A New Spatially and Temporally Continuous Antarctic Ice-Sheet Surface Temperature Retrieval Method From Passive Microwave Swath Data

Tingting Liu [ID], *Member, IEEE*, Yachao Li [ID], Zemin Wang [ID], and Qiangqiang Yuan [ID], *Member, IEEE*

Abstract—Ice surface temperature (IST) plays a fundamental role in the Antarctic ice sheet/shelf study. However, the production of spatially and temporally continuous Antarctic IST products remains a challenge. We proposed an instantaneous IST retrieval framework that can generate the spatially and temporally continuous Antarctic IST using Advanced Microwave Scanning Radiometer 2 data. To generate a temporally continuous IST product, we developed an innovative scheme, which was based on the acquisition time difference between input and output data. We considered the impact of terrain and sensor observation state. The corresponding parameters were used as the auxiliary variables to improve the model accuracy. We trained and validated nine machine learning models using the generated sample set. The Light Gradient Boosting Machine (LightGBM) model presents the best performance, and the root-mean-square error (RMSE) of the LightGBM model is only half of that of the typical linear models. The RMSE of the LightGBM model decreased with the training sample set size and stabilized at 1.67 K. Further validation using multisource data showed that the IST retrieved using the LightGBM model has RMSEs of 1.39–2.32 K (relative to IST from Landsat-8) and 3.7–5.9 K (relative to IST from Baseline Surface Radiation Network data). Compared to the commonly used ERA5 IST data, the retrieved IST in this study has higher accuracy. We retrieved Antarctic IST from 2013 to 2020. Antarctic IST decreased continuously from 2013 to 2015. After 2015, Antarctic IST increased with large fluctuations.

Index Terms—Antarctic ice surface temperature (IST), light gradient boosting machine (LightGBM) model, machine learning, passive microwave (PM).

I. INTRODUCTION

ICE surface temperature (IST) reflects the freeze–thaw state and the timing and duration of the melting of an ice sheet.

Tingting Liu, Yachao Li, and Zemin Wang are with the Chinese Antarctic Center of Surveying and Mapping, Wuhan University, Wuhan 430070, China, and also with the Key Laboratory of Polar Environment Monitoring and Public Governance, Ministry of Education, Wuhan University, Wuhan, China (e-mail: ttliu23@whu.edu.cn; yachaoli@whu.edu.cn; zmwang@whu.edu.cn).

Qiangqiang Yuan is with the School of Geodesy and Geomatics, Wuhan University, Wuhan 430070, China, and also with the Key Laboratory of Polar Environment Monitoring and Public Governance, Ministry of Education, Wuhan University, Wuhan 430070, China (e-mail: qqyuan@sgg.whu.edu.cn).

Therefore, the study of Antarctic IST is important for the understanding of the surface energy budget and the surface mass balance of the Antarctic ice sheet [1], [2], [3]. According to Giovinetto et al. [4], the surface mass balance in Antarctica is significantly correlated with IST with a correlation coefficient of 0.7–0.8. IST also plays a crucial role in the exchange of energy between the ice sheet and the atmosphere [5]. Antarctic IST varies seasonally and regionally. However, these spatiotemporal variations could change because of rapid climatic changes and frequent climate extremes [6]. Therefore, spatially and temporally continuous IST products that have high accuracy are needed to improve our understanding of the surface changes of the Antarctic ice sheet.

Antarctic IST is commonly monitored using automatic weather stations (AWSs) and remote sensing imagery [7]. AWSs provide long and continuous data records with high temporal resolution (observations every few minutes). However, the majority of AWSs measure the near-surface air temperature (NST) rather than IST; the difference between IST and NST can be substantial, especially under conditions of low wind speed and solar radiation [8]. In addition, each AWS only collects data at a specific location. AWSs are sparsely distributed over Antarctica; consequently, the IST of the entire Antarctic ice sheet cannot be monitored using only AWSs. It is possible to use thermal infrared and passive microwave (PM) remote sensing data to monitor surface temperature over a vast area [9]. Compared with PM remote sensing, thermal infrared remote sensing data have a high spatial resolution (30–1000 m) and relatively high accuracy under clear skies (1.2–2.3 K) [7], [10], [11]. However, thermal infrared remote sensing cannot penetrate clouds, resulting in the spatial discontinuity of IST products from thermal infrared data. [12]. PM remote sensing data have low spatial resolution (25 km) but can penetrate through clouds to measure surface information of almost the entire planet in a single day [13]. Consequently, PM data have been used to generate large-scale datasets. Some researchers have made much effort to generate spatially or temporally continuous LST data by using PM data. Song and Zhang [14] combined AMSR-2 and the FY-3B Microwave Radiation Imager data to enhance coverage of daily land surface temperature (LST) estimates in low latitudes. Wu et al. [15] mapped gapless all-weather LST in China using thermal infrared and PM data. Dowlin et al. [16] generated the spatially continuous LST through introducing PM data. In view of these

characteristics, PM data can be very beneficial for the monitoring of Antarctic IST.

Commonly, PM data have been used to retrieve sea surface temperature and LST, with root-mean-square errors (RMSEs) of 0–1 and 2–5 K, respectively [17], [18], [19], [20]. Generally, physical or empirical models have been used to retrieve LST from PM data. The physical models are based on the radiation transmission equation of PM; their accuracy highly relies on the land surface emissivity. However, because snow emissivity is sensitive to snow liquid water content, surface roughness, and snow grain size in the microwave range, snow emissivity varies both spatially and temporally [21], [22], [23]. It is difficult to obtain reliable snow emissivity and thus retrievals can have considerable error, limiting the application of physical models to IST retrieval [24]. Empirical models, which are based on the relationship between LST and brightness temperatures and avoid complex physical parameters, have been used widely to retrieve LST from PM data [17], [19], [25]. Therefore, in this study, we retrieved Antarctic IST using empirical models.

In 2017, Microsoft developed the Light Gradient Boosting Machine (LightGBM) algorithm on the basis of the eXtreme Gradient Boosting (XGBoost). In this study, we used LightGBM to develop a framework to retrieve IST from PM data with the goal of generating spatially continuous Antarctic IST maps with high temporal resolution. The sample set also plays a fundamental role in model construction. Its size, representativeness, and accuracy can affect the robustness and accuracy of the trained model [26]. The moderate-resolution imaging spectroradiometer (MODIS) IST product has high accuracy and large spatial and temporal coverages [27], [28]; therefore, we used it as ground truth for model training and testing. The IST retrieval model uses PM brightness temperature from the Advanced Microwave Scanning Radiometer 2 (AMSR2) as the primary input [17], [19].

We used multisource data to test our IST retrieval framework. The accuracy of different typical linear and nonlinear models was evaluated by using the testing sample set. IST retrievals from Landsat-8 thermal infrared band data have a high accuracy and spatial resolution [7]. We used IST retrieved from Landsat-8 satellite data and near-surface thermal infrared radiation data to assess the accuracy of the LightGBM model. We used the LightGBM model to retrieve IST from AMSR2 data and examined the variations of the Antarctic IST between 2013 and 2020.

The rest of this article is organized as follows. Section I introduces the background and significance of the study. The datasets that were used in the study are introduced in Section II. Section III presents the LightGBM algorithm and the details of sample set construction and model training. In Section IV, we compare the performance of different models and evaluate the accuracy of the LightGBM model using multisource data. In Section V, we discuss the error distributions of LightGBM model on month, IST, and scan time gap (STG). We examine the role of the near-surface temperature inversion in the Antarctic and discuss the variations of Antarctic IST between 2013 and 2020. Finally, Section VI concludes this article.

II. DATA

We used multisource data to establish a sample set, train models, and evaluate model accuracy. Details of the data that were used in this study are presented in this section.

A. AMSR2 Data

Global Change Observation Mission 1st - Water (GCOM-W1) satellite was launched by the Japan Aerospace Exploration Agency in 2012. It is in a sun-synchronous orbit with a satellite inclination angle of 98°. It carries AMSR2, which is the successor of the Advanced Microwave Scanning Radiometer-Earth Observing System. The AMSR2 is a multifrequency PM radiometer. It is enabled to capture approximately 29 swath granules from the polar regions because of the short revisit period of GCOM-W1 (approximately $15\times$ per day in Antarctica) [29]. Each granule is defined as a half orbit between the North Pole and the South Pole.

The AMSR2 has a conical scan mechanism. It measures microwave radiation from the Earth at 6.9, 7.3, 10.7, 18.7, 23.8, 36.5, and 89 GHz. Each frequency has two PM polarization modes (H and V). AMSR2 products are categorized according to the processing level. To capture instantaneous ice surface information, we used Level 1R swath brightness temperature data from all channels. We also used associated sensor observation information, such as latitude, longitude, Earth azimuth, Earth incidence, sun azimuth, sun elevation, area mean height, and scan time. We obtained the AMSR2 data from the official website of the Globe Portal System (https://gportal.jaxa.jp).

We used AMSR2 data covering Antarctica from July 2012 to June 2020 to generate a sample set. For the 89 GHz channel, the original sampling interval is 5 km. We aggregated the pixels into 10-km intervals to match the sampling intervals of the other channels (6.9, 7.3, 10.7, 18.7, 23.8, and 36.5 GHz).

B. MODIS IST Product

The MODIS instrument is carried by both the Terra and the Aqua satellites. It views the entire surface of the Earth every 1–2 days [30]. Its detectors collect data in 36 bands in the wavelength range of 0.4–14.4 μm; these bands include 16 thermal infrared bands and the atmospheric water vapor retrieval bands. The main inputs (i.e., atmospheric water vapor content) for IST retrieval can be estimated using these data. Retrievals of IST from MODIS data have high accuracy and are not influenced by intersatellite differences in algorithm inputs.

