Estimation and long-term trend analysis of surface solar radiation in Antarctica: A case study of Zhongshan Station

Zhaoliang ZENG[1,2], Zemin WANG[1], Minghu DING[2*], Xiangdong ZHENG[2], Xiaoyu SUN[3], Wei ZHU[4], Kongju ZHU[2], Jiachun AN[1], Lin ZANG[1], Jianping GUO[2], Baojun ZHANG[1*]

[1]Chinese Antarctic Center of Surveying and Mapping, Wuhan University, Wuhan 430079, China

[2]State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing 100081, China

[3]National Marine Environmental Forecasting Center, Beijing 100081, China

[4]Wuhan Economic & Technological Development Zone (Hannan) Meteorological Services Wuhan 430079, China

*Corresponding author: Minghu DING or Baojun ZHANG
Email: dingminghu@foxmail.com or bjzhang@whu.edu.cn

ABSTRACT

Long-term, ground-based daily global solar radiation (DGSR) at Zhongshan Station in Antarctica can quantitatively reveal the basic characteristics of Earth's surface radiation balance and validate satellite data for the Antarctic region. The fixed station was established in 1989 and conventional radiation observations started much later in 2008. In this study, a random forest (RF) model for estimating DGSR was developed using ground meteorological observation data, and a high-precision, long-term DGSR dataset was constructed. Then, the trend of DGSR from 1990 to 2019 at Zhongshan Station, Antarctica, was analyzed. The RF model, which performs better than other models, shows a desirable performance of DGSR hindcast estimation with an R^2 of 0.984, root-mean-square error of 1.377 MJ/m^2, and mean absolute error of 0.828 MJ/m^2. The trend of DGSR annual anomalies increased during 1990-2004 and then began to decrease after 2004. Note that the maximum value of annual anomalies occurred during approximately 2004/05 and is mainly related to the days with precipitation (especially those related to good weather during the polar day period) at this station. In addition to clouds and water vapor, bad weather conditions (such as snowfall, which can result in low visibility and then decreased sunshine duration and solar radiation) were the other major factors affecting solar radiation at this station. The high-precision, long-term estimated DGSR dataset enables us to further study and understand the role of Antarctica in global climate change and the interactions between snow, ice and atmosphere.

Keywords: meteorological variables, RF model, estimated historical DGSR,

long-term trend analysis

https://doi.org/10.1007/s00376-021-0386-6

Article Highlights:

- A thirty-year DGSR dataset, which was produced by combining in situ meteorological observation records with a random forest model, is presented.
- Among the considered models, the RF model shows the best performance for estimating historical DGSR with an R^2 of 0.984, root-mean-square error of 1.377 MJ/m^2, and mean absolute error of 0.828 MJ/m^2.
- The long-term DGSR trend generally increased during 1990-2004 and then began to decrease after 2004 at Zhongshan Station.
- In addition to clouds and water vapor, abnormal weather in Antarctica (such as fog, blowing snow and snowstorms) was also a major factor affecting solar radiation.

1. Introduction

Daily global solar radiation (DGSR, which includes direct and scattered radiation and refers to the total amount of downward shortwave radiation received by the surface each day) is the ultimate source of energy on Earth (Wild, 2009; Wild et al., 2005). Spatiotemporal variations in DGSR determine the climates and environments on the Earth's surface and drive the water, heat and carbon cycles of the Earth system (Che et al., 2005; Wang and Wild, 2016). Polar regions play a vital role in the Earth's surface radiation balance and climate system because there are many important, complex and interacting feedback mechanisms that closely bind the surfaces of polar regions to the global climate system (Bintanja, 1995; Braun and Hock, 2004; Park et al., 2013; Soares et al., 2019; Stanhill and Cohen, 1997). As a result, solar radiation observations and related research in the Antarctic region have received increasing attention (Choi et al., 2019; Ding et al., 2020; Garbe et al., 2020; Zhang et al., 2019).

To date, various types of data have been used to study the radiation balance in Antarctica, including reanalysis data, satellite data, and ground station data (Ding et al., 2020; Scott et al., 2017; Stanhill and Cohen, 1997; Yang et al., 2014; Zhang et al., 2016). Scott et al. (2017) used the Clouds and the Earth's Radiant Energy System (CERES) and CALIPSO-CloudSat-CERES-MODIS datasets to study the seasonal changes and spatial distribution of solar net radiation and cloud radiative forcing in southwestern Antarctica from only 2007 to 2010. Zhang et al. (2016) verified the DGSR of six reanalysis datasets by using surface stations around the Antarctic continent and the given deviation value of each reanalysis product. However, satellite-based and reanalysis-based radiation products often face the problem of deviation between the products and trusted ground observations (Jaross and Warner, 2008). Stanhill and Cohen (1997) summarized solar radiation measurements across Antarctica (from 12 stations containing 2 to 36 years of data) since 1957. However, the data from most stations used in the above study were very short-term and nonhomogeneous, and it is difficult to gain a full understanding of the long-term characteristics of DGSR in Antarctica (Lacelle et al., 2016; Stanhill and Cohen, 1997). In addition, ground-based DGSR observation sites are rare because the special geographical location and harsh natural environment of Antarctica seriously hinder the study of surface radiation balance (Aun et al., 2020).

In fact, beyond Antarctica, DGSR observation sites are similarly sparse and uneven across the world due to various problems, such as expensive instruments (He et al., 2018; He and Wang, 2020; Tang et al., 2013, 2011). However, meteorological variables (especially sunshine duration) interact with DGSR, and the number of conventional meteorological observation stations is greater than that of solar radiation stations (Tang et al., 2010; Zeng et al., 2020). Therefore, to obtain long-term, high-density, ground-based solar radiation products, many studies have estimated DGSR from conventional surface meteorological observations using traditional empirical formulas, physical models and machine learning methods (Chen et al., 2013; Huang et al., 2011; Jiang, 2009; Qin et al., 2011; Tang et al., 2018, 2013; Wang et al., 2016). However, these methods have rarely been used in the Antarctic.

The Chinese Antarctic Zhongshan Station (69°22'24.76"S, 76°22'14.28"E) is located on the southeast coast of Pritzker Bay in the Lasman Hills of East Antarctica (Figure 1a) (Ai et al., 2019; Chen et al., 2020). Its meteorological observation field is 15 m above sea level and approximately 300 m from the nearest coast (Dou et al., 2019; Yu et al., 2017). The local circulation (including valley breezes, land and sea breezes, and katabatic winds) is complicated due to the multifarious landforms and special geographical location. The climate at the station is characterized by low temperature, large temperature difference between winter and summer, low humidity and strong wind; it has the obvious characteristics of an Antarctic continental climate (Yu et al., 2017). Furthermore, as the station is located in front of the Antarctic inland ice sheet, katabatic winds are very obvious, so the DGSR in this area is affected by many factors (Ding et al., 2019). Note that the station was established in 1989, but the DGSR observations began in 2008. Therefore, we aimed to address the lack of radiation data and improve the understanding of Antarctic radiation and its response to global climate change. This study takes the Chinese Zhongshan Station in Antarctica as an example. Based on conventional ground meteorological observations and existing ground radiation observation data, DGSR is estimated using the optimal machine learning method, and then a long-term (~32 year) radiation dataset is obtained. The long-term trend of DGSR and the effects of clouds, water vapor, and visibility on DGSR were also analyzed in this study.

2. Data and methods

2.1 Meteorological observation data

Meteorological observations began in March 1989, and so far, 32 years of conventional ground meteorological observation data have been accumulated at Zhongshan Station, Antarctica. To ensure the accuracy and quality of observation data, the observation instruments and methods and the accuracy of the ground observation system have been operated in accordance with the ground meteorological observation standards of the China Meteorological Administration [Ground meteorological Observation Standards of China Meteorological Administration]. Before February 2002, instrumental observations were recorded manually four times per day (except sunshine duration, which was recorded hourly) at Zhongshan Station; after February 2002, the observation mode changed to an hourly wired telemetry automatic observation system. For more information on the meteorological variables and current sensor types, please see Table 1. All the sensor sampling intervals were changed to 1-min intervals, and the observation data were recorded 24 hours per day. During the instrument replacement period, parallel observations were performed for at least three months. After filtering the abnormal values, homogenization and quality control were also performed by the China Meteorological Administration. Then, the quality controlled and homogeneous observation data were used for data processing and model construction.

Previous work aims to establish a virtual network of DGSR datasets using the random forest (RF) model and high-density ground meteorological observations at ~2400 sites across China (Zeng et al., 2020). Due to the spatiotemporal heterogeneity of DGSR, dummy variables including day of year (DOY), latitude, longitude and altitude at all paired stations were also used as input variables in the prediction model. We used the RF model (as the optimal model) and calculated the relative importance of the variables in this study. This study aimed to construct a high-precision, long-term DGSR dataset for Zhongshan Station, Antarctica. Some dummy variables (latitude, longitude and altitude) and some meteorological variables (e.g., land surface temperature, due to the lack of observations at this station) were not used in this study. As suggested by previous studies (Wang et al., 2016; Zeng et al., 2020), meteorological elements that are highly associated with DGSR were selected as the input variables for the machine learning model. These variables include surface pressure (SP), relative humidity (RH), temperature (Tem), wind speed (WS), and sunshine duration (SSD) (corresponding short and full names are shown in Table 1). The dummy variables (i.e., DOY and month) are also used as input variables in the prediction model (similar to Zeng et al. (2020)). DGSR is affected by cloud cover, water vapor, and aerosols before reaching the surface (Che et al., 2005). However, aerosols over the Antarctic are relatively low, so they will not be discussed in this study. In addition, bad weather events (fog, snowfall, blowing snow and snowstorms) with low visibility are frequent at Zhongshan Station and also affect the DGSR. To further analyze the potential causal factors of the DGSR variations, ground vapor pressure (e, which is calculated by temperature and air pressure and represents water vapor content), cloud cover (CF), low cloud cover (LCF) and visibility (Vis) data are also used in this study (CF, LCF and Vis data were collected by manual visual observation).

