https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

Retrieval of Snow Depth in the Songhua River Basin Based on Multi-Source Data Fusion and Machine Learning

Tingyu Lu[1], Qiang Wang[1], Lijuan Zhang[2], Yufeng Zhao[2], Lei Wang[1]

[1]School of Surveying and Mapping Engineering, Heilongjiang Institute of Technology, Harbin, 150050, China
[2] College of Geographical Sciences, Harbin Normal University, Harbin 150025, China

Correspondence to: Lijuan Zhang (zhlj@hrbnu.edu.cn)

Abstract. Snow is a major surface feature of the Songhua River Basin in winter and a critical source of spring runoff. Accurately capturing its spatiotemporal distribution is crucial for hydrological modeling, water resource management, and disaster control. This study focuses on the Songhua River Basin, integrating AMSR-2 brightness temperature, ERA5-Land reanalysis, SRTM DEM, CLCD land cover data, and snow depth observations from 61 meteorological stations over ten snow seasons (2013-2023). Two ensemble learning models, XGBoost and Random Forest (RF), were developed for snow depth estimation. Four feature combination scenarios (S1-S4) were designed to investigate the impact of different input variables on model performance. The SHAP method was employed to analyze feature importance and spatiotemporal heterogeneity. Results showed that Scenario S4, integrating multi-source features, yielded optimal performance. The XGBoost model performed better on the test set (RMSE = 4.44 cm, R^2 = 0.6738), demonstrating superior generalization compared to the RF model, which exhibited noticeable overfitting. Brightness temperature difference (18.7H-36.5H), air temperature, and longitude were core features. Estimation accuracy showed significant spatiotemporal variations: highest during the stable snow cover period (December-February), peaking in January; good for snow depth ≤ 15 cm, but errors increased significantly for ≥25 cm. Accuracy was better in areas with gentle terrain and homogeneous land surface, with the best performance over cropland and worst over forests. This multi-source data fusion and XGBoost framework effectively enhances snow depth estimation accuracy in the Songhua River Basin.

1 Introduction

Snow, as one of the most active components of the cryosphere, is a critical part of the global water cycle and climate system (Armstrong and Brun, 2008; Flanner et al., 2011). Its high albedo significantly influences the surface radiation balance, while its low thermal conductivity insulates the underlying soil, thereby affecting land-atmosphere energy exchange (Marks and Dozier, 1992; Zhang, 2005). On a global scale, both the extent of Northern Hemisphere snow cover and snow water equivalent (SWE) have exhibited significant declining trends over the past decades, serving as strong evidence of climate change (Brown and Robinson, 2011; Pulliainen et al., 2020). From a hydrological perspective, seasonal snowpack acts as a crucial freshwater reservoir. Snowmelt provides drinking water, agricultural irrigation, and industrial water for over onesixth of the global population. In many arid and semi-arid regions, snowmelt is the primary source of river recharge (Barnett et al., 2005; Sturm et al., 2017). Therefore, accurate monitoring of snow depth and its spatiotemporal distribution is of paramount importance for understanding climate change, managing water resources, forecasting floods, and assessing avalanche risks (Dozier et al., 2016; Bormann et al., 2018).

Obtaining large-scale, high-accuracy snow depth data has always been a core challenge in snow remote sensing. Traditional snow depth measurement methods primarily include manual/automatic measurements at ground meteorological stations and snow depth rod networks. While these methods provide temporally continuous, high-accuracy point data, their spatial representativeness is limited. Particularly in topographically complex mountainous areas, sparse station networks struggle to capture the spatial heterogeneity of snow depth caused by drastic variations in terrain and vegetation (Miller et al., 2022; López-Moreno et al., 2011). In contrast, satellite remote sensing technology, with its advantages of macroscopic, rapid, and periodic observation, offers unprecedented opportunities for regional to global scale snow depth mapping.

Currently, mainstream satellite snow depth retrieval methods can be broadly categorized into two types: passive microwave remote sensing and active microwave remote sensing. Passive microwave remote sensing, such as AMSR-E/AMSR-2 sensors aboard Aqua satellites, exploits the sensitivity of brightness temperature (TB) at different frequencies (e.g., 18.7 GHz and 36.5 GHz) to volume scattering within the snow layer. It estimates snow depth by establishing semi-empirical models (e.g., the classic Chang algorithm) (Chang et al., 1987; Kelly et al., 2003). These methods offer advantages of high temporal resolution (daily scale) and global coverage. However, their inherent coarse spatial resolution (~25 km) leads to significant mixed pixel problems in mountainous areas, and the signal saturates under deep snow (typically >50-100 cm) and wet snow conditions, causing substantial decreases in retrieval accuracy (Tedesco and Narvekar, 2010; Vuyovich et al., 2014). Furthermore, factors like forest cover and variations in snow grain size introduce considerable uncertainty in passive microwave retrievals (Derksen, 2008; Foster et al., 2005).

Active microwave remote sensing, particularly Synthetic Aperture Radar (SAR), opens new pathways for snow monitoring due to its high spatial resolution and all-weather, day-and-night imaging capabilities. Research indicates that C-band (~5.4 GHz) SAR signals can penetrate dry snow to some extent, and the strength of volume scattering is closely related to snow depth and snowpack structure (Ulaby and Stiles, 1980; Shi and Dozier, 2000). For instance, Lievens et al. (2019, 2022) developed a snow depth retrieval algorithm based on backscatter coefficient change detection using Sentinel-1 data, successfully applying it to the European Alps. However, this method is also constrained by wet snow, forests, and steep terrain. Recent studies have further explored using dual-polarization SAR data (e.g., the polarimetric scattering angle α) to improve snow depth estimation accuracy, emphasizing the significant impact of the Local Incidence Angle (LIA) on the signal depolarization process (Mariani et al., 2026; Jans et al., 2025). Nevertheless, relying solely on SAR data remains insufficient to fully disentangle the complex relationships involving snow depth, density, grain size, and underlying surface contributions.

To overcome the limitations of single remote sensing data sources, fusing multi-source data and introducing Machine Learning methods has become a cutting-edge trend in current snow depth retrieval research. Machine learning models, such

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

as eXtreme Gradient Boosting (XGBoost) and Random Forest (RF), can effectively mine and establish complex non-linear relationships between multi-source features (e.g., passive microwave brightness temperature, SAR backscatter, topographic parameters, land cover types, meteorological reanalysis data) and in-situ snow depth measurements (Yang et al., 2020; Xiao et al., 2018; Singh et al., 2024). Existing studies have successfully applied machine learning methods in regions like the Tibetan Plateau (Zhou et al., 2024) and the European Alps (Dunmire et al., 2024), significantly improving the spatiotemporal resolution and accuracy of snow depth retrievals. These studies demonstrate that by integrating multi-source data, machine learning models can better characterize the spatial heterogeneity of snow depth and partially overcome the limitations of physical models.

The Songhua River Basin, as a crucial commodity grain base and energy base in China, relies on spring runoff primarily sourced from basin-wide snowmelt from the previous winter. Accurately understanding the spatiotemporal distribution and variation of snow cover in this basin is of significant practical importance for spring flood forecasting, agricultural water scheduling, and ecological security assessment. However, current research on snow depth retrieval for this basin remains insufficient. Traditional retrieval products based solely on single-source passive microwave data struggle to meet the demands of basin-scale applications. Therefore, this study aims to develop a high-accuracy snow depth retrieval method suitable for the Songhua River Basin. We will fully utilize multi-source data, including AMSR-2 brightness temperature data, ERA5-Land atmospheric reanalysis data, high-resolution topographic data, land cover type data, and combine these with ground-based snow depth observations to construct retrieval models based on machine learning algorithms. Through systematic feature engineering, model performance evaluation, and feature importance analysis, this study aims to: (1) Investigate the impact of different predictor variable combinations on snow depth retrieval accuracy; (2) Reveal the spatiotemporal distribution patterns of snow depth within the basin; (3) Explore the spatiotemporal differentiation patterns between snow depth estimation accuracy and snow cover distribution.

2. Materials and Methods

2.1 Study area

The Songhua River Basin (see Fig. 1) is located in the northern part of Northeast China (119°52'–132°31' E, 41°42'–51°38' N). It is one of the seven major river systems in China, with a total drainage area of approximately 556,800 km². The Songhua River originates from Heaven Lake on Changbai Mountain and is formed by the confluence of two major tributaries, the Nen River and the Second Songhua River, with a mainstream length of about 2328 km (Editorial Committee of Encyclopedia of Rivers and Lakes in China, 2014). The overall topography of the basin is characterized by higher elevations in the southeast and lower elevations in the northwest. The western and northern parts are dominated by the Greater Khingan Range and the Yilehuli Mountains, the eastern part by the Lesser Khingan Range and the Changbai Mountains, and the central part consists of the vast Songnen Plain. Snow cover is the most prominent surface feature of the basin during winter and serves as the primary source of spring runoff. The stable snow cover period within the basin

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

typically begins in late October to early November and ends in early April of the following year, lasting 5–6 months. The maximum snow depth usually occurs from late February to early March, with average snow depths reaching 30–50 cm in mountainous areas and exceeding 80 cm in localized deep snow zones (Li et al., 2020). As temperatures rise in spring, extensive snowmelt occurs from late March to April, forming a distinct spring flood.

