https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

Machine Learning and Explainable AI for Sentinel-1–Based Snow Depth Estimation: Insights from Three Different Mountainous Regions

R Chandra Prabha[1] and RAAJ Ramsankaran[1]

[1]Hydro-Remote Sensing Applications (H-RSA) Group, Department of Civil Engineering, Indian Institute of Technology Bombay, Mumbai, India.

Correspondence to: RAAJ Ramsankaran (ramsankaran@civil.iitb.ac.in)

Abstract. Mountain snow plays an important role in regulating freshwater resources and climate processes, making accurate monitoring of its depth and distribution essential for water resource planning and climate modelling. However, monitoring snow depth (SD) in mountainous regions remains challenging due to the extreme topographic and climatic conditions. Synthetic Aperture Radar (SAR)-based remote sensing is particularly useful in such regions due to its high spatial and temporal resolution and its ability to penetrate deep snow. This study focuses on improving Sentinel-1 C-band SAR backscatter-based SD estimation, which supports large-scale, continuous monitoring but faces limitations in vegetated, shallow, or wet snow conditions. To address these limitations, the present study developed a machine learning (ML) framework that integrates change detection-based backscatter indices, direct backscatter variables, and auxiliary datasets to improve SD estimation over the existing ML models. The framework is designed for implementation using preprocessed Sentinel-1 data from Google Earth Engine, thereby reducing computational requirements and enabling scalability across regions and time periods. It is implemented across three different mountainous regions, the Colorado Rocky Mountains, the European Alps and the Indian Western Himalayas, each representing diverse characteristics. Results demonstrate that the developed ML model significantly outperforms existing approaches in all regions. As the physical scattering mechanisms underlying the backscatter-based methods are not yet fully understood, explainable AI techniques are employed to interpret model behaviour and investigate how different input variables contribute under varying conditions. The analysis reveals distinct patterns across seasons, snow types, vegetation conditions, and regions, providing physical insights into C-band backscatter-based SD estimation.

1 Introduction

Seasonal snow is a critical natural resource that stores freshwater during winter and releases it during summer, sustaining rivers and groundwater systems (Barnett et al., 2005). Its high albedo also plays a key role in regulating temperature by reflecting incoming solar radiation (Qu & Hall, 2006), thereby influencing both hydrological and climatic systems. Mountainous regions

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

are particularly important in this context, as they receive substantial snowfall and serve as a major source of water for downstream ecosystems, agriculture, hydropower generation, and human consumption (Viviroli et al., 2003, 2007). Accurately quantifying this stored water and understanding its broader impacts requires reliable measurements of snow depth (SD), a fundamental parameter for estimating snow water storage and recognized as one of the Essential Climate Variables (ECVs) within the Global Climate Observing System (GCOS). However, monitoring SD in mountainous regions is extremely challenging due to complex terrain, severe weather conditions, high spatial variability, and limited accessibility, which makes ground-based observations sparse and insufficient. Remote sensing offers a better alternative, providing large-scale, spatially and temporally continuous data for improved SD monitoring (Awasthi & Varade, 2021; Gascoin et al., 2024).

Several remote sensing-based methods exist for estimating SD in mountainous regions. Among them, high-resolution optical stereo imaging-based DSM differencing provides detailed estimates but is affected by cloud cover and illumination conditions (Bühler et al., 2015; Marti et al., 2016). Airborne LiDAR provides highly accurate, high-resolution measurements, but can be operated only in favourable weather conditions and selected regions (Deems et al., 2013). Microwave remote sensing approaches overcome these limitations by providing data in all weather conditions. Currently, operational SD products are derived from passive microwave observations, which offer spatially and temporally continuous coverage at a global scale (Luojus et al., 2021; Tedesco & Jeyaratnam, 2019). However, due to its coarse resolution (typically 10–25 km) and limited penetration capability, it cannot accurately capture the depth and spatial variability of snow in mountainous regions (Tanniru & Ramsankaran, 2023). In such cases, SAR-based methods are more suitable due to their high spatial resolution and deeper penetration capability (Snehmani et al., 2015). Different SAR satellites operate at different wavelengths and polarization modes, based on which various polarimetric and interferometric approaches have been developed to estimate SD from the signal's intensity and phase information (Awasthi et al., 2021; Palomaki & Sproles, 2023). Currently, the implementation of these methods remains limited to regional scales due to restricted SAR data availability, large data volumes, high computational requirements, and their sensitivity to snowpack properties that require ground-based information (Tsang et al., 2022).

With the launch of the Sentinel-1(S1) mission, open-access, high-resolution C-band SAR data became available with continuous spatial and temporal coverage, opening new opportunities for SAR-based, large-scale, and continuous SD estimation. The study by (Lievens et al., 2019) identified that the ratio of cross-to co-polarized backscatter (VH/VV) is related to SD and used it in an empirical change detection approach to produce the first large-scale SAR-based SD product for the mountainous regions of the Northern Hemisphere. Following this work, several regional studies have been conducted in areas such as the European Alps, Rocky Mountains, Andes, and Himalayas, focusing on algorithm refinements, performance evaluation, and the use of machine learning techniques to better capture the backscatter–SD relationship (Broxton et al., 2024; Bulovic et al., 2025; Chandra Prabha et al., 2025; Daudt et al., 2023; Dunmire et al., 2024; Hoppinen et al., 2024; Jans et al., 2024; Lievens et al., 2022; Zhu et al., 2022). These studies revealed that the strength of this relationship varies by region, with some areas showing little or no correlation and others showing a moderate to good relationship, influenced by multiple

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

environmental and snowpack conditions. Despite these advances, the physical mechanisms underlying C-band backscatter in snow are not yet fully understood. Tower-based experiments have been conducted to investigate these processes, with additional experiments continuing to improve the understanding (Brangers et al., 2024). Current understanding is that the SAR signal interacts with snow grains and layers, resulting in volume scattering, with backscatter increasing as SD increases. This backscatter-based approach is influenced by snowpack properties such as wetness, density, grain size, layering and melt–freeze cycles, and is most suitable for deeper snowpacks (1–3 m) with little to no vegetation (Hoppinen et al., 2024; Lievens et al., 2022).

Recent studies have shown that in conditions where backscatter alone does not adequately capture SD variability, incorporating auxiliary variables such as snow cover duration and topographic parameters improves estimation accuracy, particularly in areas with shallow snow, wet snow, or forest cover. It has also been observed that dividing the region into zones with similar snowpack characteristics improves the estimates by accounting for the variable backscatter-SD relationship (Chandra Prabha et al., 2025; Dunmire et al., 2024). It was further observed that change detection–based backscatter indices are more strongly correlated with SD than direct backscatter variables used in existing ML-based studies (Daudt et al., 2023; Dunmire et al., 2024; Zhu et al., 2022). Moreover, besides the initially hypothesized cross- and co-polarization ratio, different backscatter variables exhibit stronger relationships with SD in specific scenarios, such as VV polarization in vegetated areas and VH polarization in regions with very deep snowpacks (Chandra Prabha et al., 2025; Lievens et al., 2022).

Building on this current understanding, the present study is motivated by two primary objectives: (i) to develop an improved, scalable SAR-based SD estimation method, and (ii) to gain a better understanding of the physical mechanisms underlying C-band backscatter–based SD estimation. The study therefore proposes a new ML-based SD estimation approach that integrates backscatter change detection–based indices with direct backscatter variables and supports cloud-based implementation on the Google Earth Engine platform, making it both scalable and computationally efficient for large-scale applications. Furthermore, explainable AI techniques are employed to evaluate the contribution of different variables to SD estimation and to examine how C-band backscatter–SD relationships vary across different conditions. The method is implemented over three contrasting mountainous regions, the Colorado Rocky Mountains, the European Alps, and the Indian Western Himalayas, representing diverse snowpack, climatic, and terrain conditions. This implementation not only aims to improve estimation accuracy but also provides insights into why and how C-band backscatter relationships differ across regions. The work is expected to advance both operational SD monitoring in mountains and the broader understanding of C-band radar–snow interactions across diverse environmental conditions.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

2 Study Area and Datasets

2.1 Study Area

This study focuses on three mountain regions: the Colorado Rocky Mountains (CRM), the European Alps (EA), and the Indian Western Himalayas (IWH) (Fig. 1a). Each region has distinct climatic, topographic, and ecological characteristics, described in detail in Sect. 2.1.1 to 2.1.3.

2.1.1 Colorado Rocky Mountains

The CRM, part of the Rocky Mountain range, are located in the interior of the North American continent covering ~133,688 km². This region serves as a critical headwaters area, where seasonal snowmelt contributes 70–80% of annual surface water, supporting agriculture, urban supply, ecosystems, and downstream states (Li et al., 2017). The region receives snowfall from October to April, driven by Pacific frontal systems (Paulson et al., 1991). The topography of the region is highly variable, with a median elevation of ~2,601 m and peaks reaching up to ~4,390 m (Fig. 1b). Climatic conditions vary with elevation, with the highest areas having cold tundra conditions and long-lasting snow cover. Mid to high elevations experience cold continental climates with snowy winters and cool summers. Lower valleys and foothills have warmer and drier semi-arid conditions (Beck et al., 2018). The type of snow present here is also diverse, with the Montane Forest class being the most widespread (46.2%), followed by Prairie (28.5%), Boreal Forest (12.8%), Tundra (6.5%), and areas with ephemeral or no snow (6%) (Fig. 1e) (Sturm & Liston, 2021). Land cover is primarily composed of rangelands and forests, with large areas of coniferous trees, small patches of deciduous trees and subalpine–alpine shrubs near the treeline (Karra et al., 2021). Overall, the snowpack in the CRM is relatively shallow compared to EA and IWH and has a good network of ground observation stations (Fig. 1b).

2.1.2 European Alps

The EA are the highest and most extensive mountain chain in central Europe, covering ~188,947 sq.km. It plays a crucial role as the "water tower of Europe," supplying meltwater to major rivers such as the Rhine, Rhône, Po, and Danube, which provide freshwater to millions of people across the continent (Viviroli et al., 2007). Snowfall occurs primarily from October to May, driven by moist air masses from the Atlantic Ocean and Mediterranean Sea (Sodemann & Zubler, 2010). The terrain is rugged, with a median elevation of ~1,178 m and peaks reaching up to ~4,700 m (Fig. 1c). Climatic conditions vary with altitude, ranging from temperate oceanic in the northern foothills and temperate humid subtropical in the southern margins to humid continental in interior valleys, transitioning to subarctic and alpine tundra above ~2,500 m (Beck et al., 2018). Snow types include tundra (11.8%), boreal forest (0.8%), maritime (20%), ephemeral or no snow (44%), prairie (4.8%), montane forest (17.6%), and ice (1%), with tundra, montane forest, and maritime snow classes being more extensive (Fig. 1f) (Sturm & Liston,

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

2021). Land cover mainly consists of forests and rangelands, followed by snow and ice-covered areas, with vegetation transitioning from broadleaf forests in lowlands to conifer forests, alpine tundra meadows, and nival barren landscapes at the highest elevations (Karra et al., 2021; Nagy, 2006). It has moderate SD, deeper than the CRM region but shallower than IWH, and has an extensive ground observation network (Fig. 1c).

