Evaluating Machine Learning and Statistical Models for Greenland Bed Topography

REU Site: Online Interdisciplinary Big Data Analytics in Science and Engineering

Katherine Yi[1], Angelina Dewar[2], Tartela Tabassum[3], Jason Lu[4], Ray Chen[5], Homayra Alam[3], Omar Faruque[3], Sikan Li[6], Mathieu Morlighem[7], Jianwu Wang[3]

[1]Department of Computer Science, Purdue University
[2]Department of Physics, University of Oregon
[3]Department of Information Systems, University of Maryland, Baltimore County
[4]Department of Information Studies, University of Maryland, College Park
[5]Marriotts Ridge High School, Maryland
[6]Texas Advanced Computing Center, University of Texas at Austin
[7]Department of Earth Sciences, Dartmouth College

Technical Report HPCF–2023–11, hpcf.umbc.edu > Publications

Abstract

The purpose of this research is to study how different machine learning and statistical models can be used to predict bed topography in Greenland using ice-penetrating radar and satellite imagery data. Accurate bed topography representations are crucial for understanding ice sheet stability, melt, and vulnerability to climate change. We explored nine predictive models including dense neural network, LSTM, variational auto-encoder (VAE), extreme gradient boosting (XGBoost), gaussian process regression, and kriging based residual learning. Model performance was evaluated with mean absolute error (MAE), root mean squared error (RMSE), coefficient of determination (R^2), and terrain ruggedness index (TRI). In addition to testing various predictive models, different interpolation methods, including Nearest Neighbor interpolation, Bilinear Interpolation, and Universal Kriging were used to obtain estimates the values of ice surface features at the ice bed observation locations. The XGBoost model with Universal Kriging interpolation exhibited strong predictive capabilities but demands extensive resources. Alternatively, the XGBoost model with bilinear interpolation showed robust predictive capabilities and required fewer resources. These models effectively captured the complexity of the Greenland ice sheet terrain with precision and efficiency, making them valuable tools for representing spatial patterns in diverse landscapes.

Key Words Greenland Topography, Interpolation, Machine Learning, Predictive Modeling

1 Introduction

Knowing the topography of glacial ice beds is critical for modeling glacial changes accurately as climate change continues. However, measuring ice bed elevation directly entails significant costs, as it necessitates radar technology to scan through the ice and measure the underlying bedrock. In contrast, surface glacier data can be readily obtained through satellite observations. Consequently, researchers are exploring models capable of estimating glacial bedrock topography from surface data. Prior efforts in this domain include the physics-based model BedMachine by Morlighem et al. [13], which predicts ice bed topography in Greenland, and the deep neural network-based DeepBedMap by Leong and Horgan [8], designed for Antarctic bed topography prediction.

To comprehensively understand the efficacy of diverse machine learning and statistical models in predicting ice bed topography, in this paper, we evaluated nine distinct models. Our exploration included statistical techniques such as gaussian process regression and universal kriging,as well as machine learning approaches such as variational autoencoder (VAE), XGBoost, VAE with XGBoost, dense layer-based neural network, and LSTM-based neural network. Besides these standalone models, we also utilized hybrid models that combine two or more standalone models using techniques such as residual learning. These hybrid models included Dense + LSTM, VAE + XGBoost, and Universal Kriging + XGBoost.

- The standalone XGBoost model yielded the best test metrics and the most detailed and realistic topographic map predictions out of the models presented in this study and the physics-based model BedMachine [13].

- Universal Kriging showed superior metrics to XGBoost for many regions of the dataset as shown in 8.4, but failed to converge properly for some pixels, making extremely poor predictions on pixels in those regions.

- To improve data integration and leverage physics-guided knowledge in preprocessing, we also conducted an ablation study to compare the effects of training our models on datasets interpolated using different approaches and the effect of incorporating additional derived features recommended based on known physics models. The results demonstrated that different interpolation techniques had varying effects on machine learning model performance, and the incorporation of our new derived features positively impacted machine learning performance.

- Besides evaluating the models on a test set with known ice bed elevation data, we also evaluated each model by making predictions on a region spanning 32,400 km^2 where true ice bed elevation has not been measured completely. Visualizing the predictions for the entire region provided a comprehensive understanding of the performance differences among the models. Furthermore, we found the terrain ruggedness index(TRI) [11]to be a valuable metric for quantifying these differences.

  In the following sections, we will delve into the details of previous research in this field, provide a brief explanation of our data sources and preprocessing techniques, and describe the architecture of the models we employed. Our results are then presented, and compared with those derived from the physics model in BedMachine [13].

2 Literature Review

Morlighem et al [13] produced a physics-based model that can predict Greenland bedrock elevation using a mass conservation optimization scheme. Morlighem et evaluated their model on the same dataset we train and evaluate our models on, which consisted of heterogenous data collected from two main sources: time series satellite data describing five ice surface features, and measurements of ice bed elevation at erratic locations obtained via ice-penetrating radar. We compare the efficacy of our models to that of Morlighem et al by comparing test metrics, and visually inspecting the predictions of all models on a 32,400 squared km region where ground truth data is not fully known.

We also took inspiration from the work of Leong and Horgan [8], who also explored machine learning approaches to predicting ice bed topography from surface features. Specifically, they trained a convolutional neural network (CNN) to predict the ice bed topography of Antarctica. Leong and Horgan used CNNs with deep learning to produce predictions with high roughness for Antarctica using sparse but with high-resolution ground truth data.

3 Background and Data Preprocessing

Addressing data gaps in bed topography is crucial due to its significance in predicting future sea level rise. The shape of the bedrock beneath thick ice holds the key to understanding the behavior of ice sheets in response to climate change. Although glacier fjords in Greenland are now exposed due to ice retreat, measuring the landscape hidden beneath thousands of meters of ice remains challenging. Scientists have developed simulation models to predict sea level rise and ice sheet responses over the years. However, uncertainties persist in forecasting potential catastrophic collapses of ice sheets in the next 100-200 years. The rate of ice sheet mass loss depends heavily on the bedrock's topography, which can lead to either slow, gradual retreat or rapid, unstoppable degradation. Identifying retrograde regions with deepening beds and areas with bumps and ridges that may impede retreat is critical. Without accurate bed representations, precise sea level rise predictions are hindered. Understanding climate change requires crucial insights into the underlying ice bed, even as thick ice poses significant obstacles to direct observation.

3.1 Problem Definition

Our main objective is to train machine learning models to forecast ice bed elevation across a vast area using solely surface data as input because the available ground truth data for the area's actual ice bed elevation is limited. So we aim to devise a method to produce reliable predictions for the entire region using only surface data. Our analysis is centered on predicting bed topographic elevation at new locations where direct measurement is unavailable, thereby enhancing the depiction of Greenland's ice deposits. Because the geolocations of the surface variables are different from the ones for the target variable, we need to first match the two sets of data points. Below, we explain the datasets used in our study and data preprocessing tasks.

3.2 Data Sources

Specifically, we selected a 32,400 km^2 square area in Upernavik, West Greenland. By splitting this area into 1200x1200 uniform grids and each grid's size is 150m x 150m, we can obtain a dataset with 1,442,401 (1201x1201) total data points. Each data point provides values for five surface variables (see rows 1-5 in Table 3.1), which are original features to be leveraged by our predictive models. Because these are uniform grids, we refer to them as *Grid Data*. Meanwhile, we also obtained 632,706 data points for ice bed elevation measurement as target values, which are rows 6-7 in Table 3.1. These data points have ground truth that can be used to measure the accuracy of our predictive models. Because these data points are only along the tracks of the flown radar sensor, we refer to them as *Track Data*.

