remote sensing

MDPI

Article

Reconstruction of Near-Surface Air Temperature over the Greenland Ice Sheet Based on MODIS Data and Machine Learning Approaches

Jiahang Che [1], Minghu Ding [2], Qinglin Zhang [1], Yetang Wang [1], Weijun Sun [1], Yuzhe Wang [1], Lei Wang [1] and Baojuan Huai [1,*]

[1] College of Geography and Environment, Shandong Normal University, Jinan 250358, China
[2] State Key Laboratory of Severe Weather, China Academy of Meteorological Sciences, Beijing 100081, China
* Correspondence: huaibaojuan@126.com or 616070@sdnu.edu.cn

Abstract: High spatial and temporal resolution products of near-surface air temperature (T2m) over the Greenland Ice Sheet (GrIS) are required as baseline information in a variety of research disciplines. Due to the sparse network of in situ data on the GrIS, remote sensing data and machine learning methods provide great advantages, due to their capacity and accessibility. The Land Surface Temperature (LST) at 780 m resolution from the Moderate Resolution Imaging Spectroradiometer (MODIS) and T2m observation from 25 Automatic Weather Stations (AWSs) are used to establish a relationship over the GrIS by comparing multiple machine learning approaches. Four machine learning methods—neural network (NN), gaussian process regression (GPR), support vector machine (SVM), and random forest (RF)—are used to reconstruct the T2m at daily and monthly scales. We develop a reliable T2m reconstruction model based on key meteorological parameters, such as albedo, wind speed, and specific humidity. The reconstructions daily and monthly products are generated on a 780 m × 780 m spatial grid spanning from 2007 to 2019. When compared with in situ observations, the NN method presents the highest accuracy, with R of 0.96, RMSE of 2.67 °C, and BIAS of −0.36 °C. Similar to the regional climate model (RACMO2.3p2), the reconstructed T2m can better reflect the spatial pattern in term of latitude, longitude, and altitude effects.

Keywords: temperature; MODIS; machine learning methods; GrIS

Citation: Che, J.; Ding, M.; Zhang, Q.; Wang, Y.; Sun, W.; Wang, Y.; Wang, L.; Huai, B. Reconstruction of Near-Surface Air Temperature over the Greenland Ice Sheet Based on MODIS Data and Machine Learning Approaches. *Remote Sens.* **2022**, *14*, 5775. https://doi.org/10.3390/rs14225775

Academic Editors: Gang Qiao, Yan Liu, Zemin Wang, Baojun Zhang and Fengming Hui

Received: 17 October 2022
Accepted: 13 November 2022
Published: 16 November 2022

1. Introduction

Due to the amplification effect in the Arctic [1], the near-surface air temperature (T2m) here rises faster than in other regions of the world [2,3]. The Greenland Ice Sheet (GrIS) is the second-largest ice sheet in the world, with an area of 1.7×10^6 km^2, being only smaller than the Antarctic Ice Sheet [4]. In recent decades, the GrIS has attracted a lot of attention due to its enhanced melting [5–7], particularly associated with unusual climatic conditions [8–10] and accelerating mass loss [11–14].

The T2m plays an important role in glaciological and climatological processes in the Arctic [15–18]. In particular, in extremely cold regions, the T2m is an important input factor for glacier or snow melting simulations [19,20], as well as the basis for global and regional climate warming analyses and an important factor in controlling the surface mass balance (SMB) [21–25]. Compared with the 20th century, the T2m has increased significantly in Greenland [21]. A common approach to obtaining T2m datasets is spatial interpolation based on Automatic Weather Station (AWS) records [26–28]. However, the interpolation approach relies on a sufficiently dense number of AWS. Simple interpolation methods over the GrIS are challenging, as the network of AWSs is sparse, due to its remoteness. At present, meteorological stations (both AWS and staffed stations), regional climate model, global re-analysis data, and remote sensing data are the four main sources for obtaining the ice surface temperature (IST) and T2m [29–32].

Although global re-analysis data and regional climate model are reliable across the GrIS, they may underestimate climate variability in regional and local environments [31,33,34]. The spatial resolution of remote sensing data is high, and it can provide IST data for the whole ice sheet, making it more advantageous than AWS and the regional climate model. Compared with limited AWSs, remote sensing IST data can provide better spatial coverage [35]. Moderate Resolution Imaging Spectroradiometer (MODIS) IST products have been successfully used to estimate the T2m in different regions of the world [11–14,36–39], on the basis of the strong correlation between T2m and IST [40–50]. Various statistical methods, including simple linear regression [32,48,51], and complex models, such as the random forest (RF) prediction model [52,53], have been used for T2m estimation. In addition to MODIS IST, some auxiliary variables, including elevation, latitude, longitude, month, windspeed, specific humidity, and so on, may also be included in the model for more accurate T2m estimation [40,52,54]. Hooker et al. [55] carried out kriging-based spatio-temporal interpolation of global daily temperature including MODIS IST as a predictor. Nielsen-Englyst et al. [32] reconstructed the Arctic T2m using a linear method and compared it with re-analysis data (ERA5, ERA-interim, etc.). Some researchers have also used non-linear methods for reconstruction; for example, Meyer et al. [44] used machine learning techniques to reconstruct the T2m over Antarctica. Zhou et al. [53] used a hybrid approach based on machine learning to estimate T2m across Israel.

However, these previous studies have not examined Greenland. Therefore, in this study, we aim to create a medium-resolution daily T2m product for the GrIS. In this context, we compare the performance of four different machine learning methods to reconstruct T2m from IST and other auxiliary variables, including specific humidity, wind speed, albedo, longitude, latitude, altitude, and month. The four machine learning methods include neural network (NN), Gaussian process regression (GPR), support vector machines (SVM), and RF.

Due to the rapid development of machine learning, these and similar algorithms have been widely used in the fields of numerical prediction and numerical simulation [56,57]. The main advantages of machine learning methods are as follows: first, the models trained by machine learning methods can adapt well to new data and have good generalization ability. Second, machine learning methods consist of optimization algorithms that can be modeled several times, until the optimal solution is found and the best reconstruction results are obtained. Finally, the use of machine learning methods, such as NN and GPR, can effectively prevent over-fitting.

An important application area of machine learning is the reconstruction of T2m. NN, RF, and SVM are commonly used methods to obtain T2m products from remote sensing LST data, which have been developed and applied in many areas. Machine learning approaches have a strong ability to solve non-linear relationship fitting problems. Meyer et al. [58] have used NDVI, elevation, and solar zenith angle as inputs to a RF model and compared them with T2m data recorded at weather stations. Choi et al. [59] used surface temperatures in bands 10 and 11 recorded by the Landsat-8 satellite, NDVI, and the automated weather observing system (ASOS) as the input layer to a deep neural network to simulate T2m. Therefore, it is feasible to apply machine learning methods to reconstruct the T2m in Greenland.

The remainder of this paper is structured as follows: The data and methods are described in Section 2. In Section 3.1, the daily data products are reconstructed. In Section 3.2, the monthly data products are reconstructed. The performance of the reconstructed products is evaluated in Section 3.3. This is followed by a discussion in Section 4, which is divided into three main subsections. The first considers the albedo scheme, the second provides a comparison with other studies, and the third details the limitations and perspectives for future research. Section 5 presents our conclusions.

2. Data and Methods

2.1. Data

2.1.1. In Situ Data

In this study, we use data from the Greenland Ice Sheet Monitoring Program (PROMICE) AWSs as observations. The data are operated by the Geological Survey of Denmark and Greenland (GEUS), in collaboration with the National Space Research Institute [60]. PROMICE AWSs are distributed mostly in the coastal regions of Greenland with an uneven spatial distribution (see Figure 1) and provide various data, such as air pressure, T2m, specific humidity, wind direction, sensible heat flux, four components of radiation, wind speed, albedo, relative humidity, and surface temperature, among others [61].

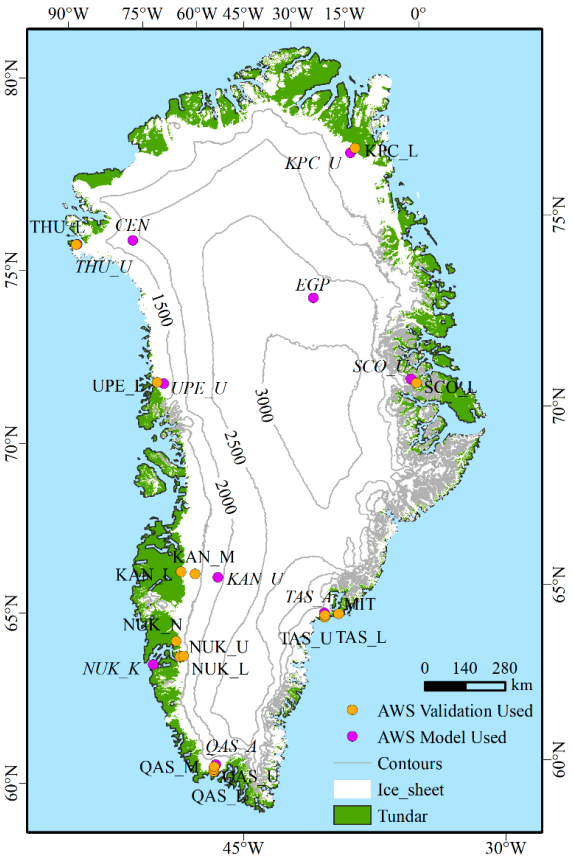

Figure 1. Overview map of the study area. The purple circles are training data set stations (marked in italics), and the orange circles are validation data set stations.

We used the daily data of surface temperature, T2m, albedo, latitude, longitude, wind speed, specific humidity, altitude, and month from 25 AWSs provided by PROMICE and divided them into training and validation data. Likewise, the monthly data were also divided into training and verification data. However, it should be noted that there is no IST in the monthly PROMICE AWSs data; therefore, it was averaged from the daily data. In this case, monthly data with less than 90% of days in a month were regarded as missing [61].