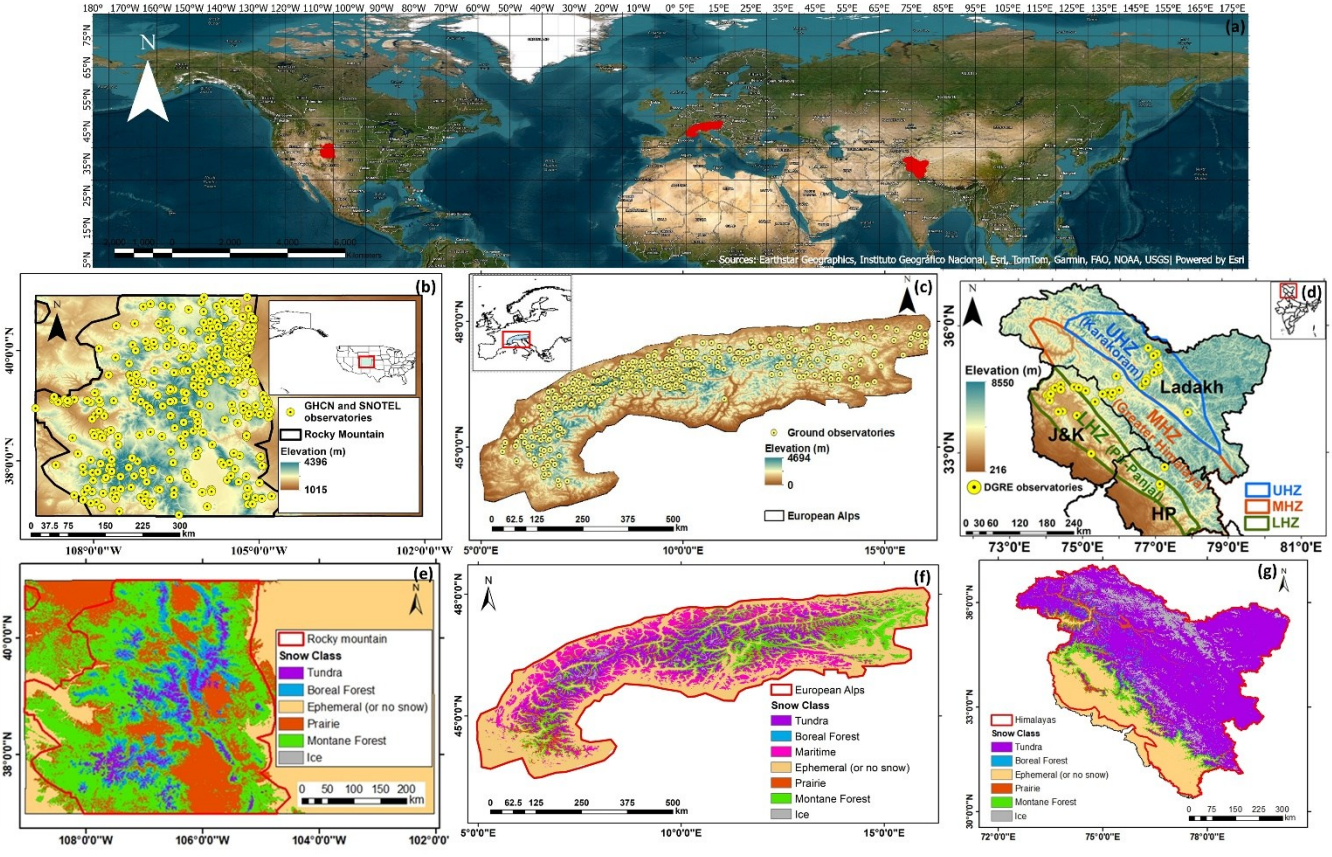


Figure 1. Global map showing the three study regions (a), with regional maps illustrating elevation characteristics and ground observation stations (b-d; d is reproduced from Singh et al. (2024)), and snow classes (e-g) for CRM, EA, and IWH. Basemap sources: Earthstar Geographics, Instituto Geográfico Nacional, Esri, Tom Tom, Garmin, FAO, NOAA, USGS | Powered by Esri. Elevation data are from SRTM Plus (Version 3.0, NASA JPL), and snow class data are from Sturm and Liston (2021).

2.1.3 Indian Western Himalayas

The IWH, covering ~270,369 sq. km, forms the western portion of the Himalayan mountains within India. This region is a critical water source for major rivers, including the Indus, Chenab, Beas, Ravi, and Ganga, with snow and ice melt contributing 50–80% of their flow, supporting agriculture, drinking water, and hydropower (Immerzeel et al., 2009; Jeelani et al., 2012). Snowfall primarily occurs from December to March, driven by western disturbances, with occasional high-elevation snowfall

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

outside this period influenced by the summer monsoon and transient convective systems. The terrain is steep and rugged, with a median elevation of ~4,399 m and peaks reaching ~8,500 m (Fig. 1d). Climate conditions vary greatly with altitude and across three broad zones: the Lower (LHZ), Middle (MHZ), and Upper Himalayan Zone (UHZ), shown in Fig. 1d. The LHZ experiences high precipitation and a temperate to cold climate, while the MHZ is generally cold with dry summers, transitioning to polar tundra conditions with extensive glacier coverage at higher elevations. The UHZ is extremely cold and arid, encompassing cold desert, steppe, polar tundra, and frost zones, with long-lasting snow and extensive glaciers (Beck et al., 2018; Sharma, 2000; S. Sharma & Ganju, 2000). Snow type is dominated by Tundra (61.3%), followed by ice (10%), montane forest (5%), prairie (3.4%), boreal forest (2.2%), and ephemeral or no snow (18.1%), shown in Fig. 1g (Sturm & Liston, 2021). Land cover consists mainly of forests in the LHZ, with subtropical broadleaf and pine species, while MHZ and UHZ are largely bare ground and rangelands (Karra et al., 2021). The IWH has the deepest snowpack among the three regions but a sparse network of ground observation stations.

2.2 Datasets

2.2.1 Sentinel-1

The S1 mission is a C-band SAR constellation operating at a wavelength of ~5.6 cm, consisting of S1-A (launched in 2014), S1-B (launched in 2016 and ceased operations in 2021), S1-C (launched in 2024), and the upcoming S1-D. The satellites primarily operate in interferometric wide swath (IW) mode over land, acquiring dual-polarization data (VV and VH) with a swath width of 250 km and a spatial resolution of 5 × 20 m. It acquires data in both ascending and descending orbits with a revisit frequency of 6 to 12 days.

For this study, S1 data was accessed through the GEE platform for the entire study period (2016–2023 for CRM and EA, and 2016–2019 for IWH). GEE provides images preprocessed for precise orbit correction, GRD border and thermal noise removal, radiometric calibration, and terrain correction. Since radiometric terrain flattening is not done in the default GEE product, which is particularly necessary for mountainous regions, it is applied in GEE following the method proposed by (Vollrath et al., 2020). The images are then resampled to 500m resolution by averaging in a linear scale, and areas affected by layover and shadow are masked.

Since differences in incidence angle between relative orbits can strongly influence backscatter, incidence angle normalization is done following the approach of (Lievens et al., 2019). For each pixel over the entire study period, the mean backscatter of each relative orbit is rescaled to the overall mean across all orbits. Outliers are then identified and replaced with the mean backscatter value of the same pixel.

From VV and VH backscatter, the ratio of VH to VV is also calculated,

$$\gamma^{\circ(t)}_{cr} = \gamma^{\circ(t)}_{vh} / \gamma^{\circ(t)}_{vv} \quad (1)$$

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

In addition to direct backscatter-based variables, three change detection-based indices are derived, adopted from (Lievens et al., 2019). Each index is initialized at 0 at the start of a water year , which begins on October 1. For each relative orbit, the difference between consecutive backscatter observations is then cumulatively added. The index is set to 0 whenever the value becomes negative or when optical data indicates no snow cover.

$$I_{vv}^{(t)} = I_{vv}^{(t-1)} + (\gamma^{\circ(t)}_{vv} - \gamma^{\circ(t-1)}_{vv}) \quad (2)$$

$$I_{vh}^{(t)} = I_{vh}^{(t-1)} + (\gamma^{\circ(t)}_{vh} - \gamma^{\circ(t-1)}_{vh}) \quad (3)$$

$$I_{cr}^{(t)} = I_{cr}^{(t-1)} + (\gamma^{\circ(t)}_{cr} - \gamma^{\circ(t-1)}_{cr}) \quad (4)$$

These indices more directly relate to the SD since it reduces the influence of terrain, incidence angle, and other static local conditions. (Chandra Prabha et al., 2025) also observed that each of these direct backscatter variables and the derived indices exhibits a variable relationship with SD, with each becoming more informative under specific conditions. Hence, all these variables are included in this study, as each can provide independent information and improve the predictive power of the model.

2.2.2 Shuttle Radar Topography Mission Digital Elevation Model

The elevation data used in this study is obtained from the Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM). Equipped with dual radar antennas, the mission acquired data in 2000 using the interferometric SAR, providing near-global coverage between 60°N and 56°S. The DEM data has a spatial resolution of 1 arc-second (~30 m) with vertical and geolocation errors of less than 16 m and 20 m, respectively. This study particularly uses the NASA JPL Version 3.0 ("SRTM Plus"), which fills data voids using ASTER GDEM2, GMTED2010, and the USGS National Elevation Dataset (NED and is accessed through the GEE platform.

2.2.3 MODIS Cloud Gap Filled Snow Cover Data

Snow cover data for this study is obtained from NASA's MODIS Cloud Gap-Filled Snow Cover product (MOD10A1F) through the NASA Earthdata portal. This global Level-3 dataset provides daily cloud-free snow cover at 500 m resolution, derived from the MODIS/Terra Snow Cover Daily product (MOD10A1) by filling cloud gaps using the previous day's observations. From this, a derived variable called consistent snow cover duration (SCD) is calculated as the number of consecutive snow-covered days since the start of each water year (October 1–September 30). The value increases with each continuous day of snow presence and resets to zero whenever snow is absent. As it correlates positively with SD and density, SCD is used as an auxiliary variable to improve model performance.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

2.2.4 Snow Class

Snow properties vary regionally due to differences in climate, precipitation patterns, and topography, which affect properties such as density, grain size, layering, and wetness, and in turn, affect the SAR backscatter. To account for this variability, this study uses the global snow classification by (Sturm & Liston, 2021) available at a resolution of 10-arc seconds (~300 m), as a categorical variable in the model. The classification is based on a binary decision tree that relates basic climate variables to field-observed snow characteristics, providing the climatologically most likely snow type for each location. It has seven classes: tundra, boreal forest, montane forest, prairie, maritime, ice, and ephemeral, each characterized by distinctive textural and stratigraphic properties.

2.2.5 MODIS Percent Tree Cover

Trees on the ground interact with SAR signals and can influence the SD estimation from SAR backscatter. To account for this effect, percent tree cover data from the MODIS Vegetation Continuous Fields (VCF) product (MOD44B, Version 6.1) is used as an auxiliary variable in the model. This annual global dataset, accessed through NASA's Earthdata portal and available since 2000 at 250 m resolution, provides sub-pixel estimates of land cover, including percent tree cover (PTC), which represents the proportion of each pixel covered by tree canopy.

2.2.6 Ground Observation Data

Ground-based SD observations used for model development and evaluation are compiled from multiple networks across the three study regions. For the CRM, data are obtained from the SNOTEL network and NOAA's Global Historical Climate Network (GHCN), comprising 433 stations for the period 2017–2023. For the EA, SD records for 2016–2023 are collected from Météo-France (France), the Zentralanstalt für Meteorologie und Geodynamik (ZAMG, Austria), the WSL Institute for Snow and Avalanche Research SLF (Switzerland; including manual plots and IMIS stations), and GHCN, totalling 483 stations. For the IWH, observations from the Defence Geoinformatics Research Establishment (DGRE) are used, covering 43 stations for the period 2016–2019.

2.2.7 C-Snow Snow Depth

C-Snow is the first SAR-based SD product available globally for mountainous regions. It is derived from S1 backscatter using an empirical change detection approach based on the VH/VV backscatter ratio (Lievens et al., 2019, 2022). The product is provided at ~1 km resolution across the Northern Hemisphere mountains from 2016 to 2020, and at 500 m resolution over selected regions from 2016 to 2021. In this study, 500 m C-Snow data from 2016 to 2021 for the IWH and CRM, and an improved version from 2017 to 2019 for the EA, are used to compare the SD estimates.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

2.2.8 C-RISE Snow Depth

The C-Band Radar-Based Improved Snow Depth Estimation (C-RISE) provides improved Sentinel-1-based SD estimates (Chandra Prabha et al., 2025). It builds on the C-Snow method by incorporating additional backscatter-based variables and accounting for diurnal, seasonal, and regional variability. It also includes auxiliary variables such as SCD and elevation to improve the estimation. SD estimates are obtained for the CRM, EA, and IWH regions for the period 2016–2023 and are used in this study for comparative analysis.

