applied to predict SD for each of the nine snow surveys, and performance metrics of predicted vs. observed SD were computed.

We then repeated this process, but now incorporating all snow survey data except those of the targeted date into the training dataset. This resulted in nine independently trained and tuned XGBoost models, each applied to predict SD for the corresponding unseen snow survey. In this approach, we also employed spatial cross-validation for training and tuning, where every training snow survey was randomly assigned to one of the five folds. This ensured that the distribution of training and validation data remained consistent across folds.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

4 Results and discussion

4.1 Performance of ML configurations with S1 PolSAR and backscatter variables

Across the three nested CV frameworks and configurations (i.e., $\text{PolSAR}_{\text{ML}}$, $\text{Backscatter}_{\text{ML}}$ and $\text{Combination}_{\text{ML}}$), performance is highest for the temporal framework (Table C1), and progressively deteriorates in the spatial and spatio-temporal frameworks. The differences between the configurations are moreover most pronounced for the temporal nested CV framework, with significant improvements in site MAE when comparing the $\text{PolSAR}_{\text{ML}}$ or $\text{Combination}_{\text{ML}}$ configurations with the $\text{Backscatter}_{\text{ML}}$ configuration ($p \ll 0.05$, 95% confidence level). While this indicates that at the training sites, PolSAR variables, or a combination of PolSAR and backscatter variables, seem to better capture the seasonal evolution of SD, the temporal framework overestimates model performance relative to the overall objective of this study, namely predicting SD at unseen locations across the European Alps.

The spatial and spatio-temporal frameworks provide a more realistic evaluation of model performance for this study. For the spatial framework, the $\text{PolSAR}_{\text{ML}}$ and $\text{Backscatter}_{\text{ML}}$ configurations display marginal differences in performance metrics (Table 3). This is likely the result of the spatial noise present in both types of S1 variables, and the similar relationship of the variables with SD. However, the $\text{Combination}_{\text{ML}}$ configuration exhibits significant ($p \ll 0.05$, 95% confidence level) improvements in site MAE, resulting in an overall MAE of 35 cm, a mean site MAE of 29 cm, and a mean site bias of 11 mm (excluding zero-measured SD). These values are slightly higher than those reported by Dunmire et al. (2024), which may stem from differences in satellite observation scaling, input feature configurations (e.g., a cumulative FSC), and the larger number of sites and years included in the present evaluation.

Additionally, all three configurations exhibit difficulties in predicting high SD values (Fig. C1a), which we attribute to the relatively small number of observation sites with a large record of observed SD > 2.5 m. Specifically, at those sites with a site bias ≤ -0.5 m (mostly located in Switzerland (Fig. C1b)) 14.2% of the observed SD measurements exceed 3 m, compared to only 1% across the entire training dataset. The limited ability of XGBoost, trained with S1 variables to explain interannual and site-specific variability, to predict high SD values must be considered when applying the model across the European Alps, especially when one wants to apply the model to predict outside the training dataset time period. For the latter, Table 3 indicates a further deterioration in model performance for the spatio-temporal nested CV framework, which must be seen as an accurate representation of the model performance when applying XGBoost to predict across the Alps outside the training period. Nonetheless, the PolSAR variables also improved model outcomes for this scenario, with significant ($p \ll 0.05$, 95% confidence interval) improvements in site MAE for both the $\text{PolSAR}_{\text{ML}}$ and $\text{Combination}_{\text{ML}}$ configurations.

4.2 Feature importance for ML configurations with S1 PolSAR and backscatter variables

Figure 3 displays the importance of the various input features across the $\text{PolSAR}_{\text{ML}}$, $\text{Backscatter}_{\text{ML}}$ and $\text{Combination}_{\text{ML}}$ configurations, as determined by (a) XGBoost's internal gain metric and (b) SHAP values computed from the outer resampling test sets. Across the configurations and nested CV frameworks, FSC and elevation consistently emerge as the most important features, similar to what has been reported by Dunmire et al. (2024). In addition, the *gain*-based FI indicates an increased

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

Table 3. Overall performance metrics for the $\text{PolSAR}_{\text{ML}}$, $\text{Backscatter}_{\text{ML}}$ and $\text{Combination}_{\text{ML}}$ configurations for the spatial and spatio-temporal nested CV frameworks. The values inside the brackets denote the metric values when zero-measured SD are included in the evaluation.

	Spatial			Spatio-temporal		
Metric	**$\text{PolSAR}_{\text{ML}}$**	**$\text{Backscatter}_{\text{ML}}$**	**$\text{Combination}_{\text{ML}}$**	**$\text{PolSAR}_{\text{ML}}$**	**$\text{Backscatter}_{\text{ML}}$**	**$\text{Combination}_{\text{ML}}$**
R [-]	0.73 (0.81)	0.73 (0.81)	0.74 (0.82)	0.70 (0.79)	0.69 (0.79)	0.70 (0.79)
MAE [m]	0.35 (0.23)	0.35 (0.23)	0.35 (0.22)	0.37 (0.24)	0.37 (0.24)	0.37 (0.24)
RMSE [m]	0.53 (0.41)	0.53 (0.42)	0.52 (0.41)	0.55 (0.43)	0.56 (0.44)	0.55 (0.43)

importance of FSC when transitioning from temporal to spatio-temporal nested CV, whereas elevation follows an opposite trend. In temporal nested CV, the ML model can leverage recurring SD patterns at sites within the training data, where — mainly at high-elevation sites — interannual variability is relatively limited and SD tends to follow similar trajectories across multiple years. As a result, the ML model may place greater emphasis on elevation to adjust the predictions. Conversely, in spatio-temporal nested CV, XGBoost may rely less on recurring SD patterns and therefore, seems to place greater emphasis on FSC and less on elevation to distinguish snow-covered from snow-free periods. Figure 3b, which displays the FI based on the training data, on the other hand, indicates a stronger importance for FSC, but shows a similar increase in fSC FI for spatio-temporal nested CV. Differently, Fig 3b displays a less clear pattern for the elevation feature. Nevertheless, non-relative mean absolute SHAP values display similar patterns as the *gain* FI.

Figure 3a also reveals that within $\text{PolSAR}_{\text{ML}}$ and $\text{Backscatter}_{\text{ML}}$ configurations, the S1 features are similarly used during model training. Specifically, S_0^{s} and $\gamma^{0,\text{s}}_{\text{VV}}$ appear to be less important than α^{s} and $\gamma^{0,\text{s}}_{\text{CR}}$ in the $\text{PolSAR}_{\text{ML}}$ and $\text{Backscatter}_{\text{ML}}$ configurations, respectively. In addition, Fig. 3b shows that α^{s} is more important in the $\text{PolSAR}_{\text{ML}}$ configuration than $\gamma^{0,\text{s}}_{\text{CR}}$ is in the $\text{Backscatter}_{\text{ML}}$ one, which indicates that XGBoost seems to effectively extract more information from α^{s} during SD prediction. Moreover, both the *gain*-based and SHAP value FI analyses reveal a less distinct separation between the S1 input features in the $\text{Backscatter}_{\text{ML}}$ configuration, with the SHAP values even suggesting that $\gamma^{0,\text{s}}_{\text{VV}}$ contributes as much as, or slightly more than, $\gamma^{0,\text{s}}_{\text{CR}}$. This implies that despite the lower importance during model training (i.e., the *gain*-based FI), XGBoost seems to rely more on $\gamma^{0,\text{s}}_{\text{VV}}$ when predicting at unseen locations and/or time periods.

Finally, both Fig. 3a and b showcase a consistent preference of the $\text{Combination}_{\text{ML}}$ configuration for α^{s} over $\gamma^{0,\text{s}}_{\text{CR}}$. Although Fig. 3 provides an aggregated view across all sites and the full snow season, including dry and wet snow conditions as well as zero-measured SD, the result aligns with findings from Jans et al. (2025), who reported stronger correlations between α and modeled SD during the accumulation (dry snow) period at most locations in their study area, compared to γ^{0}_{CR}. In contrast, our results indicate that S_0 contributes less than γ^{0}_{VV} during both training and prediction, despite its stronger correlation with dry snow reported by Jans et al. (2025). This difference may stem from several factors, including a different satellite resolution (1 km in the work of Jans et al. (2025)), the input scaling applied before model training, variations in land cover across sites, and especially the inclusion of both dry and wet snow periods.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

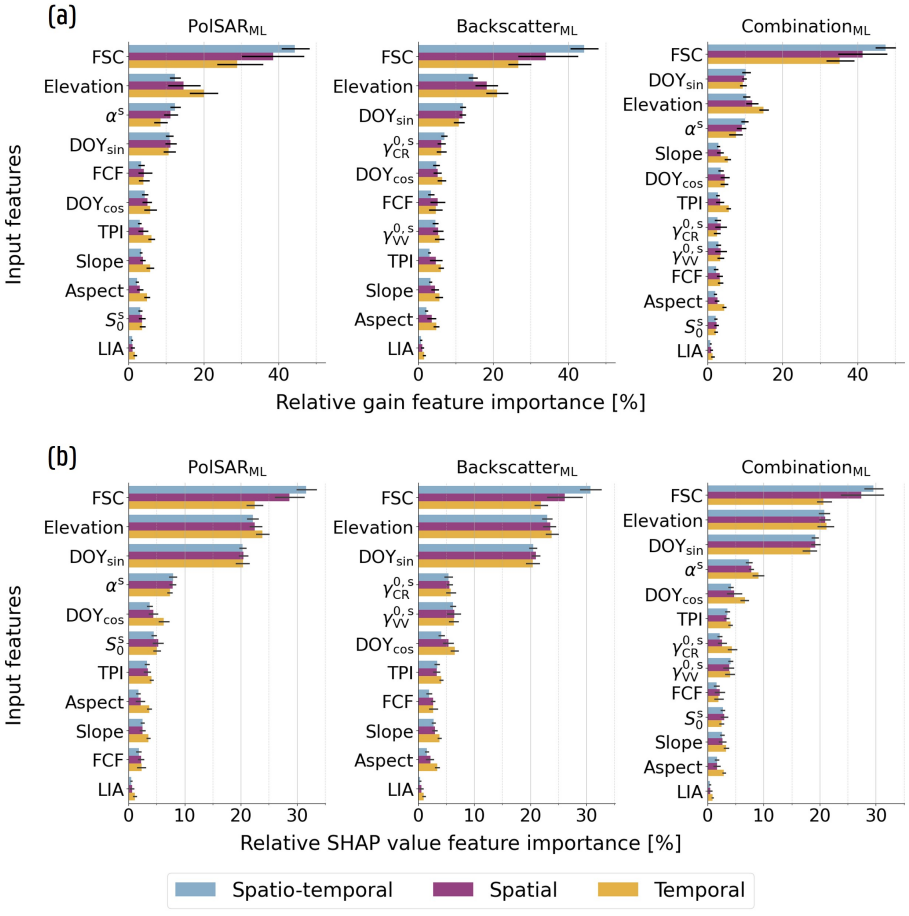

Figure 3. Upper row (a): Relative gain-based FI for the PolSAR$_{ML}$, Backscatter$_{ML}$, and Combination$_{ML}$ configurations. Bars represent the mean FI across the XGBoost models trained during inner resampling, with colors indicating the type of nested CV framework. Lower row (b): Similar to (a), but here FI is quantified as the mean absolute SHAP value per feature (relative values), computed separately for each test fold during outer resampling. The black lines in the bars denote the 95% confidence interval.