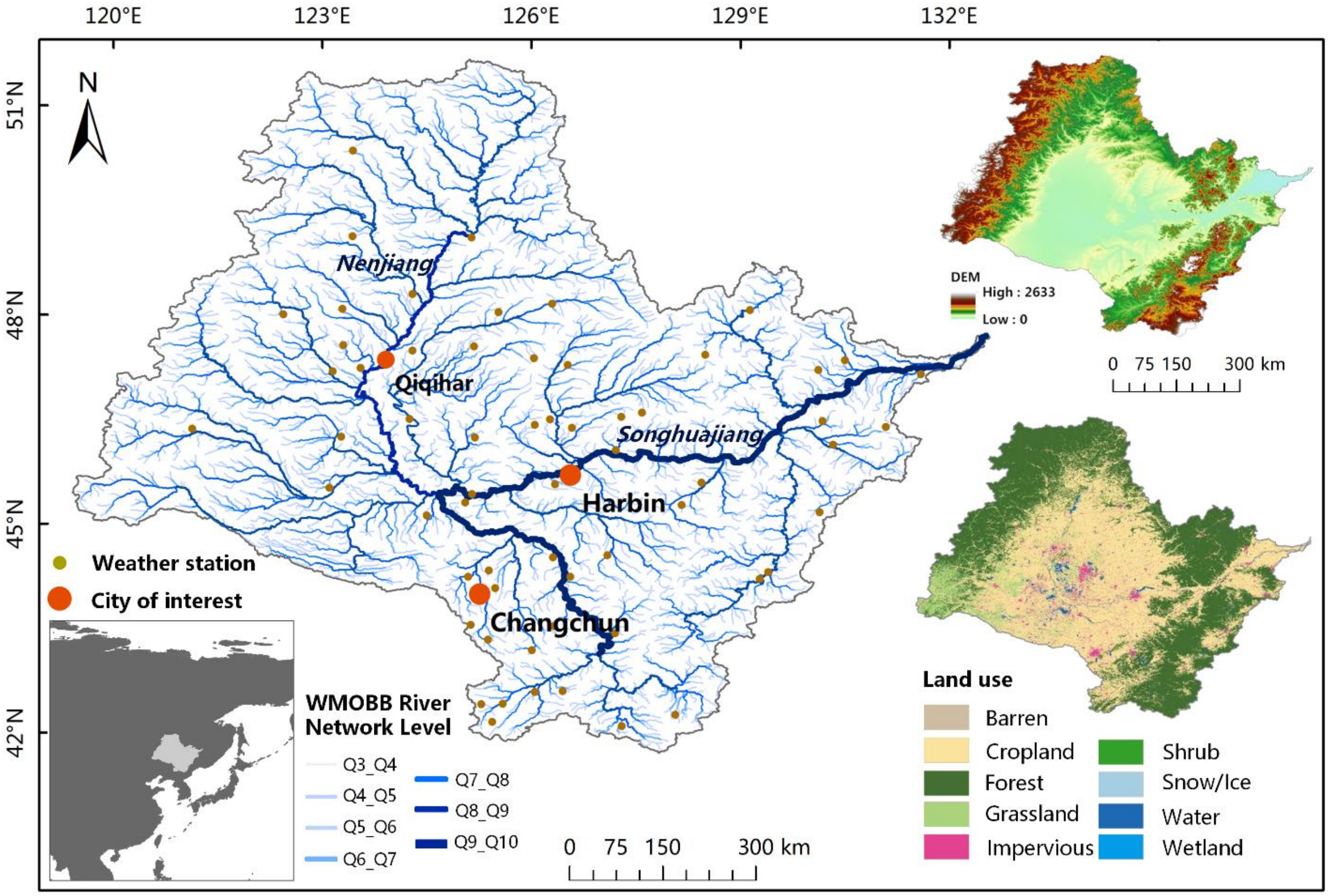


Figure 1: Spatial distribution of the weather stations in the study area.

2.2 Data

2.2.1 AMSR-2

This study utilizes the AMSR-E/AMSR2 Unified L3 Daily Product (Meier et al., 2018), released by the U.S. National Snow and Ice Data Center (NSIDC), which integrates observations from the AMSR-E and AMSR-2 sensors. This dataset has a unified spatial resolution of 12.5 km and provides global coverage using the North (EPSG:3411) and South (EPSG:3412) Hemisphere polar stereographic projections. The dataset is available from July 2, 2012, to the present and includes 8 channels: vertically polarized and horizontally polarized data at four frequency bands—18.7 GHz, 23.8 GHz, 36.5 GHz, and 89.0 GHz. Considering that seasonal snow cover in the Songhua River Basin is primarily concentrated in winter and spring,

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

this study defines each period from November 1 to March 31 of the following year as a complete snow season, used to construct the time series for spatiotemporal snow dynamics analysis. Consequently, this study ultimately acquired and processed AMSR-2 data covering a total of 10 snow seasons from 2013 to 2023.

To minimize the interference of daytime surface temperature increases and potential snowmelt processes on passive microwave signals, thereby enhancing the accuracy and stability of snow depth retrieval, this study prioritizes daily descending orbit observations (approximately 01:30 local time) as the primary input. For sporadic missing data caused by satellite orbital gaps, ascending orbit observations from the same day (approximately 13:30 local time) are used as substitutes. All raw data were downloaded via NASA Earthdata and underwent preprocessing steps including format conversion, regional cropping, and necessary quality flag screening to construct a standardized brightness temperature dataset suitable for this study.

2.2.2 ERA5-Land

ERA5-Land is a global high-resolution land reanalysis dataset released by the European Centre for Medium-Range Weather Forecasts (ECMWF). It is generated through a time-consistent downscaling of the ECMWF ERA5 climate reanalysis (Muñoz-Sabater et al., 2021). This dataset provides continuous and physically consistent time series of land surface state variables with global coverage, featuring a spatial resolution of 0.1°×0.1° (approximately 9 km) and temporal resolutions including both hourly and monthly scales, covering the period from 1950 to the present. This study selected meteorological and land surface variables from the ERA5-Land dataset that are closely related to snow processes and surface energy balance, including snow density, snow layer temperature, air temperature, and soil temperature. These variables play a crucial role in snow accumulation, ablation, and energy exchange processes.

2.2.3 Elevation data

The SRTM Digital Elevation Model (DEM) Version 004 data used in this study is one of the most widely applied elevation data products. This dataset was acquired using C-band Interferometric Synthetic Aperture Radar technology, with a spatial resolution of 90 m and available in Geo-TIFF format. It was downloaded from the Google Earth Engine platform (last accessed: December 22, 2025). This data was selected to provide crucial topographic parameters for the snow depth model, enabling analysis of the influence of terrain on snow accumulation, ablation, and redistribution processes.

2.2.4 Land cover data

Different land cover types significantly influence snow distribution as well as the emission and scattering characteristics of microwave signals. To quantify this influence, this study employs the 30 m resolution annual China Land Cover Dataset (CLCD) developed by Yang and Huang (2021) based on Landsat imagery, spanning the period from 1985 to 2025. This dataset was generated by processing over 330,000 Landsat scenes on the Google Earth Engine cloud platform and represents the first Landsat-derived annual land cover product for China, comprising nine land cover types. In terms of data processing,

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

the dataset was classified by constructing time-series indicators and inputting them into a Random Forest classifier, followed by post-processing incorporating spatiotemporal filtering and logical reasoning, which significantly enhanced the spatiotemporal consistency of the classification results. The data are available via the Zenodo platform (https://zenodo.org/record/5816591), last accessed on January 21, 2026.

2.2.5 Snow depth observation data

The observed snow depth data used in this study were obtained from 61 meteorological stations across the Songhua River Basin, covering the period from November 2013 to March 2023 (a total of 10 snow seasons). These data were sourced from the National Meteorological Science Data Center of the China Meteorological Administration (http://data.cma.cn/) and include comprehensive information such as station number, location, latitude and longitude, elevation, observation date, snow depth, and quality control codes. The spatial distribution of each station is shown in Fig. 1. Fig. 2 presents the frequency distribution histograms of snow depth observations for the training and testing sets during the 2013–2023 period. For the training set, the maximum snow depth reaches approximately 85 cm, with about 90% of samples having snow depths between 1 and 18 cm. For the testing set, the maximum snow depth reaches approximately 46 cm, with about 90% of samples having snow depths between 1 and 17 cm.

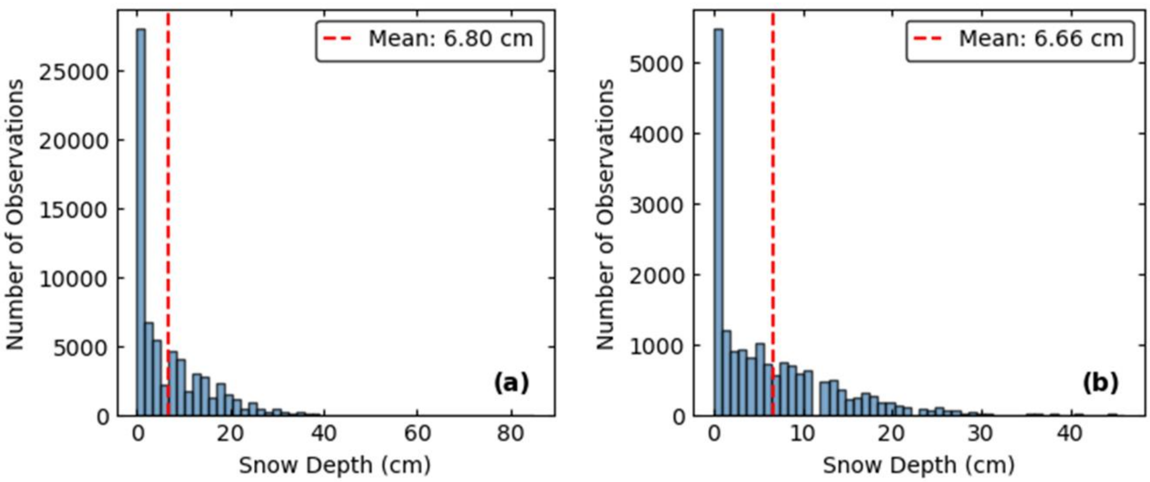


Figure 2: Histograms of snow depth observations from (a) training set (b) test set.

3. Methods

3.1 Data processing

The original AMSR-E/AMSR2 Unified L3 Daily Product (AU_SI12) data used in this study are stored in HDF-EOS5 (he5) format, with each file containing 62 sub-datasets. Using the Python programming language combined with the GDAL library, an automated batch processing routine was developed. The specific processing workflow is as follows: First, the required frequency band brightness temperature subsets were extracted from the he5 files and batch-converted into Geo-TIFF raster

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

data with complete geospatial coordinate information. Subsequently, based on the latitude and longitude coordinates and observation dates of each meteorological station, spatial matching and temporal alignment were performed on the daily brightness temperature rasters for different frequency channels, establishing the spatiotemporal correlation between stations and pixels. Finally, pixel brightness temperature values were extracted, and brightness temperature differences were further calculated to provide standardized input features for subsequent modeling.

The download and synthesis of ERA5-Land reanalysis data were completed on the Google Earth Engine (GEE) platform. Hourly-scale data were aggregated to generate daily mean values to match the daily scale of AMSR-2 data. After exporting Geo-TIFF raster data, near-surface air temperature (2 m), snow density, soil temperature (0–7 cm), and snow layer temperature data at the meteorological station locations were batch-extracted.

All terrain factor extraction and processing were completed on the GEE platform. First, the DEM data were clipped using the Songhua River Basin vector boundary. Then, based on the processed elevation data, slope, aspect, and surface roughness were further extracted. For land cover type data, annual land cover classification data from 2013 to 2023 were selected. To accurately characterize the influence of underlying surface properties on snow processes, a year-by-year matching principle was adopted when extracting the corresponding land cover types for each meteorological station. Specifically, the land cover category used for each station in each year was strictly aligned with the actual classification results of that year, ensuring that the land cover data temporally reflected the dynamic changes in surface cover surrounding the stations.

For missing data labeled as "999999" in the station snow depth observations, differentiated processing strategies were adopted based on the duration of consecutive missing periods. If the consecutive missing days did not exceed 3 days, linear interpolation based on valid observations before and after the gap was applied for infilling. If consecutive missing days exceeded 3 days, the records were considered invalid and deleted, and thus not included in subsequent analysis and modeling. Additionally, records labeled as "999990," representing observed trace snow cover, were uniformly treated as "0 cm."