3 Methodology

3.1 Data Preprocessing, Analysis, and Feature Selection

All input datasets used for model development are resampled to a common 500 m spatial resolution and reprojected to a uniform coordinate system for each study area. Ground-based SD observations are matched with corresponding remote sensing data by acquisition date and compiled into a tabular dataset. A few additional variables are derived, other than those described in Sect. 2.2. The Day of Water Year (DOW) is calculated from the date information to represent the number of days since the start of each water year (October 1), accounting for seasonality. To capture topographic and climatic variability, each study area is divided into three elevation-based zones: lower, middle, and upper, which are included as a categorical variable termed Altitude Group. In the CRM and EA, these zones are defined using elevation ranges from previous studies (Hammond et al., 2023; Rubel et al., 2017), while in the IWH, they are based on predefined boundaries (Sharma & Ganju, 2000).

A detailed exploratory data analysis is done to address missing values and outliers, and to examine pairwise correlations for assessing relationships with SD and identifying potential multicollinearity. Since no strong intercorrelations were found, all variables are retained for ML modelling. To improve numerical stability and model convergence, continuous variables are scaled using min–max normalization with realistic minimum and maximum bounds (listed in Table A1), and categorical variables (Snow Class and Altitude Group) are one-hot encoded for compatibility with ML algorithms.

Feature importance is assessed using the Mean Decrease in Impurity (MDI) from a Random Forest regressor across all three study regions, shown in Fig. 3. For categorical variables encoded using one-hot encoding, feature importance is reported for each dummy variable individually, and the overall importance of the original predictor can be interpreted as the sum of its encoded levels. Although the importance of individual predictors differed across regions, each variable contributed meaningful information in at least one region. Therefore, all variables are retained, and the same set of features is used for all study regions to maintain comparability and support the application of this framework to other areas. The final feature set includes six S1 backscatter-based parameters (γ^{o}_{vv}, γ^{o}_{vh}, $\gamma^{o}_{vh/vv}$, I_{vv}, I_{vh}, I_{cr}), SCD, PTC, DOW, snow class, elevation, and altitude group, a total of 12 independent variables and ground observation as the dependent variable for model development.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

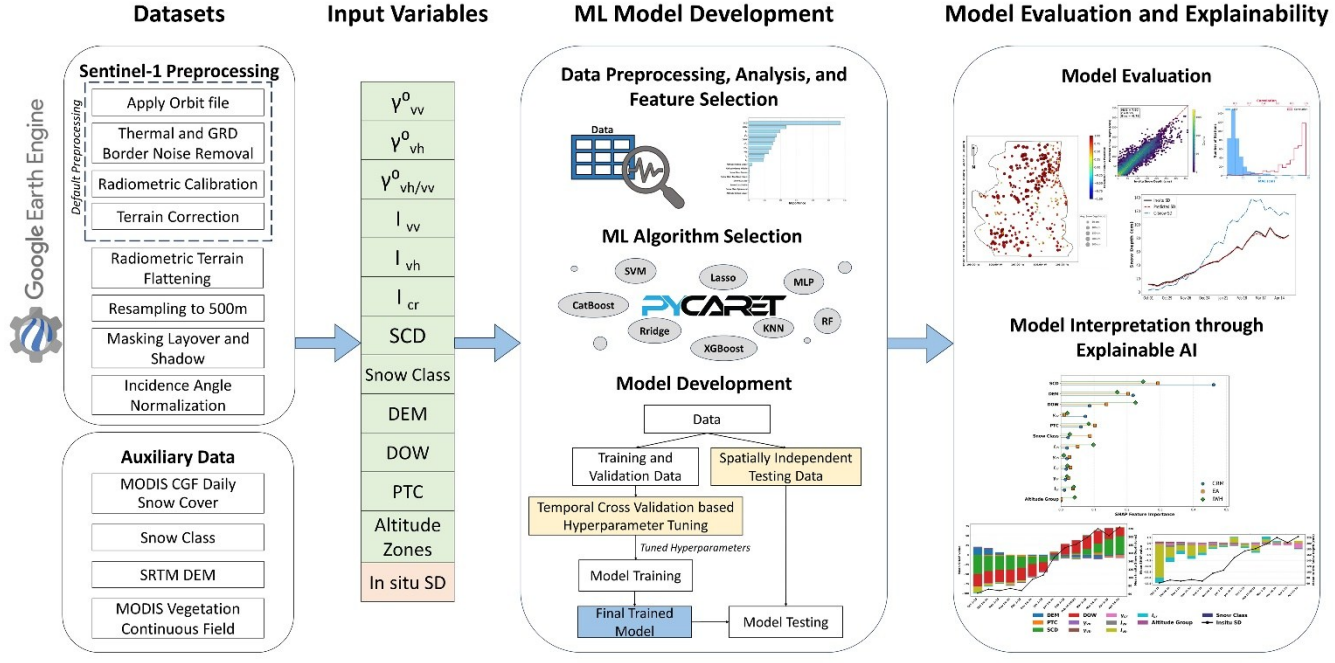


Figure 2. Overall methodology flowchart

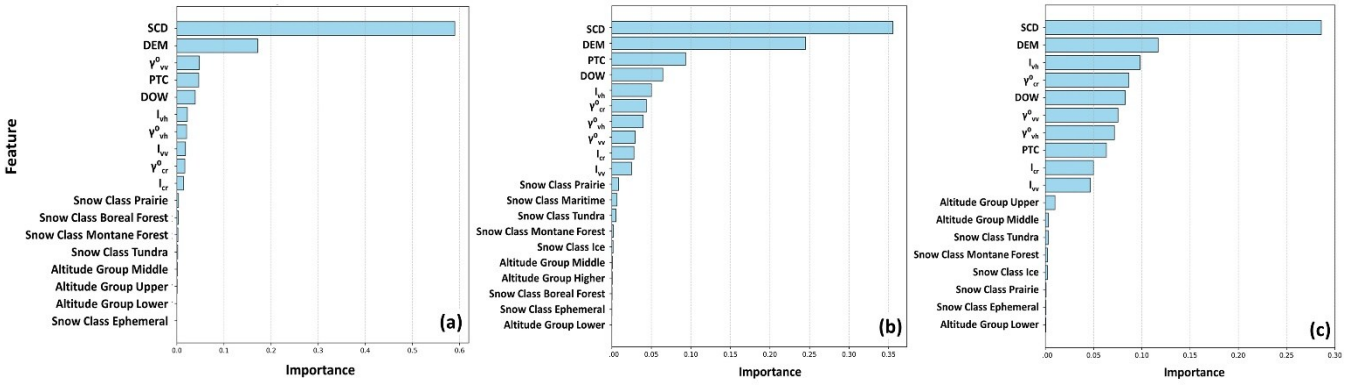


Figure 3. Feature importance of the input variables across (a) CRM, (b) EA, and (c) IWH

3.2 ML Algorithm Selection

A wide range of machine learning algorithms are initially compared to identify the most suitable model for this study. Using the PyCaret framework, around 25 commonly used algorithms (e.g., Random Forest, Support Vector Machines, Gradient Boosting methods, and others) are compared with default hyperparameters, following a similar approach to Dunmire et al., 2024. This comparison is carried out separately for each of the three study regions. Across all regions, CatBoost, XGBoost,

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

and ExtraTrees showed the best performance with comparable evaluation metrics and are selected for further development with hyperparameter tuning. The models are optimized using 10-fold cross-validation, trained, and then evaluated on separate testing data. XGBoost and CatBoost showed comparable results, but XGBoost performed slightly better in two of the three study regions, and is therefore selected as the final algorithm for this study.

3.3 Model Development

Following the selection of the ML algorithm, XGBoost-based models are developed for each of the study regions, as outlined in this subsection.

3.3.1 Training, Validation, and Testing Data Split Strategy

To assess model generalization across space and time, the datasets are partitioned in two stages. First, 20% of the stations in each study area are set aside as a spatially independent test set. These stations are selected through stratified random sampling based on snow class and altitude group, ensuring a similar distribution between the training and testing data for unbiased evaluation on unseen stations. In the remaining 80% (training and validation data), a temporally independent cross-validation strategy is adopted, where data are grouped by water year, with each year (e.g., 2016–2023, has seven water years) treated as a fold. In each iteration, one fold is held out for validation while the others are used for training, iterating through all combinations. Final validation performance is reported as the average across folds. This approach ensures evaluation on time periods unseen during training, improving its ability to generalize to new time periods. The number of samples used for each study region is provided in Table 1.

Table 1 Number of Samples used for training and testing in the three study areas

Study Area	Total No. of Samples	No. of Test Samples	No. of Training Samples
Colorado Rocky Mountains	46613	7825	38788
European Alps	186270	35874	150396
Indian Western Himalayas	3525	790	2735

3.3.2 Model Training and Evaluation

Model development follows a standard machine learning workflow, including data partitioning, hyperparameter tuning, training, and independent testing. Hyperparameters are optimized using Optuna's Bayesian optimization framework with the temporally independent cross-validation strategy described in Sect. 3.3.1. This approach models relationships between hyperparameters and performance metrics using probabilistic surrogate functions, efficiently guiding the search toward optimal

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

configurations. In each iteration, Optuna proposes a set of candidate hyperparameters within predefined ranges, which are then evaluated through temporal fold-based cross-validation and the average validation metrics across folds are used to select the best hyperparameter configuration. Once the optimal hyperparameters are identified, the models are retrained on the full training–validation dataset for each study region and then evaluated on the spatially independent test sets. The final trained models are saved and applied to the complete datasets of their respective study regions to generate SD predictions.

3.4 Model Interpretation through Explainable AI

In addition to model development, this study also focuses on understanding how different predictors influenced SD estimation across the three regions. The intention is to assess not only which variables were most influential, but also the direction and magnitude of their effects under different snowpack and environmental conditions. Such analysis is particularly important given that the physical mechanisms underlying S1 backscatter–based SD estimation are not yet fully understood. For this purpose, SHAP (SHapley Additive exPlanations) values are used. Based on cooperative game theory, SHAP provides a consistent framework to attribute model predictions to individual input features, indicating their contribution to the predicted SD. SHAP values are analysed across different stages of the winter season, as well as within snow classes and altitude zones, for all three study regions. This provides region- and condition-specific insights into the contribution of various predictors, offering a more detailed and physically interpretable understanding of the model's behaviour, while also revealing new perspectives on the role of S1 backscatter in SD estimation across the regions.

4 Results and Discussion

4.1 Model Evaluation

This section presents the evaluation of the developed models using in situ SD observations. Sect. 4.1.1 reports the model development metrics, including cross-validation and testing results for the final models. Sect. 4.1.2 evaluates the spatial performance of the developed models.

