Trends and Variability of Atmospheric Downward Longwave Radiation Over China From 1958 to 2015

Yu Wei1,2, Xiaotong Zhang1,2, Wenhong Li3, Ning Hou1,2, Weiyu Zhang1,2, Jiawen Xu1,2, Chunjie Feng1,2, Kun Jia1,2, Yunjun Yao1,2, Jie Cheng1,2, Bo Jiang1,2, Kaicun Wang4,5, and Shunlin Liang2

1 State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and Institute of Remote Sensing and Digital Earth of Chinese Academy of Sciences, Beijing, China, 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, China, 3 Nicholas School of the Environment, Duke University, Durham, NC, USA, 4 Institute of Global Change and Earth System Science, Beijing Normal University, Beijing, China, 5 Department of Geographical Sciences, University of Maryland, College Park, MD, USA

Abstract

Surface downward longwave radiation (SDLR) is a major component of the energy budget. Although studies have reported the spatiotemporal variations of SDLR in China, the spatiotemporal coverage of the situ measurements used is always limited. In this study, the gradient boosting regression tree (GBRT) was developed to reconstruct SDLR based on air temperature (T_a), relative humidity (RH), and downward shortwave radiation (DSR). Ground measurements collected at the Baseline Surface Radiation Network (BSRN) and the Arid and Semi-arid Region Collaborative Observation Project (ASRCOP) were used to build and evaluate the GBRT model. The evaluation results showed that the daily SDLR estimates are correlated well with the SDLR in situ, with an overall root mean square errors (RMSE) of 16.5 Wm⁻² and a correlation coefficient (R) value of 0.91 for the validation data set. Comparison with existing SDLR products showed that accuracy and trends of the SDLR estimates based on the GBRT method are reasonable. To obtain long-term SDLR data for spatiotemporal analysis over China, densely distributed reconstructed DSR and ground measured T_a and RH collected at 756 Chinese Meteorological Administration (CMA) stations were used as input to estimate the SDLR based on the GBRT method over China during 1958–2015. The long-term estimated SDLRs at the selected 563 stations showed that SDLR increased at an average rate of 1.3 Wm⁻² per decade over China from 1958 to 2015. The trend of SDLR is positively correlated with the trend in T_a and water vapor pressure, whereas negatively correlated with the trend in DSR.

1. Introduction

Surface downward longwave radiation (SDLR) is a fundamental component of the Earth's radiation budget (Izismon et al., 2003). Understanding of SDLR variation is vitally important for weather prediction, energy budget evaluations, and numerous applications, such as predicting evapotranspiration and temperature variations (Flerchinger et al., 2009). SDLR is mainly emitted by H₂O, CO₂, and O₃ molecules and cloud water droplets in the atmosphere (Guo et al., 2019). SDLR is considered to be an important forcing on the Earth's surface energy budget that can produce surface warming (Burt et al., 2016; Woods & Caballero, 2016; Zeppepetello et al., 2019). Additionally, it is essential for understanding the impact of increasing CO₂ and other greenhouse gases on the climate (Stephens et al., 1994).

SDLR is not conventionally measured, because it is difficult and expensive to measure directly (Duarte et al., 2006; Enz et al., 1975; Sridhar & Elliott, 2002; K. C. Wang & Liang 2009a). Consequently, different parameterization schemes have to be developed that take more readily available meteorological values as input. The weighting function of SDLR peaks near the surface, and correctly calculating the air temperature close to the surface is critical for SDLR estimation (Gupta et al., 2004). Previous studies showed that under clear sky conditions, the distributions of air temperature (T_a) and humidity are most important for SDLR estimation (Cheng et al., 2017). T_a and surface water vapor pressure measured close to the ground are used for SDLR estimation in the parameterization schemes under clear sky conditions (Brunt, 1932; Brutsaert, 1975; Idso, 1981; Swinbank, 1963). However, the coefficients for the parameterization schemes are often specific...
for a limited range of climate regions and atmosphere conditions under which they were calibrated and validated. The parameterization scheme should be redefined or calibrated before it is used in other places (Bilbao & De Miguel, 2007). Retrievals of \( T_s \) and relative humidity (RH) profiles from satellite observations are also employed to estimate SDLR (Ellingosn, 1995). Radiative transfer models (e.g. LOWTRAN or MODTRAN), named physical based methods, have been used to describe the actual emission and absorption processes in the atmosphere and to estimate SDLR (Duarte et al., 2006). Darnell et al. (1983) estimated SDLR based on Television and Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) data obtained from the National Oceanic and Atmospheric Administration (NOAA). Although these physical based methods have explicit physical basis, the accuracy of the input parameters may directly affect the estimation accuracy. It is also very difficult to obtain temperature and humidity profiles of the lower atmosphere at stations with the accuracy required for SDLR estimation (Duarte et al., 2006; Ellingosn, 1995; K. Wang & Dickinson 2013; K. C. Wang & Liang 2009a). Moreover, the physical based methods may not be suitable for cloudy sky conditions since clouds are always impenetrable in the thermal infrared spectrum as pointed by K. C. Wang and Liang (2009a). For example, most passive satellite sensors can only provide observations of the cloud top, but SDLR is more closely related to the parameters of cloud base (Ellingosn, 1995).

Most existing parameterization schemes use cloud cover fraction to quantify the contribution of clouds to SDLR, which is proportional to the total cloudiness (Aladosarboledas, 1993; Bilbao & De Miguel, 2007; Ellingosn, 1995; Niemela et al., 2001). Meteorological observations and satellite cloud detection products can provide reliable cloud cover fraction measurements (Ackerman et al., 2008). For example, Diak et al. (2000) proposed a parameterization for SDLR estimation under cloudy sky conditions, in which the cloud product collected from the Geostationary Operational Environment Satellite (GOES) was used to quantify the influence of clouds. This parameterization method showed an overall root mean square error (RMSE) value of 20 Wm\(^{-2}\) against the ground measurements. The cloud cover fraction can alternatively be represented by the ratio of the measured horizontal global solar radiation to the horizontal global solar radiance under clear sky conditions (Crawford & Duchon, 1999). Downward shortwave radiation (DSR) can reflect the contribution of cloud to the SDLR. Crawford and Duchon (1999) proposed an improved parameterization scheme for calculating SDLR based on DSR measurements. K. Yang et al. (2010) estimated the SDLR based on calculating the cloud cover fraction using DSR estimated from a hybrid model. The evaluation analysis showed that the error of four Chinese Meteorological Administration (CMA) stations in the Tibetan Plateau (TP) was less than 30 Wm\(^{-2}\). Most of the parameterization schemes under cloudy sky conditions strongly depend on the calibration data and do not fully consider the impact of cloud characteristics, such as the cloud base. Thus, it may also have larger biases outside the parameter range of their local calibration (K. Wang & Dickinson, 2013).

