Journal of Glaciology



Article

Cite this article: Wang J, Zhao L, Dong J and Moore JC (2026) Machine-learning emulation of the modeled basal thermal state for Totten Glacier. *Journal of Glaciology* **72**, e61, 1–14. https://doi.org/10.1017/jog.2026.10170

Received: 8 December 2025 UTC
Revised: 1 May 2026 UTC
Accepted: 1 May 2026 UTC

Keywords:
Antarctic glaciology; basal ice; glacier modeling; ice-sheet modeling; ice temperature

Corresponding authors: Liyun Zhao; Email: zhaoliyun@bnu.edu.cn; John C. Moore; Email: john.moore.bnu@gmail.com

© The Author(s), 2026. Published by Cambridge University Press on behalf of International Glaciological Society. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Machine-learning emulation of the modeled basal thermal state for Totten Glacier

Junshun Wang[1], Liyun Zhao[1], Jin Dong[2] and John C. Moore[3]

[1]State Key Laboratory of Earth Surface Processes and Hazards Risk Governance (ESPHR), Faculty of Geographical Science, Beijing Normal University, Beijing, China; [2]Information Center of Ministry of Natural Resources, Beijing, China and [3]Arctic Centre, University of Lapland, Rovaniemi, Finland

Abstract

The basal thermal state of the Antarctic ice sheet strongly influences ice dynamics and mass balance. Although basal thermal states can be simulated using ice-sheet models, significant uncertainties persist, with poorly constrained geothermal heat flux being one important source of uncertainty. The substantial computational cost of ice-sheet models restricts large ensemble simulations essential for uncertainty quantification and sensitivity analyses. To address this limitation, we develop machine-learning emulators for rapidly simulating the basal temperature and melt rate of Totten Glacier. Three distinct machine-learning emulators were trained using outputs of full-Stokes model simulations with different geothermal heat flux fields, achieving accurate replication (R^2 up to 0.92). At least five simulations are needed to train a reliable emulator for basal ice temperature. Model interpretability reveals complementary physics learning: Random forest and XGBoost prioritize local dynamic drivers (basal topography, ice velocity), while neural network captures lower-complexity statistical mapping that is more strongly influenced by globally predictive variables (surface temperature, ice thickness). For basal melt rate, surface velocity emerges as the primary driver. These findings highlight substantial potential for integrating machine learning with ice-sheet modeling to efficiently and better predict basal thermal states.

1. Introduction

Sea level rise projections rely on accurate predictions of ice mass loss from Antarctica (Ligtenberg and others, 2013; Dutton and others, 2015; DeConto and Pollard, 2016; Pattyn and Morlighem, 2020). Basal ice temperature is a key factor influencing basal sliding, basal melt, subglacial hydrology and the ice-flow regime (Engelhardt, 2004; Pattyn, 2010; Kyrke-Smith and others, 2014; Seroussi and others, 2017; Dawson and others, 2022), and hence mass loss (Dawson and others, 2022). If basal temperatures reach the pressure-melting point, basal sliding and basal melt may occur, allowing for much faster flow than that driven solely by internal ice deformation. These faster-flowing regions contribute disproportionately to ice discharge and dynamic mass loss and are often more responsive to external forcing. Basal thermal conditions also influence subglacial hydrology because basal temperatures at or close to the pressure-melting point favor the presence of liquid water, thereby controlling effective pressure and basal resistance. Importantly, the relevance of basal temperature extends beyond simply determining whether the bed has already reached the pressure-melting point. Even when basal ice remains below the pressure-melting point, warmer basal conditions imply softer ice and thus lower effective viscosity, as ice stiffness decreases by nearly three orders of magnitude between −20°C and 0°C (Paterson, 1994). Moreover, spatial variations in basal temperature below the melting point help distinguish cold-based, near-temperate and transitional basal regimes, and thereby identify regions that may be particularly sensitive to relatively small perturbations in geothermal heat flux (GHF) or frictional heating that could shift basal conditions toward thaw (Dawson and others, 2022). Capturing precise basal temperatures is a critical prerequisite for correctly spinning up and constraining vertical temperature profiles when solving coupled thermomechanical equations (Park and others, 2024). Finally, basal thermal conditions are important for paleoclimate applications, since the selection of suitable drilling sites for 'Oldest Ice' cores requires avoiding areas with significant basal melting that would remove the oldest ice near the bed (Fischer and others, 2013).

Despite its importance, the continental-scale distribution of basal thermal state remains poorly constrained. Our knowledge relies on three main sources: direct borehole measurements, which are accurate but sparse; geophysical inferences from radar and seismic data, which have broader but still incomplete coverage; and simulation results from ice-sheet models (MacGregor and others, 2016, 2022; Matsuoka and others, 2012). Thermodynamic ice-sheet

Downloaded from https://www.cambridge.org/core. 18 Nov 2026 at 04:51:36, subject to the Cambridge Core terms of use.

models are widely used to simulate thermal states (Budd and others, 1994; Pattyn, 2010; Seroussi and others, 2019). However, significant uncertainties persist, primarily due to poorly constrained GHF (Reading and others, 2022). While three-dimensional (3-D) full-Stokes models provide physically comprehensive descriptions of ice flow, their substantial computational cost restricts large-scale ensemble simulations needed for systematic exploration of basal temperature sensitivity to boundary condition uncertainties, particularly GHF. Machine learning (ML) encompasses a broad set of approaches that are increasingly being applied across Earth system science, including oceanography (Ducournau and Fablet, 2016; Lou and others, 2023), climatology (Rhee and Im, 2017; Jiang and others, 2018) and hydrology (Marçais and De Dreuzy, 2017; Xu and Liang, 2021). In cryosphere research, recent studies have used ML for tasks such as automated glacier mapping (Lu and others, 2021; Xie and others, 2021), ice thickness estimation (Werder and others, 2020; Jouvet and others, 2022), calving front extraction (Zhang and others, 2019; Mohajerani and others, 2021) and surface mass balance estimation (Anilkumar and others, 2023; van der Meer and others, 2025). While several studies have employed ML in related applications, such as inferring basal friction coefficients and reconstructing glacier-wide surface mass balance (Bolibar and others, 2020; Umlauft and others, 2023), the use of ML as direct emulators for basal thermal properties remains largely underexplored.

Totten Glacier (Fig. 1) is one of the most vulnerable glaciers in East Antarctica to a warming climate (Li and others, 2016; Dow and others, 2020). It has the largest ice discharge in East Antarctica (Rignot, 2006), and the drainage basin of Totten Glacier contains enough ice to raise global sea level by approximately 3.9 m. The glacier has also experienced significant thinning and mass loss during recent decades (Li and others, 2015). Recent studies indicate that Totten Glacier is losing mass at a rate of approximately 7 Gt yr^{-1}, with 73% of this loss attributed to ice-sheet dynamics (Li and others, 2016). As the primary outlet glacier of the Aurora Subglacial Basin, a marine basin below sea level, it is of particular concern because its reverse-sloping bed geometry may make it susceptible to marine ice-sheet instability (Schoof, 2007; Sergienko and others, 2026). Beyond this large-scale geometric control, the stability of Totten Glacier and ice flow dynamics are sensitive to subglacial hydrology and basal temperature. GHF, the heat flow from Earth's crust to the ice-bed interface, critically influences the ice temperature profile and consequently affects ice rheology and flow dynamics (Burton-Johnson and others, 2020; Dziadek and others, 2021). However, there are currently multiple GHF maps available for Antarctica, based on various datasets and methods, including seismic models (Shapiro and Ritzwoller, 2004; An and others, 2015; Shen and others, 2020), rock magnetic models (Purucker, 2012; Martos and others, 2017) and multivariate or ML approaches (Dziadek and others, 2017; Lösing and Ebbing, 2021; Stål and others, 2021; Haeger and others, 2022). The various spatial GHF datasets result in significantly different distributions of basal ice temperatures that then impact ice-sheet model simulations (Kang and others, 2022; Huang and others, 2024). The substantial computational cost of high-resolution full-Stokes ice-sheet models limits comprehensive basal temperature simulations across all GHF datasets, highlighting the demand for efficient predictive emulators of basal ice temperature.

In this study, we develop and evaluate an ML framework to serve as a computationally efficient emulator for a full-Stokes ice-sheet model, with the specific goal of predicting basal ice temperature for Totten Glacier. We train and compare three distinct nonlinear models—random forest (RF), XGBoost and neural network (NN)—against a traditional linear regression baseline to highlight the system's nonlinear dynamics. The emulators use bedrock elevation, ice thickness, surface temperature, observed surface velocity and GHF as inputs, and predict basal ice temperature; in additional experiments involving simultaneous prediction, basal melt rate is also included as an output. A detailed description of the ML models is provided in Section 2.2. Our analysis is multifaceted: we not only assess the models' predictive accuracy but also determine the minimum number of numerical simulations required for robust training, test the models' capacity for simultaneous prediction of both basal ice temperature and melt rate, and use state-of-the-art interpretability techniques to identify the dominant physical drivers of these subglacial conditions. Once trained on ice-sheet model outputs, this emulator enables rapid predictions of basal thermal conditions, thereby providing a computationally efficient alternative to extensive continental-scale ice dynamics simulations required for sensitivity analysis and uncertainty quantification.

2. Data and methods

2.1. Preparation of features and training dataset

In this study, the ML emulators are trained using inputs and outputs from Huang and others (2024), in which a coupled modeling workflow was used to generate final steady-state basal thermal fields for Totten Glacier under eight GHF scenarios. The emulator targets extracted from the steady-state simulations are basal ice temperature and basal melt rate. The input data of the ice-sheet model include surface and bottom elevation, ice thickness, surface temperature, surface velocity and GHF. Given that the surface topography of grounded ice can be reconstructed by bottom topography plus ice thickness and is strongly correlated with surface ice temperature (Pearson $r = -0.98$), we excluded surface topography from the input feature set. The final set of five variables represents the key quantities that govern heat and momentum conservation in the physical ice-sheet model and also serves as the input variables for the ML emulators. Bedrock elevation and ice thickness data are sourced from the MEaSUREs BedMachine Antarctica, Version 2 (Morlighem and others, 2020), consistent with the numerical workflow of Huang and others (2024), from which the emulator training data were derived. Surface temperature data are obtained from ALBMAP v1 with a resolution of 5 km (Le Brocq and others, 2010). The observed surface ice velocity data are from the MEaSUREs InSAR-Based Antarctica Ice Velocity Map, version 2, with a resolution of 450 m (Rignot and others, 2017). The GHF fields are from eight sources: Purucker (2012), Shapiro and Ritzwoller (2004), An and others (2015), Shen and others (2020), Stål and others (2021), Haeger and others (2022), Lösing and Ebbing (2021) and Martos and others (2017). As GHF does not directly influence sub-ice-shelf basal temperature, the ice shelf is excluded from this study.

