A New Long-Term Downward Surface Solar Radiation Dataset over China from 1958 to 2015

Ning Hou [1,2], Xiaotong Zhang [1,2,*], Weiyu Zhang [1,2] , Jiawen Xu [1,2], Chunjie Feng [1,2] , Shuyue Yang [1,2], Kun Jia [1,2] , Yunjun Yao [1,2] , Jie Cheng [1,2] and Bo Jiang [1,2]

[1] State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China; houning_0110@mail.bnu.edu.cn (N.H.); zhangweiyu@mail.bnu.edu.cn (W.Z.); xujw@mail.bnu.edu.cn (J.X.); fcj20190901@mail.bnu.edu.cn (C.F.); ysy_0128@mail.bnu.edu.cn (S.Y.); jiakun@bnu.edu.cn (K.J.); yaoyunjun@bnu.edu.cn (Y.Y.); Jie_Cheng@bnu.edu.cn (J.C.); bojiang@bnu.edu.cn (B.J.)

[2] Beijing Engineering Research Center for Global Land Remote Sensing Products, Institute of Remote Sensing, Science and Engineering, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China

* Correspondence: xtngzhang@bnu.edu.cn; Tel.: +86-10-5880-7712

Abstract: Downward surface solar radiation (Rs) plays a dominant role in determining the climate and environment on the Earth. However, the densely distributed ground observations of Rs are usually insufficient to meet the increasing demand of the climate diagnosis and analysis well, so it is essential to build a long-term accurate Rs dataset. The extremely randomized trees (ERT) algorithm was used to generate Rs using routine meteorological observations (2000–2015) from the Climate Data Center of the Chinese Meteorological Administration (CDC/CMA). The estimated Rs values were validated against ground measurements at the national scale with an overall correlation coefficient value of 0.97, a mean bias of 0.04 Wm^{-2}, a root-mean-square-error value of 23.12 Wm^{-2}, and a mean relative error of 9.81%. It indicates that the estimated Rs from the ERT-based model is reasonably accurate. Moreover, the ERT-based model was used to generate a new daily Rs dataset at 756 CDC/CMA stations from 1958 to 2015. The long-term variation trends of Rs at 454 stations covering 46 consecutive years (1970–2015) were also analyzed. The Rs in China showed a significant decline trend (-1.1 Wm^{-2} per decade) during 1970–2015. A decreasing trend (-2.8 Wm^{-2} per decade) in Rs during 1970–1992 was observed, followed by a recovery trend (0.23 Wm^{-2} per decade) during 1992–2015. The recovery trends at individual stations were found at 233 out of 454 stations during 1970–2015, which were mainly located in southern and northern China. The new Rs dataset would substantially provide basic data for the related studies in agriculture, ecology, and meteorology.

Keywords: downward surface shortwave radiation; extremely randomized trees; random forest; global brightening; global dimming

1. Introduction

The downward surface solar radiation (Rs) plays a dominant role in the global radiation budget as it is a basic element of energy source on the earth [1,2]. It is an important driving force of various biological, chemical, and physical processes of the Earth's system [3–6]. Therefore, understanding and determining the variability of Rs are crucial for practical applications such as environmental, hydrological, and ecological studies [7–11].

Rs is not directly measurable using the satellite sensors due to the atmospheric influences, the introduction of empirical or physical-based models for Rs estimation will induce possible uncertainties [12]. Direct ground measurements are essential for Rs quantification since it is one of the

most accurate *Rs* data sources. The monitoring of *Rs* started at limited ground stations since the early 20th century, the measurements of *Rs* were more widespread after the International Geophysical Year (1957/1958) [13]. The worldwide ground measurements of *Rs* revealed a significant declining trend from the 1950s to the 1980s (coined as 'global dimming'), as well as a recovery since the late 1980s phenomenon ('global brightening') in different regions around the world [14–18], such as Europe [19,20], USA [21,22], New Zealand [23], and Japan [24]. For example, it reported that a decline of *Rs* (-2.0 to -3.1 Wm^{-2} per decade) during 1971–1986 and a recovery of *Rs* (1.1 to 1.4 Wm^{-2} per decade) during 1987–2002 were found in Europe [19].

Many studies have also widely reported that the *Rs* in China experienced a significant decline from the 1960s and a subsequent recovery from the 1990s [25–32]. Che et al. [29] found the *Rs* over China exhibited a significant decline (-4.5 Wm^{-2} per decade) during 1961-2000 using the data at 64 stations from the Climate Data Center of the Chinese Meteorological Administration (CDC/CMA). A decreasing tendency in *Rs* (-3.1 Wm^{-2} per decade) was found based on the analysis at 85 CDC/CMA stations by Qian et al. [30]. Wang et al. [32] noticed that the *Rs* in China declined (8.1 Wm^{-2} per decade) during 1961–1989 and increased (2.7 Wm^{-2} per decade) during 1990–2009 using the ground measured *Rs* data at 52 CDC/CMA stations. Yang et al. [25] reported the *Rs* in China exhibited a decline (-4.9 Wm^{-2} per decade) during 1958-2016 based on the *Rs* measurements from 95 CDC/CMA stations. Although much effort has been conducted on the long-term *Rs* variability analysis over China, it still has fierce controversy. Moreover, the number of stations used in these studies was limited. Thus, it is still essential to reconstruct the long-term accurate *Rs* dataset at more stations for analyzing the variations of *Rs* over China.

It is known that *Rs* is influenced by temporal inhomogeneities due to the updates in the measuring in instrumentation and relocation [33,34]. The changes in the measuring equipment of the CDC/CMA may result in uncertainty of the long-term trend analysis of the *Rs* over China by introducing breaks [35–37]. The *Rs* is closely related to meteorological variables, including clouds, aerosols, temperature, precipitation, sunshine duration, and other factors [38–44]. It has been proved that it is feasible to reconstruct *Rs* dataset from those widely measured meteorological variables. Compared to the radiation observation data, the meteorological measurements have better temporal homogeneity and spatial representativeness. Currently, various methods have been widely applied to estimate *Rs* using meteorological variables, which include the empirical statistical models [45,46], physical parameterization models [42,47], and machine learning methods [38,40,48,49]. The empirical statistical models establish relationships between the meteorological variables and *Rs* observations for estimating *Rs*, but they are usually site-dependent. The physical parameterization models simulate the interactions between the solar radiation and atmosphere for estimating *Rs*. However, the proposed physical parameterization models may be directly influenced by the accuracy of the input variables, such as clouds and aerosol properties, and so on [50–52].

The machine learning method (e.g., artificial neural network (ANN), support vector machine (SVM), Random Forest (RF), etc.), which automatically learns rules from a large amount of historical data for prediction, is an alternative approach to reconstruct *Rs* accurately. For example, Tang et al. [40] applied an ANN model to develop a 50-year dataset of daily *Rs* over China, the new *Rs* data showed higher accuracy than previous estimated *Rs* data. Fan et al. [53] applied the SVM and extreme gradient boosting (XGBoost) to estimate daily *Rs* over China, it showed the XGBoost method was promising and effective to estimate *Rs*. Wang et al. [54] evaluated the Adaptive Neuro Fuzzy Inference System (ANFIS) and M5 model tree (M5Tree) method for estimating daily *Rs* in China, it indicated the ANFIS method had better performance than the M5Tree method in *Rs* estimation. Some studies also applied the machine learning methods for estimating *Rs* using satellite observations [55–60]. As far as now, RF has been widely used in remote sensing for classification and regression problems [55,56,61,62] due to its great accuracy and efficiency. Extremely randomized trees (or extra trees, ERT), considered as the further development of the RF method, can efficiently reduce the variance of models and excavate more significant information than the RF method by introducing a stronger randomization method [63].

However, studies on the *Rs* estimation based on the ERT method is rare compared to that based on the RF method.

Therefore, the main objective of this study was to develop an *Rs* estimation model based on the ERT method using quality-controlled radiation measurements and meteorological measurements from CDC/CMA. The *Rs* estimates based on the ERT approach were compared with ground measurements. Additionally, the spatial distributions, seasonal variations, and long-term trends of the estimated *Rs* over China from 1970 to 2015 were also analyzed based on the new dataset. Sections 2 and 3 introduce the data and methods used in this study. Section 4 introduces the results of the ERT-based model and presents the spatiotemporal analysis of the *Rs* estimates over China. Section 5 discusses the probable cause of variations in *Rs*. A short summary is given at the end of this paper.

2. Data

There is a total of 756 meteorological stations in China and only 96 meteorological stations have records of solar radiation since 1994. The *Rs* measurements by CDC/CMA started in 1957, and it is noted that the radiometers equipped at CDC/CMA stations had been updated during 1990–1993. Before releasing the radiation data, a quality control process was conducted by CDC/CMA. Nevertheless, some studies [43,64] suggested that the radiation data of CDC/CMA need to be inspected more strictly for further application. The method proposed by Tang et al. [43] was performed to examine the quality of radiation data. In this study, the "complete records" of meteorological data, which was defined as data that contains more than 20 days in every month, and 12 months in a year, were used for model construction at 96 CDC/CMA stations and reconstructing *Rs* at 756 CDC/CMA stations. To analyze the temporal variations of *Rs* in China, the reconstructed *Rs* data at 454 CDC/CMA stations with complete records for consecutive 46 years (1970–2015) were used for analysis.

The routine daily meteorological measurements, which include air pressure, air temperature, wind speed, relative humidity, daily precipitation, water vapor pressure, sunshine duration, and the *Rs* data collected from CDC/CMA were used to reconstruct *Rs*. In addition, the temporal information was also used to reduce the influence of the seasonal cycle of *Rs*, since *Rs* have the clear temporal variations. For each station, the cosine of the radian difference (*T*) is obtained according to Equation (1) proposed by Wei et al. [65], which is capable of minimizing the influence of the seasonal cycle.

$$T_t = \cos\left(2\pi \frac{d_t}{D}\right) \tag{1}$$

where *d* represents the Julian day of the year (DOY), *D* denotes the total number of days in a year.

Figure 1a displays the spatial distributions of 454 stations from CDC/CMA, the 96 radiation stations were denoted by the star symbols. This study also discussed the regional *Rs* trends in different climatic regions of China. Various approaches for classifying climatic types in China [66–69] were presented in previous studies. In this study, six climatic regions in China were classified according to the classification in Wang et al. [68] and Zhou et al. [69]. Figure 1b displays the six climatic zones, northeast China (NE), north China (NC), east China (EC), south China (SC), southwest China (SW), and Tibetan Plateau (TP).
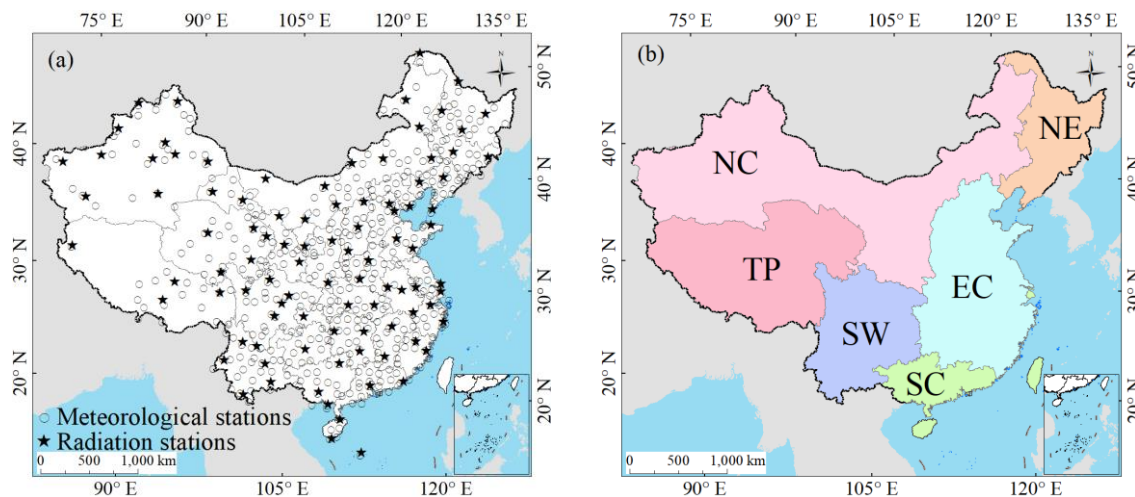
Figure 1. (**a**) The spatial distribution of 454 Climate Data Centers of the Chinese Meteorological Administration (CDC/CMA) stations, the 96 radiation stations are denoted by five-pointed star symbols, (**b**) six climatic regions of China.