Solar radiation observations started relatively late (March 2008) at Zhongshan Station, although short-term observations and research projects of solar radiation were performed during the periods of January-February 1990 (Wang and Xiong, 1991) and February 1993 to December 1994 (Bian et al., 1998). The solar radiation dataset for 2008 to 2020 was first used in this study. The observation site is located in the meteorological field north of the station (as shown in Figure 1b), where the surface is exposed rock from November to February, and there is snow for a short time in other periods, although there is usually almost no snow. A TBQ-2-B pyranometer (Figure 1c) was used to measure global solar radiation with a wavelength range of 0.3 to 3 μm and a resolution of hours at this station. The measured signal range of the TBQ-2-B pyranometer is 0-2000 W/m^2, the output signal is 0-20 mV, and the annual stability is ±2%. To ensure the accuracy of observation data, the TBQ-2-B pyranometer passed the verification and calibration of the China Meteorological Administration before installation.

2.2 Model development

2.2.1 Data processing and time matching

The meteorological observation data with a data quality code of 0 (passing all quality control checks) were extracted and time matched. The meteorological data observed at 00, 06, 12, and 18 hours (UTC) each day were then averaged to obtain the daily mean values. The DGSR and SSD were obtained as the sums of 24 hours per day. The final available data include conventional meteorological data from March 1989 to March 2020 and radiation observations from March 2008 to March 2020. Figure 3 shows the statistical properties (minimum, maximum, mean, and standard deviation) of the variables used for model training, testing and hindcast estimation at Zhongshan Station during 2010-2020. The DGSR (SSD) ranges from 0 (0) to 36.27 MJ/m^2 (21.70 hours), and the annual average value is 10.04 MJ/m^2 (4.83 hours). The annual average SP, Tem, RH, and WS are 985.08 hPa, -10.16 °C, 56.90%, and 6.49 m/s, respectively.

2.2.2 Model building

We evaluated the performances of machine learning models for estimating DGSR at Zhongshan Station, including RF (Chen et al., 2018), Light Gradient Boosting Machine (LightGBM) (Chen et al., 2019; Ke et al., 2017), decision tree (DT) (Quinlan, 1986), back propagation neural network (BPNN) (Wen et al., 2002), eXtreme Gradient Boosting (XGBoost) (Zelterman, 2015), support vector machine (SVM) (Cortes and Vapnik, 1995), multiple linear regression (MLR) (Zelterman, 2015) and Adaptive Boosting (AdaBoost) (Wang, 2012) models. The RF model is a widely used machine learning model that has a highly flexible algorithm and the capacity to analyze complex interactions of data classifications with noise or missing values (Chen et al., 2018). The RF model has also been used as a variable selection tool to select the input variables for a final model (Zeng et al., 2020). The RF model uses a bagging method to produce the training dataset. The out-of-the-bag (out-of-bag, OOB) data were used to evaluate the veracity of the regression predicted by the RF model. OOB estimation was an unbiased estimation when the number of trees was sufficient. In common statistical methods, overfitting occurs frequently when the methods have high degrees of freedom. In contrast to other models (e.g., the BPNN and SVM models), the RF model is an ensemble of random trees and basically has no issue of overfitting. LightGBM is a gradient promotion framework based on decision trees and can model complex nonlinear functions. LightGBM has the advantages of distributed and high performance in sorting, classification, and regression (Chen et al., 2019). The DT is a common and extensively researched solution to classification and prediction (Quinlan, 1986). The BPNN is a multilayer feedforward neural network based on a mathematical technique named Bayesian regularization to convert nonlinear regression into “well-posed” problems (Wen et al., 2002). The BPNN is composed of three layers: an input layer (first layer), hidden layer (middle layer), and output layer (last layer). XGBoost is a boosting algorithm with high performance for various regression and classification issues. The XGBoost method requires less training and time for prediction and can improve computing speed and accuracy (Gui et al., 2020). The SVM was developed by Vapnik-Chervonenkis dimension theory and structural analysis of the minimum risk principle. The SVM exhibits a unique advantage in dealing with small-sample problems, nonlinear cases, and high-dimensional pattern recognition problems by its kernel functions (Cortes and Vapnik, 1995). MLR is the regression analysis involving two or more independent variables. The MLR model can intuitively and quickly analyze correlations between multiple variables and dependent variables (Zelterman, 2015). Hence, MLR has been widely used in social science, economics, and technology. AdaBoost is an excellent boosting algorithm that combines multiple weak classifiers into a strong classifier. The main purpose of AdaBoost is to train different learning devices on the same training set and then combine these devices to construct a stronger final learning device (Wang, 2012).

The model with the best estimated performance will be the final model in this study. As in previous work (Zeng, et al., 2020), to obtain the optimal machine learning model, the results of 10-fold cross-validation (10-fold CV) were used to evaluate the model performances with different parameters (the final parameters of the models are shown in Table 2). In 10-fold CV, the matched pairs were partitioned into ten parts in equal proportion, with the first part as the testing subset and the other nine parts as the training subsets. This step was repeated ten times until every subset was tested, and the estimation results (mainly consisting of the coefficient of determination, R^2, root-mean-square error, RMSE, and mean absolute error, MAE) of the 10 parts were averaged and used as the accuracy of the final model. The accuracy indicators, including the R^2, RMSE, MAE, and difference (estimated minus observed), were used to assess the capabilities of the machine learning models and then obtain an optimal model (Gui et al., 2020; Zeng et al., 2020). In this study, we used the data from April 2010 to March 2020 for training, the 10-fold CV method for testing, and the period from April 2008 to March 2010 for evaluating the historic estimates.

2.2.3 Model application

The optimal model obtained from the above models was applied to estimate the DGSR using meteorological measurements recorded at Zhongshan Station from March 1989 to February 2020. The time variations in DGSR estimated by the optimal model were then compared with the observed DGSR. Finally, the long-term historical estimated DGSR at Zhongshan Station was analyzed, and then changes in the trend and the possible factors influencing these changes were further investigated.

2.3 Methods for DGSR trend analysis

Least squares regression has been applied to detect the linear trend in DGSR annual anomalies (Guo et al., 2017). Five-year running means of DGSR anomalies have been used to visually display the DGSR trend (Xue et al., 2019). In addition, the sliding trend analysis method has been used to help examine the time nodes of changes because trends often change with the span of the variable calculation period (Che et al., 2019; Gui et al., 2019). According to the method from Gui et al. (2019), we used Student's t tests to detect the robustness of each trend, and the criterion for statistical significance was set at the 95% confidence level. Since the estimated DGSR dataset and meteorological observation dataset have complete records of the whole year for each year from 1990 to 2019, the study period is set as 1990-2019 for the analyses of the monthly variations in meteorological variables and DGSR and the long-term changes in the DGSR trend.

3. Results and discussion

3.1 Validation and comparison of models

Figure 4a and 4b shows the scatterplots of the fitted model and 10-fold CV model results of the RF model from April 2010 to March 2020 at Zhongshan Station, Antarctica. Compared with those of other models, we found that the fitted and 10-fold CV results of the RF model have higher R^2 values of 0.997 and 0.988 and lower RMSE (MAE) values of 0.547 (1.189) MJ/m^2 and 0.278 (0.648) MJ/m^2, respectively. To assess the performance of the hindcast estimated by the RF model, the hindcast estimated results from April 2008 to March 2010 are shown in Figure 4c. We found that the hindcast estimated DGSR presented good consistency with the observed DGSR (R^2 = 0.984, RMSE = 1.377 MJ/m^2, and MAE=0.828 MJ/m^2). To further examine the hindcast performance of the DGSR estimated by the RF model, as an example, we selected historical estimated DGSR in 2009 (obtained from hindcast estimated results during April 2008-March 2010) for comparison with the ground observed DGSR. Figure 5a shows the daily time series (Figure 5b depicts the difference) of observed DGSR and estimated DGSR at Zhongshan Station. These two time series are highly consistent with each other, and the higher daily (monthly) mean difference values mainly occur in the summer, especially in the polar day period, but do not exceed ±5 (±0.85) MJ/m^2. Figure 5b also shows that approximately 96.7% (343 days in 2009) of the difference values (observed DGSR minus estimated DGSR) fell within the range of ±2 MJ/m^2. The results indicate that the DGSR estimated by the RF model closely fits the observed DGSR. Considering the CV results and the accuracy of the historic estimates, the RF model is highly recommended for DGSR estimation at Zhongshan Station in this study. The relative importance of the variables in the RF model is illustrated in Figure S1. As shown in Figure S1, SSD plays a dominant role in terms of the relative importance in the RF model and accounts for 60.3% of the overall importance. This result is consistent with previous studies showing that SSD is significantly correlated with DGSR (Wang et al., 2016; Zeng et al., 2020). The results also indicate two other dominant variables: DOY and Tem (accounting for 16.8% and 16.1% of the overall importance, respectively). These results suggest that DOY (seasonal effects) and Tem are also critical for DGSR estimation.

In addition to evaluating the performance of the RF model, we also evaluated the performances of other commonly used machine learning models for estimating DGSR at Zhongshan Station. From the results of the performance comparison in Table 3, the RF model performs better than the other models, and the LightGBM, XGBoost, BPNN, and gradient boosted regression tree (GBRT) models show similar historic estimation abilities, followed by those of the DT and AdBoost models. The SVM and MLR models have the worst performances. In common statistical methods, overfitting frequently occurs when the methods have high degrees of freedom. In contrast to other models, the RF model is an ensemble of random trees and has no issue of overfitting in this study.