3.2 Feature engineering and dataset splitting

To construct input data suitable for machine learning models, multiple key features were extracted from multi-source data, and the dataset was divided into training and testing sets in chronological order. The feature set constructed in this study comprises the following four categories, totaling 24 feature variables. (1) Based on AMSR-2 brightness temperature data, vertical and horizontal polarization data at four frequency bands (18.7 GHz, 23.8 GHz, 36.5 GHz, and 89.0 GHz) were extracted, and five sets of brightness temperature differences (18.7-36.5 (V/H), 23.8-89.0 (V/H), 18.7H-89.0V) were calculated, yielding 13 features in total (Wang, Y. et al., 2019; Wang, Z. et al., 2022). (2) Variables closely related to snow processes were extracted from ERA5-Land reanalysis data, including four parameters: air temperature (AT), snow density, soil temperature (0–7 cm, ST), and snow layer temperature (SLT). (3) Using SRTM DEM as the data source, DEM data for the study area were obtained, from which four terrain factors—elevation, slope, aspect, and surface roughness—were derived. (4) Using CLCD annual data, land cover types aligned year-by-year with meteorological stations were extracted and converted into categorical features.

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

The results of Pearson correlation coefficient analysis (see Fig. 3) indicate that microwave brightness temperature data have significant indicative value for predicting snow depth, consistent with findings from existing studies (Hallikainen et al., 2024). The correlations between individual features and snow depth show distinct differentiation: all brightness temperature difference features (particularly the 23.8V–89.0V difference, with the most significant correlation, $r = 0.6604$) exhibited the strongest positive correlations, indicating that brightness temperature differences are key indicators reflecting snow depth. In contrast, single-channel brightness temperatures (especially 89.0V, $r = -0.6777$) and thermal factors such as air temperature and snow layer temperature showed significant negative correlations, consistent with the fundamental physical laws governing the interaction between snow, radiation, and temperature. By comparison, although terrain features (e.g., slope, aspect, roughness) and geographical location (longitude, latitude) were statistically significant, their linear correlations with snow depth were generally weak, suggesting they may influence snow distribution through more complex nonlinear mechanisms. Snow density ($r = 0.4054$) demonstrated a certain explanatory power, while land cover type showed a weak negative correlation.

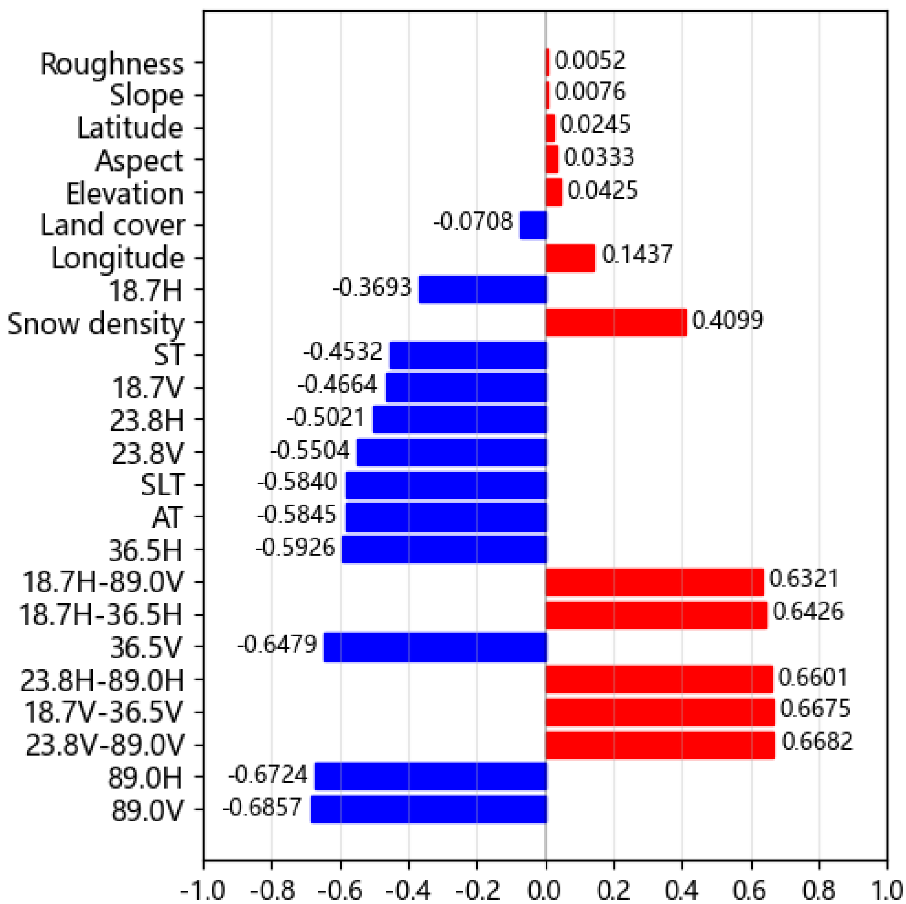


Figure 3: Ranking of Pearson correlation coefficients between features and snow depth.

Snowfall in the Songhua River Basin is primarily concentrated from November to March of the following year, occurring earlier in the northern part and slightly later in the southern part. Therefore, this study selected snow season data from 2013 to 2023 for model training and performance evaluation. This partitioning approach primarily considers the temporal dependence characteristics inherent in time-series data. Random proportional splitting methods would lead to future information leaking into the training process, which not only distorts model evaluation results but also causes overly optimistic estimates of predictive performance (Hyndman & Athanasopoulos, 2014). Consequently, maintaining the chronological order of data is crucial for dataset splitting. Furthermore, training data must always precede testing data, thereby simulating the real-world scenario of training models using historical data to predict future data. Meanwhile, to explore how different input features affect snow depth retrieval model performance, four input combination scenarios (S1–S4) were designed to systematically assess the contribution of each feature combination to retrieval accuracy. The input feature composition for each scenario is presented in Table 1.

Table 1. A detailed description of the input predictor variables based on four combination scenarios**.**

Input Scenarios	Predictor variables
S1	18.7H, 18.7V, 23.8H, 23.8V, 36.5H, 36.5V, 89.0H, 89.0V
S2	18.7H, 18.7V, 23.8H, 23.8V, 36.5H, 36.5V, 89.0H, 89.0V, 18.7V-36.5V, 18.7H-36.5H, 23.8V-89.0V, 18.7H-89.0V, 23.8H-89.0H
S3	18.7H, 18.7V, 23.8H, 23.8V, 36.5H, 36.5V,89.0H, 89.0V, 18.7V-36.5V, 18.7H-36.5H, 23.8V-89.0V, 18.7H-89.0V, 23.8H-89.0H, Snow density, AT, ST, SLT
S4	18.7H, 18.7V, 23.8H, 23.8V, 36.5H, 36.5V, 89.0H, 89.0V, 18.7V-36.5V, 18.7H-36.5H, 23.8V-89.0V,18.7H-89.0V, 23.8H-89.0H, Snow density, AT, ST, SLT, Latitude, Lon, Elevation, Slope, Aspect, Roughness, Land cover

3.3 Machine learning model

To investigate the differences in snow depth retrieval performance arising from distinct ensemble learning mechanisms, this study selected two representative machine learning models: XGBoost and Random Forest. XGBoost, based on the Boosting strategy, iteratively reduces residuals through sequential training and excels at capturing complex nonlinear feature interactions. In contrast, Random Forest, based on the Bagging strategy, effectively reduces model variance through parallel training of multiple decision trees and a voting mechanism, demonstrating strong robustness against overfitting. By comparing these two models, the effectiveness of multi-source data fusion can be more comprehensively evaluated, and the robustness of the retrieval models can be rigorously tested.

3.3.1 XGBoost

XGBoost is an efficient gradient boosting decision tree algorithm widely applied to regression problems (Chen & Guestrin, 2016). It enhances prediction accuracy by integrating multiple weak learners and minimizing a regularized loss function. Its objective function can be expressed as:

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

$$\mathcal{L}(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \tag{1}$$

where the first term is the loss function, which measures the difference between the predicted value $\hat{y}_i$ and the true value y_i; the second term is the regularization term, used to control model complexity and prevent overfitting. XGBoost optimizes the objective function by iteratively adding new trees, with each iteration fitting the negative gradient of the current model, thereby progressively reducing the error.

3.3.2 Random Forest

Random Forest regression is a data-driven model based on ensemble learning. By constructing multiple decision trees and aggregating their prediction results, it can effectively fit complex nonlinear relationships while possessing strong resistan ce to overfitting and robustness (Breiman, 2001). The core of the model lies in its dual randomness mechanism: bootstrap sampling is used to generate different training sets, and during node splitting, the optimal partition is selected only from a random subset of features, thereby enhancing tree diversity and generalization performance. In regression tasks, the final output is obtained by averaging the results from all decision trees.

$$\hat{f}_{RF}(x) = \frac{1}{B}\sum_{b=1}^{B} T_b(x) \tag{2}$$

where $T_b(x)$ is the prediction of the b-th tree for sample x. This model is widely applied in fields such as environmental modeling and hydrological forecasting due to its ability to handle high-dimensional data and its built-in feature importance evaluation.

3.4 Performance evaluation

This study employed three evaluation metrics to comprehensively assess the predictive performance of the models: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2). These metrics reflect the degree of discrepancy between the model-predicted values and the actual observed values from different perspectives. RMSE is the standard deviation of prediction errors. By squaring the errors, it imposes a higher penalty on larger errors and is sensitive to outliers in predictions. The value range of RMSE is [0, +∞), with smaller values indicating higher model prediction accuracy. In the formula, SD_i represents the actual snow depth of the i-th sample, $\widehat{SD}_i$ represents the corresponding predicted value, and n is the number of samples.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(SD_i - \widehat{SD}_i)^2} \tag{3}$$

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

MAE calculates the average of the absolute differences between predicted values and actual values. It is less sensitive to outliers compared to RMSE and can more robustly reflect the average prediction error level of the model. The value range of MAE is also [0, +∞), with smaller values indicating better model performance.