3. Methodology

3.1. Extremely Randomized Trees (ERT)

The decision tree, a nonlinear and nonparametric method, is widely used in regression and classification problems. Decision trees predict the value of a target variable using a set of values of input variables. The decision trees are capable of dealing with problems of large-scale data using plenty of training samples and input variables. The decision tree method is easy and clear to understand. The relevant variables can be recognized during the growth of trees, which provides robustness for the decision trees model [63]. However, the high sensitivity for training samples is the main inadequacy of the single tree [70,71], the low accuracy and high randomness of the single tree result from this high sensitivity restrict the application of the single tree, particularly in handle numerous datasets [63]. The ensemble of trees, such as RF and ERT, is capable of conquering the problem of single tree models.

RF, a powerful ensemble-learning method, was proposed by Breiman in 2001, which is widely applied as a classification and regression tool [72]. The RF method employs the bootstrap technique (Ibrahim and Khatib 2017), the bootstrap samples are randomly created and replaced from the training data. Using bootstrap samples in the RF model generates around one-third unused subset, named as out-of-bag data (OBB) [72], the rest data are called in-bag data. According to the minimized Gini index, the best split is determined among the subgroup of the random selection at each node.

ERT, a tree-based machine learning method, was proposed by Geurts et al. [73]. ERT is considered as the further development of the RF method. ERT has been widely used to solve diverse sequence-based prediction problems [74,75]. ERT introduces a more powerful randomization method to efficiently reduce the variance of models and excavate more significant information than other tree-based methods. There are three main parameters in ERT for regression problems including K, n_{min}, N, and M. The K denotes the number of random splits, the n_{min} indicates the minimum sample to split a node [63,73], the N is size of samples, and the M represents the number of trees of ERT. The growth of trees in ERT is conducted through exactly defining values of the K on each node to achieve pure outputs in all subsamples [63], the M controls the degree of variance reduction in the ensemble model [73]. In addition, the variance of ERT generally further decreases with the increase of M [63], but the bias may increase slightly. The ERT offers added robustness about obvious errors due to the marginal

affection of outliers on ERT prediction. The variable importance measure is also provided by the ERT method, which is defined as follows:

$$VI^{(t)}(z) = \frac{1}{N}\sum_N \left(\frac{\sum_{X_a \in B^{c(t)}}\left(I\left(L_b = c_a^{(t)}\right) - I\left(L_b = c_{a,\pi z}^{(t)}\right) \right)}{\left| B^{c(t)} \right|} \right) \tag{2}$$

where $B^{c(t)}$ is the OBB sample for a tree, X_a is sample value, t is the tree number, $c_a^{(t)}$ and $c_{a,\pi z}^{(t)}$ are the predicted values before and after change of variables, a and b are the number of samples per leave and per tree, respectively.

The ERT and RF methods are similar but different in two aspects. The ERT employs all training samples instead of the bootstrap algorithm during the growth procedure; and the ERT performs node splitting by random selections of cut-points, instead of the best node split based on the Gini index in the RF [76], which is calculated according to Equation (3):

$$I_n = 1 - \sum_{i=1}^{K}\left(\frac{M_i}{N}\right)^2 \tag{3}$$

The ERT method generate a set of independent decision trees based on the features space F. In this study, the features space X = {air pressure, air temperature, wind speed, relative humidity, water vapor pressure, daily precipitation, sunshine duration, elevation, and cosine of the radian difference}. The followed Algorithm 1 illustrates the procedure. At the begining of the training stage, the ensemble tree set is initialized as empty. Then, each decision tree is built with randomly selected features without replacement. During the testing stage, a predicted value y_i is obtained by each decision tree. The final result is the average of all the decision trees.

Algorithm 1 Extremely Randomized Trees Algorithm [73]

Given $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, with features space F where $(x_i, y_i) \in (X, Y)$ training data.
Given the N and max_depth (max depth of each tree)
procedure Train $((x_1, y_1), (x_2, y_2), \dots, (x_n, y_n))$
　　$ERT \leftarrow \{\}$
　　for i from 1 to N **do**
　　　decision tree $T_i \in X \rightarrow Y$
　　　while (do depth (T)) < max_depth)
　　　　Random select $X_i \subset X$ without replacement
　　　　Random select feature $f \in F$
　　　　　Use f as node to construct tree
　　　end while
　　　$ERT = ERT \cup \{T_i\}$
　　end for
　　return ERT
end procedure
procedure Test (x)
　　for i from 1 to N **do**
　　　Select T_i from ERT
　　　$y_i \leftarrow T_i(x)$
　　end for
　$y = \frac{\sum_{i=1}^{N} y_i}{N}$
　　return y
end procedure

3.2. Mann–Kendall (M-K) Test

The nonparametric Mann–Kendall (M-K) test [77,78] was employed for detecting the significant temporal variation of *Rs*. It is a rank-based method to identify tendencies of time-series dataset. The relative magnitudes of the samples are compared by the M-K test rather than the data values themselves. In addition, the M-K test does not require that samples conform to a certain type of distribution. The M-K test is an effective approach to identify tendencies of meteorological and other relevant variables without being affected by certain data distribution and outliers. The M-K statistics *s* is defined as:

$$s = \sum_{m=2}^{i} \sum_{n=1}^{m-1} f(T_m - T_n) \tag{4}$$

where *i* is the number of data. T_m and T_n (m>n) are observations in time series. $f()$ is given as:

$$f(T_m - T_n) = \begin{cases} 1 & T_m - T_n > 0 \\ 0 & T_m - T_n = 0 \\ -1 & T_m - T_n < 0 \end{cases} \tag{5}$$

The variance of statistics *s* is given as:

$$V(s) = i(i-1)(2i+5)/18 \tag{6}$$

The standardized test statistics *Z* is given as:

$$Z = \begin{cases} (s-1)/\sqrt{V(s)} & s > 0 \\ 0 & s = 0 \\ (s+1)/\sqrt{V(s)} & s < 0 \end{cases} \tag{7}$$

A positive Z value represents the increasing trend and a negative value represents the decreasing trend. The significance levels (p = 0.05 and 0.01) are used to identify the significant trend, corresponding $Z_{1-\alpha/2}$ = ±2.58 and $Z_{1-\alpha/2}$ = ±1.96, respectively.

4. Results and Analysis

Previous studies [35–37] reported that the update of measuring equipment of CDC/CMA could result in uncertainty in the long-term trend analysis of *Rs* in China by introducing breaks. To eliminate this effect, the *Rs* data after 2000 were applied to build the model in our study. The daily ground-measured *Rs* data collected from 96 CDC/CMA stations during 2000-2015 were used as target variables. The daily meteorological data used as predictors including air pressure, air temperature, wind speed, relative humidity, water vapor pressure, daily precipitation, and sunshine duration. The elevation data and cosine of the radian difference data at each station were also used as predictor variables. The daily *Rs* dataset was randomly split into two subsets for model construction: 80% for training the model, hence the rest 20% for assessing the performance of the ERT-based model. The k-fold cross-validation method was performed during the training procedure to evaluate the overall performance of the ERT-based model. The correlation coefficient (R), root-mean-square-error (RMSE), mean bias error (MBE), and mean relative error (MRE) were employed for evaluating the accuracy of the ERT-based model.

4.1. Evaluation Using Ground Measurements

The comparisons between the estimated daily *Rs* and corresponding ground measurements were conducted on national, regional, and station scales, respectively. The results of the ERT-based model for estimating *Rs* in the training and test stages are summarized in Table 1. The results on the national scale are displayed in Figure 2. For the training dataset, the ERT-based model obtained great performance

with an overall R of 0.99, an RMSE of 3.90 Wm^{-2}, an MBE of 0.01 Wm^{-2}, and an MRE of 1.53%. Those values were 0.97, 23.12 Wm^{-2}, 0.04 Wm^{-2}, and 9.81% for the test dataset. It is clearly showed that both the training and test datasets correlated very well with ground measurements. On the regional scale, the range of the R, RMSE, MBE, and MRE values for the training and test datasets were 0.95–0.99, 3.38–24.63 Wm^{-2}, −3.69–2.67 Wm^{-2}, and 1.42–11.30%, respectively. On the station scale, the range of the R, RMSE, MBE, and MRE values for the training and test datasets were 0.89–0.99, 2.55–35.06 Wm^{-2}, −9.19–9.00 Wm^{-2}, and 0.86–16.95%, respectively. In the training stage, the ERT-based model had the RMSE, mean absolute errors (MAE), and MRE values lower than 30 Wm^{-2}, 5 Wm^{-2}, and 12% at all stations. While in the test stage, these statistics values were lower than 30 Wm^{-2}, 5 Wm^{-2}, and 12% at 76 out of 96 stations. The test results shown in the table indicate that the ERT-based model was accurate to obtain the *Rs* estimates.

　　　Figures 3 and 4 show the estimation results using all datasets in both training and test stage to evaluate the overall performance of the ERT-based model. As shown in Figure 3, the test results suggest that the ERT-based model performed well in estimating the *Rs* in six climatic regions with reasonable accuracy. The ERT-based model in NE and TP provided slightly better accuracy than other regions, with R of 0.99 and 0.99, RMSE of 9.90 and 10.64 Wm^{-2}, MBE of 0.75 and 0.34 Wm^{-2}, and MRE of 3.06% and 2.49%, respectively. In addition to validations on national and regional scales, the validation of the ERT-based model was also conducted at individual CDC/CMA stations. As displayed in Figure 4, the daily estimated *Rs* agreed well with the *Rs* measurements at most CDC/CMA stations. For example, the great performance was found at station 52818 (Geermu) with R, RMSE, MBE, and MRE of 0.99, 7.23 Wm^{-2}, −0.14 Wm^{-2} and 1.73%, respectively. In contrast, the ERT-based model did not provide great *Rs* estimates at station 57874 (Changning) with R, RMSE, MBE and MRE of 0.98, 19.90 Wm^{-2}, −1.64 Wm^{-2}, and 7.39%, respectively. Overall, the R values were 0.99 at 91 out of 96 stations. The RMSE, and MRE, values were lower than 15 Wm^{-2} and 5% at 92 stations. The MAE were lower than 2 Wm^{-2} at 86 stations. It is obvious that the ERT-based model successfully estimated accurate daily *Rs* at most CDC/CMA stations.

Table 1. Training and test results of the ERT-based model.

| Scale.	Dataset		R	RMSE	MBE	MRE
National	Train		0.99	3.9	0.01	1.53%
	Test		0.97	23.12	0.04	9.81%
Regional	Train	NE	0.99	3.38	0.28	1.42%
		EC	0.99	3.39	−0.05	1.74%
		NC	0.99	4.02	0.2	1.42%
		SC	0.99	3.63	−0.51	1.53%
		SW	0.99	4.05	−0.28	1.64%
		TP	0.99	3.86	0.12	1.23%
	Test	NE	0.97	21.13	2.67	9.61%
		EC	0.96	23.35	−0.33	11.30%
		NC	0.97	23.41	1.2	8.91%
		SC	0.96	21.97	−3.49	9.95%
		SW	0.95	24.63	−2.04	10.64%
		TP	0.95	22.52	1.18	7.60%
Seasonal	Train	Spring	0.99	4.5	−0.21	1.53%
		Summer	0.99	4.66	0.1	1.46%
		Autumn	0.99	3.36	0.14	1.52%
		Winter	0.99	2.75	−0.01	1.66%
	Test	Spring	0.95	26.58	−1.48	9.89%
		Summer	0.94	28.47	0.91	9.66%
		Autumn	0.96	19.65	0.8	9.58%
		Winter	0.96	15.57	0.07	10.24%
Stational	Train		0.99	2.55–7.54	−1.42–1.22	0.86–3.66%
	Test		0.89–0.98	14.95–35.06	−9.19–9.00	4.95–16.95%
DOY	Train		0.99	2.23–5.38	−0.73–0.76	1.26–1.86%
	Test		0.91–0.97	11.32–32.97	−5.30–6.86	7.55–12.77%

