https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

Modelling glacier-wide annual mass balance of continental-type glaciers in China using a deep neural network

Lili Wang[1,2], Yaonan Zhang[1,2]
[1]National Cryosphere Desert Data Center, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou, People's Republic of China
[2]University of Chinese Academy of Sciences, Beijing, People's Republic of China
Correspondence to: Yaonan Zhang (yaonan@lzb.ac.cn)

Abstract. Glacier mass balance is crucial for climate and hydrological research. Although data-driven techniques have advanced mass balance estimation, their reliance on comprehensive and reliable training datasets still limits their practical application. In this study, a lightweight feed-forward fully connected neural network (FF-FCNN) was developed to simulate glacier-wide annual mass balance using multi-temporal meteorological variables from ERA5-Land, MODIS-derived summer mean albedo, and topographical attributes from ASTGTM_003 as input features, with 180 glaciological observations from ten continental-type glaciers in China as reference data. To mitigate overfitting in the "small-sample, high-dimensional" scenario, key meteorological variables were selected using the Pearson correlation analysis combined with the Random Forest (RF) algorithm, and several strategies including Gaussian noise injection, L1 regularization, and early stopping were incorporated into the model architecture. Two training dataset construction strategies were evaluated to address temporal inconsistencies in albedo data, and both results demonstrated that the FF-FCNN effectively avoids overfitting and maintains stable and reliable performance. Under the reduced-sample strategy, the FF-FCNN significantly outperformed the Random Forest model (R^2 = 0.82, RMSE = 0.19 m w.e., MAE = 0.15 m w.e.). Spatial and temporal cross-validations further confirmed the robustness and generalization capability of the proposed model. Although the dynamic loss-based weighting strategy enhanced the model's ability to capture pronounced interannual variability in glacier mass balance, reproducing extreme values remains challenging under severely limited sample conditions. Overall, the proposed framework provides a feasible pathway for estimating regional glacier mass balance in high-altitude and cold regions where observations are scarce.

1 Introduction

Glaciers serve as natural reservoirs, playing a crucial role in regulating runoff and providing stable freshwater resources for domestic use, agricultural irrigation, and hydropower generation in downstream regions (Carey et al., 2017). Glacier mass balance is a key indicator of glacier–climate interactions, and understanding its variability is fundamental for evaluating the sustainability of regional water resources under ongoing climatic change. In contrast to geometric indicators such as area or length, mass balance responds directly and immediately to climate forcing (Duan et al., 2009). The glaciological method quantifies accumulation and ablation through repeated measurements of stake exposure heights and snow pit profiles, which

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

are spatially extrapolated to derive glacier-wide mass balance (Kuhn et al., 1999; Thibert et al., 2008). Although this approach yields high-precision measurements, its long-term continuity is constrained by the logistical challenges and financial costs associated with fieldwork in harsh environments. Systematic glacier mass balance monitoring in China began in the late 1950s, yet only a limited number of glaciers possess continuous, long-term records (Xiao et al., 2007). Urumqi Glacier No. 1 provides the most complete record, with observations dating back to 1958, and represents a typical temperate continental glacier included in the World Glacier Monitoring Service (WGMS) network (Li et al., 2011). Several other glaciers—such as Qiyi, Xiao Dongkemadi, Meikuang, Hailuogou, Baishuihe No.1, Zhadang, and Kangwure—also have more than a decade of continuous observations (Su et al., 2022). However, for most glaciers in China, in situ mass balance data remain sparse, short, or discontinuous (Che et al., 2017). This scarcity limits our ability to assess long-term glacier–climate responses, predict future glacier evolution, quantify glacier contributions to basin hydrology, and refine glacier hydrological models.

Advances in remote sensing technology have established both the geodetic method based on differential DEMs (Bash et al., 2018; Rabatel et al., 2016) and the satellite gravimetry method based on temporal variations in Earth's gravitational field (Chen et al., 2007) as important techniques for observing glacier mass balance. Although these approaches provide extensive spatial coverage, they capture only discrete temporal snapshots and are therefore limited in resolving the continuous evolution of glaciers. In parallel, numerical and data-driven modeling approaches have emerged as powerful tools for simulating glacier mass balance and assessing its response to climatic forcing. Temperature-index models based on statistical relationships and physically based energy-balance models are commonly used to simulate time-series glacier mass balance. Temperature-index models take air temperature as the primary driving variable and assume a strong positive linear relationship between glacier melt and temperature (Braithwaite et al., 1995; Zhang and Liu, 2006). However, temperature-based models tend to be overly sensitive to short-term temperature fluctuations and fail to adequately represent energy components weakly related to temperature, thereby exhibiting inherent theoretical limitations. To address these limitations, researchers have incorporated additional variables—including solar radiation, surface albedo, and terrain factors—into traditional temperature-index models (Hock, 2003; Pellicciotti et al., 2005). These refinements substantially enhance the models' ability to reproduce the temporal and spatial variability of glacier mass balance (Liu et al., 2017). Energy-balance models explicitly account for exchanges of mass and energy between glaciers and the atmosphere. They can effectively elucidate the mechanisms driving changes in glacier mass balance and the associated energy–mass conversion processes, thereby providing a more accurate representation of glacier responses to climate change (Arnold, 1996; Hock, 2005; Reijmer and Hock, 2008). However, these models require a large number of input parameters and rely on complex theoretical formulations and model structures, which constrain their applicability at regional scales. Gabbi et al. (2014) evaluated the performance of five melt models (a classical temperature-index model, Hock's temperature-index model, an enhanced temperature-index model, a simplified energy-balance model, and a full energy-balance model) on the Rhonegletscher glacier over a multi-decadal period and found that the energy-balance model performed the worst. This was primarily attributed to the use of forcing data recorded outside the glacier boundary layer, which led to inaccurate heat flux estimations and consequently erroneous melt simulations.

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

The increasing availability of observational datasets, together with advances in data-driven modeling techniques, has opened new avenues for glacier mass balance estimation. Steiner et al. (2005) were the first to use a nonlinear backpropagation neural network (BPN) to reconstruct the mass balance of Grosse Aletschgletscher in Switzerland from 1500 to 2000. Their results demonstrated that nonlinear models based on artificial neural networks (ANN) can estimate glacier mass balance more accurately than stepwise multiple linear regression. Bolibar et al. (2020a) developed the ALpine Parameterized Glacier Model (ALPGM), a parameterized glacier evolution model that includes a surface mass balance (SMB) simulation component based on ANN. The authors emphasized that ANN models offer new opportunities for the glaciology community, as they capture the nonlinear relationships between glaciers and climate systems and provide a promising alternative to conventional SMB models. The SMB component was subsequently employed to reconstruct the annual glacier-wide mass balance of 661 French Alpine glaciers over 1967–2015 (Bolibar et al., 2020b) and to explore their future evolution under multiple climate scenarios (Bolibar et al., 2022). Anilkumar et al. (2023) evaluated the performance of several machine learning (ML) models in simulating point glacier mass balance in the Alps. Similarly, Ren et al. (2024) examined the applicability of various ML approaches for modeling annual glacier-wide mass balance of both maritime and continental glaciers in High Mountain Asia. Both studies yielded promising results, demonstrating the strong potential of data-driven modeling techniques for glacier mass balance estimation. Nevertheless, the limited availability of comprehensive and reliable ground-truth datasets remains a key constraint on the application of ML approaches in glacier modeling. Currently, most glacier mass balance modeling studies in China focus on individual glaciers with available in situ measurements and rely primarily on semi-empirical statistical or physically based methods, whereas ML-based regional modeling remains largely unexplored.

Glacier types differ in thermal and moisture regimes and in their sensitivity to climate forcing. Temperate glaciers, for examples, show greater interannual mass balance variability and larger annual losses compared with continental glaciers. (Fujita, 2008; Su et al., 2015). To ensure consistency in the underlying physical processes when developing data-driven models, it is therefore essential that the study glaciers belong to the same type. In China, continental glaciers account for approximately 77.8% of total glacier area and dominate the long-term observations (Li et al., 2018). Accordingly, ten continental glaciers with publicly available long-term mass balance records were selected to develop ANN models, aimed at assessing the applicability and potential advantages of data-driven methods under extremely small-sample conditions and verifying their feasibility and reliability in simulating glacier mass balance across large-scale regions. Whereas previous studies typically use monthly meteorological data as inputs and reconstructed glacier-wide annual mass balance from remote sensing or numerical modelling as reference data, this study integrates meteorological variables across multiple temporal scales (monthly, seasonal, and annual), together with each glacier's terrain factors, surface albedo, and glacier-wide annual mass balance observed by glaciological method. Feature selection was applied to identify the climate variables significantly influencing mass balance. Two training-dataset construction strategies were implemented to ensure temporal consistency in the inputs. A lightweight deep neural network was then developed to simulate glacier-wide annual mass balance and was benchmarked against a random forest model. Finally, model performance was comprehensively assessed from both temporal and spatial perspectives.

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

2 Materials and methods

2.1 Study sites

The ten glaciers selected for this study are continental-type glaciers with broadly similar accumulation and ablation regimes, making them suitable for use as a unified training dataset for model development and evaluation. These glaciers—Urumqi Glacier No. 1, Haxilegen No. 51, Laohugou No. 12, Qiyi, Shiyi, Ningchanhe No. 1, Meikuang, Xiaodongkemadi, Gurenhekou, and Muztag Ata No. 15—serve as well-studied benchmark glaciers for their respective regions in China. Their long-term monitoring records provide valuable datasets for examining glacier–climate interactions and assessing hydrological responses in arid and semi-arid environments. They are distributed across the eastern Tianshan Mountains, Qilian Mountains, eastern Kunlun Mountains, Tanggula Mountains, Nyainqêntanglha Mountains, and eastern Pamir (Fig. 1). Although regional variations exist in topography and local climatic conditions, all glaciers are situated in arid to semi-arid continental climate zones characterized by low precipitation, low temperatures, slow ice flow, and weak geomorphological activity (Liu et al., 2014). These shared characteristics, together with moderate regional differences, provide favorable conditions for evaluating the model's generalization ability across distinct environmental settings. Their spatial distribution is shown in Fig. 1, while detailed information—including geographic location, area, elevation range (from the Randolph Glacier Inventory, RGI 6.0; RGI Consortium, 2017), observation periods, and sources of mass balance data—is summarized in Table 1.