In Huang and others (2024), the basal thermal state of Totten Glacier was estimated using an off-line coupling between a forward model and an inverse model. The forward model consists of an improved shallow-ice-approximation thermomechanical model with a subglacial hydrology model (Wolovick and others, 2021), and aims to provide an appropriate vertical temperature profile for the inverse model. The full-Stokes inverse model is used to adjust the spatial distribution of the basal friction coefficient to minimize the misfit between simulated and observed surface

Downloaded from https://www.cambridge.org/core. 18 Nov 2026 at 04:51:36, subject to the Cambridge Core terms of use.

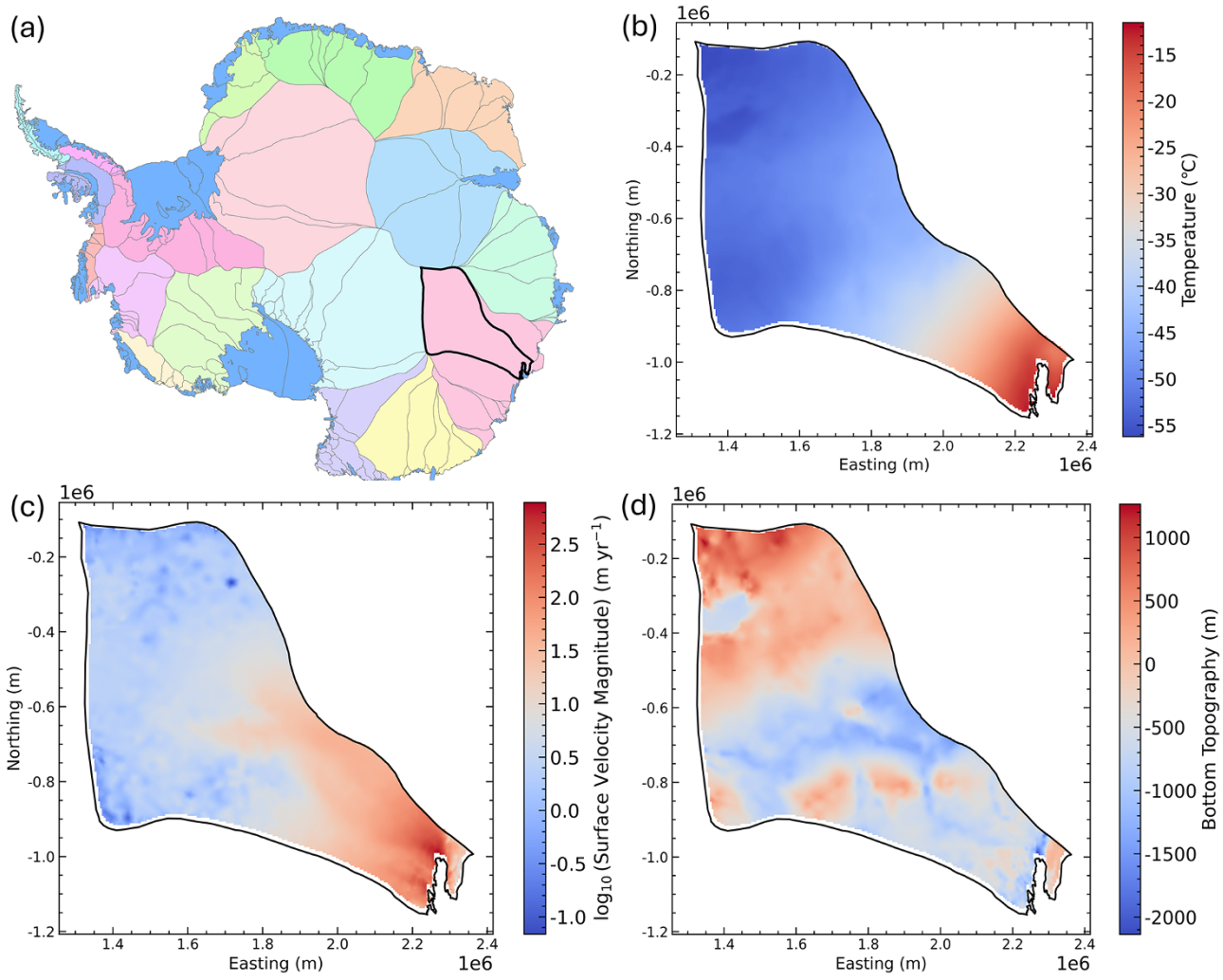


Figure 1. (a) Location of Totten Glacier in Antarctica; (b) surface ice temperature; (c) surface velocity magnitude; (d) bed elevation. The solid black line outlines the study domain, which excludes the Totten Ice Shelf.

velocities. The modeled velocity is obtained by solving a 3-D full-Stokes model subject to a basal sliding condition. Ice viscosity in the full-Stokes model is strongly dependent on ice temperature. Huang and others (2024) employed a finite-element mesh in their ice model, with mesh sizes ranging from a maximum of 20 km far inland to a minimum of 800 m in regions with fast ice flow. We take the result at every node of the 2-D horizontal footprint mesh as the data used for ML model training and evaluation. We did not perform any regular-grid interpolation of the ice-sheet model outputs because interpolating the data onto a regular grid could introduce errors that are not inherent to the ice-sheet model results. There are 11 620 nodes in the 2-D mesh. In other words, for each GHF dataset, 11 620 samples are used for ML emulator.

Huang and others (2024) generated steady-state basal thermal outputs for eight GHF scenarios. In the present study, six of these scenarios were used to form the development dataset for model training and internal validation, whereas the Purucker (2012) and Martos and others (2017) scenarios were withheld entirely and used only as an independent test set. We selected these two scenarios because they produce the lowest and highest mean basal ice temperatures, respectively, and therefore provide a deliberately challenging test of generalization across unseen GHF forcings. We note, however, that this should be interpreted as testing generalization across withheld GHF scenarios within the Totten Glacier setting, rather than as a true test of transferability to a different glacier system. Separately, to understand how the amount of training data impacts prediction accuracy, we systematically vary the number of simulations used for training from one to seven. For each training set size, we tested all possible combinations of simulation datasets, training the model on the selected simulations and evaluating it on the remaining GHF simulations that were not included in training. This process allows us to determine the minimum number of expensive ice-sheet simulations required to train an effective ML emulator.

2.2. ML emulation

In this study, we use ML to create a computationally efficient surrogate model (or 'emulator') that learns the relationship between inputs and outputs from a complex physical ice-sheet model. Unlike traditional simulations that solve explicit physical equations, a data-driven approach learns these relationships directly from simulation results (Liu and others, 2024). We employ a super-

Downloaded from https://www.cambridge.org/core. 18 Nov 2026 at 04:51:36, subject to the Cambridge Core terms of use.

vised learning framework, where an algorithm is trained on a dataset containing both the model inputs (e.g. GHF, basal topography) and outputs (the basal ice temperature calculated by the ice-sheet model). The primary goal is to determine how effectively an ML emulator can reproduce the basal thermal state simulated by the numerical ice-sheet model, but at a fraction of the computational cost. This emulator should be viewed as a diagnostic tool for estimating the present-day thermal state under different GHF fields, rather than as a prognostic module that can be directly embedded in evolving transient simulations to evaluate how the thermal state evolves, since it has been trained on steady-state velocities and temperatures. To enhance the interpretability of the ML models, we incorporated Shapley Additive Explanations (SHAP) values to explain model predictions and connect the data-driven findings back to glaciological processes. To compare model performance, we also used a baseline model: multiple linear regression.

We select three ML emulators that learn in different ways. RF is an ensemble method, meaning it combines multiple base models to improve predictive capability. The base model it uses is a binary decision tree (Supplementary Fig. S1), represented as a tree structure. By adjusting different parameter settings, RF model builds hundreds or thousands of different decision trees—an entire 'forest' (Supplementary Fig. S1). Each tree is trained by a random subset of the training data. To make a final prediction, the model asks every tree in the forest for its predicted value and calculates the average. This process, known as 'bagging', makes the final prediction more robust and significantly mitigates the risk of overfitting, although it does not make the model entirely immune to it (Breiman, 2001). The strength of RF is its ability to automatically capture complex, nonlinear interactions (e.g. the effect of GHF may differ in thick vs thin ice) without requiring pre-defined physical equations.

Like RF, XGBoost is also an ensemble of decision trees, but it builds them in a sequential way known as 'boosting' (Chen and Guestrin, 2016). XGBoost builds trees in a sequential manner, where each new tree is added to correct the errors of earlier ones, using techniques such as greedy algorithms for split point selection, which improve both model performance and computational speed. The final prediction is a weighted sum of the predictions from all the trees (Supplementary Fig. S2). This sequential, error-correcting process makes XGBoost often outperform other methods.

The NN model is inspired by the structure of the brain and is fundamentally different from tree-based models. It processes information through interconnected layers of computational 'neurons' or nodes (Dongare and others, 2012) (Supplementary Fig. S3). The input data (GHF, bottom topography, etc.) are fed into the first layer. NN models use a weighted combination of input features along with nonlinear functions (called activation functions) such as sigmoid and tanh to produce the activations of the neurons in the next layer (called a hidden layer) (Supplementary Fig. S3a). Multiple hidden layers form an NN and ultimately give the predicted output, which is basal ice temperature in our case (Supplementary Fig. S3b). NNs have been shown to approximate any continuous function even with a single hidden layer and sufficient neurons (Hornik, 1991). NNs can fit complex functional relationships through multilayer structures and nonlinear activation functions, and are suitable for various complex tasks.