The unit of RMSE and MBE is Wm^{-2}.
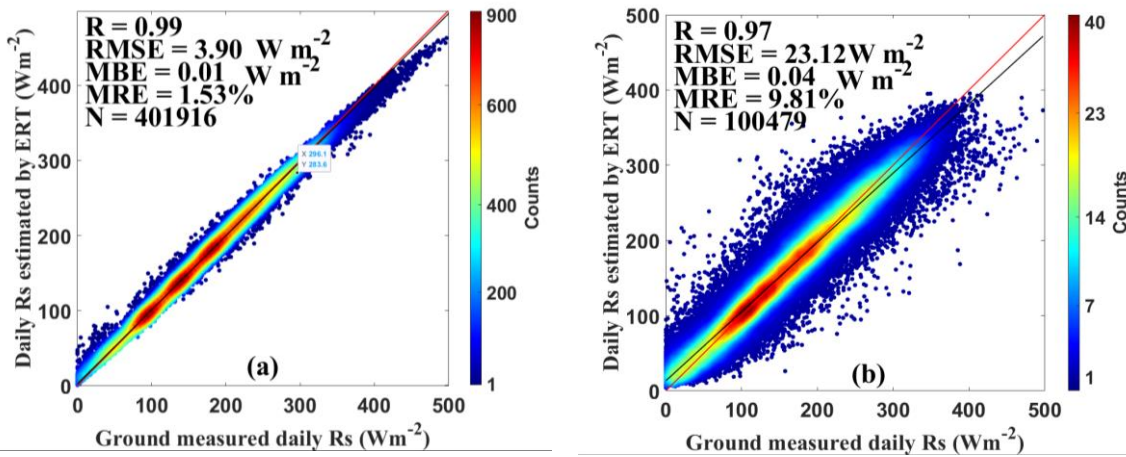
Figure 2. Evaluation results of the extremely randomized trees (ERT)-based model for the (**a**) training dataset and (**b**) test dataset against ground-measured *Rs* during 2000–2015.
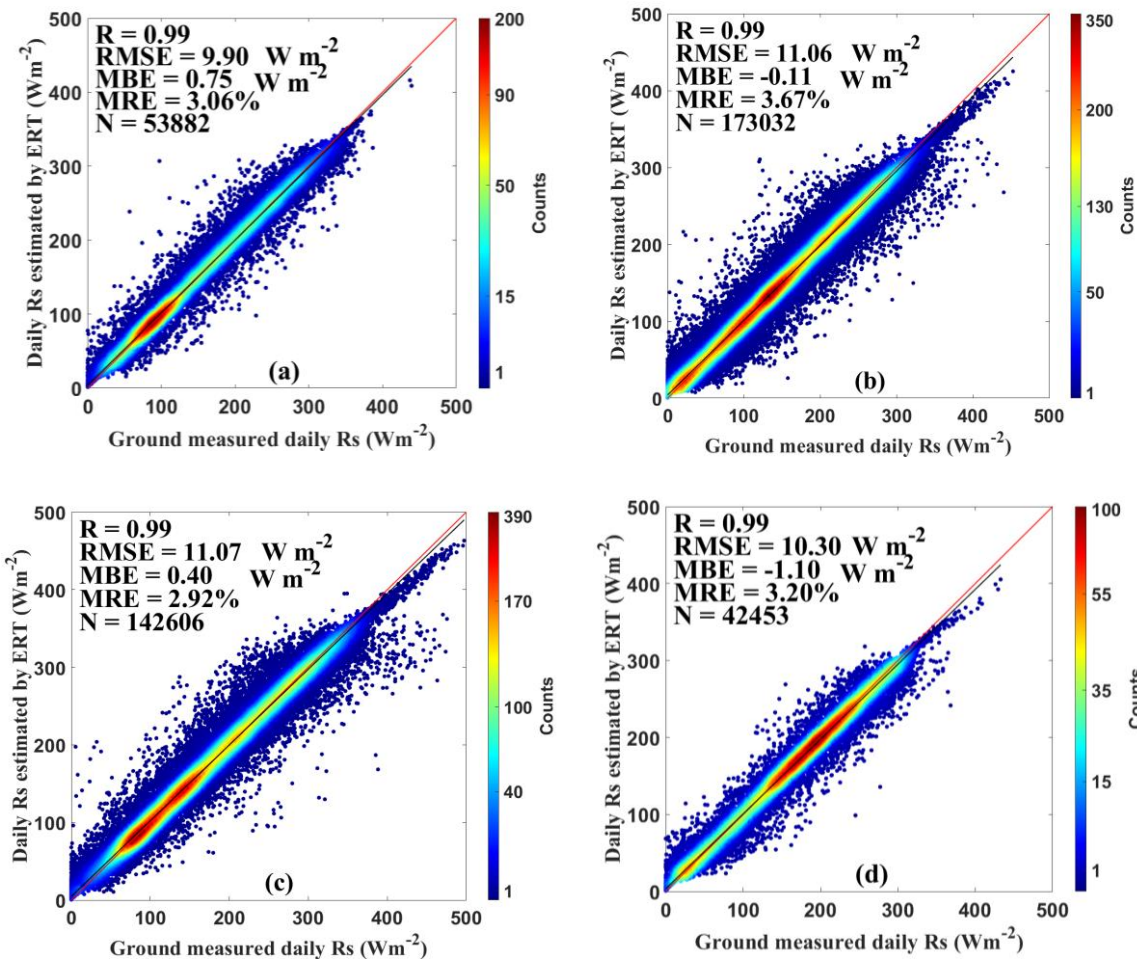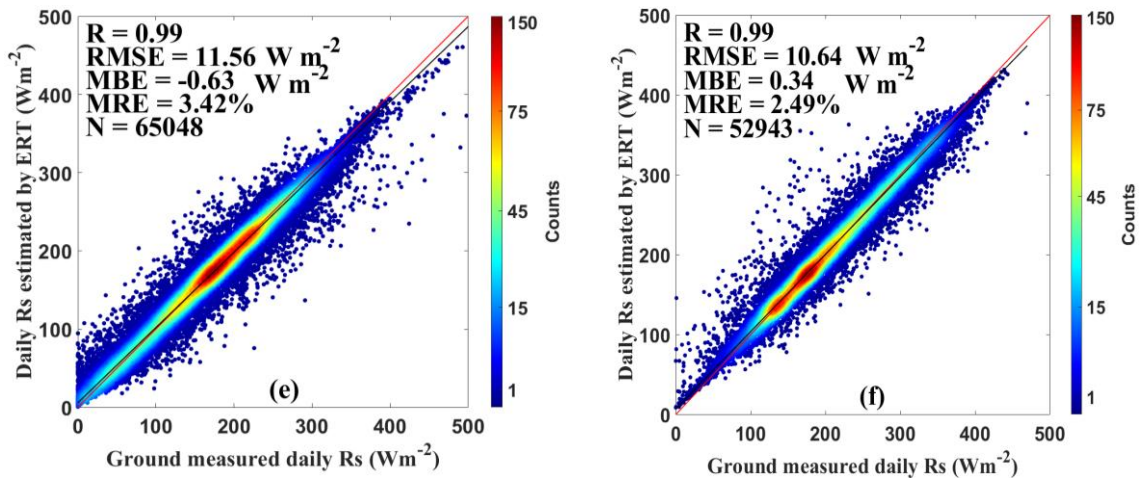

Figure 3. *Cont.*

Figure 3. Evaluation results of the ERT-based model in (**a**) northeast China (NE), (**b**) east China (EC), (**c**) north China (NC), (**d**) south China (SC), (**e**) southwest China (SW), and (**f**) Tibetan Plateau (TP) during 2000-2015.
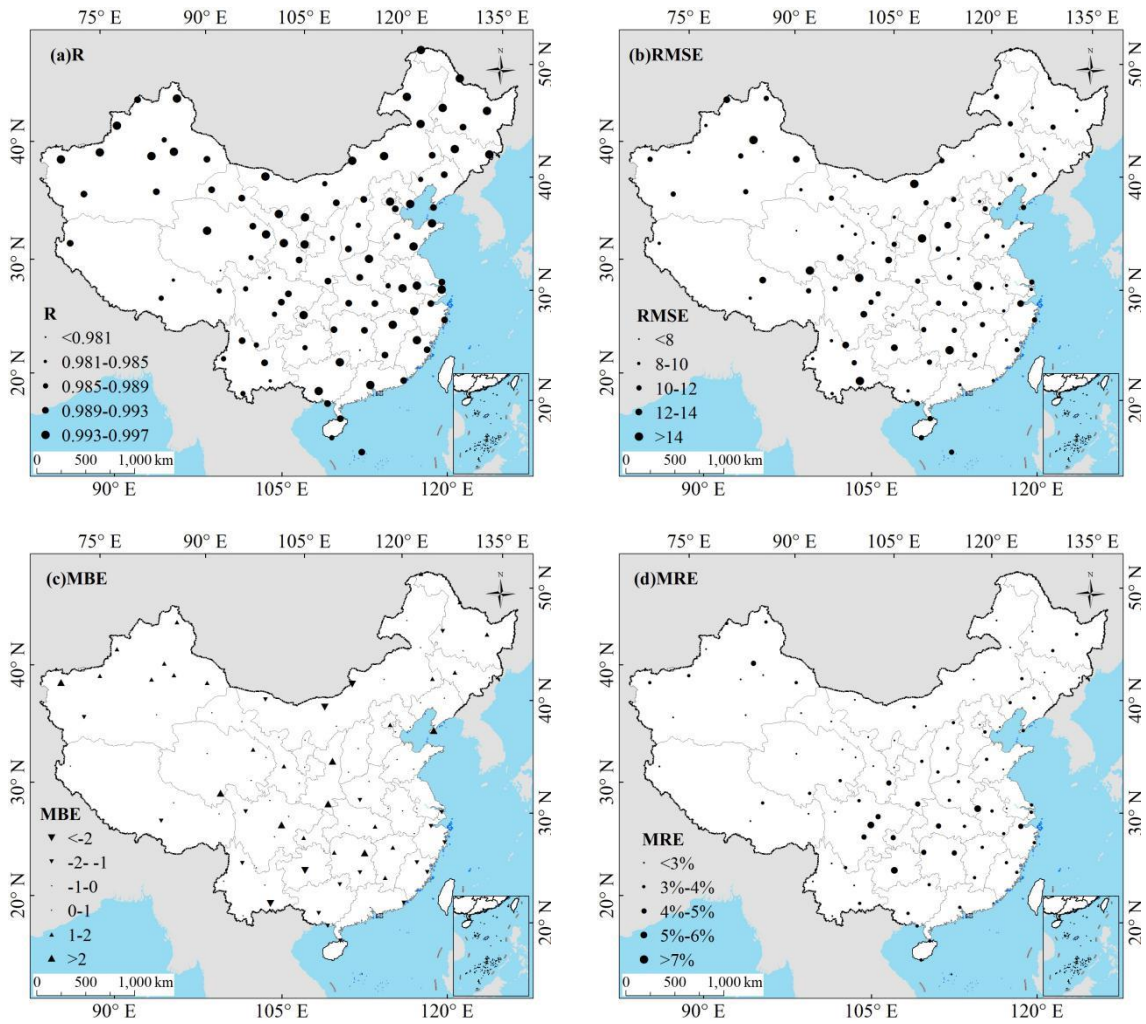

Figure 4. Spatial distribution of the ERT-based model performance on 96 CDC/CMA stations for (**a**) R, (**b**) RMSE, (**c**) mean bias error (MBE), and (**d**) mean relative error (MRE) during 2000–2015.

To assess the performance of the ERT-based model on different temporal scales, the validations on the daily and seasonal timescales were implemented as well. On the daily timescale, the range of the R, RMSE, MBE, and MRE values for the training and test datasets were 0.91–0.99, 2.23–32.97 Wm^{-2}, −5.30–6.86 Wm^{-2}, and 1.26–12.77%, respectively. The ERT-based model had the RMSE, MAE, and MRE values lower than 30 Wm^{-2}, 5 Wm^{-2}, and 12% on all days in the training stage. While in the test stage, these statistics values were lower than 30 Wm^{-2}, 5 Wm^{-2}, and 12% on more than 341 days. On the seasonal timescale, the range of the R, RMSE, MBE, and MRE values for the training and test datasets were 0.94–0.99, 2.75–28.47 Wm^{-2}, −1.48–0.91 Wm^{-2}, and 1.46–10.24%, respectively. The results in Table 1 indicate that the ERT-based model was promising to estimate the *Rs* accurately on both the daily and seasonal timescales.