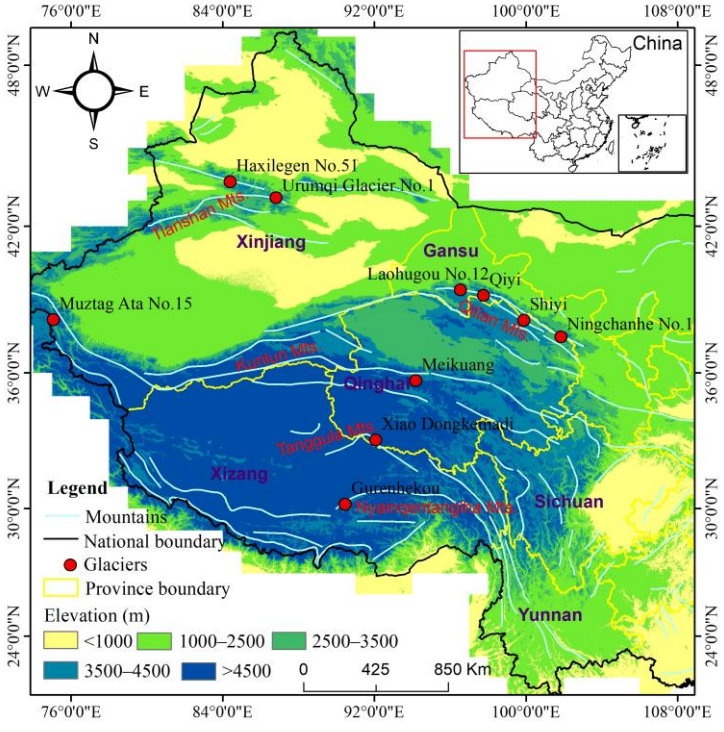


Figure 1 Spatial distribution of the ten continental-type glaciers in China.

Table 1 Overview of the ten continental-type glaciers, along with the monitoring periods and sources of the annual mass balance data.

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

Glacier	Location	Area (Km²)	Elevation range (m)	Monitoring years	Mass balance reference
Urumqi Glacier No. 1	86°49′E, 43°06′N	1.58	3773–4443	1959–2023	NCDC
Haxilegen No.51	84°23′E, 43°43′N	1.10	3495–3901	2000–2003, 2005–2006, 2011	Zhang et al. (2018)
Laohugou No.12	96°32′E, 39°26′N	20.42	4282–5439	1976, 2010–2012, 2014–2018	WGMS, TPDC, Chen et al. (2020)
Qiyi	97°45′E, 39°14′N	2.53	4317–5115	1975–1977, 1984–1988, 2002–2003, 2006–2020	WGMS, Wang et al. (2024)
Shiyi	99°53′E, 38°13′N	0.50	4328–4765	2011–2017	Zhang et al. (2021)
Ningchanhe No. 1	101°50′E, 37°30′N	0.54	4195–4611	2011–2020	WGMS, Pan et al. (2021)
Meikuang	94°11′E, 35°40′N	1.05	4836–5504	1989–1995, 1997–2001, 2012–2017	WGMS, NCDC
Xiaodongkemadi	92°05′E, 33°04′N	1.77	5285–6085	1989–2016	WGMS, NCDC
Gurenhekou	90°27′E, 30°11′N	1.33	5585–6036	2006–2010	WGMS
Muztag Ata No.15	75°40′E, 38°14′N	1.09	5233–5927	2006–2008, 2010–2012	Zhu et al. (2018)

2.2 Data

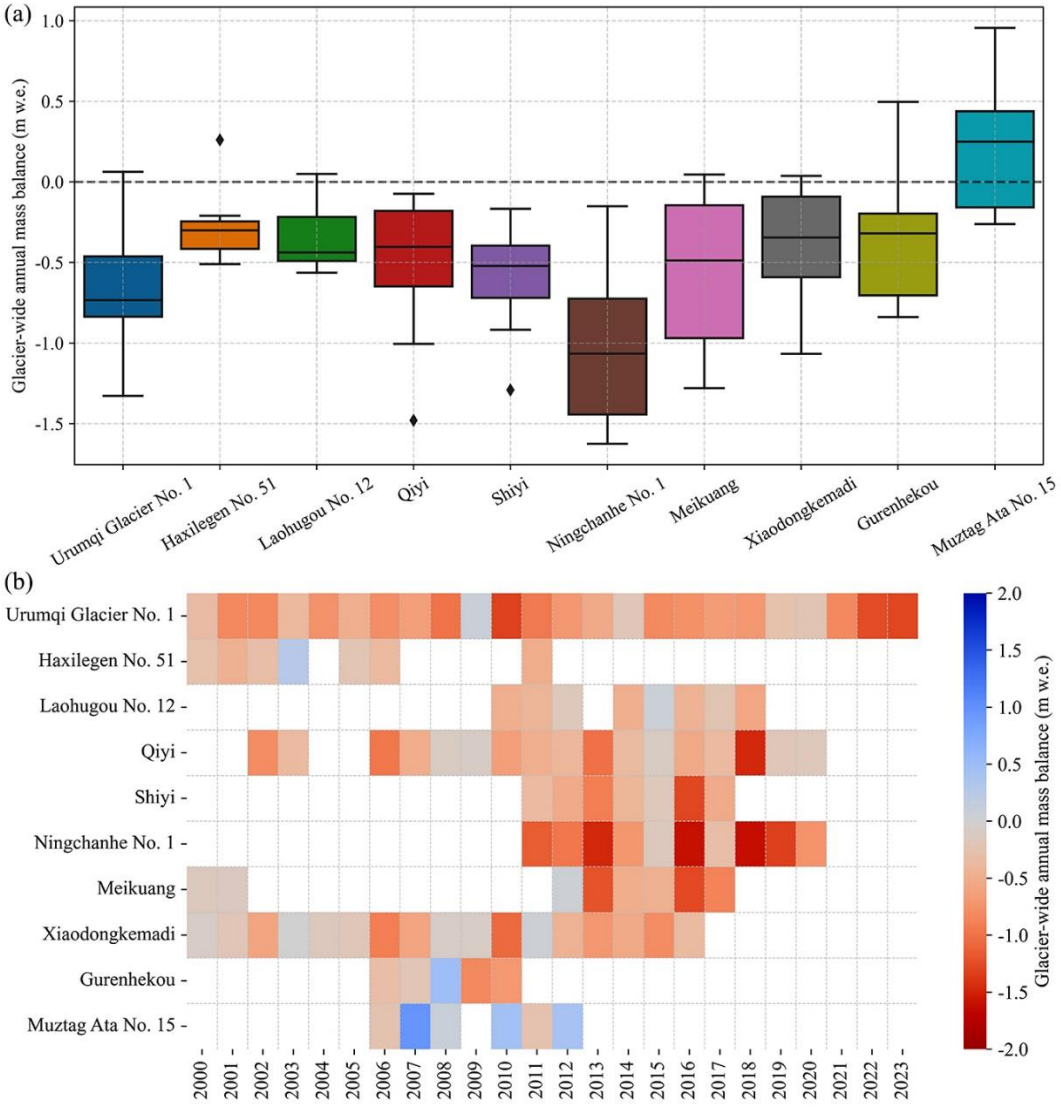
2.2.1 Glacier-wide annual mass balance data

Glacier-wide annual mass balance data for 1959–2023 were compiled from the World Glacier Monitoring Service (WGMS), the National Cryosphere Desert Data Center (NCDC), the National Tibetan Plateau Data Center (TPDC), and relevant academic literature. These datasets originate from conventional glaciological measurements and are reported on an annual basis, with values expressed in meters water equivalent (m w.e.) or millimeters water equivalent (mm w.e.). For consistency and comparability, all records were converted to m w.e. The temporal coverage varies substantially among glaciers, ranging from 5 to 64 years, yielding a total of 180 annual observations. An overview of the mass balance records—including their sources and observation periods—is provided in Table 1.

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

Because most observations were collected after 2000, this study used data from 2000 onward to minimize temporal inconsistencies and analyze the spatiotemporal variability of annual mass balance. Temporal and spatial heterogeneity among the glaciers is illustrated in the boxplots and heatmap (Fig. 2). The median mass balance values of all glaciers except Muztag Ata No. 15 are below 0 m w.e., indicating a predominant ablation trend during the observation period. In contrast, the Muztag Ata No. 15 glacier shows a positively skewed distribution, suggesting a relatively stable or slightly positive mass balance between 2006 and 2012. This pattern is particularly evident in 2007, when a substantial mass surplus of 0.96 m w.e. was recorded. Zhu et al. (2018) indicated that precipitation during the ablation season is the primary factor driving mass balance fluctuations for this glacier. Notably, the Ningchanhe No. 1 glacier exhibits a pronounced negative skew and a wide interquartile range, indicating substantial mass loss accompanied by strong interannual variability. Pan et al. (2021) reported that the mass loss of Ningchanhe No. 1 is more pronounced than that of glaciers in the central and western Qilian Mountains and the Tianshan Mountains, which may be attributed to its smaller size and higher sensitivity to climate change. Cao et al. (2017) further noted that the mean equilibrium-line altitude of this glacier over the past decade (2010–2020) was approximately 4680 m, exceeding the glacier's maximum elevation and thereby accelerating its negative mass balance. Overall, the spatially asynchronous yet regionally consistent pattern of mass loss suggests that continental-type glaciers in different climatic zones respond heterogeneously to recent warming, although a dominant mass-deficit trend persist across the region.

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026



140 **Figure 2** Spatiotemporal characteristics of glacier-wide annual mass balance for the ten continental-type glaciers. **(a)** Distribution of annual mass balance for each glacier. **(b)** Spatial and temporal variations in annual mass balance.

2.2.2 Meteorological data

Meteorological forcing was derived from the ERA5-Land reanalysis datasets produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) (Muñoz Sabater, 2019). ERA5-Land provides hourly estimates of numerous near-surface
145 climate variables from 1950 to the present with a spatial resolution of ~9 km (Muñoz-Sabater et al., 2021). This dataset was chosen for its comparatively high spatial resolution and its proven utility in glacier mass balance simulations (Anilkumar et al., 2023; Ren et al., 2024; Arndt et al., 2023; Draeger et al., 2024). The selection of climatic variables was guided by the

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

physical mechanisms governing glacier ablation and accumulation processes (Réveillet et al., 2017; Gabbi et al., 2014). Nineteen variables were extracted for each glacier's observation period using a nearest-neighbor approach: total precipitation (tp), temperature at 2m (t2m), surface net solar radiation (ssr), surface solar radiation downwards (ssrd), surface net thermal radiation (str), surface thermal radiation downwards (strd), snowfall (sf), forecast albedo (fal), surface sensible heat flux (sshf), surface latent heat flux (slhf), snow depth (sde), snow evaporation (es), snow density (rsn), temperature of snow layer (tsn), surface pressure (sp), potential evaporation (pev), 10 metre U wind component (u10), 10 metre V wind component (v10), and snowmelt (smlt). Wind speed was calculated from the vector composition of the 10m U and V wind components. Subsequently, monthly, ablation-season (May–September), accumulation-seasonal (October–March of the preceding year), and annual averages were computed for each variable. Cumulative Positive Degree Days (CPDD), a key indicator of melt energy (Braithwaite et al., 2000), were calculated from temperature data and included as an additional predictor. In total, 271 climatic variables were generated.

