A Data-Driven Model for Estimating Clear-Sky Surface Longwave Downward Radiation Over Polar Regions

Mengfei Guo, Jie Cheng, *Senior Member, IEEE*, and Qi Zeng

***Abstract*— Polar regions play a crucial role in global climate change. Surface longwave downward radiation (SLDR) is a primary energy source for the polar surface and plays an essential role in studies of polar hydrology, temperature, and climate. Therefore, accurately estimating the SLDR over polar regions is highly important. However, the accuracies of existing polar SLDR datasets and SLDR inversion methods are insufficient to meet the requirements of relevant research. In this study, we developed a data-driven model for high spatial resolution clear-sky SLDRs estimated from Moderate Resolution Imaging Spectroradiometer (MODIS) imagery in polar regions. The model comprises two layers: the first layer incorporates three machine learning models, namely, eXtreme gradient boosting (XGBoost), convolutional neural network (CNN), and transformer, while the second layer consists of a stacking meta-model. The ground measurements collected from 51 sites were used to train and validate the developed model. The bias, RMSE, and R^2 of the model training are zero, 14.15 W/m^2, and 0.95, respectively, whereas the values for the validation are 0.49, 15.35 W/m^2, and 0.9, respectively. We also compared the accuracies of the ERA5 and CERES-SYN SLDR data with the SLDR estimated by the developed model. The results indicate that the developed model is superior to the ERA5 and CERES-SYN SLDR models when evaluated at the validation sites. In addition, we analyzed the performance of the developed model under different elevations and seasons, demonstrating its robustness in different situations.**

***Index Terms*— Convolutional neural network (CNN), eXtreme gradient boosting (XGBoost), machine learning, polar region, stacking, surface longwave downward radiation (SLDR), surface radiation budget, transformer.**

I. Introduction

THE surface longwave downward radiation (SLDR, 4–100 μm) is a key parameter in the Earth's energy balance and represents the thermal energy transferred from the atmosphere to the surface [1], [2]. The SLDR plays a crucial role in various processes, including global energy distribution, climate change, hydrological cycles, and interactions within the cryosphere [3], [4]. Therefore, accurate estimates of the SLDR are essential for global climate change, hydrological cycle, and polar environmental research [5], [6].

The polar region represents a highly vulnerable area to the effects of global warming. Due to the low solar zenith angles and high albedo caused by ice and snow coverage, the amount of solar radiation absorbed by the surface in polar regions is significantly reduced compared to mid-latitude regions [7]. As a result, SLDR serves as a primary energy source in polar regions and profoundly influences ice cap variations, sea ice formation, and melting processes [8], [9], [10]. Accurate estimates of the SLDR in polar regions are critically important for observing and predicting polar climate change, determining complex hydrological cycles, and investigating the global radiation balance [11], [12]. However, owing to the unique geographical attributes of polar regions, including their high latitudes, extreme climatic conditions, and distinct physical properties of the surface and atmosphere compared to those of other regions, estimating the SLDR at the polar surface is highly challenging.

There are three primary ways to acquire SLDR data over polar regions: 1) ground measurements; 2) reanalysis or land surface model simulation; and 3) satellite remote sensing. Currently, several in situ networks, such as the baseline surface radiation network (BSRN), coordinated energy and water cycle observation project (CEOP), and AmeriFlux, have been established in polar regions [13], [14], [15]. In situ observations have the advantages of high observational accuracy and fine-grained temporal resolution. However, the installation and maintenance costs of these in situ sites are prohibitively high, necessitating substantial financial and human resources. Furthermore, the number of observation sites in polar regions is limited and unevenly distributed due to the harsh natural environments in high-latitude areas and the restrictions imposed by both nature and national policies. Consequently, this scarcity of sites results in limited SLDR data availability and poses challenges to the effective monitoring of large-scale SLDRs in polar regions.

General circulation models (GCMs) and reanalysis datasets often provide global, long-term time series of meteorological and land surface data, including polar SLDR data. The most representative of the GCM simulation datasets is the

Manuscript received 26 January 2024; revised 14 May 2024; accepted 16 June 2024. Date of publication 24 June 2024; date of current version 15 July 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 42071308 and Grant 42192581 and in part by the Second Tibetan Plateau Scientific Expedition and Research Program (STEP) under Grant 2019QZKK0206. *(Corresponding author: Jie Cheng.)*

Mengfei Guo and Jie Cheng are with the State Key Laboratory of Remote Sensing Science, Beijing Normal University and Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100875, China (e-mail: Jie_Cheng@bnu.edu.cn).

Qi Zeng is with the College of Water Science, Beijing Normal University, Beijing 100875, China.

This article has supplementary downloadable material available at https://doi.org/10.1109/TGRS.2024.3418205, provided by the authors.

Digital Object Identifier 10.1109/TGRS.2024.3418205

coupled model intercomparison project (CMIP), which was initiated by the World Climate Research Program (WCRP) in 1995 [16]. With the continuous growth of climate/Earth system model development teams worldwide, CMIP has evolved from CMIP1 to the present CMIP6. However, the spatial resolutions of GCMs are coarse (approximately $1^\circ \times 1^\circ$), which makes it challenging to meet the needs of subsequent research. In addition, studies indicate that various GCMs perform less effectively over polar regions [17]. Various reanalysis datasets, such as ERA5 and MERRA-2, have also been widely used. For example, ERA5 has a spatial resolution of $0.25^\circ \times 0.25^\circ$ and provides hourly SLDR data for polar regions from 1979 to the present [18]. Like those of GCMs, reanalysis datasets have relatively coarse spatial resolutions, and studies have shown that the accuracy of SLDR data provided by reanalysis datasets decreases with increasing latitude [19].

The advancement of remote sensing technology has opened up new possibilities for estimating SLDRs [20]. Various SLDR estimation methods have been proposed and can be categorized into four main types: physical models, parametric models, hybrid models, and data-driven models. Physical models utilize radiation transfer models or highly parameterized radiation transfer equations to calculate SLDRs based on atmospheric profiles. Physical models are grounded in solid physical foundations and employ mathematical formulae to simulate the process of longwave radiation transmission from the atmosphere to the Earth's surface [21], [22]. While physical models theoretically offer high estimation accuracy, they demand accurate atmospheric profile data inputs that are not always available through certain remote sensing observations. In addition, these models have high computational requirements. Parametric models emphasize the impact of near-surface meteorological variables, such as near-surface temperature and water vapor pressure, on the SLDR. These models establish empirical relationships between parameters and the SLDR. Hybrid models, on the other hand, are mostly built upon databases generated by radiative transfer model simulations [23], [24]. Linear and nonlinear relationships are established between various specific sensor data, such as spectral radiance, elevation, sensor zenith angle (SZA), and SLDR [25], [26]. In recent years, due to the advantages of the physical foundation and computational simplicity of models, many parametric and hybrid models have been proposed and validated. Guo et al. [27] refitted coefficients for seven parametric models. Except for the Swinbank [28] and Idso and Jackson [29] models, which performed poorly across all conditions, the remaining five models exhibited severe underestimations over the polar site validation, with biases ranging from -10.78 to -26.07 W/m^2. Wu et al. [30] assessed eight parametric models and one hybrid model for clear-sky SLDR estimation. The results indicated that all the models performed poorly over polar surfaces covered with glaciers, with relative errors exceeding 20%. Yu et al. [31] proposed a new parametric model for clear-sky SLDR estimation and compared it with two other parametric models, namely, Zhou et al. [32] and Gupta et al. [25], and two hybrid models, Tang and Li [23] and Wang and Liang [24]. The validation results for the three BSRN polar sites indicated that the newly proposed model had a bias ranging from -10.6 to -1.8 W/m^2, indicating nonnegligible underestimation. The bias values of the other four methods ranged from -17.3 to 34.0 W/m^2, with RMSE values between 17.3 and 34.8 W/m^2. In addition, Lu et al. [33] developed some surface solar radiation estimate models combining hybrid algorithms and machine learning techniques. Although the models focus on shortwave radiation, it is a good idea that can be used in the estimate of SLDR. In general, current parametric and hybrid methods exhibit limited generalizability, particularly because they fail to estimate clear-sky SLDRs over polar regions. On the one hand, this difference is primarily attributed to the development of the model, which is predominantly based on climate features from mid to low latitudes and, as a result, is inadequately suited for extremely cold and arid climate conditions in polar regions. On the other hand, observational errors and outliers significantly impact model performance, with climatic anomalies in polar regions often introducing substantial errors into the model.

With the development of versatile SLDR retrieval algorithms, several publicly available SLDR datasets, such as the Global Energy and Water Exchanges Project-Surface Radiation Budget (GEWEX-SRB), International Satellite Cloud Climatology Project-Flux Data (ISCCP-FD), Clouds and the Earth's Radiant Energy System-Synoptic Radiative Fluxes and Clouds (CERES-SYN), and Global Land Surface Satellite (GLASS), have been derived from remote sensing data collected over the past few decades [34], [35], [36], [37]. However, some of these products have limited temporal coverage and are no longer updated. For instance, GEWEX-SRB and ISCCP-FD provide data only up to 2007 and 2010, respectively. Xin et al. [38] assessed three different SLDR products derived from remote sensing data over polar regions, namely, GEWEX-SRB, ISCCP-FD, and CERES-SYN. Validation results showed that CERES-SYN has the best performance, with an RMSE of 26.9 W/m^2, compared to 35.8 and 40.5 W/m^2 for GEWEX-SRB and ISCCP-FD, respectively. This finding suggested that the accuracy of SLDR products over polar regions provided by remote sensing datasets is insufficient for meeting climate research requirements.