Besides these physical based methods, machine learning methods are alternative ways to estimate surface radiation (X. Y. Wang et al., 2017; Wei et al., 2019; L. Yang et al., 2018). Machine learning methods provide techniques that can automatically construct the relationship between input parameters and surface radiation by processing the available data and maximizing a problem dependent performance criterion. Wei et al. (2019) estimated the DSR using four machine learning methods based on Advanced Very High Resolution Radiometer (AVHRR) data. The evaluation results with ground measurements exhibited that the gradient boosting regression tree (GBRT) method was the most accurate. Unlike other machine learning methods, the GBRT method can automatically find nonlinear interaction via decision tree learning and achieve more accurate predictions (Johnson & Zhang, 2014). However, few studies have directly applied the GBRT method to estimate SDLR based on ground measurements, especially over China.

Many studies have been reported for estimating SDLR over China, including the parameterization methods (Yu et al., Yu et al., 2018), hybrid methods (B. Tang & Li, 2008; J. Wang et al., 2014; W. H. Wang & Liang, 2009b) and artificial neural networks (ANN) based methods (T. X. Wang et al., 2012; T. X. Wang et al., 2018). For example, Yu et al. (2018) compared twelve and eight parameterizations methods under clear and cloudy sky conditions over Heihe River Basin in China, respectively. It showed that the estimated SDLRs based on the proposed schemes by Idso (1981) and Dilley and O’Brien (1998), and Maykut and Church (1973) performed best for the clear and cloudy sky conditions, respectively. J. Wang et al. (2014) developed an improved hybrid method to estimate SDLR over the TP using the Moderate-resolution Imaging Spectroradiometer (MODIS) observations under clear sky conditions. The estimated SDLRs based on the
proposed hybrid method have an overall RMSE value of 25.9 and 25.7 Wm\(^{-2}\) for the MODIS observations from Terra and Aqua, respectively. The ANN-based model was used to estimate SDLR over the TP by T. X. Wang et al. (2012). The model is comparable with or even better than existing algorithms, with an overall RMSE value of 20.1 Wm\(^{-2}\) and a bias value of \(-8.8\) Wm\(^{-2}\). Although much effort has been conducted on the improvements of the methods for SDLR estimation, studies on reconstructing the long-term SDLR datasets over China is still rare. Chang and Zhang (2019) reconstructed SDLR data sets at 351 stations over China. He et al. (2020) developed the China Meteorological Forcing Data set (CMFD) based on the ground measurements at 753 CMA stations from 1979 to the present. The SDLR forcing data set was estimated based on the semi-empirical method proposed by Crawford and Duchon (1999), which calculated the atmospheric emissivity as a function of cloud amount and \(T_a\). The temporal coverage of the estimated SDLR by the released methods is limited and the spatiotemporal analysis of SDLR over China is not well discussed. To improve our understanding of the climate change, it is still necessary to reconstruct a comprehensively spatiotemporal extended SDLR data set over China and computed spatiotemporal analysis based on this data set.

The physical relationship between the \(T_a\), RH, DSR, and SDLR is fairly well known. In particular, Cheng et al. (2017) and Zeppetello et al. (2019) have shown that near surface air temperature is the prominent driver of both clear and all-sky downward longwave radiation in observations and climate models respectively. The radiative kernels presented by Previdi (2010) and Pendergrass et al. (2018) also present clear physical explanations for the relationship between SDLR and meteorological variables, including \(T_a\) and RH. SDLR estimation under cloudy sky conditions depends strongly on cloud condition. DSR can be used to quantify the contribution of clouds to SDLR under cloudy sky conditions (Crawford & Duchon, 1999; K. Yang et al., 2010). Crawford and Duchon (1999) calculated cloud fraction from DSR under clear and cloudy conditions. Thus, \(T_a\), RH, and DSR measurements are selected as input variables of the GBRT method in this study.

There are a total of 756 CMA meteorological stations where \(T_a\), RH, and other surface meteorological parameters are measured, and all these data are available to the public. Compared with 756 routine meteorological stations, only 122 have global solar radiation measurements. Among 122 radiation observation stations, only 48 have relatively complete record from 1970 to 2015 through statistics. It is clear that the current existing radiation observation stations have relatively low spatial coverage and representativeness for long-term analysis. Therefore, the objectives of this study are: (1) to estimate SDLR using ground measurements. Ground measurements collected at the Baseline Surface Radiation Network (BSRN) and the Arid and Semi-arid Region Collaborative Observation Project (ASRCOP) were used to build and evaluate the GBRT model, respectively; and (2) to analyze the spatial pattern and temporal variations of the SDLR over China. To obtain long-term and densely distributed SDLR data over China for subsequent spatiotemporal analysis, the reconstructed long-term DSR (Hou et al., 2020), and the ground measured \(T_a\) and RH collected at 756 CMA stations from 1958 to 2015 were used as input of the proposed GBRT method. The accuracy and trend evaluation results of the SDLR estimates based on the GBRT method are also compared with existing SDLR products.

This paper is organized as follows: In Section 2, the ground measurement data used in this paper are described. The machine learning method and trend test method are described in Section 3. The results are presented in Section 4. In Section 5, we discuss the correlation between the trends of estimated SDLRs and other variables (such as \(T_a\)). The conclusions are given in Section 6.

2. Data

2.1. Ground Measurements

The data records at BSRN stations have been reported to show a higher level of data quality (Liang et al., 2010). The ground measurements collected at the BSRN were used to build the model, including daily DSR (Wm\(^{-2}\)), SDLR (Wm\(^{-2}\)), air temperature (\(T_a\), °C) at 2 m height, relative humidity at 2 m height (RH, %), and elevation (m). Since the SDLR ground measurements are not provided at the CMA stations, the ground measured SDLR collected at the ASRCOP stations was used to validate the robustness and accuracy
of the model. Ground measurements collected at the CMA from 1958 to 2015 were used to derive the SDLR estimates over China in this study.

1. **BSRN**: The BSRN was initiated by the World Climate Research Programme (WCRP) to provide validation material for satellite radiometry and climate models. The BSRN operation started in 1992 at nine stations and currently consists of more than 60 operational stations (Ohmura et al., 1998). Recent reports have indicated that the BSRN measurements have the highest possible accuracy and a high temporal resolution in various climate zones and with uncertainties of approximately ±5 W m⁻² (Liang et al., 2010). Data extracted at 25 stations, which provide both DSR and SDLR records from 2000 to 2015, were used to train the model in this study. The spatial distribution of the data is shown in Figure 1.