2.2.3 Albedo data

Solar shortwave radiation is the dominant energy source driving glacier ablation (Greuell and Oerlemans, 2005). Albedo, a key glacier surface property, regulates the proportion of incoming solar radiation absorbed by the ice surface and therefore directly influences the glacier mass balance (Hock, 2003). The MOD10A1 and MYD10A1 products are global daily snow cover datasets with a spatial resolution of 500 m, providing information on snow cover, snow albedo, fractional snow cover, and associated quality metrics. These datasets, produced by the National Snow and Ice Data Center (NSIDC) from Terra and Aqua satellite observations, were used to derive glacier surface albedo. Daily albedo products from MOD10A1 and MYD10A1 were selected for the ablation seasons (June 1–August 31) from 2000 to 2023. To enhance the temporal continuity and accuracy of the albedo dataset, MOD and MYD products were first combined via daily averaging. Missing or invalid observations were then filled using a ±2-day moving average interpolation. Finally, the summer mean albedo for each glacier was extracted using glacier boundaries from RGI 6.0 and incorporated as model input feature.

2.2.4 Topographic data

Topographic conditions strongly influence glacier accumulation and ablation through their effects on energy balance, moisture, transport, and surface processes. Previous studies (e.g., Bolibar et al., 2020) have demonstrated that supplementing meteorological inputs with topographic information improves model performance, particularly for glaciers insufficiently characterized by climate data alone. In this study, topographic parameters were extracted from the ASTER Global Digital Elevation Model Version 3 (ASTGTM_003) provided by NASA's Land Processes Distributed Active Archive Center (LP DAAC). Mean elevation, slope, and aspect were derived for each glacier, while glacier longitude, latitude, and area were obtained from RGI 6.0. In total, six geometric and topographic features were included as input variables for the model.

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

2.3 Methods

2.3.1 Feature selection

With only 180 samples but 271 climatic features, the dataset represents a typical "small-sample, high-dimensional" scenario that is particularly susceptible to overfitting, model instability, limited generalization, and reduced statistical reliability. Feature selection is therefore a crucial step in the feature-engineering process. By identifying an optimal subset of predictors, it removes irrelevant or redundant variables while retaining those most informative for the task, thereby improving generalization performance, enhancing computational efficiency, and offering insights into the underlying data-generation processes (Cai et al., 2018). Feature selection first employed the Pearson correlation coefficient (Benesty et al., 2009) to identify variables significantly correlated with glacier-wide annual mass balance ($P < 0.05$), where higher correlation coefficients (R) indicate stronger linear associations. The relative contributions of these variables were subsequently assessed using permutation-based feature importance from the Random Forest (RF) algorithm (Grömping, 2009).

The Pearson correlation coefficients between the climatic variables and annual mass balance are presented in Fig. 3(a). Among all predictors, annual CPDD shows the strongest correlation with mass balance. Moreover, variables significantly correlated with mass balance generally display stronger relationships during the ablation period than during the accumulation period. This pattern reflects the characteristics of continental glaciers in China, whose annual mass balance is predominantly determined by summer ablation, with temperature-driven melt exerting a dominant influence on interannual variability (Li et al., 2003; Liang et al., 2018; Sun et al., 2018). Based on the Pearson correlation analysis, 145 climate variables showing statistically significant relationships ($P < 0.05$) were retained for RF-based feature importance ranking. As illustrated in Fig. 3(b), the top 20 predictors were ranked in descending order of relative contribution. Applying the same feature-selection process to the reduced109-sample dataset yielded 20 meteorological predictors (see Section 2.3.2).

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

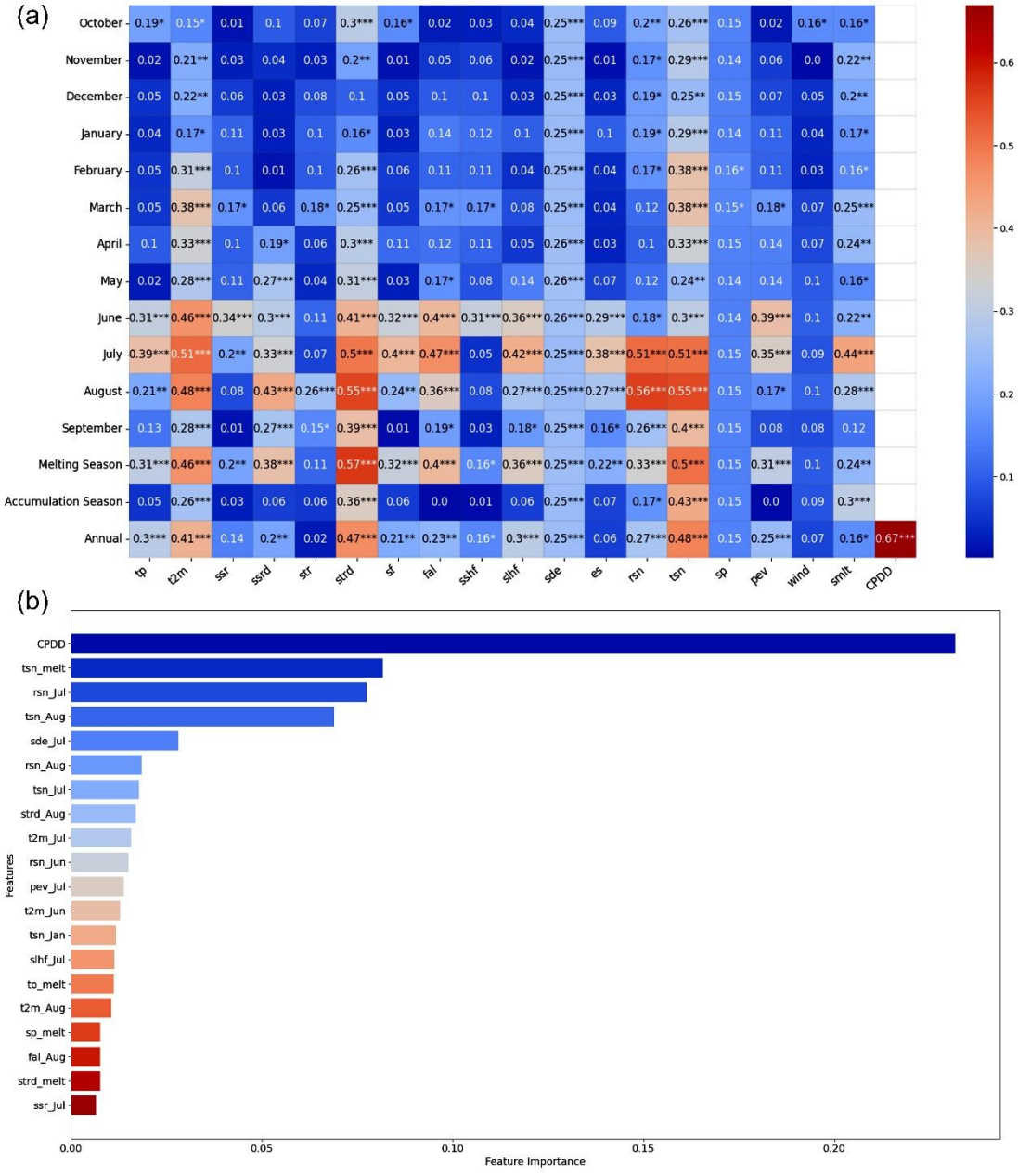


Figure 3 Feature selection results based on Pearson correlation coefficients and RF-derived variable importance. **(a)** Pearson correlation coefficients between climatic variables and glacier-wide annual mass balance. One, two, and three asterisks indicate significance at $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively. **(b)** Relative contributions of the top 20 climatic predictors (from 145 candidates).

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

2.3.2 Construction of the training dataset

Meteorological variables, surface albedo, and terrain factors were used as input features for the model. However, these datasets differ in temporal availability: while meteorological variables cover the full observation periods, albedo data are only available after 2000. To ensure temporal consistency among inputs, two strategies were adopted to construct the training dataset. The first strategy retained all 180 observations, with missing albedo values imputed using mean substitution, and an additional binary indicator variable introduced to explicitly denote missingness. The second strategy excluded glacier observations before 2000, reducing the dataset to 109 samples. As a result, the deep neural network architectures corresponding to the two dataset constructions differ slightly; the structural configurations are described in Section 2.3.3. Static features—including longitude, latitude, glacier area, mean elevation, slope, and aspect—were incorporated into each glacier's time-series feature set as auxiliary predictors. The 20 selected meteorological variables, together with summer mean albedo and the six topographic variables, were combined into a final set of 27 predictors.

2.3.3 Deep neural network construction

Artificial neural networks (ANNs) are nonlinear statistical models inspired by biological neural systems, capable of storing experiential information and applying it to interpret and solve complex problems (Hastie et al., 2009). A neural network is generally characterized by three essential components: 1) network architecture, defined by neuron connectivity and the number of layers; 2) the optimizer, which iteratively updates trainable parameters based on a specified loss function; and 3) activation functions, which introduce nonlinear transformations and enhance the model's expressiveness (Fausett, 2006). A typical ANN comprises multiple interconnected layers—an input layer, several hidden layers, and an output layer—through which raw data are progressively transformed into increasingly abstract representations before generating task-specific predictions (Goodfellow et al., 2016). The feed-forward fully connected neural network (FF-FCNN) developed in this study represents a lightweight deep-learning architecture specifically designed for limited-sample applications while retaining sufficient capacity to model the nonlinear dependencies that control glacier mass balance. In this architecture, each neuron is fully connected to all neurons in adjacent layers, and information propagates strictly in the forward direction—from input to output. This structure enables efficient transmission and transformation of information, promoting the learning of hierarchical features essential for representing climate–glacier interactions.

Hyperparameters tuning was performed using GridSearchCV tool from the scikit-learn library (Pedregosa et al., 2011) combined with cross-validation to identify the optimal configuration of neuron numbers, hidden-layer depth, activation functions, learning rate, and regularization strategies. The final network architecture consists of one input layer, four hidden layers, and one output layer. The input layer contains a number of neurons equal to the total selected predictors, while the hidden layers follow a progressively compressed structure (40, 20, 10, and 5 neurons), which enhances abstraction, stabilizes feature extraction, and reduce overfitting. The output layer contains a single neuron to generate regression-based predictions of annual glacier mass balance. Missing albedo values were imputed using the mean of available observations, and a

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

corresponding binary indicator variable was introduced to denote the presence or absence of albedo data. Both the filled albedo values and the corresponding missing indicator were incorporated into the neural network as input features. To further improve robustness and generalization, multiple optimization strategies were incorporated. Gaussian noise (standard deviation = 0.1) was added to the input layer to emulate data-augmentation effects and enhance resilience to input variability. L1 regularization constrained the effective number of active parameters and mitigated overfitting. Early stopping was employed to halt the training process when validation performance ceased to improve. The combined use of He-normal weight initialization and the LeakyReLU activation function facilitated stable gradient propagation and accelerated convergence. Batch normalization was applied before each activation function to stabilize intermediate feature distributions, thereby improving computational efficiency and training stability. During model training, the optimizer, loss function, and evaluation metrics must be specified. Given the limited and uneven distribution of samples, a loss-based dynamic weighting strategy was introduced to increase the contribution of hard-to-learn samples while suppressing the influence of easy or overrepresented samples. A dynamically weighted root-mean- square error (RMSE) was used as the loss function. Similar difficulty-aware weighting strategies have been shown to substantially improve model robustness under small-sample or imbalanced data conditions (Lin et al., 2017; Cui et al., 2019). The RMSprop optimizer was employed, and model performance was assessed using the coefficient of determination (R^2), the mean absolute error (MAE), and RMSE.