3.2.1 Ice Sheet Surface Measurement

We utilized five ice sheet surface measurements from various sources: 1) Surface Elevation is obtained from the Greenland Ice Mapping Project using Interferometry(GIMP)/Greenland Ice Mapping Project (GIMP) [15]; 2)Ice flow surface velocity data on both longitudinal and

Table 3.1: Description of Data Variables

Variable Name	Description
surf_x, surf_y	Coordinates of grids (m)
surf_vx, surf_vy	Ice flow velocity (m/yr)
surf_elv	Ice surface elevation (m)
surf_dhdt	Ice thinning rates annually (m/yr)
surf_SMB	Snow accumulation annually (m/yr)
track_bed_x, track_bed_y	Coordinates of track bed points (m)
track_bed_target	Ice bed elevation (m)

latitudinal directions is generated by integrating multiple satellite interferometry data products including Landsat-8, Sentinel-1, and RADARSAT-2 via the approach at [16]; 3)Ice thinning rates are provided by ICESat-2:(Ice, Cloud, and land Elevation Satellite-2) [17];4)Surface mass balance indicating annual snow accumulation is derived from RACMO (Regional Atmospheric Climate Model) [18]. Examples of surface measurement variables used in our study are illustrated in the left part of Figure 3.1.

3.2.2 Ice Bed Elevation Measurement

Bed topography measurements are acquired through ice-penetrating radar in NASA's Operation IceBridge [19]. A substantial radar system, mounted beneath an aircraft's wings, emits signals through the ice, but the data collection is confined to the region directly beneath the aircraft. The right part of Figure3.1 presents a visual representation of the dataset. In contrast to surface variable data that are uniformly distributed and show complete spatial patterns, our target variable data is irregularly distributed.

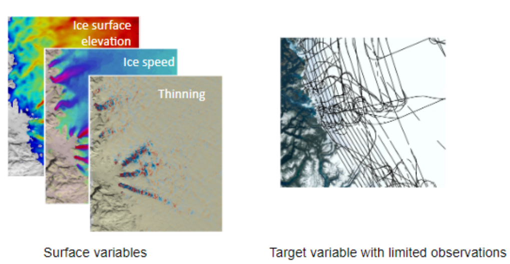

Figure 3.1: Surface variable grid data and target variable track data.

3.3 Data Preprocessing

The first data preprocessing we conducted is to estimate the values of the surface variables for each track data point based on their geolocations so every track data point has five feature variables and a target variable. It will end up with 632,706 data points for predictive models to be trained on. To achieve this goal, we employed three distinct interpolation techniques: nearest neighbor, bilinear, and universal kriging.

The first two interpolation approaches are straightforward. The first one, nearest neighbor, involved associating each radar bedrock observation with its spatially-nearest surface observation. The second approach, bilinear interpolation, predicted the value of each surface variable by calculating the weighted average of the four nearest neighbor observations of that variable. The weights of each neighbor were determined based on their distances from the prediction location.

The third interpolation technique in our study is universal kriging. Similar to bilinear interpolation, it estimated the values of surface variables using nearby observations. However, unlike bilinear interpolation, universal kriging incorporated all observations in the dataset(or a specified neighborhood due to memory constraints)to calculate a weighted average. The weights were assigned based on the autocorrelation of the observations, utilizing a variogram model. For each batch, universal kriging was fitted using different variogram models, selecting the model with the lowest capability ratio(CR). Throughout the process, an anisotropy scale of three and an anisotropy angle of ninety degrees were consistently applied to all surface variables, while the highest number of lags allowed for each batch was determined.

The application of all three interpolation methods to the original region of interest dataset yielded three distinct interpolated datasets. Each interpolated dataset comprises a total of 632,612 examples, with each example containing the interpolated values of the five surface variables, the true observed value of the bed elevation at the corresponding location, and the respective coordinates.

After obtaining our three interpolated datasets using the three interpolation methods, we conducted further preprocessing steps to prepare the data for our models. This involved dropping non-numerical values and calculating the magnitude of the ice flow velocity vector for each surface observation, which was then added as a derived feature to the dataset. Additionally, for the universal kriging and universal kriging + XGB models, we removed duplicates from the dataset to ensure proper convergence during the training process.

3.4 Assumptions

Having clarified the definitions of the fundamental data variables our focus now shifts to the underlying assumptions and methodologies involved in the data collection process. These assumptions play a pivotal role in determining the most appropriate interpolation and methods to employ in predicting Greenland's ice bed topography accurately. The first assumption is that all tools and data collected in the data collection from the Greenland Ice Sheet (GrIS) and all data sources we are using are dependable, relevant, and accurate. The second assumption is that the provided data is a normally distributed representation of the topography and does not only focus on one type of topography. For example, the known data includes a normal distribution of flat land, small mountains, and troughs which are common across Greenland. This assumption allows us to tune our models to generalize to unobserved data and derive conclusions of model accuracy. With a clear understanding of the assumptions guiding our data collection process, we now proceed to provide an overview of the methodology employed in this study.

4 Methodology Overview

In order to train a machine-learning model to make predictions at a location, it was necessary to match each bed height observation from Operation IceBridge to a set of surface feature observations. Interpolating datasets was necessary to merge these observations. Our study explored three different interpolation techniques: Nearest Neighbor (NNI), Bilinear Interpolation(BLI), and Universal Kriging(UK). Interpolated data was validated through visualization and comparison to Morlinghem's 2014 model. Next, the dataset underwent additional preprocessing to derive and select features as well as be prepared for models.

After the data completes preprocessing, it is ready for modeling. Numerous probabilistic, machine learning and deep learning models were developed and tuned for our study including Gaussian Process Regression, Spatio-Temporal Gaussian Processing, Variational Auto-Encoders, CNNs, XGBoosts, and residual learning which are described later in this paper. Models were then trained on the nearest neighbors dataset which provided a baseline for model selection. The selected model was trained on the different datasets generated from NNI, BLI, and UK.

Following training, each model was assessed individually with a supervised methodology. The model was used to predict test data, which was excluded from the training dataset. The predictions were then quantified and compared to known target bed heights using Root Mean Squared Error, Mean Average Error, Coefficient of Determination, and Terrain Ruggedness Index. In addition to numerical comparison, the predictions were visualized with validation data, a subsection of the Greenland dataset, and compared to Morlinghem's 2014 physics-based model again. Visualizations were supported by terrain ruggedness index calculations. Finally, metrics from each individual model were compared to determine model recommendations for future applications of modeling to topography predictions.

5 Preprocessing

In this section, we will outline decisions and the methodologies used to clean, normalize, and transform the raw data, ensuring that it is suitable for analysis and the subsequent application of machine learning and deep learning models.

5.1 Interpolation

As discussed in previous sections, our data comes from numerous sources. Interpolation must be leveraged to merge large surface variable datasets with the known target variable dataset. NNI, BLI, and UK were employed and tested; below is an explanation of each method.