2. **ASRCOP**: The ASRCOP provides pyranometer data from 18 China meteorological stations in the summers of 2008 and/or 2009. The observed SDLR and meteorological data at the ASRCOP stations were recorded with a temporal resolution of 10 or 30 min (Cheng & Liang, 2016; Huang et al., 2013). Data collected from nine stations were used to evaluate the accuracy of the estimation model. Figure 2 and Table 1 show the spatial distribution and detailed information of these stations.

3. **CMA**: There are a total of 756 CMA meteorological stations where daily $T_a$, the RH and other surface meteorological parameters are measured, and all these data are available to the public. Among these stations, only 122 have global solar radiation measurements. Figure 2 shows the spatial distribution of these routine meteorological and radiation stations. Solar radiation measurements at the CMA stations started in 1957. Since 1994, only 96 stations continued to measure solar radiation as the measurements at various stations stopped over the years (W. J. Tang et al., 2013). It is noted that there were two different types of radiometers equipped at the CMA stations before 1994 and afterward. Solar radiation measurements are more prone to errors and often encounter more problems, such as technical failures and operation-related problems, than other meteorological variable measurements (Moradi, 2009; W. J. Tang et al. 2010). Therefore, data quality control is indispensable for many applications. In this study, the quality control procedure proposed by Zhang et al. (2015) was performed.

### 2.2. Existing SDLR Products

Existing SDLR products are used for comparison with the SDLR estimates based on the GBRT method. Considering the time series and accuracy of different SDLR products, CERES-SYN SDLR product was used to compare accuracy on ASRCOP stations, the ERA5 and GEWEX-SRB SDLR products were used for comparison of long-term trend. The brief introduction of three SDLR products are as follows:

1. **GEWEX-SRB**: The latest version of GEWEX-SRB (v3.0) is applied in this study. The GEWEX-SRB SDLR data can be available from July 1983 to December 2007 at a 3-hourly resolution, which are then averaged.
into daily, monthly values. The data set is produced on a 1° × 1° global grid using satellite-derived cloud parameters and ozone fields, reanalysis meteorology, and a few other ancillary data sets (Cox et al., 2006). Based on their official Web site, the overall daily mean bias for GEWEX-SRB is 0.5 W m−2 and the RMSE is 21.8 W m−2 compared to BSRN measurements from 1992 through 2007. These values are −0.1 and 11.2 W m−2 at the monthly time scale.

2. CERES-SYN: CERES (Clouds and the Earth’s Radiant Energy System) SYN (Synoptic Radiation Fluxes and Clouds) product sponsored by National Aeronautics and Space Administration (NASA) were designed to study the earth’s top-of-atmosphere (TOA), on surface and within the atmosphere radiation budgets (Doelling et al., 2013; Ohmura et al., 1998). Data used in this study for comparison with the SDLR estimates based on the GBRT method are available from March 2000 to present with a spatial resolution of 1° × 1° and a daily temporal resolution.

3. ERA5: ERA5 data on single levels are the fifth-generation ECMWF atmospheric reanalysis of the global climate, covering the period from 1979 to present (Hersbach & Dee, 2016; Naseef & Kumar, 2008). The data can be available on the time resolution of hourly and monthly with the spatial resolution of 0.25° × 0.25°. Reanalysis combines model data with observations from across the world into a globally complete and consistent data set using the laws of physics.

![Figure 2](image.png)

**Figure 2.** Spatial distribution of the radiation stations provided by the ASRCOP and CMA. ASRCOP, Arid and Semi-arid Region Collaborative Observation Project; CMA, Chinese Meteorological Administration.

| Station    | Latitude (°) | Longitude (°) | Elevation (m) | Land cover         | Time period      |
|------------|--------------|---------------|---------------|-------------------|------------------|
| Arou       | 38.04        | 100.46        | 3,033         | Desert/grassland  | 2008–2009        |
| Dongsu     | 44.09        | 113.57        | 970           | Desert/grassland  | 2008–2009        |
| Jinzhou    | 41.18        | 148.48        | 22            | Farmland          | 2008–2009        |
| Maqu       | 33.89        | 102.14        | 3,423         | Wetlands          | 2008             |
| Miyun      | 40.63        | 117.32        | 350           | Farmland          | 2008–2009        |
| Tongyu grass | 44.58   | 122.92        | 184           | Grassland         | 2008–2009        |
| Tongyu farmland | 44.59 | 122.93        | 184           | Farmland          | 2008–2009        |
| Yingke     | 38.86        | 100.41        | 1,519         | Farmland/oasis    | 2008–2009        |
| Yuzhong    | 35.95        | 104.13        | 1,965         | Desert/grassland  | 2008–2009        |

**Table 1**

**Basic Summary of the ASRCOP Stations Used to Validate the Model**
2.3. Reconstructed DSR Data set over China

The reconstruction data set of long-term DSR over China from Hou et al. (2020) was used as input to obtain long-term and densely distributed SDLR data over China. This data set was generated based on the random forest (RF) algorithm using the ground measured DSR and the routine meteorological station data collected at 756 CMA stations. This data set was available from 1958 to 2015 with a daily time resolution. The DSR estimates are validated using the ground measurements with a correlation coefficient ($R$) value of 0.99, a bias value of 0.01 $\text{Wm}^{-2}$, and an RMSE value of 8.88 $\text{Wm}^{-2}$. The reconstructed DSR data set is also reasonably accurate compared to the existing reconstructed data set.

3. Method

3.1. Gradient Boosting Regression Tree

The GBRT algorithm can be considered as an improved version of boosting that is based on iteratively constructing multiple individual decision trees. Boosting is an ensemble learning algorithm which combined a series of weak classifiers into a strong classifier according to different weights. The basic idea of GBRT algorithm is to establish a new regression model in the direction of gradient reduction, and to form a regression tree model through continuous iterations. The main advantage of the GBRT algorithm is that it can automatically find nonlinear interactions via decision tree learning, and it has relatively few tuning parameters as a nonlinear learning scheme (Johnson & Zhang, 2014).

Assuming that $x_i$ is a set of predictor variables, $y_i$ is a set of response variables, and $N$ is the number of training samples. The GBRT method constructs $M$ different individual decision trees $h(x; a_1), h(x; a_2), \ldots, h(x; a_M)$, then $h(x; a_m)$ can be used as the basic function to express the approximation function $f(x)$ as follows (Ding et al. 2016):

$$f(x) = \sum_{m=1}^{M} f_m(x) = \sum_{m=1}^{M} \beta_m h(x; a_m)$$

where $\beta_m$ and $\alpha_m$ represent the weight and classifier parameter of each decision tree, respectively. The loss function $L(y, f(x))$ is used to describe the accuracy of $\beta_m$ and $\alpha_m$. Each tree partitions the input space into $J$ regions $R_{jm_1}, R_{jm_2}, \ldots, R_{jm_J}$ and each $R_{jm}$ predicts the constant $y_{jm}$. The main flowchart of the GBRT method is shown in Figure 3. In this study, the GBRT method is implemented using the scikit-learn toolbox on the Python platform (Pedregosa et al., 2012). The main flowchart of this study is shown in Figure 4.