2.3.4 Random Forest model

Random Forest (RF) is an ensemble machine learning method that constructs multiple binary decision trees and aggregates their outputs, originally proposed by Breiman (2001). It has been widely applied to both classification and regression tasks and is capable of evaluating the relative importance of input features (Meng et al., 2021; Reitz et al., 2021; Tramontana et al., 2016). RF models excel at handling complex variable relationships and capturing nonlinear characteristics and they have proven valuable for glacier mass balance prediction (Ren et al., 2024; Anilkumar et al., 2023). In this study, the RF model was implemented using the sklearn package in Python to simulate glacier-wide annual mass balance within a regression framework. GridSearchCV, combined with cross-validation, was employed to systematically explore the hyperparameter grid and identify the optimal combination that maximizes model performance, thereby improving the model's predictive capability.

2.3.5 Model accuracy evaluation

Three cross-validation strategies were implemented in this study—random, leave-one-glacier-out, and leave-one-year-out—to evaluate the model's ability for data reproduction as well as its temporal and spatial prediction performance. Fig. 4(d) displays the cross-validation strategies schematically. Each strategy first split the data into k folds and then iteratively trains the model k times, each time using all folds except one. In the random cross-validation, each sample is randomly assigned to one of the k folds. For the spatial cross-validation, samples are divided into folds according to glacier ID so that all observations from the same glacier are placed in the same fold. For the temporal cross-validation, samples are split into folds by their observation year, with each fold representing a distinct time period. Due to the differing observation periods among glaciers, the test set in

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

each fold does not necessarily include data from all glaciers. The R^2, MAE, and RMSE metrics are used to evaluate model performance and are calculated using the following equations:

$$R^2 = 1 - \sum_{i=1}^{n}(y_i - f_i) / \sum_{i=1}^{n}(y_i - \bar{y}) \quad (1)$$

$$MAE = \sum_{i=1}^{n} |y_i - f_i| / n \quad (2)$$

$$RMSE = \sqrt{\sum_{i=1}^{n}(y_i - f_i)/n} \quad (3)$$

where n represents the total number of observations, y_i denotes the observed values, $\bar{y}$ signifies the average of the observed values, and f_i stands for the corresponding simulated values.

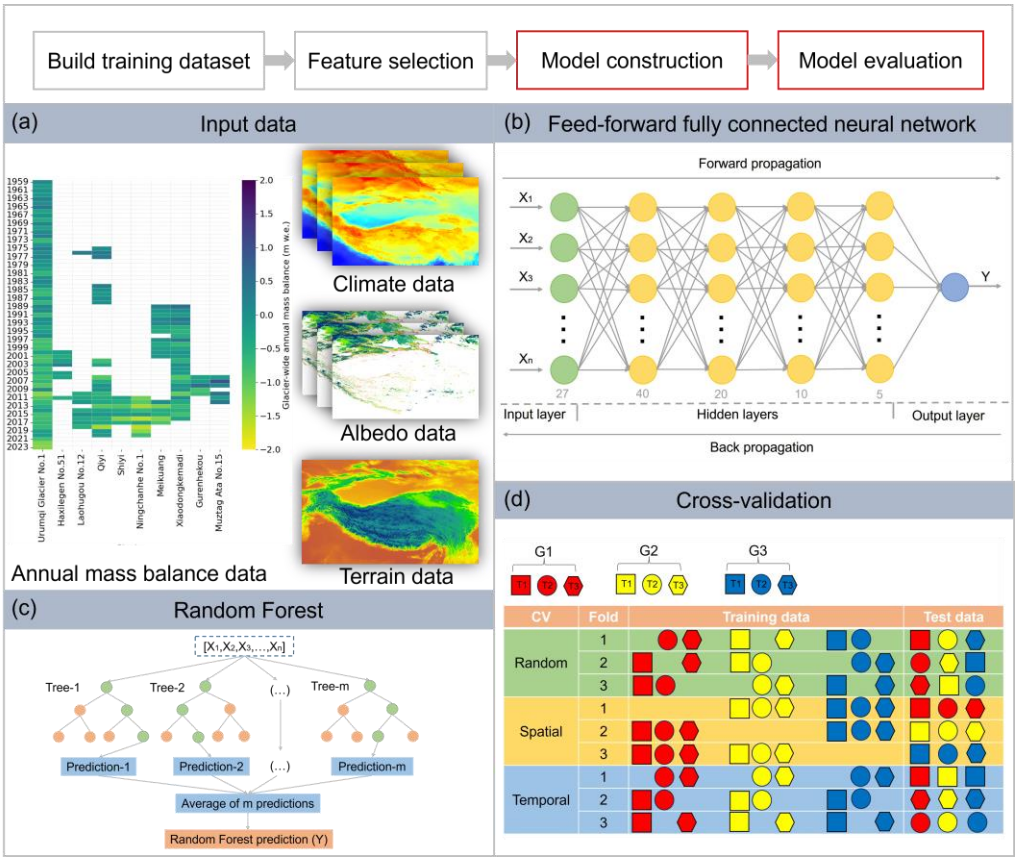


Figure 4 Schematic workflow of this study. **(a)** Input dataset used for RF and FF-FCNN models. **(b)** Architecture of the deep neural network. **(c)** RF model structure. **(d)** Schematic depiction of the model evaluation framework. The figure was recreated and modified from Meyer et al. (2018).

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

3 Results

3.1 Evaluation of training dataset construction strategies

In this study, the models were driven by meteorological, albedo, and topographic data. Because albedo data are available only from 2000 onwards, two strategies were considered for building the training dataset. In the first strategy, all 180 samples were retained by imputing missing albedo values using mean substitution and introducing a corresponding binary indicator variable as an additional input feature in the FF-FCNN. The second strategy excluded samples prior to 2000, reducing the dataset to 109 samples. Given the limited amount of data, it was necessary to evaluate potential overfitting risks to ensure model robustness and reliability. Mean learning curves from 10-fold random cross-validation were examined to compare training and validation losses, and boxplots of validation metrics (RMSE, MAE, and R^2) were used to assess model stability under both strategies. As shown in Fig. 5(a) and (b), the training and validation mean loss curves show a consistent downward trend and eventually converge, with no obvious divergence indicating that the models under both strategies were not subject to severe overfitting. The slight fluctuations observed in the validation loss likely reflect the stochastic nature of mini-batch training and the intrinsic heterogeneity of glacier-wide annual mass balance data. The boxplots of validation metrics in Fig. 5(c) and (d) further demonstrate stable model performance, characterized by low RMSE and MAE values with narrow interquartile ranges (≈0.1–0.3 m w.e.) and consistently high R^2 values (≈0.6 and 0.9). Additionally, the model trained on reduced-sample dataset achieves lower RMSE and MAE and higher R^2 values compared with the model trained on the full 180-sample dataset, indicating that the reduced-sample strategy yields better predictive performance. This conclusion is further supported by the test scatterplots (Fig. 6(a) and (b)) and the corresponding mean cross-validation metrics, both of which clearly show improved agreement between modeled and observed values. These findings suggest that masking missing albedo values does not fully eliminate the negative effects of incomplete features. Instead, it increases data noise and introduces distributional inconsistencies, ultimately degrading model performance. In contrast, restricting the training dataset to samples with complete and homogeneous feature sets yields a cleaner input, enabling the FF-FCNN model to learn more robust and generalizable relationships. Therefore, all subsequent model training was conducted on the reduced-sample dataset.