5.1.1 Nearest Neighbors

Nearest neighbor interpolation is a simple image resizing technique used in computer graphics and digital image processing. Interpolation involves estimating the value of a function at a point where the value is not explicitly known, using the information from nearby points where the function's value is given. The nearest neighbor algorithm tackles this task by selecting the value of the closest known point without taking into account the values of other nearby points. As a result, it creates a piecewise-constant approximation.

Each data grid is separated by 150 meters. For each track bed point, the other independent features must be identified by their index in their own dataset. The index of the corresponding features was identified using the following formulas. The column was identified by taking the first point of surf x and subtracting it from all the points of track bed x and dividing it by 150. The row is identified by subtracting track bed y coordinates from surf y coordinates.

$$\frac{(track\_bed_x - surf_x[0,1])}{150m} = p \tag{5.1}$$

$$\frac{(surf_y[-1,0] - track\_bed_y)}{150m} = q \tag{5.2}$$

The (p,q) index is used to map ice flow, elevation, ice thinning, and snow accumulation to create the dataset.

5.1.2 Bilinear Interpolation

Bilinear interpolation is also a commonly used image resizing and interpolation technique in computer graphics and digital image processing. It is an improvement over the nearest neighbor interpolation method and provides smoother and more accurate results. In bilinear interpolation, it considers the values of the four nearest neighboring pixels from the original image to compute the value of a pixel in the new image. The method takes the weighted average of these four pixels to estimate the value of the target pixel. This interpolation is done on a regular or rectilinear grid in arbitrary dimensions. Our dataset is defined on a rectilinear grid; that is, a rectangular grid with even spacing. After setting up the interpolator object, the bilinear interpolation method is chosen at each evaluation.

The regular grid interpolation is done on gridded data of cell center coordinates on x and y axis, the axis that constrains all the other features together. This gridded function is then interpolated for track bed data.

5.1.3 Kriging Interpolation

In the data, all five surface variables we want to interpolate have local trends although not all of the local trends are quantified. Because of this, Ordinary Kriging and Simple Kriging are not appropriate for our dataset. Instead, this study uses Universal Kriging (UK).

Universal Kriging is a variant of the geospatial probabilistic interpolation algorithm, Kriging. Like Bilinear Interpolation (BLI), UK uses the values of the surface features at nearby locations to estimate their values at each bedrock observation location. However, it differs from BLI in two crucial ways. First, unlike BLI, UK does not treat all points equally. Kriging weights nearby points based on their correlations with each other. This motivates Kriging to accurately capture more complex patterns in spatial data than Bilinear Interpolation even when lower amounts of data are present. The second difference is that UK is probabilistic rather than a deterministic interpolation algorithm. The main consequence of the algorithm's probabilistic nature is it associates a quantification of certainty for each predicted point.

One challenge faced when implementing Kriging was the large requirement of resources required when scaled to our large, diverse dataset. Many modifications were made to the traditional UK. The first alteration was applying individual variograms within a unique dynamic square for each point to be predicted. Each square was resized to capture a specific range of points. The second adjustment was allowing each variogram model to fit a linear or logarithmic trendline to better fit the data regions based on error metrics. This method will be denoted as the localized cross-validation box method.

5.2 Additional Preprocessing

After interpolation has been performed on the raw data, we apply several additional preprocessing steps to the interpolated data in order to prepare our data for the models. The interpolated data requires preprocessing before modeling. Preprocessing involves deriving additional features, feature selection, scaling values, and randomized splitting.

To incorporate relevant domain knowledge into our dataset, we collaborated with Dr. Mathieu Morlighem, an Evans Family Professor of Earth Sciences at Dartmouth University and a domain expert in ice-sheet and sea-level systems. Dr. Morlighem reviewed the features present in our nearest neighbors dataset and emphasized the importance of the velocity magnitude of ice flow as an indicator of topography. Based on his feedback, we derived the ice flow velocity magnitude at each (x,y) coordinate by calculating the standard magnitude equation applied to the scalar values of ice flow in the x direction (vx) and y direction (vy). This derived feature, denoted as "v_mag," was included in our modeling.

Table 5.1: Evaluation Metrics with XGBoost Model

	RMSE	**MAE**	$\mathbf{R^2}$
Without velocity magnitude	33.016	22.663	0.966
With velocity magnitude	32.680	22.273	0.967

Adding the velocity magnitude feature, the metrics slightly improved. With more generalized parameters, the velocity magnitude features plays a more significant role in the metrics.

After deriving the velocity magnitude feature, we performed feature selection to identify relevant and insightful features that contribute to a clear final topography map and reliable metrics. The selected features for the final model include ice flow in the x-direction, ice flow in the y-direction, surface height, ice thinning rates, snow accumulation, and the newly derived velocity magnitude of ice flow at each coordinate.

Features such as cell center coordinates used for data interpolation were not selected because they caused poor topography map predictions and did not improve the prediction error significantly. Therefore, these features were excluded from the final feature set.

Next, we scaled the target values using StandardScaler from the sklearn preprocessing module to standardize. StandardScaler takes the points and subtracts the dataset mean and divides it by the standard deviation. This makes all features have a mean of approximately zero with a standard deviation of one. The data was scaled to allow (1) a balance of feature weights in training, (2) speed up the training speed and stability of training, and (3) more easily identify patterns in the data for tuning. All data was scaled together. StandardScalar was selected over MinMax Scaling because it is not as sensitive to outliers that exist in diverse real-world data.

The final step in preprocessing involved randomly splitting the data into training, testing, and validation datasets. We used the train_test_split function from the sklearn model selection module for this purpose. The randomization seed was set to 168 to ensure reproducibility. After experimenting with different ratio splits, we found that the best metrics were obtained with a 60% training, 40% testing, and 20% validation data split. The training data was further split to allocate the validation data, maintaining the same seed.

In summary, the preprocessing steps included deriving the velocity magnitude feature, selecting relevant features, scaling the data using StandardScaler, and randomizing the data into training, testing, and validation sets.

6 Modeling

Having successfully preprocessed the data and brought it to a standardized format suitable for analysis, we now proceed to the modeling stage. In this section, we will describe the probabilistic, machine learning, and deep learning models employed to predict ice bed height in Greenland's ice bed topography.

6.1 Kriging Residual Learning with XGBoost

In addition to interpolation, Kriging was applied as a form of residual learning using the XGBoost model to predict residuals. We first used Universal Kriging to make a first-pass prediction for ice bed height using the localized cross-validation box at every point. Predictions were used to calculate residuals, the error of the prediction. XGBoost was trained to predict the residual of Kriging's first pass prediction. The residual was then subtracted from the kriging first pass prediction to yield a final prediction.

6.2 Gaussian Process Regression

Gaussian Process Regression (GPR) is a powerful probabilistic regression technique that leverages Bayesian principles to compute predictive distributions. GPR works by optimizing the log-marginal-likelihood (LML) to determine the shape and smoothness of the predictive function. To implement GPR, we utilized the 'sklearn.gaussian_process' library and defined a kernel with a constant term and radial basis function. Despite its promising properties, GPR presented challenges in terms of computational resources and storage requirements similar to the challenges of kriging. Due to the high computational cost of training GPR on large datasets, we employed mini-batch processing to manage significant time and financial investments. Unfortunately, the model continued to demand an incredible number of resources for computation and continued the challenges of applying GPR to large-scale geographical datasets.