Figures 5 and 6 demonstrate the estimation results using both the training and test datasets. As displayed in Figure 5, the performance of the ERT-based model was shown as a function of DOY at 96 CDC/CMA stations during 2000–2015. It illustrates that the ERT-based model performed well on most days. For example, the day 365 had relatively higher accuracy with R, RMSE, MBE, and MRE of 0.99, 5.53 Wm^{-2}, 0.06 Wm^{-2}, and 3.02%, respectively. In contrast, the ERT-based model did not give satisfactory *Rs* estimates on day 237 with R, RMSE, MBE, and MRE of 0.98, 13.74 Wm^{-2}, 1.43 Wm^{-2}, and 33.28%, respectively. Overall, the R values were higher than 0.98 on more than 345 days. The RMSE and MBE values were lower than 14 and 1 Wm^{-2} on more than 329 days. The MRE values were lower than 3.5% on more than 330 days. It demonstrates that the ERT-based model was capable of estimating the *Rs* with reasonable accuracy on most days in a year. Figure 6 shows the test results for the daily estimated *Rs* for different seasons during 2000–2015. The ERT-based model provided great accuracy in all seasons, especially in winter (DJF). The mean values of R, RMSE, MBE, MRE were 0.99, 7.39 Wm^{-2}, 0.01 Wm^{-2}, and 3.38% in winter, respectively.
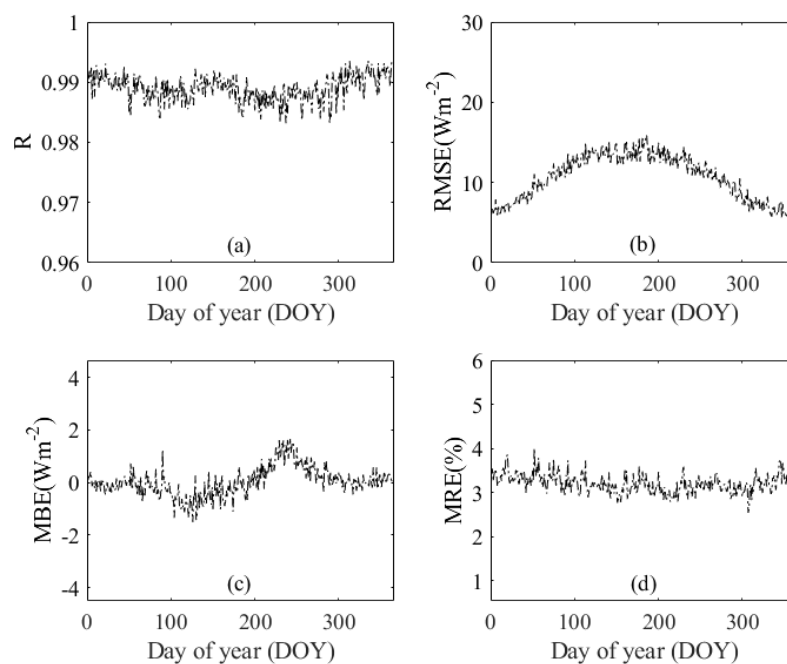

Figure 5. Daily performance of the ERT-based model (**a**) R, (**b**) RMSE, (**c**) MBE, and (**d**) MRE during 2000–2015.
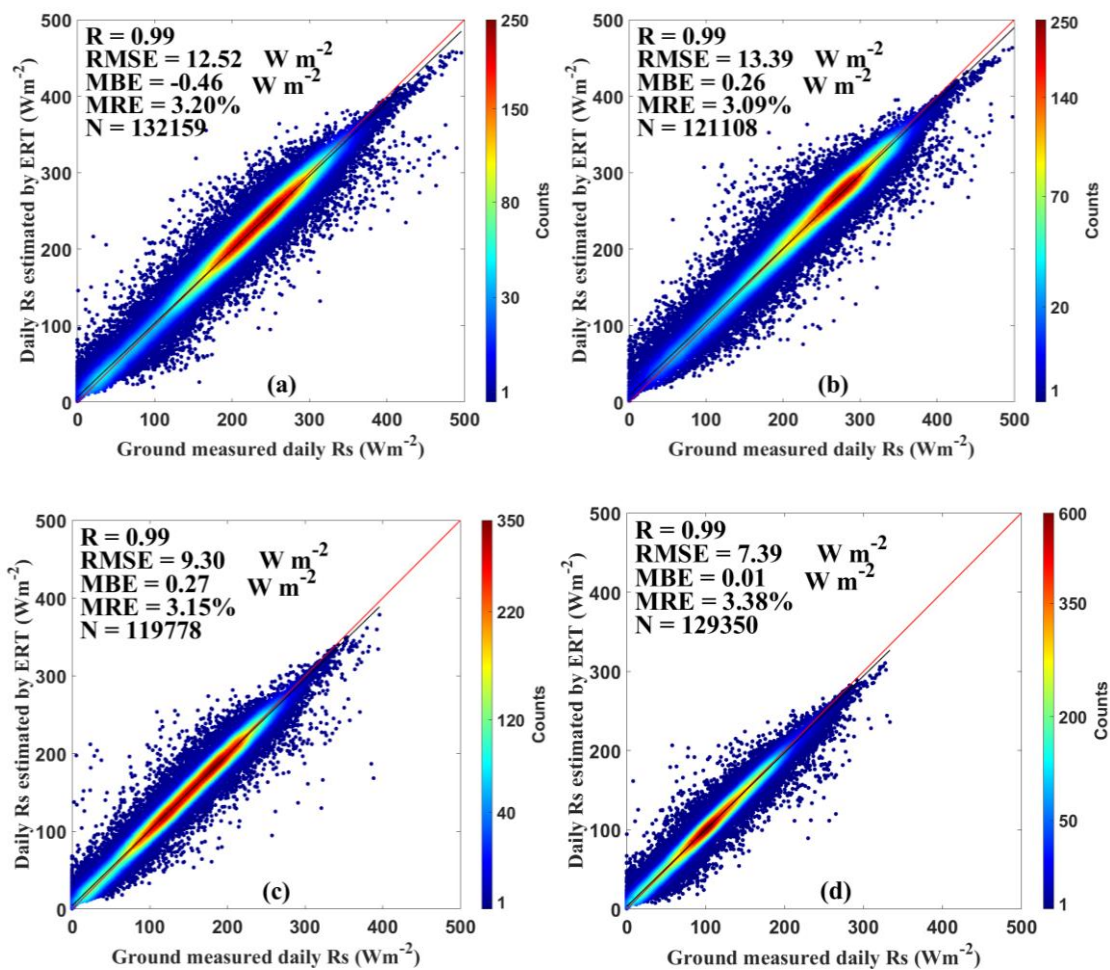
Figure 6. Evaluation results of the ERT-based model in (**a**) spring (March to May), (**b**) summer (June to August), (**c**) autumn (September to November), and (**d**) winter (December to February) over China during 2000–2015.

The validation against ground measurements using all datasets in both training and test stages was also conducted on the synthetic timescales including the monthly, seasonal, and annual timescales (Figure 7). On the monthly timescale, the *Rs* estimates had an R of 0.99, an RMSE of 4.86 Wm^{-2}, an MBE of 0.02 Wm^{-2}, and an MRE of 1.83%. Those values were 0.99, 4.20 Wm^{-2}, 0.01 Wm^{-2}, and 1.59% at the seasonal timescale, and 0.99, 3.43 Wm^{-2}, −0.03 Wm^{-2}, and 1.39% at the annual timescale. These results suggest that synthetic datasets can capture spatiotemporal variations of *Rs* in China more accurately.
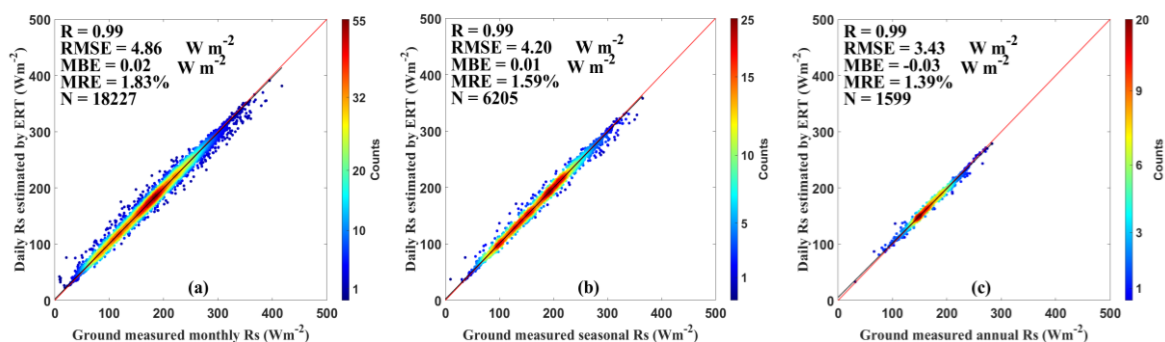
Figure 7. Evaluation results of the ERT-based model on the (**a**) monthly, (**b**) seasonal, and (**c**) annual timescales over China during 2000–2015.

4.2. Spatiotemporal Analysis

4.2.1. Spatial Variations

The above validation results show that the *Rs* estimates based on the ERT method are reasonably accurate compared to the ground measurements. Therefore, the ERT-based model was applied to reconstruct *Rs* data at 756 CDC/CMA stations during 1958–2015. The reconstructed *Rs* data at 454 CDC/CMA stations covering 46 consecutive years (1970–2015) were selected for analyzing trends of *Rs* in China.

The spatial distribution of annual mean *Rs* from 1970 to 2015 is displayed in Figure 8. It shows that *Rs* was higher in TP, NC, and NE than that in SW, EC, and SC. Apart from Sichuan and Guizhou, the *Rs* was spatially decreasing from western to eastern China, and the *Rs* in western China was decreasing with the increase of latitude. The TP had higher *Rs* values than other climatic regions. A previous study shows that the *Rs* in TP was the second highest globally and the highest in China [79]. Strong *Rs* in TP was mainly due to the small amount of cloud [29], rainfall, water vapor content [68], and good air transparency [80]. In contrast, the relatively lower *Rs* values were found in the Sichuan Basin, where the multiple hazes and low atmospheric transparency always happened due to the interactions between the cold and warm currents [81,82].
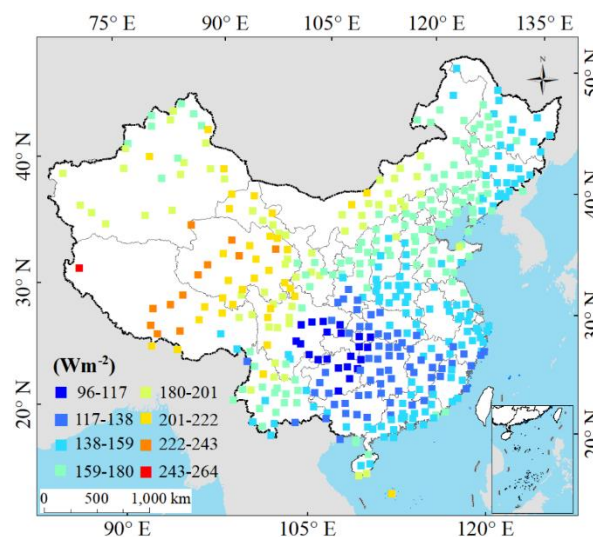

Figure 8. Spatial distribution of the annual mean *Rs* over China during 1970–2015.