4.1.1 Model Validation and Testing Results

The cross-validation metrics for the developed models, obtained using temporally independent folds and finalized hyperparameters, are presented in Table 2 for the three study regions. The three models were then evaluated using spatially independent test data (Sect.3.3.1), with the corresponding performance metrics also summarized in Table 2 and the scatter plots presented in Fig. 4.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

Table 2 Model evaluation metrics for the three study areas

Study Area	Temporally Independent Cross-Validation		Spatially Independent Testing	
	MAE (cm)	r	MAE (cm)	r
Colorado Rocky Mountains	14.7	0.9	15.7	0.9
European Alps	30.5	0.8	34.2	0.8
Indian Western Himalayas	116.6	0.6	64.3	0.6

The temporal cross-validation results show that the models perform well in the CRM (r = 0.9, MAE = 14.7 cm) and EA (r = 0.8, MAE = 30.5 cm), indicating good predictive capability across different years in these regions. In the IWH, performance is lower (MAE = 116.6 cm, r = 0.6), likely due to the high variability in SD and the limited availability of stations and temporal data. Spatially independent testing also shows a similar pattern. In the CRM, the model performs well (r = 0.9, MAE = 15.7 cm), with predictions closely following the 1:1 line (Fig. 4a), indicating low bias and strong agreement between observed and predicted SD. In the EA, performance is slightly lower (r = 0.8, MAE = 34.2 cm), with predictions generally aligning along the 1:1 line but showing a systematic underestimation above ~450 cm, likely due to limited observations in regions with deeper snowpack. In the IWH, performance is comparatively lower (r = 0.6, MAE = 64.3 cm), with the scatterplot (Fig. 4c) showing a broader dispersion of points around the 1:1 line.

Comparing temporal and spatial evaluations, the CRM and EA exhibit similar metrics, indicating good model generalization across time and space. However, IWH shows a larger difference in MAE between temporal and spatial evaluations, indicating comparatively weaker generalization. Overall, the models perform well in CRM and EA, and reasonably in IWH given the challenging conditions and limited data availability.

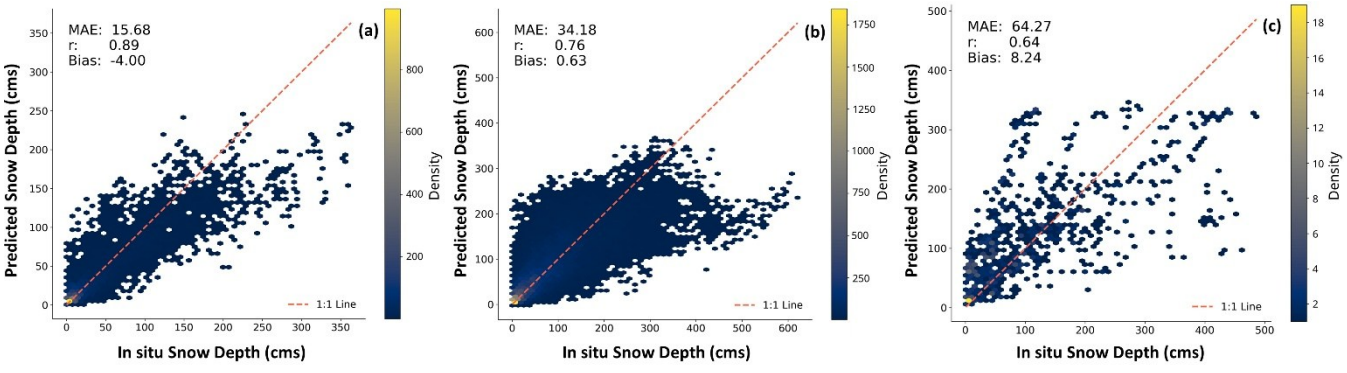


Figure 4. Scatter plots of observed versus predicted SD for the test data in the three study regions: (a) CRM, (b) EA, and (c) IWH.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

4.1.2 Overall Model Performance Across Stations

Model performance across individual stations was analysed to assess spatial consistency and identify local factors affecting the model accuracy. Figure 5 shows spatial maps of correlation, MAE, and bias for each region. Across all three regions, the models generally perform well, with higher correlations and lower errors in areas with deeper snowpacks, consistent with previous findings that C-band backscatter is more sensitive to deep snow. However, a few stations across the study regions show relatively high errors, mainly due to local conditions. Most of these are located in forested, glaciated, or areas with very deep snow. In forested regions, higher errors likely result from radar–canopy interactions, often leading to SD overestimation. Whereas stations in tundra, glaciated, or very deep snowpack regions tend to show underestimation, likely due to the limited C-band penetration into deep snow and the scarcity of ground observations in such extreme environments, which constrains the machine learning model's ability to accurately represent these environments.

In the CRM region, Fig. 5a and 5b show that the model performs consistently well across most stations, with only a few isolated exceptions. Stations in low-elevation areas with shallow snowpacks and stations corresponding to the prairie snow class show lower temporal correlations, while some high-elevation stations with deeper snow exhibit higher MAE values. As shown in Fig. 6a, station-wise temporal correlations mostly range from 0.9 to 1.0 (mean = 0.85), and MAE values from 2.5 cm to 5 cm (mean = 10 cm). Approximately 70% of stations have correlations above 0.8 and MAE below 6.5 cm, indicating good model performance across the region.

In the EA region, Fig. 5c and 5d show similarly good model performance, particularly in areas with deeper snowpacks such as Switzerland, where temporal correlations are consistently high. However, comparatively lower correlations are observed in regions with denser forest cover, such as the eastern part of the Alps and in areas characterised by ephemeral snow class. Despite these localised variations, the model performs well across most of the region, with only a few stations showing more deviations. As shown in Fig. 6b, the most frequent station-wise temporal correlation values range between 0.7 and 1.0, with a mean of 0.8. As shown in Fig. 6b, temporal correlations most frequently range between 0.7 and 1.0 (mean = 0.8), and MAE between 10 cm and 20 cm (mean = 17 cm). Around 70% of stations have correlations above 0.7 and MAE below 22.1 cm, demonstrating good performance across the EA.

In the IWH region, although the number of available stations is much less than in the other two regions, these stations represent diverse zones of the IWH. Existing studies for this region also rely on these available station datasets. The model performs well at most stations and moderately well at several others, with only two stations showing poor performance, one located in a highly forested area and another at low altitude. Spatially, the model performs reasonably well across the region (Fig. 5e, 5f). As shown in Fig. 6c, correlations typically range from 0.65 to 0.9 (mean = 0.7), and MAE from 30 cm to 50 cm (mean = 75 cm). Approximately 70% of stations have correlations above 0.6 and MAE below 68.9 cm. Although absolute errors are higher than in CRM and EA, the model still captures SD variability effectively, given the region's limited data availability and large SD variability.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

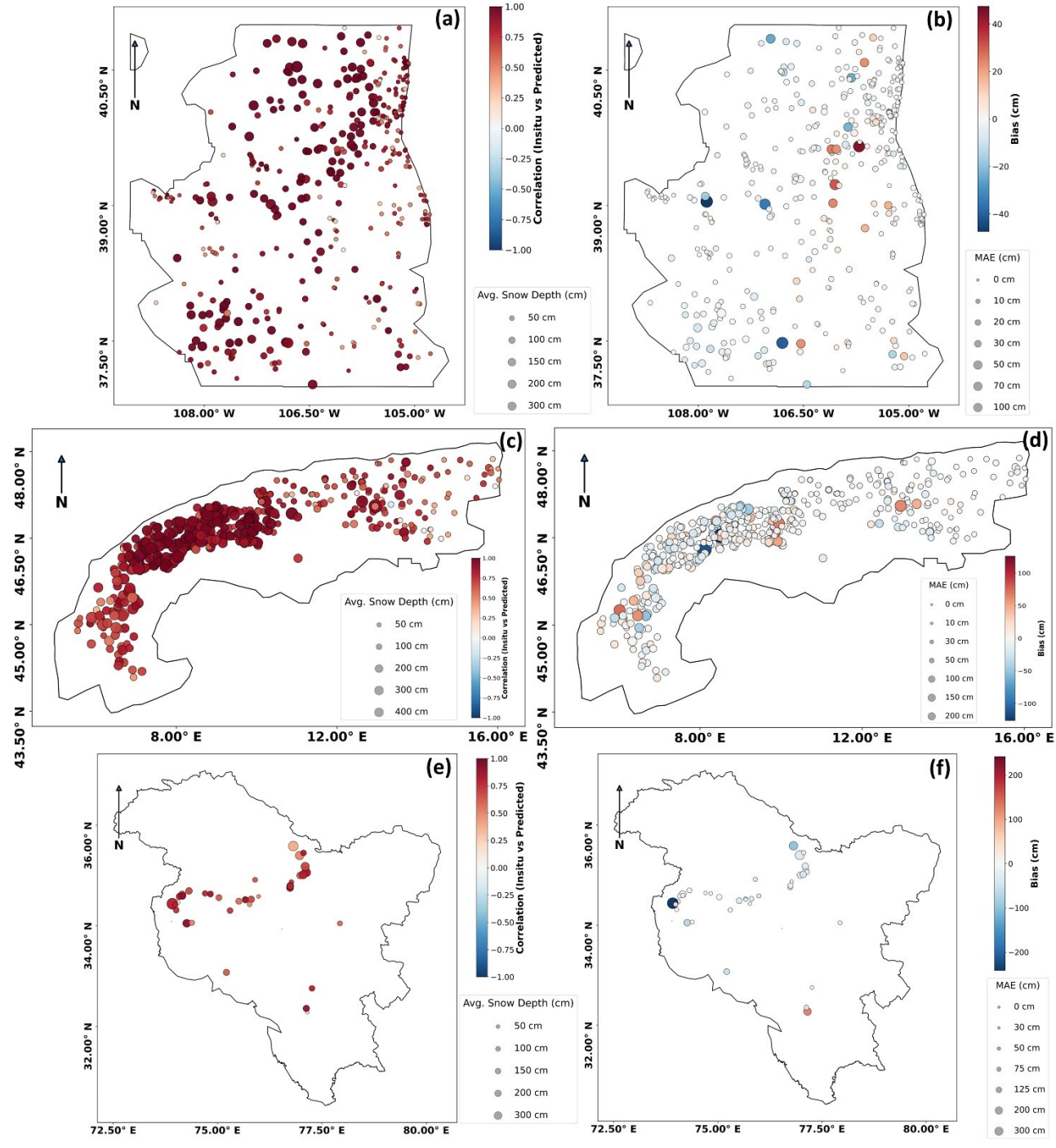


Figure 5. Spatial distribution of stations with model performance: (a, c, e) correlation and average SD, and (b, d, f) MAE and bias for CRM, EA, and IWH, respectively.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

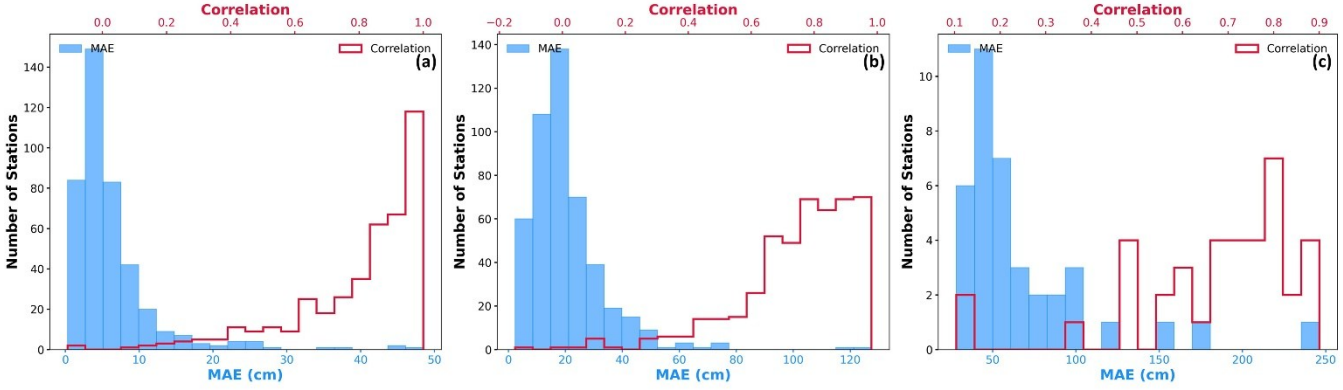


Figure 6. Histograms of station-wise correlation and MAE for (a) CRM, (b) EA, (c) IWH

4.2 Comparison with C-Snow and C-RISE Snow Depth Estimations

The SD estimates from the developed ML model are compared with the existing S1-based C-Snow and C-RISE SD estimates by evaluating them using in situ SD measurements across the three study regions. Figure 7 presents the scatterplots comparing each estimate with in situ SD, along with the corresponding average seasonal trends.