Gaussian Process Regression offered a probabilistic approach to predict topography data in Greenland, enabling us to gain valuable insights. While the model's capabilities were promising, its implementation demanded careful consideration of computational resources and the selection of appropriate subsets due to the substantial storage requirements. Because of the challenge to engage the entire topography, other models were pursued.

6.3 Spatio-Temporal Gaussian Process

We explore the Scalable Spatio-Temporal Gaussian Process (STGP) as a more efficient and accurate method for predicting topography data in Greenland. STGP enhances Gaussian process (GP) inference by combining spatial and time filtering and natural gradient variational inference. Despite not aiming to predict future landscapes explicitly, STGP's efficiency makes it relevant for analyzing current spatio-temporal data of the topography.

To fully understand STGP, we must understand variational inference (VI), a popular method in machine learning, plays a vital role in STGP. VI employs optimization techniques to estimate complex probability densities and converges faster than classical methods. By choosing a family of probability density functions and optimizing the Kullback-Leibler (KL) divergence, STGP leverages the advantages of GP while significantly reducing processing time.

As we explore the potential of STGP for predicting topography data, it is essential to acknowledge the computational challenges encountered during our experimentation with Gaussian Process Regression (GPR). It is important to note that when attempting to apply STGP to larger datasets, even in batches, we encountered similar memory issues. The resource demands for handling extensive spatio-temporal data remained a significant challenge. Addressing these memory constraints and optimizing the model's performance for larger datasets will be crucial for fully harnessing the potential of STGP in predicting topography data in Greenland.

Further tuning and allocation of additional computational resources have the potential to improve the results of STGP in predicting topography data for Greenland. However, to explore alternative avenues and mitigate the challenges of resource-intensive methods, we ventured into investigating other predictive models.

6.4 Variational Auto-Encoder

Variational Autoencoders (VAEs) have emerged as a significant and influential tool in machine learning, particularly in the domain of predicting topography data for Greenland. VAEs are designed to learn probabilistic representations of the geographical landscape by (1) reconstructing features as the encoder and (2) incorporating KL regularization to leverage existing knowledge, with the decoder aiming to accurately reconstruct the original input from the latent space.

In the VAE architecture tailored for topography data, the input representing Greenland's surface features is transformed into a probability distribution across a hidden space, and a point is sampled to reconstruct the original topography, minimizing the reconstruction error during training through a process known as backpropagation. The introduction of KL regularization ensures that the hidden space captures essential characteristics and aligns with known information about Greenland's geographical features.

To enhance the VAE's performance, a cyclical training schedule is employed, repeating a specific training pattern. This cyclic training process has proven to be effective, leading to improved representation learning and meaningful representations of Greenland's topography [20].

While VAEs present a promising approach for predicting topography data in Greenland, it's important to acknowledge that their success in this specific application may not be as high as desired. By reconstructing features and incorporating KL regularization, VAEs demonstrate their potential to capture essential information about the landscape and leverage prior knowledge. The use of cyclical training improves the model's performance, making VAEs valuable tools for analyzing and predicting geographical features not only in Greenland but also in other regions. However, further research and fine-tuning may be necessary to achieve even more accurate and reliable predictions in this challenging domain.

6.5 VAE & Extreme Gradient Boosting

The motivation behind using Variational Autoencoders (VAEs) as an encoder in combination with Extreme Gradient Boosting (XGB) as the predictor lies in their complementary strengths and the pursuit of understanding underlying patterns in Greenland topography. VAEs excel in learning meaningful and compact representations of complex data, making them ideal encoders for capturing essential features in the topographical landscape. Once trained, the VAE can encode the input data into a lower-dimensional latent space, preserving critical information while reducing the data's dimensionality. On the other hand, XGBoost has demonstrated exceptional predictive performance in the context of Greenland topography, showcasing its ability to handle non-linear relationships and complex interactions between geographical features. By employing XGB as the predictor on top of the VAE-encoded representations, we aim to leverage both models' strengths to achieve more accurate and interpretable predictions. This integrated approach fosters a deeper understanding of the underlying patterns in the topography, paving the way for more informed analysis and decision-making in studying the complex landscapes of Greenland.

6.6 MLP (Dense)

The dense layers are the most prevalent layers in deep neural networks, and they operate in a very straightforward manner. Each neuron in a thick layer receives input from every neuron in the preceding layer and aggregates its results nonlinearly. To make the MLP (Dense) model we employed three thick layers using ReLU, with a 50% dropout layer between each pair of dense layers. Finally, the prediction result is derived from the model's final layer, which contains only one neuron.

The traditional Dense model was adjusted to fit our project goals. First, the Adam optimizer was assigned to utilize the error gradient to determine the global optimum weights. The MSE loss function calculates the difference between the predicted and actual ice bed heights in the training dataset. By decreasing the MSE, the model learns to produce more accurate predictions about ice bed height. Additional parameters were put in place during training to ensure generalization ability in later performance. These parameters included random shuffling of data, 20,000 batch size, and 100 epochs of training.

Deep neural network design with Adam optimizer and dropout layers has been shown to be a reliable and efficient method for predicting ice bed height. The model exhibits improved generalization and lower prediction errors than previous methods by exploiting the momentum of the error gradient, optimizing with the MSE objective function, and 50% dropout regularization, making it a viable tool in ice bed height prediction applications.

6.7 LSTM

Motivated by the dataset's spatiotemporal elements, which included ice flow velocity measurements collected over several years, the team sought to harness the power of Long Short-Term Memory (LSTM) layers. LSTM layers were first proposed by Hochreiter and Schmidhuber in 1997. LSTM layers are well-suited for handling sequential data and have the capability to uncover hidden patterns that other types of layers might overlook. In our model, we initialized with three LSTM layers comprising 64, 32, and 16 memory cells followed by a 50% dropout layer. To arrive at the final prediction, the LSTM layers were connected to a dense layer. The Adam optimizer was chosen for its computational efficiency and effectiveness, making it ideal for processing the large datasets. Given the objective of predicting bed height as accurately as possible, mean squared error was adopted as our loss measurement. Due to the high variance in the bed height data, a large batch size of 5000 was utilized. To prevent overfitting, 30% of the testing data was reserved for validation, and the testing-validation split was shuffled each epoch. Striking a balance between accuracy and overfitting, the model underwent training for 100 epochs.

6.8 Dense+LSTM

To effectively reveal intricate relationships within the dataset and address the challenges posed by incomplete data and complex ice bed topography, we propose the fusion of dense and Long Short-Term Memory (LSTM) layers in our model. The inclusion of dense layers in deep neural networks provides a powerful and straightforward means of nonlinearly aggregating input from preceding neurons. This fusion allowed the extraction of latent features and revealed hidden patterns. The model architecture consisted of three dense layer blocks, each featuring two dense layers with sigmoid activation functions, and two LSTM layers with 128, 64, and 32 cells, respectively. A batch normalization layer was introduced, and the output of the model was generated through a single neuron with a linear activation function. The dataset was transformed into a 3D array to enable the effective application of this integrated model, capturing spatiotemporal elements in the ice flow velocity measurements.