Among them, 10 AWSs were used for training data, and the remaining 15 AWSs were used for verification data [62]. Figure 1 shows the distribution of PROMICE AWSs in Greenland (with the polar stereographic projection method). The observation period is shown in Figure 2.

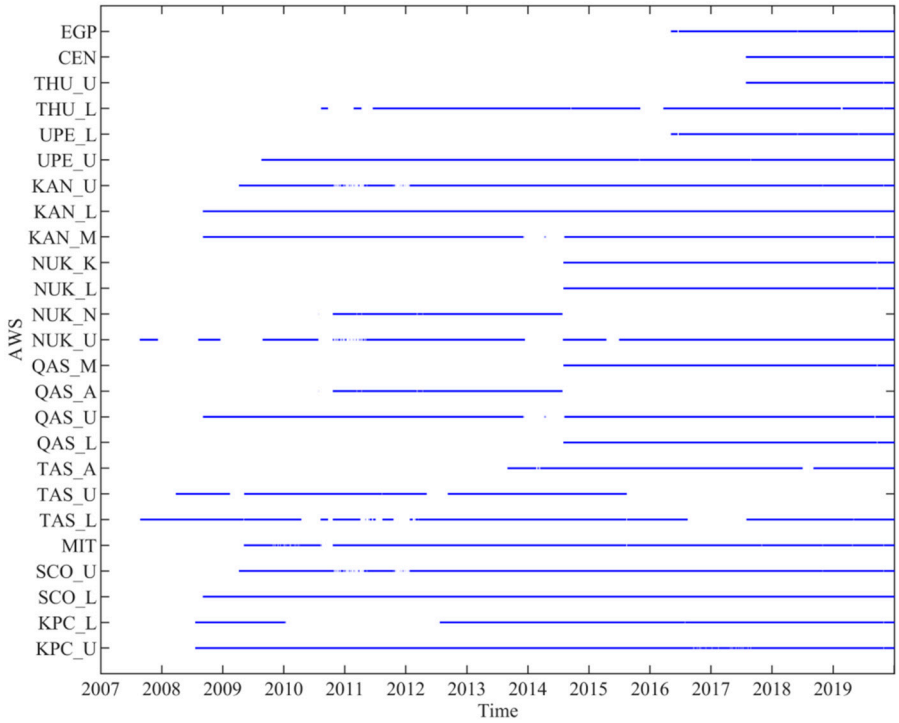

Figure 2. Observation periods for AWSs used in this study.

2.1.2. Satellite Data

MODIS is a polar-orbiting and cross-orbiting scanning spectroradiometer whose images cover the entire Earth region every day or two days [63]. The daily grid products based on GrIS strips became available in the early 2000, as MODIS began flying on the NASA Earth Observing System (EOS) Terra satellite in 2000 and on Aqua from 2002 [64]. Terra's orbit around the earth is time-limited, and it passes through Greenland in the afternoon (14:30–16:30 Universal Time). The MODIS Collection 6.1 IST and albedo used in this study come from a multi-layer daily IST albedo water vapor product, developed using the standard MODIS data set of the Terra satellite [65].

The algorithms were derived from those developed for the MODIS sea ice products MOD29 or MYD29 (MOD and MYD represent the Terra and Aqua products, respectively) [64,66,67]. From March 2000 to December 2019, MOD 29 IST and albedo products provided daily and monthly average surface temperature data on a polar stereo grid with 0.78 × 0.78 km resolution [65]. The IST accuracy of the 6.1 series products has been reported to be approximately twice that of the 5 series, due to the addition of an emissivity adjustment model into the MODIS split-window algorithm [67,68].

However, as is well-known, the surface temperature data of MODIS has a cold bias in the GrIS [69]; the colder it is, the greater the bias [70]. This cold bias in MODIS data may be due to mistaking temperatures on clouds for surface temperatures, which can lead to some extreme cases such as extremely low temperatures [64]. Cloud cover grid elements without surface temperature values were excluded from the calculation process. Due to the serious missing results of daily data reconstruction in some regions and the desire to study the effect of reconstruction at different time scales, monthly data were also reconstructed.

2.1.3. Regional Climate Model

The Regional Climate Model (RACMO2.3p2) was developed by the Royal Netherlands Meteorological Institute (KNMI) for polar glacial regions and atmospheric studies, Utrecht University (IMAU/UU), specifically for simulating climate and surface mass balance (SMB) in Antarctica and the GrIS [71]. RACMO2.3p2 is an updated version, which includes a number of changes: the glacier profile, topography, and ice albedo field have been updated; the cloud scheme parameters of increasing precipitation into the GrIS have been adjusted;

the under-estimation of inland accumulation in RACMO2.3 was corrected; the properties of snow have been improved, and so on [29,72]. RACMO2.3p2 is considered especially suitable for the large ice sheet in Greenland [73].

As a regional climate model data, RACMO2.3p2 has a high T2m accuracy [30,72,74]. Therefore, we compared the remote sensing data reconstruction results using RACMO2.3p2.

2.2. Methods

2.2.1. Machine Learning

Machine learning involves the study of how to improve the performance of a system itself through computation and experience. In computer systems, "experience" usually exists in the form of data [75]. Therefore, the main research component of machine learning is learning algorithms, which we provide with empirical data in order to generate a model. The learned model corresponds to some underlying law of the data, called the "truth," and the process of learning is the process of discovering or approaching the truth. The model must have a strong generalization ability and should be able to be applied to the whole sample space [76].

(1) Neural network (NN)

The most basic component of a neural network is the neuron model. A well-known algorithm is the Levenberg–Marquardt algorithm [77], which can provide numerical solutions for non-linear minimization (local minimization). This algorithm can achieve the advantages of combining the Gauss–Newton algorithm and the gradient descent method, improving upon the disadvantages of both by modifying the parameters during execution. In the process of building the model, training is automatically stopped when the generalization ability stops improving [78].

(2) Gaussian process regression (GPR)

The purpose of GPR is the same as the most common linear regression: to learn the training data, reconstruct the new data, and expect that the reconstructed values will be as close to the true values as possible. Gaussian regression is a non-parametric model, so it will not be over-fitted [79]. Gaussian regression can be understood from two perspectives: the weight and the function. The function perspective involves a random variable whose function obeys the Gaussian process, which can optimize the effect of the "black box" function [80].

(3) Support vector machine (SVM)

Compared with logistic regression and NN, SVM provides a clearer and more powerful way to learn complex non-linear equations [81]. An SVM models the training data, then reconstructs new data, with the expectation that the reconstructed value is as close to the real value as possible. SVM problems are usually solved using convex optimization technology [82].

(4) Random forest (RF)

RF regression is learning from the characteristic matrix. The RF algorithm is an integrated technology that combines multiple decision trees, which helps to reduce the variability of the model. Different training data subsets train each decision tree, and each node of each decision tree is divided according to random attributes in the data [83]. One of the most important problems when using the RF algorithm is parameter adjustment. The parameters to be tuned include the number of trees, leaves, and feet in the random forest. When testing for the optimal value of a single variable, the other variables are kept unchanged. After finding the optimal value of one parameter, it is used in the subsequent tests to find the optimal values for the other parameters [84].

2.2.2. Factor Determination

(1) Daily data factor determination

Previous studies have carried out a lot of work on the reconstruction of the T2m from the LST or IST, and many different influencing factors have been considered. For example, Nielsen-Englyst et al. [85] used season, wind speed, clouds, and other factors to reconstruct the T2m in the Arctic; Meyer et al. [44] used IST and month factors to reconstruct the T2m in Antarctica; Pang et al. [86] used IST, radiation factors, seasonal factors, wind speed, cloud cover, and other factors to simulate T2m in Dome A; and Zhang et al. [54] used IST, cloud cover, wind speed, and other factors to estimate T2m in the Tibetan Plateau. Therefore, according to previous experience, we used eight factors—IST, specific humidity, wind speed, albedo, elevation, longitude, month, and latitude—to reconstruct the T2m.

(2) Monthly data factor determination

Monthly data for IST and albedo are available from MODIS; however, the albedo is severely lacking. As albedo has a very important effect on T2m, the albedo from RACMO2.3p2 was used, instead of the MODIS albedo data. After pre-processing all the above data, the model was built according to the format of the monthly data, which were reconstructed using the eight factors listed above (IST, specific humidity, albedo, altitude, wind speed, longitude, latitude, and month).

2.2.3. Feature Ranking and Model Construction

By using the feature ranking method of RF in machine learning [83], the importance of the eight factors was ranked, and the model was established according to the ranking.

(1) Establishment of the daily data model

The variable importance scores for the complete model using all predictor variables indicated the importance of IST for predicting T2m (Figure 3a). In addition to IST, specific humidity was also important, followed by wind speed, albedo, elevation, longitude, month, and latitude. Through experience, we know that IST and specific humidity are the most influential factors in the reconstruction process from IST to T2m [87], consistent with the conclusions drawn in Figure 3a. Therefore, we built a daily data model according to the abovementioned order of importance.

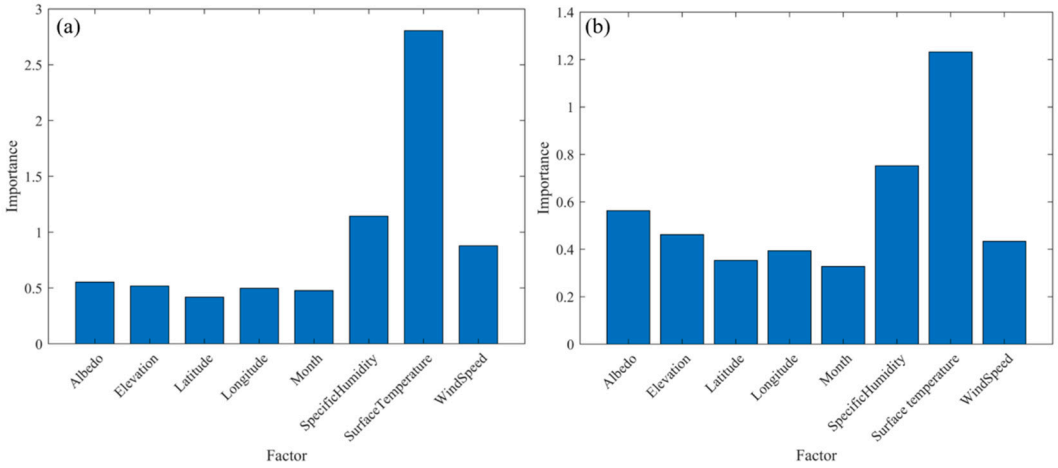

Figure 3. Daily data (**a**) and monthly data (**b**) rankings for eight factors.