The National Aeronautics and Space Administration (NASA) provides swath IST products (MOD/MYD11_L2) with 1-km resolution, which are generated using the split-windows method [31]. We obtained the data from https://ladsweb.modaps.eosdis.nasa.gov/search/. Because remote sensing data from the thermal infrared band are easily impacted by weather (e.g., clouds and aerosols), over 60% of the areas covered by the MOD/MYD11_L2 products are invalid [32]. Fig. 1 shows two randomly mosaicked Antarctic IST maps based on MOD/MYD11_L2 products for July 3, 2016 (in austral winter)
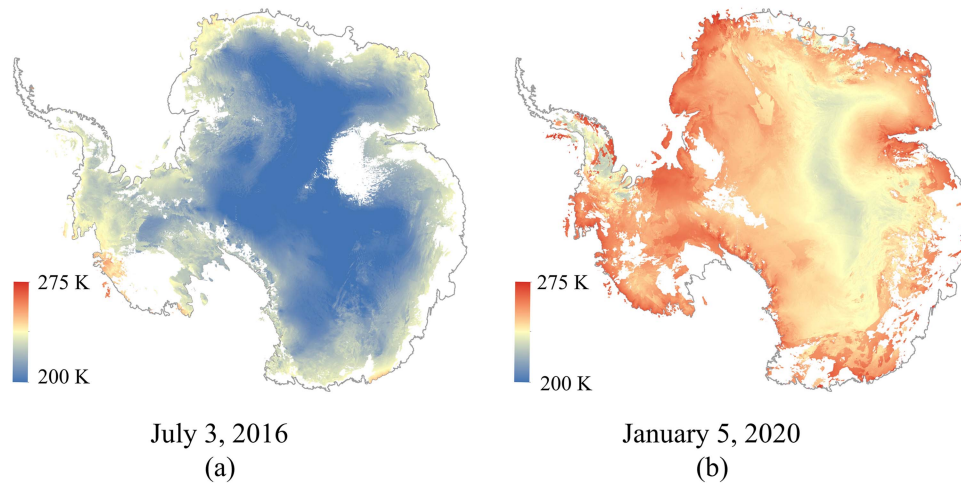
July 3, 2016
(a)

January 5, 2020
(b)

Fig. 1.   Mosaicked Antarctic IST maps from MOD/MYD11_L2 for July 3, 2016 and January 5, 2020.

and Jan. 5, 2020 (in austral summer). Although each map was mosaicked by using all the swath IST products of the day (30–40 scenes of MOD/MYD11_L2), Antarctic coverage remains incomplete in both winter and summer.

We used cloud-free MOD/MYD11_L2 products as ground truth to train and evaluate the models because they have high accuracy and large spatial and temporal coverages. According to Wan [28], there is close agreement between in situ measurements of LST and LST from MOD/MYD11_L2 under clear-sky conditions, with an RMSE of less than 2 K at most sites and an RMSE of less than 1 K at two sites in Antarctica. To allow data from different sources to be colocated in space, we projected the MOD/MYD11_L2 data to the Antarctic stereographic projection and resized the pixel size to 10 km.

C. Terrain Data

Because IST varies closely with terrain [2], we used terrain data as sample inputs for model construction. The Bedrock Mapping Project (Bedmap) has produced a suite of gridded products that describe surface elevation and other characteristics of the Antarctic [33]. We used Bedmap2, which includes more measurements from a variety of sources and has a higher quality than Bedmap1. We downloaded Bedmap2 data from https://secure.Antarctica.ac.uk/data/bedmap2/.

To obtain a fuller description of the terrain, we generated slope and aspect data from the Bedmap2 digital elevation model data, which have a resolution of 1 km. All the other inputs used in this study are on 10-km grids. Therefore, we resized the Bedmap2 elevation, slope, and aspect data onto a 10-km grid to ensure consistency between all inputs.

D. Landsat-8 Data

Landsat-8 was launched by NASA in 2013. It carries two payloads, which are the Operational Land Imager and the Thermal Infrared Sensor. Both instruments have a swath width of 190 km. The two instruments together cover a total of 11 bands. The spatial resolutions of the Operational Land Imager and the

Thermal Infrared Sensor data are 15/30 and 100 m, respectively [34]. The pixel sizes of the Landsat-8 data were downsampled to match AMSR2 data (10 km) by calculating the average pixel value in a 10-km window. We obtained Landsat-8 data from https://earthexplorer.usgs.gov/.

We used an improved single-channel algorithm proposed by Li et al. [7] to retrieve IST from Landsat-8 data for model validation. The single-channel algorithm was built specifically for Polar Regions; accuracy is relatively high and RMSE is ~1.2 K for clear conditions. In austral winter, Antarctic data are not available from Landsat-8. Therefore, we used three Landsat-8 images taken on different days in austral summer (blue rectangles in Fig. 2).

E. AWS Data

Two types of AWS data are used in this study. We used air temperature data from the Antarctic Meteorological Research Center (AMRC) of the United States Antarctic Program to compare with retrieved IST and analyze near-surface temperature inversion. The AMRC provides real-time and archived meteorological data and observations and supports a network of AWSs in Antarctica. The AWSs measure 2-m air temperature with an accuracy of approximately 0.1 K [35]. There were approximately 60 AMRC AWSs operational between 2012 and 2022; exact numbers vary because of AWS installation and retirement. AMRC air temperature data in this study are used to compare with IST rather than to validate the accuracy of the IST retrieval models. Fig. 2 shows the locations of the AMRC AWSs (yellow circles). We obtained AMRC data from https://amrc.ssec.wisc.edu/.

We also used surface thermal infrared radiation data from the Baseline Surface Radiation Network (BSRN) to validate the accuracy of the proposed model. The BSRN is a project of the Data and Assessments Panel from the Global Energy and Water Cycle Experiment, which focuses on detecting Earth's surface radiation and its effects on climate change. We obtained the BSRN data from https://www.pangaea.de/. The BSRN stations
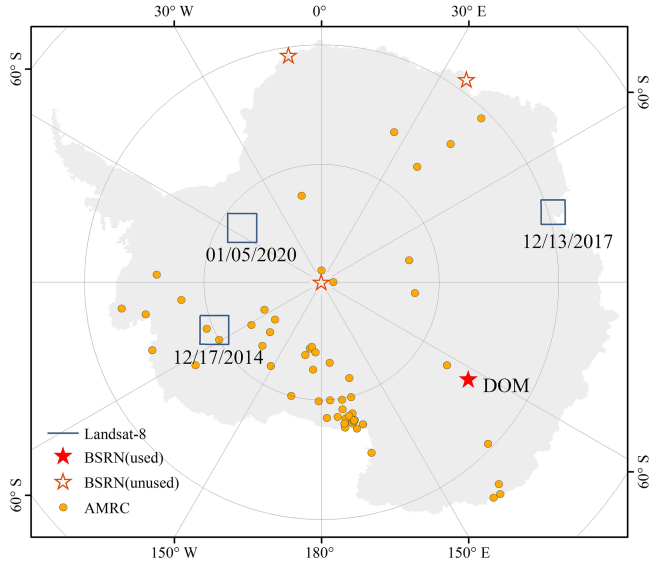
Fig. 2. Locations of the Landsat-8 images, and the BSRN and AMRC stations in Antarctica. The blue rectangles indicate the locations of the three Landsat-8 images. The yellow circles indicate the AMRC stations. The stars indicate the BSRN stations. The hollow stars indicate stations located outside our IST map (data from these stations were not used); data from the station indicated by the solid red star (DOM) were used for validation.

measure broadband thermal infrared radiation from the surface and the air rather than IST or air temperature. Following the theory of radiative transfer, IST can be retrieved from thermal infrared radiation as follows [36]:

$$\text{IST} = \sqrt[4]{\left(L_{\text{up}} - L_{\text{down}}\left(1 - \varepsilon\right)\right)/\left(\sigma\varepsilon\right)} \tag{1}$$

where L_{up} and L_{down} are upwelling and downwelling radiance, respectively; σ is the Stefan–Boltzmann constant ($\sigma = 5.67 \times 10^{-8}$ W·m^{-2}·K^{-4}); ε is the snow emissivity of broadband thermal infrared radiation, which was set to 0.985 according to Fréville et al. [37]. We used BSRN IST to evaluate model accuracy. There are four BSRN stations in the Antarctic (stars in Fig. 2). Only the DOM station (the red star in Fig. 2) lies inside our IST map. One of them is located at the South Pole, the GCOM-W1 satellite has an orbital inclination of 98° and AMSR2 data cannot cover that place. Therefore, we used the data from the DOM station for model accuracy validation. Because measurements of radiation can be affected by atmospheric emissions between the sensor and the snow surface [38], IST retrieved from BSRN data might deviate from true IST.

F. Reanalysis Data

ERA5 is the fifth-generation atmospheric reanalysis of the global climate that is provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). It has been proven to be reliable and has been used widely in climate and environmental research [39], [40], [41]. We used ERA5-Land, which is an enhanced dataset; it is forced by the atmospheric analysis of ERA5 and has a higher spatial resolution (0.1° × 0.1°) [42]. Although ERA5 assimilates AWS measurements and remote sensing data, we took the skin temperature from
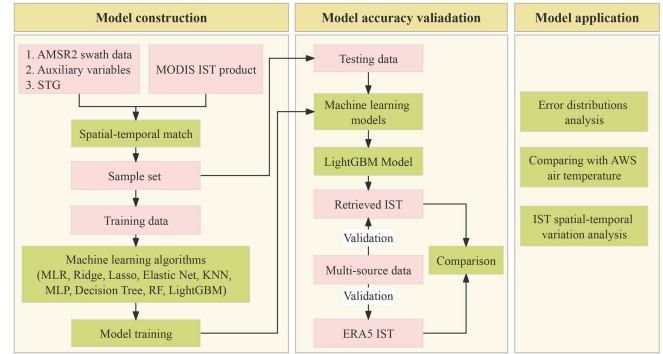
ERA5-Land as IST and compared it with the IST from AMSR2. In addition, we used the skin temperature from ERA5-Land to construct the standard Antarctic IST distribution to address the data imbalance in the sample set. The ERA5-land was reprojected to the Antarctic stereographic projection and resized to 10 km.

Fig. 3. Flowchart of the methodology in this study.