2.3. Hyperparameter fine-tuning and performance evaluation

ML emulators are inherently data-driven, and their performance is influenced not only by the internal parameters obtained automatically during training but also by hyperparameters, which are manually set before training. These hyperparameters can be regarded as analogous to prescribed constant parameters in ice-sheet models, such as the number of trees in an RF or the number of neurons within an NN model. We use a process called k-fold cross-validation on our development dataset of six simulations (all except those from Purucker (2012) and Martos and others (2017)). The dataset is randomly split into three equal parts, or 'folds'. The emulator is trained on two folds, and its performance is evaluated on the third (the validation fold). This process was repeated three times, ensuring each fold was used for validation exactly once. It is noted that the development dataset was randomly split and used only for hyperparameter selection, while the main reported model errors come from the independent withheld-GHF test set. We tested numerous combinations of hyperparameters listed in Table 1 and selected the one that yielded the best average performance across the three validation folds. Finally, the final emulators were retrained on the entire development dataset using these optimal settings. The NN model is a fully connected feed-forward model implemented in Keras, with ReLU activation in the hidden layers, Glorot uniform weight initialization and the Adam optimizer. The number of hidden layers, neurons per layer, learning rate and number of training epochs were optimized by grid search, while the batch size was fixed at 128. No additional dropout or L2 weight decay was used in the final implementation. Z-score normalization (standardization) was applied to the input features for all three ML models.

To evaluate emulator performance, we calculated three statistical metrics on the test dataset: the coefficient of determination (R^2), root mean square error (RMSE) and mean absolute error (MAE). The coefficient of determination (R^2) quantifies the proportion of variance in the target data that is explained by the model. A perfect fit yields $R^2 = 1$, whereas $R^2 = 0$ indicates no improvement over simply predicting the mean of the target values. R^2 can also be negative when the model performs worse than this mean-based reference. RMSE represents the absolute deviations between the target and predicted values. RMSE is sensitive to large errors, particularly when outliers or large deviations exist, and penalizes large discrepancies. MAE is the average of the absolute prediction errors, providing a straightforward and intuitive measure of prediction accuracy.

2.4. Interpreting model predictions using SHAP

While our ML models demonstrate high predictive accuracy, their complex internal structures can make it difficult to understand precisely how they arrive at a prediction. To overcome this 'black-box' problem and extract meaningful scientific insights, we employed SHAP, a method for interpreting the output of any ML model (Lundberg and Lee, 2017). The core idea behind SHAP is to explain an individual prediction by quantifying exactly how much each input variable contributed to pushing the prediction away from a baseline value (e.g. the average predicted basal ice temperature across the entire dataset). For every sample prediction the model makes, each input variable is assigned a SHAP value, which has two key properties:

(1) Direction: The sign of the value shows the direction of the effect. A positive SHAP value means that a variable (e.g. a high GHF value) pushed the prediction toward a warmer temperature, which is defined as a positive contribution. A negative SHAP value means it pushed the prediction toward a lower temperature, referred to as a negative contribution.

Downloaded from https://www.cambridge.org/core. 18 Nov 2026 at 04:51:36, subject to the Cambridge Core terms of use.

Table 1. Hyperparameter combinations for each model.

Machine learning model	Hyperparameter	Values	Number of combinations
Random forest	Max depth	5, 10, 20, **30**, 40	120
	Minimum samples in a leaf	**1**, 2, 4, 8	
	Number of trees	10, 20, 50, 100, 200, **300**	
XGBoost	Gamma	**0**, 0.1, 0.3	864
	Learning rate	0.01, 0.1, **0.3**, 0.5	
	Max depth	3, 5, 7, **9**	
	Minimum child weight	1, **3**, 5	
	Number of trees	10, 20, 50, 100, 200, **300**	
Neural networks	Epochs	20, 40, 60, 80, **100**	300
	Layers	2, 4, **6**, 8, 10	
	Learning rate	0.0001, **0.001**, 0.01	
	Neurons in each layer	20, 50, **100**, 150	

Note: The bold numbers in the table represent the optimal hyperparameter combination for each model.

(2) Magnitude: The size of the value reveals the importance of that variable's contribution to that specific prediction.

A major strength of SHAP is that these contributions are additive; the sum of all the SHAP values for a single sample prediction perfectly explains the difference between that prediction and the baseline average. This allows us to analyze model behavior on a case-by-case basis, for instance by assessing whether a region of warm basal ice was primarily associated with high GHF, greater ice thickness or other local boundary conditions, as inferred from the emulator. To assess which variables are most important overall, we can aggregate these individual explanations. We rank the input variables by their mean absolute SHAP value across all points in our test dataset. This metric quantifies the average impact each variable has on the model's predictions, providing a robust and consistent measure of global feature importance. By using SHAP, we can transform our ML emulators from simple prediction tools into instruments for investigating the complex physical relationships governing the basal thermal state of Totten Glacier.

3. Results

3.1. Performance of ML emulators

The RF emulator exhibits superior predictive performance on the combined test dataset (Purucker, 2012; Martos and others, 2017), achieving an RMSE of 0.73°C, a coefficient of determination (R^2) value of 0.89 and an MAE of 0.19°C (Fig. 2b). To assess the emulator's robustness, we evaluated its performance on the two test simulations individually. On the Purucker (2012) data, the RF emulator replicates the main spatial patterns of basal temperature ($R^2 = 0.87$), though it exhibits a tendency to slightly overestimate temperatures in colder regions, particularly in slow-flowing interior areas with elevated basal topography (Fig. 3e). In contrast, the RF emulator performs particularly well on the Martos and others (2017) dataset, achieving close agreement between predictions and simulation-derived data: R^2 approximately 0.97, RMSE of 0.21°C and MAE < 0.1°C (Fig. 3b). Spatial predictions closely matched the target basal ice temperatures across most of the glacier, with only minor underestimation at the southwest corner of the glacier (Fig. 3h).

The XGBoost emulator, another tree-based ML approach, yields a comparable and even marginally superior performance to the RF emulator (Fig. 2c). The emulator achieves an R^2 value of 0.91, and an RMSE of 0.65°C on the test dataset, with its regression line closely aligning with the 1:1 reference line, indicating strong agreement between predicted and target values. Spatially, the XGBoost emulator produces error patterns qualitatively similar to those of the RF emulator (Supplementary Fig. S4).

The NN emulator, while performing adequately overall, demonstrates lower predictive capability compared to the tree-based emulators (Fig. 2d). With an R^2 value of 0.86, the NN emulator's accuracy is somewhat reduced relative to both RF and XGBoost. The NN captures the large-scale distribution of basal temperature but struggles to resolve finer-scale, local variations with the same precision (Supplementary Fig. S5).

In summary, the three ML emulators (RF, XGBoost and NN) significantly outperform the linear baseline emulator in predictive performance (Fig. 2a). This result indicates that the relationships between the glaciological input variables and the resulting basal temperature are inherently complex and nonlinear, requiring the flexibility of ML approaches to be emulated accurately.

3.2. The influence of training dataset size

While our final emulators are trained on six GHF simulations, the size of the training dataset is a critical factor that can impact predictive performance. To investigate this, we systematically evaluate how emulator accuracy changes as the number of simulations in the training set is varied from one to seven. For each training-set size, the emulator is trained on the selected subset of GHF simulations and evaluated on the remaining GHF simulations that were not used for training. As illustrated by the learning curves in Fig. 4, all three ML emulators show a clear trend: performance improves as the number of training simulations increases. The R^2 value rises steeply at first, before beginning to plateau once the training set size reaches five simulations. This trend indicates that with five or six simulations, the emulators have been exposed to enough variability to learn the underlying physical relationships effectively. However, even with consistent training set sizes, emulator performance exhibits variability across different dataset combinations. Beyond six training simulations, the intervals of the R^2 plot broaden, indicating performance degradation for certain dataset combinations. We observe that the RF emulator consistently outperformed the other emulators at smaller training set sizes (e.g. one to three simulations), exhibiting relatively stable performance across different dataset combinations. Despite this relative advantage, all emulators demonstrate suboptimal performance with limited training data. As more training data were added, the performance gap between the emulators narrowed and their predictive capabilities converged. When the training set size increased to 6, RF and XGBoost performed similarly, with XGBoost even slightly outperforming RF in terms of RMSE. These findings lead

Downloaded from https://www.cambridge.org/core. 18 Nov 2026 at 04:51:36, subject to the Cambridge Core terms of use.

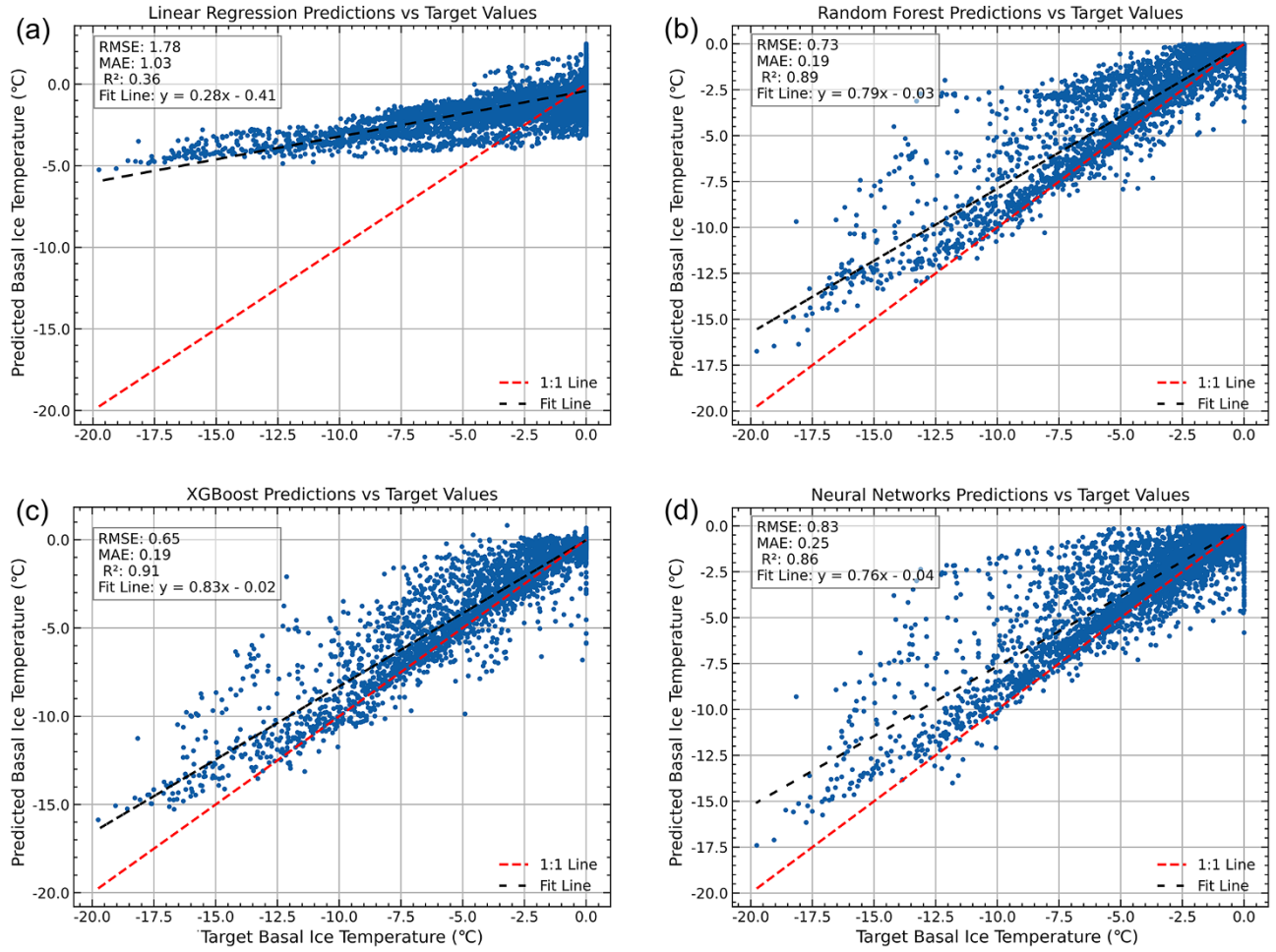


Figure 2. Scatter plot depicting prediction performance of each emulator on the two test datasets (N = 23 240), Purucker (2012) dataset and Martos and others (2017) dataset: (a) linear regression (LR); (b) random forest (RF); (c) extreme gradient boosting (XGBoost); (d) neural networks (NN). Each data point represents the result at a node of the 2-D horizontal footprint mesh used in Huang and others (2024).