3.2. Mann-Kendall (MK) Trend Test

The nonparametric MK statistical test (Kendall, 1938; Mann, 1945) has been employed to detect trends in different hydrological and meteorological time series. Compared to linear regression trend analysis, the MK trend test is more suitable for cases where the trend may be assumed as a monotonic and normal distribution (Zhou et al., 2018). The test statistic $S$ is given by (Gocić & Trajkovic, 2013):

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{sgn}(x_j - x_i),$$

where $n, x_i, x_j$ represent the number of data points, data values in the time series $i$ and $j (j > i)$, respectively.
Figure 3. The main flowchart of the GBRT method. GBRT, gradient boosting regression tree.

Figure 4. The main flowchart of this study.

Initialize \( f_1(x) = \arg \min \rho \sum_{i=1}^{N} L(y_i, \rho) \)

For \( m = 1 \) to \( M \) do
  For \( i = 1 \) to \( n \) do
    Compute the negative gradient \( \hat{y}_i = -\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)} \)
  End;

  Fit a regression tree \( h(x; \alpha_n) \) to predict the targets \( \hat{y}_i \) from covariates \( X_i \) for all training data

  Compute a gradient descent step size as \( \rho_n = \arg \min \rho \sum_{i=1}^{n} L(y_i, f_{m-1}(x_i) + \rho h(x; \alpha_n)) \)

  Update the model as \( f_m(x) = f_{m-1}(x) + \rho_n h(x; \alpha_n) \)

End;

Output the final model \( f_M(x) \)
The variance is computed as:

\[
\text{Var}(S) = \frac{n(n - 1)(2n + 5) - m \sum_{i=1}^{m} (t_i - 1)(2t_i + 5)}{18},
\]

where \(n\) and \(m\) are the number of data points and tied groups, respectively, and \(t_i\) denotes the number of ties of extent \(i\). A tied group is a set of sample data with the same value. In cases where the sample size is \(n > 10\), the standard normal test statistic \(Z_s\) is expressed as (Zhou et al., 2018):

\[
Z_s = \begin{cases} 
\frac{S - 1}{\sqrt{\text{Var}(S)}}, & \text{if } S > 0 \\
0, & \text{if } S = 0 \\
\frac{S + 1}{\sqrt{\text{Var}(S)}}, & \text{if } S < 0
\end{cases}
\]

(5)

In a two-tailed test, the null hypothesis of no trend should be accepted at a specific significance level for \(-Z_{\alpha/2} \leq Z_s \leq Z_{\alpha/2}\), where \(Z_{\alpha/2}\) is the standard score of the standard normal distribution with a cumulative probability of \(1 - \alpha / 2\). Otherwise, the null hypothesis of no trend is rejected, and a monotonic trend is identified at significance level \(\alpha\). Positive values of \(Z_s\) indicate increasing trends, while negative \(Z_s\) values indicate decreasing trends. In this study, \(\alpha = 0.05\) was taken to identify a significant trend which means that \(Z_{\alpha/2} = \pm 1.96\).

In this study, we used the Sen’s slope to describe the steepness of the trend in long time series, which is computed as (Sen, 1968):

\[
Q_i = \frac{x_j - x_k}{j - k}, \text{ for } i = 1, \ldots, N,
\]

(6)

\[
\text{Slope} = \begin{cases} 
Q_{\lfloor(N+1)/2\rfloor}, & \text{if } N \text{ is odd} \\
\frac{Q_{\lfloor(N)/2\rfloor} + Q_{\lfloor(N+1)/2\rfloor}}{2}, & \text{if } N \text{ is even}
\end{cases}
\]

(7)

where \(N\) is the number of data pairs, \(x_j\) and \(x_k\) are the data values at times \(j\) and \(k\) (\(j > k\), respectively. The \(N\) values of \(Q_i\) are ranked from smallest to largest.

3.3. Model Construction

The GBRT model can be constructed in three steps.

(1) Preparation of the training datasets: Four variables extracted at the BSRN stations were used as predictor variables, including elevation, daily \(T_a\), daily RH, and daily DSR. The daily SDLR measurements were used as target variables.

(2) Configuration of the model coefficients: We used the \(k\)-fold cross validation method to determine the optimal parameters. Each parameter is traversed in range of parameter threshold, as shown in Table 2. The error of predicted results is evaluated against ground measurements and parameters providing the highest average \(R\) in the training data set were selected as optimal parameters. The GBRT model is influenced by the number of iterations, learning rate, depth of the tree, and sampling rate. The learning rate parameter limits the contribution of each tree. A small learning rate parameter can reduce overfitting. A larger iteration number parameter means more boosting stages to perform and usually provides better performance for the training data set. The iteration number parameter should be carefully set to avoid overfitting. Moreover, there is a trade-off between the learning rate and iteration number. The model complexity and computational cost increase with increasing iteration number and decreasing learning rate, leading to a poor prediction performance. The tree depth represents the maximum depth of the individual regression estimators which can limit the number of nodes in the tree. The sampling rate parameter represents the fraction of training samples used for fitting. A subsample parameter smaller than 1.0 can prevent overfitting and reduce the variance. Successive performance testing showed that a
GBRT model with a learning rate parameter of 0.1, a sampling rate of 1, a tree depth of 5, and an iteration number of 250 was optimal to estimate the SDLR

(3) Application of the GBRT method: After determining the optimal parameters, the performance of the trained model for the SDLR estimation was evaluated using ground measurements collected from the ASRCOP stations

4. Results and Analysis

4.1. Validation Against Ground Measurements

4.1.1. Validation of the SDLR Estimates

Ground measurements collected at 25 BSRN stations were used as the training data set to determine the optimal parameters. Then, daily SDLR measurements collected at nine ASRCOP stations in the summers of 2008 and/or 2009 were used as the validation data set to evaluate the performance of the GBRT method for the SDLR estimation. The selected BSRN stations are mainly concentrated in South America, North America, and Europe. Using ASRCOP data in China as the validation data set can validate the accuracy of the GBRT method without local correction, that is, validate the robustness of the model. Three statistical measures were used to evaluate the estimates against ground measurements, including overall RMSE, R, and bias.