4.2.1 SHAP value analysis during snowpack presence

We further assessed the impact of the S1 variables during snow-covered periods, during both dry and wet snow conditions. To this end, we focused on the spatial nested CV framework, as it best aligns with the objective of this research; i.e., predicting SD at unseen locations within the S1 archive time period. Since the wet snow mask is only available from September 2016 onward, the 2015–2016 snow year was excluded from this analysis.

Fig. 4a showcases that in the PolSAR$_{ML}$ configuration, the contribution of α^s is similar under both dry and wet snow conditions (∼9%), with a slightly lower contribution observed during dry conditions. However, this result is based on a binary classification between the snow conditions that does not account for masked areas, and therefore still includes lower-elevation

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

locations and/or densely forested locations. Without the masked data, the dry-condition relative importance of α^{s} increases to approximately 13% (Fig. 4a), while the relative importance of the elevation feature decreases (not shown). Since the masked data points predominantly correspond to low SD values ($\leq$ 1 m), the SHAP value FI suggests that at bare, higher-elevation locations with deeper snowpacks, XGBoost places relatively more importance on α^{s}, a pattern also observed for $\gamma^{0,s}_{CR}$ in the Backscatter$_{ML}$ configuration.

Figure 4a also explains the more pronounced difference in SHAP value FI between the PolSAR variables, as XGBoost relies more heavily on α^{s} than on S^{s}_{0} across all snow conditions. In contrast, Fig. 4a displays that for the Backscatter$_{ML}$ configuration, XGBoost places greater emphasis on $\gamma^{0,s}_{VV}$ than on $\gamma^{0,s}_{CR}$ during wet snow conditions, which accounts for their nearly equivalent overall SHAP value FI (Fig. 3). In addition, the FI patterns observed in the Combination$_{ML}$ configuration closely resemble those of the PolSAR$_{ML}$ and Backscatter$_{ML}$ configurations, with the exception of $\gamma^{0,s}_{CR}$. We attribute this to overlapping information between $\gamma^{0,s}_{CR}$ and α^{s}, with XGBoost now relying more heavily on α^{s} at the expense of $\gamma^{0,s}_{CR}$. Additionally, internal compensations arising from the joint use of both PolSAR and backscatter inputs may further reduce the relative importance of $\gamma^{0,s}_{CR}$.

Figure 4b further illustrates the SHAP value contribution of α^{s} within the PolSAR$_{ML}$ configuration. During dry conditions, XGBoost has learned a positive relationship between α^{s} and its contribution to the SD estimates. This can be related to the findings of Jans et al. (2025), who found that at sites with medium to high LIA, α progressively augments with an increasing snowpack during the accumulation period. However, at certain medium-high elevation sites (1500 - 2500 m; not shown), mean α^{s} and corresponding SHAP value contributions remain low, even though predicted (and observed) SD are still relatively high. At these locations, there is often no clear increase in α^{s} with increasing SD; instead, a rise in $\gamma^{0,s}_{VV}$ or S^{s}_{0} may be observed. We hypothesize that this pattern is more common at sites with low LIA, where Jans et al. (2025) similarly observed decreasing values of α. Furthermore, Fig. 4b and d illustrate that, while for α^{s} values the associated SHAP values and SHAP value FI remain elevated under wet snow conditions, this is less the case for $\gamma^{0,s}_{CR}$ within the Backscatter$_{ML}$ configuration, suggesting that SD estimation based on α^{s} might be less sensitive to the presence of liquid water than when based on $\gamma^{0,s}_{CR}$.

Conversely, Fig. 4c displays that XGBoost has learned a dual relationship between the observed values for $\gamma^{0,s}_{VV}$ and the SHAP value contribution. The lower importance of $\gamma^{0,s}_{VV}$ during dry snow conditions can be explained by the findings of Lievens et al. (2019), who noted that volume scattering in VV-polarization within a dry snowpack at C-band is generally too weak to produce a significant impact. However, at certain locations – mostly with low-LIA satellite overpasses – increasing $\gamma^{0,s}_{VV}$ and associated positive SHAP values may be observed. We hypothesize that in certain similar locations, the S1 observations are more sensitive to surface scattering from snow layer interfaces, resulting in rising values for $\gamma^{0,s}_{VV}$, while both α^{s} and associated SHAP values tend to decrease. During wet snow conditions, however, the $\gamma^{0,s}_{VV}$ - SHAP value relationship is the opposite. During these conditions, XGBoost seems to leverage the information contained within the low observed $\gamma^{0,s}_{VV}$ values due to liquid water presence, to positively contribute to the predicted SD.

One of the main limitations of our current approach is that only valid wet snow classifications are available for higher-elevation and sparsely forested sites (Fig 4d), while some pixels labeled as dry may in fact correspond to wet snow conditions, or to patchy snow conditions. Although a more sophisticated classification method that incorporates meteorological data could

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

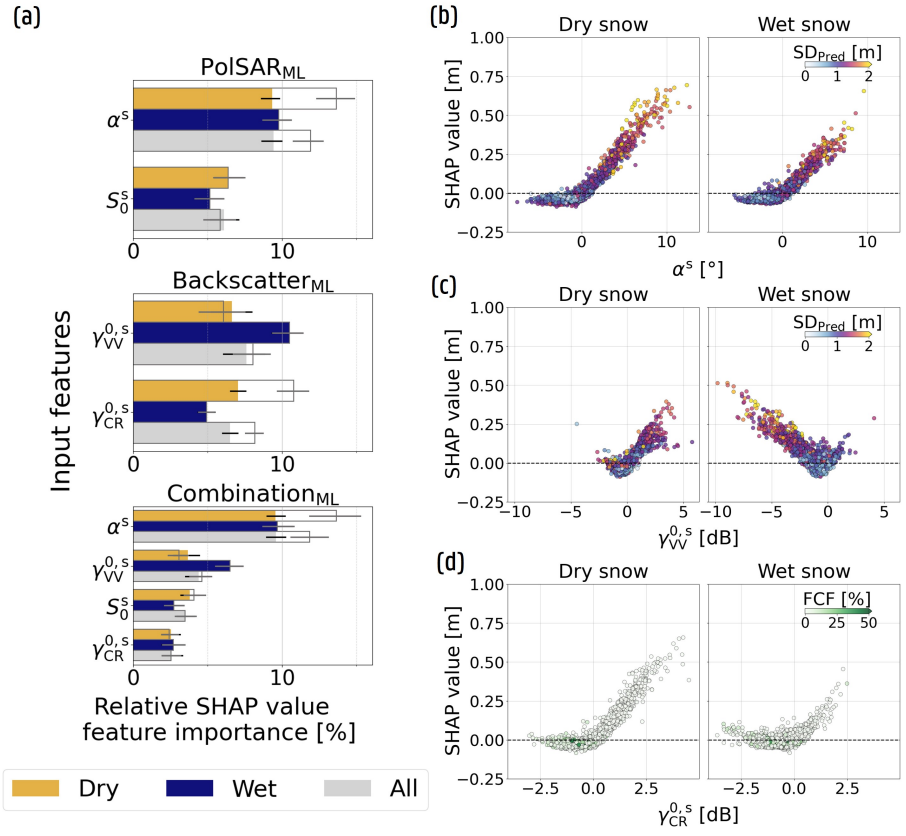

Figure 4. SHAP value analysis for the spatial nested CV framework during snow presence. (a) Relative SHAP value FI of the satellite input features for the PolSAR$_{ML}$, Backscatter$_{ML}$ and Combination$_{ML}$ configurations. The dark gray lines represent the FI after masking out non-valid dry snow classifications. (b) Scatter plot of mean α^s values and corresponding SHAP values for the PolSAR$_{ML}$ configuration, taken as the mean per site and snow season during dry and wet snow conditions (only valid classifications). The dots are colored according to the mean predicted SD (SD$_{pred}$). (c) Same as (b), but for $\gamma^{0,s}_{VV}$ within the Backscatter$_{ML}$ configuration. (d) Same as (c) for $\gamma^{0,s}_{CR}$, but colorized to the site FCF, indicating that the classification primarily targets higher-elevation non-forested areas.

provide a more accurate separation of dry and wet snow periods, we argue that the current approach remains suitable for identifying the contribution of the satellite variables within XGBoost under varying snow conditions. The wet snow mask provides a reliable classification primarily at higher-elevation sites — locations that typically exhibit deeper snowpacks — where satellite variables contribute more substantially to model performance. At masked locations, by contrast, the contribution of satellite variables is generally limited. This is consistent with findings by Dunmire et al. (2024), who showed that S1 C-band backscatter is most informative in non-forested terrain with SD $>$ 1 m, and has limited influence in areas where the wet snow mask does not provide a valid classification.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

4.2.2 Time series of SD and SHAP values

Finally, we assessed the SHAP value contribution of the S1 variables within the $\text{Combination}_{\text{ML}}$ configuration during the temporal evolution of the snowpack. To this end, Fig. 5 displays the results for two snow seasons at three distinct sites in Switzerland, all operated by SLF, whereas Fig. C2 represents an additional site near Grindelwald, Switzerland. To provide a realistic overview of the model performance within this study, the time series showcase the results of the spatial nested CV framework.