Overall, these comprehensive results further confirm that the RF model has reliable performance in estimating historical DGSR. We can expect that it will be feasible to reconstruct the historical DGSR based on meteorological observation data and the RF model. Thus, the estimated historical dataset is used to accurately describe the comprehensive characteristics and changes in the long-term trend of DGSR. Therefore, we mainly use the DGSR estimated by the RF model for 1989 to 2020 in the following trend analysis.

3.2 Monthly and annual variations in DGSR

Before analyzing the changes in the DGSR trend, we calculated the monthly and annual characteristics of DGSR estimated by the RF model and the corresponding meteorological variables over the period of 1990-2019 (as shown in Figure 6). The estimated DGSR shows significant monthly and seasonal changes, in which the DGSR during the half-year of summer (October to March) is significantly higher than that during the half-year of winter (April to September), and the higher DGSR values, up to 30 MJ/m^2, are mainly during the polar day period, while the DGSR values are 0 MJ/m^2 during the polar night period. The monthly average humidity varies from 56% to 61% from 1990 to 2019 and is higher in summer than in other seasons. The monthly mean temperature varies from -2.90 to 0.24 ℃ in summer and from -14.74 to -16.12 ℃ in winter. The monthly average temperature is the highest in January (0.24 ℃) and the lowest in July (-16.18 ℃). Variations in surface pressure at Zhongshan Station are characterized by half-year waves. From January to June, there are periods of high pressure, and other months have periods of low pressure. The highest (lowest) monthly average surface pressure occurs in June (October) and is greater than 988 hPa (lower than 981 hPa). The wind speed is the highest in winter, followed by that in autumn, spring and summer. The highest (lowest) monthly average wind speed is 7.94 m/s (5.01 m/s) in August (January).

Zhongshan Station is located on the edge of the Antarctic continent and is near the Antarctic ice sheet. The winds at this station are mainly affected by a combination of the easterly airflow in the northern part of the Antarctic continent, polar cyclones and katabatic winds. Easterly winds prevail over this station year round. In summer (December, January and February), both the Antarctic continental cold high-pressure system and the circumpolar low-pressure zone are weaker, and the smaller pressure gradient between these two synoptic systems induces a lower wind speed. In winter (April, May and June), the prevailing Antarctic continental high-pressure system strengthens, and the circumpolar low-pressure zone moves southward, which causes a larger pressure gradient and thus results in a higher wind speed. Additionally, the monthly variations in air pressure at Zhongshan Station are closely related to these two synoptic systems. Zhongshan Station is covered by snow and ice and the air has lower relative humidity in winter, while the snow and sea ice around this sation melt in summer and thus cause increased saturation of water vapor in the air and higher relative humidity. It is noted that the long sunshine duration and strong solar radiation in summer play key roles in the temperature increase at this station.

The monthly variation in SSD is basically consistent with that in DGSR, and the R^2 between SSD and DGSR is 0.88, indicating that SSD is the main input variable for DGSR in the machine learning models. Figures S2 and S3 show that the yearly average value of the estimated DGSR coincides with that of the measured DGSR from 2009 to 2019 (and also in 1994), and the differences (estimated DGSR - observed DGSR) were mainly distributed between -0.1 and 0.1 MJ/m^2. Furthermore, the annual changes in the estimated DGSR and observed SSD trends are highly consistent, suggesting that historical DGSR estimated by the RF model has high accuracy for further analysis (such as annual anomaly trends and sliding trends).

3.3 DGSR trend

Note the continuous presence and absence of solar radiation during summer and winter, respectively, because the peculiar conditions of polar day (beginning on November 23 and ending on January 21 of the following year at Zhongshan Station) and polar night (beginning on May 27 and ending on July 18 each year at this station) occur in Antarctica. In addition, the DGSR values were 0 MJ/m^2 during the polar night period. Therefore, the DGSR trend analysis was divided into the following three scenarios: all days of the year (annual, scenario 1), the polar day period (scenario 2), and the period of all days of the year except for polar day and polar night (scenario 3). Figure 7a to Figure 7c shows the DGSR anomalies and their trends for the three scenarios. It can be seen that the trend of the DGSR annual anomalies increased during 1990-2004 and then began to decrease after 2004. However, obvious differences in anomaly values exist among the three scenarios. The variation in DGSR anomalies is the greatest during the polar day period (increasing linear trend of 0.175 MJ/m^2/year and decreasing linear trend of -0.101 MJ/m^2/year, which are significant at the 95% confidence level), followed by that during all days of the year (increasing linear trend of 0.039 MJ/m^2/year and decreasing linear trend of -0.025 MJ/m^2/year, which are significant at the 95% confidence level), and is lowest during the period of all days of the year except for polar day and polar night (increasing linear trend of 0.011 MJ/m^2/year and decreasing linear trend of -0.001 MJ/m^2/year, which are not statistically significant). In general the trend of DGSR is similar to Europe (except for China, with a decreasing trend between 1990 and 2000) during the period of brightening, slightly ascend to the early 2005s, after which it shows a decrease to the present (Che et al., 2005; Ohmura, 2009).

The DGSR is strongest during the polar day period, accounting for approximately 43% of the annual global solar radiation. The surface of Earth receives more solar radiation because the sun always stays above the horizon during the polar day period.

In 1990-2019, the average DGSR during the polar day period varies between 26 MJ/m^2 and 31 MJ/m^2. However, the annual average DGSR varies from 9.6 MJ/m^2 to 10.8 MJ/m^2. Therefore, the anomaly values and the range of variation during the polar day period are higher than those during all days of the year. The anomaly values during the period of all days of the year except for polar day and polar night alternate between positive and negative, indicating that the change in total solar radiation during this period basically has no obvious trend. The maximum value of the annual anomalies occurred during approximately 2004/05 and is mainly related to the days with precipitation (such as snowfall, which can result in low visibility and then decreased sunshine duration and solar radiation) at Zhongshan Station in Antarctica. In contrast to scenario 3 (all days of the year except for polar day and polar night), scenario 2 (the polar day period) and scenario 1 (annual) have similar trends in the anomalies (and the sunshine duration during the polar day period accounts for approximately 45% of the total sunshine duration of each year), indicating that the changes in sunshine duration and DGSR during the polar day period play a leading role in the changes in the trend of the DGSR annual anomalies.

Based on Student's t tests, the sliding trends of DGSR for all situations are shown in Figure 7d to Figure 7f to present a more comprehensive analysis of the annual trends. Sliding trends were calculated for the three scenarios, starting in each year from 1990 to 2015 and ending in 2019 with increments of at least 5 years. As shown in scenario 1 (all days of the year), the trend of DGSR increased from 1990 to 2003 (although an opposite trend was found during approximately 1993 and 1995), then declined sharply after 2004 (especially in scenario 2: the polar day period) and slightly increased in 2012. The sliding trends for scenario 3 (all days of the year except for polar day and polar night) were smaller and relatively stable compared with those for the other scenarios. When the running mean window was longer than 15 years, the DGSR trends first increased and then decreased in scenario 3 (all days of the year except for polar day and polar night), and most of the trends were statistically significant.

3.4 The potential impact factors of DGSR

DGSR is affected by cloud cover, water vapor, and aerosols before reaching the surface (Che et al., 2005). However, aerosols are relatively low over the Antarctic, so they will not be discussed in this study. In contrast, bad weather events (fog, snowfall, blowing snow and snowstorms) with low visibility are frequent at Antarctic Zhongshan Station and will also affect the DGSR. Therefore, the effects of cloud fraction, low cloud fraction, ground vapor pressure (e, which represents atmospheric water vapor content), and visibility (which represents bad weather events) on the DGSR at this station are further detailed and explicitly analyzed. Note that the solar radiation is greatly affected by the solar altitude angle in the polar region (which has the phenomena of polar day and polar night). To avoid the effects of these phenomena and analyze the influence of potential factors on solar radiation, we selected only the matched samples of April and September each year for discussion in this study. Here, all the matched samples were divided into five subsets according to cloud cover

(0-20%, 20-40%, 40-60%, 60-80% and 80-100%), low cloud cover (0-20%, 20-40%, 40-60%, 60-80% and 80-100%) and ground vapor pressure (0-1 hPa, 1-2 hPa, 2-3 hPa, 3-4 hPa and >4 hPa), and then the average DGSR was calculated for each subset, and the results are shown in Figure 8a, 8b and 8c, respectively. Similarly, the matched samples were divided into six subsets according to visibility (0-5 km, 5-10 km, 10-15 km, 15-20 km, 20-25 km and >25 km), then the average DGSR was calculated for each subset (see Figure 8d for the results).

We found that the DGSR significantly decreased (significantly increased) with increasing cloud fraction and ground vapor pressure (visibility). This is because under cloudy conditions, solar radiation reflects back to the top of the atmosphere, reducing the amount of solar radiation reaching the Earth's surface. However, the DGSR did not change much when the low cloud fraction was greater than 40%. We also found that DGSR was generally low during severe weather with low visibility but significantly higher under high visibility conditions. Overall, the DGSR decreases with increasing cloud cover, low cloud cover and ground vapor pressure but increases with increasing visibility. The change in DGSR with low cloud cover is not significant with the change in cloud cover, ground vapor pressure and visibility. Meanwhile, the times series of DGSR, LCF, CF, e, Vis, and SSD and their long-term trends are examined in Fig. S4, respectively. The DGSR shows a small upward trend. Accordingly, SSD and Vis exhibit upward trends (CF, LCF and e exhibit downward trends), implying that the cloud cover, water vapor and abnormal weather (except as solar altitude angle in the polar region) are the common factors that influence the trend of DGSR at this station. To sum up, clouds and water vapor in the atmosphere are the main factors affecting solar radiation. Bad weather conditions, such as fog, blowing snow and snowstorms, are also a major factor affecting solar radiation at Zhongshan Station, Antarctica.