to a key conclusion: at least five diverse ice-sheet simulations are required to train a reliable ML emulator for Totten Glacier. Our decision to use all six available development simulations for our final emulator is therefore empirically supported, providing confidence that our emulator's performance is not limited by the quantity of training data.

Finally, we systematically evaluate how the inclusion of a specific individual dataset in the training set influenced emulator performance. The analysis reveals that training sets including the An and others (2015) simulation consistently produce more accurate emulators across all metrics (Fig. 4d). For the RF emulator, combinations including this dataset yield the highest average R^2 and the lowest RMSE and MAE. This beneficial effect is not unique to one emulator, as similar improvements are observed for both XGBoost and the NN when the An and others (2015) data are included in the training set (Supplementary Fig. S6). This suggests that the An and others (2015) simulation provides training information that is particularly useful for the emulators, possibly because it offers a broader or more representative sampling of the modeled basal thermal state and its associated feature space. For example, it may better capture the transition regions between colder and warmer basal conditions.

3.3. Physical drivers of basal ice temperature identified by the emulators

A key advantage of using ML emulators is the ability to efficiently investigate which physical variables most strongly influence basal temperature—a task that would otherwise require extensive sensitivity experiments with a numerical ice-sheet model. We use SHAP values to analyze the trained emulators and quantify the importance of the five input variables: bottom topography, surface ice velocity, surface temperature, GHF and ice thickness.

SHAP value analysis from the final RF emulator reveals that the bottom topography is the single most influential predictor of basal temperature, followed by surface ice velocity (Fig. 5a). The distribution of SHAP values (Fig. 5) shows that, generally, bottom topography demonstrates a negative contribution to basal ice temperature, where higher bedrock elevations correspond to lower predicted basal ice temperatures. Conversely, both surface ice velocity and GHF show a positive contribution, meaning higher values in these variables tend to predict warmer basal temperature.

These findings are strongly corroborated by the XGBoost emulator. Although it slightly alters the ranking of the less influential variables (swapping surface temperature and GHF), it agrees with the RF emulator that bottom topography and surface ice velocity

Downloaded from https://www.cambridge.org/core. 18 Nov 2026 at 04:51:36, subject to the Cambridge Core terms of use.

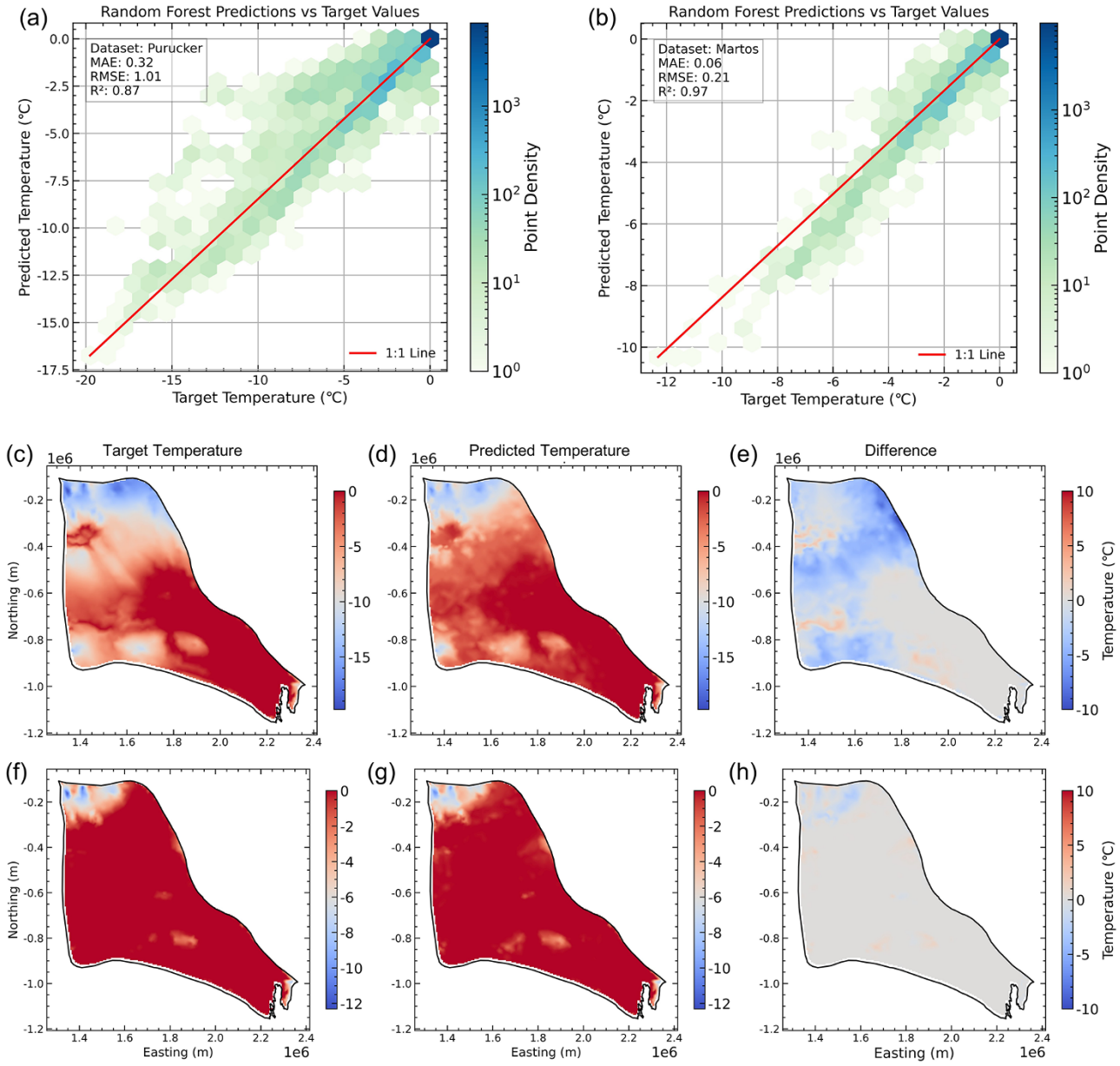


Figure 3. Prediction performance (scatter point density), target basal ice temperature, predicted basal ice temperature and their difference (target minus predicted) for RF emulator on Purucker (2012) dataset (a, c–e) and Martos and others (2017) dataset (b, f–h).

are the two dominant drivers. Crucially, the physical nature of these relationships (i.e. whether a variable's contribution is positive or negative) is identical between both tree-based emulators (Fig. 5b and d). It is important to interpret these relative rankings with caution. A lower rank for a variable like ice thickness does not imply it is physically unimportant; rather, it suggests that within the context of our training data, other variables provided stronger predictive power. Interestingly, the NN emulator diverged somewhat from the tree-based emulators, ranking ice thickness as the third most important variable. Despite these differences in ranking, the fundamental physical relationships identified by the NN—such as the negative contribution of bottom topography and the positive contribution of GHF—were consistent with the other emulators. This broad agreement across three different ML architectures provides confidence in the physical realism captured by our emulators.

3.4. Simultaneous prediction of basal ice temperature and melt rate

We next assessed the emulators' capability to perform a more complex task: predicting basal ice temperature and basal melt rate simultaneously from the same set of input variables. The tree-based emulators handled this multi-output challenge exceptionally well. As shown in Fig. 6, both the RF and XGBoost emulators maintained their high accuracy for basal temperature prediction while also delivering accurate predictions for the basal melt rate. For multi-target predictions, we utilized XGBoost's built-in multi-output regression support. Unlike traditional wrappers that train separate models for each output, this native configuration builds vector-leaf trees, in which each leaf stores a vector of predictions for both basal temperature and melt rate simultaneously. The performance of the NN emulator for temperature prediction

Downloaded from https://www.cambridge.org/core. 18 Nov 2026 at 04:51:36, subject to the Cambridge Core terms of use.