Data-driven models, which include various machine learning and deep learning algorithms, do not presuppose mathematical or physical relationships among data but rather delve deeply into uncovering correlations between input features and their intrinsic connections with the predicted values. This approach significantly reduces the impact of the aforementioned challenges on model performance, making it suitable for addressing SLDR high-accuracy estimations over polar regions. With the advancement of information science and computing capabilities, machine learning and deep learning algorithms have undergone rapid development in recent years [39], [40], [41]. For instance, Lopes et al. [42] employed the MARS machine learning model to estimate all-sky SLDRs by utilizing surface meteorological data from ERA5 and cloud data from the MeteoSat Second Generation (MSG) satellite. The model achieved a satisfactory overall estimation result, with a bias of 0.65 W/m^2 and an RMSE of 18.76 W/m^2.

TABLE I
MAIN DATA USED IN SLDR ESTIMATE

Data source	Data name	Parameter	Usage
In situ observations	-	Ta, RH, PS	Model input
		SLDR	Model training output
DEM data	GMTED2010	Elevation	Model input
MODIS products	MYD02	TOA radiance of Bands 27-29, 31-34	Model input
	MYD03	SZA	Model input
		Latitude, Longitude	Spatial matching
	MYD35	CM	Clear sky detection

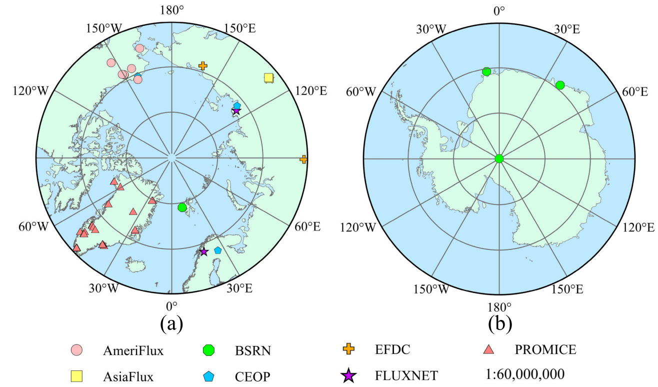

Fig. 1. Spatial distribution of the collection sites in the polar regions. (a) Arctic. (b) Antarctic.

The model was evaluated using a polar site, GVN, which exhibited a bias of -2.98 W/m^2 and an RMSE of 25.5 W/m^2 for clear-sky SLDR validation. Kim and Kim [43] addressed the uncertainty in current SLDR reanalysis products in the Arctic region by proposing optimization through the use of the CNN model. They applied the CNN to three SLDR products, namely, ERA5, polar weather research and forecasting (PWRF) improved ERA5, and PWRF ERA5 with additional data assimilation (DA). Site validation results demonstrated significantly improved accuracy after CNN optimization, with respective reductions in RMSE of 17.62%, 14.98%, and 13.14% for the three products. These studies demonstrate the potential effectiveness of data-driven models for accurate SLDR estimation over polar regions.

Above all, although scholars have recognized gaps and limitations, such as coarse spatiotemporal resolution and insufficient accuracy, in SLDR estimation over polar regions, there is still a lack of models for achieving accurate clear-sky SLDR estimation in these areas. To address this issue, this study proposes a new data-driven model in which a stacking model is established based on three different machine-learning models. The remainder of this article is organized as follows: Section II provides an introduction to the multisource data used in the study. In Section III, the three fundamental algorithms and the stacking model employed in the research are outlined. Section IV presents the model training and validation results. Section V conducts a discussion and analysis. Finally, Section VI concludes the article with a summary.

II. DATA AND MATERIALS

Three types of datasets were employed in this study, including 1) in situ observations; 2) digital elevation model (DEM) data, and (3) MODIS products. The input of the SLDR estimate model was constructed by referencing numerous widely used parameterized schemes and hybrid methods [2], [23], [24], [27]. SLDR emitted by the lower 50 hPa of the atmosphere accounts for approximately 85% of the total, whereas under inversion conditions, this proportion decreases to 63.4% [44], [45]. Therefore, the model inputs must reflect both the near-surface atmospheric conditions and the middle to upper atmosphere conditions. A total of 12 parameters were employed as input for the model, and the details regarding these parameters are provided in Table I. Among these, the air temperature at screen level (Ta) and relative humidity (RH) provided by in situ observations are the primary inputs for parameterized schemes, reflecting the SLDR emitted by the lower atmosphere near the surface. The pressure (PS) data supplements the conditions of near-surface climate. Elevation data provided by DEM, commonly used in some hybrid methods and impacting SLDR to some extent, is also included as one of the inputs. The top-of-atmosphere (TOA) radiance of MODIS bands 27–29 and 31–34 primarily captures the thermal radiation characteristics of the overall atmosphere. The SZA reflecting the path length was also input to the model as supplementary to MODIS band radiance. In addition, latitude and longitude data provided by MYD03 are used for spatial matching of MODIS products; in situ observations, DEM data, and the cloud mask (CM) data provided by MYD35 are used to identify clear and cloudy skies.

A. In Situ Observations

This study employed observational data collected from 51 sites at seven flux networks, namely, AmeriFlux, AsiaFlux, the BSRN, the Coordinated Enhanced Observing Period (CEOP), the European Fluxes Database Cluster (EFDC), FLUXNET, and the Program for Monitoring of the Greenland Ice Sheet (PROMICE). Among them, 48 sites were located in the Arctic region, and three sites were located in the Antarctic region. Fig. 1 illustrates the distribution of these sites. The extracted variables included surface air temperature (Ta), RH, surface pressure (PS), and SLDR.

The flux networks selected in this study have been extensively employed in quantitative remote sensing research. The AmeriFlux is situated in the Americas and is dedicated to measuring ecosystem carbon, water, and energy fluxes [15]. In this study, data from nine AmeriFlux sites located above 60°N latitude were used. Currently, AsiaFlux comprises 117 tower flux observation sites, two of which were utilized in this study [46]. The BSRN is part of the Global Climate Observing System (GCOS) [13]. The solar and atmospheric radiation data provided by the BSRN exhibit high precision and high temporal resolution (1 min). This study included a total of four BSRN sites. To prevent an undue impact on model training due to the excessive data volume at the BSRN sites, all the data were resampled to a 30-min resolution. In addition,

TABLE II
DETAILED INFORMATION REGARDING THE SLDR MEASUREMENT INSTRUMENTS ACROSS FLUX NETWORKS

Network	Instrument type	Instrument name	Spectral range	Effective temperature	Uncertainty
AmeriFlux	Net Radiometer	Kipp & Zonen CNR1	5 – 50μm	-40 – 70 °C	± 10 %
AsiaFlux	Pyrgeometer	EKO MS-201F	4.5 – 42μm	-40 – 80 °C	Window heating offset < 4 W/m^2 Zero offset < 1 W/m^2
BSRN	Pyrgeometer	Eppley PIR	4 – 50μm	-50 – 80 °C	5 W/m^2
CEOP	Pyrgeometer	Kipp & Zonen CGR4	4.5 – 42μm	-40 – 80 °C	Window heating offset < 4 W/m^2 Zero offset < 2 W/m^2
		Eppley PIR	4 – 50μm	-50 – 80 °C	5 W/m^2
		EKO MS-202F	3 – 50μm	-20 – 40 °C	± 5 %
		EKO MS-802F	2.85 – 30μm	-40 – 80 °C	Temperature Response < 2 %
EFDC	Net Radiometer	Kipp & Zonen CNR1	5 – 50μm	-40 – 70 °C	± 10 %
		Kipp & Zonen CNR4	4.5 – 42μm	-40 – 80 °C	Temperature Response < 4 %
	Pyrgeometer	Kipp & Zonen CG1	4.5 – 42μm	-40 – 80 °C	± 10 %
FLUXNET	Net Radiometer	Kipp & Zonen CNR1	5 – 50μm	-40 – 70 °C	± 10 %
	Pyrgeometer	Kipp & Zonen CG1	4.5 – 42μm	-40 – 80 °C	± 10 %
PROMICE	Net Radiometer	Kipp & Zonen CNR1	5 – 50μm	-40 – 70 °C	± 10 %
		Kipp & Zonen CNR4	4.5 – 42μm	-40 – 80 °C	Temperature Response < 4 %

we collected data from five CEOP sites. Apart from the STT site, which has a temporal resolution of 1 h, the other four sites have a temporal resolution of 30 min [14]. The EFDC has compiled observational data from research projects funded by the European Union since 1996 [47]. This study utilized data from three of these sites. FLUXNET is a "network of networks" that connects various regional networks, including AmeriFlux, AsiaFlux, ICOS, and NEON [48], [49]. It provides data to the public after standardized data processing. We selected three FLUXNET sites with a temporal resolution of 1.5 h. PROMICE sites are all located in Greenland and were originally set up to assess changes in the Greenland Ice Sheet by accurately monitoring surface energy and mass balances [50]. Table II summarizes the specific information on the SLDR measurement instruments used by each flux network.

All the selected sites were situated above 60° latitude, and the data collected from the sites spanned a period of 19 years, from 2002 to 2020. Detailed information on site coordinates, elevations, temporal resolutions, and other parameters can be found in Supplementary Table S1.

B. DEM Data

In this study, we used Global Multiresolution Terrain Elevation Data (GMTED2010) for elevation information, with a spatial resolution of 7.2 arc-s [51]. The global validation accuracy of GMTED2010 is approximately 6 m in terms of the root mean square error (RMSE). The GMTED2010 dataset provides elevation data only for latitudes ranging from 84°N to 56°S and does not include Greenland; the elevation data for the BSRN and PROMICE sites were obtained from site-provided sources.

C. MODIS Products

The Moderate Resolution Imaging Spectroradiometer (MODIS) is an important instrument used in the Earth Observing System (EOS) program of the United States; it is carried out onboard the Terra and Aqua polar-orbiting satellites. The Terra and Aqua satellites are placed in sun-synchronous polar orbits, ensuring that each satellite passes over the same area at approximately the same local time every day. This regular revisiting time enables MODIS to provide at least four daily updates; additionally, with higher latitudes in polar regions, there can be more frequent updates and some locations experience updates as frequently as a dozen times a day.