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

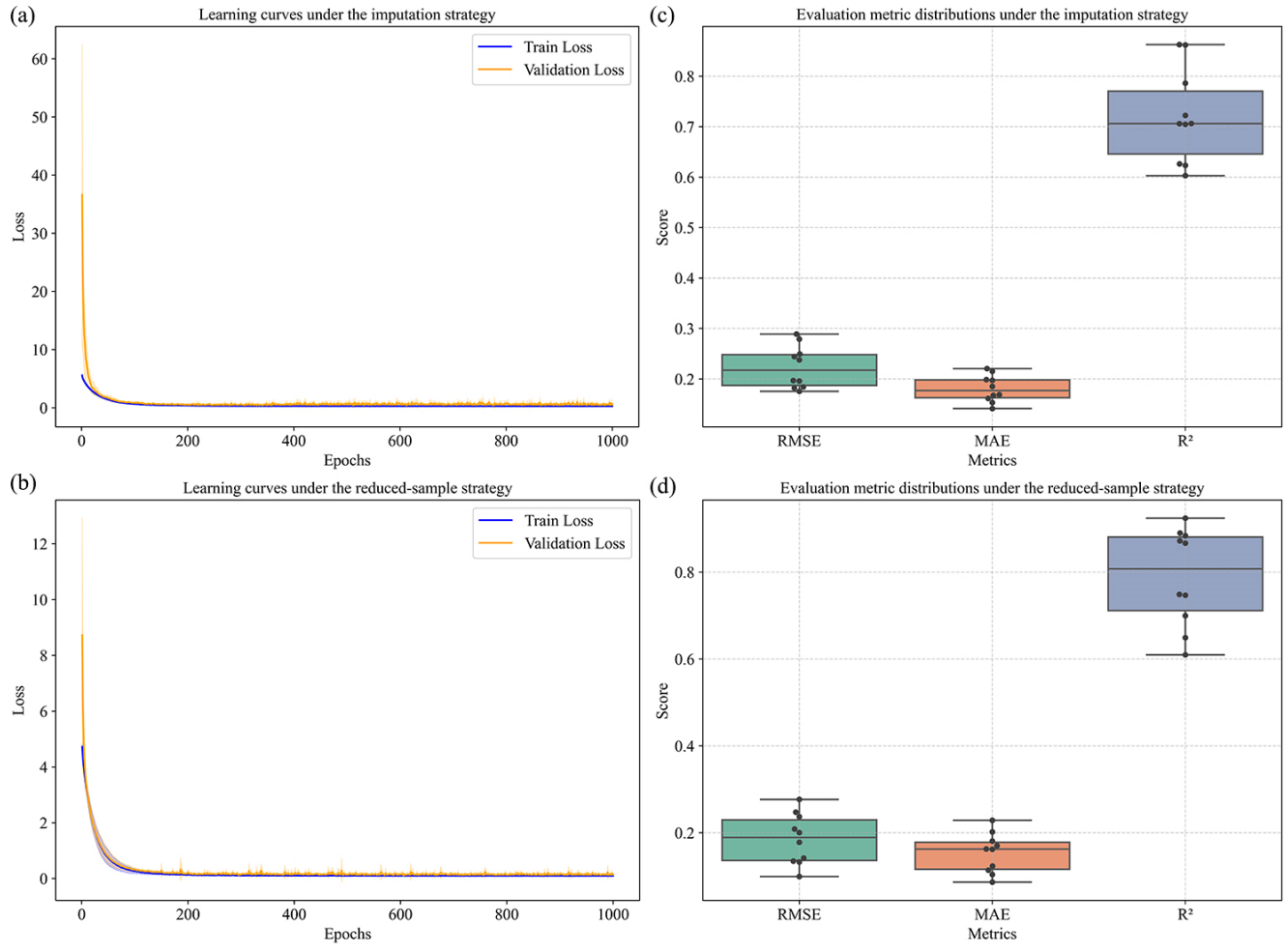
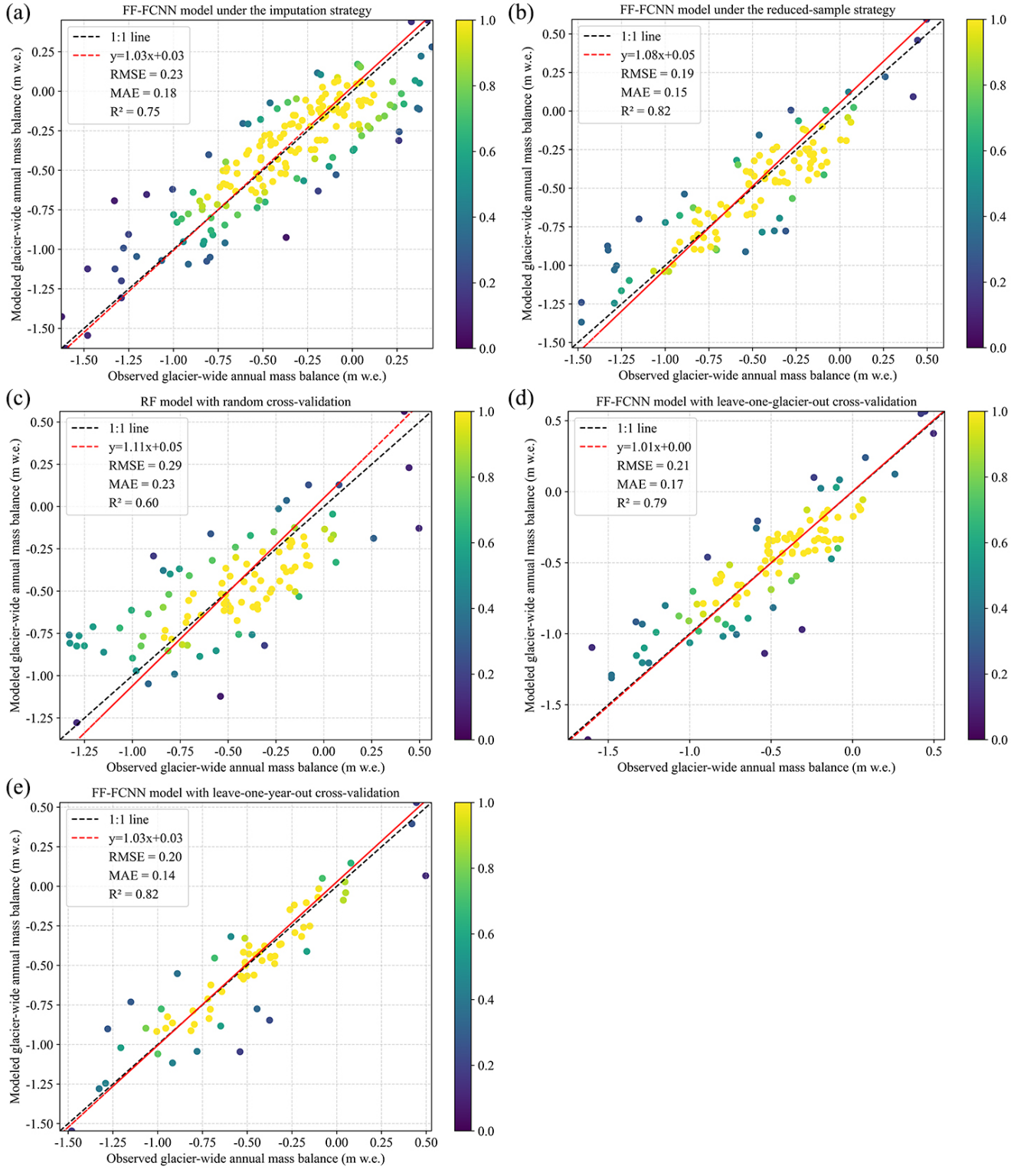


Figure 5 Assessment of overfitting risks and comparison of model performance under the two dataset construction strategies. **(a)–(b)** Training and validation mean loss curves with 10-fold random cross-validation. **(c)–(d)** Distribution of validation metrics.

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026



https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

Figure 6 Evaluation of simulated versus observed annual mass balance in the test dataset. **(a)–(b)** correspond to different training dataset construction strategies. **(b)–(c)** compare FF-FCNN and RF model performance. **(d)–(e)** show FF-FCNN model evaluations from spatial and temporal perspectives.

3.2 Performance comparison of RF and FF-FCNN models

To validate the effectiveness and advantages of the FF-FCNN model in glacier-wide annual mass balance simulation, an RF model was also implemented for comparation. As a widely recognized machine-learning algorithm, the RF model possesses strong nonlinear fitting capabilities and robustness, making it well-suited for small-to-medium-sized datasets and an appropriated benchmark for this study. Both models were trained and validated using the same dataset under a 10-fold random cross-validation scheme. The FF-FCNN model (Fig. 6(b)) achieved a mean RMSE of 0.19 m w.e., a mean MAE of 0.15 m w.e., and a mean R^2 of 0.82. In contrast, the RF model (Fig. 6(c)) yielded a mean RMSE of 0.29 m w.e., a mean MAE of 0.23 m w.e., and a mean R^2 of 0.60. These results indicate that although both models capture the overall variability of glacier-wide annual mass balance, the FF-FCNN consistently outperforms the RF model in both predictive accuracy and explanatory capability. This comparison highlights the robustness of the FF-FCNN model and underscores the advantages of deep neural networks for glacier mass balance modeling, even under limited data conditions.

3.3 Evaluation of the FF-FCNN model from spatial and temporal perspectives

The results of the 10-fold random cross-validation indicate that the FF-FCNN model achieved both high predictive accuracy and stable performance across all folds, confirming its overall reliability. To further assess the model's robustness and generalization capability, its performance was additionally evaluated from spatial and temporal perspectives. As shown in Fig. 6(d) and (e), the FF-FCNN model achieved mean R^2 values of 0.79 and 0.82 for leave-one-glacier-out and leave-one-year-out cross-validations, respectively. Corresponding mean MAE values were 0.17 and 0.14 m w.e., while the mean RMSE values were 0.21 and 0.20 m w.e. These results demonstrate that the FF-FCNN model maintains stable and reliable predictive performance under different validation schemes. A more detailed analysis at the glacier-specific and year-specific levels is presented in the following subsections.

3.3.1 Performance across different glaciers

As shown in Fig. 7, the leave-one-glacier-out cross-validation results reveal distinct spatial heterogeneity in model performance across the ten continental-type glaciers. Muztag Ata No. 15 and Ningchanhe No. 1 exhibit comparatively low predictive performance, as indicated by their higher RMSE and MAE values and notably low R^2. Both glaciers contain mass-balance outliers—an exceptionally positive year in Muztag Ata No. 15 (2007) and extreme negative years in Ningchanhe No. 1 (2016 and 2018) —leading the FF-FCNN model to systematically underestimate these outliers due to their limited representation in the training dataset. Although the introduction of a dynamic loss-based weighting strategy enhances the model's responsiveness to extreme values, the improvement remains limited under extremely small sample conditions.

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

Nevertheless, for most glaciers, the FF-FCNN demonstrates good spatial generalizability and robustness with favorable predictive performance (R^2 range: 0.64–0.94, MAE range: 0.07–0.22 m w.e., RMSE range: 0.09–0.25 m w.e.). The variation in model performance across glaciers can be partly attributed to differences in the size of the test datasets.

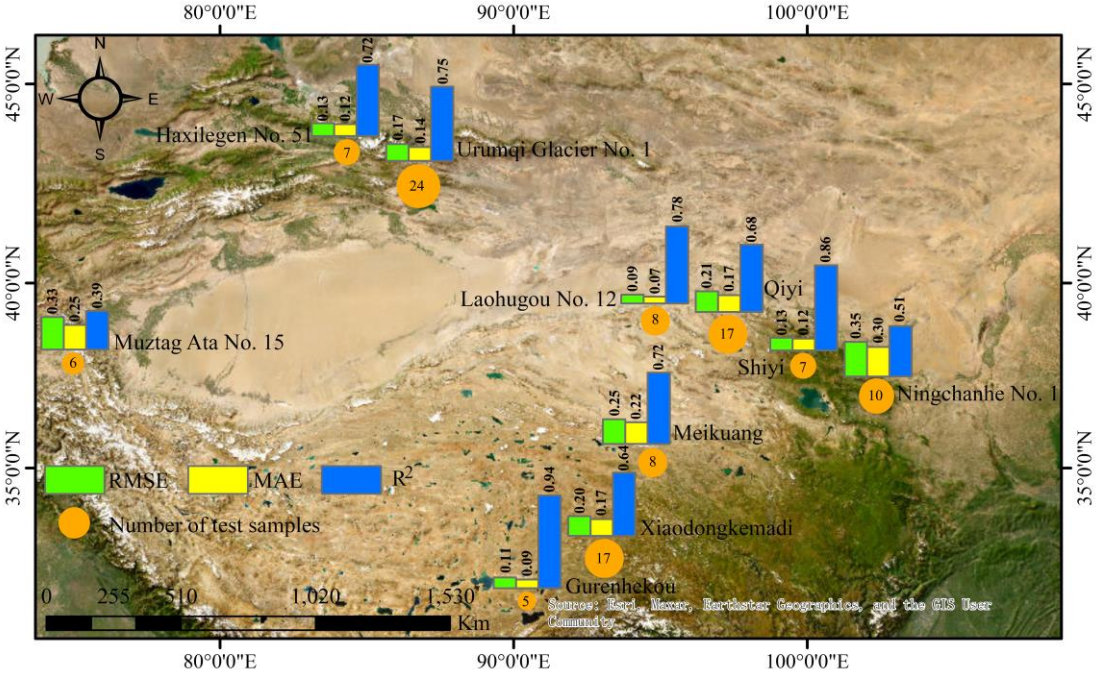


Figure 7 Accuracy evaluation of the FF-FCNN model across different glaciers. Sources: Esri, Maxar, Earthstar Geographics, and the GIS User Community | Powered by Esri.