The performance of the GBRT method for the estimation of daily SDLR are evaluated on BSRN training data set and ASRCOP validation data set, respectively. The evaluation results are shown in Figure 5. The daily SDLR estimates for the BSRN training data set have an overall RMSE value of 13.22 Wm\(^{-2}\), a bias value of 0 Wm\(^{-2}\), and an R value of 0.99, whereas these values are 16.5, 3.82 and 0.91 Wm\(^{-2}\) for the ASRCOP validation data set, respectively. The validation results for each ASRCOP station was further investigated to study the stability of the GBRT method, as shown in Figure 6. The daily SDLR estimates correlate well with the ground measurements at most ASRCOP stations, with the R values ranging from 0.76 to 0.96, the bias values ranging from −9.56 to 22.78 Wm\(^{-2}\), and overall RMSE values ranging from 10.06 to 26.11 Wm\(^{-2}\). Note that the R value is greater than 0.85 at eight out of nine stations and the absolute value of the bias is less than 10 Wm\(^{-2}\) at seven out of nine stations. The estimated SDLRs at Dongsu correlate best with the ground measurements, with an overall RMSE value of 10.06 Wm\(^{-2}\), a bias value of −0.14 Wm\(^{-2}\), and an R value of 0.96. These evaluation results further indicate that the SDLR estimates derived from the GBRT method correlate well with the ground measured SDLRs.

Figure 5. Evaluation results of the (a) BSRN training and (b) ASRCOP validation data set’s daily SDLR estimates based on the GBRT method against ground measurements. N is the number of total data points. ASRCOP, Arid and Semi-arid Region Collaborative Observation Project; BSRN, Baseline Surface Radiation Network; GBRT, gradient boosting regression tree; SDLR, Surface downward longwave radiation.
4.1.2. Validation of the Reconstructed DSR Dataset

In order to ensure the accuracy of the input variables of the machine learning method, the reconstructed long-term DSR data set over China from 1958 to 2015 was validated at the DSR ground measures collected at 122 CMA radiation stations. As shown in Figure 7, the DSR estimates from reconstructed data set have an R value of 0.95, a bias value of 1.34 Wm$^{-2}$, and an RMSE value of 27.01 Wm$^{-2}$, at a daily time scale. These values are 0.97, 15.95, and 1.34 Wm$^{-2}$, respectively, at a monthly time scale. Thus, the reconstructed DSR data set is reasonably accurate against the DSR ground measures.

4.2. Comparison with Existing SDLR Products

The CERES-SYN SDLR product was used to compare the evaluation results of SDLR estimates based on the GBRT method against ground measurements at ASRCOP stations in the summers of 2008 and/or 2009. As shown in Figure 8, the SDLR estimates based on the GBRT method correlate better with the ground measurements, with an overall RMSE value of 16.5 Wm$^{-2}$, a bias value of 3.82 Wm$^{-2}$, and an R value of 0.91.
The CERES-SYN and ERA5 SDLR products all show lower accuracy. The CERES-SYN SDLR product has an overall RMSE value of 23.93 Wm\(^{-2}\), a bias value of 7.95 Wm\(^{-2}\), and an \(R\) value of 0.84. These values were 28.38, −7.29, and 0.74 Wm\(^{-2}\), respectively, for the ERA5 SDLR product.

To further testify the SDLR estimates based on the GBRT method on 563 CMA stations, the RMSE and bias between daily SDLR estimates and ERA5 SDLR product are calculated at each CMA station from 1979 to 2015, as shown in Figures 9 and 10. The RMSE and bias range from 11.61 to 80.94 Wm\(^{-2}\) and −77.56 to 40.58 Wm\(^{-2}\), respectively. There are 292 and 278 sites whose RMSE and bias values range from 20 to 25 Wm\(^{-2}\) and −10 to 0 Wm\(^{-2}\), respectively; these are followed by 102 and 90 sites whose RMSE and bias values range from 25 to 30 Wm\(^{-2}\) and −20 to −10 Wm\(^{-2}\), respectively. The lower RMSE values are mainly found in the Northeast and South China, while the higher RMSE values are mainly distributed in the Tibet Plateau and west of Southwest China, which may be due to the high altitude and harsh environment leading to large ground observation errors. There are 27 out of 563 sites whose biases are more than 30 Wm\(^{-2}\), which may be due to the DSR estimates with relative big uncertainties at some stations. Moreover, the replacement of the CMA radiation instruments may also be a source of errors. It is worth to note that the spatial scaling issue would be another potential error sources for SLDR evaluation. We also compare the long-term trend of the SDLR estimates based on the GBRT method on 563 CMA stations with those from GEWEX-SRB and ERA5 products. The time period is set to 1984–2007 when all three SDLR datasets can be available. Figure 11 shows that the long-term trend of SDLR estimates based on the GBRT method (2.33 Wm\(^{-2}\) per decade, significant at 95% confidence) was similar to that from GEWEX-SRB (2.1 Wm\(^{-2}\) per decade, significant at 95% confidence), higher than that from ERA5 (1 Wm\(^{-2}\) per decade, significant at 95% confidence). Through the comparison of long-term trend with existing SDLR products, it is obvious that the SDLR estimates based on the GBRT showed a similar trend but different change magnitudes to existing SDLR products. Thus the temporal variations of SDLR based on the GBRT method on CMA stations are reasonable.
4.3. Spatial and Temporal Analysis of SDLR over China

The GBRT method used in this study performed well without a local correlation and only required surface meteorological and solar radiation data. Thus, we applied the GBRT method to obtain long-term and densely distributed SDLR data over China.
Meteorological measurements and reconstructed DSR data set were used to estimate SDLR based on the proposed GBRT method. The input variables of the GBRT method were $T_a$, RH measurements, elevation of the stations, and reconstructed DSR values at 756 CMA stations from 1958 to 2015. Monthly SDLR estimates were obtained by averaging the daily values over the month. If there were more than 10 missing daily SDLR estimates in a month at a station, the data for this month at this station were deleted. Then, if there was less than one missing monthly SDLR estimate at one station in a year, the missing values were obtained by piecewise cubic Hermit interpolation to calculate annual values. Meanwhile, if there were less than two missing annual values for the time period at a station, the missing annual values at this station were also obtained by piecewise cubic Hermit interpolation. Otherwise, this station would be eliminated to study the long-term trends of SDLR. Therefore, 563 stations were used to analyze the spatial pattern and temporal variations of SDLR based on the completeness of the data records.

Figure 10. The Bias between daily SDLR estimates and ERA5 SDLR products at 563 CMA stations from 1979 to 2015. CMA, Chinese Meteorological Administration; SDLR, Surface downward longwave radiation.