In this study, MODIS data from the Aqua satellites were used, including MODIS radiance data products (MYD021KM), geolocation products (MYD03), and CM products (MYD35_L2), all with a spatial resolution of 1 km. The MYD021KM product provides radiance data for various spectral bands. In this study, radiance data from bands 27–29 and 31–34 were used. The MYD03 product mainly provides latitudinal and longitudinal data, as well as SZA data. These data were used for spatial matching with site data and as supplementary information for the radiance data. The MYD35_L2 product provides CM information and was used for selecting clear-sky data in this study.

D. Data Quality Control

CM data from MYD35 were first utilized to discern clear-sky conditions. A CM value of three denoting confident clear-sky was retained for further analysis. Subsequently, site measurements were used to identify and address potential anomalies. Ta below -100 °C or above 80 °C was deemed abnormal, and RH values exceeding 90% indicated possible cloudy conditions. Consequently, we removed these data points.

Under clear-sky conditions, Ta can represent the atmospheric effective temperature. Thus, additional data quality control can be implemented based on the Stefan–Boltzmann equation [52], [53], as shown in (1) where ε_{clr} denotes clear-sky atmospheric emissivity and σ is the Stefan–Boltzmann constant. In this study, we substituted the in situ observed Ta into (1) and set the threshold range of ε_{clr} to [0.5, 1.5]. SLDR observations with ε_{clr} outside this calculated range were removed

$$\text{SLDR}_{\text{clr}} = \sigma \varepsilon_{\text{clr}} (T_a)^4. \tag{1}$$

III. Methodology

In recent years, a wealth of machine-learning algorithms have been proposed and applied to quantitative remote sensing studies. As a tree-based machine-learning algorithm, eXtreme gradient boosting (XGBoost) has demonstrated remarkable parameter inversion capabilities in soil moisture retrieval [54], land surface evapotranspiration prediction [55], surface shortwave net radiation estimation [56], and surface longwave net radiation estimation [26]. The convolutional neural network (CNN), a well-established deep learning algorithm with a relatively long development history, has also achieved promising performance in land surface temperature retrieval [57], subsurface temperature estimation [58], shortwave radiation estimation [59], and solar radiation estimation [60]. Although not yet widely used in quantitative remote sensing studies, the transformer, a newcomer to the realm of deep learning algorithms, has already demonstrated its superiority in remote sensing image change detection [61] and solar radiation time series prediction [62].

Therefore, we introduce a novel data-driven model designed for clear-sky SLDR estimates in polar regions. This model seamlessly integrates three fundamental machine-learning algorithms, namely, XGBoost, CNN, and transformer, via a stacking model. This integration leveraged the unique strengths of each foundational algorithm and heightened the overall robustness of the estimation results.

A. Stacking Model

The stacking model is an ensemble algorithm that combines multiple diverse foundational learning models to enhance the predictive performance and generalization capability [63]. When discussing ensemble learning methods, it is worth mentioning the bagging and boosting models [64]. Bagging (bootstrap aggregating) involves training independent models through multiple random data subsampling iterations and then combining their outputs, typically by averaging or voting, to improve performance. A classic example of Bagging is the random forest algorithm. In contrast, boosting is a method of sequentially training multiple models, with each model attempting to correct the errors of the previous model; for example, XGBoost and LightGBM. Both bagging and boosting typically rely on a single type of foundational model. Stacking, on the other hand, distinguishes itself from these models by focusing on aggregating insights from different foundational learners, thus achieving more accurate and robust predictions. Stacking can be viewed as a meta-model that takes the outputs from multiple foundational learning models as inputs and utilizes these outputs to train the meta-model, further enhancing its predictive performance. Prior research has demonstrated the substantial potential of stacking, as it enables multiple machine learning models to collaborate, effectively addressing complex problems and delivering optimal solutions [65], [66].

The stacking model constructed in this study consists of two layers. In the first layer, we employed three distinct types of foundational learning models, namely, XGBoost, CNN, and transformer, to conduct initial predictions on the input data. By integrating the predictive results generated by these base models, in the second layer, we further trained a meta-model to achieve precise SLDR estimation. The structure of the stacking model is depicted in Fig. 2. For the meta-model selection, we opted for the Bayesian model average (BMA) algorithm based on posterior probabilities. It is worth noting that stacking does not simply select the best model or result from the foundational models but combines the information and insights derived from each of the foundational models. This approach leverages the strengths of each model to compensate for the weaknesses of the others, addressing potential overfitting issues and enhancing the reliability of the predictions simultaneously. Therefore, through stacking, we comprehensively combined the capabilities of XGBoost, CNN, and transformer. All three models perform well yet with distinct underlying principles and strengths to achieve superior and robust performance in the task of clear-sky SLDR estimation over the polar region.

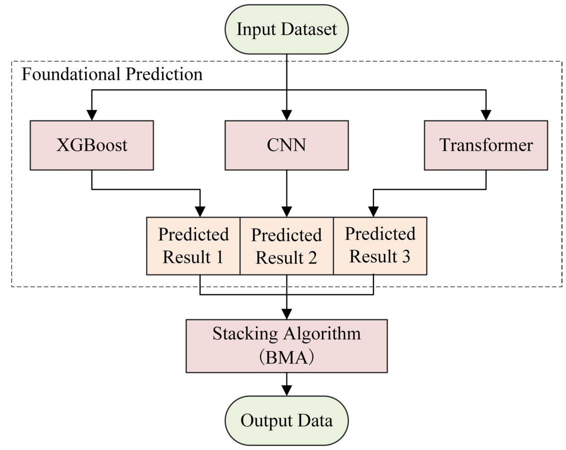

Fig. 2. Diagram of the constructed stacking model used in this study.

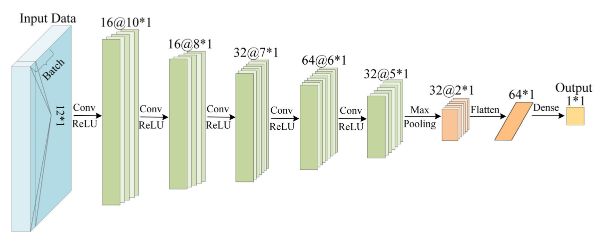

Fig. 3. Diagram of the 1-D CNN Model constructed in this study.

B. XGBoost

XGBoost is a powerful machine-learning algorithm initially proposed by Chen and Guestrin [67] for supervised learning problems. It falls within the category of gradient-boosting decision tree (GBDT) algorithms. Its fundamental principle is gradient boosting, which can be viewed from a statistical perspective as a combination of additive models and forward optimization algorithms. Mathematically, we can summarize our model using (2). In the equation, $\widehat{y_i}$ represents the predicted result of the model, x_i denotes the input features, K is the total number of trees, each f_k corresponds to an independent tree, and F represents the space of the overall regression tree. The prediction result of the gradient boosting algorithm can be explained as the sum of the predictions from each individual tree

$$\widehat{y_i} = \sum\nolimits_{k=1}^{K} f_k(x_i), \quad f_k \in \mathcal{F}. \tag{2}$$

XGBoost builds upon the gradient boosting framework and incorporates several optimization techniques to enhance its performance. The main optimization steps in XGBoost are as follows: first, XGBoost creates decision trees in a parallel manner. This differs from the sequential tree-building process in traditional GBDT, which results in faster and more efficient model training. Second, to mitigate overfitting, XGBoost introduces a regularization term in the objective function. This term prevents the model from becoming overly complex and improves its ability to generalize to unseen data. Third, XGBoost improves computational accuracy by optimizing the loss function using second-order Taylor expansion, which enhances the accuracy of calculations involved in the optimization process. The objective function of XGBoost is defined in (3) where l represents the loss function that measures the discrepancy between the predicted values and target values. N denotes the number of samples, t indicates the number of trees, Ω represents the regularization term used to penalize complex models, T represents the number of leaves, and w represents leaf weights. The goal is to minimize $\mathrm{Obj}(\Theta)$ to obtain the optimal model. In (4) $\widehat{y_i}^{(t)}$ represents the predicted value for the ith sample by the first t trees. It can be expressed as the sum of predictions for the ith sample made by the previous $t-1$ trees and the prediction made by the tth tree

$$\mathrm{Obj}(\Theta) = \sum\nolimits_{i=1}^{N} l\left(y_i, \widehat{y_i}\right) + \sum\nolimits_{j=1}^{t} \Omega\left(f_j\right), \;\; f_j \in \mathcal{F} \tag{3}$$

where $\Omega(f) = \gamma T + \frac{1}{2}\lambda \|w\|^2$

$$\widehat{y_i}^{(t)} = \sum\nolimits_{j=1}^{t} f_k(x_i) = \widehat{y_i}^{(t-1)} + f_t(x_i). \tag{4}$$

By substituting (4) into (3) and assuming that the first "$t-1$" trees are fixed, the regularization term for the previous $t-1$ trees can be treated as a constant. Therefore, the objective function can be rewritten as (5): Then, using Taylor expansion, we can derive (6) by removing the constant term. This leads to the optimization objective for the tth tree, which solely depends on the first and second derivatives of the loss function.

$$\begin{aligned}\mathrm{Obj}(\Theta) &= \sum\nolimits_{i=1}^{N} l\left(y_i, \widehat{y_i}^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) + \text{constant}\\ &= \sum\nolimits_{i=1}^{N}\left(g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)\right) + \Omega(f_t)\end{aligned} \tag{5}$$

$$\begin{aligned}\mathrm{Obj}(\Theta) &\simeq \sum\nolimits_{i=1}^{N}\left[l\left(y_i, \widehat{y_i}^{(t-1)}\right) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)\right] + \Omega(f_t)\\ &= \sum\nolimits_{i=1}^{N}\left(g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)\right) + \Omega(f_t)\end{aligned} \tag{6}$$

where $g_i = \frac{\partial l(y_i, \widehat{y_i}^{(t-1)})}{\partial \widehat{y_i}^{(t-1)}}, \;\; h_i = \frac{\partial^2 l(y_i, \widehat{y_i}^{(t-1)})}{\partial^2 \widehat{y_i}^{(t-1)}}$

The above section presents several important equations for XGBoost. For additional optimization details of XGBoost, please refer to the article by Chen and Guestrin [67]. One notable advantage of optimized XGBoost is its ability to handle large-scale datasets. It was designed with the intention of efficiently processing and learning from datasets that have a large number of samples and features. Its powerful learning and generalization capabilities make it suitable for real-world applications with complex relationships and large data volumes [26]. In addition, XGBoost provides a range of parameters and options that users can customize, allowing model fine-tuning to achieve optimal performance on specific tasks and datasets. The details and results of hyperparameter tuning for the XGBoost model in this study can be found in Supplementary Table S2.