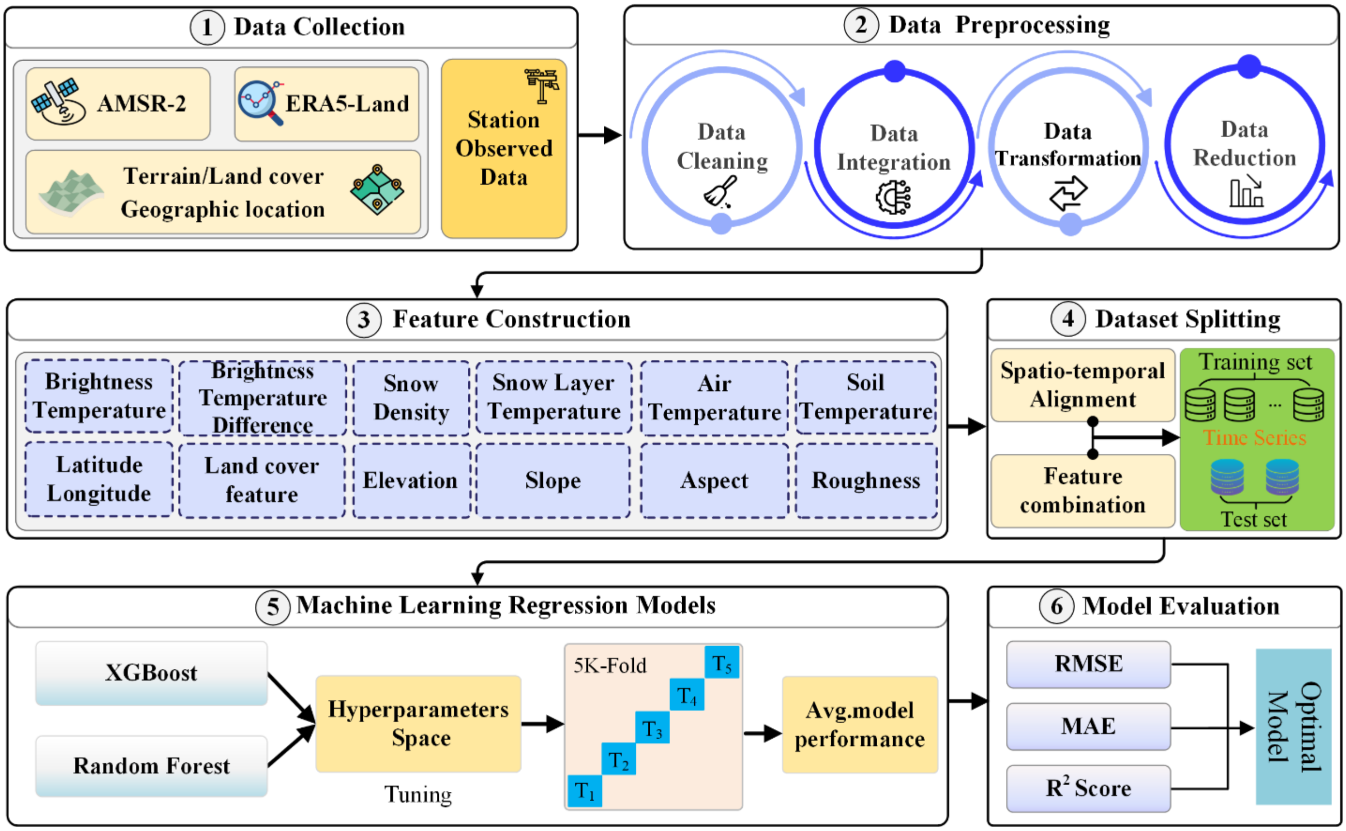
$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} | SD_i - \widehat{SD}_i | \tag{4}$$

R^2 reflects the degree to which the model explains the variability of the target variable and measures the goodness of fit between predicted values and actual values. The value range of $R^2$is (-∞, 1], with values closer to 1 indicating stronger explanatory power of the model for the data. R^2= 1 represents perfect prediction, R^2= 0 indicates that the model performs no better than simply predicting the mean, and negative values imply that the model performs worse than the simple mean baseline.

$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{n}(SD - \widehat{SD}_i)^2}{\sum_{i=1}^{n}(SD_i - \bar{SD})^2} \tag{5}$$

In the practical application of snow depth prediction, RMSE and MAE directly reflect the absolute magnitude of prediction errors in the original measurement units (cm), facilitating operational interpretation. MSE, as a derived metric, provides an overall measure of squared errors, while R^2 evaluates the overall model fit from the perspective of relative proportion. In summary, the complete methodological and technical framework for snow depth retrieval in the Songhua River Basin integrating multi-source data and machine learning is illustrated in Fig. 4.

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026



275 **Figure 4: Technical flowchart of the study.**

4. Results and Analysis

4.1 Sensitivity to Training Sample Size

To determine the appropriate sample size required for training the XGBoost and RF models while ensuring model generalization capability, this study conducted a sensitivity assessment using a gradient of "increasing sample size by 10% each time," while adopting a "time-series independent validation" strategy for dataset partitioning. Specifically, winter data from 2013–2021 were used as the training set, and temporally independent winter data from 2021–2023 were used as the testing set, with sample sizes progressively increasing from an initial 6,977 to 69,770. The evaluation results based on RMSE, MAE, and R^2 (see Fig. 5) indicate that, from the overall trend perspective, as the training sample size increased, both RMSE and MAE for the XGBoost and RF models initially decreased and then stabilized, while R^2 initially increased and then stabilized. Once the sample size exceeded 50,000, the fluctuation range of each accuracy metric significantly diminished, essentially remaining within a stable interval.

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

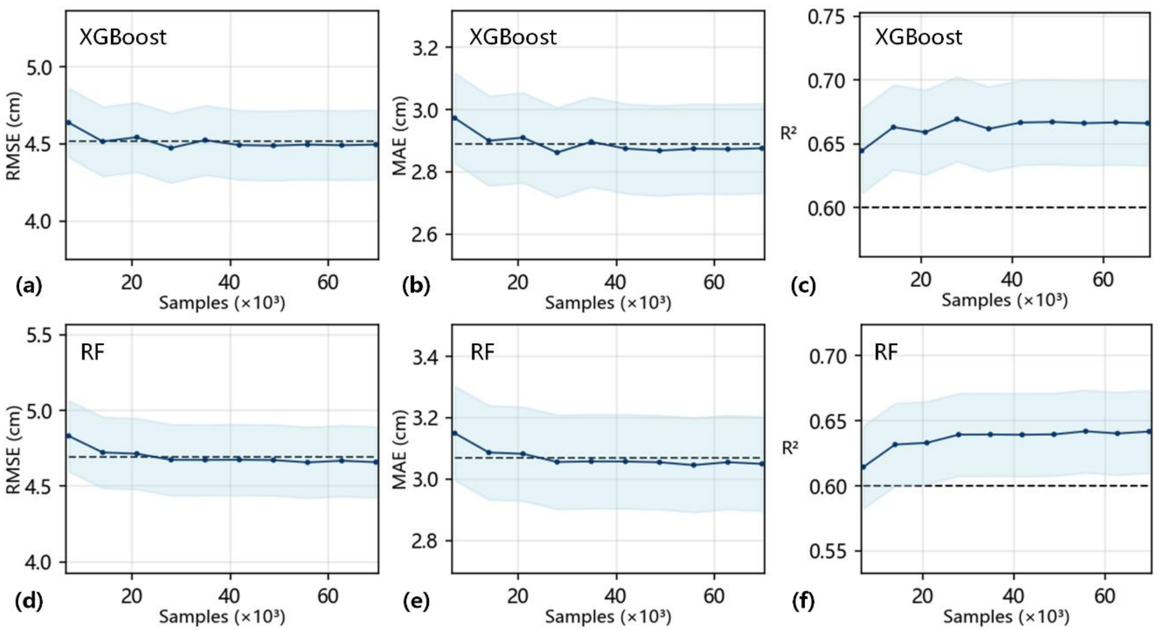


Figure 5: Performance trends of models with increasing training sample size: (a)–(c) XGBoost, (d)–(f) Random Forest.

4.2 Model validation

By evaluating the fitted XGBoost and RF algorithms under the four combinations shown in Table 2, the density scatter plots of observed versus predicted snow depth are presented (Fig. 6). The results indicate that on the training set, RF generally outperforms XGBoost, with RF exhibiting lower training errors and higher goodness-of-fit across all scenarios: the minimum training RMSE for RF was only 2.41 (S4), the minimum MAE was 1.37 (S4), and the highest R^2reached 0.9181 (S4). The optimal training performance for XGBoost was RMSE of 3.78 (S4), MAE of 2.41 (S4), and R^2 of 0.7983 (S4), all inferior to RF's performance in the same scenarios. On the test set, under the optimal scenario (S4), XGBoost achieved lower test RMSE (4.44) and MAE (2.86) compared to RF (RMSE 4.49, MAE 2.89), and a higher R^2 (0.6738) than RF (0.6661). From the perspective of feature combinations, Scenarios S1 and S2 represent low-performance scenarios, where both models performed relatively poorly. Training set R^2 values were generally below 0.6, test set R^2 values were below 0.5, and RMSE (training: 5.44–5.69, test: 5.65–5.71) and MAE (training: 3.46–3.62, test: 3.73–3.77) were significantly higher compared to S3 and S4. Compared to S1 and S2, under Scenarios S3 and S4, training set R^2improved by 0.16–0.33, and test set R^2 improved by 0.13–0.21. Scenario S4 generally outperformed S3, indicating that incorporating auxiliary geographical features can effectively enhance model generalization capability.