III. Methods

This study aims to generate spatial and temporal continuous IST data in the Antarctic. The PM data are used as the main input because it covers the Antarctic and is not affected by the cloud. The STG is used as a model input to make the temporal continuity of the retrieved IST data. The flowchart of this study is shown in Fig. 3. The flowchart includes three parts: model construction, model accuracy validation, and model application. First, we constructed the sample set before training the machine learning model. The AMSR2 swath data, auxiliary variables, and STG are model inputs, and the MODIS IST product is model output. Second, the testing data were used to validate the nine machine learning models. The LightGBM model was chosen from the nine models because of its best performance. The accuracy of the LightGBM model was also validated by using multisource data. ERA5 IST was used to compare with IST from the LightGBM model. In the last part, we have analyzed the error distribution of the LightGBM model. We have compared IST data from the LightGBM model with AWS air temperature and analyzed Antarctic IST spatial-temporal variation.

A. Machine Learning Algorithms

There are linear and nonlinear empirical models. Nonlinear models are superior in reproducing the relationship between IST and brightness temperature owing to their more complex construction. According to the PM radiation equation and assuming constant surface emissivity, the relationship between brightness temperature and surface temperature is approximately linear [17]. However, linear models cannot fully express the relationship between IST and brightness temperature because of the considerable spatiotemporal variability of snow emissivity. Among the nonlinear models, ensemble learning exhibits superior performance because this learning model improves the performance of a single model by training multiple models and
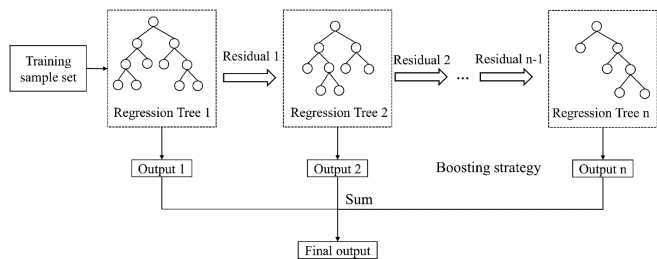
Fig. 4.   Scheme of LightGBM algorithm.

Fig. 5.   Number and date of AMSR2 and MOD/MYD11_L2 images, which are used for sample set construction. The insert figure presents the details.

combining their results [43]. RF and XGBoost [44] algorithms are two types of typical ensemble learning methods that improve the model accuracy by combining several base estimators. According to Sayed et al. [45], the performance of LightGBM is superior to that of RF or XGBoost in predicting the need for mechanical ventilation among patients with acute respiratory distress syndrome. Moreover, LightGBM requires less computer memory and has lower time complexity and higher accuracy than XGBoost [46]. In this study, we used LightGBM to develop a framework to retrieve IST from PM data to generate spatially continuous Antarctic IST maps with high temporal resolution.

Fig. 4 presents the scheme of LightGBM algorithm. LightGBM algorithm produces the final prediction model by using ensemble weak prediction models (Decision Trees). LightGBM algorithm adopts a boosting strategy. The residual from the previous tree is used as the input of the following tree. The final output is the sum of the outputs from all the trees. In addition, the LightGBM algorithm adopts two novel techniques—gradient-based one-side sampling (GOSS) and exclusive feature bundling (EFB)—to improve efficiency for cases of large data sizes and high feature dimensions [46]. The data instances with larger gradients play a more important role in the computation of information gain. Therefore, to maintain the accuracy of information gain and reduce computational cost, the data instances downsampling strategy of GOSS keeps the data instances with larger gradients. The EFB is used to bundle mutually exclusive features. In this sparse feature space, features rarely take nonzero values simultaneously. As a result, the number of dimensions is reduced and efficiency is improved while a high level of accuracy is maintained.

B. Sample Set Construction

The sample instance generally consists of inputs and an output, which is IST in this study. To avoid errors resulting from intraday variations of IST, we used swath rather than daily data for model training. We also added auxiliary inputs to the IST retrieval model to enhance model accuracy and robustness [47], [48]. These auxiliary inputs are closely related to the output and include geographical location (longitude and latitude), terrain (elevation, slope, and aspect), and sensor observation parameters (earth azimuth, earth incidence, sun azimuth, sun elevation, area mean height, and scan time). We defined the difference between the acquisition time of AMSR2 swath data and that of MODIS data as the STG and used it as a model input. Ideally, the
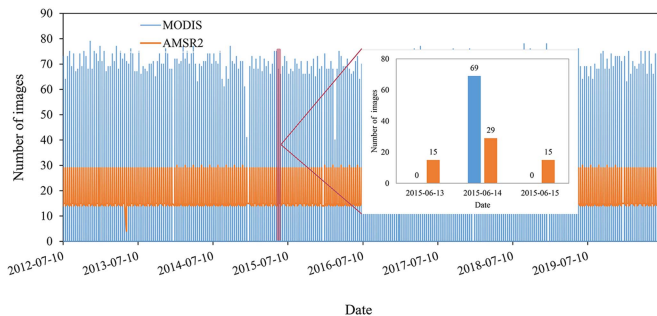
two acquisition times are the same and STG is zero. However, MODIS IST products are available only for clear-sky conditions. Near-zero STG would limit the representativeness of the sample set, which will contain no data for cloudy conditions. To provide complete coverage of Antarctica, several swath images are required and these images have different acquisition times. Near-zero STG would generate a spatially continuous product of Antarctic IST but without temporal continuity. Therefore, we extended the STG to ±12 h to enhance model robustness and to obtain a spatially and temporally continuous IST product. For time t, IST can be generated by adjusting the difference between t and the acquisition time of AMSR2 swath data.

An instance consists of a group of brightness temperatures from AMSR2 swath data, auxiliary variables from Bedmap2 and AMSR2, and the IST from MOD/MYD11_L2 products. All these data have been projected to the Antarctic stereographic projection and colocated in space. We checked the consistency of the acquisition time of the data from the different sources. The STG was set to ±12 h. There are approximately 70 scenes of MOD/MYD11_L2 data every day and approximately 59 corresponding AMSR2 swath granules every 48 h (see Fig. 5). We selected a day's MODIS data every ten days from 2012 to 2020. AMSR2 data are selected according to the MODIS data and STG. In total, we generated an initial sample set that contained data from 289 days; these included approximately 20 000 scenes from MOD/MYD11_L2 and 17 000 swath granules of AMSR2.

The distribution of the sample set is generally uneven. This data imbalance might reduce the accuracy and reliability of the trained model [49]. Therefore, we compared the IST distribution of the generated sample set in this study with a standard IST distribution. The generated sample set without being adjusted was called the original sample set and the generated sample set after adjusting was called the new sample set. We used hourly ERA5-Land Antarctic IST from 2018 as the standard IST distribution because it had complete coverage of Antarctica. The standard IST distribution [see Fig. 6(b)] has a single peak (at 245 K) and differs from that of the initial sample set, which has one peak at 210 and another peak at 240 K [see Fig. 6(a)]. Following this comparison, we abandoned a portion of the initial sample to generate a new sample set [see Fig. 6(c)]. The biggest density difference between the standard IST distribution and IST distribution of the original sample set occurs at 247 K.
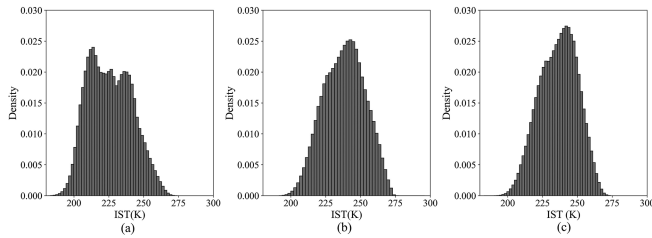
Fig. 6. IST distributions for different sample sets. (a) IST distributions for the original sample set. (b) Standard IST distribution. (c) IST distribution of the new sample set.

TABLE I
ACCURACY OF THE DIFFERENT MODELS

	Models	RMSE(K)	MAE(K)	Bias(K)
Linear	MLR	5.81	4.47	-0.05
	Ridge	5.82	4.49	0.02
	Lasso	6.09	4.74	-0.04
	Elastic Net	6.09	4.75	0.03
Nonlinear	KNN	3.88	2.75	-0.27
	MLP	3.83	2.82	0.15
	Decision Tree	4.62	3.45	-0.01
	RF	3.71	2.76	-0.02
	LightGBM	**2.96**	**2.13**	**-0.01**

We used the number of sample instances at 247 K as the new sample set number at 247 K. Then, we can calculate the sample instances number of the new sample set and abandon some sample instances based on the standard IST distribution to generate a new sample set. The new sample set [see Fig. 6(c)], with an IST distribution close to that of the new sample set [see Fig. 6(b)], was used for model training.

C. Model Training

Approximately 80% of 1.3 million randomly selected samples were used for model construction and the remaining 20% were used for model testing. In addition to the LightGBM algorithm, some typical algorithms (including multiple linear regression—MLR, ridge regression, lasso regression, elastic net regression, decision tree regression, RF, k-nearest neighbor—KNN, and Multilayer perceptron—MLP) were also used to produce results for comparison. These algorithms differ in complexity but were applied to the same training and testing sets.

The 80% sample set was split into training and validation sets using a fivefold cross validation. The training and validation sets were used for model parameter determination and hyper-parameter optimization, respectively. We used the k-fold cross validation to optimize model hyperparameters. It is a typical method that is used to effectively avoid overfitting. Generally, k exceeds 2 [50] and a high k value (e.g., 10) is used for small sample sets. We took into consideration the size of the sample set and computation time and set k to 5. In a fivefold cross validation, 80% of the sample instances for model construction are used for model training and 20% are used for model validation. For the model with multiple hyperparameters, we adopted the greedy strategy to determine the hyperparameters [51]. We used the testing set to evaluate model accuracy. Because hyperparameters are absent in the MLR model, validation data were not used for MLR model training.

IV. RESULTS

We used the testing sample set to evaluate the performance of the nine machine learning models and compare their performances. In addition, we used the LightGBM model to retrieve hourly IST for the 15th day of every month between 2013 and 2020 and used multisource data to evaluate model accuracy. This section focuses on the model accuracy evaluation and the spatial and temporal IST results will be shown in Section V.