4.2.2. Seasonal Variations

The annual and seasonal mean *Rs* were obtained from individual stations with complete records observations. Figure 9 shows that the monthly *Rs* increased slowly from January to June then it declined gradually from July to December. This may relate to the changes of the annual cycle of solar zenith angle and the maximum sunshine duration over China. The monthly *Rs* in July was the largest with the value of 216.52 Wm^{-2}, and the *Rs* in December was the smallest with the value of 92.73 Wm^{-2}. The spatial distribution of the seasonal mean *Rs* over China between 1970 and 2015 is shown Figure 10. The average *Rs* in spring, summer, autumn, and winter were 187.22, 213.28, 143.99 and 105.06 Wm^{-2}, respectively. The seasonal *Rs* trends detected by the M-K test in four seasons from 1970 to 2015 are shown in Table 2. The *Rs* showed an increasing trend in spring and decline trends in summer ($p < 0.01$), autumn ($p < 0.01$), and winter over China. In spring, the *Rs* in NE exhibited a significant decline ($p < 0.01$). Similarly, decline of *Rs* was also found in SW and TP. The seasonal *Rs* showed increasing trends in EC, NC, and SC. In summer, the significant decreasing trends were detected in *Rs* in EC ($p < 0.01$) and NE ($p < 0.05$). The decline tendencies of *Rs* were also found in NC and SC. The seasonal *Rs* showed increasing trends in SW and TP but were not significant. In autumn, the significant decline

of *Rs* was found in EC ($p < 0.01$) and NC ($p < 0.05$). Similarly, the decline tendencies of *Rs* were also found in NE and SC but not significant. The increasing trends of *Rs* were detected in SW and TP. In winter, the *Rs* showed decline in all six climatic regions.
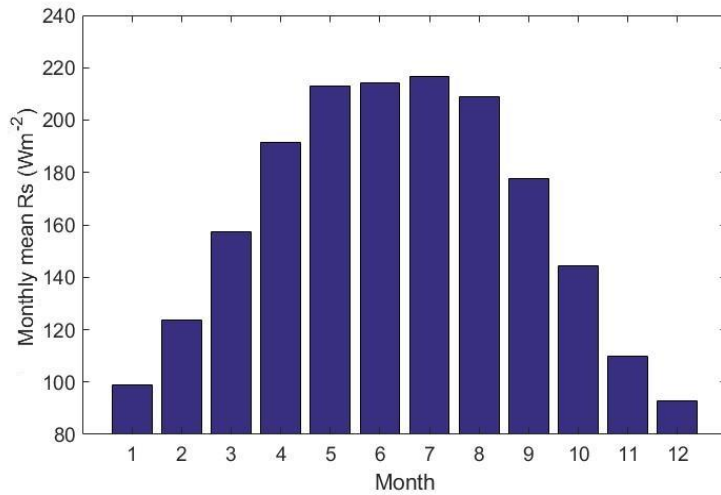

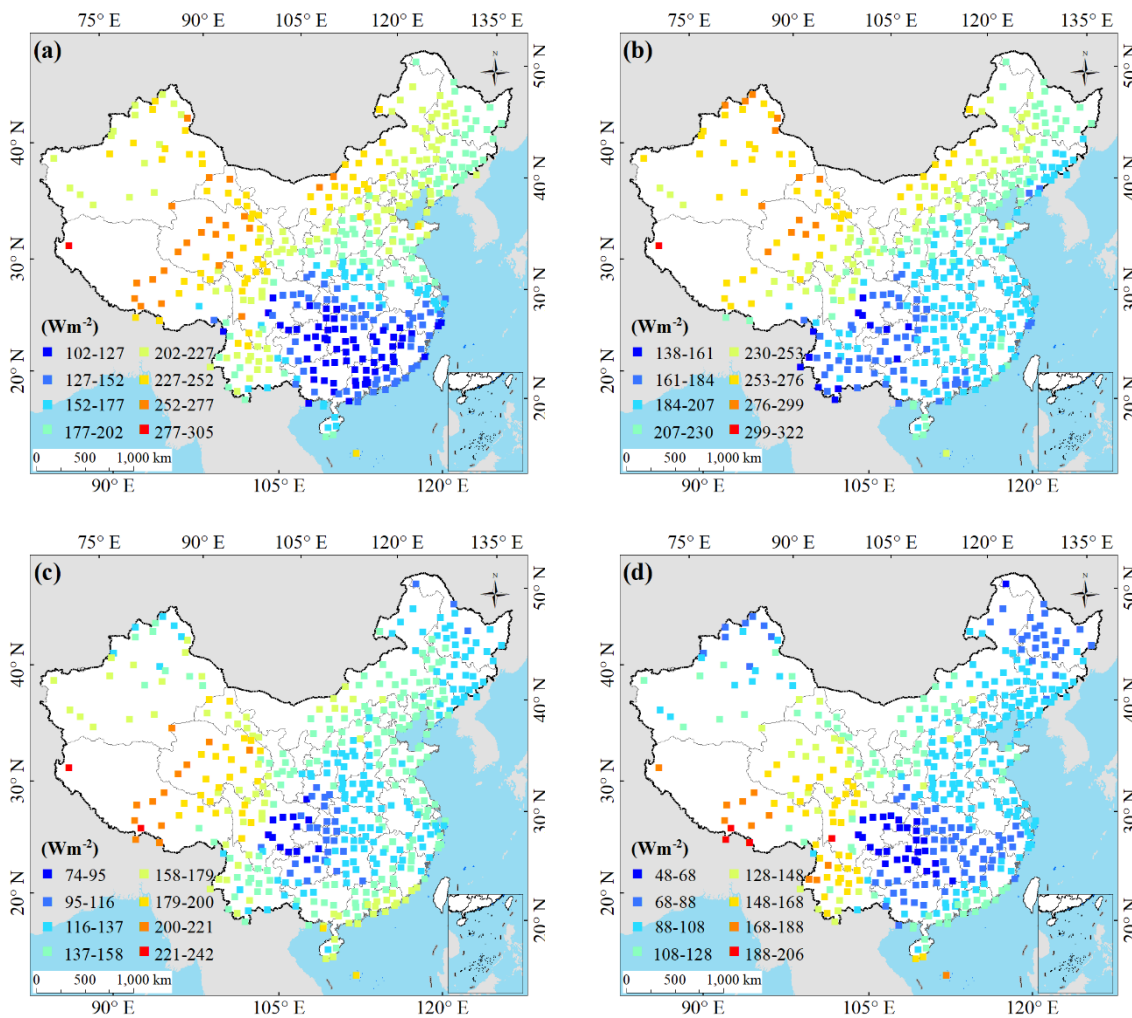**Figure 9.** The monthly mean *Rs* over China during 1970–2015.

Figure 10. Spatial distribution of the seasonal mean *Rs* over China in (**a**) spring, (**b**) summer, (**c**) autumn, and (**d**) winter during 1970–2015.

Table 2. Values of Z for the Mann–Kendall (M-K) test.

Region	Annual (1970–2015)	Annual (1970–1992)	Annual (1992–2015)	Spring	Summer	Autumn	Winter
China	−3.79 **	−3.43 **	0.12	0.09	−4.60 **	−2.63 **	−1.76
EC	−4.11 **	-1.90	−1.81	0.68	−4.56 **	−2.63 **	−1.65
NC	−1.69	−3.27 **	0.77	0.32	−1.72	−2.25 *	−1.93
NE	−3.03 **	−2.59 **	−0.62	−2.58 **	−2.42 *	−1.12	−1.86
SC	−1.16	−1.00	0.87	0.83	−1.31	−0.83	−0.66
SW	1.38	−2.32 *	2.46 *	−0.38	1.40	1.04	−0.49
TP	1.26	−1.11	2.51 *	−0.86	0.73	1.88	−0.57

* Trend at the 95% significance level ($p < 0.05$); ** Trend at the 99% significance level ($p < 0.01$).

4.2.3. Decadal Variations

Previous studies show that the *Rs* in China showed significant decline trends before the 1980s by 2 to 8 Wm^{-2} per decade [25–32], which did not persist into the 1990s. However, the magnitudes of variations in *Rs* were still controversial. The decadal variations of the *Rs* over China during 1970–2015 were analyzed based on the reconstructed *Rs* data at 454 CDC/CMA stations. As given in Figure 11, the *Rs* in China significantly decreased by −1.1 Wm^{-2} per decade ($p < 0.01$) during 1970–2015. The largest annual mean *Rs* value of 169.51 Wm^{-2} appeared in 1978, while the smallest value of 158.72 Wm^{-2} appeared in 1993. The trends of *Rs* during 1970–1992 and 1992-2015 over China were also analyzed. The *Rs* significantly decreased ($p < 0.01$) from 1970 to 1992, by −2.8 Wm^{-2} per decade. While during 1992–2015, the *Rs* exhibited a recovery of 0.23 Wm^{-2} per decade.
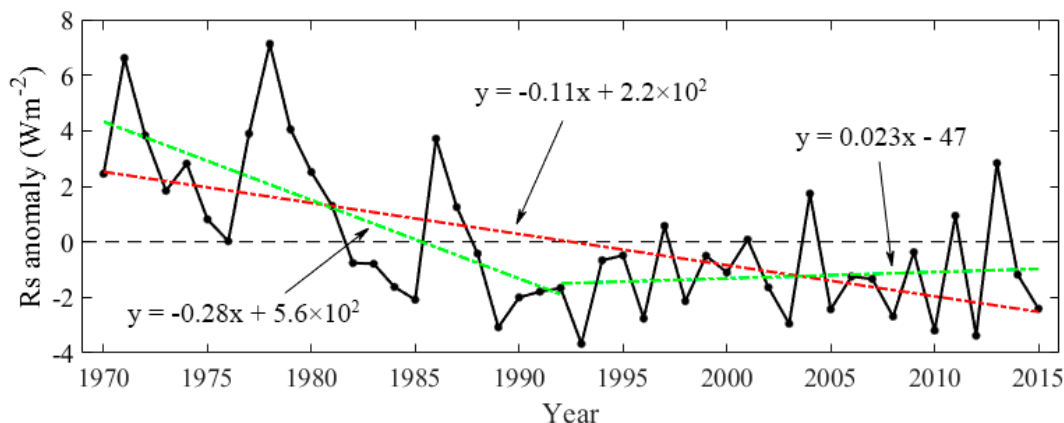

Figure 11. The anomaly of the annual mean *Rs* in China during 1970–2015. The green dashed lines are linear trends of 1970–1992 and 1992–2015.

The decadal variations of *Rs* in six climatic regions during 1970–2015, 1970–1992, and 1992–2015 were also analyzed based on the reconstructed *Rs* dataset. As shown in Table 2 and Figure 12, the significant decreasing trends ($p < 0.01$) were detected in EC and NE during 1970–2015, decreased by −1.8 and −1.2 Wm^{-2} per decade. The *Rs* in NC and SC also showed the decreasing trends (−0.6 and −0.7 Wm^{-2} per decade, respectively) during 1970–2015. Moreover, the increasing trends were detected in SW (0.1 Wm^{-2} per decade) and TP (0.01 Wm^{-2} per decade) during 1970–2015. Between 1970 and 1992, the *Rs* showed significant decline trends ($p < 0.05$) in NC, NE, and SW (−3.3, −2.6, and −1.4 Wm^{-2} per decade, respectively). The decline trends of *Rs* were also detected in EC, SC, and TP (−3.0, −1.7, and −1.0 Wm^{-2} per decade, respectively) but not significant. From 1992 to 2015, the *Rs* in SW and TP showed significant increasing trends (2.4 and 3.6 Wm^{-2} per decade, respectively). Similarly, increasing trends were also found in NC and SC (0.93 and 1.3 Wm^{-2} per decade) but were not significant. However, the *Rs* showed decreasing tendencies of −1.2 and −0.75 Wm^{-2} per decade in EC and NE, respectively.
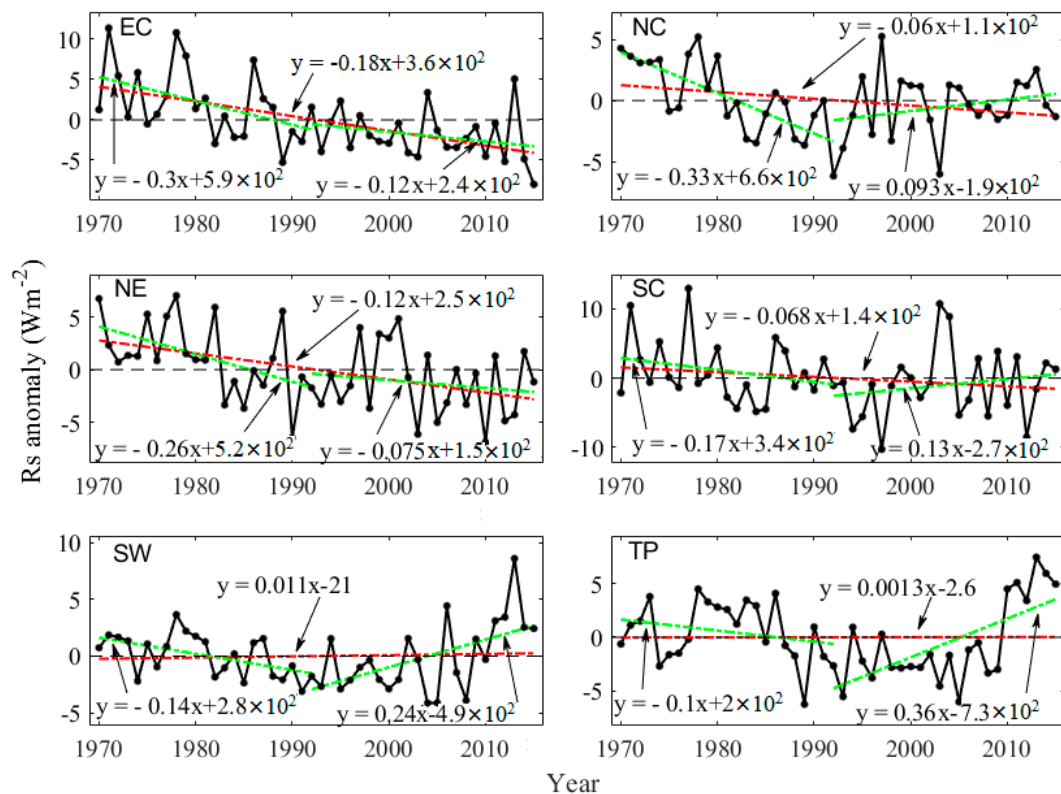
Figure 12. The anomaly of the annual mean *Rs* in six climatic regions during 1970–2015. The green dashed lines are linear trends of 1970–1992 and 1992–2015.