Figure 5a shows the time series for a measurement site (46.41°N, 8.92°E) near Arolla, Valais, located in a bare rock area with a 0% forest cover fraction and surrounded by mountains to the east, north, and west. At this high-elevation and bare site, the added value of α^{s} under dry snow conditions and $\gamma^{0,\text{s}}_{\text{VV}}$ under wet snow conditions becomes clear. During the 2017–2018 snow season, namely, a clear increase in α^{s} is observed during the accumulation phase. After reaching peak values around late January, associated with a large snowfall event, α^{s} and its associated SHAP values remain elevated until the snowpack starts to wet near the end of March. However, the peak occurs before the maximum measured SD in early April, a phenomenon not unique to this site and also reported by Jans et al. (2025), who found an average lag of around 60 days between peak α and peak SD in their study area. In contrast, the observed $\gamma^{0,\text{s}}_{\text{VV}}$ values do not increase during the accumulation phase, consistent with the findings of Lievens et al. (2019), and contribute minimally to the SD predictions. At certain sites — usually with low-LIA satellite overpasses — however, $\gamma^{0,\text{s}}_{\text{VV}}$ and it's SHAP values do increase during the accumulation period (e.g., Fig. C2). As such, depending on the LIA, α^{s} and $\gamma^{0,\text{s}}_{\text{VV}}$ might contribute more or less to the predictions.

During the ablation phase, marked by a decline in $\gamma^{0,\text{s}}_{\text{VV}}$ in early April, the contribution of α^{s} drops sharply, in line with decreasing observed values. In contrast, $\gamma^{0,\text{s}}_{\text{VV}}$ emerges as the most informative S1 feature under these wet snow conditions, with positive SHAP values peaking near its lowest observed backscatter values. Beyond this point, as $\gamma^{0,\text{s}}_{\text{VV}}$ increases due to the growing influence of superficial scattering processes (Marin et al., 2020), the contribution gradually declines. Similar patterns have been observed at other sites (e.g., 46.42°N, 8.23°E and 46.39°N, 7.97°E) and is also evident from Fig. 4c. In contrast, the behavior of α^{s} and its contribution to the SD estimates during the ablation period is less consistent. At some sites (e.g., 46.08°N, 7.92°E), the drop in $\gamma^{0,\text{s}}_{\text{VV}}$ coincides with a sudden increase in α^{s}, whereas at other sites, no such increase is observed or an increase is linked to late-season snowfall events. Further research is required to better understand the interaction between wet snow and α, though this is beyond the scope of this study.

The Arolla site (Fig. 5a) shows a similar seasonal evolution of $\gamma^{0,\text{s}}_{\text{VV}}$ in 2018–2019 as in 2017–2018, further supporting the findings of Fig. 4 on the role of this S1 feature under different snow conditions. In contrast, although α^{s} increases slightly in mid-March, it begins to decline as early as late February, whereas no corresponding sharp decrease is observed in $\gamma^{0,\text{s}}_{\text{VV}}$ (not shown). While difficult to confirm, this may be related to melt-freeze events and/or metamorphism processes happening in the snowpack starting in early March. Nevertheless, α is also here strongly contribution during snow accumulation, when it is most informative.

The potentially added value of α^{s} to the SD predictions at higher-elevated, bare locations is further supported by Fig. 5b, that displays the time series at a measurement site located near Prato, Ticino (46.47°N, 8.72°E), with a similar 0% forest

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

cover fraction but lower elevation (2222 m above sea level). Unlike the Arolla site, however, α^{s} is not only mainly informative during snow accumulation, but remains relatively elevated throughout February and March for both snow seasons, when the snowpack exceeds 2 m of snow. In addition, following the drop in $\gamma^{0,s}_{VV}$ near the beginning of the accumulation period, a (very) slight increase of $\gamma^{0,s}_{VV}$ is observed toward the ablation phase at this site (not shown), a pattern also observed by Jans et al. (2025) at some locations with very deep (and dry) snowpacks. In addition, Fig. 5b helps to explain the relative SHAP value FI of $\gamma^{0,s}_{CR}$ in Fig. 4a. Although XGBoost captures the increase in $\gamma^{0,s}_{CR}$ during the accumulation period to contribute positively to the predicted SD, the SHAP value contribution is way smaller compared to α^{s}, indicating a preference for α^{s} over $\gamma^{0,s}_{CR}$ by XGBoost.

Finally, the bottom time series (Fig. 5c) corresponds to a forested (63% FCF, 1868 m above sea level) site in Sobrio, Ticino (46.41°N, 8.92°E). Despite being consistently classified as dry snow in the time series, this area is actually masked as forest in the CLMS SWS wet snow product. This site was also analyzed by Dunmire et al. (2024), who concluded that, given the low FI of $\gamma^{0,s}_{VV}$ and $\gamma^{0,s}_{CR}$ in their study, S1 satellite observations should not be used for this location. Our findings support the limited contribution of the S1 features at similar lower and/or forested areas, as we observe minimal added value from both α^{s} and $\gamma^{0,s}_{VV}$ at this site, and the predictions from XGBoost trained without satellite observations are similar to, or even outperform, those based on satellite inputs for the time series at this site (not shown).

4.3 Added value of meteorological forcings and *Snowclim* SD estimates

Figure 6 illustrates the improvements when incorporating either regionally downscaled meteorological forcing data (Fig. 6b) or *Snowclim* SD estimates (Fig. 6c), with the most important improvements for high observed SD values ($\geq$ 2.5 m) when predicting at unseen locations (and time periods). As an example, the bias for SD $\geq$ 2.5 m improves by 27 cm when comparing the Weather_{ML} to the Combination_{ML} configuration within the spatial nested CV framework, decreasing from -1.47 m to -1.20 m. Incorporating *Snowclim* SD estimates further reduces this bias to -1.12 m, likely reflecting a stronger correspondence between the modeled and observed high SD values than what is captured through derived meteorological variables. Despite these gains, the number of sites with a mean bias ≤ -0.5 m remains similar in both the Weather_{ML} and $\textit{Snowclim}_{ML}$ configurations, with most of these sites again located in Switzerland (not shown). Although a smaller clustering radius during fold creation could mitigate this underestimation through more balanced fold SD distributions, the current clustering offers a fair evaluation of model performance, indicating a need for expanded data collection of high observed SD values.

The meteorological forcing data and *Snowclim* SD estimates also help XGBoost better estimate low-snow and snow-free conditions at unseen locations, particularly at low-elevation sites ($<$ 1000 m) near the end of the snow season. While the $\textit{Snowclim}_{ML}$ configuration performs best at high-elevation sites ($>$ 2500 m) during peak snow depth and the ablation period, it tends to overestimate SD at mid-elevation sites (1000 – 2500 m) near the end of the season, and at high-elevation sites at the beginning of the season, compared to the Combination_{ML} and Weather_{ML} configurations (not shown). Overall, however, the $\textit{Snowclim}_{ML}$ configuration yields the best performance (Fig. 6c). Under the spatio-temporal nested CV framework, site MAE during snow-covered periods decreases by at least 2.5 cm at 55% of sites, while only 16% exhibit an increase of 2.5 cm or more, relative to the Combination_{ML} configuration. Compared to the Weather_{ML} configuration, the difference in site MAE is minimal,

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

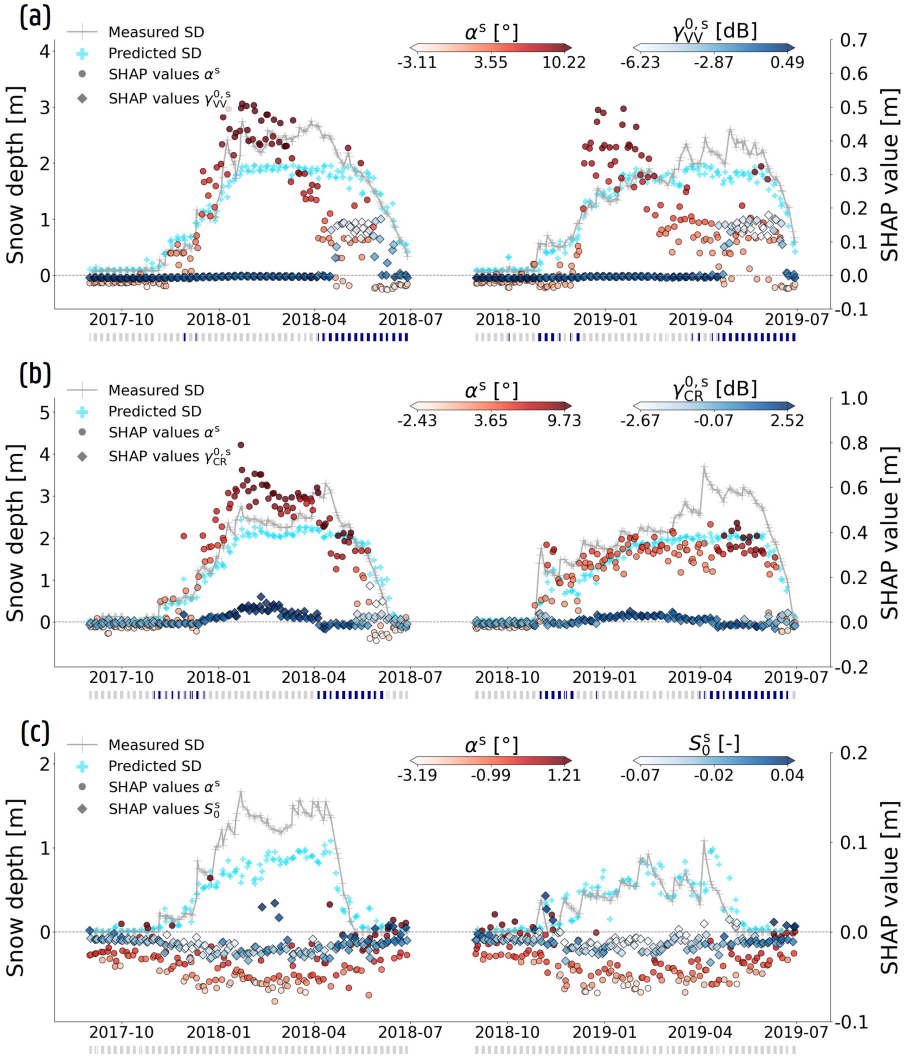

Figure 5. Time series for the Combination$_{ML}$ configuration within the spatial nested CV framework for three sites within Switzerland. The time series display measured (gray) and predicted (light blue) SD for the 2017-2018 and 2018-2019 snow seasons. SHAP values for two satellite variables, representing their contribution in time, are plotted as well, and colored according to their input value in XGBoost. The stripes below the time series indicate dry (light gray) or wet (dark blue) snow conditions according to CLMS SWS wet snow product.