C. Convolutional Neural Network

With the advancement of modern technology and the support of massive amounts of data and powerful computing capabilities, deep learning has emerged as a subfield of machine learning. Deep learning can be understood as various types of neural networks that mimic the perceptual process of the human brain, constructing hierarchical models to extract features layer by layer. CNNs are among the most important deep learning models and have been widely applied in quantitative remote sensing research in recent years [68], [69].

CNN models exhibit a well-defined layer structure, including convolutional layers, pooling layers, flattening layers, and fully connected layers [70]. In this study, a 1-D CNN model is constructed to estimate the clear-sky SLDR over polar regions. The model structure is depicted in Fig. 3.

Among the layer structures, the convolutional layer plays a crucial role because it performs convolution operations on the input data using a sliding window, followed by the application of activation functions for nonlinear feature extraction and propagation to the next layer. With this model, we constructed a total of 5 consecutive convolutional layers, applying the widely used rectified linear unit (ReLU) activation function. ReLU has several advantages, such as simplicity, high computational efficiency, and ease of use of gradient descent algorithms for optimization. Pooling layers, another important component of CNNs, aim to capture important features and reduce the spatial dimension of network parameters; this leads to feature compression, which not only decreases computational complexity but also mitigates the risk of overfitting. Common pooling operations include max pooling and average pooling. In our model construction, we employed max pooling, where the maximum value within each local region was selected as the output. Before entering the fully connected layers, the outputs from the convolutional or pooling layers need to be flattened into a 1-D representation. This flattening process, performed by the flattened layer, arranges all the elements of the matrix in sequential order, reducing the data dimensions for further processing by the fully connected layers. The fully connected layers in CNNs, also known as dense layers, are typically positioned toward the end of the CNN model and are responsible for generating the final regression or classification results. Each output neuron in the fully connected layers is connected to all the inputs from the previous layer, creating a fully connected network structure. In our model, the last layer is a fully connected layer with only one output neuron, representing the estimation of the SLDR.

After defining the CNN model structure, the training process can be divided into several main steps: 1) network initialization: the number of training epochs, batch size, learning rate, loss function, and other related hyperparameters are set. 2) Feedforward pass: the training samples are input into the network and passed through the layers to obtain the output values while calculating the loss functions at each layer. 3) Backpropagation pass: by calculating the difference between the predicted values of the model and the given actual values, the top-level loss is obtained. The loss gradient is subsequently propagated backward, and the weights and biases of the convolutional filters are optimized layer by layer via the gradient descent algorithm. 4) Iterate the above steps until the specified number of iterations is reached or the stopping criterion is satisfied. The important hyperparameters of the CNN model, along with their tuning ranges and Optimum value, can be found in Supplementary Table S3.

D. Transformer

The transformer model was proposed by Vaswani et al. [71] from Google with the primary objective of addressing challenges in natural language processing (NLP) and related tasks. Subsequently, significant advancements have been made in research applications such as image recognition, time-series forecasting, and change detection [61], [62], [72]. The transformer model stands out from traditional deep learning models such as CNNs and RNNs because it eliminates convolutional and recurrent operations and introduces the self-attention mechanism. This innovative mechanism enables the transformer model to effectively capture the relationships between input elements, leading to significant improvements in model performance.

The self-attention mechanism can be explained using (7). In this mechanism, the input feature matrix X is transformed into query matrix Q, key matrix K, and value matrix V by multiplying it by three weight matrices W^Q, W^K, and W^V. The key aspect is the calculation of similarity scores between all keys, which is achieved by taking the dot product of matrix Q and K. To ensure stable gradients, the scores are divided by the square root of the dimension of the key vectors (scaling factor $(d)_k^{1/2}$) and then normalized. The final step involves multiplying the normalized scores by the value matrix V to obtain the output of the self-attention layer

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{7}$$

where $Q = XW^Q$, $K = XW^K$, $V = XW^V$.

In this study, we developed a simple transformer model consisting of an encoder and a decoder, as illustrated in Fig. 4. The input data are first processed through the encoder layer, which comprises four sublayers: one self-attention layer, two feedforward layers with residual connections, and one fully connected feedforward network. The fully connected feed-forward network can be viewed as a basic multilayer perceptron (MLP) consisting of multiple fully connected layers and an activation function layer. In our model, we also chose the "ReLU" activation function. The output of the encoder layer serves as the input to the decoder layer. The structure of the decoder layer is similar to that of the encoder layer, but it contains only three sublayers. In the decoder layer, the output of the last fully connected layer in the feed-forward network is used as the prediction output. Although our transformer model has a relatively simple structure, the stacked encoder and decoder effectively enable the encoding and decoding of input features. The combination of sublayers allows the model to capture the relationships within the input data, which leads this transformer model to achieve accurate predictions in the task of SLDR estimation. In addition, detailed information on hyperparameter fine-tuning for the transformer model can be found in Supplementary Table S4.

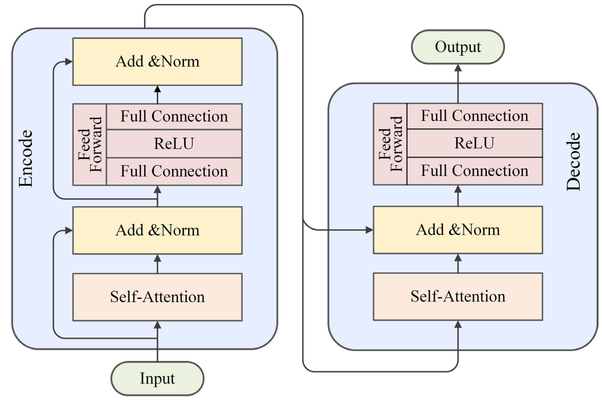

Fig. 4. Diagram of the constructed transformer model.

IV. Results

We randomly divided the collected data in Table I into two sets according to the site: two-thirds (34 sites) for training and one-third (17 sites) for validation. Three evaluation metrics are used to evaluate the model performance, namely, bias, RMSE, and determination coefficient (R^2). The formulae are shown as follows:

$$\text{Bias} = \frac{1}{N}\sum\nolimits_{i=1}^{N}(y_s - y_o) \tag{8}$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum\nolimits_{i=1}^{N}(y_s - y_o)^2} \tag{9}$$

$$R^2 = \frac{\sum_{i=1}^{N}(y_o)^2}{\sum_{i=1}^{N}(y_s)^2 - \sum_{i=1}^{N}(y_s - y_o)^2} \tag{10}$$

where y_s represents the model-retrieved SLDR value, y_o represents the observed SLDR, and N represents the number of samples. The bias reflects the mean difference between the simulated values and the observed true values and provides an assessment of the overall deviation of the model. The RMSE, on the other hand, is a metric that quantifies the difference between simulated values and observed true values. This approach provides an objective evaluation of the dispersion of the model's simulated values in relation to the true values. R^2 is a relative measure that indicates how well the model fits the true values, and it ranges between 0 and 1, where a value closer to 1 indicates a stronger relationship between the model's predictions and the true values.

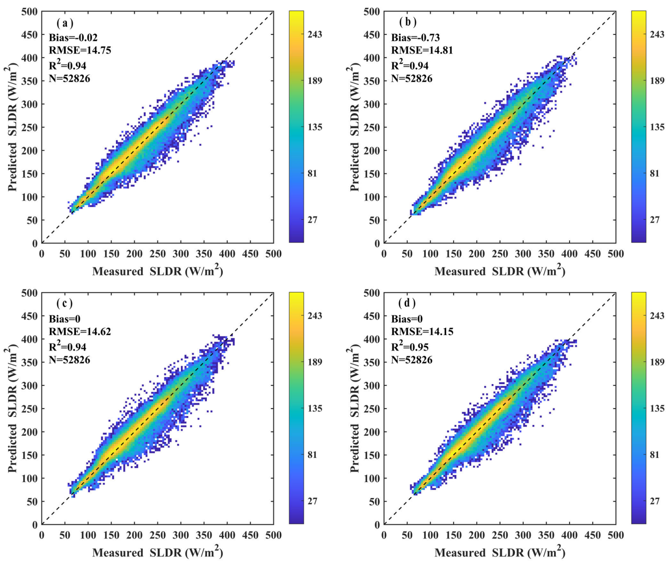

Fig. 5. Training results for (a) XGBoost, (b) CNN, (c) transformer, and (d) stacking.

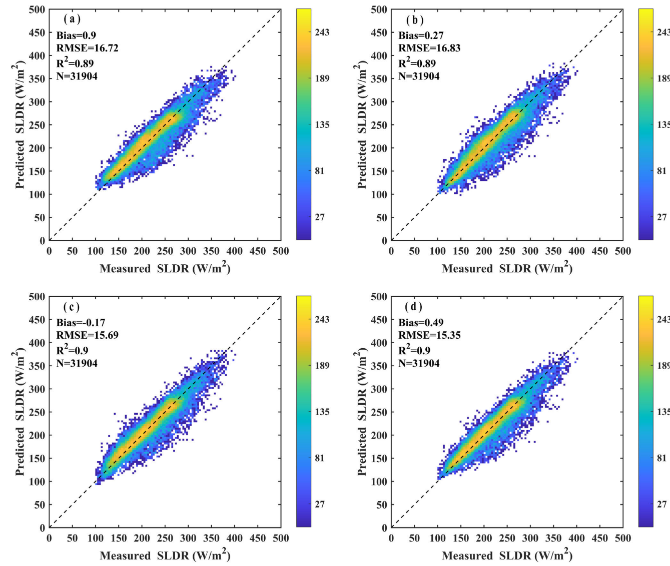

Fig. 6. Scatter density plots of the test results. (a) XGBoost model. (b) CNN model. (c) Transformer model. (d) Stacking model.