The decadal variations of *Rs* at 454 CDC/CMA stations were shown in Figure 13. Among 454 CDC/CMA stations used in this study, there were 63, 147, 95, 42, 66, and 41 stations in NE, EC, NC, SC, SW, and TP, respectively. The *Rs* exhibited the decreasing trends at 317 (70%) stations during 1970-2015 (Figure 13a). During 1970-2015, the *Rs* showed decline at 52 (83%), 125 (85%), 61 (64%), 30 (71%), 32 (48%), and 17 (41%) stations in NE, EC, NC, SC, SW, and TP, respectively. The *Rs* exhibited decline at more than 64% of the stations in NE, EC, NC, SC. On the contrary, the *Rs* showed the increasing trends at more than 50% of the stations in SW and TP. From 1970 to 1992 (Figure 13b), the *Rs* showed the decreasing trends at 378 (83%) stations. The *Rs* showed the decline trends at 58 (92%), 128 (87%), 78 (82%), 27 (64%), 59 (89%), and 28 (68%) stations in NE, EC, NC, SC, SW, and TP, respectively. The *Rs* exhibited the decline trends at more than 64% of the stations in six climatic regions. Between 1992 and 2015 (Figure 13c), the *Rs* showed the increasing trends at 233 (51%) stations over China, including 20 (32%), 55 (37%), 60 (63%), 27 (64%), 51 (77%), and 20 (49%) stations in NE, EC, NC, SC, SW, and TP, respectively. It shows that the increasing trends of *Rs* mainly concentrated in NC, SC, SW, and TP during 1992–2015. All in all, the *Rs* showed the decline trends at more than 70% of the CDC/CMA stations during 1970–1992 and 1970–2015, while it showed the increasing trends at more than 51% of the stations during 1992–2015.
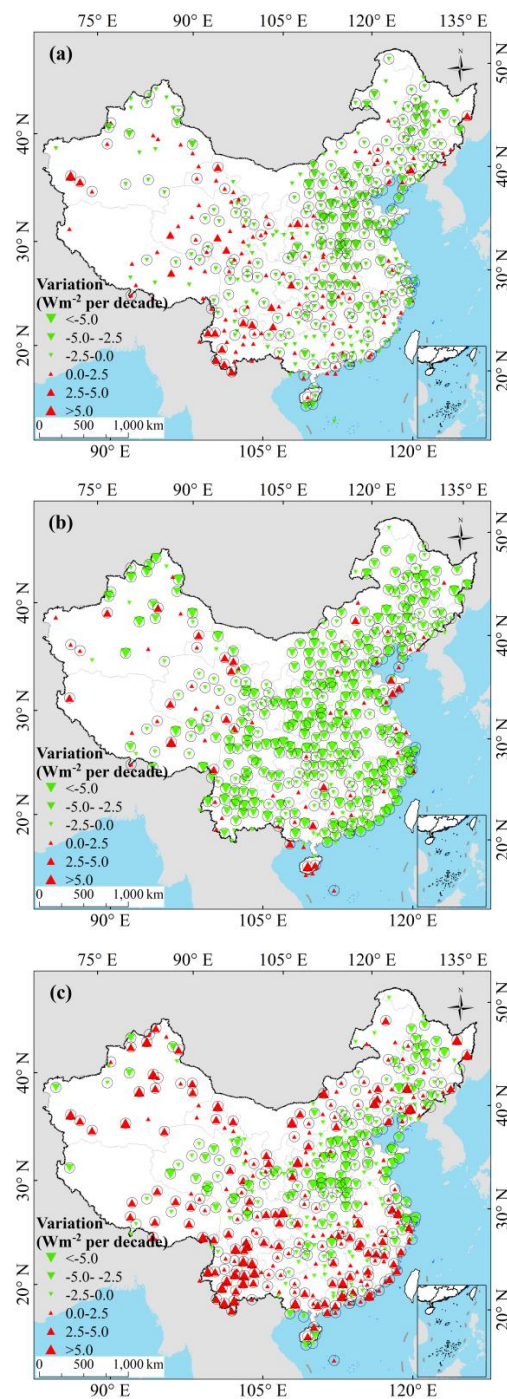
Figure 13. Decadal variation of the annual mean *Rs* on 454 stations during (**a**) 1970–2015, (**b**) 1970–1992, and (**c**) 1992–2015. Trends at the 95% significance level ($p < 0.05$) are denoted by five-pointed black circle symbols.

5. Discussion

To investigate the influence of input variables of the ERT-based model on the estimation results, the variable importance measures provided by the ERT method was conducted. As shown in Table 3, the importance of the input variables of the ERT-based model was in the order of sunshine duration, cosine of the radian difference, air pressure, water vapor pressure, relative humidity, elevation, air temperature, wind speed and daily precipitation. The sunshine duration had significant influence

on estimating *Rs*, while the wind speed and daily precipitation had relative low influence on the *Rs* estimation.

Table 3. Variable importance measure.

Variable	Importance	Variable	Importance
Sunshine duration	0.47	Elevation	0.04
Cosine of the radian difference	0.16	Air temperature	0.04
Air pressure	0.13	Wind speed	0.01
Water vapor pressure	0.07	Daily precipitation	0.01
Relative humidity	0.07		

The above analysis reveals significant declining tendencies of the *Rs* in China during 1970–2015. Meanwhile, there was a significant decline during 1970–1992, and a recovery during 1992–2015. Previous studies pointed that the sunshine duration decreased by −0.17 h per decade in China during 1960–2000 [83,84], the variations of the *Rs* in China may relate to the variations of sunshine duration. The variable importance measures indicate that the sunshine duration had significant influence on the *Rs* estimation. To further investigate probable causes of changes in *Rs*, the comparison of the variations of sunshine duration and *Rs* was conducted. Figure 14 shows the anomalies of the annual mean *Rs* and sunshine duration in China during 1970–2015, 1970–1992, and 1992–2015, respectively. It is clear that annual variations of the anomaly of *Rs* were almost in line with variations of sunshine duration from 1970 to 2015, before 1992 the sunshine duration exhibited a decreasing trend (0.18 h per decade), and an increasing trend (0.18 h per decade) afterward. As shown in Figure 15, it illustrates that the trends of sunshine duration were consistent with the trends of the *Rs* in six climatic regions from 1970 to 2015. The consistency of variations in sunshine duration and *Rs* indicates that the variations of *Rs* were likely due to long-term variations of sunshine duration. In the most recent years, a slight difference between the sunshine duration and the *Rs* was observed. This probably contributes to the influence of the aerosol variations on *Rs* according to the previous studies [37,48,85–88]. For instance, Jia et al. [87] reported that the variations in *Rs* is likely due to the variations in the aerosols, cloud, and water vapor in north and south China. Wang and Pinker [85] found that the global average aerosol showed an increasing trend during 1979–2006, particularly in east and south Asia (including China). The increase in global average aerosol could be associated with a continuous decrease of *Rs* in China from 1979 to 2006. Qin et al. [86] pointed out that *Rs* over China declined gradually between 1980 and 2015, which may relate to the increasing aerosol radiative forcing effects over China in recent decades. It is clear that the spatiotemporal variations in *Rs* over China were complicated due to the effects of natural factors and human activities. The influence of other factors (e.g., aerosols, cloud, and topography) on *Rs* estimation are still needed to be investigated in future research.
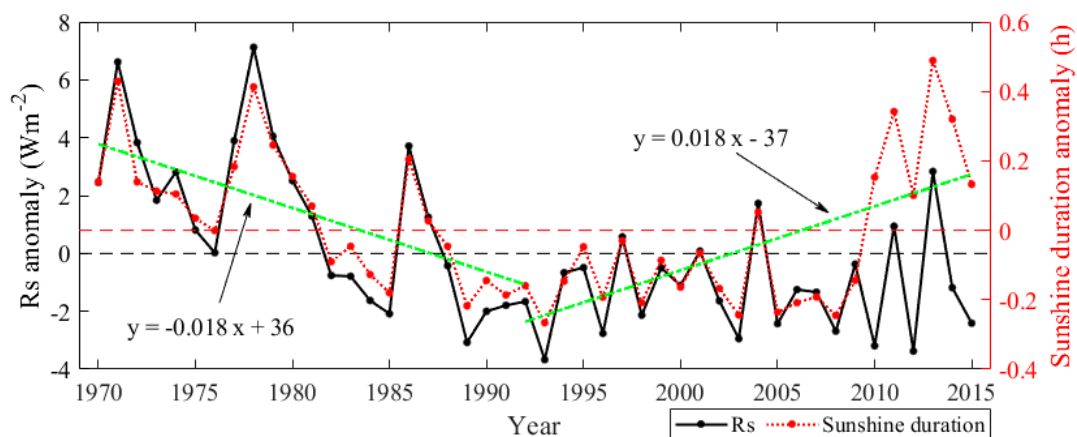

Figure 14. Comparison of the anomaly of the annual mean *Rs* and sunshine duration between 1970 and 2015. The green dashed lines are linear trends of 1970–1992 and 1992–2015.
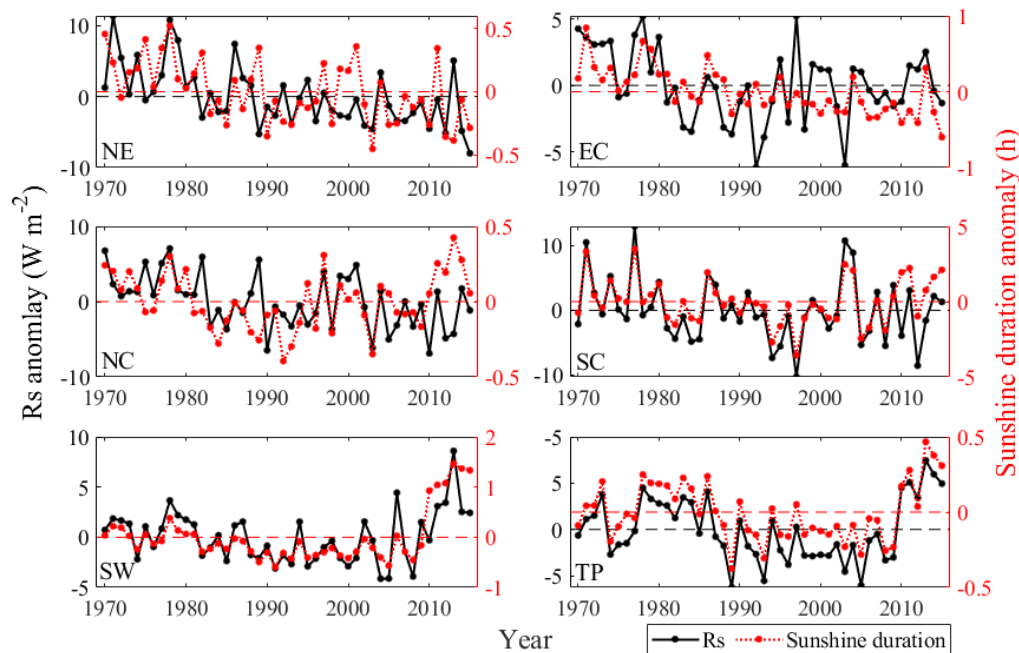
Figure 15. Comparison of the anomaly of the annual mean *Rs* and sunshine duration between 1970 and 2015 in six climatic regions.