4. Concluding remarks

Based on ground meteorological observation data, an RF model was developed to estimate DGSR, and a high-precision, long-term DGSR dataset was constructed for 1989 to 2020 at Zhongshan Station, Antarctica. Long-term trends and the potential impact factors of DGSR were then analyzed in this study. Compared with those of other models, we found that the fitted and 10-fold CV results of the RF model have higher R^2 values and lower RMSE and MAE, and the hindcast estimated DGSR presents good consistency with the observed DGSR (R^2 = 0.984, RMSE = 1.377 MJ/m^2, and MAE=0.828 MJ/m^2). The RF model is better than other models for reconstructing the historical DGSR based on the meteorological observations in this study. The DGSR trends were very consistent in all situations, and DGSR generally increased during 1989-2004 and then began to decrease after 2004. The sliding trend of DGSR in the all days of the year except for polar night period and the polar day period increased from 1990 to 2003 (although an opposite trend was found during approximately 1993 and 1995 for the all days of the year except for polar night period), then declined sharply after 2004 and slightly increased in 2012, while the sliding trends for the period of all days of the year except for polar day and polar night were smaller and relatively stable. The DGSR decreases with increasing cloud cover, low cloud cover and ground vapor pressure but increases with increasing visibility. The results show that clouds and water vapor are the main factors affecting solar radiation in Antarctica. Meanwhile, bad weather conditions, such as fog, blowing snow and snowstorms, are also a major factor affecting solar radiation at Zhongshan Station, Antarctica. Based on the DGSR estimation method in this study, our plan for future work can be divided into two parts: the first part is to construct a virtual DGSR observation network across the Antarctic region, and the second part is to reconstruct historical site-scale DGSR concentrations through this newly constructed virtual DGSR observation network. It is worth noting that some sites in remote areas of Antarctica lack DGSR datasets. Therefore, these high-precision, long-term DGSR datasets can be used to study the radiation balance and the ultimate source of solar energy in Antarctica.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (41941010 and 41771064), the National Basic Research Program of China (2016YFC1400303), and the Basic Fund of the Chinese Academy of Meteorological Sciences (2018Z001). We greatly appreciate the help from the Polar Research Institute of China and the Antarctic expeditioners at the Chinese Zhongshan Station. The long-term (March 1989-February 2020) estimated DGSR dataset can be acquired in the Mendeley data repository (data identification number: http://dx.doi.org/10.17632/2y2mmhzvcx.1. Direct URL to data: https://data.mendeley.com/2y2mmhzvcx).

REFERENCES

Ai, S., S. Wang, Y. Li, G. Moholdt, C. Zhou, L. Liu, and Y. Yang, 2019: High-precision ice-flow velocities from ground observations on Dalk Glacier, Antarctica. *Polar Sci.,* **19**, 13-23, https://doi.org/10.1016/j.polar.2018.09.003.

Aun, M., K. Lakkala, R. Sanchez, E. Asmi, F. Nollas, O. Meinander, L. Sogacheva, V. D. Bock, A. Arola, and G. d. Leeuw, 2020: Solar UV radiation measurements in Marambio, Antarctica, during years 2017-2019. *Atmos. Chem. Phys.,* **20**, 6037-6054, https://doi.org/10.5194/acp-20-6037-2020.

Bian, L., and Coauthors, 1998: A study of radiative features at the Great Wall and Zhongshan Stations of Antarctica, Quarterly journal of applied meteorology, **9**, 160-168. (in Chinese)

Bintanja, R., 1995: The Local Surface Energy Balance of the Ecology Glacier, King George Island, Antarctica: Measurements and modelling. *Antarct. Sci.,* **7** (03), 315-325, https://doi.org/10.1017/S0954102095000435.

Braun, M., and R. Hock, 2004: Spatially distributed surface energy balance and ablation modelling on the ice cap of King George Island (Antarctica), *Glob. Planet. Change,* **42**, 45-58, https://doi.org/10.1016/j.gloplacha.2003.11.010.

Che, H., Gui, K., Xia, X., Wang, Y., Holben, B.N., Goloub, P., Cuevas-Agulló, E., Wang, H., Zheng, Y., Zhao, H., Zhang, X., 2019. Large contribution of meteorological factors to inter-decadal changes in regional aerosol optical depth. *Atmos. Chem. Phys.,* **19**, 10497-10523 https://doi.org/10.5194/acp-2019-360.

Che, H., and Coauthors, 2005: Analysis of 40 years of solar radiation data from China, 1961-2000. *Geophys. Res. Lett.,* **32**, 1–5.

https://doi.org/10.1029/2004GL022322.

Chen, C., Q. Zhang, Q. Ma, and B. Yu, 2019: LightGBM-PPI: Predicting protein-protein interactions through LightGBM with multi-information fusion. *Chemom. Intell. Lab. Syst.,* **191**, 54-64, https://doi.org/10.1016/j.chemolab.2019.06.003.

Chen, G., and Coauthors, 2018: Spatiotemporal patterns of PM10 concentrations over China during 2005–2016: A satellite-based estimation using the random forests approach. *Environ. Pollut.,* **242**, 605–613. https://doi.org/10.1016/j.envpol.2018.07.012.

Chen, J.-L., G.-S. Li, and S.-J. Wu, 2013: Assessing the potential of support vector machine for estimating daily solar radiation using sunshine duration. *Energy Convers. Manag.,* **75**, 311-318, https://doi.org/10.1016/j.enconman.2013.06.034.

Chen, Y., C. Zhou, S. Ai, Q. Liang, L. Zheng, R. Liu, and H. Lei, 2020: Dynamics of Dalk Glacier in East Antarctica Derived from Multisource Satellite Observations Since 2000. *Remote Sens.,* **12**, 1809, https://doi.org/10.3390/rs12111809.

Choi, T. i., S. J. Kim, J. H. Kim, H. Kwon, and M. A. Lazzara, 2019: Characteristics of Surface Meteorology at Lindsey Islands, Amundsen Sea, West Antarctica. *J. Geophys. Res. Atmos.,* 124, 6294–6306. https://doi.org/10.1029/2018JD029556

Cortes, C., and V. Vapnik, 1995: Support-vector networks. *Mach. Learn.,* **20**, 273-297, https://doi.org/10.1023/A:1022627411411.

Ding, M., D. Yang, M. R. van den Broeke, I. Allison, C. Xiao, D. Qin, and B. Huai, 2020: The Surface Energy Balance at Panda 1 Station, Princess Elizabeth Land:

A Typical Katabatic Wind Region in East Antarctica. *J. Geophys. Res. Atmos.,* **125**. https://doi.org/10.1029/2019JD030378.

Dou, Y., G. Zuo, X. Chang, and Y. Chen, 2019: A Study of a Standalone Renewable Energy System of the Chinese Zhongshan Station in Antarctica. *Appl. Sci.,* 9,1968.

Garbe, J., T. Albrecht, A. Levermann, J. F. Donges, and R. Winkelmann, 2020: The hysteresis of the Antarctic Ice Sheet. *Nature,* **585**, 538-544. https://doi.org/10.1038/s41586-020-2727-5.

Gui, K., and Coauthors, 2019: Satellite-derived PM2.5 concentration trends over Eastern China from 1998 to 2016: Relationships to emissions and meteorological parameters. *Environ. Pollut.,* **247**, 1125–1133. https://doi.org/10.1016/j.envpol.2019.01.056.

Gui, K., and Coauthors, 2020: Construction of a virtual PM2.5 observation network in China based on high-density surface meteorological observations using the Extreme Gradient Boosting model. *Environ. Int.,* **141**, 105801. https://doi.org/10.1016/j.envint.2020.105801.

Guo, J., and Coauthors, 2017: Declining frequency of summertime local-scale precipitation over eastern China from 1970 to 2010 and its potential link to aerosols. *Geophys. Res. Lett.,* **44**, 5700–5708. https://doi.org/10.1002/2017GL073533.

He, Y., and K. Wang, 2020. Variability in Direct and Diffuse Solar Radiation Across China From 1958 to 2017. *Geophys. Res. Lett.,* **47**.

https://doi.org/10.1029/2019GL084570.

He, Y., K. Wang, C. Zhou, and M. Wild, 2018: A Revisit of Global Dimming and Brightening Based on the Sunshine Duration. *Geophys. Res. Lett.,* **45**, 4281–4289. https://doi.org/10.1029/2018GL077424.

Huang, G., M. Ma, S. Liang, S. Liu, and X. Li, 2011: A LUT-based approach to estimate surface solar irradiance by combining MODIS and MTSAT data. *J. Geophys. Res. Atmos.,* **116** (D22), https://doi.org/10.1029/2011JD016120.

Jaross, G., and J. Warner, 2008: Use of Antarctica for validating reflected solar radiation measured by satellite sensors. *J. Geophys. Res. Atmos.,* **113**, https://doi.org/10.1029/2007JD008835.

Jiang, Y., 2009: Computation of monthly mean daily global solar radiation in China using artificial neural networks and comparison with other empirical models. *Energy,* **34**, 1276-1283, https://doi.org/10.1016/j.energy.2009.05.009.

Ke, G., and Coauthors, 2017: LightGBM: A highly efficient gradient boosting decision tree, in: *Advances in Neural Information Processing Systems,* 3146-3154.