The model was meticulously fine-tuned using a batch size of 5000 and 200 epochs as hyperparameters for training. The optimization process revolved around the Adam method, with mean squared error serving as the objective function to be minimized. During each epoch, 60% of the training data was utilized, with the remaining 40% reserved for model validation. This approach allowed for an accurate assessment of the model's performance, promoting generalization and preventing overfitting.

6.9 XGBoost

After testing various models, we decided to adopt the Extreme Gradient Boosting (XGB) supervised machine learning algorithm. Known for its effectiveness in regression tasks and computational efficiency, XGB is well-suited for processing large datasets. The model combines multiple decision trees, optimizing them to minimize errors, which enhances flexibility and makes it adaptable to diverse topographic data. XGB incorporates techniques to prevent overfitting, ensuring reliable and accurate predictions, while its ability to aggregate predictions from individual trees solidifies its standing as the preferred choice for our objectives.

In justifying our choice of the XGB algorithm, we carefully considered the importance of finding the right balance among several key parameters. To understand the parameters the model was tuned to extremely overfit the data. Each parameter was then tested individually to identify how it affected the metrics and clarity of the predictions – through visualization. After understanding the effects of each parameter, they were adjusted according to the project scope to best represent the data predictions and emphasize domain specifics. These parameters significantly impact the model's performance and adaptability, making it crucial to achieve optimal results.

Parameter settings were carefully chosen to optimize the model's performance while avoiding overfitting. The depth of the decision trees was set to a balanced value of 7 to capture intricate patterns without overfitting. The number of boosting rounds and XGBoost trees was set at 350 to ensure comprehensive learning while maintaining training efficiency. The minimum child weight, set at 0.25, controlled overfitting by requiring a minimum number of samples to split a node. The subsample parameter, set at 0.8, is balanced incorporating diverse samples and providing sufficient dataset coverage. Finally, the learning rate (eta) of 0.25 facilitated stable performance and convergence without overshooting.

Finding the right balance among these parameters is crucial. Adjusting these parameters too drastically may lead to either overly complex or oversimplified models, impacting performance and pattern capture. By carefully fine-tuning these parameter values, we can strike the optimal balance, maximizing the model's performance, capturing important patterns, and avoiding overfitting. This enables us to achieve reliable predictions tailored to our project's needs.

7 Metrics

To validate the results, we considered a range of metrics to comprehensively assess the performance and accuracy of the model. While all metrics provide valuable insights, four key metrics were given importance: the coefficient of determination (R^2), the root mean squared error (RMSE), the mean absolute error (MAE), and the terrain ruggedness index (TRI). These metrics play a critical role in evaluating performance and drawing meaningful conclusions.

R^2 is a crucial metric as it measures the proportion of the variance in the dependent variable that is predictable from the independent variables. A high R^2 value indicates a strong relationship between the predicted and actual values, providing a measure of how well the model fits the observed data. Focusing on R^2 allows for an understanding of the achieved predictability and an assessment of the goodness of fit.

Similarly, RMSE provides an essential measure of the average magnitude of the residuals or errors between the predicted values and the actual values. By calculating the square root of the average squared differences, RMSE offers a comprehensive evaluation of the model's predictive accuracy. Emphasizing RMSE allows for a focus on the precision of predictions and an understanding of the typical magnitude of errors.

Furthermore, MAE is given equal importance as it provides insights into the average magnitude of errors, independent of their direction. By calculating the average absolute difference between the predicted and actual values, MAE offers a robust measure of the model's accuracy. It allows for a clear assessment of the typical error magnitude and helps quantify the overall performance of the model.

The Terrain Ruggedness Index (TRI) is a valuable tool for quantifying terrain characteristics, expressing the elevation difference between adjacent cells of a Digital Elevation Model (DEM). It measures topographic heterogeneity by calculating the difference between the center cell and its eight surrounding cells, squaring and averaging these differences, and taking the square root to obtain the TRI value. Higher TRI values indicate greater terrain ruggedness or complexity compared to lower values which indicate smoothness. By incorporating TRI into our analysis, we can evaluate how well our predictions capture the terrain's variability and roughness, ensuring an accurate representation of the study area's complex topographic features.

While R^2, RMSE, MAE, and TRI are important in this evaluation, additional metrics such as cosine similarity and the Pearson correlation coefficient have also been considered. Cosine similarity assesses the similarity between predicted and actual values, regardless of their magnitude, while the Pearson correlation coefficient measures the linear relationship between predicted values and known values. These metrics contribute to the comprehensive evaluation of the model's performance and provide valuable insights into its predictive capabilities.

To validate the results, we considered a range of metrics to comprehensively assess the performance and accuracy of the model. These metrics play a critical role in evaluating performance and drawing meaningful conclusions. With a comprehensive evaluation of various metrics, we now delve into the results to gain insights into the accuracy and efficacy of our integrated model, showcasing its predictive capabilities and ability to capture intricate topographical characteristics.

8 Results

In this study, we evaluated the performance of several predictive models to uncover the most effective approach for capturing and representing spatial patterns in complex topography. The models considered include the XGBoost model on Kriging residual learning, Gaussian Process Regression, Spatio-Temporal Gaussian Process, Variational Auto-encoders (VAE), VAE with XGBoost, Dense, LSTM, Dense+LSTM, and XGBoost on kriging interpolated data as well as nearest neighbors interpolated data. Each model and interpolation set was subjected to rigorous testing using RMSE, MAE, and R^2 evaluation, with a particular focus on their predictive accuracy and ability to represent the underlying topography.

8.1 XGBoost for Kriging Residual Learning

The XGBoosting model, trained for predicting Kriging's first guess residuals using the same features as other models, exhibited notably poor performance compared to the other models. It produced an RMSE of 136.650, an MAE of 8.122, and an R^2 value of 0.424. The results accentuate Kriging's first pass prediction results which produce RMSE of 188.948, MAE of 7.801, and R^2 value of -0.103 which means the predictions are not strongly correlated with the correct, known values. The large difference in RMSE to MAE can be explained by Kriging's ability to predict dense regions well (R^2 of 0.998) and sparse regions poorly (R^2 of -0.5). These results indicate its inability to effectively capture the relationships across all the data and render it unsuitable for practical use.

8.2 GPR Model

The application of GPR to predict topography data in Greenland produced interesting results. The GPR model demonstrated an RMSE of 150.086 and an MAE of 97.855. Additionally, the coefficient of determination was found to be 0.293. The probabilistic nature of GPR allowed for estimating predictive uncertainties, providing valuable information about the uncertainties associated with the topographical predictions. However, the results revealed some challenges in accurately capturing the complex and diverse geographical features of Greenland. The results show that the predictions made using GPR may not have been very accurate, as indicated by the relatively high errors and lower R^2 value. Additionally, due to the considerable time and expensive computational resources required for training and fine-tuning the GPR model, further optimization and tuning were not pursued. The resource-intensive nature of GPR made it impractical to explore extensive hyperparameter searches and exhaustive refinements. As a result, the focus shifted towards alternative modeling approaches that could provide more efficient and cost-effective solutions for predicting topography data in Greenland.