A. Validation Using Testing Data

We evaluated the accuracy of the nine models by using the same testing sample set (see Table I). The biases of all the models are close to zero, which indicates the absence of system bias. Among the linear models, the performance of MLR exceeds the performance of the Ridge, Lasso, or Elastic Net regression model. In the Ridge, Lasso, and Elastic Net regression models, overfitting is avoided by the addition of regularization terms. However, the addition of regularization terms reduces model accuracy in the case of underfitting [52]. This indicates that the expressiveness of the linear models is insufficient to fit the training set. The performance of the nonlinear models exceeds that of the linear models. Among the nonlinear models, the Decision Tree model has the lowest accuracy because it has a relatively simple structure. The KNN, MLP, and RF models have similar levels of accuracy; RMSE is approximately 3.8 K. The performance of LightGBM is superior with an RMSE of 2.96 K and mean absolute error (MAE) of 2.13 K. The KNN model is easy to use but its sensitivity to noise in the training set might limit its performance [53]. MLP is a typical feed forward neural network trained by backpropagation methods; it is prone to falling into the local optimum rather than the global optimum [54]. Both RF and LightGBM are ensemble learning algorithms that use Decision Tree as the base regressor. The ensemble strategy of RF is bagging and that of LightGBM is boosting. In addition, LightGBM adopts two novel techniques (GOSS and EFB), which address the problems of efficiency and scalability for large data sizes and high feature dimensions [46]. Our results show that the performance of LightGBM is superior to that of RF; this is in agreement with previous studies [28], [55]. Because the LightGBM model has the best performance, we only used the LightGBM model for the remaining analyses.

Training set size is an essential quantitative factor that affects model accuracy [56]. When the sample set size is small, the model accuracy increases with the increase of the sample set size. When there are sufficient sample instances, the sample set size hardly changes the model accuracy and the sample set quality dominates the model accuracy. To figure out the appropriate sample set size for this study, we trained the LightGBM model using five different training sets and calculated the RMSEs. The numbers of sample instances for the five sample sets are 2, 4, 6, 8, and 10 million, respectively. The sample instances in each set were chosen randomly. The corresponding model accuracies of the five sample sets are 2.17, 1.97, 1.78, 1.67,

20141217   20171213   20200105

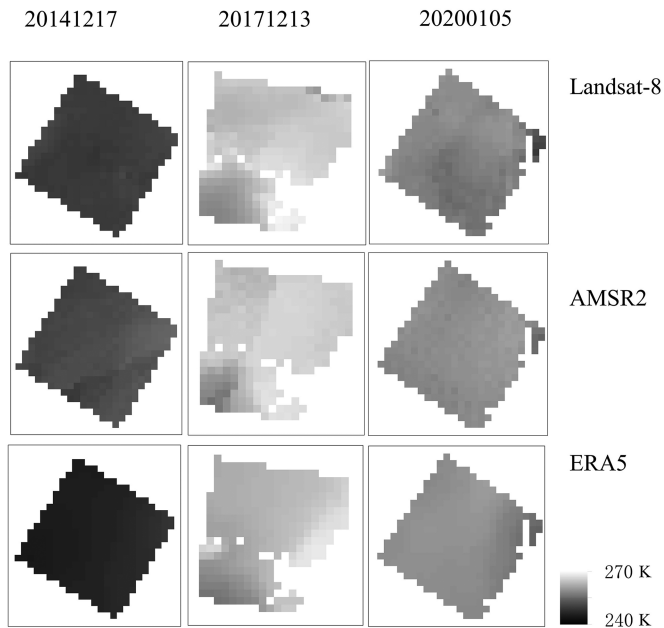Landsat-8

AMSR2

ERA5

270 K

240 K

Fig. 7. Maps of IST from AMSR2, ERA5-land, and Landsat-8 data. The pixel size of all maps is 10 km. The different columns denote the IST maps on different dates. The different rows denote the IST maps from different sources.

TABLE II
COMPARISON OF ISTs FROM AMSR2, ERA5-LAND, AND LANDSAT-8

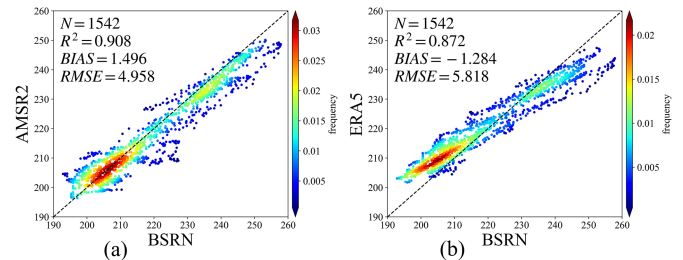| Date	AMSR2 - Landsat-8		ERA5-Land - Landsat-8	
	RMSE(K)	Bias(K)	RMSE(K)	Bias(K)
20141217	2.32	2.07	3.80	-3.73
20171213	1.39	0.43	2.18	-1.36
20200105	1.99	1.44	1.92	1.30

Fig. 8. ISTs from the DOM station of BSRN and (a) AMSR2 and (b) ERA5-Land; N is the number of data samples; R^2 is Pearson's correlation coefficient; BIAS is the mean bias; RMSE is the root-mean-square error.

and 1.65 K, respectively. Therefore, the number of sample set instances exceeding 8 million is appropriate and adopted for the following study.

B. Cross Validation Using Multisource Data

We evaluated the accuracy of the LightGBM model using IST from Landsat-8 thermal infrared data and IST from BSRN broadband thermal infrared radiation data. Skin temperature from ERA5-Land was also used for comparison.

Fig. 7 shows maps of IST from Landsat-8, AMSR2, and ERA5-Land for three different days. The pixel sizes of the Landsat-8 and ERA5-Land maps were resized to match those of the AMSR2 data (10 km). ERA5-Land is forced by the atmospheric analysis of ERA5, which adopts interpolation methods during data assimilation [57]. The AMSR2 and Landsat-8 maps were generated pixel by pixel and, therefore, have more random noise. As a result, the appearance of the ERA5-Land maps is smoother than that of the other maps. Visual inspection of the three maps reveals clear agreement between the Landsat-8, AMSR2, and ERA5-Land maps for December 13, 2017 and January 5, 2020. However, for September 17, 2014, ERA5-Land IST is considerably lower than AMSR2 IST and Landsat-8 IST. We calculated the accuracy of both AMSR2 IST and ERA5-Land IST in comparison with Landsat-8 IST, which we took as ground truth (see Table II). The accuracy of AMSR2 IST is higher than that of ERA5-Land; RMSEs are 1.39–2.32 K.

We evaluated the accuracy of AMSR2 IST and ERA5-Land IST using data from a BSRN station that is in the inland area of East Antarctica (DOM in Fig. 2). To match the temporal resolution of ERA5-Land (1 h), we used the LightGBM model to retrieve hourly IST for the 15th day of every month between

2013 and 2020 from the AMSR2 swath data. This sampling provides 2304 instances. However, because of missing BSRN data, we only had 1542 data pairs.

Fig. 8 shows ISTs from BSRN, AMSR2, and ERA5-Land and the number of data samples (N), Pearson's correlation coefficient (R2), mean bias, and RMSE. The ISTs range from 200 to 255 K; ASMR2 IST has a larger R2 and smaller RMSE than ERA5-Land IST. This indicates that the correlation between ASMR2 IST and BSRN IST is higher than that between ERA5-Land IST and BSRN IST. The mean bias shows that AMSR2 IST is lower than BSRN IST and that ERA5-Land IST is higher than BSRN IST. This indicates that the mean AMSR2 IST is lower than the mean ERA5-Land IST. We infer that the performance of AMSR2 IST is higher in cold inland areas.

Fig. 9 shows the accuracy of retrievals of seasonal IST. The accuracy of AMSR2 IST is the highest in austral autumn (RMSE of 3.732 K). The accuracy of ERA5-Land IST is the highest in spring (RMSE of 4.655 K). In spring (Sep., Oct., and Nov.), summer (Dec., Jan., and Feb), and winter (Jun., Jul., and Aug.), BSRN IST is higher than AMSR2 IST. In summer, autumn (Mar., Apr., and May), and winter, BSRN IST is lower than ERA5-Land IST. The accuracy of AMSR2 IST is higher than that of ERA5-Land IST in summer, autumn, and winter. In spring, the accuracy of ERA5-Land IST is higher than that of AMSR2 IST. There is little consistency in the seasonal variations of the IST accuracy of ERA5 and AMSR2.

V. DISCUSSION

A. Error Distributions of the LightGBM Model

We analyzed the error distributions of LightGBM model based on STG, month, and IST. First, the STG is a novel parameter for the LightGBM model, and our concern focuses on the magnitude of the error that might be caused by the STG. Second, the ice surface features of the Antarctic vary substantially according to
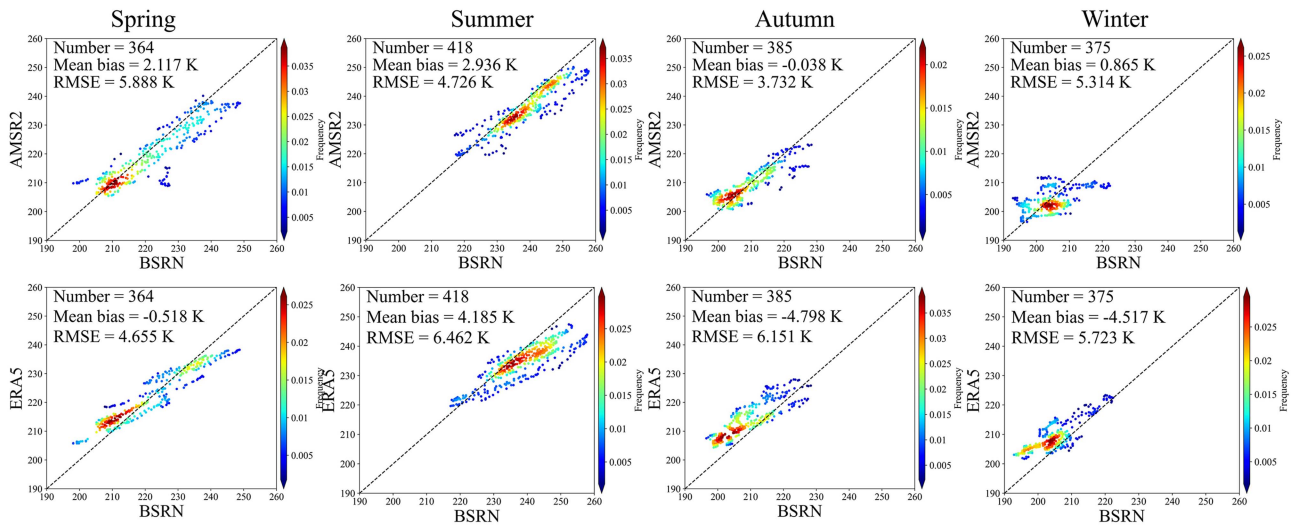
Fig. 9.   ISTs from the DOM station of BSRN, AMSR2, and ERA5-Land for different seasons. Spring is shown in the leftmost column; summer is shown in the second column to the right; autumn is shown in the third column to the right; winter is shown in the rightmost column. The top row shows BSRN IST and AMSR2 IST. The bottom row shows BSRN IST and ERA5-Land IST.

month, which might induce changes in PM radiance intensity. This could further impact the accuracy of the model; therefore, the month is also considered in the discussion. In addition, IST is also adopted because it is the output of the model. We calculated the RMSE and mean bias using the testing sample set. Fig. 10 shows the error distributions of the LightGBM model based on month, IST, and STG.