Lacelle, D., C. Lapalme, A. F. Davila, W. Pollard, M. Marinova, J. Heldmann, and C. P. McKay, 2016: Solar Radiation and Air and Ground Temperature Relations in the Cold and Hyper-Arid Quartermain Mountains, McMurdo Dry Valleys of Antarctica. *Permafr. Periglac. Process,* **27**, 163-176, https://doi.org/10.1002/ppp.1859.

Liu, J., F. Weng, Z. Li, and M. C. Cribb, 2019: Hourly PM 2.5 Estimates from a

Geostationary Satellite Based on an Ensemble Learning Algorithm and Their Spatiotemporal Patterns over Central East China. *Remote Sens.,* **11**, 2120.

Ohmura, A., 2009: Observed decadal variations in surface solar radiation and their causes. *J. Geophys. Res. Atmos.*, **114,** https://doi.org/10.1029/2008JD011290, 2009.

Park, S.-J., T.-J. Choi, and S.-J. Kim, 2013: Heat flux variations over sea ice observed at the coastal area of the sejong station, Antarctica. *Asia-Pacific J. Atmos. Sci.,* **49**, 443-450, https://doi.org/10.1007/s13143-013-0040-z.

Qin, J., Z. Chen, K. Yang, S. Liang, and W. Tang, 2011: Estimation of monthly-mean daily global solar radiation based on MODIS and TRMM products. *Appl. Energy,* **88**, 2480-2489, https://doi.org/10.1016/j.apenergy.2011.01.018.

Scott, R. C., D. Lubin, A. M. Vogelmann, and S. Kato, 2017: West antarctic ice sheet cloud cover and surface radiation budget from NASA A-Train satellites. *J. Clim.,* **30**, 6151-6170, https://doi.org/10.1175/JCLI-D-16-0644.1.

Soares, J., M. Alves, F. N. D. Ribeiro, and G. Codato, 2019: Surface radiation balance and weather conditions on a non-glaciated coastal area in the Antarctic region. *Polar Sci.,* **20**, 117-128, https://doi.org/10.1016/j.polar.2019.04.001.

Stanhill, G., and S. Cohen, 1997: Recent changes in solar irradiance in Antarctica. *J. Clim.,* **10**, 2078-2086, https://doi.org/10.1175/1520-0442(1997)010<2078:RCISII>2.0.CO;2.

Tang, W., K. Yang, J. He, and J. Qin, 2010: Quality control and estimation of global solar radiation in China. *Sol. Energy.,* **84**, 466-475, https://doi.org/10.1016/j.solener.2010.01.006.

Tang, W., K. Yang, J. Qin, M. Min, and X. Niu, 2018: First Effort for Constructing a Direct Solar Radiation Data Set in China for Solar Energy Applications. *J. Geophys. Res. Atmos.,* **123**, 1724-1734, https://doi.org/10.1002/2017JD028005

Tang, W., K. Yang, J. Qin, C. C. Cheng, and J. He, 2011: Solar radiation trend across China in recent decades: A revisit with quality-controlled data. *Atmos. Chem. Phys.,* **11**, 393-406, https://doi.org/10.5194/acp-11-393-2011.

Tang, W., K. Yang, J. Qin, and M. Min, 2013: Development of a 50-year daily surface solar radiation dataset over China. *Sci. China Earth Sci.,* **56**, 1555-1565, https://doi.org/10.1007/s11430-012-4542-9.

Wang, G., Xiong X, 1991: Analysis of some characteristics of solar radiation at Zhongshan Station, Antarctica. *Antarctic Research*, **3**, 64-68. (in Chinese)

Wang, L., O. Kisi, M. Zounemat-Kermani, G. A. Salazar, Z. Zhu, and W. Gong, 2016: Solar radiation prediction using different techniques: Model evaluation and comparison. *Renew. Sustain. Energy Rev.,* **61**, 384-397, https://doi.org/10.1016/j.rser.2016.04.024.

Wang, R., 2012: AdaBoost for Feature Selection, Classification and Its Relation with SVM, A Review. *Phys. Procedia*, **25**, 800-807, https://doi.org/10.1016/j.phpro.2012.03.160.

Wang, Y., and M. Wild, 2016: A new look at solar dimming and brightening in China. *Geophys. Res. Lett.,* **43**, 11,777-11,785. https://doi.org/10.1002/2016GL071009

Wen J., J. L. Zhao S. W. Luo, and Z. Han, 2002: The improvements of BP neural network learning algorithm. Preprints, *WCC 2000-ICSP 2000. 2000 5th international conference on signal processing proceedings. 16th world computer congress 2000*, 1647-1649.

Wild, M., 2009. Global dimming and brightening: A review. *J. Geophys. Res. Atmos.*, **114**, https://doi.org/10.1029/2008JD011470.

Wild, M., and Coauthors, 2005: From dimming to brightening: Decadal changes in solar radiation at Earth's surface. *Science,* **308**, 847-850, https://doi.org/10.1126/science.1103215.

Xue, W., and Coauthors, 2019: Declining diurnal temperature range in the North China Plain related to environmental changes. *Clim. Dyn.*, **52**, 6109–6119. https://doi.org/10.1007/s00382-018-4505-8.

Yang, Y., S. P. Palm, A. Marshak, D. L. Wu, H. Yu, and Q. Fu, 2014: First satellite-detected perturbations of outgoing longwave radiation associated with blowing snow events over Antarctica. *Geophys. Res. Lett,* **41**, 730-735, https://doi.org/10.1002/2013GL058932.

Yu, L., and Coauthors, 2019: The variability of surface radiation fluxes over landfast sea ice near Zhongshan Station, east Antarctica during austral spring. *Int. J. Digit. Earth,* **12**, 860-877, https://doi.org/10.1080/17538947.2017.1304458.

Zelterman, D., 2015: Applied Multivariate Statistics with R. Appl. Multivar. Stat. with R *ed. Springe*, https://doi.org/10.1007/978-3-319-14093-3.

Zeng, Z., and Coauthors, 2020: Daily Global Solar Radiation in China Estimated From High-Density Meteorological Observations: A Random Forest Model

Framework. *Earth Sp. Sci*, **7**, https://doi.org/10.1029/2019EA001058.

Zhang, T., C. Zhou, and L. Zheng, 2019: Analysis of the temporal–spatial changes in surface radiation budget over the Antarctic sea ice region. *Sci. Total Environ*, **666**, 1134-1150, https://doi.org/10.1016/j.scitotenv.2019.02.264.

Zhang, X., S. Liang, G. Wang, Y. Yao, B. Jiang, and J. Cheng, 2016: Evaluation of the reanalysis surface incident shortwave radiation products from NCEP, ECMWF, GSFC, and JMA using satellite and surface observations. *Remote Sens*. **8**, https://doi.org/10.3390/rs8030225.

Table 1. The short name, definition and sensor type of variables.

Variable	Unit	Definition	Sensor type
SP	hPa	Atmospheric pressure	Vaisala PTB220
RH	%	Relative humidity	Vaisala HMP35D
Tem	°C	Surface air temperature	Vaisala HMP35D
WS	m/s	Wind speed	XFY3-1
SSD	h	Sunshine duration	Australia DSU12
GSR	MJ/m^2	Global solar radiation	TBQ-2-B
CF	%	Cloud fraction	-
LCF	%	Low cloud fraction	-
Vis	km	Visibility	-
WVP	hPa	Ground vapor pressure	-
DGSR	MJ/m^2	Sum of global solar radiation over a day	-
DOY	-	Day of year	-
Month	-	Month of year	-

Table 2. The final selected values of the main parameters for each model.

Model name	Parameter	Final value
RF	n_estimators	500
	oob_score	True
	n_jobs	2
LightGBM	n_estimators	3500
	num_leaves	600
	learning_rate	0.05
	max_depth	18
XGBoost	max_depth	16
	learning_rate	0.1
	n_estimators	700
GBRT	n_estimators	750
	Learning_rate	0.5
	Max_depth	3
BP	Solver	adam
	Alpha	1e-5
	hidden-layer-sizes	1000,500,100
SVM	tol	0.000001
AdBoost	n_estimators	500
MLR	--	--
DT	--	--

--: Indicates the model parameter value was set as the default.

Table 3. The results of the fitted model, 10-fold CV model and historical estimation power by different machine learning models.

Model name	Fitted model			10-fold CV model			Historical estimation power		
	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE
RF	0.998	0.434	0.238	0.988	1.183	0.648	0.984	1.494	0.846
LightGBM	0.999	0.039	0.027	0.987	1.22	0.78	0.982	1.481	0.884
XGBoost	0.999	0.001	0.013	0.987	1.236	0.678	0.982	1.469	0.845
BP	0.982	1.581	1.132	0.982	1.604	1.151	0.978	1.861	1.370
GBRT	0.999	0.398	0.185	0.98	1.541	0.98	0.974	1.541	0.980
DT	0.998	0.533	0.314	0.976	1.692	0.919	0.967	1.999	1.146
AdBoost	0.962	3.487	3.063	0.959	3.535	3.097	0.961	3.417	3.097
SVM	0.892	5.561	4.755	0.892	5.498	4.687	0.897	3.865	2.980
MLR	0.895	3.498	2.773	0.895	3.500	2.777	0.882	3.778	2.988