8.3 STGP Model

The results of the STGP in predicting topography data for Greenland yielded an RMSE of 225.772 and an MAE of 126.819. However, the coefficient of determination displayed a negative value of -0.580, indicating that the current STGP model might not effectively capture the data's variability around the mean and perform poorly compared to a simple horizontal line. Despite these initial outcomes, with the acknowledgment that resource constraints may have influenced the results, there remains hope that dedicating more time and computational resources for further tuning and optimization could significantly enhance the model's performance. By continuing to explore and fine-tune STGP, in conjunction with other predictive models, we aim to unlock valuable insights and develop more robust tools for accurately predicting and understanding the complex geographical features of Greenland.

8.4 Universal Kriging

Universal kriging obtained very poor metrics for R^2 and RMSE (-0.103 and 188.948) yet a very low MAE score depicted in 8.4. Upon further investigation, it became clear that this was because of a small number of outliers with very poor predictions; and that Kriging performed very well on the majority of the dataset. 83.4% of predictions had a residual of 10 or less, 81.4% had an RMSE of 25 or less, and 79.6% had an R^2 score of 0.97 or higher. But because RMSE is more sensitive to outliers than MAE, Kriging's RMSE is poor.

Table 8.1: Proportion of UK predictions where $R^2 >=$ threshold

Threshold	**Proportion**
.6	0.896
.9	0.841
.97	0.796
.99	0.416

Table 8.2: Proportion of UK predictions where $RMSE <=$ threshold

Threshold	**Proportion**
50	0.872
25	0.814
20	0.691

Table 8.3: Proportion of UK predictions where *AbsoluteResidual* <= threshold

Threshold	**Proportion**
20	0.925
10	0.834
5	0.703
2	0.510
1	0.387

8.5 VAE & XGB Model

The combination of VAEs as an encoder, and XGB as the predictor aimed to provide a robust and interpretable approach for capturing patterns and predicting topography data in Greenland. The results of this combined methodology revealed an RMSE of 129.760, an MAE of 100.035, and an R^2 of 0.481. The methodology appeared promising due to the VAEs' ability to capture essential features, leveraging the encoder, and XGB's powerful predictive capabilities in handling complex interactions between geographical features. However, the relatively high error metrics and lower R^2 value indicate that the combined approach may not have achieved the desired levels of accuracy and precision in predicting topographical features in Greenland.

8.6 VAE Model

The application of VAEs as an encoder and decoder to predict topography data in Greenland resulted in an RMSE of 106.884, an MAE of 83.798, and an R^2 of 0.648. While the VAEs demonstrated promise in capturing essential information about the topography, the relatively high error metrics, and R^2 suggest that their predictive performance may not be as accurate as desired compared to other models. The complex nature of Greenland's geographical features presents challenges, highlighting the need for further refinement to enhance the VAEs' effectiveness in predicting topographical data in this region and other diverse landscapes.

8.7 Multilayer Perceptron (MLP) Dense Model

The MLP (Dense) model demonstrated predictive capabilities in predicting ice bed height with an RMSE of 104.475, an MAE of 74.381, and an R^2 of 0.663. Despite not being the best-performing model, there is reason to be hopeful that the ideas and insights gained from this model will contribute to the development of a more powerful and effective LSTM+Dense model in future iterations. The model's architecture, optimization techniques, and regularization approaches serve as valuable building blocks for further enhancing predictive performance in ice bed height prediction applications.

8.8 LSTM Model

The Long Short-Term Memory (LSTM) model displayed commendable predictive capabilities in predicting ice bed height, with an RMSE of 101.630, an MAE of 74.522, and an R^2 of 0.682. These results indicate a notable performance that outperforms the Dense model, making the LSTM model a strong contender among the evaluated methods. LSTM's ability to capture long-term dependencies and handle sequential data allowed it to effectively exploit temporal patterns within the ice bed height dataset. The model's performance showcases its potential as a reliable tool for accurate ice bed height prediction. Considering the promising results achieved with LSTM, there is optimism that further refinement and combination of ideas from both LSTM and the DNN model can lead to even more improved and precise predictions in future iterations. By combining the strengths of LSTM and Dense layers, there is potential to create a more sophisticated and accurate model for this critical task.

8.9 Dense+LSTM Model

The Dense+LSTM model, with the nearest neighbors interpolated dataset, displayed respectable predictive accuracy, but it did not outperform the XGBoosting model. It achieved an RMSE of 81.174, an MAE of 57.993, and an R^2 value of 0.797. When visualizing the predictions on the additional validation data, we can see clearly low areas in the topography, but the predictions do not capture detail in other sections of the land. While this model provided reasonably accurate predictions, it did not match the performance of the leading XGBoosting model.

8.10 XGBoost with Kriging Interpolation

In contrast, the XGBoost model with kriging interpolated data demonstrated the best metrics among the alternative models, achieving an RMSE of 27.099, an MAE of 17.947, and an R^2 value of 0.977, showcasing its predictive capabilities. The TRI analysis yielded a mean TRI score of 169.498, with an R^2 value of 0.951, further validating its ability to capture the roughness and complexity of the terrain. However, it is important to note that the kriging method demands extensive computational resources, taking an estimated 6 hours with five nodes and 320GB memory on high-performance computing clusters for interpolation. While the numerical evaluation of chosen metrics indicates strong performance, the method's total training time requirement poses a limitation in certain applications.

8.11 XGBoost with Nearest Neighbors Interpolation

The XGBoost model with nearest neighbors interpolated data showcased strong predictive performance, achieving an RMSE of 32.680, an MAE of 22.273, and an impressive R^2 value of 0.967. The TRI analysis further solidified its effectiveness, yielding a mean TRI score of 168.384, with an R^2 value of 0.932, attesting to its ability to predict the topography with precision. Notably, this model required an estimated one hour for interpolation steps to predict models, utilizing less than 12GB of memory. When applied to a larger dataset for additional validation, the plotted topography predictions exhibited clear differentiation in target heights, capturing intricate details of the underlying patterns in the data and topography. These exceptional metrics highlight the effectiveness of the trained XGBoost model in accurately predicting the target variable and capturing the complexity of the Greenland Ice Sheet terrain with precision and efficiency.

8.12 XGBoost with Bilinear Interpolation

Finally, the XGBoost model with bilinear interpolation also demonstrated excellent performance, achieving an RMSE of 28.085, an MAE of 18.549, and an impressive R^2 value of 0.976, affirming its strong predictive capabilities. The TRI analysis further supported its effectiveness, yielding an R^2 value of 0.950 and a mean TRI score of 168.384, which is remarkably close to the known validation TRI mean of 173.404. This model achieved these outstanding results with a memory usage of only 12GB and completed in just 15 minutes, demonstrating both efficiency and accuracy in its predictions. The XGBoost model with bilinear interpolation proves to be a valuable addition to our evaluation, showcasing its capabilities in predicting the target variable while efficiently processing large datasets with complex topography.

8.13 Summary

Through a comprehensive evaluation of various predictive models, we identified the XGBoost model with kriging interpolated data as the best performer with strong predictive capabilities with an RMSE of 27.099, an MAE of 17.947, and an R^2 of 0.977. However, due to its extensive computational requirements, the XGBoost model with bilinear interpolation emerged as a close contender, achieving an RMSE of 28.085, MAE of 18.549, and R^2 of 0.976, with a remarkable TRI validation result. Overall, these XGBoost models stood out, efficiently capturing complex topography patterns and providing valuable tools for representing spatial patterns in the Greenland Ice Sheet and beyond.