3.3.2 Performance across different years

To evaluate the model's temporal performance, prediction accuracy was assessed over time for the test dataset. Due to discrepancies in observation periods among glaciers, the distribution of observed mass balance between 2000 and 2023 was highly uneven. To ensure a representative and reliable evaluation, only 11 years with at least five glacier records were included in the test dataset. As illustrated in Fig. 8, the model exhibited excellent generalization capability in seven testing years (R^2 range: 0.74–0.95, MAE range: 0.07–0.14 m w.e., RMSE range: 0.08–0.21 m w.e.). Hydrological years with relatively power performance, characterized by lower R^2 or higher RMSE and MAE, were 2007, 2011, 2016, and 2017. The reduced predictive performance in 2007 and 2016 can be partly attributed to the presence of extreme values, a pattern that is also evident in the spatial cross-validation results presented in Section 3.3.1. To further diagnose the underlying causes of performance degradation, a binary classifier (RandomForestClassifier) was trained to discriminate between samples from the training and validation periods using only the input features. Its performance was quantified using the area under the receiver operating characteristic curve (AUC). An AUC value close to 0.5 indicates similar feature distributions between the two periods, whereas higher values signify increasingly pronounced distribution shifts. Table 2 summarizes the diagnostic analysis for years with

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

degraded performance under temporal cross-validation. In 2007, prediction errors are primarily driven by extreme mass-balance values rather than distribution shift, as indicated by relatively high RMSE (0.32 m w.e.) and MAE (0.24 m w.e.) together with a low AUC (0.411). In 2011, the reduced performance is more likely attributable to weak signal separability and limited feature representativeness. In 2016, extreme values and a pronounced distribution shift jointly contribute to accuracy loss, reflacted by a low R^2 (0.36) and a high AUC (0.919). In 2017, strong distribution shift (AUC = 0.871) dominates the performance degradation despite the absence of extreme values. These results highlight the multifaceted origins of temporal performance degradation under non-stationary conditions.

Table 2 Performance degradation diagnosis for specific years in temporal cross-validation, including accuracy metrics, identified issues, distribution shift quantified by AUC, and dominant error sources.

Year	Accuracy metrics (R^2/MAE/RMSE)	Identified issues	Distribution shift (AUC)	Dominant error sources
2007	0.71/0.24/0.32	Extreme positive values	0.411	Extreme value
2011	0.55/0.17/0.24	No extreme values	0.403	No clear distribution shift
2016	0.36/0.31/0.36	Extreme negative values	0.919	Extreme value +Distribution shift
2017	0.38/0.14/0.18	No extreme values	0.871	Significant Distribution shift

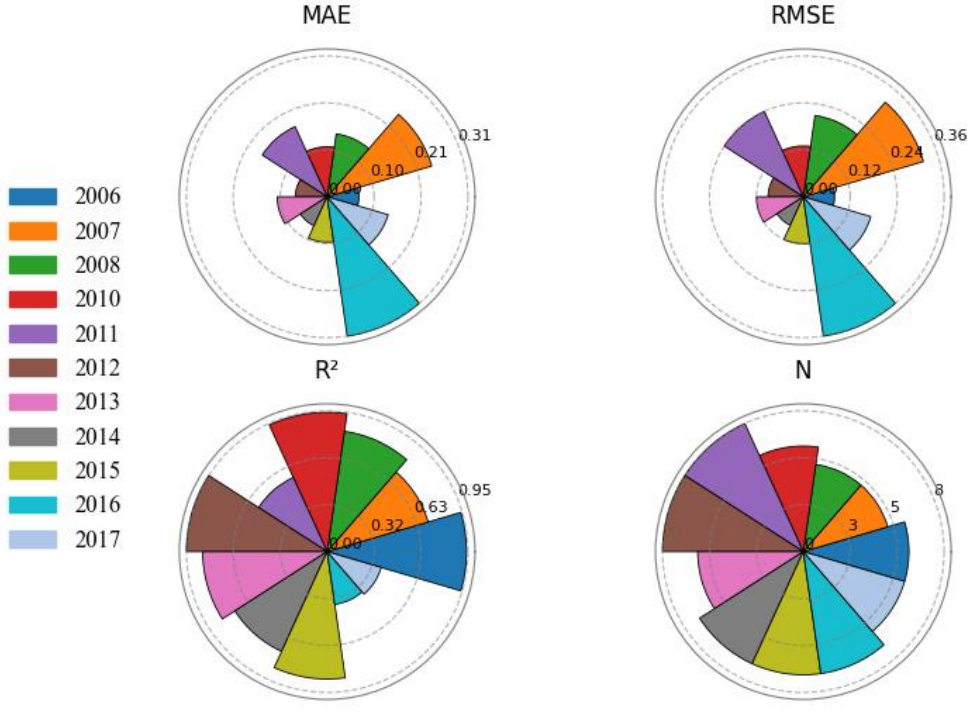


Figure 8 Accuracy evaluation of the FF-FCNN model across different years.

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

4 Discussion

4.1 Influence of topography and surface albedo

In this study, meteorological variables at monthly, seasonal, and annual scales were obtained from the ERA5-Land dataset and used as the primary driving inputs for the model. Pearson correlation analysis and the RF algorithm were applied to identify the most influential and non-redundant features (Fig. 2). Terrain and surface albedo data were incorporated during the training dataset construction stage due to their critical roles in modulating glacier accumulation and ablation. Further analysis was conducted to examine the influence of topography and surface albedo on model performance. Three feature combination schemes were tested within the FF-FCNN framework, using identical training and test samples but different sets of input variables: (1) only meteorological features from ERA5-Land, (2) meteorology combined with topographical features, and (3) meteorology combined with surface albedo features. The evaluation results for the test samples are presented in Fig. 9.

Compared with the model trained solely on meteorological features, incorporating terrain factors led to a slight improvement in overall simulation accuracy, with R^2 increasing by 0.02 and both RMSE and MAE decreasing by 0.01 m w.e. Topographical data provide essential contextual information for differentiating glacier-specific responses to climatic forcing and enable the model to better capture the coupling between topography and local climate. In contrast, incorporating surface albedo data resulted in a substantial improvement in overall simulation performance, with R^2 increasing by 0.05, RMSE and MAE decreasing by 0.02 m w.e. Albedo provides direct radiative feedback for glacier mass balance modeling and effectively compensates for the omission of glacier surface process information inherent in meteorology-only approaches. The combined inclusion of climate forcings, terrain factors, and surface radiative conditions enabled the model to achieve optimal predictive performance. These results indicate that while meteorological variables serve as the primary drivers of glacier-wide annual mass balance, the addition of topographical and surface albedo data provides complementary constraints that help distinguish glacier-specific responses that cannot be fully captured by climate data alone.

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

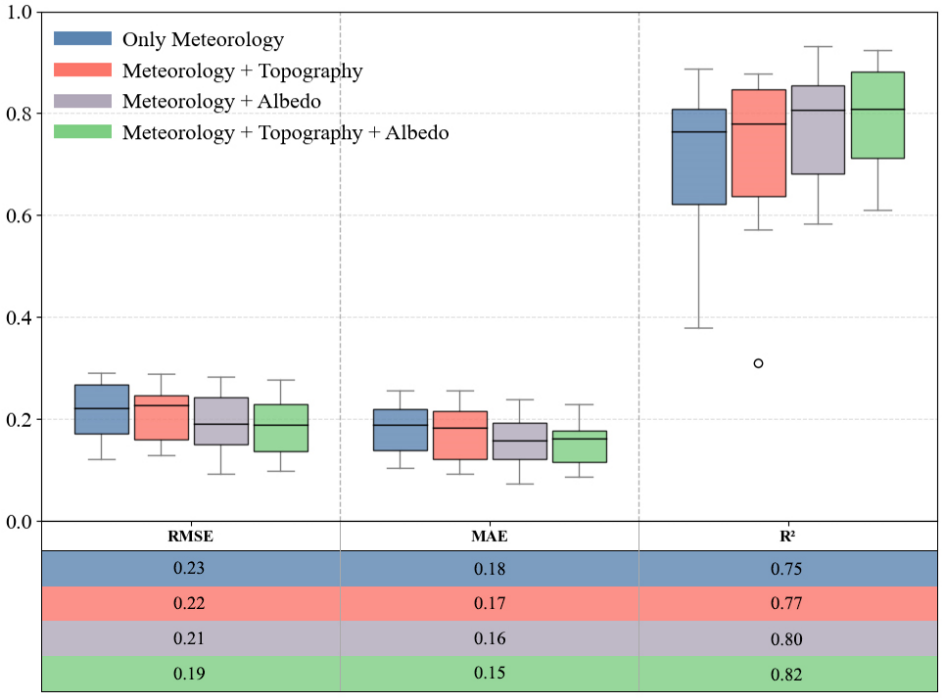


Figure 9 Comparison of FF-FCNN model performance with different input features.

4.2 Uncertainty of reanalysis meteorological data

The ERA5-Land reanalysis product with its relatively high spatiotemporal resolution and comprehensive meteorological variables, has been widely applied in studies such as glacier mass balance modeling (Arndt et al., 2023; Anilkumar et al., 2023), snow cover monitoring (Blau et al., 2024; Faal et al., 2025), and runoff simulation (Li et al., 2024). Mihalevich et al. (2022) demonstrated that spatially coarsened ERA5-Land data can still support reliable simulations of large-scale physical processes, owing to its ability to represent the spatial variability of meteorological fields over extensive regions. However, the use of reanalysis meteorological inputs may introduce biases, particularly in regions with sparse or absent ground-based observations (Zandler et al., 2019; Guidicelli et al., 2023). Wu et al. (2023) reported that while ERA5-Land effectively captures the spatial and temporal patterns of precipitation across the Tibetan Plateau, it substantially overestimates both precipitation amount and frequency. Similarly, Zhao and He (2022) showed that ERA5-Land successfully reproduces temperature trends in the Qilian Mountains but systematically overestimates observed temperatures. These biases can propagate through the modeling framework, thereby affecting the reliability of simulations.

Moreover, the relatively coarse spatial resolution of ERA5-Land tends to smooth complex terrain and weakens elevation-dependent climatic gradients, potentially underrepresenting local atmospheric variability in mountainous regions (Muñoz-Sabater et al., 2021). Approaches such as statistical or dynamical downscaling, lapse-rate corrections or data fusion with high-resolution DEMs have been applied to improve the spatial representation of climate inputs (Immerzeel et al., 2014; Sebbar et

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

al., 2023; Wang et al., 2021). However, most studies have primarily focused on precipitation and temperature, while substantial challenges remain in bias-correcting other meteorological variables that influence glacier mass balance due to the scarcity of observations. Consequently, the systematic biases and coarse spatial resolution of reanalysis meteorological datasets inevitably propagate into the uncertainties of most large-scale modeling efforts, influencing both the accuracy and generalization of simulation results.

4.3 Limitations and prospects of the study

Glacier mass balance observations remain scarce due to the harsh and inaccessible conditions of mountainous regions. Although the FF-FCNN model demonstrates generally satisfactory predictive performance, the limited sample size restricts its ability to fully capture the variability of glacier mass balance, particularly under extreme accumulation or ablation conditions. The model performs less accurately for glaciers with extreme values, such as Muztagh Ata No. 15 and Ningchanhe No. 1, because such patterns are underrepresented in the training dataset, limiting effective generalization. Although the dynamic loss-based weighting strategy partially alleviates the impact of large interannual variability, pronounced volatility combined with extremely limited sample sizes still limits the model's ability to accurately reproduce extreme mass-balance values. To further enhance model robustness, future work should prioritize enlarging the dataset and exploring approaches such as transfer learning and advanced data augmentation.