6. Conclusions

In this study, an *Rs* estimation model was constructed employing the ERT algorithm based on the quality-controlled daily *Rs* data and routine meteorological data. The accuracy and applicability of the ERT-based model in estimating daily *Rs* over China were investigated. A new daily *Rs* dataset was developed at 756 CDC/CMA stations from 1958 to 2015. Moreover, the spatial and temporal variations of the *Rs* at 454 CDC/CMA stations covering 46 consecutive years (1970–2015) were analyzed. The results indicate that the ERT-based model estimated well the *Rs* at the national scale with R of 0.97, RMSE of 23.12 Wm^{-2}, MBE of 0.04 Wm^{-2}, and MRE of 9.81%. The ERT-based model also performed well in estimating the *Rs* at most CDC/CMA stations and individual days. Furthermore, the spatiotemporal variations in the *Rs* over China were investigated using the M-K test based on the reconstructed dataset. During 1970–2015, the annual mean *Rs* in China showed a significant decline ($p < 0.01$) by −1. 1 Wm^{-2} per decade. The *Rs* decreased significantly ($p < 0.01$) in EC and NE by −1.8 and −1.2 Wm^{-2} per decade. The *Rs* in NC and SC also showed decreasing trends while the *Rs* in SW and TP showed increasing trends. The *Rs* showed decreasing trends at 317 out of 454 CDC/CMA stations. During 1970–1992, the *Rs* showed a significant decreasing trend (−2.8 Wm^{-2} per decade) in China. The declined trends of *Rs* were detected at 378 out of 454 CDC/CMA stations. During 1992–2015, the *Rs* in China showed a recovery (0.23 Wm^{-2} per decade). The increasing trends of *Rs* were detected at 233 out of 454 CDC/CMA stations, which are mainly concentrated in NC, SC, SW, and TP. The comparison of *Rs* and sunshine duration illustrates that the variations of the *Rs* in China may relate to the variations in sunshine duration. More ground measurements of *Rs* and other important factors including aerosols, clouds, water vapor should be collected for further studies to investigate the variations of *Rs* over China.

Overall, the reconstruction of new daily long-term *Rs* with high accuracy would be a useful data source for the related climate change studies. More attention needs to be paid to quantitative analysis of *Rs* and other climatic factors.

References

1. Gupta, S.K.; Ritchey, N.A.; Wilber, A.C.; Whitlock, C.H.; Gibson, G.G.; Stackhouse, P.W. A climatology of surface radiation budget derived from satellite data. *J. Clim.* **1999**, *12*, 2691–2710. [CrossRef]

2. Liang, S.; Wang, K.; Zhang, X.; Wild, M. Review on Estimation of Land Surface Radiation and Energy Budgets from Ground Measurement, Remote Sensing and Model Simulations. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2010**, *3*, 225–240. [CrossRef]

3. Hinkelman, L.M.; Jr. Stackhouse, P.W.; Wielicki, B.A.; Zhang, T.; Wilson, S.R. Surface insolation trends from satellite and ground measurements: Comparisons and challenges. *J. Geophys. Res. Atmos.* **2009**, *114*. [CrossRef]

4. Mueller, R.W.; Matsoukas, C.; Gratzki, A.; Behr, H.D.; Hollmann, R. The CM-SAF operational scheme for the satellite based retrieval of solar surface irradiance-A LUT based eigenvector hybrid approach. *Remote Sens. Environ.* **2009**, *113*, 1012–1024. [CrossRef]

5. Wild, M.; Folini, D.; Schaer, C.; Loeb, N.; Dutton, E.G.; Koenig-Langlo, G. The global energy balance from a surface perspective. *Clim. Dyn.* **2013**, *40*, 3107–3134. [CrossRef]

6. Feng, F.; Wang, K. Determining Factors of Monthly to Decadal Variability in Surface Solar Radiation in China: Evidences from Current Reanalyses. *J. Geophys. Res. Atmos.* **2019**, *124*, 9161–9182. [CrossRef]

7. Benghanem, M.; Mellit, A. Radial Basis Function Network-based prediction of global solar radiation data: Application for sizing of a stand-alone photovoltaic system at Al-Madinah, Saudi Arabia. *Energy* **2010**, *35*, 3751–3762. [CrossRef]

8. Cline, D.W.; Bales, R.C.; Dozier, J. Estimating the spatial distribution of snow in mountain basins using remote sensing and energy balance modeling. *Water Resour. Res.* **1998**, *34*, 1275–1285. [CrossRef]

9. Cooter, E.J.; Dhakhwa, G.B. A solar radiation model for use in biological applications in the South and Southeastern USA. *Agric. For. Meteorol.* **1996**, *78*, 31–51. [CrossRef]

10. Hoogenboom, G. Contribution of agrometeorology to the simulation of crop production and its applications. *Agric. For. Meteorol.* **2000**, *103*, 137–157. [CrossRef]

11. Pohlert, T. Use of empirical global radiation models for maize growth simulation. *Agric. For. Meteorol.* **2004**, *126*, 47–58. [CrossRef]

12. Wild, M.; Ohmura, A.; Schar, C.; Muller, G.; Folini, D.; Schwarz, M.; Hakuba, M.Z.; Sanchez-Lorenzo, A. The global energy balance archive (GEBA) version 2017: A database for worldwide measured surface energy fluxes. *Earth Syst. Sci. Data* **2017**, *9*, 601–613. [CrossRef]

13. Stanhill, G. The Distribution of Global Solar-Radiation over the Land Surfaces of the Earth. *Solar Energy* **1983**, *31*, 95–104. [CrossRef]

14. Gilgen, H.; Roesch, A.; Wild, M.; Ohmura, A. Decadal changes in shortwave irradiance at the surface in the period from 1960 to 2000 estimated from Global Energy Balance Archive Data. *J. Geophys. Res. Atmos.* **2009**, *114*. [CrossRef]

15. Pinker, R.T.; Zhang, B.; Dutton, E.G. Do satellites detect trends in surface solar radiation? *Science* **2005**, *308*, 850–854. [CrossRef] [PubMed]

16. Wild, M. Global dimming and brightening: A review. *J. Geophys. Res. Atmos.* **2009**, *114*. [CrossRef]

17. Liepert, B.G. Observed reductions of surface solar radiation at sites in the United States and worldwide from 1961 to 1990. *Geophys. Res. Lett.* **2002**, *29*. [CrossRef]

18. Wild, M.; Gilgen, H.; Roesch, A.; Ohmura, A.; Long, C.N.; Dutton, E.G.; Forgan, B.; Kallis, A.; Russak, V.; Tsvetkov, A. From dimming to brightening: Decadal changes in solar radiation at Earth's surface. *Science* **2005**, *308*, 847–850. [CrossRef]

19. Norris, J.R.; Wild, M. Trends in aerosol radiative effects over Europe inferred from observed cloud cover, solar "dimming" and solar "brightening". *J. Geophys. Res. Atmos.* **2007**, *112*. [CrossRef]

20. Kambezidis, H.D. The solar radiation climate of Athens: Variations and tendencies in the period 1992-2017, the brightening era. *Sol. Energy* **2018**, *173*, 328–347. [CrossRef]

21. Long, C.N.; Dutton, E.G.; Augustine, J.A.; Wiscombe, W.; Wild, M.; McFarlane, S.A.; Flynn, C.J. Significant decadal brightening of downwelling shortwave in the continental United States. *J. Geophys. Res. Atmos.* **2009**, *114*. [CrossRef]

22. Augustine, J.A.; Dutton, E.G. Variability of the surface radiation budget over the United States from 1996 through 2011 from high-quality measurements. *J. Geophys. Res. Atmos.* **2013**, *118*, 43–53. [CrossRef]

23. Ohmura, A. Observed decadal variations in surface solar radiation and their causes. *J. Geophys. Res. Atmos.* **2009**, *114*. [CrossRef]

24. Norris, J.R.; Wild, M. Trends in aerosol radiative effects over China and Japan inferred from observed cloud cover, solar "dimming" and solar "brightening". *J. Geophys. Res.* **2009**, *114*. [CrossRef]

25. Yang, S.; Wang, X.L.L.; Wild, M. Homogenization and trend analysis of the 1958–2016 in situ surface solar radiation records in China. *J. Clim.* **2018**, *31*, 4529–4541. [CrossRef]

26. Wang, Y.W.; Trentmann, J.; Pfeifroth, U.; Yuan, W.P.; Wild, M. Improvement of air pollution in china inferred from changes between satellite-based and measured surface solar radiation. *Remote Sens.* **2019**, *11*, 2910. [CrossRef]

27. Yang, S.; Wang, X.L.L.; Wild, M. Causes of dimming and brightening in china inferred from homogenized daily clear-sky and all-sky in situ surface solar radiation records (1958–2016). *J. Clim.* **2019**, *32*, 5901–5913. [CrossRef]

28. Schwarz, M.; Folini, D.; Yang, S.; Allan, R.P.; Wild, M. Changes in atmospheric shortwave absorption as important driver of dimming and brightening. *Nat. Geosci.* **2020**, *13*, 110–115. [CrossRef]

29. Che, H.Z.; Shi, G.Y.; Zhang, X.Y.; Arimoto, R.; Zhao, J.Q.; Xu, L.; Wang, B.; Chen, Z.H. Analysis of 40 years of solar radiation data from China, 1961-2000. *Geophys. Res. Lett.* **2005**, *32*. [CrossRef]

30. Qian, Y.; Kaiser, D.P.; Leung, L.R.; Xu, M. More frequent cloud-free sky and less surface solar radiation in China from 1955 to 2000. *Geophys. Res. Lett.* **2006**, *33*. [CrossRef]

31. Xia, X. A closer looking at dimming and brightening in China during 1961–2005. *Ann. Geophys.* **2010**, *28*, 1121–1132. [CrossRef]

32. Wang, K.; Ye, H.; Chen, F.; Xiong, Y.; Wang, C. Urbanization effect on the diurnal temperature range: Different roles under solar dimming and brightening*. *J. Clim.* **2012**, *25*, 1022–1027. [CrossRef]

33. Wang, K.C. Measurement biases explain discrepancies between the observed and simulated decadal variability of surface incident solar radiation. *Sci. Rep. UK* **2014**, *4*. [CrossRef] [PubMed]

34. Manara, V.; Beltrano, M.C.; Brunetti, M.; Maugeri, M.; Sanchez-Lorenzo, A.; Simolo, C.; Sorrenti, S. Sunshine duration variability and trends in Italy from homogenized instrumental time series (1936–2013). *J. Geophys. Res. Atmos.* **2015**, *120*, 3622–3641. [CrossRef]

35. Wang, K.; Dickinson, R.E.; Ma, Q.; Augustine, J.A.; Wild, M. Measurement Methods Affect the Observed Global Dimming and Brightening. *J. Clim.* **2013**, *26*, 4112–4120. [CrossRef]

36. Wang, K.; Dickinson, R.E.; Wild, M.; Liang, S. Atmospheric impacts on climatic variability of surface incident solar radiation. *Atmos. Chem. Phys.* **2012**, *12*, 9581–9592. [CrossRef]

37. Zhang, X.; Liang, S.; Wild, M.; Jiang, B. Analysis of surface incident shortwave radiation from four satellite products. *Remote Sens. Environ.* **2015**, *165*, 186–202. [CrossRef]

38. Kumar, R.; Aggarwal, R.K.; Sharma, J.D. Comparison of regression and artificial neural network models for estimation of global solar radiations. *Renew. Sustain. Energy Rev.* **2015**, *52*, 1294–1299. [CrossRef]

39. Meza, F.; Varas, E. Estimation of mean monthly solar global radiation as a function of temperature. *Agric. Forest Meteorol.* **2000**, *100*, 231–241. [CrossRef]

40. Tang, W.; Yang, K.; Qin, J.; Min, M. Development of a 50-year daily surface solar radiation dataset over China. *Sci. China Earth Sci.* **2013**, *56*, 1555–1565. [CrossRef]