The performance of the LightGBM model is lower in austral winter and higher in austral summer [see Fig. 10(a)]. This is because PM penetration depends on snow characteristics, which vary with season. In winter, low temperature and relatively dry snow result in large PM penetration depth [22]. In contrast, the snow becomes wet and even melts during the summer; as a result, PM radiation signals from below the surface are weak, and satellite data contain most information from the surface. In addition, the LightGBM model presents slightly cold biases in austral summer and autumn, and warm biases in austral winter and spring.

Fig. 10(b) shows the error distribution of the LightGBM model based on IST. The performance of the model is higher at higher ISTs and lower at lower ISTs; this is consistent with the variation of model accuracy with month [see Fig. 10(a)] because there is a strong correlation between IST and month. Moreover, the mean bias and RMSE of the LightGBM model are especially high for ISTs of 0–200 and 200–210 K. The warm bias might be caused by a small number of sample instances with cloudy conditions in the sample set. For cloudy conditions, IST in the MOD/MYD11_L2 products is extremely low. The low accuracy of LightGBM IST for ISTs of 0–200 and 200–210 K could also be caused by high PM penetration at low IST. However, the impact of these sample instances at 0–210 K on the overall LightGBM model accuracy is not obvious because the number of these sample instances is small [see Fig. 6(c)]. Variations of RMSE with IST [see Fig. 10(b)] are larger than the
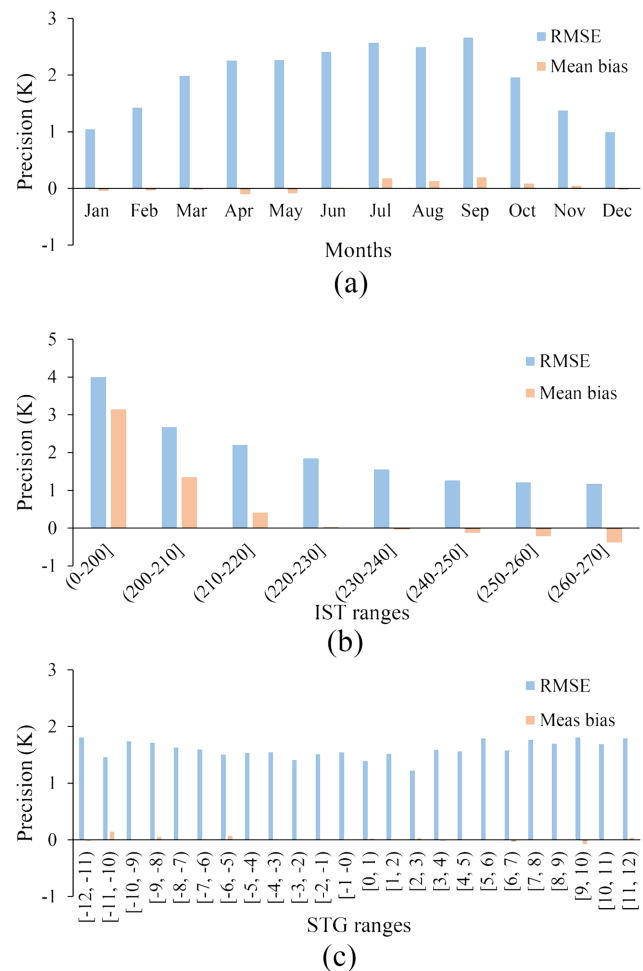

Fig. 10.   Error distributions of the LightGBM model based on (a) month, (b) IST, and (c) STG.

variations of RMSE with month [see Fig. 10(a)]; this indicates that model accuracy is more sensitive to IST than to month. We infer that low-temperature areas in inland areas in summer and high-temperature areas in coastal areas in winter reduce the intermonth variation in model accuracy and decrease model sensitivity to month. We suggest that LightGBM model accuracy may be improved by using sample sets with different IST ranges for model training. However, this might be difficult to achieve because of the need to match input data with the correct IST retrieval model.

Fig. 10(c) shows the error distribution of the LightGBM model based on STG. On the *x*-axis, negative values indicate that the scan time of AMSR2 is before that of MODIS, and positive values indicate the reverse. Model accuracy is slightly higher for STG values in the middle range and slightly lower for low and high STGs. This indicates that the LightGBM model is insensitive to STG and that the introduction of STG hardly affects model accuracy. Mean bias varies little with STG and is near-zero for all STGs.

Overall, model accuracy is the most sensitive to IST and the least sensitive to STG. The performance of the LightGBM model is superior for specific IST ranges. Therefore, the accuracy of retrieved IST may be improved if the models are trained by sample sets with different IST ranges and the input data are assigned to the correct retrieval models.

B. Bias Between the IST and the NST

AWSs have been collecting NST in Antarctica for many years. We used this long-term data to examine the bias between the IST and the NST. According to Adolph et al. [8], it is not recommended to validate IST retrievals using AWS air temperatures because of the uncertainty induced by near-surface temperature inversion, which is controlled mainly by wind speed, solar radiation, and other factors [58]. In addition, the near-surface temperature inversion might not fully account for the bias between the NST and IST because of errors in the retrieved IST, which are difficult to eliminate. We analyzed the bias between AMSR2 IST and AWS NST. Hourly AMSR IST on the 15th of each month from 2013 to 2020 is used.

Fig. 11 shows the spatial distribution of the biases between NST and AMSR2 IST through the seasons. In austral autumn and winter, there is a distinct spatial heterogeneity with larger biases in West Antarctica and smaller biases in East Antarctica. Spatial heterogeneity is smaller in austral spring and summer and larger in austral autumn and winter. As mentioned above, the bias between NST and ASMR2 IST is dominated by the near-surface inversion and IST error. The LightGBM model has a warm bias in austral winter and a near-zero bias in austral summer [see Fig. 10(a)]. The warm bias in ASMR2 IST reduces the bias between the ASMR2 IST and NST because the IST is generally lower than NST; a cold bias in ASMR2 IST causes the bias between the ASMR2 IST and NST to increase. The biases between IST and NST in austral summer [see Fig. 11(b)] are clearly lower than those in austral winter [see Fig. 11(d)]. Therefore, we infer that the near-surface temperature inversion in austral summer is weaker than that in austral winter. Our results are in agreement
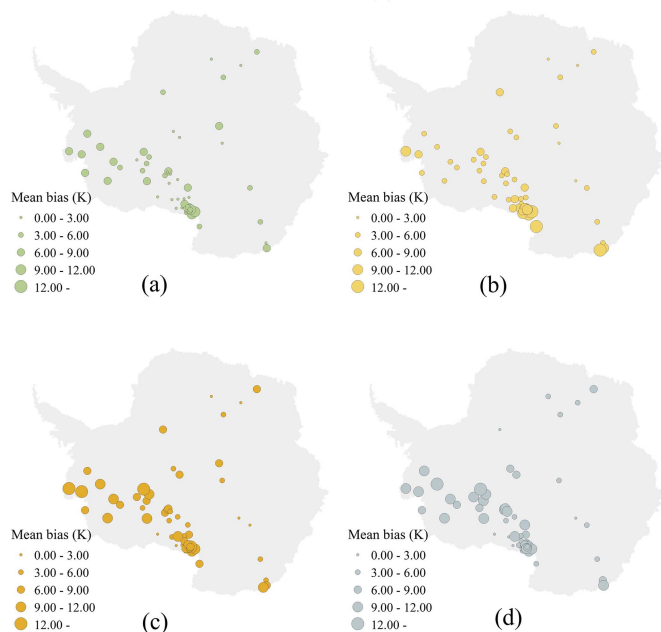

Fig. 11. Mean bias between AWS NST and AMSR2 IST for different seasons. AWS NST and hourly AMSR2 IST data on the 15th of each month from 2013 to 2020 are used.
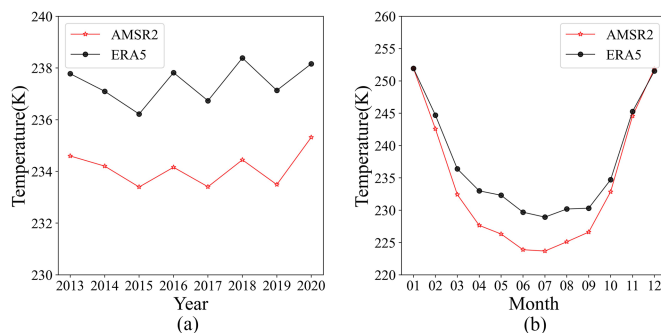

Fig. 12. Variations of (a) annual and (b) monthly AMSR2 IST and ERA5-Land IST. Hourly AMSR IST and ERA5 IST on the 15th of each month from 2013 to 2020 are used.

with those from Summit station in Greenland. According to Miller et al. [59], near-surface temperature inversions at Summit station are prevalent in the winter with decreasing values in the summer months. Similarly, the LightGBM model has a warm bias in austral spring and a cold bias in austral autumn [see Fig. 10(a)]. Therefore, the bias between the IST and NST in austral autumn [see Fig. 11(c)] and spring [see Fig. 11(a)] is larger and smaller than the near-surface temperature inversion, respectively.