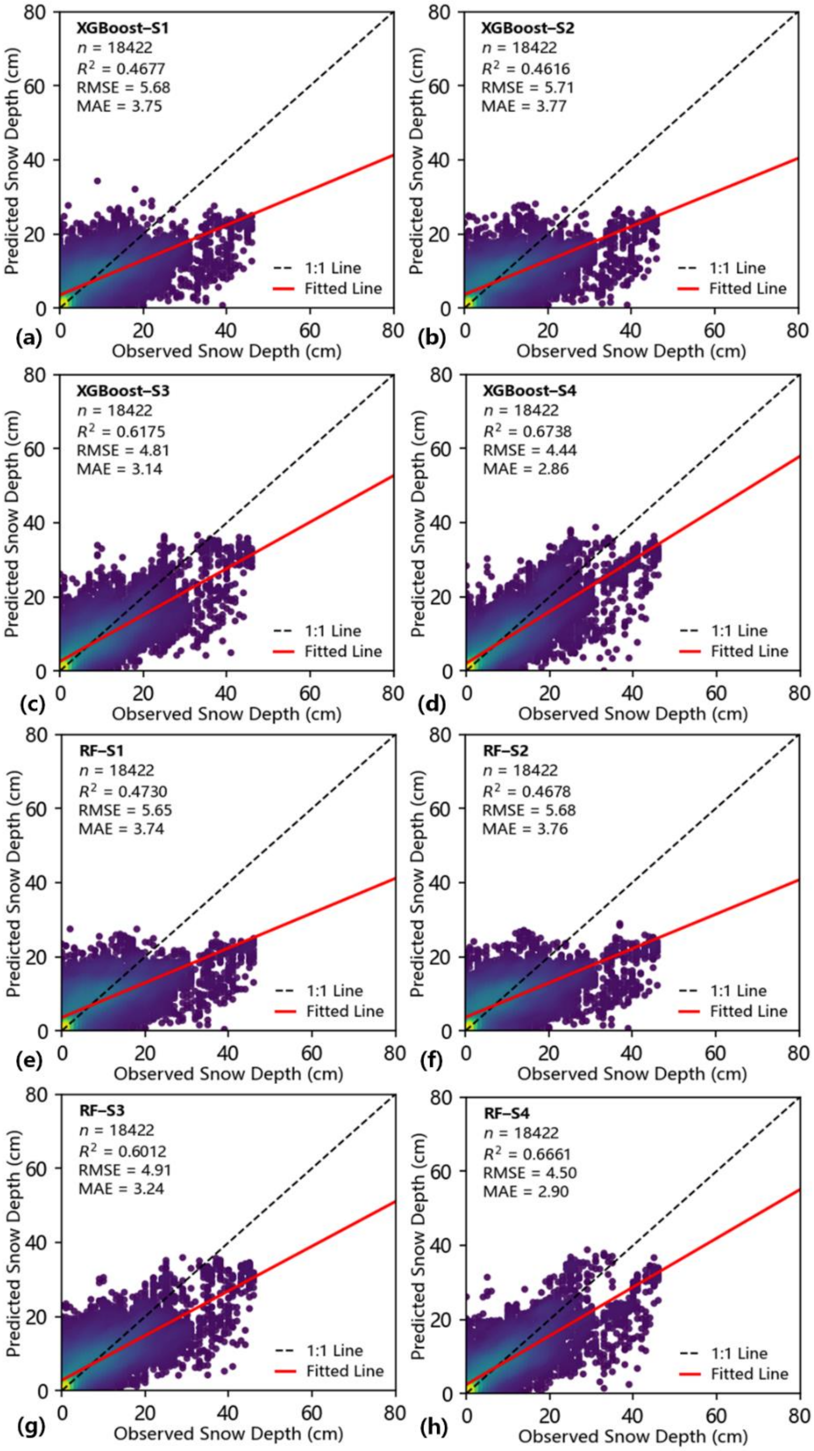
In summary, RF exhibits a more pronounced overfitting tendency, with a larger performance gap between the training and test sets: for example, in Scenario S4, the difference between RF's training R^2 (0.9181) and test R^2 (0.6661) reached 0.252, while the difference for XGBoost in the same scenario was only 0.1245. XGBoost based on Scenario S4 achieved the highest accuracy and superior generalization stability, making it the optimal model in this study.

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

Table 2. Statistical metrics for snow depth prediction results.

Models	Scenarios	Training set			Testing set		
		RMSE	MAE	R2	RMSE	MAE	R2
XGBoost	S1	5.51	3.51	0.5723	5.68	3.75	0.4677
	S2	5.69	3.62	0.5427	5.71	3.77	0.4616
	S3	4.06	2.60	0.7670	4.81	3.14	0.6175
	S4	3.78	2.41	0.7983	4.44	2.86	0.6738
RF	S1	5.44	3.47	0.5820	5.65	3.73	0.4730
	S2	5.66	3.46	0.5838	5.67	3.75	0.4698
	S3	2.56	1.61	0.9074	4.91	3.23	0.6012
	S4	2.41	1.37	0.9181	4.49	2.89	0.6661

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026



310 **Figure 6: Scatter plots of snow depth estimated by the two machine learning models versus ground-based observations under different feature combinations: (a)–(d) XGBoost, (e)–(h) Random Forest.**

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

4.3 SHAP-Based feature importance analysis

Although XGBoost and RF can provide powerful predictive capabilities, their complexity makes it difficult to reveal the specific impact of features on prediction results by relying solely on feature importance measures such as feature frequency or gain. To address this issue, we introduce SHAP, an interpretability method based on Shapley values, which can be used to measure the contribution of each feature to the model's predictions (Lundberg & Lee, 2017).

Experimental results indicate that under Scenario S4, which incorporates all input features, both the XGBoost and RF models achieved optimal snow depth retrieval performance. Therefore, based on this optimal scenario, this section employs the SHAP method to further investigate the contribution and importance ranking of different features for snow depth prediction. The SHAP bee swarm plot illustrates the magnitude and direction of the impact of sample feature values on snow depth predictions (Fig. 7). In the plot, the vertical axis represents feature importance, ranked from top to bottom in descending order of importance. Feature values are color-coded, ranging from blue (low values) to red (high values). The horizontal axis represents the SHAP value magnitude: values less than 0 indicate that the sample feature has a negative impact on snow depth prediction, while values greater than 0 indicate a positive impact.

Feature importance analysis based on the XGBoost model combined with the SHAP method reveals that snow depth retrieval results from the combined effects of three categories of factors: microwave signals, thermal conditions, and geospatial elements (Fig. 7a). Among these, the brightness temperature difference 18.7H-36.5H dominates model predictions with a SHAP value of 2.12, demonstrating that low-frequency brightness temperature differences serve as effective indicators for eliminating surface temperature influences and highlighting snow layer scattering signals. Three temperature-related features—air temperature (AT, 1.95), soil temperature (ST, 1.37), and snow layer temperature (SLT, 0.38)—all rank within the top ten (total SHAP value ≈ 3.70), revealing the decisive influence of the snow-air-soil thermal coupling system on snow depth distribution. Longitude (1.74) shows significantly higher importance than latitude (0.40), indicating that snow depth in the study area exhibits a spatial distribution pattern with a more pronounced east-west gradient compared to the north-south gradient. Other features such as elevation and snow density have relatively limited impacts on model output. For the RF model (Fig. 7b), spatial distribution, underlying surface temperature, and brightness temperature differences serve as the core driving factors. Longitude emerges as the primary predictor with the highest SHAP importance score (1.3809), exerting the most significant influence on snow depth prediction, consistent with the west-to-east spatial distribution pattern of snow cover in the Songhua River Basin. Soil temperature (ST) and microwave brightness temperature differences such as 18.7H-36.5H follow closely, contributing prominently to snow depth retrieval as underlying surface factors and microwave polarization differences, respectively. Features such as air temperature (AT), snow density, and elevation exhibit relatively lower importance, while terrain factors including slope and aspect rank as secondary influencing factors.

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

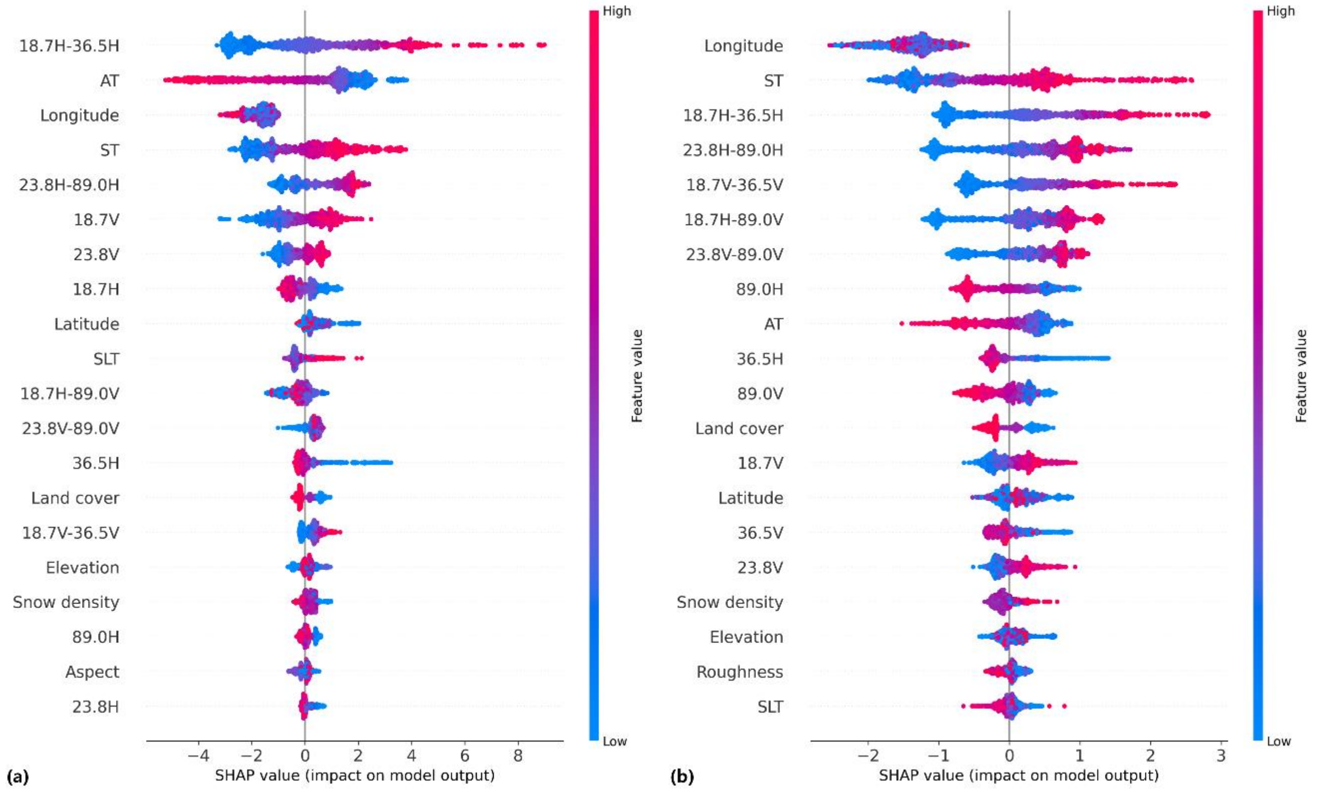


Figure 7: Feature importance analysis based on Scenario S4 and the SHAP method: (a) XGBoost, (b) Random Forest.

4.4 Spatiotemoral analysis of snow depth inversion

Influenced by the combined effects of snow accumulation and ablation patterns, underlying surface properties, and topographic relief, the model retrieval accuracy exhibits significant spatiotemporal heterogeneity. This section aims to quantitatively analyze the distribution patterns of snow depth retrieval errors under different time periods, snow depth ranges, land cover types, and terrain factors.

Figure 8 indicates that both models exhibit highly consistent seasonal fluctuations. From December to February, the R^2 values for XGBoost and RF remained stable between 0.60 and 0.71, peaking in January (XGB: 0.6963, RF: 0.7137). This suggests that during the stable snow cover period, the models can explain approximately 70% of the snow depth variability. Despite the relatively high R^2 values, the absolute errors (MAE) from December to February were generally above 3–4 cm, particularly in February, where the MAE for both models exceeded 3.4 cm and RMSE approached 5 cm. This indicates that during periods of greater snow depth, the absolute prediction bias of the models also increases. In November and March, relatively higher air temperatures lead to the presence of liquid water within the snowpack, altering its scattering characteristics. Particularly in March, as the snow cover duration increases, snow grain size and snow density rapidly

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

increase. These changes affect the accuracy of AMSR-2 snow depth products, consequently impacting model performance (Wei et al., 2021a).