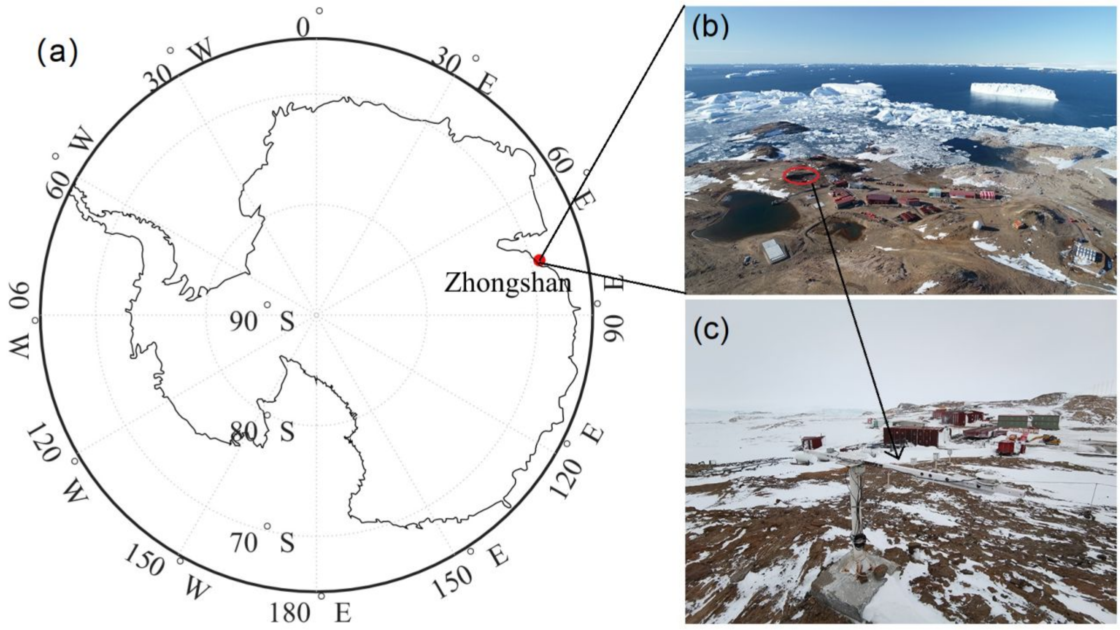

Fig. 1. (a) Map showing the location of Zhongshan Station, (b) image of Zhongshan Station area, and (c) the solar radiation instrument in the meteorological observation field.

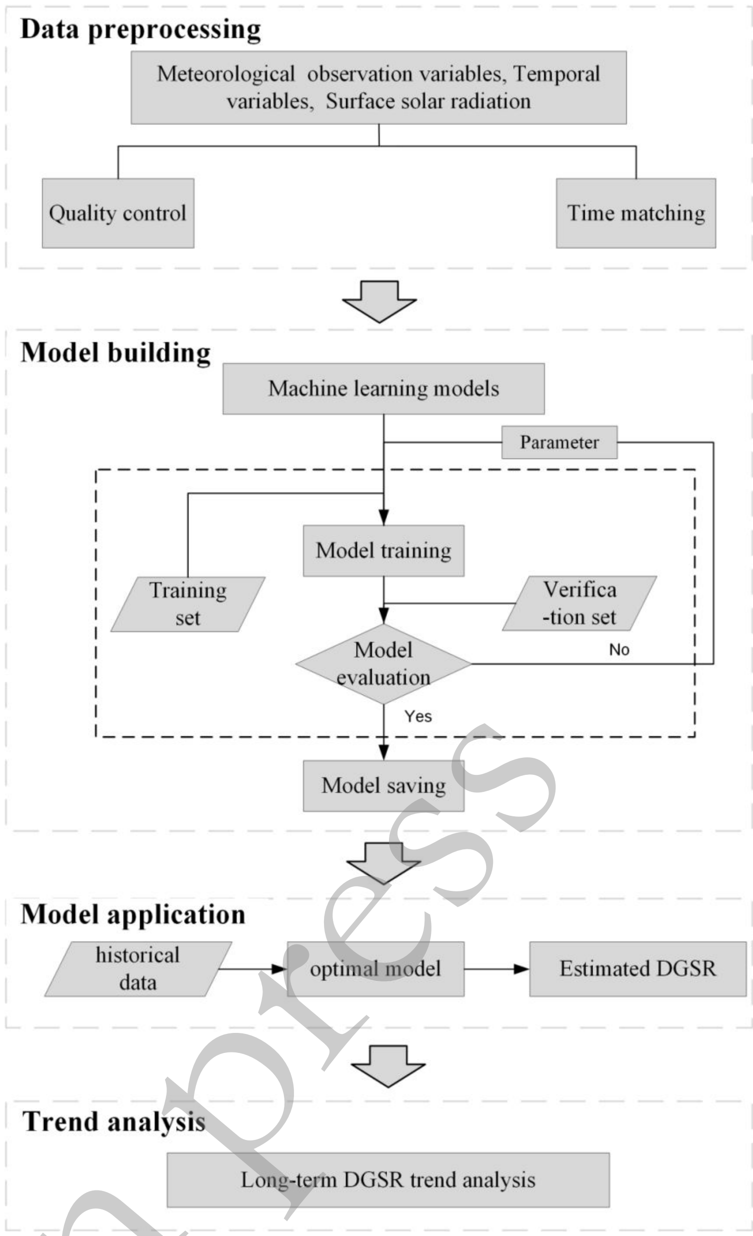

Fig. 2. Steps of historical estimation and long-term trend analysis of DGSR at Zhongshan Station, Antarctica.

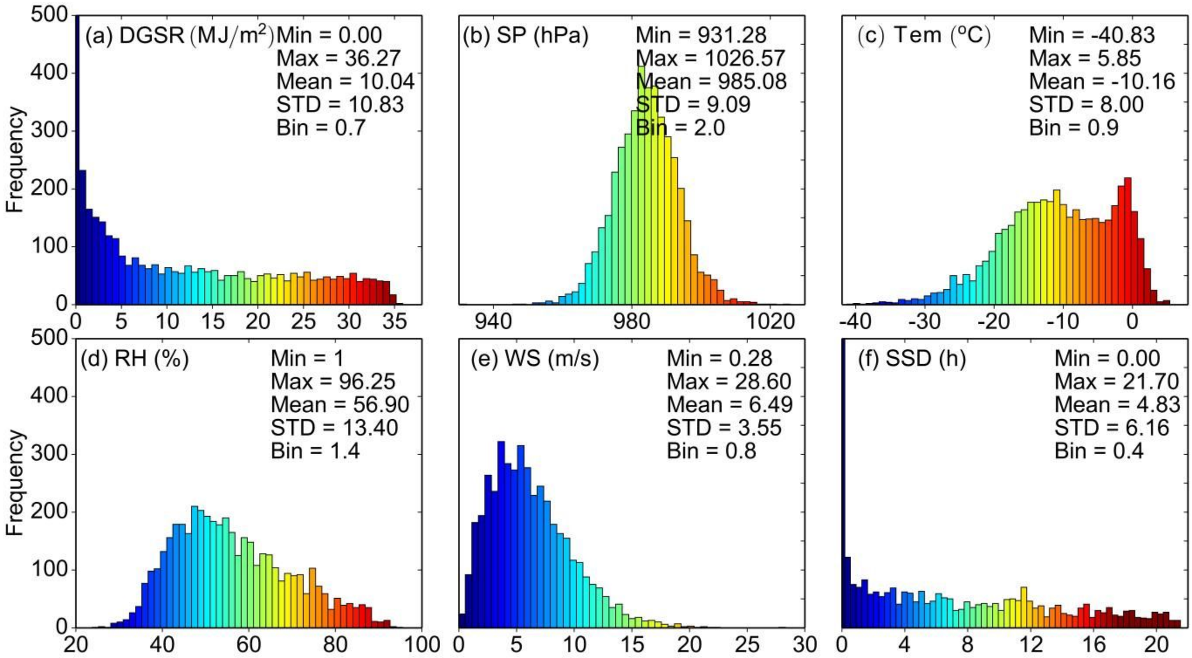

Fig. 3. The frequency distribution of (a) observed DGSR, (b) SP, (c) Tem, (d) RH, (e) WS, and (f) SSD at Zhongshan Station, Antarctica during 2010-2020 for model training and cross-validation.

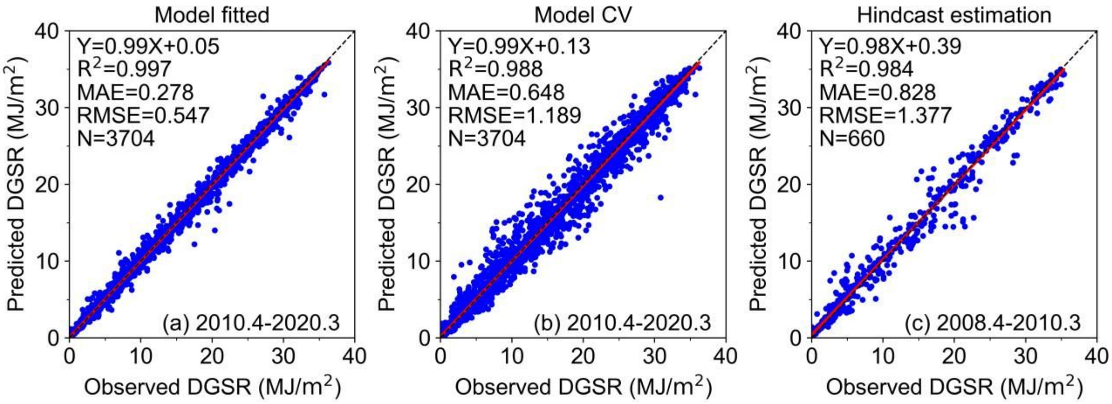

Fig. 4. Scatterplots of the (a) fitted model, (b) CV model and (c) hindcast estimation results of the RF model at Zhongshan Station, Antarctica.