A. Training Results

In total, we obtained 52 826 samples from the training dataset. The training results are shown in Fig. 5. The majority of the data points are concentrated around the 1:1 line. The outcomes of the training phase indicate that all three fundamental models, XGBoost, CNN, and transformer, achieved commendable levels of accuracy. Specifically, the XGBoost model exhibited a bias of -0.02 W/m^2 and an RMSE of 14.75 W/m^2, with an R^2 of 0.94. The CNN model has a bias of -0.73 W/m^2 and an RMSE of 14.81 W/m^2, with an R^2 of 0.94. For the transformer model, the bias is 0 W/m^2, the RMSE is 14.62 W/m^2, and the R^2 is 0.94.

The three fundamental models have comparable performances, and stacking demonstrates a slight improvement. Compared with those of the three fundamental models, the RMSE of the proposed model decreases by 0.47–0.66 W/m^2 and the R^2 increases by 0.01. In addition, the scatter density plot of the stacking training results exhibits reduced dispersion, indicating an enhanced ability to handle outliers.

B. Validation Results

We extracted 31 904 samples from the 17 sites collected; these samples were not included in the model training. Fig. 6 shows the validation results. In terms of validation metrics, the transformer outperforms XGBoost and the CNN, which achieved a bias of -0.17 W/m^2, an RMSE of 15.69 W/m^2, and an R^2 of 0.9. In contrast, the XGBoost model and the CNN model had biases of 0.9 and 0.27 W/m^2, RMSEs of 16.72 and 16.83 W/m^2, and R^2 values of 0.89. The machine learning models XGBoost and the deep learning model CNN perform similarly, which implies that for the task of SLDR estimation in polar regions, convolutional models do not necessarily outperform GBDT models. In addition, despite the relatively simple architecture of the transformer model employed in this study, which is composed of only one encoder layer and one decoder layer, it still outperforms conventional machine learning and convolutional models during validation; this is because the self-attention mechanism within its sublayers effectively considers the interrelationships between input features, leading to more accurate SLDR estimations.

Subsequently, the predictions from these three fundamental models were stacked using the BMA model. The bias and RMSE were 0.49 and 15.35 W/m^2, respectively, with an R^2 of 0.9. In terms of the evaluation metrics, the stacking model exhibited significant improvements compared to the XGBoost and CNN models, performing comparably to the best-performing transformer model. In addition, the scatter density plot distribution revealed a reduction in dispersion after stacking, implying that this approach can enhance the reliability and robustness of the fundamental models during practical validation. In comparison to the training results, there is a slight decrease in the model's accuracy or performance. However, in general, the overall performance remains satisfactory, indicating that the models exhibit good generalizability.

C. Geographical Distribution of the Estimated SLDR

To further demonstrate the ability of the data-driven model to estimate practical SLDRs, the model was applied to real Aqua/MODIS imagery utilizing near-surface meteorological data from ERA5. Fig. 7 displays the estimated clear-sky SLDRs for areas above 60° north and south latitude on January 1, 2020 and July 1, 2020, respectively, and the blank areas indicate regions covered by clouds. The image resolution matches that of MODIS, with a consistent 1×1 km grid, providing finer details than products such as ERA5 and CERES. On January 1, the Northern Hemisphere was in winter, and the Southern Hemisphere was in summer. Therefore, the overall SLDRs in the Arctic region are lower than those in the Antarctic region. Conversely, on July 1st, when the Northern Hemisphere was in summer and the Southern Hemisphere was in winter, the overall SLDRs in the Arctic region were greater

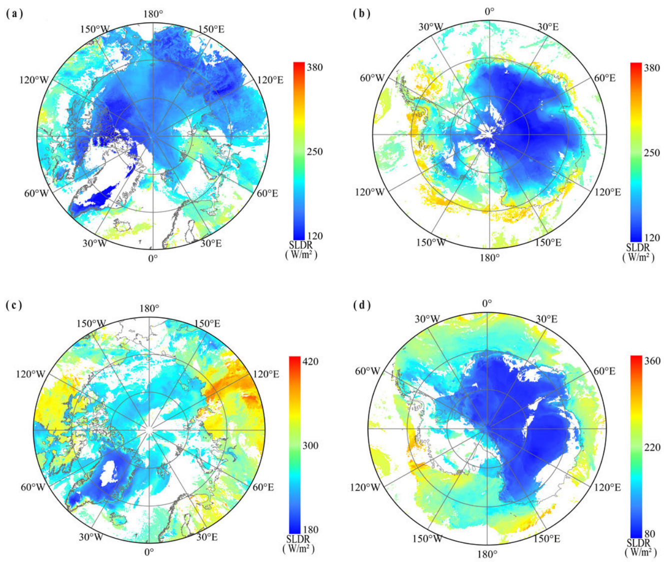

Fig. 7. Geographical distribution of the estimated clear-sky SLDRs in the Arctic and Antarctic regions. (a) Arctic, January 1, 2020. (b) Antarctic, January 1, 2020. (c) and (d) Spatial distributions of SLDRs estimated over the Arctic and Antarctic on July 1, 2020.

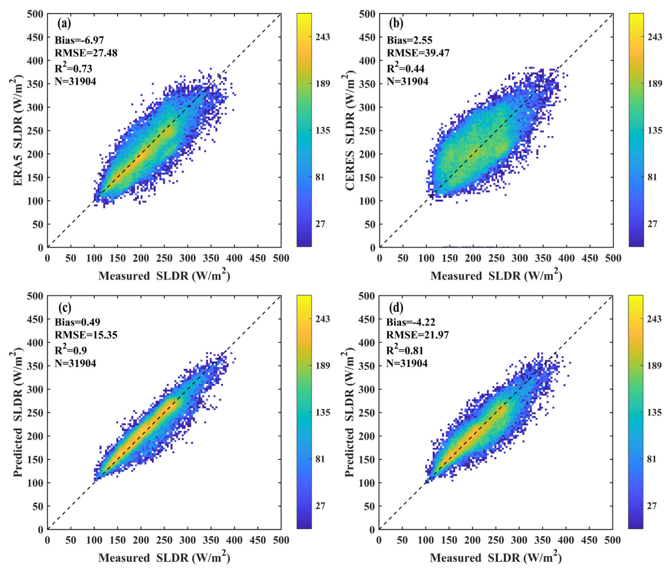

Fig. 8. Scatter density plots of SLDR product evaluation. (a) ERA5. (b) CERES-SYN. (c) Stacking. (d) Stacking_era5.

than those in the Antarctic region. Notably, the application of the data-driven model proposed in this study is not confined to surface types; it can simultaneously estimate the SLDR for both oceanic and terrestrial areas, which has significant practical implications for polar regions with extensive marine and sea ice distributions. The spatial distribution of clear-sky SLDRs in the Arctic and Antarctic regions reveals a general trend toward higher SLDRs in the Arctic and lower SLDRs in the Antarctic, with higher SLDRs in low-latitude areas and lower SLDRs in high-latitude areas. Moreover, over glacier surfaces, such as Greenland and the Antarctic continent, the SLDR tends to have lower values. Overall, the spatial distribution of clear-sky SLDR estimation results over polar regions aligns with the expected patterns. Furthermore, although the development and implementation of the data-driven model are based on MODIS sensor data, the model will not be restricted to only MODIS sensor data in the future. The model can be extended to other observation satellites equipped with thermal infrared sensors, such as GOES-16, Himawari-8, and FY-3, demonstrating its expansive potential for various applications.

V. Discussion

A. Comparison With the Existing SLDR Products

To further evaluate the performance of the proposed model, the accuracy of two widely used SLDR products was assessed. One is the ERA5 reanalysis data product provided by the European Centre for Medium-Range Weather Forecasts (ECMWFs) [18], and the other is the CERES-SYN global radiation data product provided by the Clouds and the Earth's Radiant Energy System (CERES) of the National Aeronautics and Space Administration (NASA) [73], [74]. Both products offer long-term time series of hourly SLDR data over polar regions. ERA5 utilizes the ECMWF Integrated Forecast System (IFS) CY41R2 with 4D-Var DA and model forecasts to provide hourly SLDR data with a spatial resolution of $0.25° \times 0.25°$ for polar regions. The CERES-SYN product, derived from the CERES sensors onboard the TERRA and AQUA satellites, applies improved SLDR simulation algorithms based on geostationary (GEO) infrared channel data and multichannel cloud property data. It offers high-accuracy hourly SLDR data with a spatial resolution of $1° \times 1°$.

ERA5 and CERES-SYN products provide cloud cover data and can identify clear sky conditions. To ensure the consistency of the comparison results, we used the MYD35 CM to identify the clear-sky ERA5 and CERES-SYN grid. Note that a scale effect issue exists between ERA5, CERES-SYN data, and in situ observations, which may introduce additional errors in the validation and intercomparison results. The footprint of the flux measurement typically ranges from tens to hundreds of meters, much smaller than the grid size of ERA5 and CERES-SYN products, inevitably introducing representativeness errors [75], [76]. Although a few researchers, like Wang et al. [77], have attempted to address this issue, no mature scheme is currently available to mitigate these representativeness errors in validating SLDR. Thus, we employed the nearest neighbor method to temporally and spatially match the products with the validation dataset in this study.

Fig. 8 presents the scatter density plots. The ERA5 SLDR product underestimates the clear-sky SLDR values in polar regions with a bias of -6.97 W/m^2, an RMSE of 27.48 W/m^2, and an R^2 of 0.73. The CERES product exhibits an overall minor bias, with a bias of 2.55 W/m^2. However, its RMSE is relatively high and reaches 39.47 W/m^2. Moreover, the bias and RMSE values of the proposed data-driven model are 0.49 and 15.35 W/m^2, respectively. This result demonstrated a clear advantage over the other two products in the polar regions.