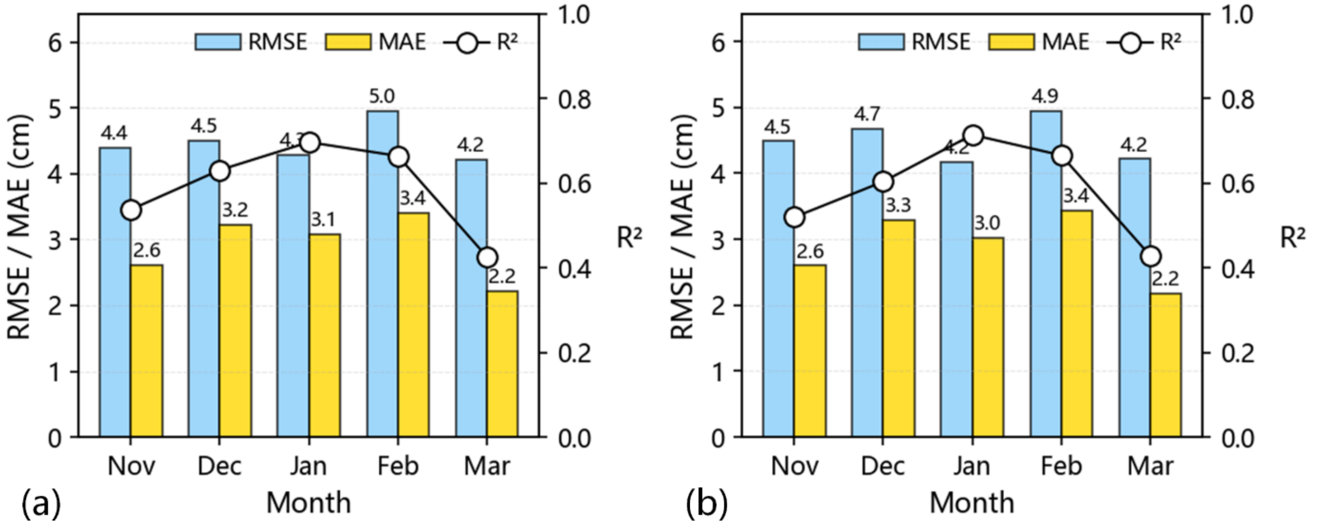


Figure 8: Monthly performance comparison: (a) XGBoost and (b) Random Forest.

Existing studies have shown that passive microwave remote sensing signals exhibit significant differences in response to varying snow depth ranges. Signals are sensitive and stable in shallow snow intervals, while signal saturation effects tend to occur in deep snow intervals, potentially leading to distinct differentiation in model retrieval performance under different snow depth conditions (Chang et al., 1987; Wu, 2025). In the comparison of RMSE performance across different snow depth intervals (Fig.9), the retrieval accuracy of both XGBoost and RF models showed a significant decreasing trend as snow depth increased, with notable interval-specific differences in accuracy between the two models. In the shallow snow interval (≤5 cm), XGBoost achieved better accuracy (RMSE = 3.18 cm) compared to RF (3.30 cm). In the 5–15 cm snow depth interval, the two models exhibited similar performance, with RF achieving slightly better retrieval results (RMSE = 3.57 cm) compared to XGBoost (3.78 cm). In the 15–25 cm snow depth interval, RF maintained slightly superior accuracy (RMSE = 6.10 cm), while XGBoost achieved 6.39 cm. In the deep snow interval (≥25 cm), the accuracy of both models decreased substantially, with XGBoost performing relatively better (RMSE = 12.55 cm) compared to RF (13.37 cm). Overall, both models demonstrated good retrieval accuracy in the ≤15 cm snow depth interval, with RMSE controlled within 4 cm. However, accuracy declined significantly in the ≥15 cm snow depth interval, with particularly pronounced error increases in the ≥25 cm deep snow interval.

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

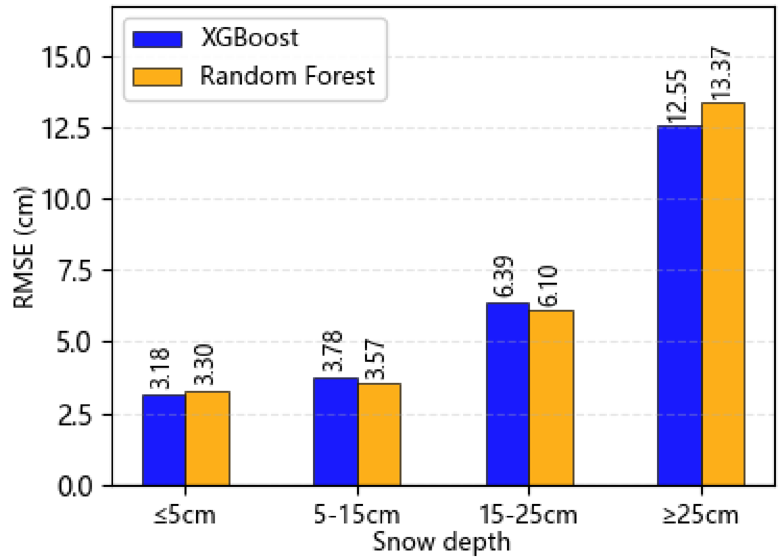
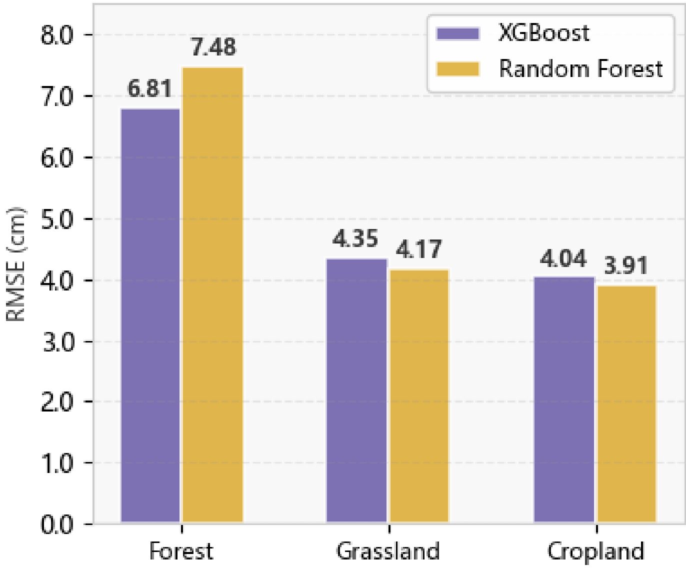


Figure 9: Comparison of retrieval RMSE between XGBoost and RF models under different snow depth intervals.

To investigate the influence of underlying surface on the performance of snow depth retrieval models, we compared the RMSE variations under different land cover types. As shown in Fig. 10, in forest-covered areas, both models exhibited the highest retrieval errors among the three underlying surface types, with XGBoost performing relatively better (RMSE = 6.80 cm) compared to RF (7.47 cm). In grassland areas, RF achieved slightly superior retrieval accuracy (RMSE = 4.17 cm) compared to XGBoost (4.35 cm). In cropland areas, both models achieved the lowest errors among the three underlying surface types, with RF performing slightly better (RMSE = 3.91 cm) compared to XGBoost (4.04 cm). Overall, both models achieved the best snow depth retrieval accuracy in cropland areas, followed by grassland areas, with the poorest performance in forest-covered areas. The underlying reason is that different land cover types exhibit significant differences in their scattering, absorption, and shielding effects on passive microwave remote sensing signals. The complex canopy structure of forests substantially interferes with microwave signal transmission and retrieval. In contrast, areas with homogeneous underlying surfaces and minimal vegetation interference, such as croplands, allow microwave signals to more accurately reflect snow characteristics, thereby facilitating better model retrieval performance.

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026



390 **Figure 10: Comparison of retrieval RMSE between XGBoost and RF models under different land cover types.**

Considering terrain factors as auxiliary features for snow depth retrieval, we analyzed the RMSE variations of the XGBoost model under different terrain conditions in Scenario S4 (see Fig. 11). The statistical results indicate that the model's retrieval accuracy exhibits certain differences across spatial variations in terrain features, with accuracy decreasing progressively as elevation increases, slope steepens, and surface roughness increases. In areas with elevation ≤500 m, the retrieval RMSE was 4.218 cm, significantly lower than the 5.575 cm observed in areas with elevation >500 m. In areas with slope ≤10°, the retrieval RMSE was 4.375 cm, outperforming the 4.88 cm in areas with slope >10°. In areas with surface roughness ≤0.5, the retrieval RMSE was 4.409 cm, also lower than the 4.617 cm in areas with surface roughness >0.5. These findings confirm that complex terrain conditions significantly interfere with snow depth retrieval, while areas with gentle terrain and homogeneous surface conditions are more conducive to improving model retrieval accuracy.

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

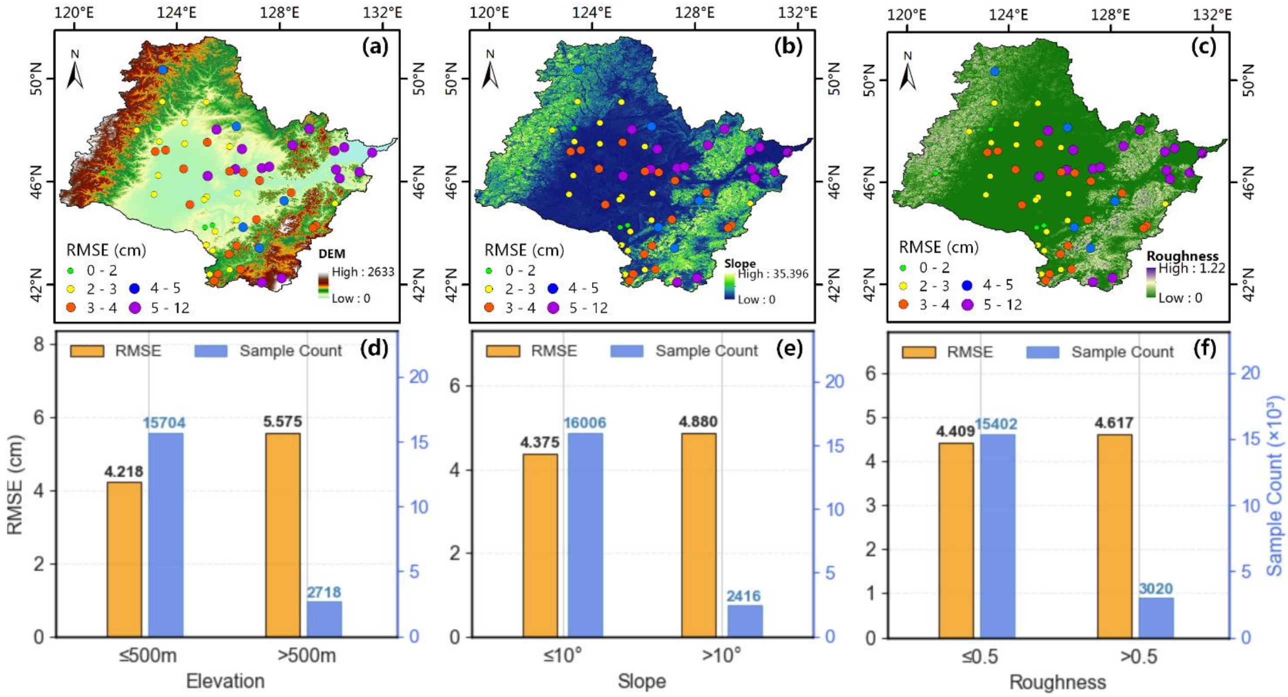


Figure 11: RMSE of snow depth retrieval results from XGBoost under different terrain conditions: Elevation (a) and (d), Slope (b) and (e), Surface Roughness (c) and (f).