41. Tang, W.J.; Qin, J.; Yang, K.; Niu, X.L.; Zhang, X.T.; Yu, Y.; Zhu, X.D. Reconstruction of daily photosynthetically active radiation and its trends over China. *J. Geophys. Res. Atmos.* **2013**, *118*, 13292–13302. [CrossRef]

42. Yang, K.; Koike, T.; Ye, B.S. Improving estimation of hourly, daily, and monthly solar radiation by importing global data sets. *Agric. For. Meteorol.* **2006**, *137*, 43–55. [CrossRef]

43. Tang, W.; Yang, K.; He, J.; Qin, J. Quality control and estimation of global solar radiation in China. *Sol. Energy* **2010**, *84*, 466–475. [CrossRef]

44. Ghayekhloo, M.; Ghofrani, M.; Menhaj, M.B.; Azimi, R. A novel clustering approach for short-term solar radiation forecasting. *Sol. Energy* **2015**, *122*, 1371–1383. [CrossRef]

45. Dos Santos, C.M.; De Souza, J.L.; Ferreira Junior, R.A.; Tiba, C.; de Melo, R.O.; Lyra, G.B.; Teodoro, I.; Lyra, G.B.; Lemes, M.A.M. On modeling global solar irradiation using air temperature for Alagoas State, Northeastern Brazil. *Energy* **2014**, *71*, 388–398. [CrossRef]

46. Katiyar, A.K.; Pandey, C.K. Simple correlation for estimating the global solar radiation on horizontal surfaces in India. *Energy* **2010**, *35*, 5043–5048. [CrossRef]

47. Qin, J.; Tang, W.J.; Yang, K.; Lu, N.; Niu, X.L.; Liang, S.L. An efficient physically based parameterization to derive surface solar irradiance based on satellite atmospheric products. *J. Geophys. Res. Atmos.* **2015**, *120*, 4975–4988. [CrossRef]

48. Piri, J.; Shamshirband, S.; Petkovic, D.; Tong, C.W.; Rehman, M.H.U. Prediction of the solar radiation on the Earth using support vector regression technique. *Infrared Phys. Technol.* **2015**, *68*, 179–185. [CrossRef]

49. Yao, W.; Zhang, C.; Hao, H.; Wang, X.; Li, X. A support vector machine approach to estimate global solar radiation with the influence of fog and haze. *Renew. Energy* **2018**, *128*, 155–162. [CrossRef]

50. Zhou, Z.; Lin, A.; Wang, L.; Qin, W.; Zhong, Y.; He, L. Trends in downward surface shortwave radiation from multi-source data over China during 1984–2015. *Int. J. Climatol.* **2019**, *40*, 3467–3485. [CrossRef]

51. Ruiz-Arias, J.A.; Gueymard, C.A.; Santos-Alamillos, F.J.; Quesada-Ruiz, S.; Pozo-Vázquez, D. Bias induced by the AOD representation time scale in long-term solar radiation calculations. Part 2: Impact on long-term solar irradiance predictions. *Sol. Energy* **2016**, *135*, 625–632. [CrossRef]

52. Ruiz-Arias, J.A. Bias in modeled solar radiation by non-resolved intra-daily AOD variability. *Sol. Energy* **2020**, *205*, 221–229. [CrossRef]

53. Fan, J.L.; Wang, X.K.; Wu, L.F.; Zhou, H.M.; Zhang, F.C.; Yu, X.; Lu, X.H.; Xiang, Y.Z. Comparison of Support Vector Machine and Extreme Gradient Boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in China. *Energy Convers. Manag.* **2018**, *164*, 102–111. [CrossRef]

54. Wang, L.C.; Kisi, O.; Zounemat-Kermani, M.; Zhu, Z.M.; Gong, W.; Niu, Z.G.; Liu, H.F.; Liu, Z.J. Prediction of solar radiation in China using different adaptive neuro-fuzzy methods and M5 model tree. *Int. J. Climatol.* **2017**, *37*, 1141–1155. [CrossRef]

55. Wei, Y.; Zhang, X.; Hou, N.; Zhang, W.; Jia, K.; Yao, Y. Estimation of surface downward shortwave radiation over China from AVHRR data based on four machine learning methods. *Sol. Energy* **2019**, *177*, 32–46. [CrossRef]

56. Hou, N.; Zhang, X.T.; Zhang, W.Y.; Wei, Y.; Jia, K.; Yao, Y.J.; Jiang, B.; Cheng, J. Estimation of surface downward shortwave radiation over china from himawari-8 ahi data based on random forest. *Remote Sens.* **2020**, *12*, 181. [CrossRef]

57. Yang, L.; Zhang, X.T.; Liang, S.L.; Yao, Y.J.; Jia, K.; Jia, A.L. Estimating surface downward shortwave radiation over china based on the gradient boosting decision tree method. *Remote Sens.* **2018**, *10*, 185. [CrossRef]

58. Wang, T.; Yan, G.; Chen, L. Consistent retrieval methods to estimate land surface shortwave and longwave radiative flux components under clear-sky conditions. *Remote Sens. Environ.* **2012**, *124*, 61–71. [CrossRef]

59. Rodríguez-Benítez, F.J.; Arbizu-Barrena, C.; Huertas-Tato, J.; Aler-Mur, R.; Galván-León, I.; Pozo-Vázquez, D. A short-term solar radiation forecasting system for the Iberian Peninsula. Part 1: Models description and performance assessment. *Sol. Energy* **2020**, *195*, 396–412. [CrossRef]

60. Cornejo-Bueno, L.; Casanova-Mateo, C.; Sanz-Justo, J.; Salcedo-Sanz, S. Machine learning regressors for solar radiation estimation from satellite data. *Sol. Energy* **2019**, *183*, 768–775. [CrossRef]

61. Hao, D.; Asrar, G.R.; Zeng, Y.; Zhu, Q.; Wen, J.; Xiao, Q.; Chen, M. Estimating hourly land surface downward shortwave and photosynthetically active radiation from DSCOVR/EPIC observations. *Remote Sens. Environ.* **2019**, *232*, 111320. [CrossRef]

62. Ibrahim, I.A.; Khatib, T. A novel hybrid model for hourly global solar radiation prediction using random forests technique and firefly algorithm. *Energy Convers. Manag.* **2017**, *138*, 413–425. [CrossRef]

63. Wehenkel, L.; Ernst, D.; Geurts, P. Ensembles of extremely randomized trees and some generic applications. In Proceedings of the Robust Methods for Power System State Estimation and Load Forecasting, Versailles, France, 29–30 May 2006.

64. Moradi, I. Quality control of global solar radiation using sunshine duration hours. *Energy* **2009**, *34*, 1–6. [CrossRef]

65. Wei, J.; Li, Z.; Cribb, M.; Huang, W.; Xue, W.; Sun, L.; Guo, J.; Peng, Y.; Li, J.; Lyapustin, A.; et al. Improved 1 km resolution PM2.5 estimates across China using enhanced space–time extremely randomized trees. *Atmos. Chem. Phys.* **2020**, *20*, 3273–3289. [CrossRef]

66. Lam, J.C.; Wan, K.K.W.; Yang, L. Solar radiation modelling using ANNs for different climates in China. *Energy Convers. Manag.* **2008**, *49*, 1080–1090. [CrossRef]

67. Lau, C.C.S.; Lam, J.C.; Yang, L. Climate classification and passive solar design implications in China. *Energy Convers Manag.* **2007**, *48*, 2006–2015. [CrossRef]

68. Wang, C.H.; Zhang, Z.F.; Tian, W.S. Factors affecting the surface radiation trends over China between 1960 and 2000. *Atmos. Env.* **2011**, *45*, 2379–2385. [CrossRef]

69. Zhou, Z.G.; Wang, L.C.; Lin, A.W.; Zhang, M.; Niu, Z.G. Innovative trend analysis of solar radiation in China during 1962-2015. *Renew. Energy* **2018**, *119*, 675–689. [CrossRef]

70. Geurts, P.; Wehenkel, L. Investigation and reduction of discretization variance in decision tree induction. *Lect. Notes Artif. Int.* **2000**, *1810*, 162–170.

71. Wehenkel, L. Discretization of continuous attributes for supervised learning. variance evaluation and variance reduction. *Proc. Seventh Int. Fuzzy Syst. Assoc. World Congr.* **2012**, *1*, 413–418.

72. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

73. Geurts, P.; Ernst, D.; Wehenkel, L. Extremely randomized trees. *Mach. Learn.* **2006**, *63*, 3–42. [CrossRef]

74. Basith, S.; Manavalan, B.; Shin, T.H.; Lee, G. iGHBP: Computational identification of growth hormone binding proteins from sequences using extremely randomised tree. *Comput. Struct. Biotechnol. J.* **2018**, *16*, 412–420. [CrossRef]

75. Manavalan, B.; Govindaraj, R.G.; Shin, T.H.; Kim, M.O.; Lee, G. iBCE-EL: A new ensemble learning framework for improved linear b-cell epitope prediction. *Front. Immunol.* **2018**, *9*, 1695. [CrossRef] [PubMed]

76. Manavalan, B.; Basith, S.; Shin, T.H.; Wei, L.; Lee, G. AtbPpred: A robust sequence-based prediction of anti-tubercular peptides using extremely randomized trees. *Comput. Struct. Biotechnol. J.* **2019**, *17*, 972–981. [CrossRef] [PubMed]

77. Mann, H.B. Non–parametric tests against trend. *Econometrica* **1945**, *13*, 245–259. [CrossRef]

78. Kendall, M.G. *Rank Correlation Methods*; Charles Griffin: London, UK, 1975; p. 220.

79. Li, H.S.; Ma, W.B.; Lian, Y.W.; Wang, X.L.; Zhao, L. Global solar radiation estimation with sunshine duration in Tibet, China. *Renew. Energy* **2011**, *36*, 3141–3145. [CrossRef]

80. Zhang, X.Z.; Zhang, Y.G.; Zhoub, Y.H. Measuring and modelling photosynthetically active radiation in Tibet Plateau during April-October. *Agric. Forest Meteorol.* **2000**, *102*, 207–212. [CrossRef]

81. Feng, Y.; Li, Y.T. Estimated spatiotemporal variability of total, direct and diffuse solar radiation across China during 1958–2016. *Int. J. Climatol.* **2018**, *38*, 4395–4404. [CrossRef]

82. Yang, H.B.; Yang, D.W.; Hu, Q.F.; Lv, H.F. Spatial variability of the trends in climatic variables across China during 1961–2010. *Theor. Appl. Climatol.* **2015**, *120*, 773–783. [CrossRef]

83. Ren, G.; Guo, J.; Xu, M.; Chu, Z.; Zhang, L.; Zou, X.; Li, Q.; Liu, X. Climate changes of china's mainland over the past half century. *Acta Meteorol. Sin.* **2005**, *63*, 942–956.

84. Li, H.; Fu, Z.; Wen, X.; Huang, J. Characteristic Analysis of Sunshine Duration Change in China during the Last 50 Years. *Clim. Environ. Res.* **2013**, *18*, 203–209.

85. Wang, H.; Pinker, R.T. Shortwave radiative fluxes from MODIS: Model development and implementation. *J. Geophys. Res. Atmos.* **2009**, *114*. [CrossRef]

86. Qin, W.; Wang, L.; Lin, A.; Zhang, M.; Bilal, M. Improving the Estimation of Daily Aerosol Optical Depth and Aerosol Radiative Effect Using an Optimized Artificial Neural Network. *Remote Sens.* **2018**, *10*, 1022. [CrossRef]

87.　Jia, B.; Xie, Z.; Dai, A.; Shi, C.; Chen, F. Evaluation of satellite and reanalysis products of downward surface solar radiation over East Asia: Spatial and seasonal variations. *J. Geophys. Res. Atmos.* **2013**, *118*, 3431–3446. [CrossRef]

88.　Randles, C.A.; Kinne, S.; Myhre, G.; Schulz, M.; Stier, P.; Fischer, J.; Doppler, L.; Highwood, E.; Ryder, C.; Harris, B.; et al. Intercomparison of shortwave radiative transfer schemes in global aerosol modeling: Results from the AeroCom Radiative Transfer Experiment. *Atmos. Chem. Phys.* **2013**, *13*, 2347–2379. [CrossRef]