We employed in situ measured near-surface meteorological data such as Ta, RH, and PS to establish the data-driven model, which ensures that our model will not excessively rely on specific meteorological data sources in later applications. To confirm the transferability of the stacking model and

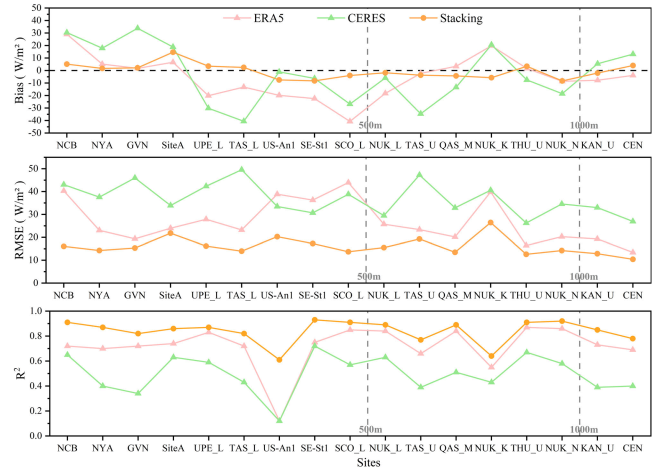

Fig. 9. Line charts of the validation results across different sites: analysis by bias, RMSE, and R^2.

investigate the potential errors caused by the scale effects in model input data when comparing the stacking model and ERA5 SLDR product, we conducted an additional experiment by substituting the site-observed meteorological data with near-surface 2 m temperature, 1000 hPa RH, and surface pressure (PS) data provided by ERA5 reanalysis products. Fig. 9(d) shows the evaluation results. The bias and RMSAE are -4.22 and 21.97 W/m^2, and R^2 is 0.81. The results indicate that although the performance of the stacking model with ERA5 input (stacking_era5) is slightly decreased, it still outperforms the two SLDR products, indicating the established Stacking model is robust and the impact of model input is weak.

In addition, to provide a more comprehensive analysis, we conducted a site-per-site comparison. The comparison results are shown in Fig. 9. The ERA5 and CERES-SYN SLDR products exhibit significant differences in validation accuracy across different sites, while the stacking model consistently demonstrates robust and stable validation accuracy at each site. For instance, at site NCB, ERA5, and CERES-SYN both exhibit significant overestimations, with biases of 29.3 and 30.39 W/m^2, respectively, whereas the stacking model has a bias of only 5.12 W/m^2. At site SCO_L, ERA5, and CERES-SYN both show substantial underestimations, with biases of -40.7 and -26.84 W/m^2, respectively, while the stacking model achieves a bias of only -3.98 W/m^2. Furthermore, even at sites where both products perform relatively well, the stacking model still exhibited better performance. For instance, at site CEN, which has an elevation of 1880 m, ERA5 has a bias and RMSE of -3.88 and 13.42 W/m^2, respectively; CERES-SYN has a bias and RMSE of 13.13 and 26.94 W/m^2. The stacking model has a bias and RMSE of 4.04 and 10.43 W/m^2, respectively. In summary, the stacking model surpasses existing products in terms of estimation accuracy, model stability, and robustness. The detailed site-specific validation results can be found in Supplementary Table S5.

B. Influence of Surface Elevation

The polar region has complex terrain and significant variations in surface elevation. Previous studies have indicated

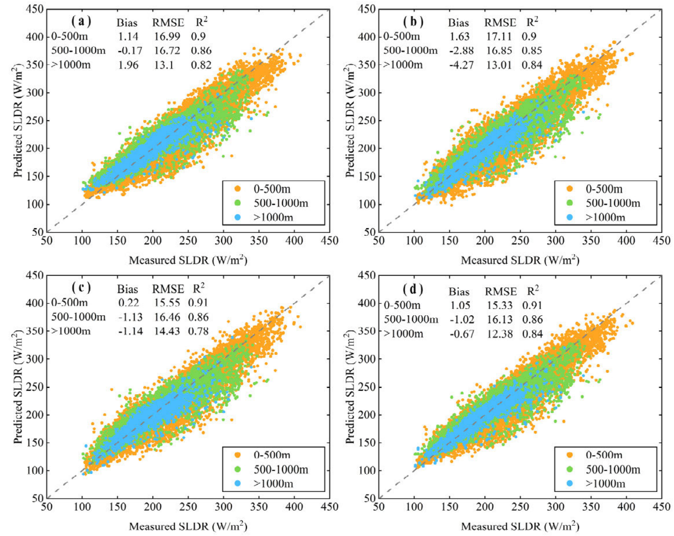

Fig. 10. Scatter plots of the SDLR estimation evaluation for the different elevation ranges. (a) XGBoost model. (b) CNN model. (c) Transformer model. (d) Stacking model.

that elevation can influence the accuracy of SLDR products [3], [24]. In the construction of the data-driven model, elevation was incorporated as part of the model's input features. To evaluate the impact of elevation on the accuracy of the data-driven model's SLDR estimation, we performed a detailed analysis.

The elevation of the validation sites ranged from 6 to 1880 m. We categorized the elevations into three subranges: 0, 500 m, 500, 1000 m, and 1000 m and above. The evaluation results are shown in Fig. 10. The accuracy of SLDR estimation by the three models exhibited variations with elevation fluctuations. With increasing elevation, the bias of the CNN model increased from 1.63 to -4.27 W/m^2, and R^2 decreased from 0.9 to 0.84. Similarly, the transformer model exhibited an increase in bias from 0.22 to -1.14 W/m^2 and a decrease in R^2 from 0.91 to 0.78 with increasing elevation. In contrast, the XGBoost model achieved the lowest bias in the 500–1000 m range, at -0.17 W/m^2, and the highest R^2 in the 0–500 m range, at 0.9. This indicates that due to factors such as atmospheric conditions and terrain features, the performance of individual models varies at different elevations.

By integrating the predictions of the three fundamental models, with increasing elevation, the bias values of the final stacking model decreased from 1.05 to -0.67 W/m^2, and the highest RMSE was achieved in the 500–1000 m range at 16.13 W/m^2, and the lowest RMSE was achieved in regions above 1000 m at 12.38 W/m^2, with R^2 values ranging from 0.91 to 0.84. This finding suggested that the stacking model exhibited improved RMSE and R^2 values compared to those of the other three fundamental models, revealing the individual weaknesses of the individual models. In addition, for elevations above 1000 m, the stacking model demonstrated the best SLDR estimation performance, with a bias of -0.67 W/m^2, an RMSE of 12.38 W/m^2, and an R^2 of 0.84. Compared to the individual model, the stacking model has successfully reduced RMSE, constrained bias, and ensured

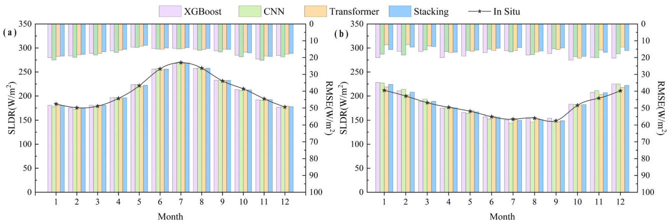

Fig. 11. Bar plots of the seasonal variation of the estimated SLDRs in polar regions generated by different models. (a) Arctic. (b) Antarctic.

the stability of estimation results. In other words, the stacking model effectively combines the strengths of individual models, enhancing its generalization and robustness.

C. Seasonal Variations

Fig. 11 illustrates the model performance during different seasons in both the northern and southern polar regions. The detailed SLDR monthly mean values and RMSE data can be found in Supplementary Tables S6 and S7. To better showcase the performance of the stacking model, the results from the three fundamental models are also presented in Fig. 11.

In the Arctic region, pronounced seasonal variations in the SLDR can be observed. In July, the highest average SLDR was observed, with site observations reaching 269.56 W/m^2. In February, the lowest average SLDR is observed, with a value of 175.61 W/m^2. Seasonal variations in the RMSE show that the proposed data-driven model performs slightly less effectively in winter. This could be attributed to the complex winter climate in the polar region, which includes phenomena such as polar nights and ice storms, leading to the occurrence of SLDR outliers, which are challenging for model estimation. Nevertheless, all three fundamental models established in the present study effectively capture the seasonal changes in the SLDR in the Arctic region, contributing to the excellent performance of the stacking model in the Arctic region.

In the Antarctic region, the seasonal variation in the SLDR is less obvious than that in the Arctic region, but it still follows the general pattern of higher SLDR values in summer and lower values in winter. In January, the highest observed monthly average SLDR is 211.81 W/m^2, while in September, the lowest observed monthly average SLDR is 148.85 W/m^2. The SLDRs estimated by the three fundamental models are influenced by seasonal changes. All three models tend to overestimate SLDRs during the summer in the Antarctic region, especially the CNN model. Stacking integration helps mitigate the extent of overestimation in the final SLDR estimations by the models. However, in general, the performance of the data-driven model in the Antarctic region is not as strong as its performance in the Arctic region, possibly due to the limited number of Antarctic sites used during model validation. There are fewer established sites in the Antarctic region, and some site data are not publicly available. The quality of the data provided by certain Antarctic sites is subpar. As a result, only data from one Antarctic site were used for validation in this study. In the next steps, we will continue to gather Antarctic data and expand the model's validation coverage.

VI. Conclusion

Currently, various reanalysis and remote sensing SLDR products are available at global and regional scales. However, these products often suffer from issues such as coarse resolution and poor accuracy over polar regions. Moreover, the complex surface environment in polar regions poses a challenge to existing clear-sky SLDR estimation methods. To address these challenges, this study develops a data-driven model for estimating the clear-sky SLDR over polar regions.