Figure 11. The comparison of anomalous annual mean SDLR estimates (unit: W m$^{-2}$) averaged over mainland China from GBRT-based estimates, GEWEX-SRB and ERA5 SDLR products during 1984–2007. GBRT, gradient boosting regression tree; SDLR, Surface downward longwave radiation.
4.3.1. Comparison with the SDLR Estimates Based on DSR Ground Measures

The SDLR estimates based on the reconstructed DSR data set was compared with which based on DSR ground measures at 122 CMA radiation stations from 1958 to 2015 to ensure the feasibility of spatiotemporal analysis. As shown in Figure 12, the SDLR estimates based on the reconstructed DSR data set correlate well with which based on the ground measures, with an $R$ value of 1, a bias value of $-0.42 \, \text{Wm}^{-2}$, and an RMSE value of $6.65 \, \text{Wm}^{-2}$, at a daily time scale. These values are 1, $-0.42$, and $5.17 \, \text{Wm}^{-2}$, respectively, at a monthly time scale. Thus the error of the reconstructed data set has little effect on the accuracy of SDLR estimates based on the GBRT method. We also compare the long-term trend between the SDLR estimates based on the reconstructed and ground measured DSRs. Regarding the completeness of the DSR ground measures, the time series is determined to be 1970–2015 for comparison. 48 and 563 CMA stations were used for long-term trend analysis from 1970 to 2015 based on the reconstructed and ground measured DSRs, respectively. As shown in Figure 13, the SDLR estimates at CMA stations based on the reconstructed and ground measured DSRs show consistent trends from 1970 to 2015. The difference between anomalous annual mean SDLR estimates based on two DSR datasets was range from $-0.02$ to $3.09 \, \text{Wm}^{-2}$, and the absolute values were within $1 \, \text{Wm}^{-2}$ in most years. The SDLR estimates used the DSR ground measurements as input showed significant increasing trends at a rate of $0.98 \, \text{Wm}^{-2}$ per decade from 1970 to 2015, while the value was $1.25 \, \text{Wm}^{-2}$ per decade for the SDLR estimates based on the reconstructed DSR data set. Thus, the SDLR estimates using reconstructed DSR data set as input can be used to perform spatiotemporal analysis of SDLR over China.

4.3.2. Spatial Distribution and Seasonal Variations of SDLR

According to the classification method of climatic types in China proposed by Zhou et al. (2018) and Liu et al. (2018), this study divided mainland China into six regions to compute spatiotemporal analysis, including East China (EC), North China (NC), Northeast China (NE), Southwest China (SW), South China (SC), and the TP. The spatial distribution and annual mean SDLR estimates of each region are presented in Figure 14. The different sizes and colors of the stations indicate the magnitude of the annual mean SDLR estimates at 563 CMA stations during 1958–2015. As it is shown in the figure, the annual mean SDLR estimates shows a pronounced latitudinal dependency except over the TP, with larger values at low latitudes and smaller values at high latitudes. This may be due to the low values of aerosol optical thickness and
atmospheric water vapor content over the TP. Table 3 shows the annual mean SDLR estimates over China and six climate regions. EC, SC, and SW show higher annual mean SDLR than the other regions in China. The maximum annual mean SDLR estimates occur in SC, whereas the minimum value occurs in TP. The difference between the annual mean SDLR estimates in SC and TP is up to 100 Wm$^{-2}$. Figures 15 and 16 show the monthly and seasonal mean SDLR estimates over mainland China and six subregions during 1958–2015, respectively. The monthly SDLR estimates gradually increase from January to July and then gradually decrease from July to December over both mainland China and six subregions. The highest seasonal mean SDLR typically occurs in summer, with areal averaged SDLR estimates of 379.20 Wm$^{-2}$. Following the same logic, the minimum SDLR occurs in winter, with values of 255.60 Wm$^{-2}$. The SDLR estimates show similar seasonal variation trends in all six subregions. The maximum seasonal

**Figure 13.** Comparison between the anomalous annual mean SDLR estimates (unit: Wm$^{-2}$) based on the DSR estimates and the DSR ground measures during 1970–2015. DSR, downward shortwave radiation; SDLR, Surface downward longwave radiation.

**Figure 14.** The annual mean SDLR estimates (unit: Wm$^{-2}$) at 563 CMA stations during 1958–2015. CMA, Chinese Meteorological Administration; SDLR, Surface downward longwave radiation.
differences occur in NE (147.7 Wm⁻²), while smaller values occur in SC and SW (66.9 and 68.3 Wm⁻², respectively).

### 4.3.3. Long-Term Trends

Figure 17 demonstrates the trends of the SDLR estimates at 563 CMA stations during 1958–2015. The size of each triangle represents the magnitude of the trend, and the red and green triangles indicate increasing and decreasing trends, respectively. Stations with a circle indicate that the trend detected by the MK test is significant at a 95% confidence level. The SDLR estimates at stations in central and northern China all showed increasing trends from 1958 to 2015, whereas the SDLR estimates at some stations in southern China showed decreasing trends. 549 out of 563 CMA stations show positive trends, and 476 stations increasing significantly during 1958–2015; 14 stations show negative trends, with one station decreasing significantly. The stations with negative SDLR trends were mainly concentrated in SC and SW.

Figure 18 shows the anomalous annual mean SDLR estimates averaged over mainland China and the six subregions from 1958 to 2015. A line was fitted to the anomalous annual mean SDLR estimates using linear regression. In mainland China, the maximum values occurred in 1998, and the minimum values occurred in 1969. A 1.04 Wm⁻² per decade increasing linear trend of SDLR is observed during the period of 1958–2015. The trends in EC, NC, NE, SW, and TP regions are similar to those over mainland China. In SC, the trend is insignificant but exhibits large interannual variations during the period of 1958–2015.

Since the MK test is more statistically rigorous than the regression method (Mann, 1945), the MK test is used to further analyze the long-term trend. The annual mean SDLR estimates over mainland China show significant increasing trends at a rate of 1.02 Wm⁻² per decade detected by the MK test. The increasing trends are comparable to those from previous studies (Prata, 2008; K. C. Wang & Liang, 2009a). CO₂ is another dominant emitter of SDLR, hence the effect of CO₂ on SDLR should be considered. The global atmospheric CO₂ concentration has increased by an average of 1.5 ppm per year from 1958 to 2015, which was calculated based on globally averaged marine surface data from the National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratory (ESRL) flask network (Laboratory, 2019). The CO₂ concentration increase in China approximately at the same rate as that of global (Administration, 2018). Increasing CO₂ concentration amount by 10% causes ~0.2% (~0.6 Wm⁻²) increase in SDLR (Prata, 2008). If the CO₂ concentration in the atmosphere increases at a rate of 1.5 ppm yr⁻¹, it will result in a corresponding increase in SDLR of 0.28 Wm⁻² per decade. Therefore, the increasing trend of the SDLR estimates would be 1.3 Wm⁻² per decade considering the variability of CO₂ concentration over China. Table 3 also shows that positive values are dominated in most regions. The annual mean SDLR estimates in EC, NC, NE, SW and

| Region | Annual mean SDLR (Wm⁻²) | 1958–2015 | 1958–1990 | 1991–2015 |
|--------|-------------------------|------------|-----------|-----------|
| China  | 316.14                  | 1.02*      | 0.59      | 0.59      |
| EC     | 336.74                  | 1.07*      | 0.44      | 0.73      |
| NC     | 277.06                  | 1.30*      | 0.51      | 0.57      |
| NE     | 278.31                  | 1.31*      | 0.41      | 0.94      |
| SC     | 384.22                  | 0.61       | 0.53      | 0.10      |
| SW     | 347.74                  | 0.60*      | 0.43      | −0.68     |
| TP     | 271.55                  | 1.18*      | 0.62      | 0.25      |

*Trend at the 5% significant level (p < 0.05). SDLR, Surface downward longwave radiation; EC, East China; NC, North China; NE, Northeast China; SC, South China; SW, Southwest China; TP, Tibetan Plateau.