The IST and albedo in the model were obtained from MODIS, while the other variables came from the observation data of the AWSs. In the process of reconstruction, the specific humidity and wind speed data were obtained from RACMO2.3p2, as MODIS data cannot provide specific humidity and wind speed data. MODIS provided IST, albedo, elevation, longitude, month, and latitude data.

(2) Establishment of the monthly data model

The order of monthly data importance was (from high to low): IST, specific humidity, albedo, elevation, wind speed, longitude, latitude, and month. As can be seen from Figure 3b, IST, specific humidity, albedo, and elevation were the four most relatively important influencing factors, followed by the other four factors. Similarly, we established a model for the monthly data, according to the order of importance. The IST and albedo were obtained from MODIS, while the other variables were obtained from AWS observations. In the process of reconstruction, the albedo, specific humidity, and wind speed data were obtained from RACMO2.3p2, as MODIS data cannot provide complete albedo, specific humidity, and wind speed data. According to this model, the T2m monthly data for the GrIS from 2007 to 2019 were reconstructed.

Comparison of the daily data model with the monthly data model indicated that the difference was not significant, which is in line with meteorological theory. We can see, from the importance ranking of the eight monthly data variables, that albedo is a reasonably essential factor in the process of monthly data reconstruction, so we will continue to analyze it in the following parts. To assess the significance of lacking albedo, we reconstructed the T2m from monthly data without albedo. Thus, IST, specific humidity, wind speed, latitude, longitude, elevation, and month are the seven factors employed in this reconstruction procedure. This issue will be examined in detail in Section 4.1. Figure 4 shows the T2m reconstruction process for the GrIS based on the four machine learning methods used in this paper.

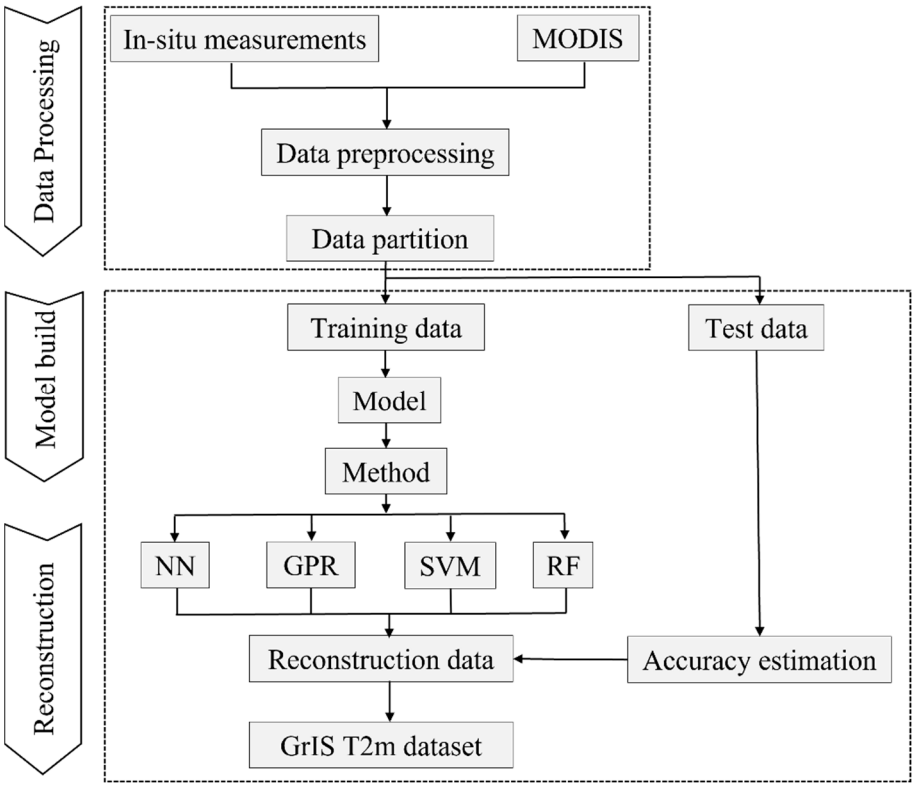

Figure 4. The GrIS T2m reconstruction process based on four machine learning methods.

3. Results

3.1. Daily Data Reconstruction Products

The fitting effects of the four models are shown in Figure 5. The models were evaluated using R, RMSE, and BIAS, compared with the in-situ data. Based on the reconstruction results for the whole GrIS, the scatter plot was obtained by extracting the reconstruction data near the 15 AWSs and comparing them with the observed data.

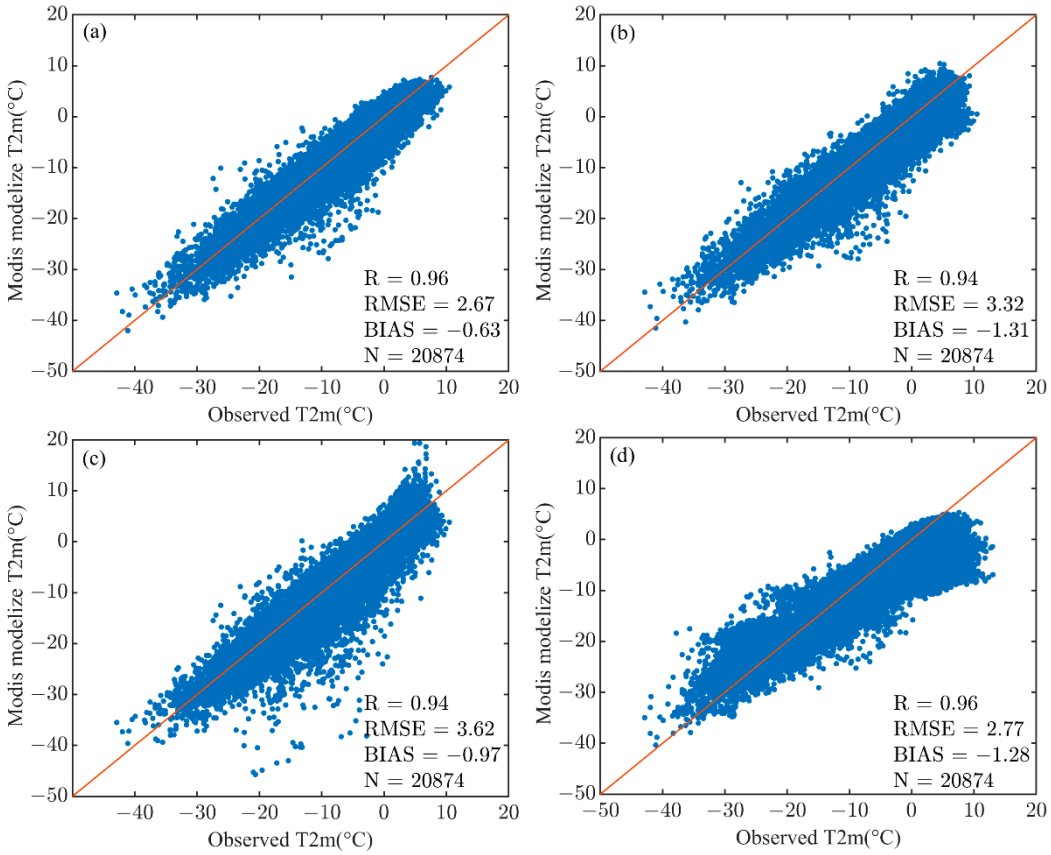

Figure 5. Comparison of reconstruction results of NN (**a**), GPR (**b**), SVM (**c**), and RF (**d**) against the observed T2m. R, correlation coefficient; RMSE (°C), root mean square error; BIAS (°C), mean difference; and N, number of samples.

The NN was considered to be the best-performing algorithm (see Figure 5a) with an R value of 0.96, an RMSE of 2.67 °C, and BIAS of −0.63 °C. Therefore, the tuned and trained NN model was selected as the final model to create the T2m product. The RF performed slightly worse than the NN (R = 0.96, RMSE = 2.77 °C, BIAS = −1.28 °C) and showed little difference from the GPR (R = 0.94, RMSE = 3.32 °C, BIAS = −1.31 °C) model. SVM performed slightly worse than the other algorithms, presenting the lowest performance (R = 0.94, RMSE = 3.62 °C, BIAS = −0.97 °C). The derived data set included daily reconstructs of T2m on a 780 m regular latitude and longitude grid for the period 2017–2019. Our daily data reconstructions surpassed the accuracy of previous reconstructions, such as those of Meyer et al. [44] and Nielsen-Englyst et al. [32].

As can be seen from Figure 5, each method had some outliers, and the reconstructed T2m was generally lower than the observed T2m, which may be due to the abnormal MODIS IST data [69]. Extreme outliers occurred in the SVM method, and the temperature difference may reach more than 20 °C; this may be due to the algorithm itself. When the error function is reduced to the minimum value, the phenomenon of over-fitting can be effectively prevented; however, the accuracy will inevitably be affected, leading to extreme values in the reconstruction results [88]. Evaluation of the RF reconstruction results in Figure 5d eliminated the filled missing values.

To more clearly compare the reconstruction quality of each reconstruction method, we constructed box plots, which are shown in Figure 6. Figure 6 shows a plot of the reconstruction results for each method minus the AWS observed data. Through comparative analysis of the box diagrams, it can be seen that the difference between the results of the NN model reconstruction was closest to 0 °C, which also suggests that the reconstruction accuracy of the NN model was the best.