C. Antarctic IST From 2013 to 2020

We analyzed the variation of hourly AMSR2 IST and ERA5-Land IST from the 15th day of each month for 2013–2020. Fig. 12 shows the variations of annual and monthly AMSR2 IST and ERA-5 Land IST. Variations of annual AMSR2 IST were similar to those of annual ERA5-Land IST. Annual IST
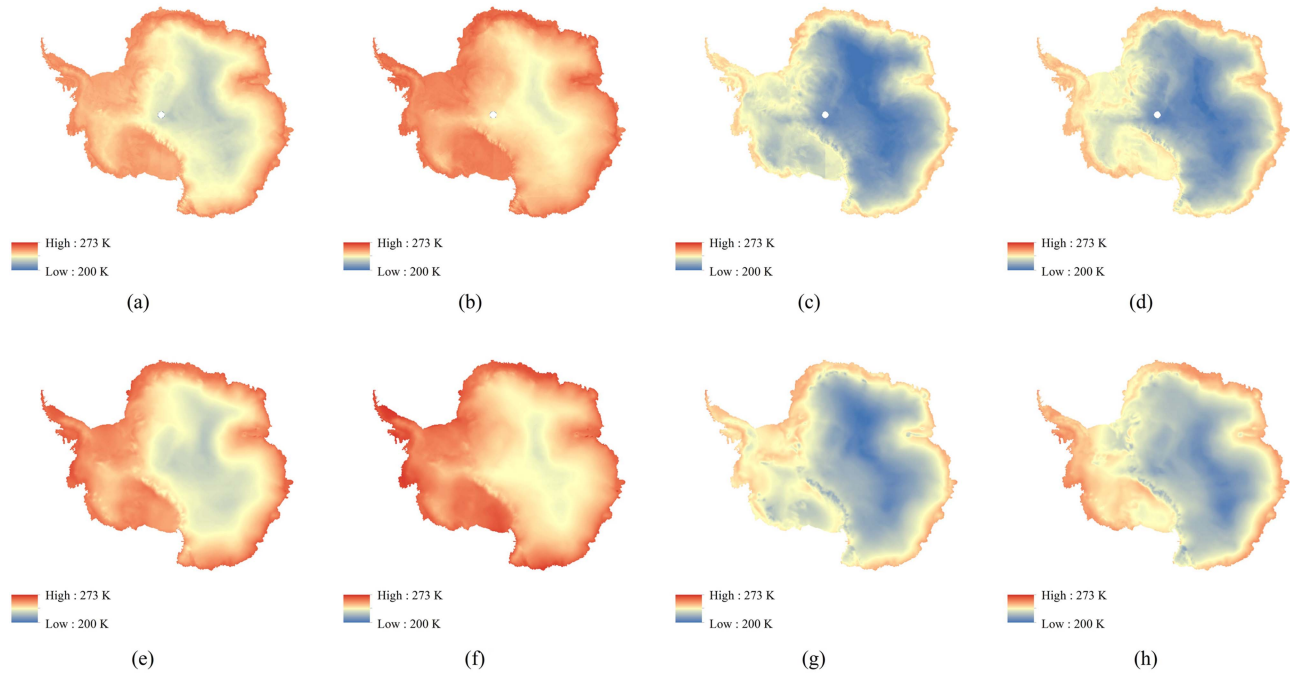
Fig. 13. IST maps from (top row) AMSR2 and (bottom row) ERA5-Land for (a) and (e) austral summer 2015, (b) and (f) summer 2020, (c) and (g) austral winter 2015, and (d) and (h) winter 2020.

decreased continuously between 2013 and 2015. Between 2015 and 2020, annual IST increased with large fluctuations. Both ERA5-Land IST and AMSR2 IST were at their minima in 2015. AMSR2 IST was at its maximum in 2018 and ERA5-Land IST was at its maximum in 2020. Annual ERA5-Land IST exceeded annual AMSR2 IST by approximately 3–4 K [see Fig. 12(a)]. Furthermore, variations of mean monthly AMSR2 IST were similar to those of mean monthly ERA5-Land IST [see Fig. 12(b)]. Both ERA5-Land IST and AMSR2 IST were at their minima in the same month; both datasets were also at their maxima in the same month. There are considerable monthly variations in the difference between AMSR2 IST and ERA5-Land IST (ΔT_{source}). In January, ΔT_{source} was negligible. It increased after January and peaked in May. We think two reasons led to the difference between AMSR2 IST and ERA5-Land IST. According to Fréville et al. [37], there is a warm bias in the IST from ERA-Interim reanalysis in Antarctica, which is mainly caused by the overestimation of the surface turbulent fluxes under very stable conditions. Cao et al. [57] also reported a warm bias in ERA5-Land soil temperature in permafrost regions, especially in winter. Therefore, we infer that there is a warm bias in ERA5-Land skin temperature, especially in austral winter. In addition, the cold bias in MODIS IST products may cause the cold bias in AMSR2 IST. Our results indicate consistent annual and monthly IST variations, which prove the reliability of the LightGBM model for IST retrieval.

Fig. 13 shows the spatial distribution of IST in austral summer and winter for 2015 (year of annual AMSR2 IST maximum) and 2020 (year of annual AMSR2 IST minimum). Clearly, the IST in West Antarctica exceeded that in East Antarctica. The IST in coastal areas was higher than that in inland areas; ΔT_{source} was
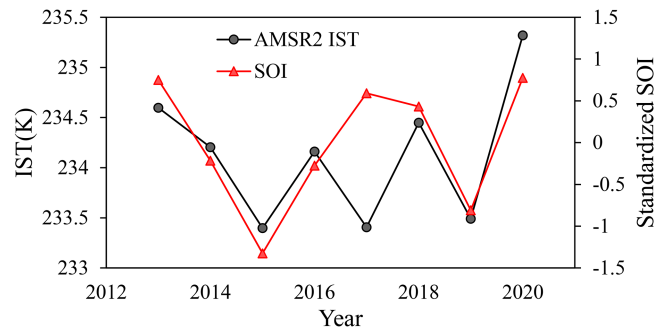

Fig. 14. Variations of Antarctic IST from AMSR2 and SOI between 2013 and 2020.

small in austral summer and obvious in austral winter. This is consistent with variations of monthly IST [see Fig. 12(b)]. We examined the difference between IST in 2020 and IST in 2015 (ΔT_{year}). In summer, ΔT_{year} was positive over almost all of Antarctica. In winter, clearly positive ΔT_{year} was found along the coast and in West Antarctica. According to many studies, the West Antarctic Ice Sheet is involved in more processes of change than the East Antarctic Ice Sheet, including ice shelf calving, elevation reduction, and surface melting [60], [61], [62]. There is a step of approximately 2 K in IST across longitude $\pm 180°$ in the area of the Ross Ice Shelf. This is an artifact of longitude $\pm 180°$ and a result of missing MOD/MYD11_L2 data in the training set in this area.

Fig. 14 shows variations of the southern oscillation index (SOI) (from https://www.cpc.ncep.noaa.gov/data/indices/) and Antarctic IST between 2013 and 2020. The IST minimum in

2015 and maximum in 2020 could be explained by the El Niño–Southern Oscillation, which has considerable influence on the Antarctic environment and climate [63], [64]. In 2015, SOI and IST are extremely low and Antarctic sea ice extent is at its maximum since 1974 (data from Bremen University https://seaice.uni-bremen.de/sea-ice-concentration/amsr-amsr2/time-series/). There is a positive correlation between SOI and Antarctic IST (Pearson's correlation coefficient = 0.664, significance level $p = 0.05$). However, the SOI and Antarctic IST variations from 2016 to 2018 are opposite. We infer that the Antarctic IST is impacted only when the SOI is extremely low or high (El Niño–Southern Oscillation event).

VI. Conclusion

We used AMSR2 swath data and a robust ensemble machine learning algorithm to develop an instantaneous IST retrieval framework. We used the MODIS IST product as ground truth for model training. Auxiliary model inputs include terrain parameters and satellite observation information. To generate a temporally continuous IST product, we developed an innovative scheme based on the difference between the acquisition time of AMSR2 swath data and that of MODIS data (i.e., STG). We used a testing sample set to evaluate the accuracy of ISTs retrieved using typical machine learning models. The performance of nonlinear models is superior to that of linear models. The LightGBM model has the best performance in IST retrieval and the RMSE of the lightGBM model is half that of the linear model. For sets with more than 8 million samples, set size has a negligible effect on model accuracy; RMSE stabilizes at 1.67 K. We compared LightGBM IST with the IST from Landsat-8, which we considered as ground truth. LightGBM IST is consistent with Landsat-8 IST; RMSE is in the range of 1.39–2.32 K. The RMSE of IST from ERA5-Land reanalysis ranges from 1.92 to 3.80 K. We compared LightGBM IST and ERA5-Land IST with IST from a BSRN station. The difference between LightGBM IST and BSRN IST is smaller than that between ERA5-Land IST and BSRN IST.

We examined the variations of LightGBM IST RMSE and mean bias with month, IST, and STG. Model accuracy is sensitive to both month and IST and insensitive to STG. This indicates that the introduction of STG hardly affects model accuracy. We suggest that the accuracy of retrieved IST may be improved if the models are trained using sample sets with different IST ranges and the input data are assigned to the correct retrieval models. We analyzed the bias between retrieved IST and NST. The mean bias between IST and NST is larger in West Antarctica and smaller in East Antarctica. In addition, the near-surface temperature inversion in austral summer is weaker than that in austral winter. Between 2013 and 2020, annual ASMR2 IST and annual ERA5-Land IST exhibited similar variations. Annual ASMR2 and ERA5-Land ISTs both decreased continuously between 2013 and 2015. Between 2015 and 2020, Annual IST increased with large fluctuations. In austral winter, AMSR IST was lower than ERA5-Land IST. This might be a result of the warm bias in ERA5-Land data. Pearson's correlation coefficient

between SOI and IST is 0.664; SOI and IST minima both occurred in 2015.