In CRM (Fig. 7a-d), the ML model predicted SD with an MAE(r) of 7.9 cm (0.96), substantially outperforming C-Snow and C-RISE, which have an MAE(r) of 38 cm (0.65) and 20.25 cm (0.82), respectively. The ML model has reduced the MAE by approximately 79.2 % relative to C-Snow and 61.0 % relative to C-RISE. The average seasonal trend (Fig. 7d) shows that the ML model closely follows in situ SD throughout the season, while C-RISE captures the general pattern but deviates during early accumulation and late winter.

In EA also (Fig. 7e-h), the ML model performs better than both C-Snow and C-RISE, with an MAE (r) of 22.27 cm (0.91), compared to 37.25 cm (0.79) for C-Snow and 40.13 cm (0.78) for C-RISE, a MAE reduction of about 40.2 % and 44.5 %, respectively. A recent study by (Dunmire et al., 2024), using a similar Sentinel-1 backscatter-based ML approach, reported an MAE (r) of 33 cm (0.78) compared to 37 cm (0.75) for C-Snow, a 10.8% improvement. In contrast, the present study achieves a 40.2% reduction, likely due to using change detection-based indices instead of direct backscatter values. The seasonal trend (Fig. 7h) further confirms that the ML model best follows the in-situ SD variation, with C-Snow performing next best.

In IWH, the ML model predicted SD with an MAE (r) of 68.38 cm (0.72), a reduction in MAE by 29.7% and 20.4% compared to C-Snow and C-RISE, which have MAE (r) values of 97.23 cm (0.53) and 85.9 cm (0.60), respectively. While the overall performance is good, the scatterplot (Fig. 7i) shows flattening beyond ~350 cm, likely due to limited in situ observations in very deep snow areas, despite weighted sampling during training. The average seasonal trend (Fig. 7l) shows that the ML estimates follow the in-situ SD observations more closely than C-Snow or C-RISE, with C-RISE following reasonably well but less accurately than ML estimates.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

The performance of the ML model predicted SD, C-Snow, and C-RISE SD was further evaluated under different scenarios, including different months of the season, snow classes, percentages of tree cover, and SD ranges, using bias distributions, shown in Fig. 8.

Across different months of the season, the ML-predicted SD in CRM exhibits minimal bias with a narrow range throughout the winter (Fig. 8a). In EA and IWH (Fig. 8b and 8c), although the ML predictions show increasing bias as winter progresses, both the mean bias and the variability remain substantially lower than those of C-Snow and C-RISE.

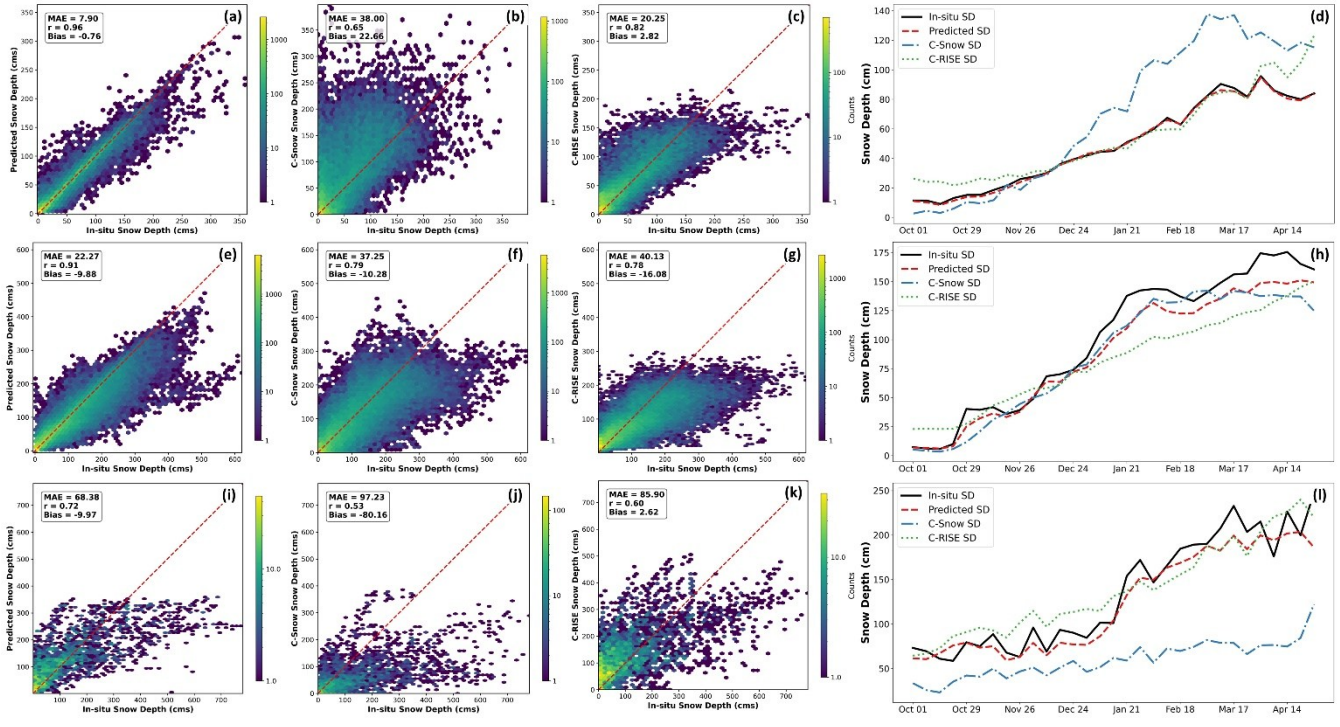


Figure 7. Comparison of SD estimates from the ML model with C-Snow and C-RISE. Scatterplots of ML-predicted SD versus in situ SD (a, e, i), C-Snow versus in situ SD (b, f, j), and C-RISE versus in situ SD (c, g, k) are shown for CRM, EA, and IWH, respectively. The average seasonal SD trends for the ML model, C-Snow, C-RISE, and in situ measurements are shown in (d, h, l) for each region.

When analysed across snow classes, the ML model generally outperforms the other two. In CRM (Fig. 8d), a slight underestimation is observed for the tundra class. In EA (Fig. 8e), underestimation is observed in both tundra and ice classes. In IWH (Fig. 8f), the mean bias does not indicate systematic over- or underestimation, however, the ice class shows a wide bias range with numerous negative outliers. These observations in general suggest potential challenges in glaciated areas that require further investigation.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

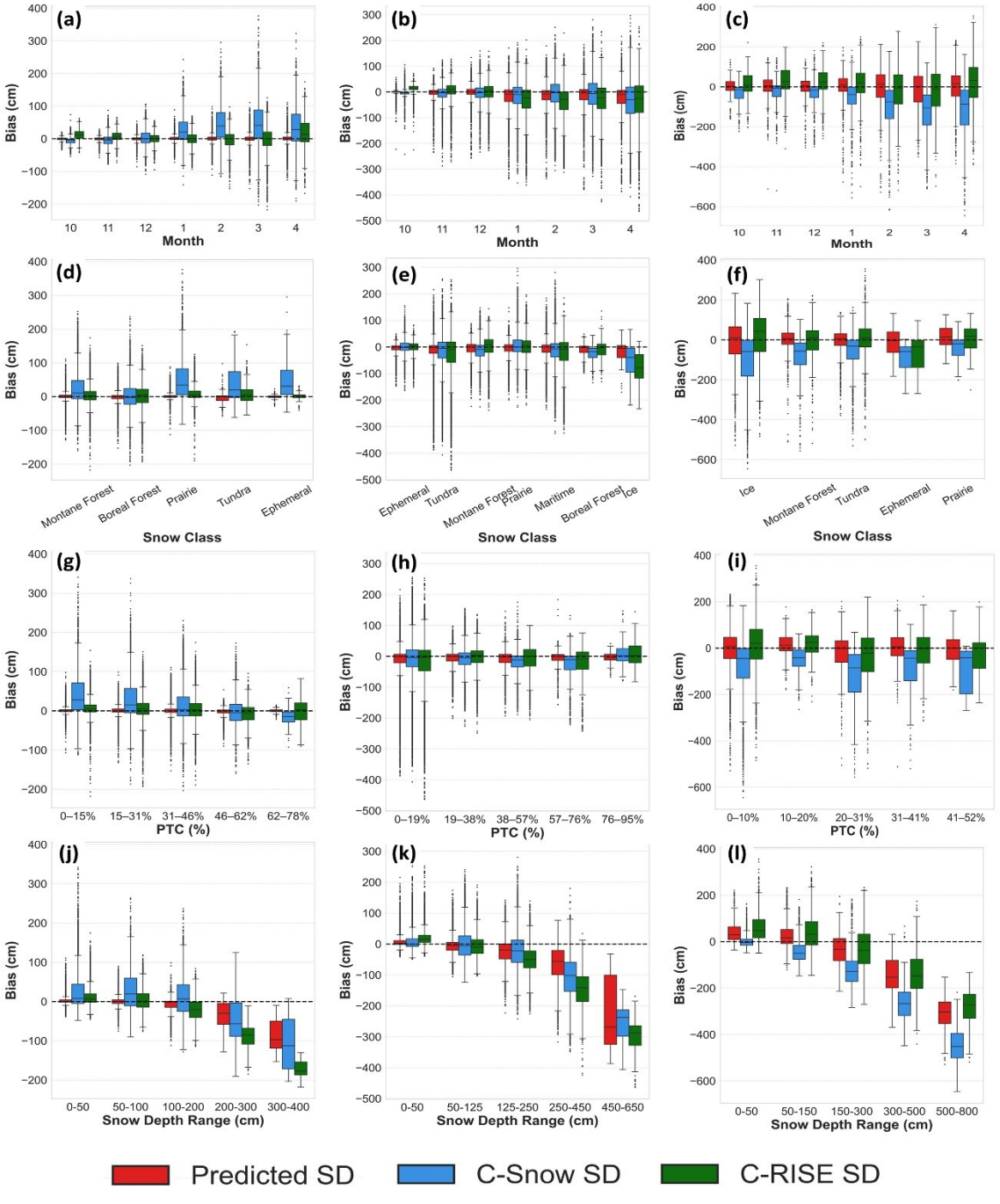


Figure 8. Bias comparison of ML model predicted SD, C-Snow, and C-RISE across different conditions. Bias distributions are shown for different months of the season (a–c), snow classes (d–f), percentages of tree cover (PTC) (g–i), and SD ranges (j–l) in CRM, EA, and IWH, respectively.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

When assessed across different PTC, the ML model again performs best. In CRM (Fig. 8g), a slight increase in bias range is seen with higher PTC. In EA and IWH (Fig. 8h and 8i), no clear trend is observed.

Across SD ranges, all three SD estimates tend to underestimate deeper snowpacks, likely due to limited C-band signal penetration or sparse ground observations in deeper snowpack regions. The ML model performs best across all SD ranges in CRM (Fig. 8j) and most ranges in EA (Fig. 8k), except for the 450–650 cm range, where C-Snow performs better. In IWH (Fig. 8l), ML estimates outperform others, except in the deepest 500–800 cm range, where C-RISE performs better. This indicates that the ML model requires sufficient training data representing very deep snowpack regions to improve its performance.

Overall, the ML model provides substantially improved SD estimates compared to both C-Snow and C-RISE in all three regions.

4.3 Physical Insights through Explainable AI

This section focuses on model interpretation and the physical insights obtained from the SHAP value, which quantifies the contribution of each input variable to individual SD predictions. Sect. 4.3.1 discusses the overall SHAP values, and Sect. 4.3.2 examines scenario-specific SHAP behaviour.