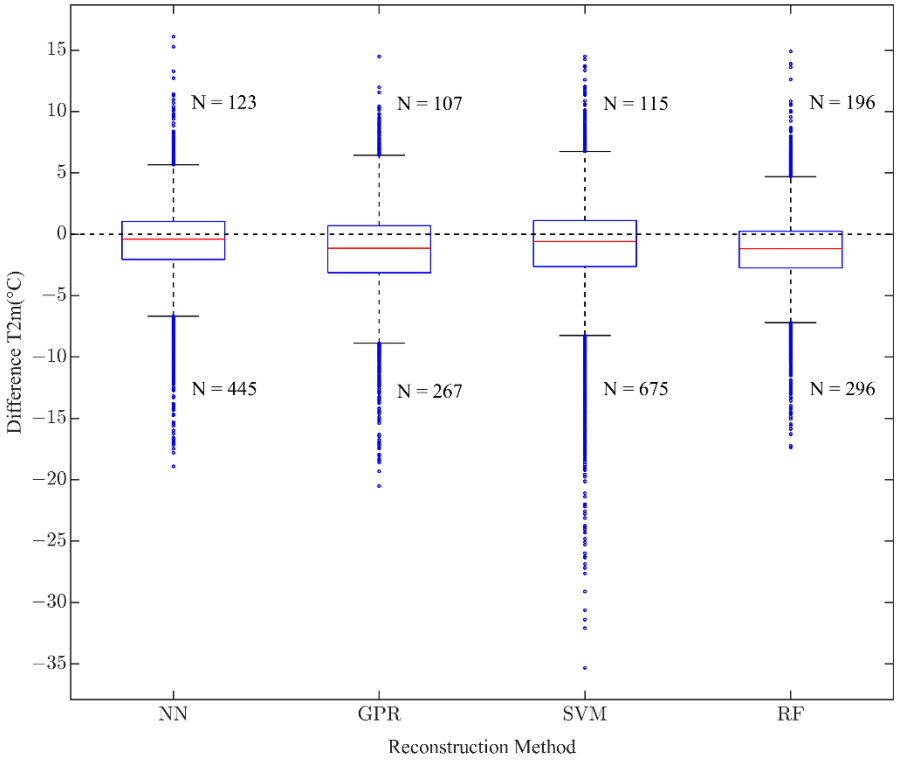

Figure 6. Difference between reconstruction results and observed data at AWSs. The red horizontal line in the rectangular box indicates the median; the closer it is to 0 °C, the more accurate the reconstruction results. The upper end of the box represents the quartile data (i.e., the first 25% of the data). The lower end of the box represents three-quarters of the data (i.e., the first 75% of the data). The upper black horizontal line indicates the upper edge of the data, and the lower black horizontal line indicates the lower edge of the data. The combination of the upper and lower edges indicates the normal distribution range for this data set. The blue dots indicate outliers, and N represents the number of outliers.

Notably, there were many outliers in the reconstruction using the SVM model, but most of the data were within a reasonable range. Compared with the NN reconstruction results, the T2m in the GPR reconstruction results was lower, and the reconstruction results of the RF method also tended to be cold. It may be that the overall reconstruction temperature is low, as there were more outliers with extremely low temperatures than extremely high temperatures. The number of outlier samples was small, accounting for only approximately 3% of the total number of samples.

3.2. Monthly Data Reconstruction Products

On the basis of our reconstruction, we investigated the spatial variation of the GrIS T2m during 2018. The reconstructions were compared with data from RACMO2.3p2 in order to further confirm the reliability of our reconstructions. The NN reconstruction results were close to the RACMO2.3p2, and even local application of the NN for T2m reconstruction better reflected the changes.

In particular, the T2m in northern Greenland was obviously lower than that in southern Greenland. Furthermore, in July, the T2m in southwest Greenland was above 0 °C, which was preserved in our reconstructions. However, in the northern part of Greenland in January and December, the reconstruction of low temperature data was not as accurate, which may be related to the data and methods used [44]; although MODIS IST products include cloud cover masking, we realize that cirrus clouds cannot be reliably removed from the data. Cloud classification over the GrIS is challenging in winter, due to "cold-on-cold" and "white-on-white" problems [89]. This leads to a more error-prone T2m reconstruction

in winter than in summer. In addition, Janatian et al. [90] reported that the performance of a model for predicting T2m at low temperatures may be significantly decreased.

From Figure 7, we can see that the reconstruction results for monthly data wellreflected the latitude zonality of T2m. At the same time, they also reflect the T2m difference between coastal and inland regions, with higher temperature in coastal areas and lower temperature in inland areas. The longitude zonation and the spatial distribution of T2m are also shown. In all months, the northern and inner parts of the GrIS are generally colder than other regions, while the warmest areas are the coastal areas and the melting zone of the GrIS.

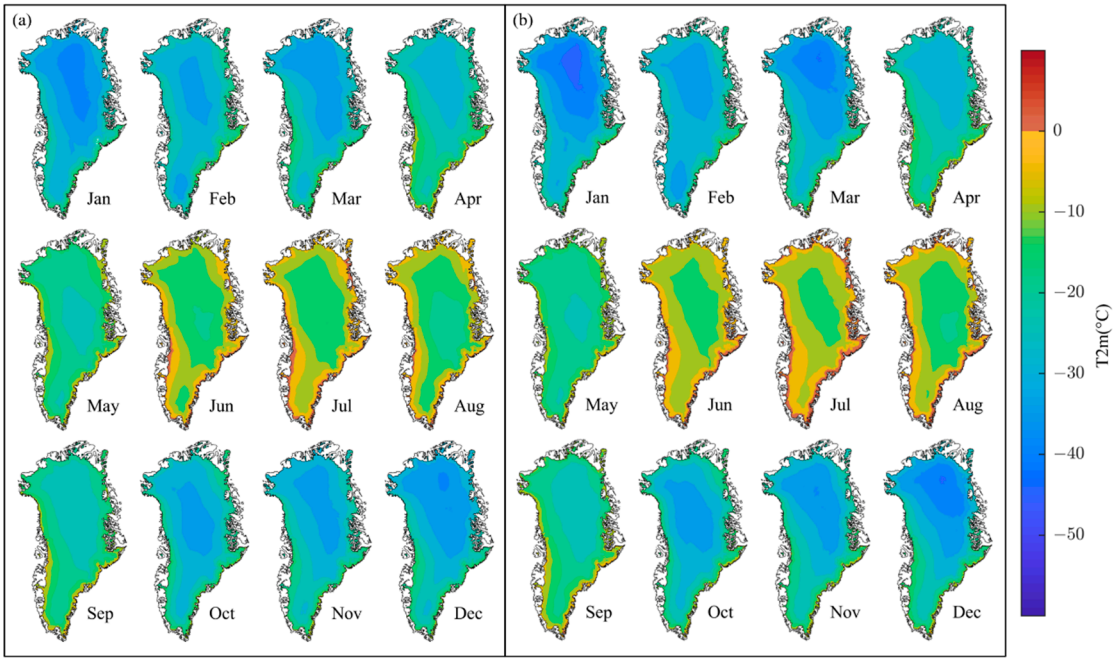

Figure 7. T2m of the reconstruction results of NN (**a**) model for monthly data in 2018, and T2m of monthly data of RACMO2.3p2 (**b**) for 2018.

Figure 8a,b show the spatial distribution of our reconstructed monthly T2m linear trends and RACMO2.3p2 for 2007–2019, respectively. The trend values of mean monthly temperatures varied from somewhat to very different between regions. The clear cooling trends could be observed for February, July, and December, which were almost the same as RACMO2.3p2. However, in January, RACMO2.3p2 showed a significant warming trend, while the reconstructed trend was not show. The trends of T2m in any month between 2007 and 2019 were spatially heterogeneous (Figure 8). Both the reconstruction and RACMO2.3p2 showed that the trend had the largest significant warming patch in April and showed that the cooling trend occurs mainly in northern and central Greenland in December, while there is a warming trend in parts of the south. The trends of reconstructions and RACMO2.3p2 were very similar in the spatial distribution. For example, in February and December, the majority of grids experienced a cooling trend, while, in April, May, October, and November, the opposite happens.

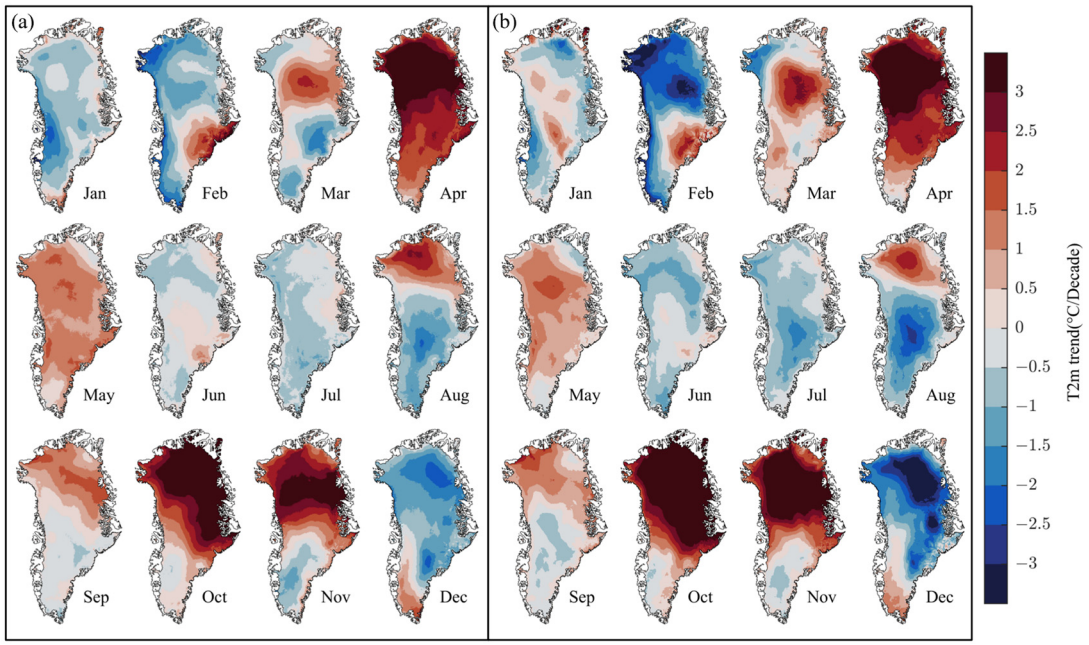

Figure 8. Monthly T2m trend (°C decade^{-1}) from 2007 to 2019 from NN Reconstruction (**a**) and RACMO2.3p2 (**b**).