The focus of this study is to establish an empirical IST retrieval model, which has the ability to generate spatially and temporally continuous Antarctic IST data. On the contrary, this study lacks research on the physical process of PM radiation transfer, especially the interaction among PM, snow, and atmosphere. Because model accuracy is influenced by PM radiation transfer, we will focus on deriving snow surface information by combining physical clues with the empirical model in our future study. In addition, MOD/MYD11_L2 is more uncertain during polar nighttime than during polar daytime due to the lack of visible channels during polar nighttime for cloud detection. Therefore, more reliable thermal infrared IST products can improve the empirical model accuracy.

References

[1] N. P. M. van Lipzig, E. van Meijgaard, and J. Oerlemans, "Temperature sensitivity of the Antarctic surface mass balance in a regional atmospheric climate model," *J. Climate*, vol. 15, no. 19, pp. 2758–2774, 2002, doi: 10.1175/1520-0442(2002)015<2758:TSOTAS>2.0.CO;2.

[2] E. J. Steig, D. P. Schneider, S. D. Rutherford, M. E. Mann, J. C. Comiso, and D. T. Shindell, "Warming of the Antarctic ice-sheet surface since the 1957 international geophysical year," *Nature*, vol. 457, no. 7228, pp. 459–462, 2009, doi: 10.1038/nature07669.

[3] Q. Ding, E. J. Steig, D. S. Battisti, and M. Küttel, "Winter warming in West Antarctica caused by central tropical Pacific warming," *Nature Geosci.*, vol. 4, no. 6, pp. 398–403, 2011, doi: 10.1038/ngeo1129.

[4] M. B. Giovinetto, N. M. Waters, and C. R. Bentley, "Dependence of Antarctic surface mass balance on temperature, elevation, and distance to open ocean," *J. Geophys. Res.*, vol. 95, no. D4, 1990, Art. no. 3517, doi: 10.1029/JD095iD04p03517.

[5] R. J. Braithwaite, "Positive degree-day factors for ablation on the Greenland ice sheet studied by energy-balance modelling," *J. Glaciol.*, vol. 41, no. 137, pp. 153–160, 1995, doi: 10.3189/S0022143000017846.

[6] S. Wang et al., "New record of explosive warmings in East Antarctica," *Sci. Bull.*, vol. 68, no. 2, pp. 129–132, 2023, doi: 10.1016/j.scib.2022.12.013.

[7] Y. Li, T. Liu, M. E. Shokr, Z. Wang, and L. Zhang, "An improved single-channel polar region ice surface temperature retrieval algorithm using landsat-8 data," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 8557–8569, Nov. 2019, doi: 10.1109/TGRS.2019.2921606.

[8] A. C. Adolph, M. R. Albert, and D. K. Hall, "Near-surface temperature inversion during summer at summit, Greenland, and its relation to MODIS-derived surface temperatures," *Cryosphere*, vol. 12, no. 3, pp. 907–920, 2018, doi: 10.5194/tc-12-907-2018.

[9] A. Ciappa, L. Pietranera, and G. Budillon, "Observations of the Terra Nova Bay (Antarctica) polynya by MODIS ice surface temperature imagery from 2005 to 2010," *Remote Sens. Environ.*, vol. 119, pp. 158–172, 2012, doi: 10.1016/j.rse.2011.12.017.

[10] A. Gillespie, S. Rokugawa, T. Matsunaga, J. S. Cothern, S. Hook, and A. B. Kahle, "A temperature and emissivity separation algorithm for advanced spaceborne thermal emission and reflection radiometer (ASTER) images," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 4, pp. 1113–1126, Jul. 1998, doi: 10.1109/36.700995.

[11] T. Liu, Z. Wang, X. Huang, L. Cao, M. Niu, and Z. Tian, "An effective Antarctic ice surface temperature retrieval method for MODIS," *Photogrammetric Eng. Remote Sens.*, vol. 81, no. 11, pp. 861–872, 2015, doi: 10.14358/PERS.81.11.861.

[12] Z. L. Li et al., "Satellite remote sensing of global land surface temperature: Definition, methods, products, and applications," *Rev. Geophys.*, vol. 61, no. 1, pp. 1–77, 2023, doi: 10.1029/2022RG000777.

[13] M. Owe and A. A. van de Griend, "On the relationship between thermodynamic surface temperature and high-frequency (37 GHz) vertically polarized brightness temperature under semi-arid conditions," *Int. J. Remote Sens.*, vol. 22, no. 17, pp. 3521–3532, 2001, doi: 10.1080/01431160110063788.

[14] P. Song and Y. Zhang, "An improved non-linear inter-calibration method on different radiometers for enhancing coverage of daily LST estimates in low latitudes," *Remote Sens. Environ.*, vol. 264, 2021, Art. no. 112626, doi: 10.1016/j.rse.2021.112626.

[15] P. Wu et al., "A two-step deep learning framework for mapping gapless all-weather land surface temperature using thermal infrared and passive microwave data," *Remote Sens. Environ.*, vol. 277, 2022, Art. no. 113070, doi: 10.1016/j.rse.2022.113070.

[16] T. P. F. Dowling et al., "An improved cloud gap-filling method for longwave infrared land surface temperatures through introducing passive microwave techniques," *Remote Sens.*, vol. 13, no. 17, 2021, Art. no. 3522, doi: 10.3390/rs13173522.

[17] M. J. McFarland, R. L. Miller, and C. M. U. Neale, "Land surface temperature derived from the SSM/I passive microwave brightness temperatures," *IEEE Trans. Geosci. Remote Sens.*, vol. 28, no. 5, pp. 839–845, Sep. 1990, doi: 10.1109/36.58971.

[18] H. Gao, R. Fu, R. E. Dickinson, and R. I. Negron Juarez, "A practical method for retrieving land surface temperature from AMSR-E over the Amazon forest," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp. 193–199, Jan. 2008, doi: 10.1109/TGRS.2007.906478.

[19] T. R. H. Holmes, R. A. M. de Jeu, M. Owe, and A. J. Dolman, "Land surface temperature from ka band (37 GHz) passive microwave observations," *J. Geophys. Res.*, vol. 114, no. D4, pp. 1–15, 2009, doi: 10.1029/2008JD010257.

[20] F. J. Wentz, C. Gentemann, D. Smith, and D. Chelton, "Satellite measurements of sea surface temperature through clouds," *Science*, vol. 288, no. 5467, pp. 847–850, 2000, doi: 10.1126/science.288.5467.847.

[21] J. L. Foster, D. K. Hall, A. T. C. Chang, and A. Rango, "An overview of passive microwave snow research and results," *Rev. Geophys.*, vol. 22, no. 2, 1984, Art. no. 195, doi: 10.1029/RG022i002p00195.

[22] A. Kontu and J. Pulliainen, "Simulation of spaceborne microwave radiometer measurements of snow cover using in situ data and brightness temperature modeling," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1031–1044, Mar. 2010, doi: 10.1109/TGRS.2009.2030499.

[23] H. Norouzi, M. Temimi, and R. Khanbilvardi, "Global microwave land surface emissivity retrieval at the AMSR-E microwave frequencies," in *Proc. 11th Spec. Meeting Microw. Radiometry Remote Sens. Environ.*, 2010, pp. 90–95.

[24] A. Basist, N. C. Grody, T. C. Peterson, and C. N. Williams, "Using the special sensor microwave/imager to monitor land surface temperatures, wetness, and snow cover," *J. Appl. Meteorol. Climatol.*, vol. 37, no. 9, pp. 888–911, 1998, doi: 10.1175/1520-0450(1998)037<0888:UTSSMI>2.0.CO;2.

[25] P. Song and Y. Zhang, "An improved non-linear inter-calibration method on different radiometers for enhancing coverage of daily LST estimates in low latitudes," *Remote Sens. Environ.*, vol. 264, 2021, Art. no. 112626, doi: 10.1016/j.rse.2021.112626.

[26] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, 2015, doi: 10.1126/science.aaa8415.

[27] W. Tang, D. Xue, Z. Long, X. Zhang, and J. Zhou, "Near-real-time estimation of 1-km all-weather land surface temperature by integrating satellite passive microwave and thermal infrared observations," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 7001305, doi: 10.1109/LGRS.2021.3067908.

[28] Z. Wan, "New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product," *Remote Sens. Environ.*, vol. 140, pp. 36–45, 2014, doi: 10.1016/j.rse.2013.08.027.

[29] T. Maeda, Y. Taniguchi, and K. Imaoka, "GCOM-W1 AMSR2 level 1R product: Dataset of brightness temperature modified using the antenna pattern matching technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 2, pp. 770–782, Feb. 2016, doi: 10.1109/TGRS.2015.2465170.

[30] W. L. Barnes, X. Xiong, and V. V. Salomonson, "Status of terra MODIS and aqua MODIS," *Adv. Space Res.*, vol. 32, no. 11, pp. 2099–2106, 2003, doi: 10.1016/S0273-1177(03)90529-1.

[31] Z. Wan and J. Dozier, "A generalized split-window algorithm for retrieving land-surface temperature from space," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 4, pp. 892–905, Jul. 1996, doi: 10.1109/36.508406.

[32] K. Mao, J. Shi, Z. Li, Z. Qin, M. Li, and B. Xu, "A physics-based statistical algorithm for retrieving land surface temperature from AMSR-E passive microwave data," *Sci. China D*, vol. 50, no. 7, pp. 1115–1120, 2007, doi: 10.1007/s11430-007-2053-x.

[33] P. Fretwell et al., "Bedmap2: Improved ice bed, surface and thickness datasets for Antarctica," *Cryosphere*, vol. 7, no. 1, pp. 375–393, 2013, doi: 10.5194/tc-7-375-2013.

[34] T. R. Loveland and J. R. Irons, "Landsat 8: The plans, the reality, and the legacy," *Remote Sens. Environ.*, vol. 185, pp. 1–6, 2016, doi: 10.1016/j.rse.2016.07.033.