Based on the inner physics of the parametric and hybrid models and considering the unique characteristics of the polar region, 12 input parameters of the data-driven model were determined. With extensive data gathered from 51 polar sites, we trained and validated the clear-sky SLDR estimate data-driven model. The initial relationships between the SLDR and the input parameters were established using three machine learning models, namely, XGBoost, CNN, and transformer, from the data randomly selected from two-thirds of the sites. The final SLDR was estimated through the stacking model. Validation results from one-third of the site data showed that the proposed data-driven model significantly improved the estimation accuracy of clear-sky SLDRs in polar regions when compared to that of the ERA5 and CERES products. The proposed stacking model yields a bias of 0.49 W/m^2 and an RMSE of 15.34 W/m^2, with an R^2 of 0.9, whereas ERA5 and CERES yield biases of -6.97 and 2.55 W/m^2, RMSEs of 27.48 and 39.47 W/m^2, and R^2 values of 0.73 and 0.44, respectively. This suggests that the stacking model is a promising method for clear-sky SLDR estimation in polar regions.

In conclusion, this study reveals that data-driven models are effective approaches for addressing SLDR estimation challenges in polar regions. In particular, the successful application of the stacking model based on various machine learning models provides us with a new direction. In our next step, we will explore the implementation of the data-driven models and the stacking model for accurate cloud-sky SLDR estimation over polar regions.

References

[1] J. E. Sicart, J. W. Pomeroy, R. L. H. Essery, and D. Bewley, "Incoming longwave radiation to melting snow: Observations, sensitivity and estimation in northern environments," *Hydrolog. Processes*, vol. 20, no. 17, pp. 3697–3708, Nov. 2006.

[2] F. Carmona, R. Rivas, and V. Caselles, "Estimation of daytime downward longwave radiation under clear and cloudy skies conditions over a sub-humid region," *Theor. Appl. Climatol.*, vol. 115, nos. 1–2, pp. 281–295, Apr. 2013.

[3] J. Cheng, F. Yang, and Y. Guo, "A comparative study of bulk parameterization schemes for estimating cloudy-sky surface downward longwave radiation," *Remote Sens.*, vol. 11, no. 5, p. 528, Mar. 2019.

[4] J. Cheng and S. Liang, "Global estimates for high-spatial-resolution clear-sky land surface upwelling longwave radiation from MODIS data," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 7, pp. 4115–4129, Jul. 2016.

[5] R. P. Allan et al., "Changes in global net radiative imbalance 1985–2012," *Geophys. Res. Lett.*, vol. 41, no. 15, pp. 5588–5597, Aug. 2014.

[6] N. G. Loeb et al., "Observed changes in top-of-the-atmosphere radiation and upper-ocean heating consistent within uncertainty," *Nature Geosci.*, vol. 5, no. 2, pp. 110–113, Feb. 2012.

[7] Z. Wang, M. Zhang, L. Wang, and W. Qin, "A comprehensive research on the global all-sky surface solar radiation and its driving factors during 1980–2019," *Atmos. Res.*, vol. 265, Jan. 2022, Art. no. 105870.

[8] Z. Bo-Tao and Q. Jin, "Changes of weather and climate extremes in the IPCC AR6," *Adv. Climate Change Res.*, vol. 17, no. 6, p. 713, 2021.

[9] S. A. Robinson, "Climate change and extreme events are changing the biology of polar regions," *Global Change Biol.*, vol. 28, no. 20, pp. 5861–5864, Oct. 2022.

[10] K. Gao, A. Duan, D. Chen, and G. Wu, "Surface energy budget diagnosis reveals possible mechanism for the different warming rate among earth's three poles in recent decades," *Sci. Bull.*, vol. 64, no. 16, pp. 1140–1143, Aug. 2019.

[11] S. Lee, T. Gong, S. B. Feldstein, J. A. Screen, and I. Simmonds, "Revisiting the cause of the 1989–2009 Arctic surface warming using the surface energy budget: Downward infrared radiation dominates the surface fluxes," *Geophys. Res. Lett.*, vol. 44, no. 20, p. 10, Oct. 2017.

[12] R. L. Raddatz et al., "Downwelling longwave radiation and atmospheric winter states in the western maritime Arctic," *Int. J. Climatol.*, vol. 35, no. 9, pp. 2339–2351, Jul. 2015.

[13] A. Driemel et al., "Baseline surface radiation network (BSRN): Structure and data description (1992–2017)," *Earth Syst. Sci. Data*, vol. 10, no. 3, pp. 1491–1501, 1992.

[14] K. Tamagawa, M. Kitsuregawa, E. Ikoma, T. Ohta, S. Williams, and T. Koike, "An advanced quality control system for the CEOP/CAMP in-situ data management," *IEEE Syst. J.*, vol. 2, no. 3, pp. 406–413, Sep. 2008.

[15] K. A. Novick et al., "The AmeriFlux network: A coalition of the willing," *Agricult. Forest Meteorol.*, vol. 249, pp. 444–456, Feb. 2018.

[16] M. D. Palmer and D. J. McNeall, "Internal variability of earth's energy budget simulated by CMIP5 climate models," *Environ. Res. Lett.*, vol. 9, no. 3, Mar. 2014, Art. no. 034016.

[17] J. Xu et al., "Assessment of surface downward longwave radiation in CMIP6 with comparison to observations and CMIP5," *Atmos. Res.*, vol. 270, Jun. 2022, Art. no. 106056.

[18] H. Hersbach et al., "The ERA5 global reanalysis," *Quart. J. Roy. Meteorol. Soc.*, vol. 146, no. 730, pp. 1999–2049, Jul. 2020.

[19] C. Feng et al., "Comprehensive assessment of global atmospheric downward longwave radiation in the state-of-the-art reanalysis using satellite and flux tower observations," *Climate Dyn.*, vol. 60, nos. 5–6, pp. 1495–1521, Mar. 2023.

[20] S. Liang, K. Wang, X. Zhang, and M. Wild, "Review on estimation of land surface radiation and energy budgets from ground measurement, remote sensing and model simulations," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 3, no. 3, pp. 225–240, Sep. 2010.

[21] D. Schweizer and C. Gautier, "Validation of downwelling longwave computations with surface measurements during FIFE 89," *J. Geophys. Res., Atmos.*, vol. 100, no. D6, pp. 11569–11579, Jun. 1995.

[22] A. Viúdez-Mora, M. Costa-Surós, J. Calbó, and J. A. González, "Modeling atmospheric longwave radiation at the surface during overcast skies: The role of cloud base height," *J. Geophys. Res., Atmos.*, vol. 120, no. 1, pp. 199–214, Jan. 2015.

[23] B. Tang and Z.-L. Li, "Estimation of instantaneous net surface longwave radiation from MODIS cloud-free data," *Remote Sens. Environ.*, vol. 112, no. 9, pp. 3482–3492, Sep. 2008.

[24] W. Wang and S. Liang, "Estimation of high-spatial resolution clear-sky longwave downward and net radiation over land surfaces from MODIS data," *Remote Sens. Environ.*, vol. 113, no. 4, pp. 745–754, Apr. 2009.

[25] S. K. Gupta, D. P. Kratz, P. W. Stackhouse, A. C. Wilber, T. Zhang, and V. E. Sothcott, "Improvement of surface longwave flux algorithms used in CERES processing," *J. Appl. Meteorol. Climatol.*, vol. 49, no. 7, pp. 1579–1589, Jul. 2010.

[26] J. Cheng, Q. Zeng, and J. Shi, "A direct algorithm for estimating clear-sky surface longwave net radiation (SLNR) from MODIS imagery," *Int. J. Remote Sens.*, vol. 43, no. 5, pp. 1655–1683, Mar. 2022.

[27] Y. Guo, J. Cheng, and S. Liang, "Comprehensive assessment of parameterization methods for estimating clear-sky surface downward longwave radiation," *Theor. Appl. Climatol.*, vol. 135, nos. 3–4, pp. 1045–1058, Feb. 2019.

[28] W. C. Swinbank, "Long-wave radiation from clear skies," *Quart. J. Roy. Meteorolog. Soc.*, vol. 89, no. 381, pp. 339–348, 2007.

[29] S. B. Idso and R. D. Jackson, "Thermal radiation from the atmosphere," *J. Geophys. Res.*, vol. 74, no. 23, pp. 5397–5403, Oct. 1969.

[30] H. Wu, X. Zhang, S. Liang, H. Yang, and G. Zhou, "Estimation of clear-sky land surface longwave radiation from MODIS data products by merging multiple models," *J. Geophys. Res., Atmos.*, vol. 117, no. D22, Nov. 2012, Art. no. D22107.

[31] S. Yu, X. Xin, Q. Liu, H. Zhang, and L. Li, "An improved parameterization for retrieving clear-sky downward longwave radiation from satellite thermal infrared data," *Remote Sens.*, vol. 11, no. 4, p. 425, Feb. 2019.

[32] Y. Zhou, D. P. Kratz, A. C. Wilber, S. K. Gupta, and R. D. Cess, "An improved algorithm for retrieving surface downwelling longwave radiation from satellite measurements," *J. Geophys. Res., Atmos.*, vol. 112, no. D15, Aug. 2007, Art. no. D15102.

[33] Y. Lu et al., "Predicting surface solar radiation using a hybrid radiative transfer–machine learning model," *Renew. Sustain. Energy Rev.*, vol. 173, Mar. 2023, Art. no. 113105.

[34] S. Liang et al., "The global land surface satellite (GLASS) product suite," *Bull. Amer. Meteorolog. Soc.*, vol. 102, no. 2, pp. E323–E337, Feb. 2021.

[35] D. A. Rutan et al., "CERES synoptic product: Methodology and validation of surface radiant flux," *J. Atmos. Ocean. Technol.*, vol. 32, no. 6, pp. 1121–1143, Jun. 2015.