505 and only significant for the spatio-temporal framework. These results thus indicate that directly using meteorological forcings can achieve comparable predictive performance, underscoring the potential of machine learning as a valuable complement to physically-based snow models, provided sufficient training data are available.

Although the inclusion of meteorological forcing data or modeled SD reduces the relative SHAP value FI of the S1 input features, these features still influence the SD predictions. In the spatial and spatio-temporal nested CV frameworks, S1 features

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

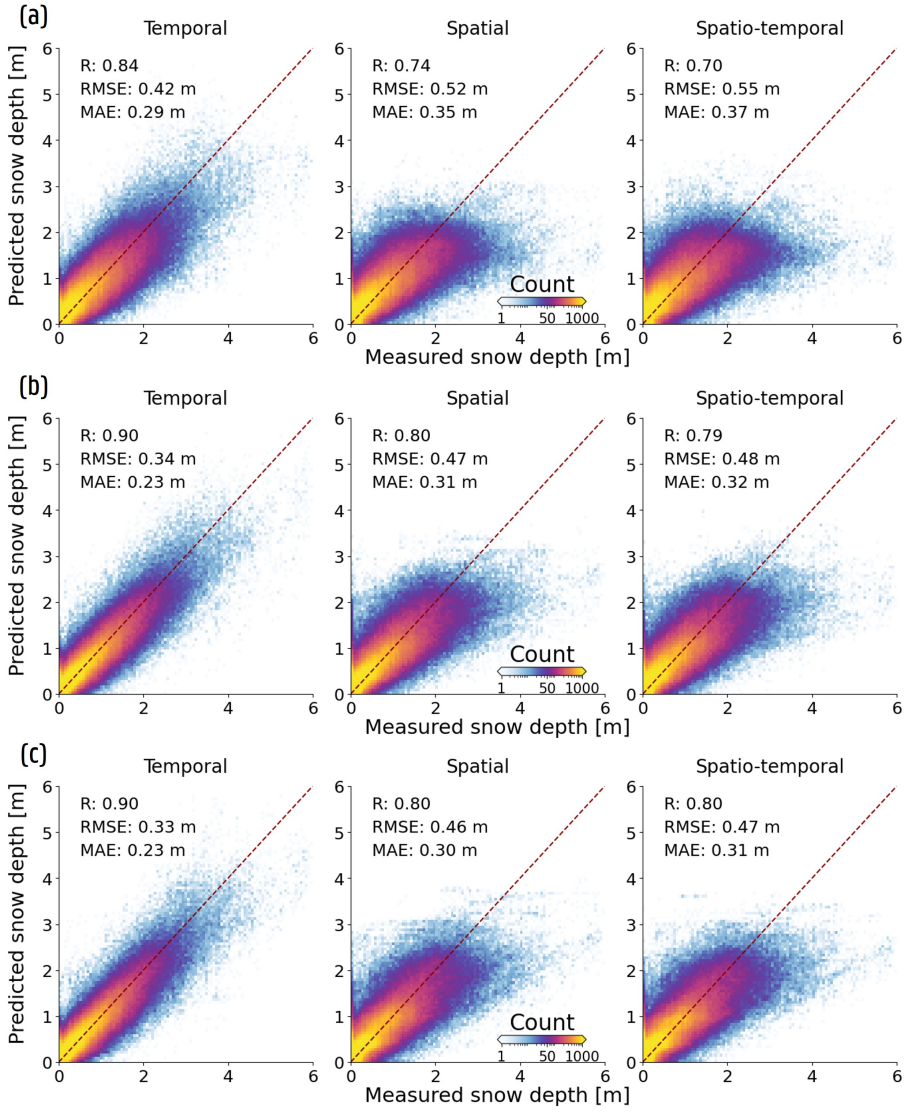

Figure 6. 2D histograms of measured vs. predicted SD for different ML configurations and nested CV frameworks: (a) Combination$_{\mathrm{ML}}$, (b) Weather$_{\mathrm{ML}}$, and (c) *Snowclim*$_{\mathrm{ML}}$. Performance metrics, excluding zero-measured SD, are shown in the left-upper corner of the subplots.

account for approximately 9% and ∼7% of the total SHAP value FI for the Weather$_{\mathrm{ML}}$ and *Snowclim*$_{\mathrm{ML}}$ configurations, respectively, with α^{s} and $\gamma^{0,\mathrm{s}}_{\mathrm{VV}}$ contributing the most. Similarly, the relative gain-based FI reveals a non-negligible role for the S1 features, with contributions of around 12% for both configurations and nested CV frameworks.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

4.4 Spatial snow depth prediction

López-Moreno et al. (2015) and Miller et al. (2022) emphasized that stationary site SD measurements are insufficient to represent the spatial variability within the snowpack. To address this limitation, Fig. 7 displays the SD predictions for the nine Dischma valley snow surveys, made with (a and e) the Combination$_{\mathrm{ML}}$ (b and f) the Weather$_{\mathrm{ML}}$ and (c and g) the *Snowclim*$_{\mathrm{ML}}$ configurations, all trained with spatial SD data included. Among these, the Combination$_{\mathrm{ML}}$ configuration exhibits the highest MAE for the 16 March 2017 snow survey (Fig. 7a), also when trained solely with stationary site data (Fig. C3). However, by including spatial training data, mountain ridges appear more distinctly in the predictions, and high SD predictions at steep locations and near ridges are more accurately corrected. This effect is also reflected in Fig. 7h, which shows that including spatial training data leads to a generally lower SHAP value FI for the topographic input features, with a pronounced drop for the slope. When trained without spatial data, XGBoost namely had incorrectly established a relationship between increasing slopes and predicted SD – explaining the previously high slope FI – leading to high SD at high-elevation areas with (relatively) steep slopes (Fig. C3a). This misrepresentation likely stemmed from the training dataset, that contains few high-elevation stationary measurement sites in steep-slope terrain – sites that moreover display deep snow accumulation during March (data not shown).

SD predictions further improve for both the Weather$_{\mathrm{ML}}$ and *Snowclim*$_{\mathrm{ML}}$ configurations, as shown not only in the 16 March 2017 snow survey, but also in the 2D histograms (Fig. 7b, c, f and g), and what was already clear from Fig. 6. Including spatial SD training data mainly leads to better estimates for lower observed SD, whereas the differences with the same configurations, trained without spatial data, are less clear for high ($\geq$ 2.5 m) SD (Fig. C3f and g). Although high SD values are generally better captured by including *Snowclim* SD estimates than by using meteorological forcings (Fig. 7f and g), their inclusion also leads to substantial overestimation for the 16 March 2017 snow survey, particularly at locations $\geq$ 1500 m. The overestimation introduced by the *Snowclim* SD estimates – on average 53 cm higher than observed values within the 1-2 m range – may also explain its 6 cm higher overall MAE compared to the Weather$_{\mathrm{ML}}$ configuration (Fig. 7b and c). Between approximately 1500 and 2500 m, the Weather$_{\mathrm{ML}}$ configuration shows a slightly negative bias of 4 cm, whereas *Snowclim*$_{\mathrm{ML}}$ maintains a consistent positive bias of around 17 cm. In the 2500 to 3200 m elevation range a similar pattern is noted, with the Weather$_{\mathrm{ML}}$ configuration underestimating the observed values by approximately 7 cm, while *Snowclim*$_{\mathrm{ML}}$ overestimates them by around 13 cm. At the highest elevations, both configurations display less consistent patterns, with instances of both under- and overestimation.

Nevertheless, even with spatial training data included, the Weather$_{\mathrm{ML}}$ and *Snowclim*$_{\mathrm{ML}}$ configurations still overestimate observed values, leading to overall biases of 7 cm and 17 cm across the nine snow surveys, respectively. Albeit it is harder to pinpoint the underlying cause for the Weather$_{\mathrm{ML}}$ configuration, the (strong) positive discrepancies may be driven by an overestimation of $P_{\mathrm{s,c}}$. Alternatively, it is also possible that XGBoost is unable to (indirectly) account for increases in snow bulk density, such that $P_{\mathrm{s,c}}$ may be accurate, but SD is nonetheless overestimated due to an implicit underestimation of snow density. Finally, large elevation differences within the relative coarse 500 m forcing data might also lead to inaccurate temperature representations, resulting in incorrect snowfall estimates or an under-/overestimation of the number of melt days.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