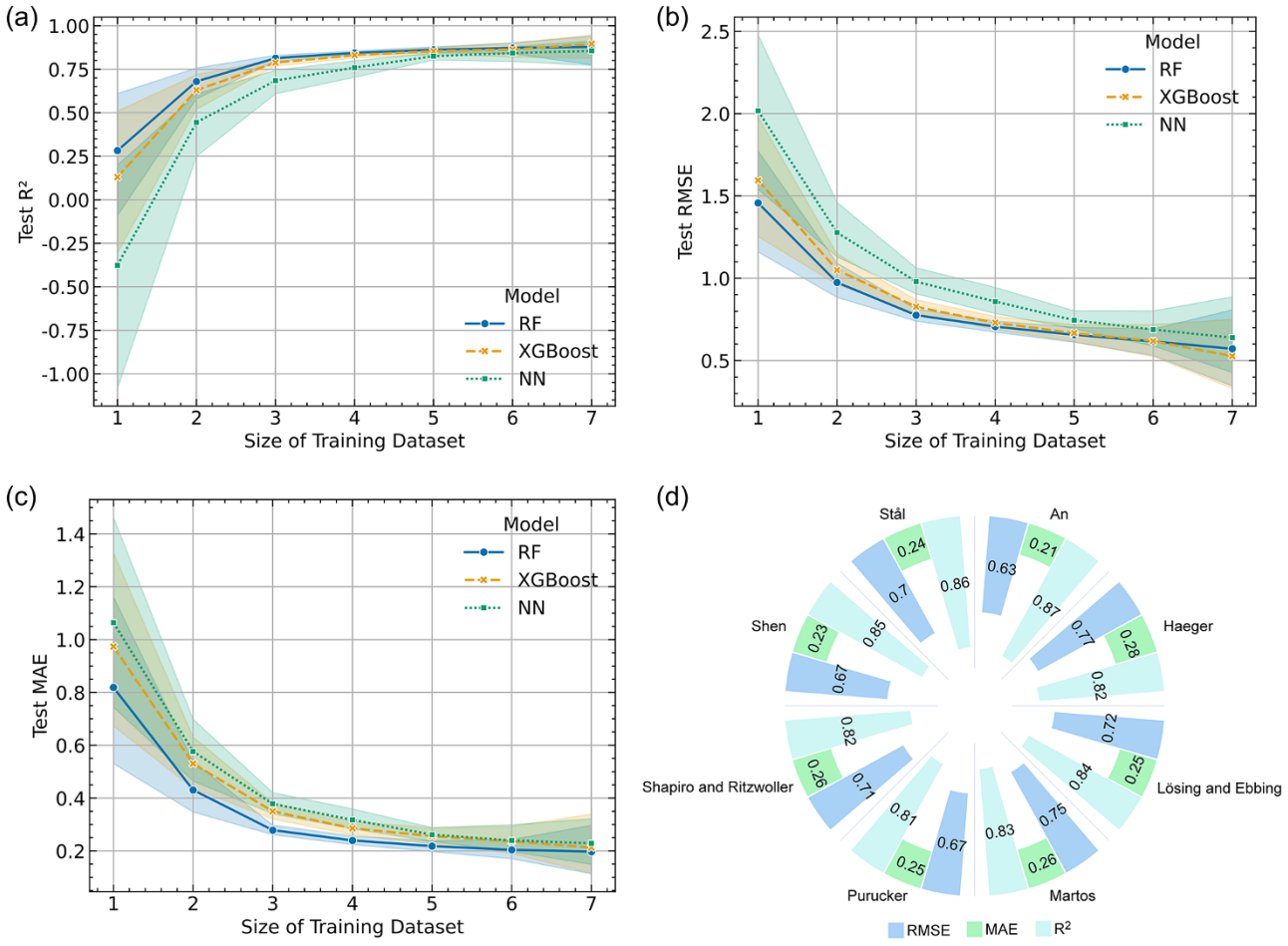


Figure 4. Emulator-wise testing performance varying the training dataset size: (a) R^2; (b) RMSE; (c) MAE. Shaded regions indicate the range of performance metrics across all test set permutations at a given training set size. The average prediction performance across all combinations that include a specific individual dataset in the training set of RF emulator is depicted in plot (d).

shows a minor degradation when it is also tasked with predicting melt rate. This slight trade-off is expected, as the emulator must optimize its internal parameters to minimize the combined prediction error for two distinct variables. SHAP value analysis for the basal melt rate predictions reveals a consistent and physically intuitive driver across all three emulators. Surface ice velocity is identified as the overwhelmingly dominant predictor, with higher velocities consistently corresponding to higher predicted melt rates (Supplementary Fig. S7). This finding aligns with the physical understanding that frictional heat, a key component of basal melt, is directly related to the sliding velocity at the ice-bed interface.

4. Discussion

In our final prediction experiments, we evaluate the performance of three ML emulators using the independent test set that includes the Purucker (2012) and Martos and others (2017) datasets. Our results confirm the potential of ML emulators to efficiently reproduce computationally expensive ice-sheet simulations. The XGBoost emulator yields a marginally superior performance, but all three nonlinear emulators (XGBoost, RF and NN) substantially outperform the baseline linear regression. This finding strongly suggests that the relationship between glaciological inputs and basal temperature is fundamentally nonlinear, making sophisticated, data-driven approaches essential for accurate prediction. The results align with the conclusions of Anilkumar and others (2023), who similarly advocated for complex emulators in glacier mass balance prediction.

A closer analysis of emulator errors reveals a key challenge for any data-driven emulator in the geosciences: extrapolation. All emulators, whether linear or ML-based, show some overestimation on the test set, as evidenced by the distribution of fitted lines compared to the 1:1 line (Fig. 2). Upon splitting the test set into individual datasets, we observe that the overestimated predictions predominantly originated from the Purucker (2012) dataset (Fig. 3a and Supplementary Figs. S4a and S5a). The Purucker (2012) dataset includes basal ice temperatures lower than those in the training data, which have temperatures ranging from −17.11°C to 0°C, whereas the minimum temperature in the Purucker (2012) dataset is −19.7°C. This discrepancy likely contributes to the overestimation of basal ice temperature by the emulator. The emulators, having never seen such cold examples, struggle to predict accurately outside their training domain. This highlights a limitation not of the emulators themselves, but of the representativeness of the training data.

Furthermore, the predictive challenge is compounded by a data imbalance issue. The basal thermal state of Totten Glacier

Downloaded from https://www.cambridge.org/core. 18 Nov 2026 at 04:51:36, subject to the Cambridge Core terms of use.

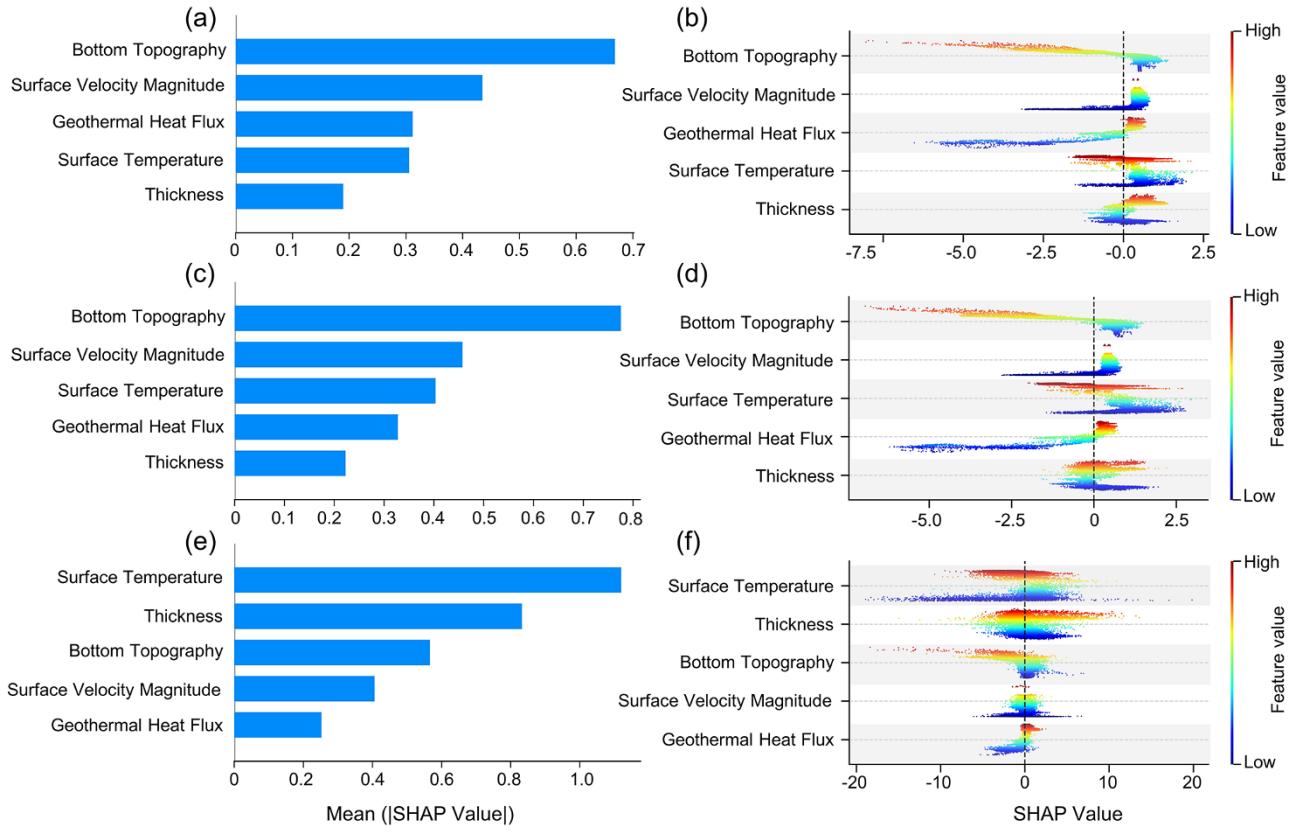


Figure 5. The mean of absolute SHAP values of each contributor (left column) and distribution of SHAP values (right column) in the emulator RF (a, b), XGBoost (c, d) and NN (e, f). The larger the SHAP values magnitude, the larger contribution to the predicted basal ice temperature. A positive/negative SHAP value means that the factor tends to increase/decrease the predicted basal ice temperature. Each cluster of scattered points ($N = 23\,240$) corresponds to the distribution of SHAP values for the specific feature, where the points are vertically jittered slightly around a dotted horizontal line to avoid overlap and facilitate the visualization of data density. We adopt alternating gray-white background shades in the right column to enhance the visual separation between different features.

is highly skewed, with a large fraction of the domain at or near the pressure-melting point. This skewed distribution limits the ability of data-driven emulators to effectively learn from lower-temperature samples (Ribeiro and Moniz, 2020; Dolar and others, 2025). While applying weighted balancing techniques to different samples may address this issue, such an approach may not universally improve prediction performance across all scenarios. Additionally, the linear regression model emulator demonstrates poor performance on both the training and test datasets. This underperformance could be attributed to two primary factors: (1) the absence of a linear relationship between input features and basal ice temperature; (2) the skewed distribution of basal ice temperatures in Totten Glacier violates the linear regression model's assumption of normality. These combined factors likely contribute to its suboptimal predictive capabilities.