3.3. Evaluation of the Reconstruction Products Performance

Based on the R, RMSE, and BIAS from the individual AWS daily and monthly data, it can be seen that it is feasible and accurate to reconstruct T2m using an NN. All R values were greater than 0.9, the RMSE values were less than 3.5 °C, and the BIAS values were less than 2 °C and more than −2.5 °C (mostly near 0 °C). Overall, the precision between observed and reconstructed values was significantly lower for the daily data than for the monthly data (Figure 9). Only a few AWSs had lower precision.

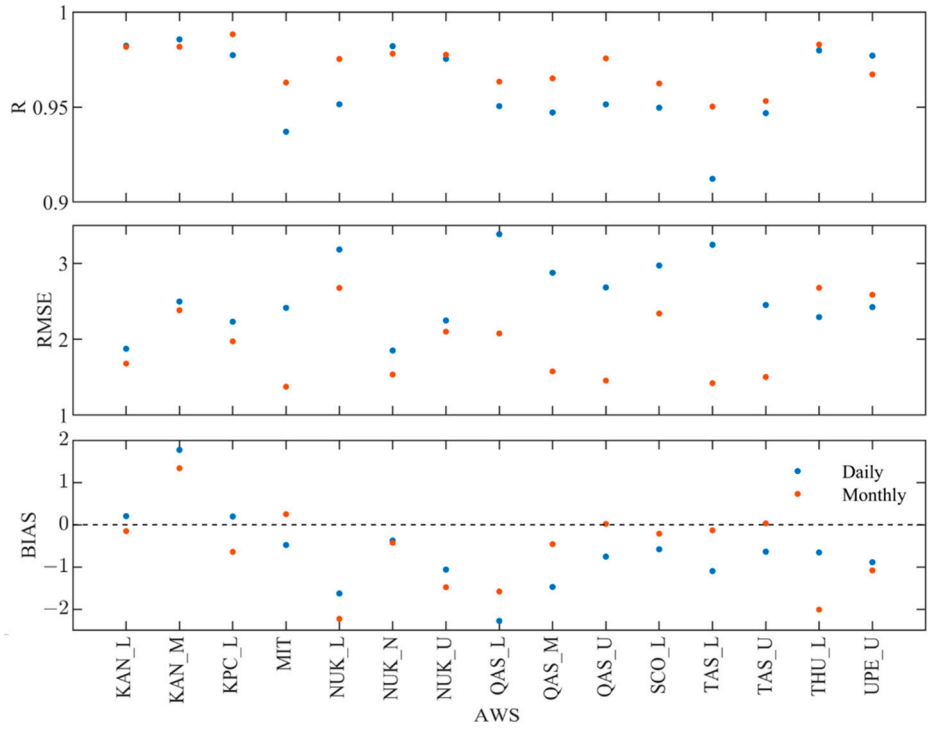

Figure 9. Evaluation of reconstruction results for each AWS, with RMSE and BIAS in °C.

The overall reconstruction accuracy for the NUK_L and QAS_L AWSs was low, while the reconstruction accuracy for the KAN_L and NUK_N AWSs was relatively high. The reconstruction accuracy for the monthly data was higher for the MIT and QAS_U stations, while the reconstruction accuracy for the daily data is significantly better than that of the monthly data for the UPE_U station. The nearest neighbor interpolation method that we used may have had some influence on the verification of the reconstruction results.

4. Discussion

4.1. The Albedo Scheme

The NN taking IST, specific humidity, wind speed, albedo, latitude, longitude, elevation, and month as reconstruction variables was the best-performing model. Due to the serious lack of monthly albedo data from MODIS data, the albedo of RACMO2.3p2 was used to replace the MODIS albedo when reconstructing monthly data; however, the reconstruction results were worth comparing. Therefore, the influence of albedo was decisively abandoned, and a combination of the remaining seven factors was also considered.

According to the importance ranking in Section 2.2.3, we removed the albedo and sorted the remaining seven factors, as shown in Figure 10. The order of importance was (from high to low): IST, specific humidity, wind speed, longitude, elevation, latitude, and month. Among them, the IST data were obtained from MODIS, while other data were obtained from the AWSs. In the reconstruction process, wind speed and specific humidity data were obtained from RACMO2.3p2. Among the seven factors, IST, wind speed, and specific humidity were the three most important factors, differing from the results obtained for the previous eight factors. This indicates that a lack of albedo may have a great impact on the reconstruction results [87,91].

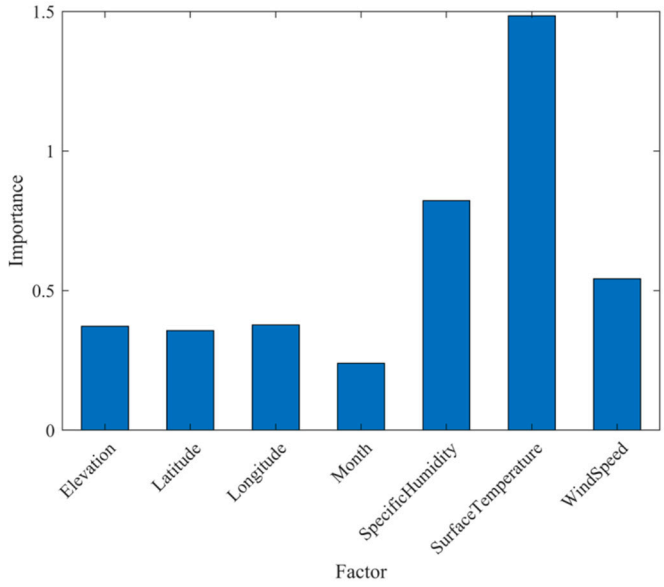

Figure 10. Importance ranking of seven factors (without albedo): IST, specific humidity, wind speed, month, elevation, latitude, and longitude.

It is clear from Figure 11 that the reconstruction accuracy decreased without albedo. The reconstruction at very low temperatures was poor. The T2m was lower in the middle than in the north, which is obviously unrealistic. It does not conform to the law of latitudinal zonation. Simulations of very high temperatures were even worse; for example, values above 0 °C in June, July, and August were not reconstructed. In addition, the spatial pattern of T2m was problematic. Regardless of the season, the reconstructions without the albedo were always wrong and the reconstruction results were problematic, both in general and in local areas. All of these problems are likely due to the lack of albedo in the

reconstruction process, resulting in poor reconstruction results. This also suggests that albedo is an important factor in T2m reconstruction.

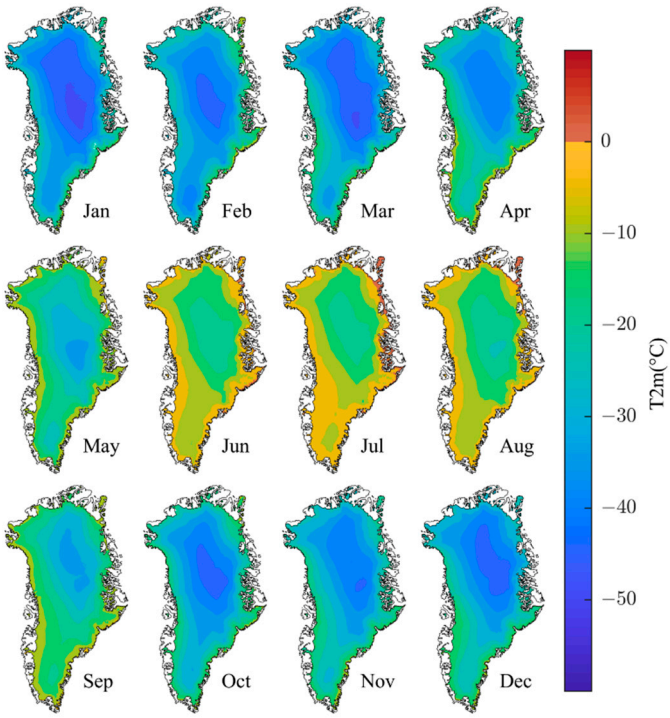

Figure 11. T2m of the reconstruction results for the NN model of monthly data in 2018.

4.2. Comparison with Other Studies

Direct comparison of the accuracy of our reconstructed T2m products for the GrIS with those presented in other studies is difficult, not only because of the different environments and time scales, but also (and most notably) due to the differences in the validation methods used. The verification method in this paper was cross-validation, which can prevent overfitting by dividing the data set into several folds and estimating the accuracy of each fold [44].

Recursive feature elimination is the most commonly used method for feature selection, using the importance scores of the variables calculated from the training subset [92]. Pichierri et al. [93] used stepwise regression with forward selection to find the best estimation algorithm. The selection of factors was carried out using this method. The approach used in the present paper involved ranking the machine learning methods for RF feature selection, according to the importance of the relationship between each factor and the T2m. This ranking allowed us to derive which factors are of high importance and which are of low importance. Models were then built based on the importance of each factor, which is more reasonable and plausible.

Jang et al. [77] used multi-layer feed-forward (MLF) neural networks to estimate air temperatures over southern Canada using Advanced Very High Resolution Radiometer (AVHRR) data. Compared with previous results, the quality of the daily data in this paper was improved significantly, with increased R, decreased RMSE and BIAS, and a significant improvement in the overall results. Compared with the results of Nielsen-Englyst et al. [32], our RMSE was significantly lower and the resolution was improved. Compared with the results of Meyer et al. [44], the R value was greatly improved, the RMSE was greatly reduced, and the reconstruction accuracy was greatly improved. Compared with the results of Bai et al. [91] and Mao et al. [87], the R value was greatly improved in this paper. In conclusion, our reconstruction accuracy is higher than that obtained by our predecessors.