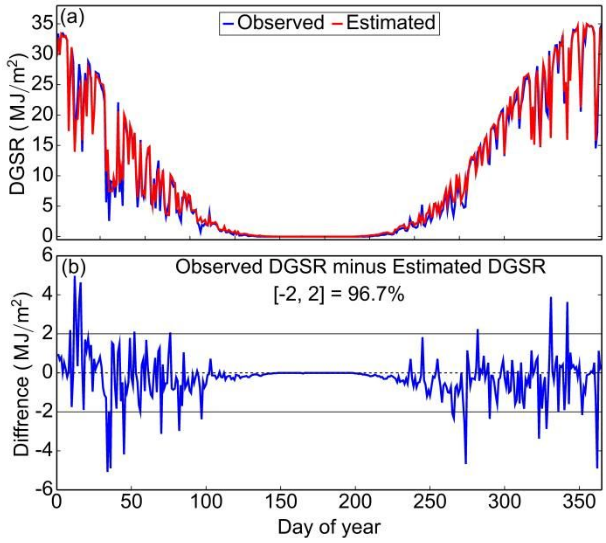

Fig. 5. The (a) time series of observed (blue) versus estimated (red) DGSR and the (b) corresponding difference (observed DGSR - estimated DGSR) in 2009 at Zhongshan Station, Antarctica. [-2, 2] indicates a difference within ±2 MJ/m^2.

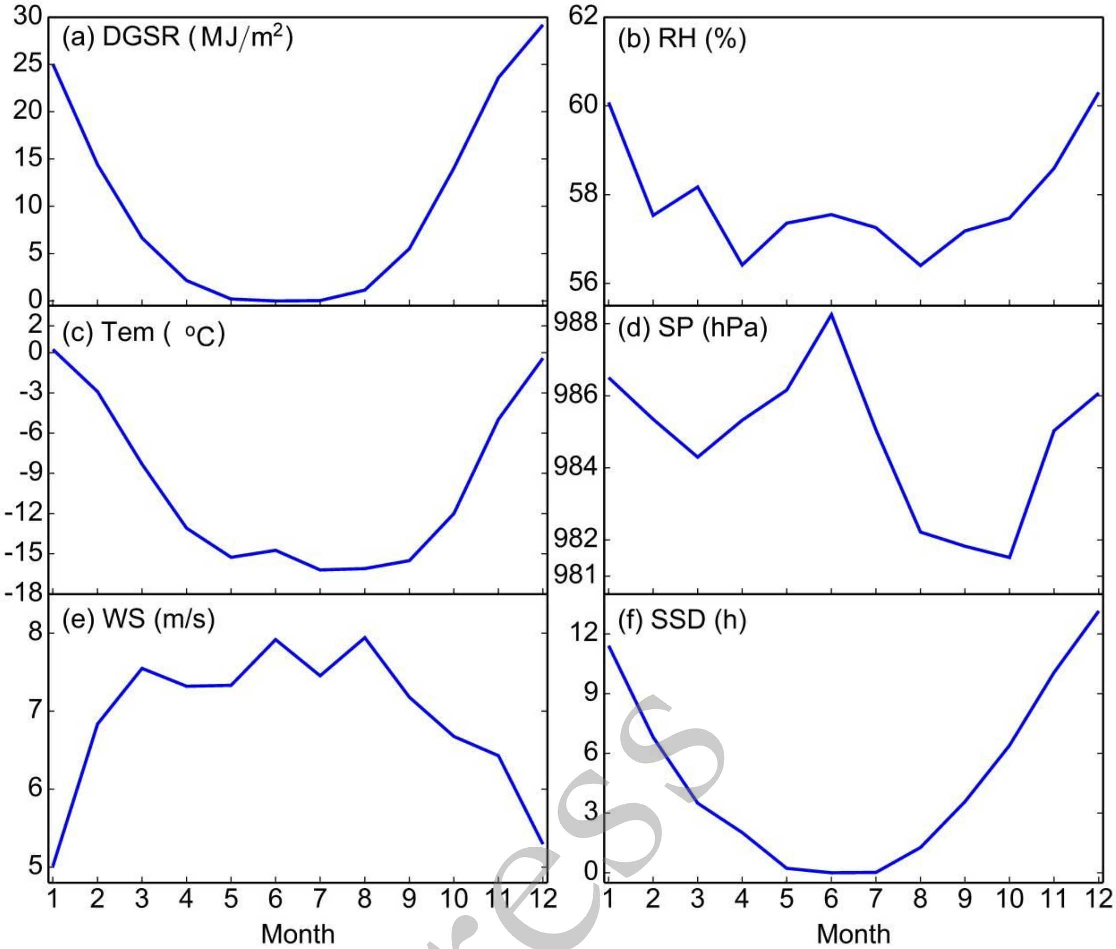

Fig. 6. The monthly variation in (a) DGSR (estimated), (b) SP, (c) Tem, (d) RH, (e) WS, and (f) SSD at Zhongshan Station, Antarctica during 1990-2019.

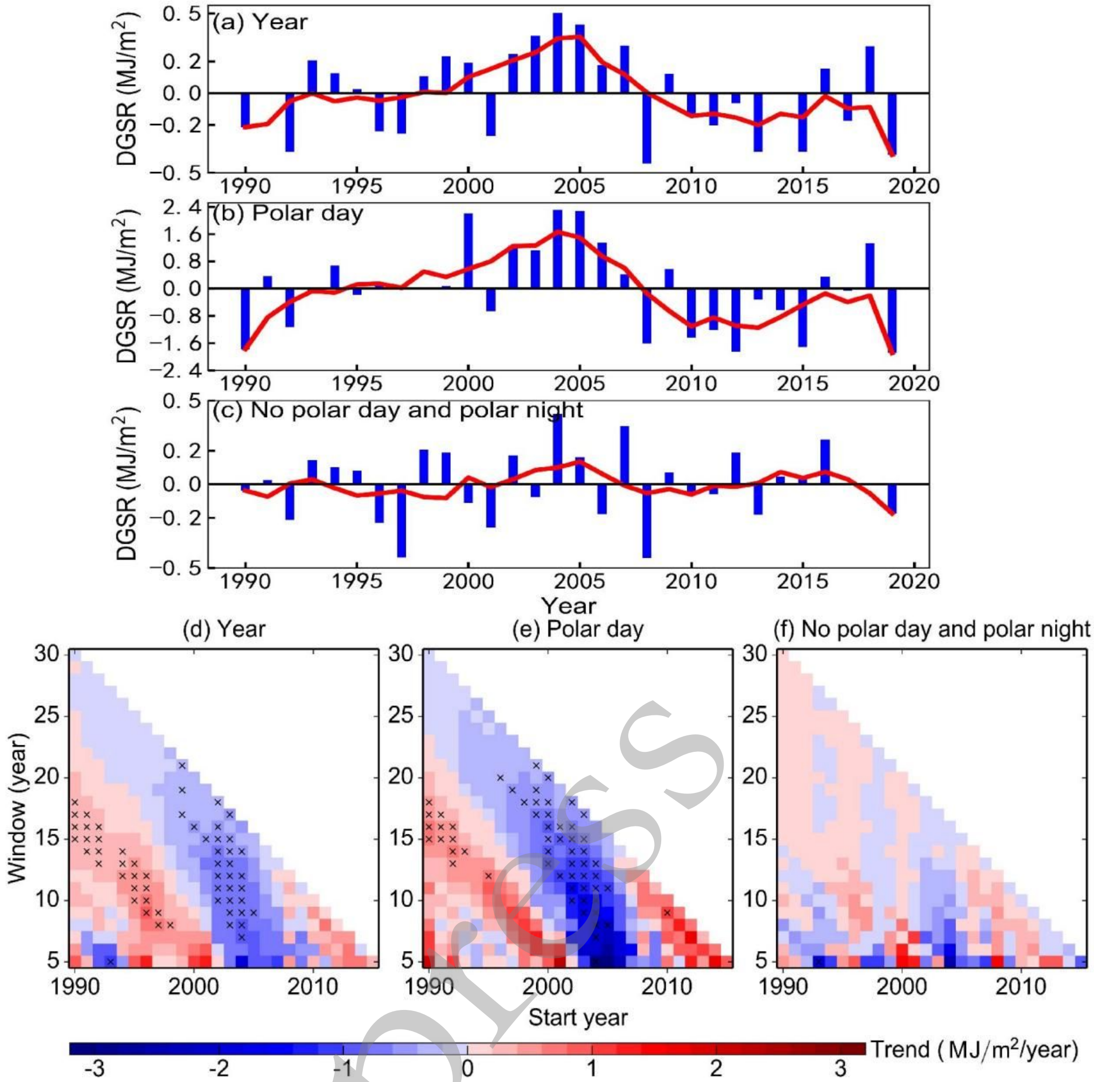

Fig. 7. Time series of the annual mean anomalies of estimated DGSR: (a) annual mean but no polar night, (b) polar day and (c) no polar day and no polar night. The red lines indicate the 5-year running means of the DGSR anomalies. Sliding-window trend analyses of annual mean estimated DGSR at Zhongshan Station, Antarctica, from 1990 to 2019 for (d) all years but no polar night, (e) polar day and (f) no polar day and no polar night.