Table 8.4: Summary of Results (Sorted by Best Metrics)

Model	**RMSE**	**MAE**	$\mathbf{R^2}$
XGBoost (Kriging)	27.099	17.947	0.977
XGBoost (Bilinear)	28.085	18.549	0.976
XGBoost (Nearest Neighbors)	32.680	22.273	0.967
Physics BedMachine	71.554	50.422	0.842
Dense + LSTM	81.174	57.993	0.797
LSTM	101.630	74.522	0.682
Dense	104.475	74.381	0.663
Variational AutoEncoder	106.884	83.798	0.648
VAE + XGBoost	129.760	100.035	0.481
Kriging Only	188.948	7.801	-0.103
XGBoost on Kriging Residual	136.650	8.122	0.424
Gaussian Process Regression	150.086	97.855	0.293
Spatio-Temporal GP	225.772	126.819	-0.580

9 Conclusions

Motivated by the significance of understanding complex topography and its implications in various fields, such as climate change research, environmental modeling, and glaciology, this study aimed to identify the most effective predictive models for accurately capturing and representing spatial patterns in intricate terrain. Leveraging insights from past literature and

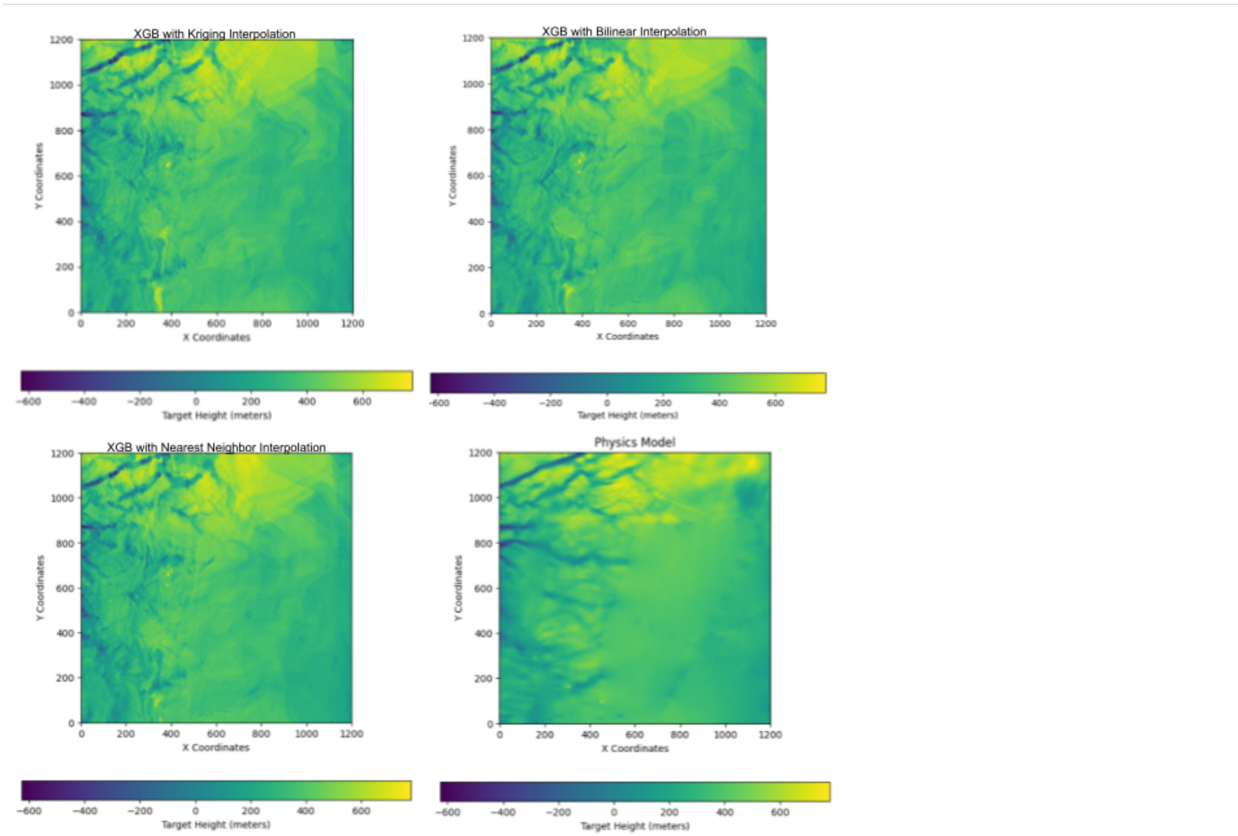

Figure 8.1: Validation Data Visualized for XGBoost Models

a comprehensive understanding and preprocessing of our data, we embarked on a rigorous evaluation of several predictive models, employing well-established evaluation metrics like RMSE, MAE, R^2, and TRI with a specific focus on predictive accuracy and topography representation.

The results revealed distinct performance characteristics among the evaluated models. The XGBoost model with kriging interpolated data exhibited the best metrics, achieving an RMSE of 27.099, an MAE of 17.947, and an R^2 value of 0.977, indicating its strong predictive capabilities. However, the method's computational requirements, taking an estimated 6 hours with extensive resources, limit its practicality in certain applications.

The XGBoost model with bilinear interpolation also demonstrated excellent predictive capabilities, achieving an RMSE of 28.085, an MAE of 18.549, and an R^2 value of 0.976. The TRI analysis further supported its effectiveness, with an R^2 value of 0.950 and a mean TRI score of 168.384, remarkably close to the known validation TRI mean of 173.404. Notably, this model required a mere 15 minutes and 12GB of memory, showcasing both efficiency and accuracy in its predictions.

The other evaluated models exhibited varying degrees of predictive performance, with some showing promising results and potential for further refinement. However, the XGBoost models with kriging and bilinear interpolated data stood out as the top performers, effectively capturing the complexity of the Greenland Ice Sheet terrain with precision and efficiency.

In conclusion, the XGBoost models with kriging and bilinear interpolated data proved to be highly effective in predicting topographical features in complex terrains. These models demonstrated exceptional predictive accuracy while efficiently processing large datasets, making them valuable tools for capturing and representing spatial patterns in the Greenland Ice Sheet and potentially other diverse landscapes.

10 Future Work

The comprehensive evaluation of predictive models in this study has provided valuable insights into the most effective approach for capturing and representing complex topography. Moving forward, there are several exciting avenues for future research that can further enhance predictive accuracy and efficiency in modeling topographical features in diverse landscapes.

The results of this study highlight the promising potential of incorporating the Terrain Ruggedness Index (TRI) as an early stopping method in model training. By leveraging TRI as a metric during the training process, we can effectively identify when the model has reached a satisfactory level of capturing the terrain's complexity and variability. This approach could save computational resources and time. As a future direction, further research can explore the integration of TRI-based early stopping in different predictive models and assess its impact on training efficiency and predictive accuracy.

Another avenue for future work involves dedicating more computational resources to explore the Spatio-Temporal Gaussian Process (STGP) model in greater depth. The initial results from the STGP model demonstrated its potential to predict topographical features in complex terrains. However, the limitations imposed by resource constraints may have influenced the model's performance. By allocating additional resources, such as computational power and memory, researchers can conduct more extensive hyperparameter searches and fine-tuning, leading to a better understanding of the model's capabilities and potential for accurate predictions. With the availability of larger datasets and advancements in computing technologies, exploring the STGP model's full potential becomes an essential future research direction.