4.3. Limitations and Future Perspectives

The machine learning approach has some limitations in that it requires a large amount of training data [75]. Due to the harsh climate conditions in Greenland, the observations are difficult to access and have more data gaps. Therefore, training data usually cannot meet the need of machine learning [94], limiting the further improvement of the T2m reconstruction. In the future, with more available observations, the accuracy of machine learning reasonably will be improved. In addition, the simulation of extreme T2m by machine learning methods is limited [44]. For example, our reconstructions do not show low temperatures in March in northern Greenland (Figure 7), while RACMO2.3p2 shows low temperatures. Moreover, for machine learning, there is still a subject of debate regarding how the temporal arrangement of training data affect the reconstruction, potentially causing the bias [53]. In the future, we will attempt to disrupt the time series of data for comparison to better reconstruct the T2m. Notably, many machine learning methods are black-box models not to be clearly understood [80] and we could not analyze the impact of every selected climate factor [86]. The unreasonable selection of characteristic variables can also cause errors; therefore, more relevant variables need to be introduced into the model. In this study, the result of reconstruction using eight factors is significantly better than that using seven factors. Nevertheless, a study suggested that inclusion of counterproductive variables could cause over-fitting and worsen reconstruction performance [95], so we cannot just blindly add variables either. In the future, we will attempt to increase relevant influencing factors to explore whether the reconstruction accuracy will be improved.

Previous studies proved the IST from satellite datasets, such as thermal infrared satellite sensors, AVHRR, and Landsat, can be used to reconstruct T2m using linear and machine learning [32,59,83]. This means we can use a variety of remote sensing data to reconstruct T2m on the basis of machine learning methods. By comparing the reconstruction results from different satellite datasets, it is possible to obtain more accurate T2m than MODIS. In addition, T2m is considered as baseline data for understanding GrIS climate variability and reconstructing SMB [21–25,96], suggesting the possibility of SMB reconstruction over the GrIS based on multi-source data and machine learning methods.

5. Conclusions

T2m is one of the key indicators of climate change in the Arctic. In this study, we proposed the use of four machine learning methods—NN, GPR, SVM, and RF—to reconstruct the T2m from MODIS IST and other auxiliary variables. IST, specific humidity, wind speed, albedo, altitude, longitude, month, and latitude were found to be the most suitable and reliable influencing factors for reconstructing T2m. In terms of model accuracy, the NN algorithm performed the best, with R of 0.96, RMSE of 2.67 °C, and BIAS of −0.63 °C. With the developed model, we can provide the daily and monthly data for the period 2007–2019 over the GrIS with 780 m resolution. We used MODIS data for T2m reconstruction, which can greatly improve the spatial resolution of the reconstructed data and make up for the shortage of in situ observations. These T2m data can provide data support for other subsequent research considering the GrIS (e.g., SMB, energy balance, surface melting, etc.). Future research should focus on improving the accuracy of the reconstructions then extending the product to the entire lifetime of the MODIS sensor.

Author Contributions: Conceptualization, B.H. and M.D.; data curation, Q.Z. and B.H.; investigation, Y.W. (Yetang Wang) and M.D.; methodology, Y.W. (Yuzhe Wang) and L.W.; writing—original draft, J.C.; software, Q.Z. and W.S.; writing—review and editing, B.H., W.S. and Y.W. (Yuzhe Wang). All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Natural Science Foundation of China (42171121, 42271145, 41901088, and 42101120) and the open fund of Key Laboratory of Oceanic Atmospheric Chemistry and Global Change, Ministry of Natural Resources (GCMAC2206).

Data Availability Statement: The PROMICE observations are available online (http://promice.org/download-data/ (accessed on 7 September 2021). MODIS data are obtained from National Snow and Ice Data Center (https://nsidc.org/data/modis (accessed on 7 September 2021).

Acknowledgments: The authors gratefully acknowledge support from data availability from PROMICE and MODIS. The authors thank Brice Noël (Utrecht University) for RACMO2.3p2 data support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Serreze, M.C.; Francis, J.A. The Arctic Amplification Debate. *Clim. Chang.* **2006**, *76*, 241–264. [CrossRef]
2. Pithan, F.; Mauritsen, T. Arctic amplification dominated by temperature feedbacks in contemporary climate models. *Nat. Geosci.* **2014**, *7*, 181–184. [CrossRef]
3. Graversen, R.G.; Mauritsen, T.; Tjernstrom, M.; Kallen, E.; Svensson, G. Vertical structure of recent Arctic warming. *Nature* **2008**, *451*, 53–56. [CrossRef] [PubMed]
4. Steffen, K.; Box, J. Surface climatology of the Greenland Ice Sheet: Greenland Climate Network 1995–1999. *J. Geophys. Res. Atmos.* **2001**, *106*, 33951–33964. [CrossRef]
5. Griffiths, M.; Wise, S.; Irvine-Fynn, T.; Shuman, C.; Huff, R.; Cappelen, J.; Steffen, K.; Huybrechts, P.; Hanna, E. Increased Runoff from Melt from the Greenland Ice Sheet: A Response to Global Warming. *J. Clim.* **2008**, *21*, 331–341. [CrossRef]
6. Tedesco, M.; Serreze, M.; Fettweis, X. Diagnosing the extreme surface melt event over southwestern Greenland in 2007. *Cryosphere* **2008**, *2*, 159–166. [CrossRef]
7. Mernild, S.H.; Mote, T.L.; Liston, G.E. Greenland ice sheet surface melt extent and trends: 1960–2010. *J. Glaciol.* **2017**, *57*, 621–628. [CrossRef]
8. Nghiem, S.V.; Hall, D.K.; Mote, T.L.; Tedesco, M.; Albert, M.R.; Keegan, K.; Shuman, C.A.; DiGirolamo, N.E.; Neumann, G. The extreme melt across the Greenland ice sheet in 2012. *Geophys. Res. Lett.* **2012**, *39*, 20. [CrossRef]
9. Hall, D.K.; Comiso, J.C.; DiGirolamo, N.E.; Shuman, C.A.; Box, J.E.; Koenig, L.S. Variability in the surface temperature and melt extent of the Greenland ice sheet from MODIS. *Geophys. Res. Lett.* **2013**, *40*, 2114–2120. [CrossRef]
10. Hanna, E.; Fettweis, X.; Mernild, S.H.; Cappelen, J.; Ribergaard, M.H.; Shuman, C.A.; Steffen, K.; Wood, L.; Mote, T.L. Atmospheric and oceanic climate forcing of the exceptional Greenland ice sheet surface melt in summer 2012. *Int. J. Climatol.* **2014**, *34*, 1022–1037. [CrossRef]
11. Rignot, E.; Velicogna, I.; van den Broeke, M.R.; Monaghan, A.; Lenaerts, J.T.M. Acceleration of the contribution of the Greenland and Antarctic ice sheets to sea level rise. *Geophys. Res. Lett.* **2011**, *38*, 5. [CrossRef]
12. Enderlin, E.M.; Howat, I.M. Submarine melt rate estimates for floating termini of Greenland outlet glaciers (2000–2010). *J. Glaciol.* **2017**, *59*, 67–75. [CrossRef]
13. Fettweis, X.; Hanna, E.; Lang, C.; Belleflamme, A.; Erpicum, M.; Gallée, H. Brief communication: "Important role of the mid-tropospheric atmospheric circulation in the recent surface melt increase over the Greenland ice sheet". *Cryosphere* **2013**, *7*, 241–248. [CrossRef]
14. Wouters, B.; Bamber, J.L.; van den Broeke, M.R.; Lenaerts, J.T.M.; Sasgen, I. Limits in detecting acceleration of ice sheet mass loss due to climate variability. *Nat. Geosci.* **2013**, *6*, 613–616. [CrossRef]
15. Cai, D.; You, Q.; Fraedrich, K.; Guan, Y. Spatiotemporal Temperature Variability over the Tibetan Plateau: Altitudinal Dependence Associated with the Global Warming Hiatus. *J. Clim.* **2017**, *30*, 969–984. [CrossRef]
16. Hänninen, H.; Zhang, G.; Rikala, R.; Luoranen, J.; Konttinen, K.; Repo, T. Frost hardening of Scots pine seedlings in relation to the climatic year-to-year variation in air temperature. *Agric. For. Meteorol.* **2013**, *177*, 1–9. [CrossRef]
17. Kollas, C.; Randin, C.F.; Vitasse, Y.; Körner, C. How accurately can minimum temperatures at the cold limits of tree species be extrapolated from weather station data? *Agric. For. Meteorol.* **2014**, *184*, 257–266. [CrossRef]
18. Wang, L.; Sun, L.; Shrestha, M.; Li, X.; Liu, W.; Zhou, J.; Yang, K.; Lu, H.; Chen, D. Improving snow process modeling with satellite-based estimation of near-surface-air-temperature lapse rate. *J. Geophys. Res. Atmos.* **2016**, *121*, 12005–12030. [CrossRef]
19. Kang, S.; Xu, Y.; You, Q.; Flügel, W.-A.; Pepin, N.; Yao, T. Review of climate and cryospheric change in the Tibetan Plateau. *Environ. Res. Lett.* **2010**, *5*, 015101. [CrossRef]
20. Zhang, F.; Zhang, H.; Hagen, S.C.; Ye, M.; Wang, D.; Gui, D.; Zeng, C.; Tian, L.; Liu, J. Snow cover and runoff modelling in a high mountain catchment with scarce data: Effects of temperature and precipitation parameters. *Hydrol. Process.* **2015**, *29*, 52–65. [CrossRef]
21. Zhang, Q.; Huai, B.; van den Broeke, M.R.; Cappelen, J.; Ding, M.; Wang, Y.; Sun, W. Temporal and Spatial Variability in Contemporary Greenland Warming (1958–2020). *J. Clim.* **2022**, *35*, 2755–2767. [CrossRef]
22. Liu, X.; Chen, B. Climatic warming in the Tibetan Plateau during recent decades. *Int. J. Climatol.* **2000**, *20*, 1729–1742. [CrossRef]
23. Xu, W.; Liu, X. Response of vegetation in the Qinghai-Tibet Plateau to global warming. *Chin. Geophys. sci.* **2007**, *17*, 151–159. [CrossRef]
24. Pepin, N.; Bradley, R.S.; Diaz, H.F.; Baraer, M.; Caceres, E.B.; Forsythe, N.; Fowler, H.; Greenwood, G.; Hashmi, M.Z.; Liu, X.D.; et al. Elevation-dependent warming in mountain regions of the world. *Nat. Clim. Chang.* **2015**, *5*, 424–430. [CrossRef]