[35] M. A. Lazzara, G. A. Weidner, L. M. Keller, J. E. Thom, and J. J. Cassano, "Antarctic automatic weather station program: 30 years of polar observation," *Bull. Amer. Meteorological Soc.*, vol. 93, no. 10, pp. 1519–1537, 2012, doi: 10.1175/BAMS-D-11-00015.1.

[36] P. Wu et al., "A two-step deep learning framework for mapping gapless all-weather land surface temperature using thermal infrared and passive microwave data," *Remote Sens. Environ.*, vol. 277, 2022, Art. no. 113070, doi: 10.1016/j.rse.2022.113070.

[37] H. Fréville et al., "Using MODIS land surface temperatures and the Crocus snow model to understand the warm bias of ERA-Interim reanalyses at the surface in Antarctica," *Cryosphere*, vol. 8, no. 4, pp. 1361–1373, 2014, doi: 10.5194/tc-8-1361-2014.

[38] J. Schmetz, "Towards a surface radiation climatology: Retrieval of downward irradiances from satellites," *Atmos. Res.*, vol. 23, no. 3/4, pp. 287–321, 1989, doi: 10.1016/0169-8095(89)90023-9.

[39] B. Bell et al., "The ERA5 global reanalysis: Preliminary extension to 1950," *Quart. J. Roy. Meteorological Soc.*, vol. 147, no. 741, pp. 4186–4227, 2021, doi: 10.1002/qj.4174.

[40] B. McNicholl, Y. H. Lee, A. G. Campbell, and S. Dev, "Evaluating the reliability of air temperature from ERA5 reanalysis data," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 1004505, doi: 10.1109/LGRS.2021.3137643.

[41] J. M. van Wessem, M. R. van den Broeke, B. Wouters, and S. Lhermitte, "Variable temperature thresholds of melt pond formation on Antarctic ice shelves," *Nature Climate Change*, vol. 13, no. 2, pp. 161–166, 2023, doi: 10.1038/s41558-022-01577-1.

[42] J. Muñoz-Sabater et al., "ERA5-Land: A state-of-the-art global reanalysis dataset for land applications," *Earth Syst. Sci. Data*, vol. 13, no. 9, pp. 4349–4383, 2021, doi: 10.5194/essd-13-4349-2021.

[43] O. Sagi and L. Rokach, "Ensemble learning: A survey," *WIREs Data Mining Knowl. Discov.*, vol. 8, no. 4, 2018, Art. no. e1249, doi: 10.1002/widm.1249.

[44] T. Chen and C. Guestrin, "XGBoost," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2016, pp. 785–794.

[45] M. Sayed, D. Riaño, and J. Villar, "Predicting duration of mechanical ventilation in acute respiratory distress syndrome using supervised machine learning," *J. Clin. Med.*, vol. 10, no. 17, 2021, Art. no. 3824, doi: 10.3390/jcm10173824.

[46] G. Ke et al., "LightGBM: A highly efficient gradient boosting decision tree," in *Proc 31st Int. Conf. Neural Inf. Process. Syst.*, 2017, pp. 3149–3157.

[47] J. Wang et al., "Estimating snow depth by combining satellite data and ground-based observations over Alaska: A deep learning approach," *J. Hydrol.*, vol. 585, 2020, Art. no. 124828, doi: 10.1016/j.jhydrol.2020.124828.

[48] Y. Zhong, L. Meng, Z. Wei, J. Yang, W. Song, and M. Basir, "Retrieval of all-weather 1 km land surface temperature from combined MODIS and AMSR2 data over the Tibetan plateau," *Remote Sens.*, vol. 13, no. 22, 2021, Art. no. 4574, doi: 10.3390/rs13224574.

[49] Y. Yang, K. Zha, Y.-C. Chen, H. Wang, and D. Katabi, "Delving into deep imbalanced regression," Feb. 2021. [Online]. Available: http://arxiv.org/pdf/2102.09554v2

[50] P. Burman, "A comparative study of ordinary cross-validation, v -fold cross-validation and the repeated learning-testing methods," *Biometrika*, vol. 76, no. 3, pp. 503–514, 1989, doi: 10.1093/BIOMET/76.3.503.

[51] G. Li, X. Zhao, K. Du, F. Ru, and Y. Zhang, "Recognition and evaluation of bridge cracks with modified active contour model and greedy search-based support vector machine," *Automat. Construction*, vol. 78, pp. 51–61, 2017, doi: 10.1016/j.autcon.2017.01.019.

[52] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *J. Roy. Stat. Soc. B*, vol. 67, no. 2, pp. 301–320, 2005, doi: 10.1111/j.1467-9868.2005.00503.x.

[53] J. Park and K.-Y. Kim, "Instance variant nearest neighbor using particle swarm optimization for function approximation," *Appl. Soft Comput.*, vol. 40, pp. 331–341, 2016, doi: 10.1016/j.asoc.2015.10.055.

[54] R. K. Yadav and Anubhav, "PSO-GA based hybrid with Adam optimization for ANN training with application in medical diagnosis," *Cogn. Syst. Res.*, vol. 64, pp. 191–199, 2020, doi: 10.1016/j.cogsys.2020.08.011.

[55] M. Saber et al., "Examining LightGBM and CatBoost models for Wadi flash flood susceptibility prediction," *Geocarto Int.*, vol. 37, no. 25, pp. 7462–7487, 2022, doi: 10.1080/10106049.2021.1974959.

[56] Z. Cui and G. Gong, "The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features," *NeuroImage*, vol. 178, pp. 622–637, 2018, doi: 10.1016/j.neuroimage.2018.06.001.

[57] B. Cao, S. Gruber, D. Zheng, and X. Li, "The ERA5-land soil temperature bias in permafrost regions," *Cryosphere*, vol. 14, no. 8, pp. 2581–2595, 2020, doi: 10.5194/tc-14-2581-2020.

[58] D. Hall, J. Box, K. Casey, S. Hook, C. Shuman, and K. Steffen, "Comparison of satellite-derived and in-situ observations of ice and snow surface temperatures over Greenland," *Remote Sens. Environ.*, vol. 112, no. 10, pp. 3739–3749, 2008, doi: 10.1016/j.rse.2008.05.007.

[59] N. B. Miller et al., "Surface-based inversions above central Greenland," *J. Geophys. Res. Atmos.*, vol. 118, no. 2, pp. 495–506, 2013, doi: 10.1029/2012JD018867.

[60] B. Zhang et al., "Elevation changes of the Antarctic ice sheet from joint Envisat and CryoSat-2 radar altimetry," *Remote Sens.*, vol. 12, no. 22, 2020, Art. no. 3746, doi: 10.3390/rs12223746.

[61] C. L. Jakobs et al., "A benchmark dataset of in situ Antarctic surface melt rates and energy balance," *J. Glaciol.*, vol. 66, no. 256, pp. 291–302, 2020, doi: 10.1017/jog.2020.6.

[62] C. Baumhoer, A. Dietz, S. Dech, and C. Kuenzer, "Remote sensing of Antarctic glacier and ice-shelf front dynamics—A review," *Remote Sens.*, vol. 10, no. 9, 2018, Art. no. 1445, doi: 10.3390/rs10091445.

[63] L. Yu et al., "Influence of the Antarctic oscillation, the Pacific–South American modes and the El Niño–southern oscillation on the Antarctic surface temperature and pressure variations," *Antarctic Sci.*, vol. 24, no. 1, pp. 59–76, 2012, doi: 10.1017/S095410201100054X.

[64] X. Yuan, "ENSO-related impacts on Antarctic sea ice: A synthesis of phenomenon and mechanisms," *Antarctic Sci.*, vol. 16, no. 4, pp. 415–425, 2004, doi: 10.1017/S0954102004002238.

Yachao Li received the B.S. degree in geodesy and geomatics engineering from Wuhan University, Wuhan, China, in 2017, and the M.S. degree in photogrammetry and remote sensing in 2020 from the Chinese Antarctic Center of Surveying and Mapping, Wuhan University, where he is currently working toward the Ph.D. degree in polar remote sensing and environmental change.

His research interests include global warming, polar ice sheets, and land surface temperature retrieval.

Zemin Wang received the Ph.D. degree in geodesy and surveying engineering from Wuhan University, Wuhan, China, in 2001.

He is currently a Professor with the Chinese Antarctic Center of Surveying and Mapping, Wuhan University. His research focuses on the study of changes in mass balance of polar ice sheets and shelves, as well as remote sensing monitoring of polar snowmelt.

Qiangqiang Yuan (Member, IEEE) received the B.S. degree in surveying and mapping engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2006 and 2012, respectively.

In 2012, he joined the School of Geodesy and Geomatics, Wuhan University, where he is currently a Professor. He has authored or coauthored more than 90 research articles, including more than 70 peer-reviewed articles in international journals, such as *Remote Sensing of Environment*, *ISPRS Journal of Photogrammetry and Remote Sensing*, IEEE TRANSACTIONS ON IMAGE PROCESSING, and IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. His research interests include image reconstruction, remote sensing image processing and application, and data fusion.

Dr. Yuan was a recipient of the Youth Talent Support Program of China in 2019, the Top-Ten Academic Star of Wuhan University in 2011, and the recognition of Best Reviewers of *IEEE Geoscience and Remote Sensing Letters* in 2019. In 2014, he received the Hong Kong Scholar Award from the Society of Hong Kong Scholars and the China National Postdoctoral Council. He is an Associate Editor for five international journals and he has frequently served as a referee for more than 40 international journals for remote sensing and image processing.

Tingting Liu (Member, IEEE) received the B.S. degree in computer science and technology and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2004 and 2009, respectively.

She conducted postdoctoral research with the Institute of Space and Earth Information Science, The Chinese University of Hong Kong, from December 2009 to July 2011. Since 2011, she has been with the Chinese Antarctic Center of Surveying and Mapping, Wuhan University, where she is currently a Full Professor. She has participated in several field campaigns in Antarctica and the Arctic to study ice surface properties. She has been a PI and a Co-PI on a number of projects in new image processing technologies for retrieving sea ice parameters, and their variations and impacts. Her research interests include remote sensing of Polar Regions and climate change, specifically image processing technologies for sea ice and ice sheet mapping.