Additionally, an intriguing direction for future work involves combining the Spatio-Temporal Gaussian Process (STGP) model's pattern capturing with the powerful computational speed and architecture of the XGBoost model. Leveraging the complementary strengths of STGP and XGBoost has the potential to unlock new possibilities in accurately predicting and understanding spatial patterns in diverse landscapes, paving the way for more efficient and effective applications in terrain modeling and related fields.

Furthermore, extending the investigation to consider other interpolation techniques and data preprocessing methods can provide valuable comparisons and insights. While the XGBoost models with kriging and bilinear interpolated data demonstrated impressive performance, exploring other interpolation methods, such as inverse distance weighting, radial basis functions, or splines, could uncover additional nuances and optimizations for predicting complex topography. Fine-tuning the parameters and configurations of the selected interpolation techniques may further enhance predictive accuracy and efficiency.

Another potential avenue for future research involves exploring ensemble approaches that combine multiple models to harness their collective predictive power. Ensemble methods, such as bagging or boosting, can help mitigate the weaknesses of individual models and capitalize on their strengths, leading to improved overall performance and robustness. By leveraging the complementary nature of different predictive models, ensemble techniques can provide more reliable and accurate predictions, especially in complex terrains where capturing spatial patterns requires a multifaceted approach.

This study represents a valuable step forward in understanding the predictive capabilities of various models for complex topography representation. The integration of the Terrain Ruggedness Index as an early stopping method and the exploration of the Spatio-Temporal Gaussian Process model with more resources and alternative architecture offers future directions to enhance predictive accuracy and efficiency. By continuously refining and advancing our predictive approaches, we can deepen our understanding of complex topographical features and their implications, ultimately contributing to a wide range of applications, from environmental modeling to climate change research.

Acknowledgments

This work is supported by the grants “REU Site: Online Interdisciplinary Big Data Analytics in Science and Engineering (grant no. OAC–2050943)” and “HDR Institute: HARP- Harnessing Data and Model Revolution in the Polar Regions (grant no. OAC–2118285)” from the National Science Foundation. The hardware used in the computational studies is part of the UMBC High-Performance Computing Facility (HPCF). The facility is supported by the U.S. National Science Foundation through the MRI program (grant nos. CNS–0821258, CNS–1228778, OAC–1726023, and CNS–1920079) and the SCREMS program (grant no. DMS–0821311), with additional substantial support from the University of Maryland, Baltimore County (UMBC). See hpcf.umbc.edu for more information on HPCF and the projects using its resources.

References

[1] Ali, S. A. (2023, March 27). sea-ice-prediction/probabilistic-modeling at main · big-data-lab-umbc/sea-ice-prediction. GitHub. https://github.com/big-data-lab-umbc/sea-ice-prediction/tree/main/probabilistic-modeling

[2] Chollet, F. C. (2020, May 5). Keras documentation: Variational AutoEncoder. https://keras.io/examples/generative/vae/

[3] Denton, E. (2016, November 19). Semi-Supervised Learning with Context-Conditional Generative Adversarial Networks. arXiv.org. https://arxiv.org/abs/1611.06430

[4] Faruque, O., Alam, H., Hossain, E., & Hussein, A. (n.d.). From Data to Topography: Deep Learning Approach to Predict Ice Bed. Department of Information Systems, UMBC.

[5] Hamelijnck, O. (2021, November 2). Spatio-Temporal variational gaussian processes. arXiv.org. https://arxiv.org/abs/2111.01732

[6] Han, D. (2013). Comparison of commonly used image interpolation methods. In Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013). https://doi.org/10.2991/iccsee.2013.391

[7] Kleijnen, J. P. C. (2017). Kriging: Methods and Applications. SSRN Electronic Journal, 10(2139). https://doi.org/10.2139/ssrn.3075151

[8] Leong, W. L., & Horgan, H. J. (2020). DeepBedMap: a deep neural network for resolving the bed topography of Antarctica. The Cryosphere, 14(11), 3687–3705. https://doi.org/10.5194/tc-14-3687-2020

[9] Liu-Schiaffini, M., Ng, G., Grima, C., Young, D. (2022). Ice thickness from deep learning and conditional random fields: application to Ice-Penetrating radar data with radiometric validation. IEEE Transactions on Geoscience and Remote Sensing, 60, 1–14. https://doi.org/10.1109/tgrs.2022.3214147

[10] LoPresti, L. (2019, June 26). Less pain, more gain: A simple method for VAE training with less of that KL-vanishing agony - Microsoft Research. Microsoft Research. https://www.microsoft.com/en-us/research/blog/less-pain-more-gain-a-simple-method-for-vae-training-with-less-of-that-/kl-vanishing-agony/#:%7E:text=VAEs%20aim%20to%20learn%20probabilistic%20representations%20z%20of,leverage%20the%20prior%20knowledge%20to%20modulate%20language%20understanding

[11] Riley, Shawn J and DeGloria, Stephen D and Elliot, Robert Index that quantifies topographic heterogeneity. Intermountain Journal of sciences, 5, 23-27.

[12] Morlighem, M., Rignot, E., Mouginot, J., Seroussi, H., Larour, E. (2014) Deeply incised submarine glacial valleys beneath the Greenland ice sheet. Nature Geosci, 7, 418–422. https://doi.org/10.1038/ngeo2167

[13] Morlighem M. et al., (2017), BedMachine v3: Complete bed topography and ocean bathymetry mapping of Greenland from multi-beam echo sounding combined with mass conservation, Geophys. Res. Lett., 44, doi:10.1002/2017GL074954, http://onlinelibrary.wiley.com/doi/10.1002/2017GL074954/full

[14] Leong, W.J. and Horgan, H.J., 2020. DeepBedMap: A deep neural network for resolving the bed topography of Antarctica. The Cryosphere, 14(11), pp.3687-3705.

[15] Howat, I., A. Negrete, and B. Smith. MEaSUREs Greenland Ice Mapping Project (GIMP) Digital Elevation Model from GeoEye and WorldView Imagery, Version 1. 2017, Distributed by NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/H0KUYVF53Q8M. Date Accessed 08-30-2023.

[16] https://earth.esa.int/eogateway/catalog/radarsat-2-esa-archive.

[17] Kwok, R., A. A. Petty, G. Cunningham, T. Markus, D. Hancock, A. Ivanoff, J. Wimert, M. Bagnardi, N. Kurtz, and the ICESat-2 Science Team. ATLAS/ICESat-2 L3A Sea Ice Height Quick Look, Version 6. 2023, Distributed by NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/ATLAS/ATL07QL.006. Date Accessed 08-30-2023.

[18] Wessem,J. M. Van, Laffin,M. K. (2020). Regional Atmospheric Climate Model 2 (RACMO2), version 2.3p2 (2.3p2) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.3677642.

[19] MacGregor, J.A., Boisvert,L.N., Medley, B.,Petty, A.A.,Harbeck, J.P., Bell, R.E., Blair, J.B., Blanchard-Wrigglesworth, E., Buckley, E.M., Christoffersen, M.S. and Cochran, J.R., 2021. The scientific legacy of NASA's Operation IceBridge.

[20] Kingma, D.P. and Welling, M., 2013. Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114.