Our analysis of training set size reveals that the RF emulator was the most robust when data were limited, consistently outperforming the other emulators with small training sets. This accords with existing literature suggesting that RF's bagging approach is highly resistant to overfitting, making it a reliable choice when data are scarce (Moghaddam and others, 2020; Ramezan and others, 2021). Perhaps more importantly, we found that emulator performance is sensitive not just to the quantity of training data, but to its quality and diversity. The inclusion of the An and others (2015) simulation consistently improved the performance of all emulators. As shown in Fig. 4 and Supplementary Fig. S6, this dataset provides a uniquely comprehensive sample of the glacier's thermal state, spanning both warm and cold regimes. Its inclusion likely provides the emulator with a more complete and representative picture of the system's physics. In contrast, the inclusion of the Purucker (2012) dataset decreased the R^2 scores for RF and XGBoost in the test set. This suggests that basal ice temperature data from this dataset—particularly in the southwestern region of Totten Glacier—exhibit poor alignment with the other datasets, likely attributable to its lower simulated ice temperatures (Huang and others, 2024). This underscores that for developing robust ML emulators, selecting a diverse set of training simulations that covers the full range of expected behaviors is as important as the ML architecture itself.

While all three ML emulators achieve high predictive accuracy, their internal mechanisms differ, leading to a divergence in what they appear to have learned about the glacier's physics. Tree-based ensemble emulators, whether they employ the bagging strategy of RF or the boosting strategy of XGBoost, are fundamentally built on decision trees. This architecture excels at learning piecewise, non-linear relationships, making it adept at capturing phenomena governed by thresholds. In contrast, NN learns through a series of linear transformations and nonlinear activation functions. While NNs are universal function approximators, the structure inherently favors learning smooth, continuous functions. This architectural difference is reflected in the feature importance results from our SHAP analysis. For both the RF and XGBoost emulators, bottom topography emerges as the dominant factor, exhibiting a consistent

Downloaded from https://www.cambridge.org/core. 18 Nov 2026 at 04:51:36, subject to the Cambridge Core terms of use.

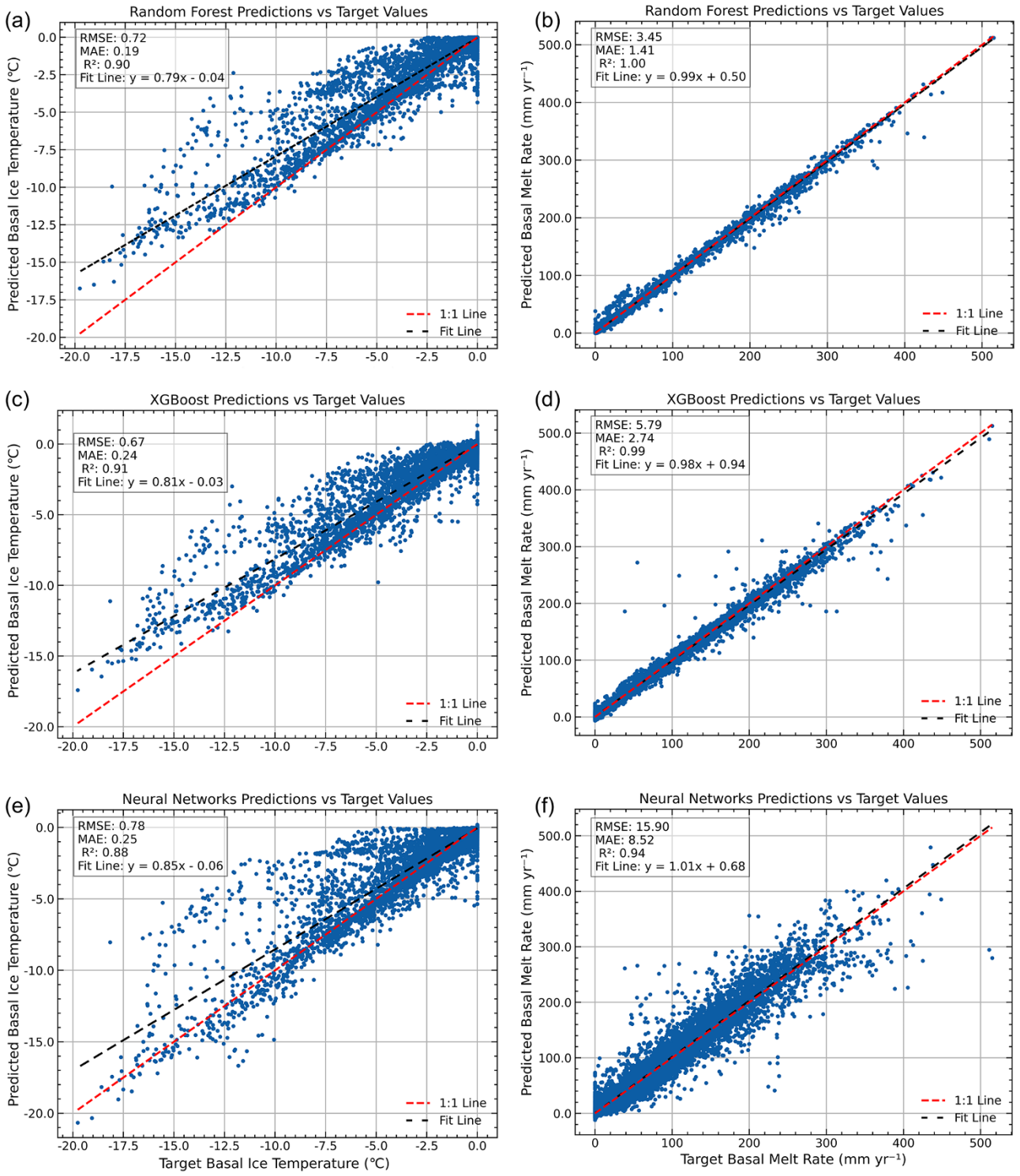


Figure 6. Simultaneous prediction performance of both basal ice temperature and melt rate on the test set of RF emulator (a, b), XGBoost emulator (c, d) and NN emulator (e, f), $N = 23\,240$.

negative contribution to basal ice temperature. This relationship is physically grounded: given the generally smooth glacier surface, higher basal topography corresponds to thinner overlying ice. Thinner ice provides less insulation from cold surface temperatures, leading to a cooler ice base (Fig. 6). Bottom topography demonstrates a greater influence than ice thickness in RF and XGBoost emulators, but a smaller influence than ice thickness in the NN emulator. Consistent across all emulators, GHF exerts a positive contribution to basal temperature. Similarly, surface ice speed is an important predictor, particularly in the RF and XGBoost emulators. Its positive contribution is well-understood: higher surface velocities are often associated with greater heat

Downloaded from https://www.cambridge.org/core. 18 Nov 2026 at 04:51:36, subject to the Cambridge Core terms of use.

dissipation from ice deformation and sliding, which can contribute to higher basal temperatures.

The feature rankings reveal that the emulators have developed different physical sensitivities. The tree emulators (RF and XGBoost) prioritize features driving local, dynamic heat production at the ice-bed interface: basal topography (which influences pressure and stress concentrations) and ice velocity (which generates friction). Conversely, the NN assigns greater importance to large-scale predictor variables such as surface temperature and ice thickness, which are associated with broad thermal structure in the modeled outputs. This suggests that the NN appears to capture a lower-complexity statistical mapping that is more strongly influenced by globally predictive variables. In contrast, the tree-based emulators appear to rely more strongly on predictors associated with local variability, leading to a different but still physically interpretable predictor ranking. Interestingly, the positive and negative contributions of the variables to basal ice temperature are largely consistent, although there are differences in the ranking of the most influential variables. It is noted that some of the predictor variables are correlated, and this can influence SHAP-based attribution by redistributing importance among related features. For example, geometric variables such as bedrock elevation and ice thickness are not fully independent, so part of the importance attributed to one variable may reflect shared information with another correlated predictor. In the present study, the differing feature sensitivities of the three emulator types suggest that relying on a single pointwise emulator may provide only a partial view of the modeled relationships. For the specific model classes considered here, comparing multiple architectures can therefore provide a more balanced interpretation when combined with domain expertise.

We further demonstrate that ML emulators can emulate the simultaneous prediction of both basal temperature and melt rate, like ice-sheet simulations. This is not merely a technical exercise; it serves a vital scientific purpose. Totten Glacier has extensive warm-bed areas where the basal ice temperature reaches the pressure-melting point. In contrast, the basal melting rate varies significantly when the basal ice temperature reaches pressure-melting point, which can thus demonstrate the prediction performance of the ML emulator across the warm-bed areas to a certain extent. From the simultaneous prediction results, both RF and XGBoost emulators accurately predict basal melting rate without significantly compromising basal ice temperature predictions. This is due to the tree-based nature of these emulators, which can handle multiple targets simultaneously. This near-perfect agreement is also physically consistent with the expected energy budget: melt rates of order 100 mm yr^{-1}, as seen in Fig. 6, require latent heat fluxes of approximately 1 W m^{-2}, which are at least an order of magnitude larger than typical GHF values. This suggests that, for these relatively high melt rates, variations among GHF datasets make only a secondary contribution compared with dynamical heat sources such as basal friction and internal deformation. Thus, variations among the geothermal datasets have only a minor effect on the simulated basal melting rates. The two target variables were standardized before training in the simultaneous-prediction experiments. The NN's temperature predictions are slightly degraded in the simultaneous-prediction setting, indicating that greater care is required when training a shared model for multiple outputs. In the present implementation, we set the loss weighting ratio of the two target variables to 1:2 (basal temperature to melt rate), which is a reasonable empirical choice. This configuration produced acceptable results, but it should not be regarded as uniquely optimal, and further improvements may be possible using more advanced multitask training strategies. We nevertheless retained the shared multi-output architecture because our aim was to test whether a single emulator could simultaneously reproduce multiple basal thermal quantities generated together by the underlying ice-sheet model, and to potentially capture their statistical interdependencies. In the simultaneous prediction emulator, the surface ice velocity is the main factor affecting the basal melting rate, which is reflected in the emulator's SHAP calculation (Supplementary Fig. S7). This is because the surface ice velocity and basal ice velocity are often positively correlated, and higher basal ice velocity will produce higher frictional heat, which is the primary driver of basal melting in Totten Glacier.