---

**Figure 15.** Monthly mean SDLR estimates (unit: Wm⁻²) over the six regions and mainland China during 1958–2015. SDLR, Surface downward longwave radiation.
TP show significant increasing trends at a rate of 1.07, 1.30, 1.31, 0.60, and 1.18 Wm\(^{-2}\) per decade, respectively. SC shows insignificant increasing trends from 1958 to 2015, at a rate of 0.61 Wm\(^{-2}\) per decade. It can be seen that most regions over China showed a significant increasing trend from 1958 to 2015, except in SC and SW.

Given the long-term variability in SDLR, characterizing various time periods separately may be more useful than linearly fitting the entire time period. Pinker et al. (2005) claimed that the DSR showed a decrease until about 1990 and then a sustained increase. Therefore, the annual mean SDLR estimates series was divided, with break at 1990. A 0.59 Wm\(^{-2}\) per decade insignificant increasing between 1958 and 1990, followed by a 0.59 Wm\(^{-2}\) per decade insignificant increasing from 1991 to 2015, was found in anomalous annual mean SDLR estimates over mainland China. Annual mean SDLR estimates in EC, NC, NE, SC, and TP had very similar trends to mainland China in the specified three time periods, but with different magnitudes. The corresponding values were 0.44 and 0.73 Wm\(^{-2}\) for EC, 0.51 and 0.57 Wm\(^{-2}\) for NC, 0.41 and 0.94 Wm\(^{-2}\) for NE, 0.53 and 0.1 Wm\(^{-2}\) for SC, and 0.62 and 0.25 Wm\(^{-2}\) for TP. The trends in SW were different from other regions. In SW, there were a 0.43 Wm\(^{-2}\) per decade insignificant increasing over 1958–1990, and a 0.68 Wm\(^{-2}\) per decade insignificant decreasing over 1991–2015. Before 1990, the anomalous annual mean SDLR was negative in most years, but mostly positive after 1990. Therefore, the trends in these two time periods over all subregions are insignificant.

Table 4 summarizes the trends of the seasonal mean SDLR estimates over mainland China detected by the MK test from 1958 to 2015. The mean SDLR estimates exhibits a positive trend in the four seasons, with the highest rising rate in winter and lowest in summer. The seasonal mean SDLR in four season all show insignificant increasing trends from 1958 to 2015, with rates of 0.36, 0.38, 0.45, and 0.99 Wm\(^{-2}\) per decade, respectively. We also divided the time period with break at 1990. The seasonal mean SDLR estimates in four seasons all showed insignificant increasing over 1958–1990 and significant decreasing over 1991–2015. In spring, there were a 0.13 Wm\(^{-2}\) per decade increasing over 1958–1990, and a 3.13 Wm\(^{-2}\) per decade increasing over 1991–2015. In spring, there were a 0.13 Wm\(^{-2}\) per decade increasing over 1958–1990, and a 3.13 Wm\(^{-2}\) per decade increasing over 1991–2015.
Figure 18. The anomalous annual mean SDLR estimates (unit: Wm$^{-2}$) averaged over each region and mainland China during 1958–2015. SDLR, Surface downward longwave radiation.
Table 4
The Trends in Seasonal Mean SDLR Estimates of Each Season Over Mainland China Detected by MK Test

| Season | Seasonal mean SDLR (Wm⁻²) | 1991–2015 | 1958–1990 |
|--------|---------------------------|-----------|-----------|
| Spring | 309.38                    | 0.36      | 0.13      |
| Summer | 379.20                    | 0.38      | 0        |
| Autumn | 320.52                    | 0.45      | 0.90      |
| Winter | 255.60                    | 0.99      | 1.15      |

*Trend at the 5% significant level (p < 0.05). SDLR, Surface downward longwave radiation.

5. Discussion

Previous studies suggest that long-term SDLR variation is often determined by \( T_a \) and atmospheric water vapor concentration (K. C. Wang & Liang, 2009a). In this session, we investigate the correlation between the long-term variation of SDLR and other parameters over China. Near surface temperature and water vapor are used to calculate SDLR based on the Stefan-Boltzmann equation:

\[
SDLR = (eT_a, e) \sigma T_a^4
\]

where \( \sigma \) is the Stefan-Boltzmann constant \((5.67 \times 10^{-8} \text{ Wm}^{-2} \text{K}^{-4})\). \( e \) is the atmospheric effective emissivity under clear sky conditions. \( e \) can be modeled as a function of \( T_a \), water vapor pressure (e). RH is the ratio of water vapor pressure and saturation water vapor pressure, which can be calculated by \( T_a \) using the following equations. We choose water vapor pressure rather than RH to investigate the correlation between the trend in SDLR. Under cloudy sky conditions, the cloud cover fraction can also be estimated by the ratio of the measured horizontal global solar radiation to the horizontal global solar radiance under clear sky conditions. Thus, the DSR can be used to reflect cloud conditions. Next, the correlation between the trend in SDLR and the trend in \( T_a \), water vapor pressure, DSR over China from 1958 to 2015 are further explored:

\[
RH = \frac{e}{e_s} \times 100\% 
\]

\[
e_s = 6.11 \exp \left( \frac{L_v}{R_v} \left( \frac{1}{273} - \frac{1}{T_a} \right) \right).
\]

where \( e \) and \( e_s \) are the water vapor pressure and the saturation water vapor pressure, respectively.