25. Rückamp, M.; Greve, R.; Humbert, A. Comparative simulations of the evolution of the Greenland ice sheet under simplified Paris Agreement scenarios with the models SICOPOLIS and ISSM. *Polar Sci.* **2019**, *21*, 14–25. [CrossRef]
26. Appelhans, T.; Mwangomo, E.; Hardy, D.R.; Hemp, A.; Nauss, T. Evaluating machine learning approaches for the interpolation of monthly air temperature at Mt. Kilimanjaro, Tanzania. *Spat. Stat.* **2015**, *14*, 91–113. [CrossRef]
27. Hofstra, N.; Haylock, M.; New, M.; Jones, P.; Frei, C. Comparison of six methods for the interpolation of daily, European climate data. *J. Geophys. Res.* **2008**, *113*, D21. [CrossRef]
28. Stahl, K.; Moore, R.D.; Floyer, J.A.; Asplin, M.G.; McKendry, I.G. Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. *Agric. For. Meteorol.* **2006**, *139*, 224–236. [CrossRef]
29. Noël, B.; van de Berg, W.J.; van Wessem, J.M.; van Meijgaard, E.; van As, D.; Lenaerts, J.T.M.; Lhermitte, S.; Kuipers Munneke, P.; Smeets, C.J.P.P.; van Ulft, L.H.; et al. Modelling the climate and surface mass balance of polar ice sheets using RACMO2–Part 1: Greenland (1958–2016). *Cryosphere* **2018**, *12*, 811–831. [CrossRef]
30. Noël, B.; van de Berg, W.J.; Lhermitte, S.; van den Broeke, M.R. Rapid ablation zone expansion amplifies north Greenland mass loss. *Sci. Adv.* **2019**, *5*, eaaw0123. [CrossRef]
31. Huai, B.; van den Broeke, M.R.; Reijmer, C.H. Long-term surface energy balance of the western Greenland Ice Sheet and the role of large-scale circulation variability. *Cryosphere* **2020**, *14*, 4181–4199. [CrossRef]
32. Nielsen-Englyst, P.; Høyer, J.L.; Madsen, K.S.; Tonboe, R.T.; Dybkjær, G.; Skarpalezos, S. Deriving Arctic 2 m air temperatures over snow and ice from satellite surface temperature measurements. *Cryosphere* **2021**, *15*, 3035–3057. [CrossRef]
33. Leeson, A.A.; Van Wessem, J.M.; Ligtenberg, S.R.M.; Shepherd, A.; Van Den Broeke, M.R.; Killick, R.; Skvarca, P.; Marinsek, S.; Colwell, S. Regional climate of the Larsen B embayment 1980–2014. *J. Glaciol.* **2017**, *63*, 683–690. [CrossRef]
34. Medley, B.; Joughin, I.; Das, S.B.; Steig, E.J.; Conway, H.; Gogineni, S.; Criscitiello, A.S.; McConnell, J.R.; Smith, B.E.; van den Broeke, M.R.; et al. Airborne-radar and ice-core observations of annual snow accumulation over Thwaites Glacier, West Antarctica confirm the spatiotemporal variability of global and regional atmospheric models. *Geophys. Res. Lett.* **2013**, *40*, 3649–3654. [CrossRef]
35. Li, Z.; Tang, B.; Wu, H.; Ren, H.; Yan, G.; Wan, Z.; Trigo, I.F.; Sobrino, J.A. Satellite-derived land surface temperature: Current status and perspectives. *Remote Sens. Environ.* **2013**, *131*, 14–37. [CrossRef]
36. Comiso, J.C. Detection of change in the Arctic using satellite and in situ data. *J. Geophys. Res.* **2003**, *108*, C12. [CrossRef]
37. Wang, X.; Key, J.R. Recent trends in Arctic surface, cloud, and radiation properties from space. *Science* **2003**, *299*, 1725–1728. [CrossRef]
38. Wang, X.; Key, J.R. Arctic Surface, Cloud, and Radiation Properties Based on the AVHRR Polar Pathfinder Dataset. Part I: Spatial and Temporal Characteristics. *J. Clim.* **2005**, *18*, 2558–2574. [CrossRef]
39. Wang, X.; Key, J.R. Arctic Surface, Cloud, and Radiation Properties Based on the AVHRR Polar Pathfinder Dataset. Part II: Recent Trends. *J. Clim.* **2005**, *18*, 2575–2593. [CrossRef]
40. Benali, A.; Carvalho, A.C.; Nunes, J.P.; Carvalhais, N.; Santos, A. Estimating air surface temperature in Portugal using MODIS LST data. *Remote Sens. Environ.* **2012**, *124*, 108–121. [CrossRef]
41. Cristóbal, J.; Ninyerola, M.; Pons, X. Modeling air temperature through a combination of remote sensing and GIS data. *J. Geophys. Res.* **2008**, *113*, D13. [CrossRef]
42. Kilibarda, M.; Hengl, T.; Heuvelink, G.B.M.; Gräler, B.; Pebesma, E.; Perčec Tadić, M.; Bajat, B. Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution. *J. Geophys. Res. Atmos.* **2014**, *119*, 2294–2313. [CrossRef]
43. Lin, X.; Zhang, W.; Huang, Y.; Sun, W.; Han, P.; Yu, L.; Sun, F. Empirical Estimation of Near-Surface Air Temperature in China from MODIS LST Data by Considering Physiographic Features. *Remote Sens.* **2016**, *8*, 629. [CrossRef]
44. Meyer, H.; Katurji, M.; Appelhans, T.; Müller, M.; Nauss, T.; Roudier, P.; Zawar-Reza, P. Mapping Daily Air Temperature for Antarctica Based on MODIS LST. *Remote Sens.* **2016**, *8*, 732. [CrossRef]
45. Park, S. Integration of satellite-measured LST data into cokriging for temperature estimation on tropical and temperate islands. *Int. J. Climatol.* **2011**, *31*, 1653–1664. [CrossRef]
46. Peón, J.; Recondo, C.; Calleja, J.F. Improvements in the estimation of daily minimum air temperature in peninsular Spain using MODIS land surface temperature. *Int. J. Remote Sens.* **2014**, *35*, 5148–5166. [CrossRef]
47. Pepin, N.C.; Maeda, E.E.; Williams, R. Use of remotely sensed land surface temperature as a proxy for air temperatures at high elevations: Findings from a 5000 m elevational transect across Kilimanjaro. *J. Geophys. Res. Atmos.* **2016**, *121*, 9998–10015. [CrossRef]
48. Zhang, W.; Huang, Y.; Yu, Y.; Sun, W. Empirical models for estimating daily maximum, minimum and mean air temperatures with MODIS land surface temperatures. *Int. J. Remote Sens.* **2011**, *32*, 9415–9440. [CrossRef]
49. Liu, C.; Cao, G.; Zhang, M.; Niu, X.; Xu, W.; Fan, J. Influence of Temporal and Spatial Uariability on Estimation of Air Temperatures from MODIS Land Surface Temperatures. *Remote Sens. Technol. Appl.* **2014**, *28*, 831–835.
50. Qu, P.; Shi, R.; Liu, C.; Zhong, H. The Evaluation of MODIS Data and Geographic Data for Estimating Near Surface Air Temperature. *Remote Sens. Nat. Resour.* **2011**, *23*, 78–82. [CrossRef]
51. Fu, G.; Shen, Z.; Zhang, X.; Shi, P.; Zhang, Y.; Wu, J. Estimating air temperature of an alpine meadow on the Northern Tibetan Plateau using MODIS land surface temperature. *Acta. Ecol. Sin.* **2011**, *31*, 8–13. [CrossRef]