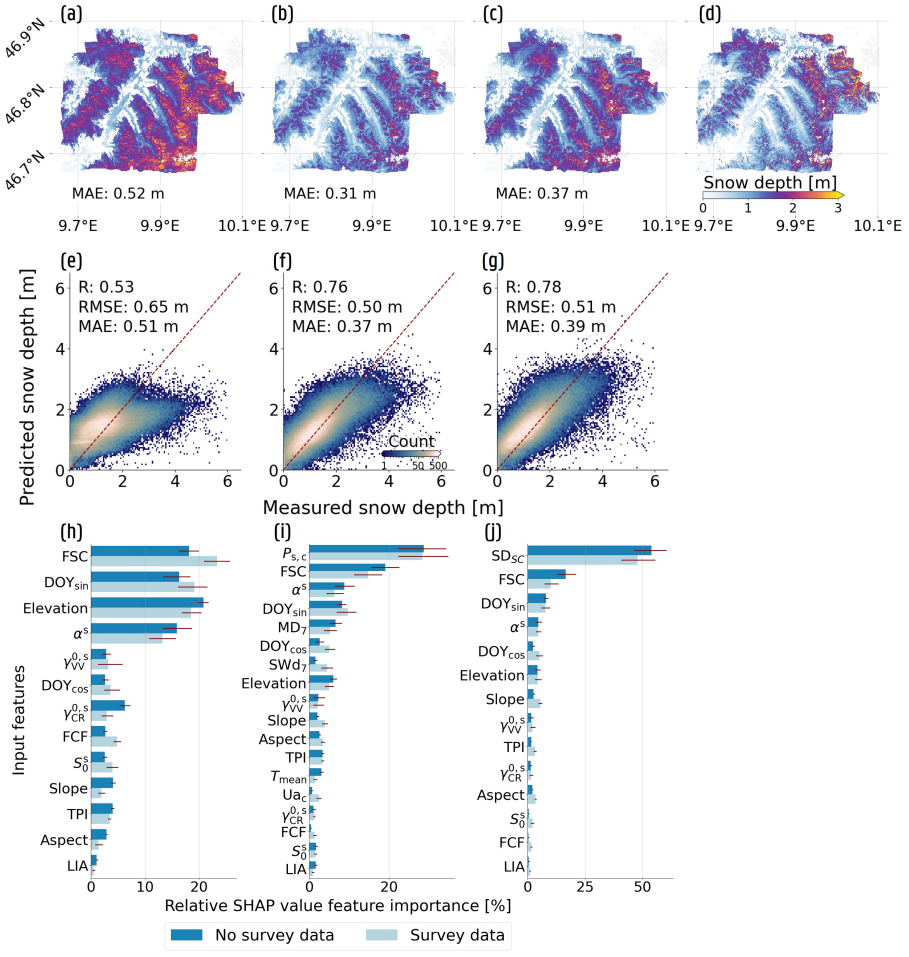

Figure 7. XGBoost spatial prediction performance for the Dischma valley snow surveys. The first column represents the Combination$_{ML}$ configuration, while the second and third column show the results for the Weather$_{ML}$ and *Snowclim*$_{ML}$ configurations, respectively. (a), (b) and (c) Display the spatial predictions for the 2017-03-16 snow survey, with (d) representing the measured SD. (e), (f) and (g) Show 2D histograms and performance metrics across all nine conducted snow surveys. Finally, (h), (i) and (j) display relative SHAP value FI, both when trained with or without spatial SD (survey) data. The 95% confidence interval is indicated with a red bar.

Similar as for the Combination$_{ML}$ configuration, model training with spatial SD data decreases the elevation SHAP value FI for both the Weather$_{ML}$ and *Snowclim*$_{ML}$ configurations (Fig. 7i and j). However, here, this reduction appears to be driven by the increased topographic complexity in the training dataset, which weakens the relationship between elevation and SD. Indeed, for both configurations, elevation contributes less consistently across space, while other topographic features play a greater role in adjusting predictions for slopes, mountain ridges and/or local depressions (not shown). In addition, α^s still emerges among the more important features, indicating that despite the decreased importance of this S1 variable, including it (and/or other S1 variables) can still influence the final predictions.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

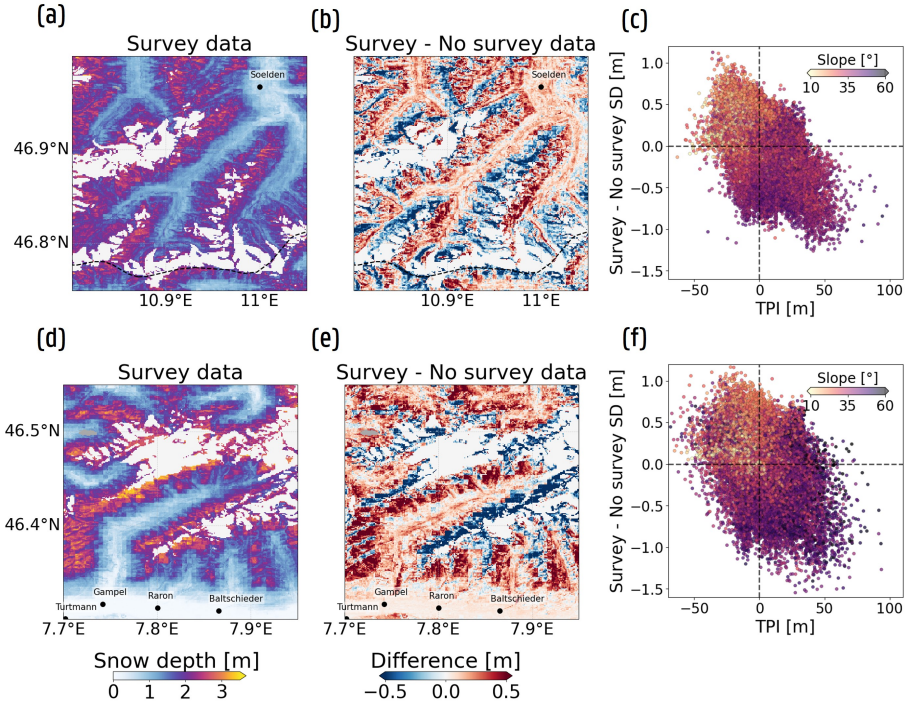

Figure 8. XGBoost SD predictions on 29 January 2018 over the Ötztal, Austria (top row) and the Lötschental, Switzerland (bottom row), using the $\text{Weather}_{\text{ML}}$ configuration. (a) and (d) mean SD estimates predicted with the nine separately trained XGBoost models utilizing spatial (survey) training data. (b) and (e) Difference maps between the mean SD prediction, trained using survey data and SD estimates predicted with XGBoost trained without spatial (survey) data. (c) and (f) Scatter plots of TPI and SD differences, colored according to aspect. Glaciers and lakes located within the valleys are colored in white and dark gray, respectively, and are excluded from the analyses.

Finally, Fig. 8 illustrates how the estimates are adjusted outside the Dischma valley, following the inclusion of spatial training data. The top row displays the SD estimates across the Ötztal region (Austria) on 29 January 2018, while the bottom row showcases the results for the Lötschental region (Switzerland), on the same date. For both valleys, training XGBoost with spatial SD data leads to improved spatial patterns, manifested by the appearance of mountain ridges and local depressions (Fig. 8a and d). These patterns also appear in the difference maps (Fig. 8b and e), and seem to be related to the TPI (Fig. 8c and f). Specifically, it appears that there is negative relationship between increasing TPI, a proxy for the microtopography present in the area, and the adjusted SD estimates. This indicates that after incorporating spatial training data, XGBoost assigns higher SD values to areas with negative TPI (i.e., locations situated below their surroundings) than it would when trained solely on stationary site data. On the contrary, places with high TPI values (i.e., mountain ridges) have been negatively adjusted now. While the relationship with the other topographic features (e.g., aspect) is less clear, Fig. 8c and f also indicate a dependency with the slope of the area. Steeper slopes appear to be corrected more negatively, while the opposite is true for more flat locations.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

5 Conclusions

In this study, we applied an extreme gradient boosting model (XGBoost) that integrates Sentinel-1 (S1) C-band synthetic aperture radar (SAR) backscatter or dual-polarized SAR polarimetry (PolSAR) observations, or both, along with regionally downscaled meteorological forcing data, and snow depth (SD) estimates from the physically-based model *Snowclim*, to further advance SD predictions across the European Alps. We implemented a threefold nested cross-validation (nested CV) framework to analyse model performance when predicting at either unseen time periods, unseen locations, or at unseen locations during time periods not covered by the training data. We also evaluated XGBoost's ability to generate spatial predictions, assessed the impact of spatial training data on SD estimates, and used feature importance analysis to better understand the relationship between the input features and model output.

Our results show that incorporating S1 C-band PolSAR observations, either in addition to or in place of backscatter, can slightly improve SD predictions compared to utilizing an ML model trained with only backscatter as S1 satellite variables. Although modest — likely due to the similar informational content and noise in both types of S1 variables — the improvements are most evident when predicting at known locations for unseen time periods, and less pronounced when generalizing to unseen locations. Thereby, feature importance analysis indicates that among the S1 variables, the polarimetric scattering angle (α) contributes the most. In addition, FI reveals that XGBoost consistently relies more on α than on the cross-polarization ratio (γ^0_{CR}) under both wet and dry snow conditions when both PolSAR and backscatter observations are included. XGBoost moreover actively exploits the drop in co-polarized (VV) backscatter (γ^0_{vv}) during the ablation period to refine its SD estimates. These results showcase the potential of using PolSAR observations within non-machine learning applications, e.g., within a conceptual model.

Next, we revealed that incorporating meteorological forcing data or *Snowclim* SD estimates substantially improve the predictions, because both help to capture interannual and site-specific variability. Besides improving the predictions for low-snow and snow-free conditions at low-elevation sites, particularly near the beginning and end of the snow season, these additional input features can help mitigate the underestimation of high observed ($\geq$ 2.5 m) SD values. Nonetheless, our results reveal that high SD values remain systematically underestimated, primarily due to their limited representation in the training dataset, and that the S1 C-band observations may still influence the predictions, reflected through their FI.

Finally, our results underscore the importance of incorporating spatial training data to more accurately capture topographic variability in SD estimates. This is primarily reflected in a negative relationship between topographic position index (TPI) and adjustments in SD estimates, relative to predictions from the same model trained without spatial data. With the inclusion of spatial training data, and meteorological forcings or modeled SD as additional input features, XGBoost predictions achieve a spatiotemporal Pearson correlation coefficient of 0.77 and a mean absolute error of approximately 40 cm when compared to airborne photogrammetry snow surveys conducted in the Dischma Valley, Switzerland.

Despite these advances, SD values above 3 m remain underestimated, emphasizing the need for more extensive snow monitoring and additional spatial snow surveys. Furthermore, XGBoost operates at the pixel level, limiting its ability to incorporate information from neighboring pixels. Finally, while improving SD estimates is a step toward addressing snow mass knowledge

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

gaps, future research should focus directly on snow mass estimation and integrating spatial dependencies through methods that incorporate information from adjacent pixels.

Code availability. The datasets with the ML experiments, used for visualization and analyses, will be made freely available after review.

Data availability. The Sentinel-1 data is freely available on the Alaska Satellite Facility's Vertex data portal (https://search.asf.alaska.edu). IMS and MODIS snow cover data can be downloaded at https://cmr.earthdata.nasa.gov/search/concepts/C1386246258-NSIDCV0.html and by using the earthaccess Python package ("MOD10A1F" and "MYD10A1F" collections), respectively. The SAR Wet Snow product can be sourced from https://land.copernicus.eu/en/products/snow/high-resolution-sar-wet-snow. Instructions to download the MSWX and MSWEP meteorological data can be found at https://www.gloh2o.org/. The Copernicus 30 m digital elevation model can be retrieved from the STAC collection described at https://hda.data.destination-earth.eu/ui/dataset/EO.DEM.DAT.COP-DEM_GLO-30-DTED. Finally, land cover data can be sourced from the Copernicus Land Monitoring Service via: https://land.copernicus.eu/en/products/global-dynamic-land-cover/copernicus-global-land-service-land-cover-100m-collection-3-epoch-2018-globe.