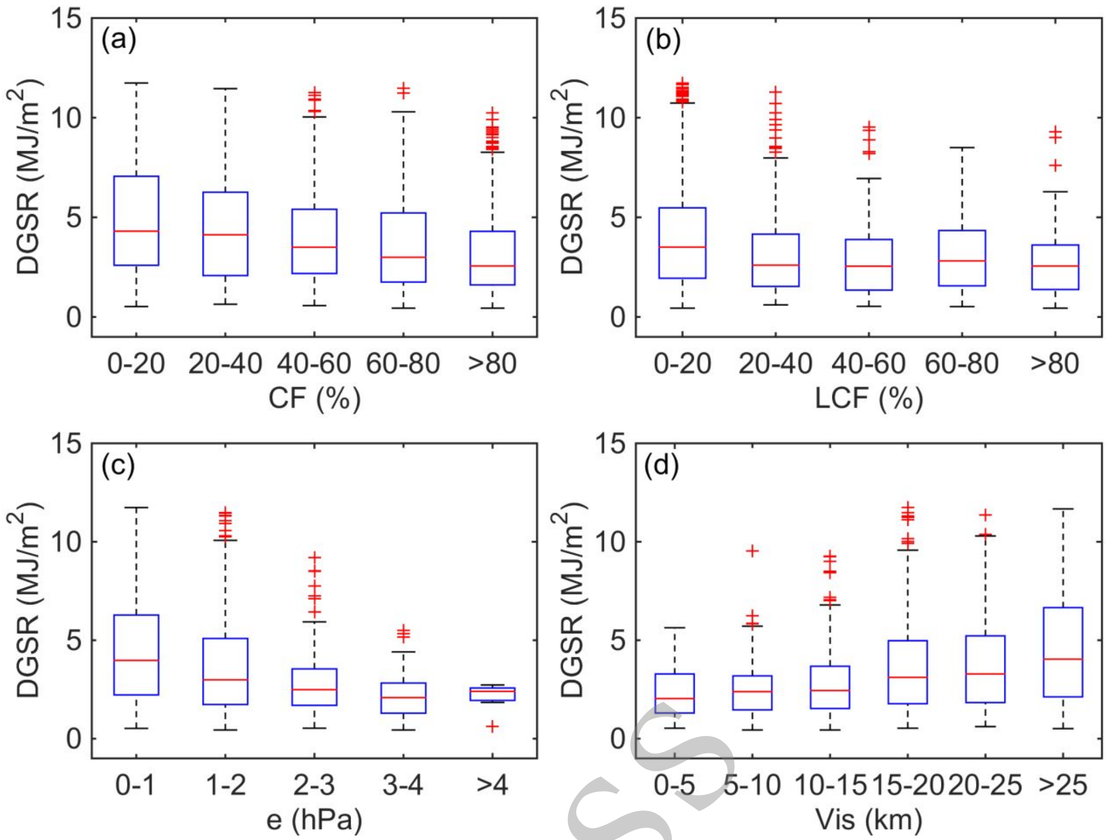

Fig. 8. The effect on the estimated DGSR by the different (a) CF, (b) LCF, (c) e, and (d) Vis conditions at Zhongshan Station, Antarctica.

Supplementary Information

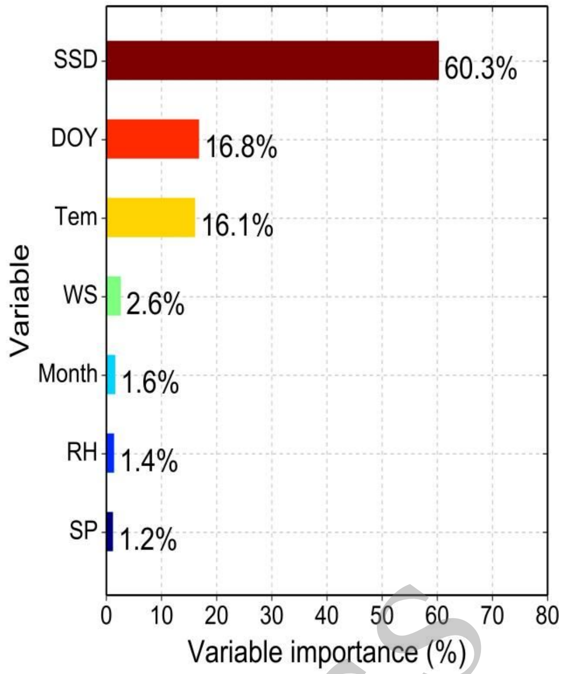

Figure S1. Evaluation of the relative importance of the variables used in the RF models. SSD: sunshine duration; DOY: day of year; Tem: air temperature; WS: wind speed; RH: relative humidity; SP: surface air pressure.

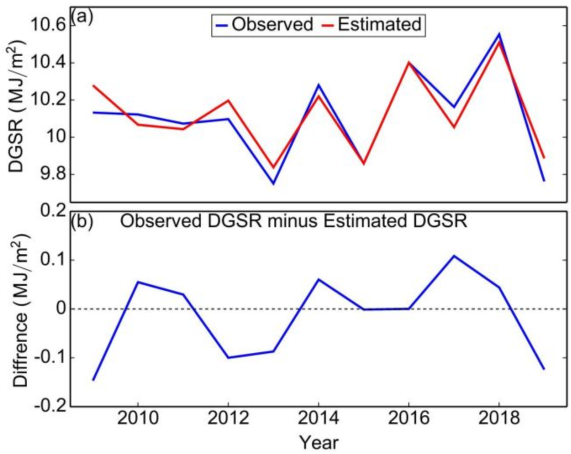

Fig. S2. The (a) annual change in observed (blue) versus estimated (red) DGSR and the (b) corresponding difference (observed DGSR - estimated DGSR) from 2009 to 2019 at Zhongshan Station, Antarctica.

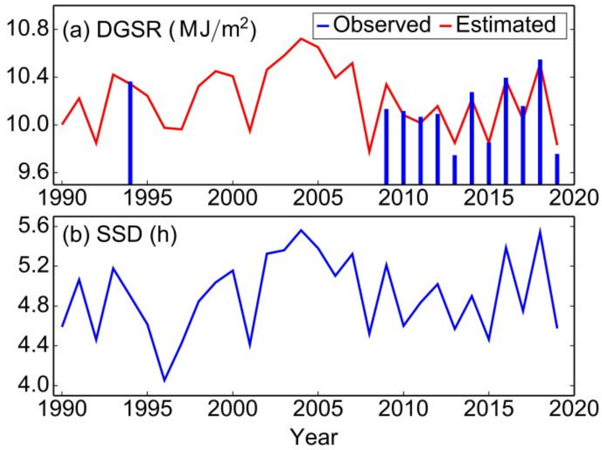

Fig. S3. The (a) annual change in observed (blue) versus estimated (red) DGSR and the (b) annual change in SSD from 1990 to 2019 at Zhongshan Station, Antarctica.

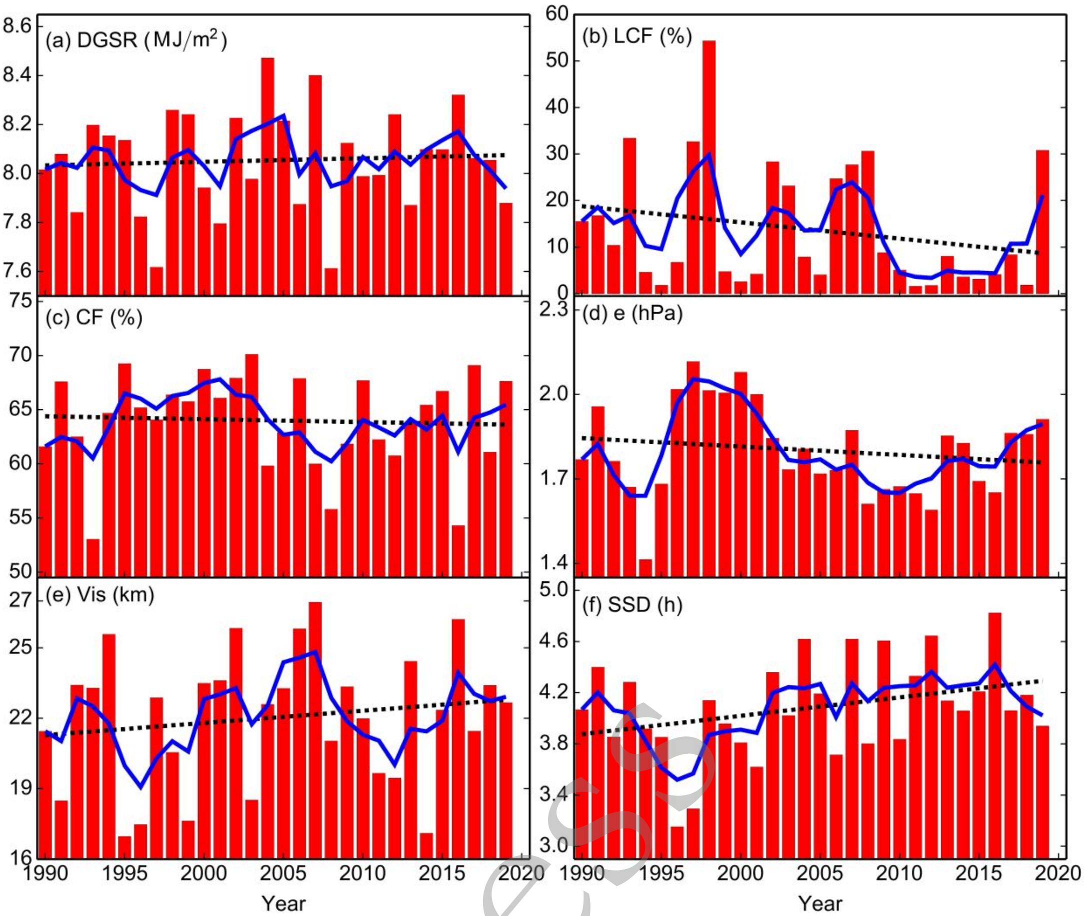

Fig. S4. The yearly series of (a) DGSR (estimated), (b) LCF, (c) CF, (d) e, (e) Vis, and (f) SSD at Zhongshan Station, Antarctica from 1990 to 2019 for no polar day and no polar night. The blue lines indicate their corresponding 5-year running mean values and straight lines indicate their corresponding linear trends.