4.3.1 Overall SHAP values

SHAP values were averaged and normalized to derive feature importance for each study area (Fig. 9). SCD emerged as the most influential variable across all regions, with its influence particularly dominant in CRM, where it was much higher than that of the other variables. DEM was the next most important variable across the regions. In CRM, unlike the other regions, γ^{o}_{vv} was the most influential backscatter-based variable, likely due to higher vegetation cover. As noted by (Lievens et al., 2022), in vegetated regions, vv-polarisation tends to be more strongly related to SD, though the underlying scattering mechanisms require further investigation.

In EA, the snow class variable was more important than in CRM and IWH, likely because EA contains almost an equal proportion of montane forest, tundra, and maritime snow classes, unlike the largely homogeneous snow classes in CRM (montane forest) and IWH (tundra). In IWH, SCD and DOW were almost equally important, suggesting that seasonal variations in snowpack properties strongly influence SD estimation. Interestingly, while I_{cr} has often been used for Sentinel-1 backscatter-based SD estimation, I_{vh} showed higher importance, particularly in IWH, where deeper snowpacks enhance volume scattering, making I_{vh} more strongly related to SD. Additionally, the altitude group variable, which divides the region into lower, middle, and upper zones, was more influential in IWH, highlighting strong altitudinal and regional variability in this region.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

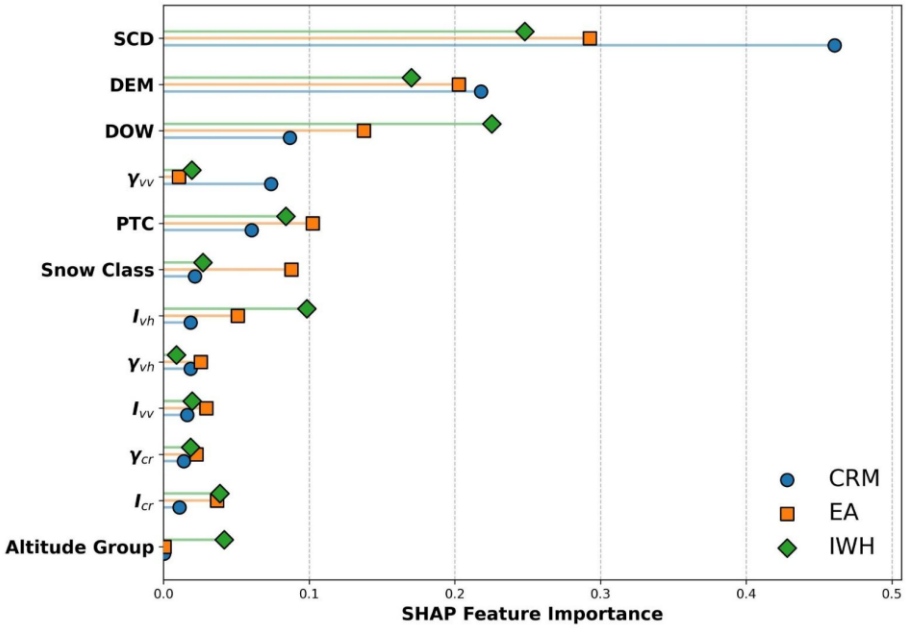


Figure 9. Feature importance based on SHAP values in CRM, EA, and IWH

4.3.2 Analysis of SHAP values in different scenarios

SHAP values quantify the contribution of each variable to individual model predictions, showing how much each variable shifts the model's output from its base value, either positively or negatively (temporal variation of the actual SHAP values is provided in Fig. B1 for reference). To facilitate interpretation, relative SHAP values (the absolute SHAP value of each variable divided by the sum for all variables) were analysed across four scenarios: across the season at 15-day intervals (Fig. 10a, c, e), across stations grouped by average SD (Fig. 10b, d, f), across different ranges of PTC (Fig. 11a, c, e), and across different snow classes (Fig. 11b, d, f), for the three study regions CRM, EA, and IWH.

Across the season (Fig. 10a, c, e), the relative contribution of backscatter-based variables increased through early winter, peaked mid-season, and declined after February—likely due to wet snow affecting the backscatter signal. In CRM and EA, the contribution of the DEM variable closely followed the average SD pattern, indicating an elevation-dependent effect. In EA, the contribution of the snow class variable peaks in mid-December and remains stable from then, indicating that as winter progresses, the increasingly distinct snowpack properties across classes influence the backscatter signal. The PTC variable also exhibited a seasonal variation, although the physical reasons underlying this behaviour require further investigation.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

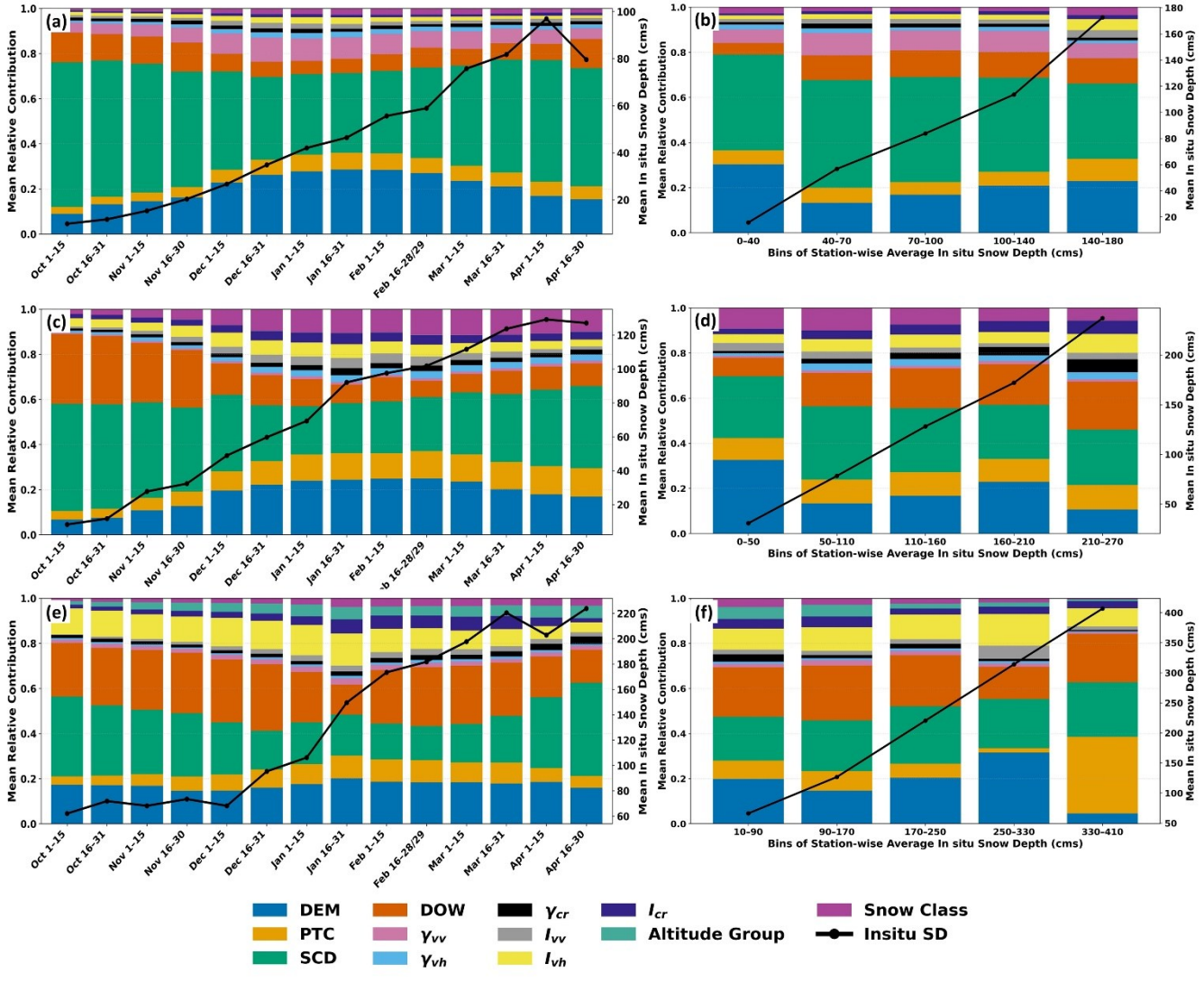
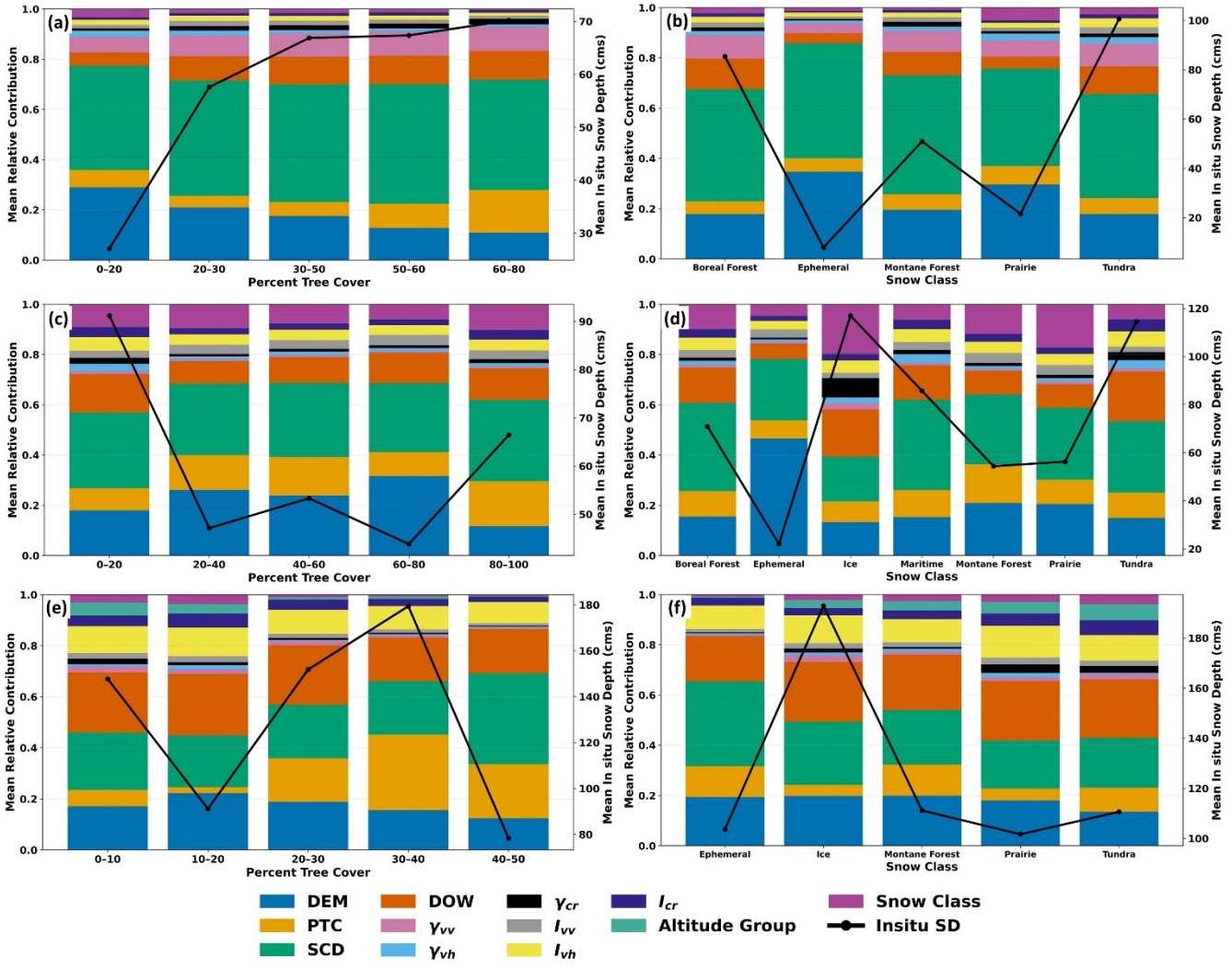


Figure 10. Average relative contribution of each variable based on SHAP values of model predictions: (a, c, e) across the season and (b, d, f) across stations with different average SD ranges in CRM, EA, and IWH, respectively.