The majority of the in situ snow depth measurements can be freely downloaded from the various providers across the European Alps:

- https://www.doi.org/10.16904/envidat.406 (Switzerland)
- https://www.doi.org/10.16904/envidat.408 (Switzerland)
- https://donneespubliques.meteofrance.fr/?fond=recherche (France)
- http://meteotrentino.it/index.html#!/content?menuItemDesktop=111 (Italy)
- https://data.civis.bz.it//dataset/1512fa49-3e97-40d7-9bdb-2fc76c9efe3c/resource/9ca68cd2-2060-4a02-8a04-09b9d4acac40/download/dokumentationopendatameteode.pdf (Italy)
- https://dataset.api.hub.geosphere.at/app/frontend/station/historical/klima-v2-1d (Austria)

Additional snow depth measurements were collected through personal communication.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

Appendix A: Combining Terra and Aqua fractional snow cover data

To generate a combined daily fractional snow cover (FSC) product from the MODIS Terra and Aqua satellites, we applied a weighted averaging approach to merge the individual satellite observations. As such, we calculated for each location i and day t the weights of the individual FSC data ($\mathrm{W}_{\mathrm{sat}_{i,t}}$) using the following formula:

$$\mathrm{W}_{\mathrm{sat}_{i,t}} = \frac{1}{\mathrm{W}_{\mathrm{QA}}\mathrm{W}_{\mathrm{CP}}\mathrm{W}_{\mathrm{DD}}} \tag{A1}$$

in which W_{QA} is the weight associated with the quality flag (QA) of the satellite data, the latter ranging from 0 to 2 (with 0 indicating the highest quality); W_{CP} is the weight based on the number of days since the last cloud-free observation (CP; values between 0 and ∞); and W_{DD} is the weight accounting for the difference in days between the most recent cloud-free Terra and Aqua observations (DD; values between 0 and ∞), where 0 corresponds to the satellite with the most recent cloud-free data. The individual weights were calculated using the following formulas:

$$\mathrm{W}_{\mathrm{QA}} = \frac{1}{\mathrm{QA} + \frac{(2-\mathrm{QA})}{2}} \tag{A2}$$

$$\mathrm{W}_{\mathrm{CP}} = \frac{1}{(\mathrm{CP}+1)^2} \tag{A3}$$

$$\mathrm{W}_{\mathrm{DD}} = \frac{1}{(\mathrm{DD}+1)} \tag{A4}$$

Appendix B: *Snowclim* parameter setting and performance

Snowclim model parameters were calibrated with a two-step approach, involving a subset of the snow dataset. First, we identified the best working downscaling techniques by assessing model performance using standard parameters. Next, we used the downscaled forcing data to select the optimal parameters out of 1276 tested combinations, by applying a similar approach as described in Lute et al. (2022). A list of the final calibrated parameters is provided in Table B1.

Figure B1 shows the performance across the snow depth dataset, incorporating both data from the stationary sites and the photogrammetry snow surveys. Performance metrics are computed excluding zero-measured SD. When including zero values, the Pearson correlation coefficient increases to 0.85, while the RMSE and MAE decrease to 0.60 m and 0.36 m, respectively. Compared to the stationary sites alone, the *Snowclim* SD estimates show higher accuracy, with a MAE of 0.22 m (excluding zero-measured SD). In contrast, when only the nine photogrammetry snow surveys are compared with the corresponding *Snowclim* estimates, the MAE deteriorates to 0.81 m.

Appendix C: Additional tables and figures

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

Table B1. Parameter settings used to generate snow depth estimates across the European Alps. For the detailed description of the separate parameters, we refer to Lute et al. (2022).

Parameter	Abbreviated name	Calibrated values/methods	Units
Albedo algorithm	*albedo_opt*	Essery[a]	-
Momentum roughness length	z_0	10^{-5}	m
Heat and vapor roughness length	z_h	$z_0/10$	m
Maximum albedo	*albedo_max*	0.90	-
Maximum liquid water fraction	*lw_max*	0.1	-
Windless heat exchange coefficient	*E0*	0,1	$\mathrm{Wm^{-2}K^{-1}}$
Windless heat exchange coefficient flux application	*E0_app*	1	-
Windless heat exchange coefficient stability condition	*E0_stability*	2	-
Cold content threshold at which to start energy tax	$\mathrm{cc_0}$	-5000	$\mathrm{kJm^{-2}}$
Cold content range to tax	$\mathrm{cc_1}$	-5000	$\mathrm{kJm^{-2}}$
Maximum tax to apply to snow cover energy	maxtax	0.9	-
Snow cover energy flux smoothing window	*smooth_hrs*	24	h
Snow surface temperature augmentation	T_{add}	2	$^\circ$ C

[a] Essery et al. (2013).

Table C1. Overall performance metrics for the $\mathrm{PolSAR_{ML}}$, $\mathrm{Backscatter_{ML}}$ and $\mathrm{Combination_{ML}}$ configurations for the temporal nested CV framework. The values inside the brackets denote the metric values when zero-measured SD are included in the evaluation.

Metric	**$\mathrm{PolSAR_{ML}}$**	**$\mathrm{Backscatter_{ML}}$**	**$\mathrm{Combination_{ML}}$**
R [-]	0.83 (0.88)	0.83 (0.88)	0.84 (0.89)
MAE [m]	0.30 (0.19)	0.30 (0.19)	0.29 (0.18)
RMSE [m]	0.43 (0.33)	0.44 (0.34)	0.42 (0.33)

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

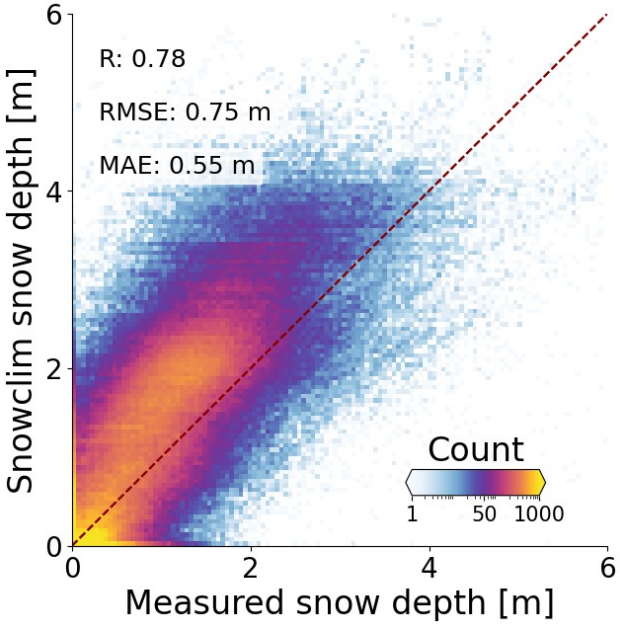

Figure B1. 2D histogram comparing measured SD (stationary sites and photogrammetry snow surveys) with SD estimates of the physically-based model *Snowclim*. Performance metrics do not incorporate zero-measured SD.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

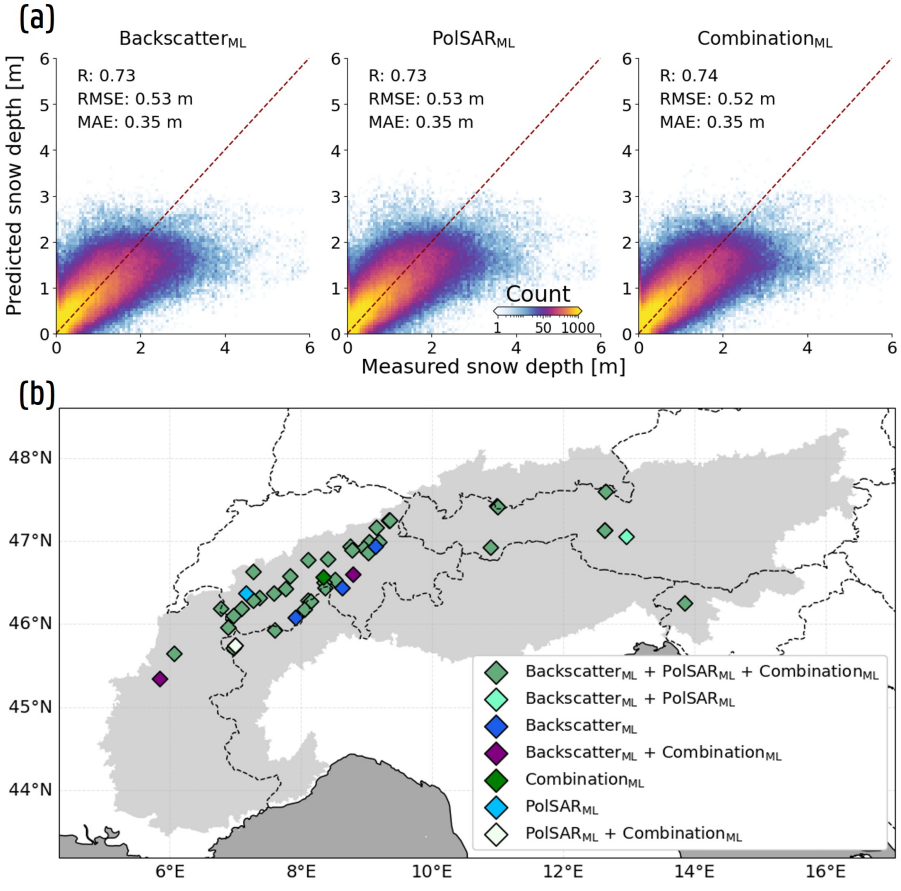

Figure C1. XGBoost prediction performance at unseen locations (spatial nested CV framework) across the Backscatter$_{ML}$, PolSAR$_{ML}$, and Combination$_{ML}$ configurations. (a) 2D histograms of observed versus predicted SD. (b) Sites with a mean site bias (observed minus predicted SD) $>$ 0.5 m. Measurement stations with fewer than 10 observations are excluded. Colors indicate the configurations for which the site bias exceeds 0.5 m.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

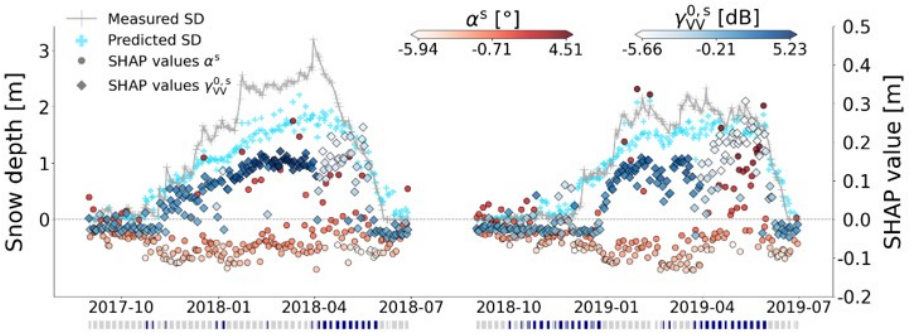

Figure C2. Time series for a measurement site located in Switzerland (46.67°N, 8.06°E). Predictions are made with the Combination$_{ML}$ configuration within the spatial nested CV framework. Measured (gray) and predicted (light blue) SD are displayed, along with SHAP values for two satellite variables, representing their contribution in time. The SHAP values are colored according to their input value in the model. The stripes below the time series indicate dry (light gray) or wet (dark blue) snow conditions according to the CLMS SWS wet snow product.