The transferability of these emulators remains an important caveat. The emulators in this study learn a specific mapping from a small set of steady-state variables to basal temperatures under a specific model configuration (Huang and others, 2024). They cannot be assumed to be directly applicable to other glacier systems. Furthermore, the emulators are constructed in a pointwise manner without incorporating spatial context. Consequently, the emulators are best interpreted as learning local statistical relationships within Totten Glacier, and are unlikely to represent processes governed by long-range physical connections. As an additional sensitivity analysis, we performed a spatial-block validation experiment to assess the potential influence of spatial autocorrelation and nonuniform mesh density on node-wise performance metrics. This analysis was conducted using only the six development GHF scenarios, while the Purucker (2012) and Martos and others (2017) scenarios remained reserved for the main withheld-GHF scenario evaluation. Spatial blocks were generated from the common mesh-node coordinates and grouped into five spatial cross-validation folds. This spatial-block validation was not used for hyperparameter tuning or model selection, but served as a conservative sensitivity test of model generalization across both geographic space and associated feature-space regimes, that is, spatially distinct combinations of physical conditions. The spatial-block validation produced lower but still moderate predictive performance compared with random node-based validation, with R^2 values of approximately 0.7 for multiple ML emulators (Supplementary Fig. S8). This result supports the interpretation that the emulators show reduced generalization across spatially distinct physical regimes than across randomly mixed nodes. Nevertheless, the retained performance under spatial-block validation suggests that the emulators capture meaningful local statistical relationships in the input variables. These results do not alter the main model comparison, which is based on the independent withheld GHF scenarios, but provide a more conservative assessment of spatial generalization.

This work demonstrates the potential of ML emulators in glaciology but also highlights areas for future refinement. Future work should also explore techniques to address the data imbalance issue, such as applying different weights in the loss function to force the emulator to pay more attention to the rare, cold-based samples. For multi-output prediction with NNs, further investigation into optimal loss function weighting is warranted. By addressing these limitations, future ML emulators can become more reliable tools that mitigate the high computational costs and other limitations associated with ice-sheet dynamic models.

5. Conclusion

In this study, we demonstrate that ML can create computationally efficient and accurate emulators of a sophisticated, full-Stokes ice dynamics model for predicting basal ice temperature across

Downloaded from https://www.cambridge.org/core. 18 Nov 2026 at 04:51:36, subject to the Cambridge Core terms of use.

the Totten Glacier basin. All three ML emulators (RF, XGBoost and NN) substantially outperform a traditional linear regression, confirming that the physical relationships governing basal temperature are inherently complex and nonlinear. The XGBoost emulator yields the highest overall accuracy ($R^2 \approx 0.91$). The emulators identify physically meaningful drivers. The SHAP analysis identifies basal topography as the dominant control on basal ice temperature and surface ice velocity as the primary driver of basal melt rate. We establish that a minimum of five diverse numerical simulations is sufficient to train a robust emulator that closely reproduces the basal ice temperature simulated by the ice dynamics model for Totten Glacier. Furthermore, the emulators can simultaneously predict basal temperature and melt rate, providing a diagnostic for warm-based ice where temperature alone is an insufficient metric. Our interpretability analysis reveals that the tree-based emulators (RF and XGBoost) capture localized statistical relationships by prioritizing local, dynamic features at the ice-bed interface, such as friction (proxied by surface velocity) and insulation effects from basal topography. In contrast, the NN acts like a lower-complexity statistical emulator, prioritizing globally predictive inputs like surface temperature and overall ice thickness. This finding underscores that a multi-emulator approach can offer a more balanced physical interpretation for the specific model classes considered here. The primary implication of this work is that the ML emulation framework provides a powerful and computationally inexpensive complementary approach to running large, computationally prohibitive ensembles of numerical simulations for sensitivity analysis and uncertainty quantification. While the trained emulators presented here are specific to Totten Glacier, the methodology itself is transferable to other glacial systems provided the models are retrained and reevaluated. It offers a promising pathway for conducting sensitivity analyses and uncertainty quantification for complex glaciological systems within a diagnostic framework. Ultimately, we emphasize that while the superior predictive performance of ML methods can help mitigate the high computational cost of numerical simulations, their interpretation must be combined with deep domain expertise to produce robust and physically meaningful scientific insights.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/jog.2026.10170.

Data availability statement. MEaSUREs BedMachine Antarctica, version 2, is available at https://doi.org/10.5067/E1QL9HFQ7A8M (Morlighem and others, 2020). MEaSUREs InSAR-Based Antarctic Ice Velocity Map, version 2, is available at https://doi.org/10.5067/D7GK8F5J8M8R (Rignot and others, 2017). MEaSUREs Antarctic Boundaries for IPY 2007–2009 from Satellite Radar, version 2, is available at https://doi.org/10.5067/AXE4121732AD (Mouginot and others, 2017). ALBMAP v1 and the GHF dataset of Shapiro and Ritzwoller (2004) are available at https://doi.org/10.1594/PANGAEA.734145 (Le Brocq, Payne, Vieli, 2010b). The GHF dataset of An and others (2015) is available at http://www.seismolab.org/model/antarctica/lithosphere/AN1-HF.tar.gz (last access: 11 April 2023). The GHF dataset of Shen and others (2020) is available at https://sites.google.com/view/weisen/research-products?authuser=0 (last access: 11 April 2023). The GHF dataset of Martos and others (2017) is available at https://doi.org/10.1594/PANGAEA.882503. The GHF dataset of Purucker (2012) is available at https://core2.gsfc.nasa.gov/research/purucker/heatflux_mf7_foxmaule05.txt (last access: 11 April 2023).

Funding statement. This work was supported by the National Natural Science Foundation of China (grant no. 42576280) and the Finnish Academy (grant no. 355572).

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

References

An M and 8 others (2015) Temperature, lithosphere-asthenosphere boundary, and heat flux beneath the Antarctic plate inferred from seismic velocities. *Journal of Geophysical Research: Solid Earth* **120**(12), 8720–8742. doi:10.1002/2015JB011917.

Anilkumar R, Bharti R, Chutia D and Aggarwal SP (2023) Modelling point mass balance for the glaciers of the Central European Alps using machine learning techniques. *The Cryosphere* **17**(7), 2811–2828. doi:10.5194/tc-17-2811-2023.

Bolibar J, Rabatel A, Gouttevin I, Galiez C, Condom T and Sauquet E (2020) Deep learning applied to glacier evolution modelling. *The Cryosphere* **14**(2), 565–584. doi:10.5194/tc-14-565-2020.

Breiman L (2001) Random forests. *Machine Learning* **45**(1), 5–32. doi:10.1023/A:1010933404324.

Budd WF, Jenssen D and Coutts B (1994) Modelling the Antarctic ice-sheet changes through time. *Annals of Glaciology* **20**, 291–297. doi:10.3189/1994AoG20-1-291-297.

Burton-Johnson A, Dziadek R and Martin C (2020) Review article: Geothermal heat flow in Antarctica: Current and future directions. *The Cryosphere* **14**(11), 3843–3873. doi:10.5194/tc-14-3843-2020.

Chen T and Guestrin C (2016) XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM. doi:10.1145/2939672.2939785

Dawson EJ, Schroeder DM, Chu W, Mantelli E and Seroussi H (2022) Ice mass loss sensitivity to the Antarctic ice sheet basal thermal state. *Nature Communications* **13**(1), 4957. doi:10.1038/s41467-022-32632-2.

DeConto RM and Pollard D (2016) Contribution of Antarctica to past and future sea-level rise. *Nature* **531**(7596), 591–597. doi:10.1038/nature17145.

Dolar T, Chen J and Chen W (2025) Uncertainty quantification driven machine learning for improving model accuracy in imbalanced regression tasks. *Expert Systems with Applications* **261**, 125526. doi:10.1016/j.eswa.2024.125526.

Dongare AD, Kharde RR and Kachare AD (2012) Introduction to artificial neural network. *International Journal of Engineering and Innovative Technology* **2**(1), 189–194.

Dow CF, McCormack FS, Young DA, Greenbaum JS, Roberts JL and Blankenship DD (2020) Totten Glacier subglacial hydrology determined from geophysics and modeling. *Earth and Planetary Science Letters* **531**, 115961. doi:10.1016/j.epsl.2019.115961.

Ducournau A and Fablet R (2016) Deep learning for ocean remote sensing: An application of convolutional neural networks for super-resolution on satellite-derived SST data. In *2016 9th IAPR Workshop on Pattern Recognition in Remote Sensing (PRRS)* (pp. 1–6). IEEE. doi:10.1109/PRRS.2016.7867019

Dutton A and 8 others (2015) Sea-level rise due to polar ice-sheet mass loss during past warm periods. *Science* **349**(6244), aaa4019. doi:10.1126/science.aaa4019.

Dziadek R, Ferraccioli F and Gohl K (2021) High geothermal heat flow beneath Thwaites Glacier in West Antarctica inferred from aeromagnetic data. *Communications Earth & Environment* **2**(1), 162. doi:10.1038/s43247-021-00242-3.

Dziadek R, Gohl K, Diehl A and Kaul N (2017) Geothermal heat flux in the Amundsen Sea sector of West Antarctica: New insights from temperature measurements, depth to the bottom of the magnetic source estimation, and thermal modeling. *Geochemistry Geophysics Geosystems* **18**(7), 2657–2672. doi:10.1002/2016GC006755.

Engelhardt H (2004) Thermal regime and dynamics of the West Antarctic ice sheet. *Annals of Glaciology* **39**, 85–92. doi:10.3189/172756404781814203.

Fischer H and 29 others (2013) Where to find 1.5 million yr old ice for the IPICS "oldest-ice" ice core. *Climate of the Past* **9**(6), 2489–2505. doi:10.5194/cp-9-2489-2013.

Haeger C, Petrunin AG and Kaban MK (2022) Geothermal heat flow and thermal structure of the Antarctic lithosphere. *Geochemistry Geophysics Geosystems* **23**(10), e2022GC010501. doi:10.1029/2022GC010501.

Downloaded from https://www.cambridge.org/core. 18 Nov 2026 at 04:51:36, subject to the Cambridge Core terms of use.

Hornik K (1991) Approximation capabilities of multilayer feedforward networks. *Neural Networks* **4**(2), 251–257. doi:10.1016/0893-6080(91)90009-T.

Huang Y, Zhao L, Wolovick M, Ma Y and Moore JC (2024) Using specularity content to evaluate eight geothermal heat flow maps of Totten Glacier. *The Cryosphere* **18**(1), 103–119. doi:10.5194/tc-18-103-2024.

Jiang G, Xu J and Wei J (2018) A deep learning algorithm of neural network for the parameterization of typhoon-ocean feedback in typhoon forecast models. *Geophysical Research Letters* **45**(8), 3706–3716. doi:10.1002/2018GL077004.

Jouvet G, Cordonnier G, Kim B, Lüthi M, Vieli A and Aschwanden A (2022) Deep learning speeds up ice flow modelling by several orders of magnitude. *Journal of Glaciology* **68**(270), 651–664. doi:10.1017/jog.2021.120.