When analysed with respect to station-wise average SD range representing regions with shallow to deep snowpack (Fig. 10b, d, f), backscatter-based variable contributions increased with SD range across all three regions, confirming the S1 backscatter method's suitability for deeper snowpacks. In CRM (Fig. 10b), I_{vh} shows a relatively higher contribution at the maximum SD range. In EA (Fig. 10d), the contributions of γ^{o}_{cr}, I_{cr}, and I_{vh} increased for stations with deeper snow. In IWH (Fig. 10f), I_{vh} was more influential than in the other regions, and increased with SD, suggesting stronger cross-polarisation relevance in deep snow conditions.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026



455 **Figure 11. Average relative contribution of each variable based on SHAP values of model predictions: (a, c, e) across different PTC ranges and (b, d, f) across different snow classes in CRM, EA, and IWH, respectively.**

Across PTC ranges (Fig. 11a, c, e), backscatter-based variables contributed more at lower PTC values, especially cross-polarisation variables. In CRM, however (Fig. 11a), γ^0_{vv} contribution increased with PTC, suggesting its sensitivity to SD in forested regions. In contrast, in IWH (Fig. 11e), the contribution of backscatter-based variables decreased with increasing PTC, 460 suggesting that the method is generally more effective in areas with sparse vegetation. Still, γ^0_{vv} may relate to SD in forested regions, but further radar-based experiments are needed to better understand the underlying scattering mechanisms.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

The relative contributions were then analysed across different snow classes (Fig. 11b, d, f). In the ephemeral snow class, the contributions of backscatter-based variables were generally lower, whereas in the tundra snow class, the contributions were comparatively higher, especially for I_{vh} and I_{cr}.

Overall, the SHAP analysis highlights that SCD and DEM are more influential in SD estimation, and the contribution of backscatter-based variables is highly region- and condition-dependent due to varying radar–snow interactions. Backscatter variables contribute most in deeper or tundra snowpacks and are sensitive to vegetation and evolving snowpack conditions. These findings provide insights into the mechanisms of S1-based SD estimation and highlight the need to account for these factors to improve the SD estimates. However, detailed radar experiments are required to corroborate these observations.

5 Conclusion

This study presents an ML framework for improving SD estimation using S1 backscatter data and attempts to understand the physical mechanisms behind this C-band backscatter-based SD estimation. It advances existing approaches by incorporating change detection–based indices along with direct backscatter variables at different polarisations and other auxiliary parameters. The inclusion of these indices enables the model to better capture temporal variations in the snowpack across diverse conditions, resulting in improved SD estimates. All variables are derived and processed within the cloud-based GEE platform, reducing the computational requirements and allowing efficient, scalable implementation across different regions and time periods.

The framework was implemented across three different mountainous regions, CRM, EA and IWH, each characterised by distinct snow and terrain characteristics. Across all regions, the proposed ML model performed better than the existing S1-based C-Snow SD product and the C-RISE SD estimates. In the EA region, the model achieved a 40.2% reduction in MAE compared to C-Snow, whereas the most recent ML study across this region reported only a 10.8% reduction. The key difference between the two models is the incorporation of change detection–based indices, which more effectively capture SD variations than direct backscatter variables. While other factors like weighted sampling and the choice of auxiliary variables may have had some effect, the use of change detection–based indices is the primary reason for improved performance.

Explainable AI techniques, specifically SHAP values, were then used to interpret model behaviour and quantify the contribution of each variable across different scenarios, providing new insights into the C-band backscatter-based SD estimation from the three study regions. Results show that change detection–based indices contribute more strongly to SD estimation than direct backscatter variables, with their influence varying by region. The contribution of backscatter variables follows regional SD characteristics, being relatively low in CRM with shallow snowpacks, to higher in EA with deeper snowpacks, and highest in IWH with very deep snowpacks. Spatial and environmental factors further influence these relationships, with VV-polarised backscatter contributing more in areas of dense vegetation and VH-polarised backscatter

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

contributing more in regions with very deep snowpacks. Temporally, backscatter contributions increase through early and mid-winter and decrease toward late winter as snow becomes wetter or begins to melt. Overall, these variations highlight the complex interactions between radar signals and snowpack properties, emphasising the need for a deeper physical understanding for more accurate SD estimation.

Although the developed ML models perform better than the existing approaches, there remains considerable scope for further improvement. The models in this study were trained using point-based ground observations, which may not adequately represent the spatial variability of SD in mountainous regions. Future work should integrate high-resolution spatial datasets, such as airborne LiDAR or photogrammetric measurements, to better represent the spatial variability. However, such datasets are scarce in CRM and EA and are unavailable for the IWH. Further, in IWH, the model was developed using the very limited ground observation station data available, due to which the ML model was not well generalised as in CRM and EA. For such data-scarce regions, alternative strategies such as data augmentation, transfer learning, or incorporating physical constraints should be explored to overcome these limitations.

This study also emphasises the need to better understand C-band radar interactions with snow under varying environmental conditions. More detailed tower-based radar and field experiments across different snow and terrain types are essential to characterise these processes more accurately and inform the radiative transfer models. Integrating such models with ML can enable physics-guided machine learning (PGML) approaches, enhancing accuracy, generalisation, and physical consistency while reducing dependence on extensive ground data. With the launch of S1-C and S1-D, the continued availability of preprocessed S1 data through the GEE platform presents an excellent opportunity to further refine and expand the C-band backscatter-based SD estimation framework, supporting large-scale and long-term snow monitoring.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

.

Appendix A

Table A1 Minimum and maximum bounds used for min-max scaling of numerical variables

Sl. No	Numerical Variable	Minimum Bound	Maximum Bound
1	Elevation (m)	100	4000
2	PTC (%)	0	100
3	SCD	0	215
4	DOW	0	215
5	γ^{o}_{vv} (dB)	-30	5
6	γ^{o}_{vh} (dB)	-30	5
7	$\gamma^{o}_{vh/vv}$ (dB)	-30	5
8	I_{vv}	0	15
9	I_{vh}	0	15
10	I_{cr}	0	15

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

Appendix B

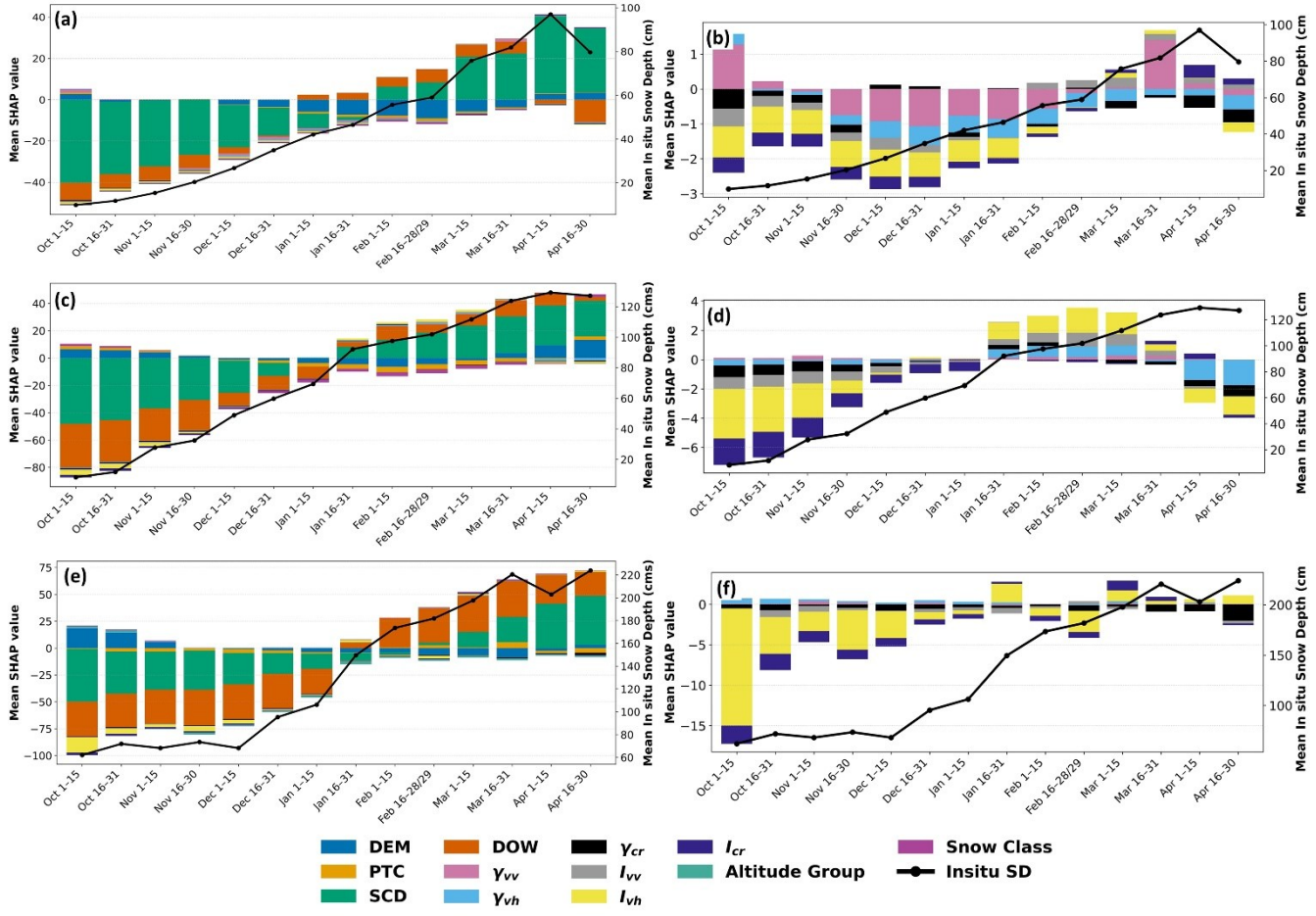


Figure B1. Average Temporal SHAP values (stacked bar plot) along with average in situ SD across the season for (a) all the variables, (b) backscatter-based variables

Code availability

The scripts and the final machine learning model developed for snow depth estimation will be made publicly available upon completion of the peer-review process.

Data availability

The Sentinel-1 SAR data and SRTM Digital Elevation Model used in this study are freely available through Google Earth Engine (GEE) under the identifiers ee.ImageCollection("COPERNICUS/S1_GRD") and ee.Image("USGS/SRTMGL1_003"), respectively.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

MODIS Snow Cover data and the MODIS Vegetation Continuous Fields (VCF) percent tree cover product are accessible through NASA's Earthdata portal. The Snow Class dataset is available from the National Snow and Ice Data Center (NSIDC) at https://nsidc.org/data/nsidc-0768/versions/1.

In situ snow depth data for the Colorado region are freely available from the SNOTEL network (https://www.nrcs.usda.gov/) and NOAA's Global Historical Climatology Network (GHCN) Daily dataset (https://www.ncei.noaa.gov/products/land-based-station/global-historical-climatology-network-daily).

For the European Alps, in situ snow depth data can be obtained from Météo-France (https://donneespubliques.meteofrance.fr/), the Zentralanstalt für Meteorologie und Geodynamik (ZAMG, Austria) (https://data.hub.geosphere.at/), and the WSL Institute for Snow and Avalanche Research SLF (https://www.slf.ch/en/).

For the Indian Himalayas, in situ snow depth data can be requested from the Defence Geoinformatics Research Establishment (DGRE).

The snow depth estimates generated in this study are available from the corresponding author upon request.

Author contributions

RCP: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization Writing – original draft; RR: Conceptualization, Writing – review and editing, Supervision.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

The authors would like to thank the Defence Geoinformatics Research Establishment (DGRE), DRDO, for providing access to ground observation station data for Western Himalayas through the Contract for Acquisition of Research Services (CARS) project carried out during 2021-2022.