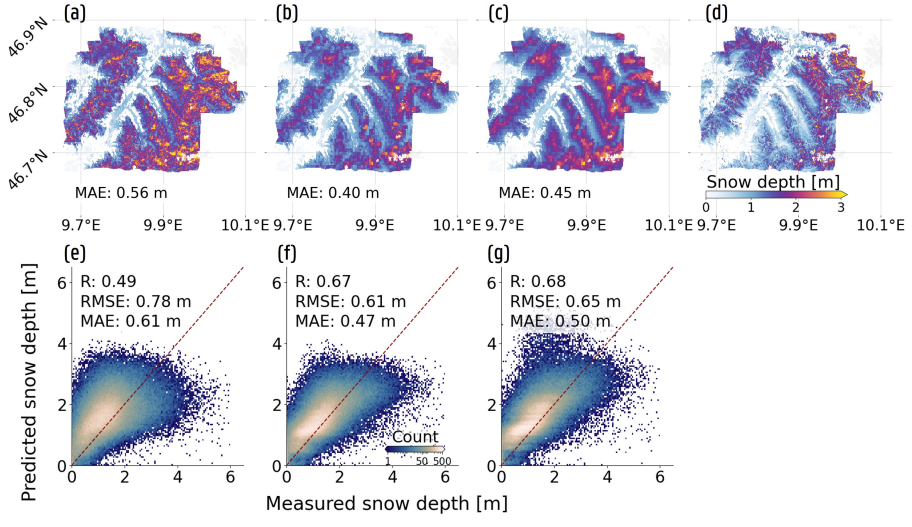

Figure C3. XGBoost spatial prediction performance for the Dischma valley snow surveys with no spatial data included during model training and tuning. The first column represents the Combination$_{ML}$ configuration, while the second and third column show the results for the Weather$_{ML}$ and *Snowclim*$_{ML}$ configurations, respectively. (a), (b) and (c) Display the spatial predictions for the 2017-03-16 snow survey, with (d) representing the measured SD. (e), (f) and (g) Show 2D histograms and performance metrics across all nine conducted snow surveys.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

Author contributions. LB, HL, DD, WW and GDL designed the study. LB and EB worked on the data preparation, and LB conducted the experiments discussed in the manuscript. The analyses, original draft preparation and visualization was performed by LB. All other authors reviewed and edited the manuscript before first submission.

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

Acknowledgements. This work was funded by the SNOWTRANE project (SR/00/407) of the Belgian Science Policy (Belspo). The Flemish Supercom-410 puter Center, funded by FWO and the Flemish Government, was used for computational resources and services. Next, the authors employed chatGPT to further improve the readability and language of the manuscript. Before submission, the authors edited and reviewed the manuscript, and take full responsibility for its content.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

References

Abdulaal, M. J., Casson, A. J., and Gaydecki, P.: Performance of nested vs. non-nested SVM cross-validation methods in visual BCI: Validation Study, in: 2018 26th European Signal Processing Conference (EUSIPCO), pp. 1680–1684, https://doi.org/10.23919/EUSIPCO.2018.8553102, 2018.

Aureliano, P.: Perimeter of the Alpine Convention, https://www.atlas.alpconv.org/layers/geonode:Alpine_Convention_Perimeter_2018_v2#more, 2020.

Bales, R. C., Molotch, N. P., Painter, T. H., Dettinger, M. D., Rice, R., and Dozier, J.: Mountain hydrology of the western United States, Water Resources Research, 42, https://doi.org/10.1029/2005WR004387, 2006.

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

Bartelt, P. and Lehning, M.: A physical SNOWPACK model for the Swiss avalanche warning: Part I: numerical model, Cold Regions Science and Technology, 35, 123–145, https://doi.org/https://doi.org/10.1016/S0165-232X(02)00074-5, 2002.

Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., van Dijk, A. I. J. M., McVicar, T. R., and Adler, R. F.: MSWEP V2 global 3-Hourly 0.1° precipitation: methodology and quantitative assessment, Bulletin of the American Meteorological Society, 100, 473 – 500, https://doi.org/10.1175/BAMS-D-17-0138.1, 2019.

Beck, H. E., van Dijk, A. I. J. M., Larraondo, P. R., McVicar, T. R., Pan, M., Dutra, E., and Miralles, D. G.: MSWX: global 3-Hourly 0.1° bias-corrected meteorological data including near-real-time updates and forecast ensembles, Bulletin of the American Meteorological Society, 103, E710 – E732, https://doi.org/10.1175/BAMS-D-21-0145.1, 2022.

Blancas, E.: Model selection done right: A gentle introduction to nested cross-validation, https://ploomber.io/blog/nested-cv/, accessed: 2025-06-02, 2021.

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

Brangers, I., Lievens, H., Getirana, A., and De Lannoy, G. J. M.: Sentinel-1 snow depth assimilation to improve river discharge estimates in the western European Alps, Water Resources Research, 60, e2023WR035 019, https://doi.org/https://doi.org/10.1029/2023WR035019, 2024a.

Brangers, I., Marshall, H.-P., De Lannoy, G., Dunmire, D., Mätzler, C., and Lievens, H.: Tower-based C-band radar measurements of an alpine snowpack, The Cryosphere, 18, 3177–3193, https://doi.org/10.5194/tc-18-3177-2024, 2024b.

Broxton, P., Ehsani, M. R., and Behrangi, A.: Improving mountain snowpack estimation using machine learning with Sentinel-1, the Airborne Snow Observatory, and University of Arizona snowpack data, Earth and Space Science, 11, e2023EA002 964, https://doi.org/10.1029/2023EA002964, 2024.

Buchhorn, M., Smets, B., Bertels, L., De Roo, B., Lesiv, M., Tsendbazar, N.-E., Herold, M., and Fritz, S.: Copernicus global land service: land cover 100 m: collection 3: epoch 2018: Globe, https://doi.org/10.5281/zenodo.3518038, 2020.

Bühle, L. J., Marty, M., Eberhard, L. A., Stoffel, A., Hafner, E. D., and Bühler, Y.: Spatially continuous snow depth mapping by aeroplane photogrammetry for annual peak of winter from 2017 to 2021 in open areas, The Cryosphere, 17, 3383–3408, https://doi.org/10.5194/tc-17-3383-2023, 2023.

Burakowski, E. and Magnusson, M.: Climate impacts on the winter tourism economy in the United States, 2012.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

Chen, T. and Guestrin, C.: XGBoost: A scalable tree boosting system, in: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pp. 785–794, Association for Computing Machinery, New York, NY, USA, https://doi.org/10.1145/2939672.2939785, 2016.

Cloude, S. and Pottier, E.: An entropy based classification scheme for land applications of polarimetric SAR, IEEE Transactions on Geoscience and Remote Sensing, 35, 68–78, https://doi.org/10.1109/36.551935, 1997.

Copernicus Land Monitoring Service: SAR wet snow 2016-present (raster 60 m), Europe, daily, https://land.copernicus.eu/en/products/snow/high-resolution-sar-wet-snow, 2024.

Craymer, M.: Geodetic toolbox, MATLAB Central File Exchange, 2022.

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

De Lannoy, G. J. M., Bechtold, M., Busschaert, L., Heyvaert, Z., Modanesi, S., Dunmire, D., Lievens, H., Getirana, A., and Massari, C.: Contributions of irrigation modeling, soil moisture and snow data assimilation to high-resolution water budget estimates over the Po basin: Progress towards digital replicas, Journal of Advances in Modeling Earth Systems, 16, e2024MS004 433, https://doi.org/https://doi.org/10.1029/2024MS004433, e2024MS004433 2024MS004433, 2024.

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

Dora, L., Agrawal, S., Panda, R., and Abraham, A.: Nested cross-validation based adaptive sparse representation algorithm and its application to pathological brain classification, Expert Systems with Applications, 114, 313–321, https://doi.org/https://doi.org/10.1016/j.eswa.2018.07.039, 2018.

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

Duan, S., Ullrich, P., Risser, M., and Rhoades, A.: Using temporal deep learning models to estimate daily snow water equivalent over the Rocky Mountains, Water Resources Research, 60, e2023WR035 009, https://doi.org/10.1029/2023WR035009, 2024.

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.

Essery, R., Morin, S., Lejeune, Y., and B Ménard, C.: A comparison of 1701 snow models using observations from an alpine site, Advances in Water Resources, 55, 131–148, https://doi.org/https://doi.org/10.1016/j.advwatres.2012.07.013, 2013.

Fiddes, J. and Gruber, S.: TopoSCALE v. 1.0: Downscaling gridded climate data in complex terrain, Geoscientific Model Development, 7, 387–405, https://doi.org/10.5194/gmd-7-387-2014, 2014.

Gascoin, S., Luojus, K., Nagler, T., Lievens, H., Masiokas, M., Jonas, T., Zheng, Z., and De Rosnay, P.: Remote sensing of mountain snow from space: Status and recommendations, Frontiers in Earth Science, 12, 1381 323, https://doi.org/10.3389/feart.2024.1381323, 2024.

Groisman, P. Y., Karl, T. R., Knight, R. W., and Stenchikov, G. L.: Changes of snow cover, temperature, and radiative heat balance over the Northern Hemisphere, Journal of Climate, 7, 1633–1656, https://doi.org/10.1175/1520-0442(1994)007<1633:COSCTA>2.0.CO;2, 1994.