Additionally, we also analyzed the distribution of RMSE across different aspects. As shown in Fig.12, RMSE values were higher on west-, south-, and east-facing slopes, while lower RMSE values were observed on other aspects. This distribution pattern may be related to solar radiation conditions: west-, south-, and east-facing slopes receive more solar radiation during daytime, which alters the physical properties of snow such as density and grain size, consequently modifying the microwave scattering characteristics of the snowpack. This reduces the correlation between microwave brightness temperature signals and snow depth, ultimately decreasing snow depth retrieval accuracy on these aspects.

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

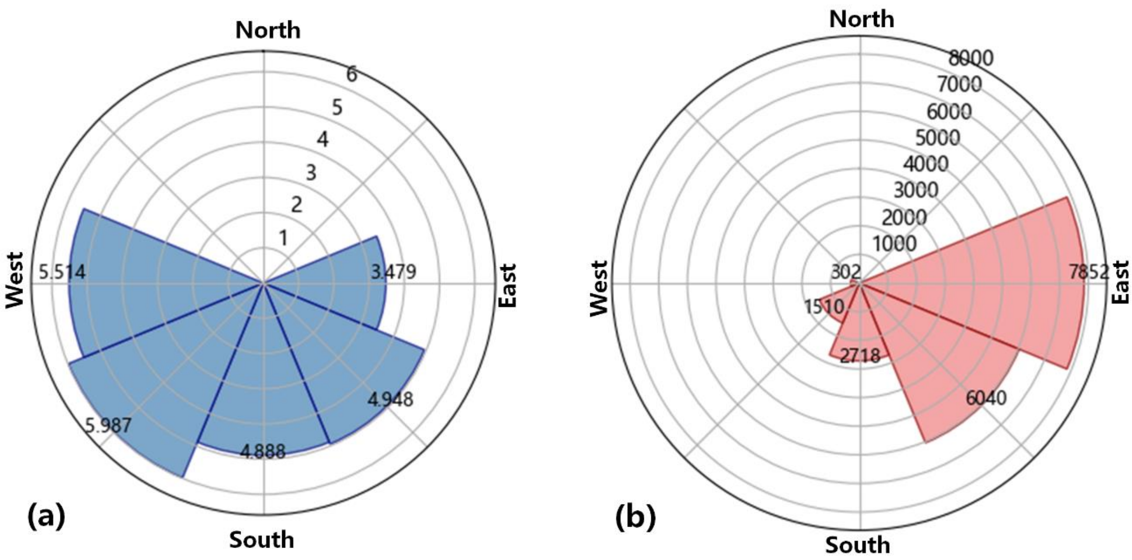


Figure 12: RMSE of snow depth on different aspects, (a) RMSE; (b) Sample size.

5. Discussion

5.1 Model Performance

The XGBoost and RF models constructed in this study exhibited significant performance differences in snow depth retrieval over the Songhua River Basin, with model performance closely related to feature combination scenarios and dataset partitioning methods. From the perspective of feature combinations, under Scenarios S1 and S2, which only utilized single-band microwave brightness temperatures, both models achieved relatively low accuracy on both training and test sets, with test set R^2 values below 0.5 and RMSE exceeding 5.6 cm. This indicates that single microwave brightness temperature features alone are insufficient to fully characterize the complex variation patterns of snow depth. In contrast, under Scenario S3, which incorporated snow physical parameters, and Scenario S4, which integrated comprehensive geographical features, model accuracy improved substantially, with test set R^2 increasing by up to 0.21 and RMSE decreasing by up to 1.27 cm. This demonstrates that the introduction of auxiliary features such as snow physical parameters, terrain factors, and land cover types can effectively compensate for information deficiencies in microwave signals and enhance the model's explanatory power for snow depth.

Comparing the performance of the two machine learning models, the RF model exhibited superior fitting capability on the training set, achieving a training $R^2$of 0.9181 and RMSE of only 2.41 cm under Scenario S4, significantly outperforming the XGBoost model. However, on the temporally independent test set, its accuracy degradation was substantially greater than that of XGBoost, with a training-test R^2 difference of 0.252, compared to only 0.1245 for XGBoost under the same scenario. This indicates that the RF model has a pronounced overfitting tendency and weaker generalization capability for unseen time-series data. This phenomenon is related to the algorithmic characteristics of the RF model. Although its parallel training

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

based on the Bagging strategy can reduce model variance, it tends to overfit local features in the training set when dealing with snow cover sequence data that exhibit strong temporal dependence, and cannot effectively extrapolate to samples beyond the training range. In contrast, XGBoost, through iterative training that minimizes a regularized loss function, effectively controls model complexity while capturing nonlinear feature interactions. Ultimately, under Scenario S4, it achieved optimal performance with a test RMSE of 4.44 cm and R^2 of 0.6738, with a smaller training-test accuracy difference and superior generalization stability, establishing it as the optimal model for snow depth retrieval in the Songhua River Basin.

5.2 Feature Contribution Analysis

The interpretability analysis based on the SHAP method revealed the contribution patterns of different features to snow depth retrieval under the optimal Scenario S4. The core driving features of the XGBoost and RF models exhibited both commonalities and distinct model-specific characteristics, with feature contribution patterns highly consistent with the snow physical properties and spatial distribution patterns of the Songhua River Basin.

The common features of the two models are reflected in the core role of brightness temperature differences and underlying surface temperature factors. The 18.7H-36.5H brightness temperature difference was the most important variable in both the XGBoost (SHAP value 2.12) and RF (SHAP value 0.8245) models. This is because this low-frequency brightness temperature difference can effectively eliminate the interference of surface temperature on microwave signals, highlight the scattering signals from the snow layer, and thus serve as a key indicator for characterizing snow depth. Soil temperature (ST) also ranked among the top ten important features in both models. By influencing snow ablation and accumulation processes, it becomes an important underlying surface factor for snow depth retrieval, confirming the decisive influence of the snow-soil thermal coupling system on snow distribution.

Model specificity is reflected in the differential weighting of features. In the XGBoost model, air temperature (AT, SHAP value 1.95) was second only to the 18.7H-36.5H brightness temperature difference in importance, whereas in the RF model, longitude (SHAP value 1.3809) emerged as the dominant predictor. This disparity stems from the distinct algorithmic characteristics of the two models: XGBoost excels at capturing nonlinear interactions among features, enabling it to effectively identify the direct influence of air temperature on snow ablation processes. Conversely, the RF model exhibits greater sensitivity to spatial distribution features, with the high importance of longitude reflecting the pronounced east-west gradient in snow depth across the Songhua River Basin. This pattern aligns consistently with the basin's topographic and climatic gradients from west to east. In contrast, terrain factors including slope, aspect, and roughness ranked as secondary contributors in both models. Pearson correlation coefficient analysis further confirmed their weak linear relationship with snow depth, indicating that terrain factors primarily influence snow distribution through complex nonlinear mechanisms rather than serving as direct predictors of snow depth variation.

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

5.3 Limitations and future research directions

The validation of the snow depth retrieval model in this study primarily relies on in-situ observations from meteorological stations within the Songhua River basin. However, the spatial distribution of these stations is notably uneven: as shown in Fig. 1, the meteorological stations are mainly concentrated in the central Songnen Plain, while stations in mountainous areas are extremely sparse, with some regions even lacking any meteorological coverage. This directly leads to insufficient representativeness of the model validation results for mountainous areas.

To address the above limitations, future research should focus on the following improvements: (1) conducting intensive ground-based snow depth observations in representative mountainous areas (e.g., Changbai Mountains, Lesser Khingan Range) by deploying temporary snow depth measurement transects or automated snow sensor networks to obtain high-quality independent validation data; (2) incorporating higher spatial resolution remote sensing validation data sources, such as snow depth products retrieved from Sentinel-1 SAR backscatter coefficients or snow cover area products derived from optical remote sensing (e.g., MODIS, Landsat), to indirectly evaluate model performance in regions without station coverage.

6. Conclusions

This study focused on the Songhua River Basin as the study area, integrating multi-source data including AMSR-2 microwave brightness temperature, ERA5-Land meteorological data, topography, and land cover to construct two machine learning models—XGBoost and Random Forest. The study investigated the impacts of different feature combinations, sample sizes, and spatiotemporal conditions on snow depth retrieval performance, and employed the SHAP method to elucidate feature contribution mechanisms.

Feature combinations substantially affect model retrieval performance. The S4 scenario, which integrated brightness temperature differences, snow physical parameters, topographic factors, and land cover data, yielded the optimal results. Compared to scenarios relying solely on single-band brightness temperatures, this configuration improved test set R^2 by up to 0.21 and reduced RMSE by up to 1.27 cm, demonstrating the effectiveness of multi-source data fusion. The XGBoost model exhibited superior generalization stability relative to the Random Forest model. Under the S4 scenario, XGBoost achieved a test RMSE of 4.44 cm and an R^2 of 0.6738, with a smaller gap between training and test performance and no evident overfitting, making it the optimal model for snow depth retrieval in the Songhua River Basin. In contrast, while the RF model showed better fitting performance on the training set, its generalization capability was comparatively weaker. Compared with existing studies, the XGBoost model in this study achieves higher accuracy than traditional AMSR-2-based snow depth products in Northeast China, and is comparable to similar machine learning studies conducted on the Qinghai–Tibet Plateau, though slightly lower than their optimal performance, primarily due to the influence of forest cover and topographic complexity within the basin. The generalizability of the model is limited: it is applicable to mid-latitude monsoon climate basins, and requires retraining and recalibration when applied to regions with significantly different climatic backgrounds or markedly different land cover types.