[36] P. W. Stackhouse et al., "GEWEX (Global energy and water exchanges Project): Surface radiation budget (SRB) release 4 integrated product (IP4)—Algorithm theoretical basis document and evaluation," NASA Langley Research Center, Hampton, VA, USA, Tech. Rep. 20210018620, 2021.

[37] Y. Zhang, W. B. Rossow, A. A. Lacis, V. Oinas, and M. I. Mishchenko, "Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data," *J. Geophys. Res., Atmos.*, vol. 109, no. D19, Oct. 2004, Art. no. D19105.

[38] X. Xin, S. Yu, D. Sun, H. Zhang, L. Li, and B. Zhong, "Assessment of three satellite-derived surface downward longwave radiation products in polar regions," *Atmosphere*, vol. 13, no. 10, p. 1602, Sep. 2022.

[39] J. Xu, S. Liang, H. Ma, and T. He, "Generating 5 km resolution 1981–2018 daily global land surface longwave radiation products from AVHRR shortwave and longwave observations using densely connected convolutional neural networks," *Remote Sens. Environ.*, vol. 280, Oct. 2022, Art. no. 113223.

[40] C. Feng et al., "Estimating surface downward longwave radiation using machine learning methods," *Atmosphere*, vol. 11, no. 11, p. 1147, Oct. 2020.

[41] Y. Cao, M. Li, and Y. Zhang, "Estimating the clear-sky longwave downward radiation in the Arctic from FengYun-3D MERSI-2 data," *Remote Sens.*, vol. 14, no. 3, p. 606, Jan. 2022.

[42] F. M. Lopes, E. Dutra, and I. F. Trigo, "Integrating reanalysis and satellite cloud information to estimate surface downward long-wave radiation," *Remote Sens.*, vol. 14, no. 7, p. 1704, Apr. 2022.

[43] D.-H. Kim and H. M. Kim, "Deep learning for downward longwave radiative flux forecasts in the Arctic," *Expert Syst. Appl.*, vol. 210, Dec. 2022, Art. no. 118547.

[44] J. Cheng, S. Liang, and J. Shi, "Impact of air temperature inversion on the clear-sky surface downward longwave radiation estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 4796–4802, Jul. 2020.

[45] S. K. Gupta, "A parameterization for longwave surface radiation from sun-synchronous satellite data," *J. Climate*, vol. 2, pp. 305–320, Apr. 1989.

[46] Y. Mizoguchi, A. Miyata, Y. Ohtani, R. Hirata, and S. Yuta, "A review of tower flux observation sites in Asia," *J. Forest Res.*, vol. 14, no. 1, pp. 1–9, Feb. 2009.

[47] R. Valentini, *Fluxes of Carbon, Water and Energy of European Forests*. Cham, Switzerland: Springer, 2003.

[48] D. Baldocchi et al., "FLUXNET: A new tool to study the temporal and spatial variability of ecosystem–scale carbon dioxide, water vapor, and energy flux densities," *Bull. Amer. Meteorolog. Soc.*, vol. 82, no. 11, pp. 2415–2434, Nov. 2001.

[49] G. Pastorello et al., "The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data," *Sci. Data*, vol. 7, no. 1, p. 225, Jul. 2020.

[50] R. S. Fausto et al., "Programme for monitoring of the Greenland ice sheet (PROMICE) automatic weather station data," *Earth Syst. Sci. Data*, vol. 13, no. 8, pp. 3819–3845, Aug. 2021.

[51] J. J. Danielson and D. B. Gesch, "Global multi-resolution terrain elevation data 2010 (GMTED2010)," U.S. Geological Surv., Earth Resour. Observ. Sci. (EROS) Center, Sioux Falls, SD, USA, Tech. Rep. 2011-1073, 2010.

[52] R. D. García, E. Cuevas, R. Ramos, V. E. Cachorro, A. Redondas, and J. A. Moreno-Ruiz, "Description of the baseline surface radiation network (BSRN) station at the Izaña observatory (2009–2017): Measurements and quality control/assurance procedures," *Geosci. Instrum., Methods Data Syst.*, vol. 8, no. 1, pp. 77–96, Feb. 2019.
[53] F. Zhu et al., "Integration of multisource data to estimate downward longwave radiation based on deep neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4103015.
[54] Y. Jia et al., "GNSS-R soil moisture retrieval based on a XGboost machine learning aided method: Performance and validation," *Remote Sens.*, vol. 11, no. 14, p. 1655, Jul. 2019.
[55] A. El Bilali, T. Abdeslam, N. Ayoub, H. Lamane, M. A. Ezzaouini, and A. Elbeltagi, "An interpretable machine learning approach based on DNN, SVR, extra tree, and XGBoost models for predicting daily pan evaporation," *J. Environ. Manage.*, vol. 327, Feb. 2023, Art. no. 116890.
[56] Y. Wang et al., "Surface shortwave net radiation estimation from Landsat TM/ETM+ data using four machine learning algorithms," *Remote Sens.*, vol. 11, no. 23, p. 2847, Nov. 2019.
[57] S. Wang et al., "Estimating land surface temperature from satellite passive microwave observations with the traditional neural network, deep belief network, and convolutional neural network," *Remote Sens.*, vol. 12, no. 17, p. 2691, Aug. 2020.
[58] H. Su, A. Wang, T. Zhang, T. Qin, X. Du, and X.-H. Yan, "Super-resolution of subsurface temperature field from remote sensing observations based on machine learning," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 102, Oct. 2021, Art. no. 102440.
[59] M. Krinitskiy, V. Koshkina, M. Borisov, N. Anikin, S. Gulev, and M. Artemeva, "Machine learning models for approximating downward short-wave radiation flux over the ocean from all-sky optical imagery based on DASIO dataset," *Remote Sens.*, vol. 15, no. 7, p. 1720, Mar. 2023.
[60] S. Ghimire, B. Bhandari, D. Casillas-Pérez, R. C. Deo, and S. Salcedo-Sanz, "Hybrid deep CNN-SVR algorithm for solar radiation prediction problems in Queensland, Australia," *Eng. Appl. Artif. Intell.*, vol. 112, Jun. 2022, Art. no. 104860.
[61] Q. Li, R. Zhong, X. Du, and Y. Du, "TransUNetCD: A hybrid transformer network for change detection in optical remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5622519.
[62] S. P. Simon, "Time series prediction of solar radiation using transformer neural networks," in *Proc. IEEE 19th India Council Int. Conf. (INDICON)*, Nov. 2022, pp. 1–5.
[63] B. Pavlyshenko, "Using stacking approaches for machine learning models," in *Proc. IEEE 2nd Int. Conf. Data Stream Mining Process. (DSMP)*, Aug. 2018, pp. 255–258.
[64] I. Syarif et al., "Application of bagging, boosting and stacking to intrusion detection," in *Proc. Int. Workshop Mach. Learn. Data Mining Pattern Recognit.*, 2012, pp. 593–602.
[65] S. Cui, Y. Yin, D. Wang, Z. Li, and Y. Wang, "A stacking-based ensemble learning method for earthquake casualty prediction," *Appl. Soft Comput.*, vol. 101, Mar. 2021, Art. no. 107038.
[66] N. Kardani, A. Zhou, M. Nazem, and S.-L. Shen, "Improved prediction of slope stability using a hybrid stacking ensemble method based on finite element analysis and field data," *J. Rock Mech. Geotech. Eng.*, vol. 13, no. 1, pp. 188–201, Feb. 2021.
[67] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. ACM SIGKDD*, 2016, pp. 785–794.
[68] T. Kattenborn, J. Leitloff, F. Schiefer, and S. Hinz, "Review on convolutional neural networks (CNN) in vegetation remote sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 173, pp. 24–49, Mar. 2021.
[69] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, "1D convolutional neural networks and applications: A survey," *Mech. Syst. Signal Process.*, vol. 151, Apr. 2021, Art. no. 107398.
[70] A. Ajit, K. Acharya, and A. Samanta, "A review of convolutional neural networks," in *Proc. Int. Conf. Emerg. Trends Inf. Technol. Eng.*, Feb. 2020, pp. 1–5.
[71] A. Vaswani et al., "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 5998–6008.
[72] F. Giuliari, I. Hasan, M. Cristani, and F. Galasso, "Transformer networks for trajectory forecasting," in *Proc. 25th Int. Conf. Pattern Recognit. (ICPR)*, Jan. 2021, pp. 10335–10342.
[73] D. R. Doelling et al., "Geostationary enhanced temporal interpolation for CERES flux products," *J. Atmos. Ocean. Technol.*, vol. 30, no. 6, pp. 1072–1090, Jun. 2013.
[74] D. R. Doelling et al., "Advances in geostationary-derived longwave fluxes for the CERES synoptic (SYN1deg) product," *J. Atmos. Ocean. Technol.*, vol. 33, no. 3, pp. 503–521, Mar. 2016.
[75] G. Huang, X. Li, C. Huang, S. Liu, Y. Ma, and H. Chen, "Representativeness errors of point-scale ground-based solar radiation measurements in the validation of remote sensing products," *Remote Sens. Environ.*, vol. 181, pp. 198–206, Aug. 2016.
[76] S. Gui, S. Liang, and L. Li, "Evaluation of satellite-estimated surface longwave radiation using ground-based observations," *J. Geophys. Res., Atmos.*, vol. 115, no. D18, Sep. 2010, Art. no. D18214.
[77] Z. Wang, M. Zhang, H. Li, L. Wang, W. Gong, and Y. Ma, "Bias correction and variability attribution analysis of surface solar radiation from MERRA-2 reanalysis," *Climate Dyn.*, vol. 61, nos. 11–12, pp. 5613–5628, Dec. 2023.

Mengfei Guo received the bachelor's degree in geographical information science from Hohai University, Nanjing, Jiangsu, China, in 2021. She is currently pursuing the master's degree with Beijing Normal University, Beijing, China.

Her research interests include thermal infrared quantitative remote sensing and surface radiation budget.

Jie Cheng (Senior Member, IEEE), photograph and biography not available at the time of publication.

Qi Zeng, photograph and biography not available at the time of publication.