Hall, D. and Riggs, G.: MODIS/Aqua CGF Snow Cover Daily L3 Global 500m SIN Grid, Version 61, https://doi.org/10.5067/MODIS/MYD10A1F.061, 2020a.

Hall, D. and Riggs, G.: MODIS/Terra CGF Snow Cover Daily L3 Global 500m SIN Grid, Version 61, https://doi.org/10.5067/MODIS/MOD10A1F.061, 2020b.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

Hill, D. F., Burakowski, E. A., Crumley, R. L., Keon, J., Hu, J. M., Arendt, A. A., Wikstrom Jones, K., and Wolken, G. J.: Converting snow depth to snow water equivalent using climatological variables, The Cryosphere, 13, 1767–1784, https://doi.org/10.5194/tc-13-1767-2019, 2019.

Jans, J.-F., Beernaert, E., De Breuck, M., Brangers, I., Dunmire, D., De Lannoy, G., and Lievens, H.: Sensitivity of Sentinel-1 C-band SAR backscatter, polarimetry and interferometry to snow accumulation in the Alps, Remote Sensing of Environment, 316, 114 477, https://doi.org/10.1016/j.rse.2024.114477, 2025.

Jennings, K. S., Winchell, T. S., Livneh, B., and Molotch, N. P.: Spatial variation of the rain–snow temperature threshold across the Northern Hemisphere, Nature Communications, 9, 1148, 2018.

Landschoot, S., Waegeman, W., Audenaert, K., Vandepitte, J., Haesaert, G., and De Baets, B.: Toward a reliable evaluation of forecasting systems for plant diseases: A case study using Fusarium head blight of wheat, Plant Disease, 96, 889–896, https://doi.org/10.1094/PDIS-08-11-0665, 2012.

Lievens, H., Demuzere, M., Marshall, H.-P., Reichle, R. H., Brucker, L., Brangers, I., de Rosnay, P., Dumont, M., Girotto, M., Immerzeel, W. W., Jonas, T., Kim, E. J., Koch, I., Marty, C., Saloranta, T., Schöber, J., and De Lannoy, G. J. M.: Snow depth variability in the Northern Hemisphere mountains observed from space, Nature Communications, 10, 1–12, https://doi.org/10.1038/s41467-019-12566-y, 2019.

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

Liston, G. E. and Elder, K.: A meteorological distribution system for high-resolution terrestrial modeling (MicroMet), Journal of Hydrometeorology, 7, 217–234, https://doi.org/10.1175/JHM486.1, 2006.

Luce, C., Staab, B., Kramer, M., Wenger, S., Isaak, D., and McConnell, C.: Sensitivity of summer stream temperatures to climate variability in the Pacific Northwest, Water Resources Research, 50, 3428–3443, https://doi.org/10.1002/2013WR014329, 2014.

Lundberg, S. M. and Lee, S.-I.: A unified approach to interpreting model predictions, in: Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS'17, p. 4768–4777, Curran Associates Inc., Red Hook, NY, USA, ISBN 9781510860964, 2017.

Lundblad, E. R., Wright, D. J., Miller, J., Larkin, E. M., Rinehart, R., Naar, D. F., Donahue, B. T., Anderson, S. M., and Battista, T.: A benthic terrain classification scheme for American Samoa, Marine Geodesy, 29, 89–111, https://doi.org/10.1080/01490410600738021, 2006.

Lute, A. C., Abatzoglou, J., and Link, T.: SnowClim v1.0: High-resolution snow model and data for the western United States, Geoscientific Model Development, 15, 5045–5071, https://doi.org/10.5194/gmd-15-5045-2022, 2022.

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

Marin, C., Bertoldi, G., Premier, V., Callegari, M., Brida, C., Hürkamp, K., Tschiersch, J., Zebisch, M., and Notarnicola, C.: Use of Sentinel-1 radar observations to evaluate snowmelt dynamics in alpine regions, The Cryosphere, 14, 935–956, https://doi.org/10.5194/tc-14-935-2020, 2020.

Marks, D. and Dozier, J.: Climate and energy exchange at the snow surface in the alpine region of the Sierra Nevada: 2. Snow cover energy balance, Water Resources Research, 28, 3043–3054, https://doi.org/10.1029/92WR01483, 1992.

Marty, M., Bühler, Y., and Ginzler, C.: Snow depth mapping, https://doi.org/http://dx.doi.org/10.16904/envidat.62, 2019.

Matiu, M., Crespi, A., Bertoldi, G., Carmagnola, C. M., Marty, C., Morin, S., Schöner, W., Cat Berro, D., Chiogna, G., De Gregorio, L., Kotlarski, S., Majone, B., Resch, G., Terzago, S., Valt, M., Beozzo, W., Cianfarra, P., Gouttevin, I., Marcolini, G., Notarnicola, C.,

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

Petitta, M., Scherrer, S. C., Strasser, U., Winkler, M., Zebisch, M., Cicogna, A., Cremonini, R., Debernardi, A., Faletto, M., Gaddo, M., Giovannini, L., Mercalli, L., Soubeyroux, J., Sušnik, A., Trenti, A., Urbani, S., and Weilguni, V.: Observed snow depth trends in the European Alps: 1971 to 2019, The Cryosphere, 15, 1343–1382, https://doi.org/10.5194/tc-15-1343-2021, 2021.

Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., and Nauss, T.: Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation, Environmental Modelling & Software, 101, 1–9, https://doi.org/10.1016/j.envsoft.2017.12.001, 2018.

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

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

Notarnicola, C.: Overall negative trends for snow cover extent and duration in global mountain regions over 1982–2020, Scientific Reports, 12, 13 731, https://doi.org/10.1038/s41598-022-16743-w, 2022.

Painter, T. H., Berisford, D. F., Boardman, J. W., Bormann, K. J., Deems, J. S., Gehrke, F., Hedrick, A., Joyce, M., Laidlaw, R., Marks, D., et al.: The Airborne Snow Observatory: Fusion of scanning lidar, imaging spectrometer, and physically-based modeling for mapping snow water equivalent and snow albedo, Remote Sensing of Environment, 184, 139–152, https://doi.org/10.1016/j.rse.2016.06.018, 2016.

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

Parvandeh, S., Yeh, H.-W., Paulus, M. P., and McKinney, B. A.: Consensus features nested cross-validation, Bioinformatics, 36, 3093–3098, https://doi.org/10.1093/bioinformatics/btaa046, 2020.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V.and Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine learning in Python, Journal of Machine Learning Research, 12, 2825–2830, 2011.

Pomarol Moya, O., Nussbaum, M., Mehrkanoon, S., Kraaijenbrink, P. D. A., Gouttevin, I., Karssenberg, D., and Immerzeel, W. W.: Improving forecasts of snow water equivalent with hybrid machine learning, EGUsphere, 2025, 1–23, https://doi.org/10.5194/egusphere-2025-1845, 2025.

PyCaret: PyCaret: An open-source, low-code machine learning library, Version 1.0.0, Available at https://www.pycaret.org/, 2020.

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

RGI 7.0 Consortium: Randolph glacier inventory - A dataset of global glacier uutlines, Version 7.0, https://doi.org/10.5067/f6jmovy5navz, online access: https://doi.org/10.5067/f6jmovy5navz, 2023.

Schwanghart, W. and Scherler, D.: TopoToolbox 2–MATLAB-based software for topographic analysis and modeling in Earth surface sciences, Earth Surface Dynamics, 2, 1–7, https://doi.org/10.5194/esurf-2-1-2014, 2014.

Seeherman, J. and Liu, Y.: Effects of extraordinary snowfall on traffic safety, Accident Analysis & Prevention, 81, 194–203, https://doi.org/10.1016/j.aap.2015.04.029, 2015.

Small, D.: Flattening gamma: Radiometric terrain correction for SAR imagery, IEEE Transactions on Geoscience and Remote Sensing, 49, 3081–3093, https://doi.org/10.1109/TGRS.2011.2120616, 2011.

https://doi.org/10.5194/egusphere-2025-3327
Preprint. Discussion started: 30 July 2025

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

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

Tsai, Y.-L. S., Dietz, A., Oppelt, N., and Kuenzer, C.: Wet and dry snow detection using Sentinel-1 SAR data for mountainous areas with a machine learning technique, Remote Sensing, 11, 895, https://doi.org/10.3390/rs11080895, 2019.

Tsang, L., Durand, M., Derksen, C., Barros, A. P., Kang, D.-H., Lievens, H., Marshall, H.-P., Zhu, J., Johnson, J., King, J., Lemmetyinen, J., Sandells, M., Rutter, N., Siqueira, P., Nolin, A., Osmanoglu, B., Vuyovich, C., Kim, E., Taylor, D., Merkouriadi, I., Brucker, L., Navari, M., Dumont, M., Kelly, R., Kim, R. S., Liao, T.-H., Borah, F., and Xu, X.: Review article: Global monitoring of snow water equivalent using high-frequency radar remote sensing, The Cryosphere, 16, 3531–3573, https://doi.org/10.5194/tc-16-3531-2022, 2022.

U.S. National Ice Center: IMS daily Northern Hemisphere snow and ice analysis at 4 km and 24 km resolution, National Snow and Ice Data Center, 2008.

Varade, D., Singh, G., Dikshit, O., and Manickam, S.: Identification of snow using fully polarimetric SAR data based on entropy and anisotropy, Water Resources Research, 56, e2019WR025 449, https://doi.org/10.1029/2019WR025449, 2020.

Wang, F. and Tian, D.: Multivariate bias correction and downscaling of climate models with trend-preserving deep learning, Climate Dynamics, 62, 9651–9672, https://doi.org/10.1007/s00382-024-07406-9, 2024.

Warren, S. G.: Optical properties of snow, Reviews of Geophysics, 20, 67–89, https://doi.org/10.1029/RG020i001p00067, 1982.

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

World Meteorological Organization: Rapid changes in cryosphere demand urgent, coordinated action, https://wmo.int/news/media-centre/rapid-changes-cryosphere-demand-urgent-coordinated-action, 2023.

Xiong, C., Yang, J., Pan, J., Lei, Y., and Shi, J.: Mountain snow depth retrieval from optical and passive microwave remote sensing using machine learning, IEEE Geoscience and Remote Sensing Letters, 19, 1–5, https://doi.org/10.1109/LGRS.2022.3226204, 2022.