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

The brightness temperature difference 18.7H-36.5H, air temperature, soil temperature, and longitude emerged as the core features for snow depth retrieval, with their contributions closely aligned with the snow physical properties and spatial distribution patterns of the Songhua River Basin. Topographic and land cover factors primarily influenced retrieval accuracy indirectly through nonlinear mechanisms. Model retrieval accuracy exhibited significant spatiotemporal heterogeneity, with higher accuracy during the stable snow cover period compared to the early melting period, better performance in the ≤15 cm snow depth interval than in deep snow intervals, superior accuracy in cropland areas relative to grassland and forest areas, and enhanced performance in gentle terrain compared to complex terrain. These patterns are fundamentally associated with microwave signal response characteristics and the interference mechanisms of underlying surface and topographic conditions. The multi-source data fusion-based snow depth retrieval framework established in this study offers technical support for the dynamic monitoring of snow cover in the Songhua River Basin. Furthermore, the findings serve as a reference for model selection and feature optimization in snow depth retrieval across analogous cold-region basins. The identified deficiencies in retrieval accuracy during periods of deep snow and snowmelt should be prioritized as key directions for future model enhancement.

Data availability. Satellite passive microwave measurement data and snow depth parameter data are available for download from https://10.5281/zenodo/20076684 (Lu et al.). The daily station snow depth dataset from the China Meteorological Administration (CMA) can be accessed by scientific researchers through the submission of an application (http://data.cma.cn/en). Land use and land cover (CLCD) data are available for download from https://zenodo.org/records/12779975.

Author contributions. TL: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Conceptualization. QW: Methodology, Data curation. LZ: Writing – review & editing, Supervision, Conceptualization. YZ: Software, Formal analysis. LW: Validation, Resources, Conceptualization.

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

Acknowledgements. The authors thank National Meteorological Information Center for providing ground observations.

Financial support. This research was supported by the National Natural Science Foundation of China (Grant No. 42271136), the 2025 Longjiang Project “Young Goose Innovation Team” Program (Grant No. 2025CYLJ01), and the 2025 Fundamental Research Funds for Provincial Universities of Heilongjiang Province (Grant No. 2025-KYYWF-ZR0497).

References

Armstrong, R. L., & Brun, E. (Eds.). (2008). Snow and Climate: Physical Processes, Surface Energy Exchange and Modeling. Cambridge University Press.

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. Nature, 438(7066), 303-309.

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

Bormann, K. J., Brown, R. D., Derksen, C., & Painter, T. H. (2018). Estimating snow-cover trends from space. Nature Climate Change, 8(11), 924-928.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.

Brown, R. D., & Robinson, D. A. (2011). Northern Hemisphere spring snow cover variability and change over 1922–2010 including an assessment of uncertainty. The Cryosphere, 5(1), 219-229.

Chang, A. T., Foster, J. L., & Hall, D. K. (1987). Nimbus-7 SMMR derived global snow cover parameters. Annals of Glaciology, 9, 39-44.

Chen, T., & Guestrin, C. (2016, August). 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.

Derksen, C. (2008). The contribution of AMSR-E 18.7 and 10.7 GHz measurements to improved boreal forest snow water equivalent retrievals. Remote Sensing of Environment, 112(5), 2701-2710.

Dozier, J., Bair, E. H., & Davis, R. E. (2016). Estimating the spatial distribution of snow water equivalent in the world's mountains. Wiley Interdisciplinary Reviews: Water, 3(3), 461-474.

Dunmire, D., Lievens, H., Boeykens, L., & De Lannoy, G. J. (2024). A machine learning approach for estimating snow depth across the European Alps from Sentinel-1 imagery. Remote Sensing of Environment, 314, 114369.

Editorial Committee of Encyclopedia of Rivers and Lakes in China. (2014). Encyclopedia of Rivers and Lakes in China: Volume of Heilongjiang and Liaohe Rivers. China Water & Power Press.

Flanner, M. G., Shell, K. M., Barlage, M., Perovich, D. K., & Tschudi, M. A. (2011). Radiative forcing and albedo feedback from the Northern Hemisphere cryosphere between 1979 and 2008. Nature Geoscience, 4(3), 151-155.

Foster, J. L., Sun, C., Walker, J. P., Kelly, R., Chang, A., Dong, J., & Powell, H. (2005). Quantifying the uncertainty in passive microwave snow water equivalent observations. Remote Sensing of Environment, 94(2), 187-203.

Hallikainen, M., Lemmetyinen, J., & Jiang, L. (2024). Snow depth and water equivalent from passive microwave. In Reference Module in Earth Systems and Environmental Sciences. Elsevier. https://doi.org/10.1016/b978-0-443-13220-9.00036-6

Hyndman, R. J., & Athanasopoulos, G. (2014). Forecasting: Principles and practice. Bowker-Saur. https://doi.org/10.1007/978-1-4612-0211-0_10

Jans, J. F., Beernaert, E., De Breuck, M., Brangers, I., Dunmire, D., De Lannoy, G., & Lievens, H. (2025). Sensitivity of Sentinel-1 C-band SAR backscatter, polarimetry and interferometry to snow accumulation in the Alps. Remote Sensing of Environment, 316, 114477.

Kelly, R. E., Chang, A. T., Tsang, L., & Foster, J. L. (2003). A prototype AMSR-E global snow area and snow depth algorithm. IEEE Transactions on Geoscience and Remote Sensing, 41(2), 230-242.

Li, X., Liang, S., Zhao, K., Wang, J., Che, T., & Li, Z. (2020). Snow cover classification based on climate variables and its distribution characteristics in China. Journal of Glaciology and Geocryology, 42(1), 62-71. https://doi.org/10.7522/j.issn.1000-0240.2019.0050

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. The Cryosphere, 16(1), 159-177.

Lievens, H., Demuzere, M., Marshall, H. P., Reichle, R. H., Brucker, L., Brangers, I., ... & De Lannoy, G. J. (2019). Snow depth variability in the Northern Hemisphere mountains observed from space. Nature Communications, 10(1), 1-12.

López-Moreno, J. I., Fassnacht, S. R., Beguería, S., & Latron, J. B. P. (2011). Variability of snow depth at the plot scale: implications for mean depth estimation and sampling strategies. The Cryosphere, 5(3), 617-629.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems (pp. 4768–4777). Curran Associates Inc.

Mariani, A., Borsotti, J., Livio, F., Villa, G., Metzger, M., & Monti, F. (2026). Influence of snowpack properties and local incidence angle on SAR signal depolarization: a mathematical model for high-resolution snow depth estimation. The Cryosphere, 20, 963-982.

Marks, D., & Dozier, J. (1992). 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(11), 3043-3054.

Meier, W. N., Markus, T., & Comiso, J. C. (2018). AMSR-E/AMSR2 Unified L3 Daily 12.5 km Brightness Temperatures, Sea Ice Concentration, Motion & Snow Depth Polar Grids (AU_SI12, Version 1) [Data Set]. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/RA1MIJOYPK3P

Miller, Z. S., Peitzsch, E. H., Sproles, E. A., Birkeland, K. W., & Palomaki, R. T. (2022). Assessing the seasonal evolution of snow depth spatial variability and scaling in complex mountain terrain. The Cryosphere, 16(12), 4907-4930.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., ... & Thépaut, J. N. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. Earth System Science Data, 13(9), 4349-4383. https://doi.org/10.5194/essd-13-4349-2021

Pulliainen, J., Luojus, K., Derksen, C., Mudryk, L., Lemmetyinen, J., Salminen, M., ... & Norberg, J. (2020). Patterns and trends of Northern Hemisphere snow mass from 1980 to 2018. Nature, 581(7808), 294-298.

Shi, J., & Dozier, J. (2000). Estimation of snow water equivalence using SIR-C/X-SAR, Part I: Inferring snow density and subsurface properties. IEEE Transactions on Geoscience and Remote Sensing, 38(6), 2465-2474.

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. The Cryosphere, 18(1), 451-474.

Sturm, M., Goldstein, M. A., & Parr, C. (2017). Water and life from snow: A trillion dollar science question. Water Resources Research, 53(5), 3534-3544.

https://doi.org/10.5194/egusphere-2026-2646
Preprint. Discussion started: 15 July 2026

Tedesco, M., & Narvekar, P. S. (2010). Assessment of the NASA AMSR-E SWE product. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 3(1), 141-159.

Ulaby, F. T., & Stiles, W. H. (1980). The active and passive microwave response to snow parameters: 2. Water equivalent of dry snow. Journal of Geophysical Research: Oceans, 85(C2), 1045-1049.

Vuyovich, C. M., Jacobs, J. M., & Daly, S. F. (2014). Comparison of passive microwave and modeled estimates of total watershed SWE in the continental United States. Water Resources Research, 50(11), 9088-9102.

Wang, Y., Huang, X., Wang, J., Zhou, M., & Liang, T. (2019). AMSR2 snow depth downscaling algorithm based on a multi-factor approach over Tibetan Plateau, China. Remote Sensing of Environment, 231, 111268.

Wang, Z., Gan, F., Yan, B., et al. (2022). Inversion of snow depth and snow water equivalent based on passive microwave remote sensing and its application progress. Remote Sensing for Natural Resources, 34(3), 1-9.

Wei, P., Zhang, T., Zhou, X., et al. (2021). Reconstruction of snow depth data at moderate spatial resolution (1km) from remotely sensed snow data and multiple optimized environmental factors: A case study over the Qinghai-Tibetan Plateau. Remote Sensing, 13, 1-21.

Wu, M. (2025). Assessing the sensitivity of snow depth retrieval algorithms to inter-sensor brightness temperature differences. Remote Sensing, 17(19), 3355. https://doi.org/10.3390/rs17193355

Xiao, X., Zhang, T., Zhong, X., Shao, W., & Li, X. (2018). Support vector regression snow-depth retrieval algorithm using passive microwave remote sensing data. Remote Sensing of Environment, 210, 48-64.

Yang, J., & Huang, X. (2021). The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019. Earth System Science Data, 13(8), 3907-3925.

Yang, J., Jiang, L., Luojus, K., Pan, J., Lemmetyinen, J., Takala, M., & Wu, S. (2020). Real-Time Snow depth estimation and historical data reconstruction over China based on a random forest machine learning approach. The Cryosphere, 14(6), 1763-1778.

Zhang, T. (2005). Influence of the seasonal snow cover on the ground thermal regime: An overview. Reviews of Geophysics, 43(4).

Zhou, Z., Zhu, L., Zhang, Y., et al. (2024). Downscaling machine learning snow depth inversion on the Qinghai-Xizang Plateau based on FY-3B passive microwave data. Journal of Glaciology and Geocryology, 46(2), 539–554.