Kang H, Zhao L, Wolovick M and Moore JC (2022) Evaluation of six geothermal heat flux maps for the Antarctic Lambert–Amery glacial system. *The Cryosphere* **16**(9), 3619–3633. doi:10.5194/tc-16-3619-2022.

Kyrke-Smith TM, Katz RF and Fowler AC (2014) Subglacial hydrology and the formation of ice streams. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **470**(2161), 20130494. doi:10.1098/rspa.2013.0494.

Le Brocq AM, Payne AJ and Vieli A (2010) An improved Antarctic dataset for high resolution numerical ice sheet models (ALBMAP v1). *Earth System Science Data* **2**(2), 247–260. doi:10.5194/essd-2-247-2010.

Li X, Rignot E, Morlighem M, Mouginot J and Scheuchl B (2015) Grounding line retreat of Totten Glacier, East Antarctica, 1996 to 2013. *Geophysical Research Letters* **42**(19), 8049–8056. doi:10.1002/2015GL065701.

Li X, Rignot E, Mouginot J and Scheuchl B (2016) Ice flow dynamics and mass loss of Totten Glacier, East Antarctica, from 1989 to 2015. *Geophysical Research Letters* **43**(12), 6366–6373. doi:10.1002/2016GL069173.

Ligtenberg SRM, van de Berg WJ, van den Broeke MR, Rae JGL and van Meijgaard E (2013) Future surface mass balance of the Antarctic ice sheet and its influence on sea level change, simulated by a regional atmospheric climate model. *Climate Dynamics* **41**(3–4), 867–884. doi:10.1007/s00382-013-1749-1.

Liu Z, Koo Y and Rahnemoonfar M (2024, April 30). Physics-informed machine learning on polar ice: A survey. [Preprint]. arXiv. http://arxiv.org/abs/2404.19536

Lösing M and Ebbing J (2021) Predicting geothermal heat flow in Antarctica with a machine learning approach. *Journal of Geophysical Research: Solid Earth* **126**(6), e2020JB021499. doi:10.1029/2020JB021499.

Lou R, Lv Z, Dang S, Su T and Li X (2023) Application of machine learning in ocean data. *Multimedia Systems* **29**(3), 1815–1824. doi:10.1007/s00530-020-00733-x.

Lu Y, Zhang Z, Shangguan D and Yang J (2021) Novel machine learning method integrating ensemble learning and deep learning for mapping debris-covered Glaciers. *Remote Sensing* **13**(13), 2595. doi:10.3390/rs13132595.

Lundberg SM and Lee SI (2017) A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems* **30**, 4765–4774.

MacGregor JA and 7 others (2022) GBaTSv2: A revised synthesis of the likely basal thermal state of the Greenland ice sheet. *The Cryosphere* **16**(8), 3033–3049. doi:10.5194/tc-16-3033-2022.

MacGregor JA and 11 others (2016) A synthesis of the basal thermal state of the Greenland Ice Sheet. *Journal of Geophysical Research: Earth Surface* **121**(7), 1328–1350. doi:10.1002/2015JF003803.

Marçais J and De Dreuzy J (2017) Prospective interest of deep learning for hydrological inference. *Groundwater* **55**(5), 688–692. doi:10.1111/gwat.12557.

Martos YM, Catalán M, Jordan TA, Golynsky A, Golynsky D, Eagles G and Vaughan DG (2017) Heat flux distribution of Antarctica unveiled. *Geophysical Research Letters* **44**(22), 11,417–11,426. doi:10.1002/2017GL075609.

Matsuoka K, MacGregor JA and Pattyn F (2012) Predicting radar attenuation within the Antarctic ice sheet. *Earth and Planetary Science Letters* **359–360**, 173–183. doi:10.1016/j.epsl.2012.10.018.

Moghaddam DD and 8 others (2020) The effect of sample size on different machine learning models for groundwater potential mapping in mountain bedrock aquifers. *CATENA* **187**, 104421. doi:10.1016/j.catena.2019.104421.

Mohajerani Y, Jeong S, Scheuchl B, Velicogna I, Rignot E and Milillo P (2021) Automatic delineation of glacier grounding lines in differential interferometric synthetic-aperture radar data using deep learning. *Scientific Reports* **11**(1), 4992. doi:10.1038/s41598-021-84309-3.

Morlighem M and 36 others (2020) Deep glacial troughs and stabilizing ridges unveiled beneath the margins of the Antarctic ice sheet. *Nature Geoscience* **13**(2), 132–137. doi:10.1038/s41561-019-0510-8.

Park I-W, Jin EK, Morlighem M and Lee K-K (2024) Impact of boundary conditions on the modeled thermal regime of the Antarctic ice sheet. *The Cryosphere* **18**(3), 1139–1155. doi:10.5194/tc-18-1139-2024.

Paterson WSB (1994) *The Physics of Glaciers*. Butterworth-Heinemann.

Pattyn F (2010) Antarctic subglacial conditions inferred from a hybrid ice sheet/ice stream model. *Earth and Planetary Science Letters* **295**(3–4), 451–461. doi:10.1016/j.epsl.2010.04.025.

Pattyn F and Morlighem M (2020) The uncertain future of the Antarctic Ice Sheet. *Science* **367**(6484), 1331–1335. doi:10.1126/science.aaz5487.

Purucker M (2012) Geothermal heat flux data set based on low resolution observations collected by the CHAMP satellite between 2000 and 2010, and produced from the MF-6 model following the technique described in Fox Maule et al. (2005), Interactive System for Ice sheet Simulation [Data set]. https://core2.gsfc.nasa.gov/research/purucker/heatflux_mf7_foxmaule05.txt

Ramezan CA, Warner TA, Maxwell AE and Price BS (2021) Effects of training set size on supervised machine-learning land-cover classification of large-area high-resolution remotely sensed data. *Remote Sensing* **13**(3), 368. doi:10.3390/rs13030368.

Reading AM and 8 others (2022) Antarctic geothermal heat flow and its implications for tectonics and ice sheets. *Nature Reviews Earth & Environment* **3**(12), 814–831. doi:10.1038/s43017-022-00348-y.

Rhee J and Im J (2017) Meteorological drought forecasting for ungauged areas based on machine learning: Using long-range climate forecast and remote sensing data. *Agricultural and Forest Meteorology* **237–238**, 105–122. doi:10.1016/j.agrformet.2017.02.011.

Ribeiro RP and Moniz N (2020) Imbalanced regression and extreme value prediction. *Machine Learning* **109**(9–10), 1803–1835. doi:10.1007/s10994-020-05900-9.

Rignot E, Mouginot J and Scheuchl B (2017) *MEaSUREs InSAR-Based Antarctica Ice Velocity Map, Version 2*. NASA National Snow and Ice Data Center Distributed Active Archive Center [Data set]. doi:10.5067/D7GK8F5J8M8R.

Rignot E (2006) Changes in ice dynamics and mass balance of the Antarctic ice sheet. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **364**(1844), 1637–1655. doi:10.1098/rsta.2006.1793.

Schoof C (2007) Ice sheet grounding line dynamics: Steady states, stability, and hysteresis. *Journal of Geophysical Research: Earth Surface* **112**(F3), 2006JF000664. doi:10.1029/2006JF000664.

Sergienko O, Haseloff M, Robel A and Wingham D (2026) A new paradigm for understanding earth's marine ice sheets. *Nature Geoscience* **19**(4), 374–383. doi:10.1038/s41561-026-01941-2.

Seroussi H, Ivins ER, Wiens DA and Bondzio J (2017) Influence of a West Antarctic mantle plume on ice sheet basal conditions. *Journal of Geophysical Research: Solid Earth* **122**(9), 7127–7155. doi:10.1002/2017JB014423.

Seroussi H and 38 others (2019) initMIP-Antarctica: An ice sheet model initialization experiment of ISMIP6. *The Cryosphere* **13**(5), 1441–1471. doi:10.5194/tc-13-1441-2019.

Shapiro NM and Ritzwoller MH (2004) Inferring surface heat flux distributions guided by a global seismic model: Particular application to Antarctica. *Earth and Planetary Science Letters* **223**(1–2), 213–224. doi:10.1016/j.epsl.2004.04.011.

Shen W, Wiens DA, Lloyd AJ and Nyblade AA (2020) A geothermal heat flux map of Antarctica empirically constrained by seismic structure. *Geophysical Research Letters* **47**(14), e2020GL086955. doi:10.1029/2020GL086955.

Downloaded from https://www.cambridge.org/core. 18 Nov 2026 at 04:51:36, subject to the Cambridge Core terms of use.

Stål T, Reading AM, Halpin JA and Whittaker JM (2021) Antarctic geothermal heat flow model: Aq1. *Geochemistry Geophysics Geosystems* **22**(2), e2020GC009428. doi:10.1029/2020GC009428.

Umlauft J and 13 others (2023) Mapping glacier basal sliding applying machine learning. *Journal of Geophysical Research: Earth Surface* **128**(8), e2023JF007280. doi:10.1029/2023JF007280.

van der Meer M, Zekollari H, Huss M, Bolibar J, Sjursen KH and Farinotti D (2025) A minimal machine-learning glacier mass balance model. *The Cryosphere* **19**(2), 805–826. doi:10.5194/tc-19-805-2025.

Werder MA, Huss M, Paul F, Dehecq A and Farinotti D (2020) A Bayesian ice thickness estimation model for large-scale applications. *Journal of Glaciology* **66**(255), 137–152. doi:10.1017/jog.2019.93.

Wolovick MJ, Moore JC and Zhao L (2021) Joint inversion for surface accumulation rate and geothermal heat flow from ice-penetrating radar observations at Dome A, East Antarctica. Part II: Ice sheet state and geophysical analysis. *Journal of Geophysical Research: Earth Surface* **126**(5), e2020JF005936. doi:10.1029/2020JF005936.

Xie Z, Asari VK and Haritashya UK (2021) Evaluating deep-learning models for debris-covered glacier mapping. *Applied Computing and Geosciences* **12**, 100071. doi:10.1016/j.acags.2021.100071.

Xu T and Liang F (2021) Machine learning for hydrologic sciences: An introductory overview. *WIREs Water* **8**(5), e1533. doi:10.1002/wat2.1533.

Zhang E, Liu L and Huang L (2019) Automatically delineating the calving front of Jakobshavn Isbræ from multitemporal TerraSAR-X images: A deep learning approach. *The Cryosphere* **13**(6), 1729–1741. doi:10.5194/tc-13-1729-2019.

Downloaded from https://www.cambridge.org/core. 18 Nov 2026 at 04:51:36, subject to the Cambridge Core terms of use.