Financial support

This work was supported by the Prime Minister Research Fellowship (PMRF) awarded by the Ministry of Education (MoE), Government of India.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

References

Awasthi, S., Kumar, S., Thakur, P. K., Jain, K., Kumar, A., & Snehmani. (2021). Snow depth retrieval in North-Western Himalayan region using pursuit-monostatic TanDEM-X datasets applying polarimetric synthetic aperture radar interferometry-based inversion Modelling. *Https://Doi.Org/10.1080/01431161.2020.1862439*, *42*(8), 2872–2897. https://doi.org/10.1080/01431161.2020.1862439

Awasthi, S., & Varade, D. (2021). Recent advances in the remote sensing of alpine snow: a review. *GIScience & Remote Sensing*, *58*(6), 852–888. https://doi.org/10.1080/15481603.2021.1946938

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature 2005 438:7066*, *438*(7066), 303–309. https://doi.org/10.1038/nature04141

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future köppen-geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*(1), 1–12. https://doi.org/10.1038/SDATA.2018.214;SUBJMETA

Brangers, I., Marshall, H. P., De Lannoy, G., Dunmire, D., Mätzler, C., & Lievens, H. (2024). Tower-based C-band radar measurements of an alpine snowpack. *Cryosphere*, *18*(7), 3177–3193. https://doi.org/10.5194/TC-18-3177-2024

Broxton, P., Ehsani, M. R., & Behrangi, A. (2024). 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*(3). https://doi.org/10.1029/2023EA002964

Bühler, Y., Marty, M., Egli, L., Veitinger, J., Jonas, T., Thee, P., & Ginzler, C. (2015). Snow depth mapping in high-alpine catchments using digital photogrammetry. *Cryosphere*, *9*(1), 229–243. https://doi.org/10.5194/TC-9-229-2015

Bulovic, N., Johnson, F., Lievens, H., Shaw, T. E., McPhee, J., Gascoin, S., Demuzere, M., & McIntyre, N. (2025). Evaluating the Performance of Sentinel-1 SAR Derived Snow Depth Retrievals Over the Extratropical Andes Cordillera. *Water Resources Research*, *61*(2), e2024WR037766. https://doi.org/10.1029/2024WR037766

Chandra Prabha, R., Tanniru, S., & Ramsankaran, R. (2025). C-Band Radar-Based Improved Snow Depth Estimation (C-RISE) in the Indian Western Himalayas and Colorado Rocky Mountains. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *18*, 11787–11802. https://doi.org/10.1109/JSTARS.2025.3563462

Daudt, R. C., Wulf, H., Hafner, E. D., Bühler, Y., Schindler, K., & Wegner, J. D. (2023). 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/10.1016/J.ISPRSJPRS.2023.01.017

Deems, J. S., Painter, T. H., & Finnegan, D. C. (2013). Lidar measurement of snow depth: a review. *Journal of Glaciology*, *59*(215), 467–479. https://doi.org/10.3189/2013JOG12J154

Dunmire, D., Lievens, H., Boeykens, L., & De Lannoy, G. J. M. (2024). A machine learning approach for estimating snow depth across the European Alps from Sentinel-1 imagery. *Remote Sensing of Environment*, *314*, 114369. https://doi.org/10.1016/J.RSE.2024.114369

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

Gascoin, S., Luojus, K., Nagler, T., Lievens, H., Masiokas, M., Jonas, T., Zheng, Z., & De Rosnay, P. (2024). Remote sensing of mountain snow from space: status and recommendations. *Frontiers in Earth Science*, *12*, 1381323. https://doi.org/10.3389/FEART.2024.1381323/FULL

Hammond, J. C., Sexstone, G. A., Putman, A. L., Barnhart, T. B., Rey, D. M., Driscoll, J. M., Liston, G. E., Rasmussen, K. L., McGrath, D., Fassnacht, S. R., & Kampf, S. K. (2023). High Resolution SnowModel Simulations Reveal Future Elevation-Dependent Snow Loss and Earlier, Flashier Surface Water Input for the Upper Colorado River Basin. *Earth's Future*, *11*(2), e2022EF003092. https://doi.org/10.1029/2022EF003092

Hoppinen, Z., Palomaki, R. T., Brencher, G., Dunmire, D., Gagliano, E., Marziliano, A., Tarricone, J., & Marshall, H. P. (2024). Evaluating snow depth retrievals from Sentinel-1 volume scattering over NASA SnowEx sites. *Cryosphere*, *18*(11), 5407–5430. https://doi.org/10.5194/TC-18-5407-2024

Immerzeel, W. W., Droogers, P., de Jong, S. M., & Bierkens, M. F. P. (2009). Large-scale monitoring of snow cover and runoff simulation in Himalayan river basins using remote sensing. *Remote Sensing of Environment*, *113*(1), 40–49. https://doi.org/10.1016/J.RSE.2008.08.010

Jans, J. F., Beernaert, E., De Breuck, M., Brangers, I., Dunmire, D., De Lannoy, G., & Lievens, H. (2024). Sensitivity of Sentinel-1 C-band SAR backscatter, polarimetry and interferometry to snow accumulation in the Alps. *Remote Sensing of Environment*, *316*, 114477. https://doi.org/10.1016/J.RSE.2024.114477

Jeelani, G., Feddema, J. J., Van Der Veen, C. J., & Stearns, L. (2012). Role of snow and glacier melt in controlling river hydrology in Liddar watershed (western Himalaya) under current and future climate. *Water Resources Research*, *48*(12). https://doi.org/10.1029/2011WR011590

Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J. C., Mathis, M., & Brumby, S. P. (2021). Global land use / land cover with Sentinel 2 and deep learning. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, *2021-July*, 4704–4707. https://doi.org/10.1109/IGARSS47720.2021.9553499

Li, D., Wrzesien, M. L., Durand, M., Adam, J., & Lettenmaier, D. P. (2017). How much runoff originates as snow in the western United States, and how will that change in the future? *Geophysical Research Letters*, *44*(12), 6163–6172. https://doi.org/10.1002/2017GL073551

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

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., & De Lannoy, G. J. M. (2019). Snow depth variability in the Northern Hemisphere mountains observed from space. *Nature Communications 2019 10:1*, *10*(1), 1–12. https://doi.org/10.1038/s41467-019-12566-y

Luojus, K., Pulliainen, J., Takala, M., Lemmetyinen, J., Mortimer, C., Derksen, C., Mudryk, L., Moisander, M., Hiltunen, M., Smolander, T., Ikonen, J., Cohen, J., Salminen, M., Norberg, J., Veijola, K., & Venäläinen, P. (2021). GlobSnow v3.0

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

Northern Hemisphere snow water equivalent dataset. *Scientific Data*, *8*(1). https://doi.org/10.1038/S41597-021-00939-2

Marti, R., Gascoin, S., Berthier, E., De Pinel, M., Houet, T., & Laffly, D. (2016). Mapping snow depth in open alpine terrain from stereo satellite imagery. *The Cryosphere*, *10*(4), 1361–1380. https://doi.org/10.5194/TC-10-1361-2016

Nagy, L. (2006). European high mountain (alpine) vegetation and its suitability for indicating climate change impacts. *Biology and Environment*, *106*(3), 335–341. https://doi.org/10.3318/BIOE.2006.106.3.335

Palomaki, R. T., & Sproles, E. A. (2023). Assessment of L-band InSAR snow estimation techniques over a shallow, heterogeneous prairie snowpack. *Remote Sensing of Environment*, *296*, 113744. https://doi.org/10.1016/J.RSE.2023.113744

Paulson, R. W., Chase, E. B., Roberts, R. S., & Moody, D. W. (1991). National water summary 1988–89 — Hydrologic events and floods and droughts. *United States Geological Survey Water Supply Paper 2375, U.S. Government Printing Office*, *591 pp*. https://doi.org/10.3133/WSP2375

Qu, X., & Hall, A. (2006). Assessing Snow Albedo Feedback in Simulated Climate Change. *Journal of Climate*, *19*(11), 2617–2630. https://doi.org/10.1175/JCLI3750.1

Rubel, F., Brugger, K., Haslinger, K., & Auer, I. (2017). The climate of the European Alps: Shift of very high resolution Köppen-Geiger climate zones 1800–2100. *Meteorologische Zeitschrift*, *26*(2), 115–125. https://doi.org/10.1127/METZ/2016/0816

Sharma, S. (2000). *An Overview of Snow and Avalanche Research in Indian Himalaya*.

Sharma, S. S., & Ganju, A. (2000). Complexities of avalanche forecasting in Western Himalaya — an overview. *Cold Regions Science and Technology*, *31*(2), 95–102. https://doi.org/10.1016/S0165-232X(99)00034-8

Singh, D. K., Tanniru, S., Singh, K. K., Negi, H. S., & Ramsankaran, R. (2024). Passive microwave remote-sensing-based high-resolution snow depth mapping for Western Himalayan zones using multifactor modeling approach. *Cryosphere*, *18*(1), 451–474. https://doi.org/10.5194/TC-18-451-2024

Snehmani, Singh, M. K., Gupta, R. D., Bhardwaj, A., & Joshi, P. K. (2015). Remote sensing of mountain snow using active microwave sensors: a review. *Geocarto International*, *30*(1), 1–27. https://doi.org/10.1080/10106049.2014.883434

Sodemann, H., & Zubler, E. (2010). Seasonal and inter-annual variability of the moisture sources for alpine precipitation during 1995-2002. *International Journal of Climatology*, *30*(7), 947–961. https://doi.org/10.1002/JOC.1932;PAGEGROUP:STRING:PUBLICATION

Sturm, M., & Liston, G. E. (2021). Revisiting the Global Seasonal Snow Classification: An Updated Dataset for Earth System Applications. *Journal of Hydrometeorology*, *22*(11), 2917–2938. https://doi.org/10.1175/JHM-D-21-0070.1

Tanniru, S., & Ramsankaran, R. A. A. J. (2023). Passive Microwave Remote Sensing of Snow Depth: Techniques, Challenges and Future Directions. *Remote Sensing*, *15*(4). https://doi.org/10.3390/RS15041052

Tedesco, M., & Jeyaratnam, J. (2019). *AMSR-E/AMSR2 Unified L3 Global Daily 25 km EASE-Grid Snow Water Equivalent (AU_DySno, Version 1)*. NASA National Snow and Ice Data Center Distributed Active Archive Center.

https://doi.org/10.5194/egusphere-2026-1207
Preprint. Discussion started: 8 April 2026

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., … Xu, X. (2022). Review article: Global monitoring of snow water equivalent using high-frequency radar remote sensing. *The Cryosphere*, *16*(9), 3531–3573. https://doi.org/10.5194/TC-16-3531-2022

Viviroli, D., Dürr, H. H., Messerli, B., Meybeck, M., & Weingartner, R. (2007). Mountains of the world, water towers for humanity: Typology, mapping, and global significance. *Water Resources Research*, *43*(7). https://doi.org/10.1029/2006WR005653

Viviroli, D., Weingartner, R., & Messerli, B. (2003). Assessing the hydrological significance of the world's mountains. *Mountain Research and Development*, *Vol. 23, No. 1*, 32–40. https://doi.org/10.1659/0276

Vollrath, A., Mullissa, A., & Reiche, J. (2020). Angular-Based Radiometric Slope Correction for Sentinel-1 on Google Earth Engine. *Remote Sensing 2020, Vol. 12, Page 1867*, *12*(11), 1867. https://doi.org/10.3390/RS12111867

Zhu, X., Gu, L., Li, X., & Jiang, T. (2022). Snow Depth Estimation Based on Parameter Combinations Selection and Machine Learning Algorithm Using C-Band SAR Data in Northeast China. *IEEE Geoscience and Remote Sensing Letters*, *19*. https://doi.org/10.1109/LGRS.2021.3129998