Although nearly all physical drivers of glacier mass balance were obtained from the ERA5-Land dataset, the correlation analysis (Fig. 2(a)) showed that only a few variables exhibited significant correlations with glacier-wide annual mass balance. While the selected continental-type glaciers share similar ablation and accumulation mechanisms, individual differences in topography, glacier geometry, and debris-cover extent result in heterogeneous responses to meteorological forcing, reducing overall correlations. Moreover, as discussed in Section 4.2, the relatively coarse spatial resolution and potential systematic biases of ERA5-Land limit the representativeness of the driving factors for glacier mass balance. Future studies should therefore focus on spatial downscaling and bias correction of reanalysis meteorological data to enhance the accuracy and reliability of glacier mass balance modeling.

Compared with previous machine learning approaches for glacier mass balance, which were primarily temperature-index models based only on temperature and precipitation (Bolibar et al., 2020; Ren et al., 2024), the FF-FCNN model was trained with variables typically used in more sophisticated energy-mass balance models, including radiation, turbulent fluxes, and albedo. This approach significantly improves predictive performance and provides a more comprehensive representation of the underlying physical processes. However, despite the strong nonlinear representation capabilities of deep learning models, their inherent "black box" nature poses challenges for interpreting physical mechanisms. Recently, hybrid approaches combining physical modeling with deep learning have been increasingly applied in geoscientific research (Fuchs et al., 2023; Steidl et al., 2025; Teufel et al., 2023;). These approaches can be implemented in two main ways: (1) embedding physical regularization constraints directly into the loss function of the deep learning model, or (2) running numerical simulations with the deep learning model and subsequently assimilating or calibrating the results using a physical model. Such strategies not

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

only enhance predictive accuracy but also ensure that model outputs remain physically consistent. Future research could incorporate physical processes such as energy balance components and snow accumulation and melt processes into the FF-FCNN framework, establishing a hybrid modeling approach that integrates data-driven and physically-based methods.

5 Conclusions

This study developed a lightweight feed-forward fully connected neural network (FF-FCNN) trained on limited glacier-wide mass balance observations to simulate annual glacier mass balance. Annual, monthly, and seasonal meteorological variables from ERA5-Land, summer mean albedo from MODIS, and terrain factors from ASTGTM_003 and RGI 6.0 were integrated as model inputs, with 180 glacier-wide annual mass balance observations from ten continental-type glaciers in China serving as reference data. Twenty climate variables were identified using Pearson correlation analysis combined with the RF model. Among these, annual CPDD and melt-season features emerged as the dominant controls on annual mass balance. Two training dataset construction strategies were adopted to ensure temporal consistency among all input features: (1) retaining all 180 samples by imputing missing albedo values using mean substitution and introducing a corresponding binary missingness indicator, and (2) using only the 109 samples with available albedo data after 2000. The final set of 27 predictors—20 meteorological variables, summer mean albedo, and six terrain attributes (longitude, latitude, area, mean altitude, slope, and aspect)—was used for model training. The FF-FCNN model exhibited no signs of overfitting and maintained stable performance under both strategies, although the 109-sample strategy achieved higher predictive accuracy. The RF model and FF-FCNN were trained on the same dataset to assess performance differences under the extremely limited sample conditions. Contrary to prevailing views that traditional machine learning models outperform deep learning under small-sample conditions, the FF-FCNN model demonstrated superior performance in this study. Spatial and temporal generalization capabilities were further evaluated using leave-one-glacier-out and leave-one-year-out cross-validation schemes. The results indicate that the model achieves satisfactory generalization performance for most glaciers and years. However, under conditions of extremely limited samples and the presence of extreme mass-balance values, the model exhibits reduced generalization capability for Muztag Ata No. 15 and Ningchanhe No. 1 glaciers. Moreover, temporal cross-validation reveals that, beyond the influence of extreme values, distribution shift in the input features also contributes to a degradation in predictive accuracy.

Data availability

Glacier mass balance observations used in this study were obtained from publicly available data centers (e.g., the World Glacier Monitoring Service, the National Cryosphere Desert Data Center, the National Tibetan Plateau Data Center). These datasets are available directly from the respective repositories under their access policies and are therefore not redistributed here. ERA5-Land reanalysis data are available from the Copernicus Climate Data Store (CDS). MODIS snow products (MOD10A1 and MYD10A1) are available from NASA LP DAAC. The processed datasets generated during this study, including glacier-wide

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

albedo time series, compiled meteorological variables and terrain attributes are publicly available at Zenodo (DOI: 10.5281/zenodo.18676994).

Author contributions

Lili Wang developed the deep learning and machine learning models, performed the analyses, and wrote the manuscript. Yaonan Zhang conducted the glaciological analysis and contributed data and instrumentation. All authors discussed the results and contributed to the development and revision of the manuscript.

Competing interests

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

Disclaimer

Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. The authors bear the ultimate responsibility for providing appropriate place names. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Acknowledgements

We would like to thank European Centre for Medium-Range Weather Forecasts (ECMWF) for providing the ERA5-Land data, National Aeronautics and Space Administration (NASA) for providing the MODIS data, National Cryosphere Desert Data Center (NCDC), National Tibetan Plateau Data Center (TPDC), and World Glacier Monitoring Service (WGMS) for providing in situ glacier mass balance data, NASA's Land Processes Distributed Active Archive Center (LP DAAC) for providing ASTGTM_003 data, and WGMS for providing RGI 6.0 data. We also thank both anonymous reviewers and the editor for their constructive comments that improve the quality of this paper.

Financial support

This work was supported by the Research of Key technologies and applications of data mining on cryosphere big data of National Key R&D Program of China [grant no 2022YFF0711704], the Research on theoretical system and application strategy of data engineering discipline of Information consulting project of CAS [grant no CAS-WX2023ZX02-01], the National Cryosphere Desert Data Center [grant no 2020000339].

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

References

Anilkumar, R., Bharti, R., Chutia, D., & Aggarwal, S. P. (2023). Modelling point mass balance for the glaciers of the Central European Alps using machine learning techniques. The Cryosphere, 17(7), 2811-2828.

Arnold, N. S., Willis, I. C., Sharp, M. J., Richards, K. S., & Lawson, W. J. (1996). A distributed surface energy-balance model for a small valley glacier. I. Development and testing for Haut Glacier d'Arolla, Valais, Switzerland. Journal of Glaciology, 42(140), 77-89.

Arndt, A., & Schneider, C. (2023). Spatial pattern of glacier mass balance sensitivity to atmospheric forcing in High Mountain Asia. *Journal of Glaciology*, *69*(278), 1616-1633.

Bash, E. A., Moorman, B. J., & Gunther, A. (2018). Detecting short-term surface melt on an Arctic glacier using UAV surveys. Remote Sensing, 10(10), 1547.

Benesty, J., Chen, J., Huang, Y., & Cohen, I. (2009). Pearson correlation coefficient. In *Noise reduction in speech processing* (pp. 1-4). Berlin, Heidelberg: Springer Berlin Heidelberg.

Blau, M. T., Kad, P., Turton, J. V., & Ha, K. J. (2024). Uneven global retreat of persistent mountain snow cover alongside mountain warming from ERA5-land. *npj Climate and Atmospheric Science*, *7*(1), 278.

Braithwaite, R. J. (1995). Positive degree-day factors for ablation on the Greenland ice sheet studied by energy-balance modelling. *Journal of Glaciology*, *41*(137), 153-160.

Braithwaite, R. J., & Zhang, Y. (2000). Sensitivity of mass balance of five Swiss glaciers to temperature changes assessed by tuning a degree-day model. *Journal of Glaciology*, *46*(152), 7-14.

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

Bolibar, J., Rabatel, A., Gouttevin, I., Galiez, C., Condom, T., & Sauquet, E. (2020). Deep learning applied to glacier evolution modelling. *The Cryosphere*, *14*(2), 565-584.

Bolibar, J., Rabatel, A., Gouttevin, I., & Galiez, C. (2020). A deep learning reconstruction of mass balance series for all glaciers in the French Alps: 1967–2015. *Earth System Science Data Discussions*, *2020*, 1-15.

Bolibar, J., Rabatel, A., Gouttevin, I., Zekollari, H., & Galiez, C. (2022). Nonlinear sensitivity of glacier mass balance to future climate change unveiled by deep learning. Nature communications, 13(1), 409.

Cai, J., Luo, J., Wang, S., & Yang, S. (2018). Feature selection in machine learning: A new perspective. *Neurocomputing*, *300*, 70-79.

Cao, B., Pan, B., Guan, W., Wang, J., & Wen, Z. (2017). Changes in ice volume of the Ningchan No. 1 Glacier, China, from 1972 to 2014, as derived from in situ measurements. *Journal of Glaciology*, *63*(242), 1025-1033.

Carey, M., Molden, O. C., Rasmussen, M. B., Jackson, M., Nolin, A. W., & Mark, B. G. (2017). Impacts of glacier recession and declining meltwater on mountain societies. *Annals of the American Association of Geographers*, *107*(2), 350-359.

Cui, Y., Jia, M., Lin, T. Y., Song, Y., & Belongie, S. (2019). Class-balanced loss based on effective number of samples. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 9268-9277).

Che, Y., Zhang, M., Li, Z., Li, H., Wang, S., Sun, M., & Zha, S. (2017). Glacier mass-balance and length variation observed in China during the periods 1959–2015 and 1930–2014. *Quaternary International*, *454*, 68-84.

Chen, J. L., Wilson, C. R., Tapley, B. D., Blankenship, D. D., & Ivins, E. R. (2007). Patagonia icefield melting observed by gravity recovery and climate experiment (GRACE). *Geophysical Research Letters*, *34*(22).

Chen, J., Qin, X., Kang, S., Du, W., Sun, W., & Liu, Y. (2020). Potential effect of black carbon on glacier mass balance during the past 55 years of Laohugou Glacier No. 12, western Qilian Mountains. *Journal of Earth Science*, *31*(2), 410-418.

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

Duan, J. P., Wang, L. L., Ren, J. W., & Li, L. (2009). Progress in glacier variations in China and its sensitivity to climatic change during the past century. *Progress in Geography*, *28*(2), 231-237.

Draeger, C., Radić, V., White, R. H., & Tessema, M. A. (2024). Evaluation of reanalysis data and dynamical downscaling for surface energy balance modeling at mountain glaciers in western Canada. *The Cryosphere*, *18*(1), 17-42.

Faal, R., Saboori, M., Patro, E. R., Ala-Aho, P., & Haghighi, A. T. (2025). A novel spatiotemporal analysis of snow cover pattern over Finland based on ERA5-land reanalysis. *Journal of Hydrology*, 134264.

Fausett, L. V. (2006). *Fundamentals of neural networks: architectures, algorithms and applications*. Pearson Education India.

Fujita, K. (2008). Influence of precipitation seasonality on glacier mass balance and its sensitivity to climate change. *Annals of Glaciology*, *48*, 88-92.

Fuchs, D., Sherwood, S. C., Prasad, A., Trapeznikov, K., & Gimlett, J. (2023). TorchClim v1. 0: A deep-learning framework for climate model physics. *EGUsphere*, *2023*, 1-25.

Gabbi, J., Carenzo, M., Pellicciotti, F., Bauder, A., & Funk, M. (2014). A comparison of empirical and physically based glacier surface melt models for long-term simulations of glacier response. *Journal of Glaciology*, *60*(224), 1140-1154.