52. Xu, Y.; Knudby, A.; Ho, H.C. Estimating daily maximum air temperature from MODIS in British Columbia, Canada. *Int. J. Remote Sens.* **2014**, *35*, 8108–8121. [CrossRef]
53. Zhou, B.; Erell, E.; Hough, I.; Rosenblatt, J.; Just, A.C.; Novack, V.; Kloog, I. Estimating near-surface air temperature across Israel using a machine learning based hybrid approach. *Int. J. Climatol.* **2020**, *40*, 6106–6121. [CrossRef]
54. Zhang, H.; Zhang, F.; Ye, M.; Che, T.; Zhang, G. Estimating daily air temperatures over the Tibetan Plateau by dynamically integrating MODIS LST data. *J. Geophys. Res. Atmos.* **2016**, *121*, 11425–11441. [CrossRef]
55. Hooker, J.; Duveiller, G.; Cescatti, A. A global dataset of air temperature derived from satellite remote sensing and weather stations. *Sci. Data* **2018**, *5*, 180246. [CrossRef]
56. Arévalo, A.; Niño, J.; Hernández, G.; Sandoval, J. High-Frequency Trading Strategy Based on Deep Neural Networks. *Int. Conf. Intell. Comput.* **2016**, *9773*, 424–436.
57. Amodei, D.; Ananthanarayanan, S.; Anubhai, R.; Bai, J.; Battenberg, E.; Case, C.; Casper, J.; Catanzaro, B.; Cheng, Q.; Chen, G.; et al. Deep Speech 2: End-to-End Speech Recognition in English and Mandarin. In Proceedings of the 33rd International Conference on Machine Learning, New York, NY, USA, 19–24 June 2016; Volume 48, pp. 173–182.
58. Meyer, H.; Schmidt, J.; Detsch, F.; Nauss, T. Hourly gridded air temperatures of South Africa derived from MSG SEVIRI. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *78*, 261–267. [CrossRef]
59. Choi, S.; Jin, D.; Seong, N.-H.; Jung, D.; Sim, S.; Woo, J.; Jeon, U.; Byeon, Y.; Han, K.-S. Near-Surface Air Temperature Retrieval Using a Deep Neural Network from Satellite Observations over South Korea. *Remote Sens.* **2021**, *13*, 4334. [CrossRef]
60. Ahlstrøm, A.P. A new programme for monitoring the mass loss of the Greenland ice sheet. *GEUS Bull.* **2008**, *15*, 61–64. [CrossRef]
61. Fausto, R.S.; van As, D.; Mankoff, K.D.; Vandecrux, B.; Citterio, M.; Ahlstrøm, A.P.; Andersen, S.B.; Colgan, W.; Karlsson, N.B.; Kjeldsen, K.K.; et al. Programme for Monitoring of the Greenland Ice Sheet (PROMICE) automatic weather station data. *Earth Syst. Sci. Data.* **2021**, *13*, 3819–3845. [CrossRef]
62. Zhang, X.; Dong, X.; Zeng, J.; Hou, S.; Smeets, P.C.J.P.; Reijmer, C.H.; Wang, Y. Spatiotemporal Reconstruction of Antarctic Near-Surface Air Temperature from MODIS Observations. *J. Clim.* **2022**, *35*, 5537–5553. [CrossRef]
63. Barnes, W.L.; Pagano, T.S.; Salomonson, V.V. Prelaunch characteristics of the moderate resolution imaging spectroradiometer (MODIS) on EOS-AM1. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1088–1100. [CrossRef]
64. Comiso, J.C.; Hall, D.K.; DiGirolamo, N.E.; Shuman, C.A.; Key, J.R.; Koenig, L.S. A Satellite-Derived Climate-Quality Data Record of the Clear-Sky Surface Temperature of the Greenland Ice Sheet. *J. Clim.* **2012**, *25*, 4785–4798. [CrossRef]
65. Hall, D.K.; Cullather, R.I.; DiGirolamo, N.E.; Comiso, J.C.; Medley, B.C.; Nowicki, S.M. A Multilayer Surface Temperature, Surface Albedo and Water Vapor Product of Greenland from MODIS. *Remote Sens.* **2018**, *10*, 555. [CrossRef] [PubMed]
66. Key, J.; Haefliger, M. Arctic ice surface temperature retrieval from AVHRR thermal channels. *J. Geophys. Res. Atmos.* **1992**, *97*, 5885–5893. [CrossRef]
67. Key, J.R.; Collins, J.B.; Fowler, C.; Stone, R.S. High-latitude surface temperature estimates from thermal satellite data. *Remote Sens. Environ.* **1997**, *61*, 302–309. [CrossRef]
68. Duan, S.-B.; Li, Z.-L.; Wu, H.; Leng, P.; Gao, M.; Wang, C. Radiance-based validation of land surface temperature products derived from Collection 6 MODIS thermal infrared data. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *70*, 84–92. [CrossRef]
69. Adolph, A.C.; Albert, M.R.; Hall, D.K. Near-surface temperature inversion during summer at Summit, Greenland, and its relation to MODIS-derived surface temperatures. *Cryosphere* **2018**, *12*, 907–920. [CrossRef]
70. Zikan, K.H.; Adolph, A.C.; Brown, W.P.; Fausto, R.S. Comparison of MODIS surface temperatures to in situ measurements on the Greenland Ice Sheet from 2014 to 2017. *J. Glaciol.* **2022**, 1–12. [CrossRef]
71. Noël, B.; van de Berg, W.J.; van Meijgaard, E.; Kuipers Munneke, P.; van de Wal, R.S.W.; van den Broeke, M.R. Evaluation of the updated regional climate model RACMO2.3: Summer snowfall impact on the Greenland Ice Sheet. *Cryosphere* **2015**, *9*, 1831–1844. [CrossRef]
72. van Wessem, J.M.; van de Berg, W.J.; Noël, B.P.Y.; van Meijgaard, E.; Amory, C.; Birnbaum, G.; Jakobs, C.L.; Krüger, K.; Lenaerts, J.T.M.; Lhermitte, S.; et al. Modelling the climate and surface mass balance of polar ice sheets using RACMO2—Part 2: Antarctica (1979–2016). *Cryosphere* **2018**, *12*, 1479–1498. [CrossRef]
73. Reijmer, C.H. Evaluation of temperature and wind over Antarctica in a Regional Atmospheric Climate Model using 1 year of automatic weather station data and upper air observations. *J. Geophys. Res.* **2005**, *110*, D4. [CrossRef]
74. Huai, B.; van den Broeke, M.R.; Reijmer, C.H.; Cappellen, J. Quantifying rainfall in Greenland: A combined observational and modelling approach. *J. Appl. Meteorol. Climatol.* **2021**, *60*, 1171–1188. [CrossRef]
75. Berral-García, J.L. A quick view on current techniques and machine learning algorithms for big data analytics. In Proceedings of the 2016 18th International Conference on Transparent Optical Networks (ICTON), Trento, Italy, 10–14 July 2016; pp. 1–4.
76. Mitchell, T.; Buchanan, B.; DeJong, G.; Dietterich, T.; Rosenbloom, P.; Waibel, A. Machine learning. *Annu. Rev. Comput. Sci.* **1990**, *4*, 417–433. [CrossRef]
77. Jang, J.D.; Viau, A.A.; Anctil, F. Neural network estimation of air temperatures from AVHRR data. *Int. J. Remote Sens.* **2010**, *25*, 4541–4554. [CrossRef]
78. Wilamowski, B.M.; Yu, H. Improved computation for Levenberg-Marquardt training. *IEEE Trans. Neural Netw.* **2010**, *21*, 930–937. [CrossRef] [PubMed]
79. Seeger, M. Gaussian processes for machine learning. *Int. J. Neural Syst.* **2004**, *14*, 69–106. [CrossRef] [PubMed]

80. Cai, H.; Jia, X.; Feng, J.; Li, W.; Hsu, Y.-M.; Lee, J. Gaussian Process Regression for numerical wind speed prediction enhancement. *Renew. Energy* **2020**, *146*, 2112–2123. [CrossRef]
81. Huang, C.; Davis, L.; Townshend, J. An assessment of support vector machines for land cover classification. *Int. J. Remote Sens.* **2002**, *23*, 725–749. [CrossRef]
82. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2006; Available online: https://link.springer.com/book/9780387310732 (accessed on 10 June 2022).
83. Livingston, F. Implementation of Breiman's random forest machine learning algorithm. *Mach. Learn. J. Pap.* **2005**, 1–13.
84. Rodriguez-Galiano, V.; Sanchez-Castillo, M.; Chica-Olmo, M.; Chica-Rivas, M. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geol. Rev.* **2015**, *71*, 804–818. [CrossRef]
85. Nielsen-Englyst, P.; Høyer, J.L.; Madsen, K.S.; Tonboe, R.; Dybkjær, G.; Alerskans, E. In situ observed relationships between snow and ice surface skin temperatures and 2 m air temperatures in the Arctic. *Cryosphere* **2019**, *13*, 1005–1024. [CrossRef]
86. Pang, X.; Liu, C.; Zhao, X.; He, B.; Fan, P.; Liu, Y.; Qu, M.; Ding, M. Application of Machine Learning for Simulation of Air Temperature at Dome A. *Remote Sens.* **2022**, *14*, 1045. [CrossRef]
87. Mao, K.; Shi, J.; Li, Z.; Tang, H. An RM-NN algorithm for retrieving land surface temperature and emissivity from EOS/MODIS data. *J. Geophys. Res.* **2007**, *112*, D21. [CrossRef]
88. Shlens, J. A tutorial on principal component analysis. *arXiv* **2014**, arXiv:1404.1100.
89. Bromwich, D.H.; Nicolas, J.P.; Monaghan, A.J.; Lazzara, M.A.; Keller, L.M.; Weidner, G.A.; Wilson, A.B. Central West Antarctica among the most rapidly warming regions on Earth. *Nat. Geosci.* **2012**, *6*, 139–145. [CrossRef]
90. Janatian, N.; Sadeghi, M.; Sanaeinejad, S.H.; Bakhshian, E.; Farid, A.; Hasheminia, S.M.; Ghazanfari, S. A statistical framework for estimating air temperature using MODIS land surface temperature data. *Int. J. Climatol.* **2017**, *37*, 1181–1194. [CrossRef]
91. Bai, L.; Xu, Y.; He, M.; Li, N. Remote sensing inversion of near surface air temperature based on random forest. *J. Geo-Inf. Sci.* **2017**, *19*, 390–397. [CrossRef]
92. Kuhn, M.; Johnson, K. *Appl. Predict. Model*; Springer: New York, NY, USA, 2013; Volume 26.
93. Pichierri, M.; Bonafoni, S.; Biondi, R. Satellite air temperature estimation for monitoring the canopy layer heat island of Milan. *Remote Sens. Environ.* **2012**, *127*, 130–138. [CrossRef]
94. Demšar, J. Statistical comparisons of classifiers over multiple data sets. *J. Mach. Learn. Res.* **2006**, *7*, 1–30.
95. Meyer, H.; Reudenbach, C.; Hengl, T.; Katurji, M.; Nauss, T. Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Environ. Model. Softw.* **2018**, *101*, 1–9. [CrossRef]
96. Bolibar, J.; Rabatel, A.; Gouttevin, I.; Galiez, C.; Condom, T.; Sauquet, E. Deep learning applied to glacier evolution modelling. *Cryosphere* **2020**, *14*, 565–584. [CrossRef]