Goodfellow, I., Bengio, Y., Courville, A. 2016. Deep learning, MIT press.

Greuell, W., & Oerlemans, J. (2005). Assessment of the surface mass balance along the K-transect (Greenland ice sheet) from satellite-derived albedos. *Annals of Glaciology*, *42*, 107-117.

Grömping, U. (2009). Variable importance assessment in regression: linear regression versus random forest. *The American Statistician*, *63*(4), 308-319.

Guidicelli, M., Huss, M., Gabella, M., & Salzmann, N. (2023). Spatio-temporal reconstruction of winter glacier mass balance in the Alps, Scandinavia, Central Asia and western Canada (1981–2019) using climate reanalyses and machine learning. *The Cryosphere*, *17*(2), 977-1002.

Hastie, T. (2009). The elements of statistical learning: data mining, inference, and prediction.

Hock, R. (2003). Temperature index melt modelling in mountain areas. *Journal of hydrology*, *282*(1-4), 104-115.

Hock, R. (2005). Glacier melt: a review of processes and their modelling. *Progress in physical geography*, *29*(3), 362-391.

Immerzeel, W. W., L. Petersen, S.Ragettli, and F. Pellicciotti (2014), Theimportance of observed gradients ofair temperature and precipitation formodeling runoff from a glacierizedwatershed in the Nepalese Himalayas,Water Resour. Res., 50, 2212–2226,doi:10.1002/2013WR014506.

Kolmogorov, A. (1992). On the empirical determination of a distribution function. In *Breakthroughs in statistics: Methodology and distribution* (pp. 106-113). New York, NY: Springer New York.

Kuhn, M., Dreiseitl, E., Hofinger, S., Markl, G., Span, N., & Kaser, G. (1999). Measurements and models of the mass balance of Hintereisferner. Geografiska Annaler: Series A, Physical Geography, 81(4), 659-670.

Li, Z., Li, H., & Chen, Y. (2011). Mechanisms and simulation of accelerated shrinkage of continental glaciers: a case study of Urumqi Glacier No. 1 in eastern Tianshan, central Asia. Journal of Earth Science, 22(4), 423-430.

Li Z., Han T., Jin Z., Yang H., & Jiao K. (2003). A summary of 40-year observed variation facts of climate and Glacier No. 1 at headwater of Ürümqi River, Tianshan, China. *Journal of Glaciology and Geocryology*, *25*(2), 117-123.

Li K M, Chen S F, Kang L F, et al. (2018). Variation of Continental Glacier and Temperate Glacier in China:A Case Study of Glacier No.1 at the Headwaters of the Urumqi River and Baishui Glacier No.1. Arid Zone Research,35(01): 12-19. DOI: 10.13866/j.azr.

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

Li, Z., Mu, Z., & Gao, R. (2024). Applicability of ERA5 Reanalysis Precipitation Data in Runoff Modeling in China's Ili River Basin. *Journal of Hydrologic Engineering*, *29*(5), 04024036.

Liang, L., Cuo, L., & Liu, Q. (2018). The energy and mass balance of a continental glacier: Dongkemadi Glacier in central Tibetan Plateau. *Scientific Reports*, *8*(1), 12788.

Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision* (pp. 2980-2988).

Liu J., Zhang W., Xia J., Shen Y., & Kand S. (2017). Study of degree-day model from 2000 to 2016: the main progress and key issues. Journal of Glaciology and Geocryology, 39(4), 801-810.

Liu, S., Pu, J., Deng, X., et al. (2014). Glaciers and glacier landscapes in China. Shanghai: Shanghai Popular Science Press.

Meng, Y., Yang, M., Liu, S., Mou, Y., Peng, C., & Zhou, X. (2021). Quantitative assessment of the importance of bio-physical drivers of land cover change based on a random forest method. *Ecological Informatics*, *61*, 101204.

Mihalevich, B. A., Neilson, B. T., & Buahin, C. A. (2022). Evaluation of the ERA5-Land reanalysis data set for process-based river temperature modeling over data sparse and topographically complex regions. Water Resources Research, 58, e2021WR031294. https://doi.org/10.1029/2021WR031294.

Muñoz Sabater, J.: ERA5-Land hourly data from 1950 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [data set], https://doi.org/10.24381/cds.e2161bac, 2019.

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.

Pan B T, Cao B, Guan W J. (2021). Changes of Ningchan No. 1 Glacier in Lenglongling,eastern Qilian Mountains from 2010 to 2020 based on observation. Journal of Glaciology and Geocryology, 43(03):864-873.

Pellicciotti, F., Brock, B., Strasser, U., Burlando, P., Funk, M., & Corripio, J. (2005). An enhanced temperature-index glacier melt model including the shortwave radiation balance: development and testing for Haut Glacier d'Arolla, Switzerland. *Journal of glaciology*, *51*(175), 573-587.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. the Journal of machine Learning research, 12, 2825-2830.

Rabatel, A., Dedieu, J. P., & Vincent, C. (2016). Spatio-temporal changes in glacier-wide mass balance quantified by optical remote sensing on 30 glaciers in the French Alps for the period 1983–2014. *Journal of Glaciology*, *62*(236), 1153-1166.

Reijmer, C.H., Hock, R., 2008. Internal accumulation on Storglaciären, Sweden, in a multi-layer snow model coupled to a distributed energy- and mass-balance model. Journal of Glaciology, 54(54): 61-72.

Reitz, O., Graf, A., Schmidt, M., Ketzler, G., & Leuchner, M. (2021). Upscaling net ecosystem exchange over heterogeneous landscapes with machine learning. Journal of Geophysical Research: Biogeosciences, 126(2), e2020JG005814.

Réveillet, M., Vincent, C., Six, D., & Rabatel, A. (2017). Which empirical model is best suited to simulate glacier mass balances ?. *Journal of Glaciology*, *63*(237), 39-54.

Ren, W., Zhu, Z., Wang, Y., Su, J., Zeng, R., Zheng, D., & Li, X. (2024). Comparison of machine learning models in simulating glacier mass balance: insights from Maritime and Continental Glaciers in High Mountain Asia. *Remote Sensing*, *16*(6), 956.

RGI Consortium. GLIMS: Global Land Ice Measurements from Space, A Dataset of Global Glacier Outlines Version 6.0 Technical Report, Colorado, USA. 2017. Available online: https://www.glims.org/RGI/randolph60.html (accessed on 10 January 2024).

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

Sebbar, B. E., Khabba, S., Merlin, O., Simonneaux, V., Hachimi, C. E., Kharrou, M. H., & Chehbouni, A. (2023). Machine-learning-based downscaling of hourly ERA5-Land air temperature over mountainous regions. *Atmosphere*, *14*(4), 610.

Steiner, D., Walter, A., & Zumbühl, H. J. (2005). The application of a non-linear back-propagation neural network to study the mass balance of Grosse Aletschgletscher, Switzerland. *Journal of Glaciology*, *51*(173), 313-323.

Steidl, V., Bamber, J. L., & Zhu, X. X. (2025). Physics-aware machine learning for glacier ice thickness estimation: a case study for Svalbard. *The Cryosphere*, *19*(2), 645-661.

Su B., LI Z., Zhang M., Guo R., Sun M., Che Y., Ying X. (2015). A comparative study on mass balance between the continental glaciers and the temperate glaciers: Taking the typical glaciers in the Tianshan Mountains and the Alps as examples. Journal of Glaciology and Geocryology, 37(5), 1131-1140.

Su, B., Xiao, C., Chen, D., Huang, Y., Che, Y., Zhao, H., ... & Yao, T. (2022). Glacier change in China over past decades: Spatiotemporal patterns and influencing factors. *Earth-Science Reviews*, *226*, 103926.

Sun, W., Qin, X., Wang, Y., Chen, J., Du, W., Zhang, T., & Huai, B. (2018). The response of surface mass and energy balance of a continental glacier to climate variability, western Qilian Mountains, China. *Climate Dynamics*, *50*(9), 3557-3570.

Thibert, E., Blanc, R., Vincent, C., & Eckert, N. (2008). Glaciological and volumetric mass-balance measurements: error analysis over 51 years for Glacier de Sarennes, French Alps. *Journal of Glaciology*, *54*(186), 522-532.

Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., ... & Papale, D. (2016). Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms. *Biogeosciences*, *13*(14), 4291-4313.

Teufel, B., Carmo, F., Sushama, L., Sun, L., Khaliq, M. N., Bélair, S., ... & Vaze, J. (2023). Physics-informed deep learning framework to model intense precipitation events at super resolution. *Geoscience Letters*, *10*(1), 19.

Wang, S., Wang, J., Zhu, M., Yao, T., Pu, J., & Wang, J. (2024). Long-term glacier variations and the response to climate fluctuation in Qilian Mountains, China. *Journal of Geographical Sciences*, *34*(10), 1904-1924.

Wang, X., Tolksdorf, V., Otto, M., & Scherer, D. (2021). WRF-based dynamical downscaling of ERA5 reanalysis data for High Mountain Asia: Towards a new version of the High Asia Refined analysis. *International Journal of Climatology*, *41*(1), 743-762.

Wu, X., Su, J., Ren, W., Lü, H., & Yuan, F. (2023). Statistical comparison and hydrological utility evaluation of ERA5-Land and IMERG precipitation products on the Tibetan Plateau. *Journal of hydrology*, *620*, 129384.

Xiao, C., Liu, S., Zhao, L., Wu, Q., Li, P., Liu, C., ... & Pu, J. (2007). Observed changes of cryosphere in China over the second half of the 20th century: an overview. Annals of Glaciology, 46, 382-390.

Yong, Z., & Shi-yin, L. I. U. (2006). Progress of the application of degree-day model to study glaciers and snow cover. Journal of Glaciology and Geocryology, 28(1), 101-107.

Zandler, H., Haag, I., & Samimi, C. (2019). Evaluation needs and temporal performance differences of gridded precipitation products in peripheral mountain regions. *Scientific reports*, *9*(1), 15118.

Zhang, H., Li, Z., Zhou, P., Zhu, X., & Wang, L. I. N. (2018). Mass-balance observations and reconstruction for Haxilegen Glacier No. 51, eastern Tien Shan, from 1999 to 2015. Journal of Glaciology, 64(247), 689-699.

Zhang, H., Li, Z., & Zhou, P. (2021). Mass balance reconstruction for Shiyi Glacier in the Qilian Mountains, Northeastern Tibetan Plateau, and its climatic drivers. Climate Dynamics, 56(3), 969-984.

Zhao, P., & He, Z. (2022). A first evaluation of ERA5-Land reanalysis temperature product over the Chinese Qilian Mountains. *Frontiers in Earth Science*, *10*, 907730.

https://doi.org/10.5194/egusphere-2026-333
Preprint. Discussion started: 24 February 2026

Zhu, M., Yao, T., Yang, W. E. I., Xu, B., Wu, G., Wang, X., & Xie, Y. (2018). Reconstruction of the mass balance of Muztag Ata No. 15
645 glacier, eastern Pamir, and its climatic drivers. *Journal of Glaciology*, *64*(244), 259-274.