Figure 19 shows the time series of anomalous annual mean SDLR estimates, measured \( T_a \) calculated water vapor pressure and measured DSR from 1958 to 2015 over China. It is shown that the trend of SDLR is generally consistent with the trend of \( T_a \) and water vapor pressure, whereas the trend of SDLR is opposite to the trend of DSR. To fully assess the causes of changes in SDLR, we further quantitatively investigated the correlation between the SDLR and other variables. Figure 20 is the scatterplots of the trend in SDLR estimates detected by the MK test as a function of the trends in \( T_a \), water vapor pressure, and DSR at the 563 CMA stations. One point in the figure represents one station. The trend of SDLR is positively correlated with the trends in \( T_a \) and water vapor pressure, and the R values between SDLR with \( T_a \) and vapor pressure are 0.62 and 0.60, respectively. The trend of SDLR is negatively correlated with the trend in DSR, with an R value of −0.16. SDLR is not strongly correlated with DSR over mainland China during the period of 1958–2015.

In order to study the characteristics of parameters which controlling the long-term variation of SDLR in different regions over China, the correlation of the trends in SDLR with the trends in \( T_a \), water vapor pressure and DSR over the six regions is also shown in Figure 21. The trend of SDLR is positively correlated with the trend in \( T_a \) and water vapor pressure, whereas it is negatively correlated with the trend in DSR in all sub-regions. In EC and TP, the trend of SDLR is highly correlated with the trend of \( T_a \), with R values of 0.65. In EC, NE, SC, SW, and TP, the trend of SDLR is highly correlated with the trend of water vapor pressure, with an R value of 0.63, 0.72, 0.70, 0.75, and 0.69, respectively. In TP, the trend of SDLR exhibits a relatively high negative correlation with the trend of DSR, with R values of −0.57. The trend of SDLR has no significant correlation with the trend of \( T_a \), water vapor pressure and DSR in NC. These results suggest that the primary controlling factors of the SDLR long-term variation for six climatic zones were different: the increases in water vapor pressure results to the rising trend over most subregions, the rising trend over TP mainly results from both increases in \( T_a \) and water vapor pressure and decreases in DSR, the rising trend over NC has no
significant correlation with those of other three variables. It is noted that the number of samples over each region is quite limited based on the completeness of the data records, and the points are scattered and uneven, which may lead to errors in the results.

6. Conclusions

SDLR is a major component of the Earth's climate system. However, SDLR is not conventionally observed due to the high cost and difficulty of a direct measurement. It has great significance to generate a comprehensively spatiotemporal extended SDLR dataset over China based on more readily available data. In this study, we reconstructed SDLR based on the GBRT method using $T_a$, RH and DSR. Daily ground measurements collected at the BSRN and ASRCOP stations were used to build and validate the GBRT model, respectively. The evaluation results showed that the estimated SDLRs using the GBRT method correlate well with the SDLR in situ, with an overall RMSE of 16.5 Wm$^{-2}$ and an $R$ value of 0.91 at a daily time scale. Thus, applying the GBRT method to estimate SDLR provides reasonable and realistic radiation quantity and its variation without a local correlation.
Figure 21. The scatterplots of trends in SDLR as a function of trends in $T_a$, water vapor pressure and DSR over six regions. DSR, downward shortwave radiation; SDLR, Surface downward longwave radiation.
To obtain long-term SDLR data for subsequent spatiotemporal analysis based on the proposed method over China, the densely distributed reconstructed DSR and ground measured $T_a$ and RH collected at 756 CMA stations were used as input to estimate the SDLR based on the GBRT method over China from 1958 to 2015. We also analyzed the spatial pattern and temporal variations of the estimated SDLRs at 563 CMA stations over China where the data were relatively complete during the period of 1958–2015. The maximum annual mean SDLR occurred in SC, whereas the minimum value occurred in TP. The seasonal mean SDLR estimates were highest in summer and lowest in winter. The spatial distribution of the estimated SDLRs in each season was similar to that in the whole year. It was found that SDLR increased significantly at an average rate of 1.3 Wm$^{-2}$ per decade from 1958 to 2015 as detected by the MK test. The long-term trends in most regions were consistent with those in the whole China area, except for SC. In SC, the annual mean SDLR exhibited insignificant increasing trends at a rate of 0.61 Wm$^{-2}$ per decade. We also compared the accuracy and trends of the SDLR estimates based on the GBRT method between those from existing SDLR products. The comparison result showed that accuracy and trends of the estimated SDLRs of the GBRT method are reasonable.

The primary controlling factors of the SDLR long-term variation was investigated in mainland China by analyzing the correlation between the trend of SDLR and the trends of $T_a$, water vapor pressure, and DSR at the 563 CMA stations. The trend of SDLR was generally positively correlated with the trend in $T_a$ and water vapor pressure, negatively correlated with the trend in DSR. The primary controlling factors of the SDLR long-term variation for six climatic zones were different.

Although the GBRT method is robust to outliers in output space, and has been efficient and practical for many research applications, the GBRT method also has some disadvantages. First, the GBRT method has poor scalability due to the order nature of its promotion. Second, the training procedure is sensitive to the choice of parameters. There is a trade-off between overfitting and computational cost. The step size of learning rate parameter may need to be small to avoid overfitting. However, the small learning rate parameter may imply a high computational cost of applications. Thus other machine learning methods or deep learning methods can be further explored to improve accuracy and efficiency of SDLR estimation.

The density of the SDLR measurements is sparser than that of the meteorological and DSR measurements. The SDLR can be estimated and easily extended to more stations and over longer time periods using the GBRT method without a local correlation. This study only applies the GBRT method at stations using ground measurements. However, the number and spatial distribution of the training samples may have influence on SDLR estimation. We plan to extend the GBRT method for SDLR estimation from stations to surface, using reanalysis data and/or retrievals from satellite observations.

Data Availability Statement
The ground measurements at CMA stations are available at http://cdc.nmic.cn/home.do; the ground measurements at BSRN stations are available at https://dataportals.pangaea.de/bsrn; the CERES-SYN data are available at https://ceres.larc.nasa.gov/; the ERA5 data are available at https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5; the GEWEX-SYN data are available at http://www.gewex.org/data-sets-surface-radiation-budget-srb/.

Acknowledgments
The authors would like to thank the Arid and Semi-arid Region Collaborative Observation Project (ASRCOP) program, which provided ground measurements for our validation. This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFA0603002, in part by the National Natural Science Foundation of China under Grants 41571340.

References
Ackerman, S. A., Holz, R. E., Frey, R., Eitaranta, E. W., Maddux, B. C., & McGill, M. (2008). Cloud detection with MODIS. Part II: Validation. *Journal of Atmospheric and Oceanic Technology*, 25(7), 1073–1086. https://doi.org/10.1175/2007jtech1053.1
Aladosarboledas, L. (1993). Estimation of hourly values of downward atmospheric radiation under cloudless skies during daytime and nighttime conditions. *Theoretical and Applied Climatology*, 48(2-3), 127–131. https://doi.org/10.1007/bf00864919
Bilbao, J., & De Miguel, A. H. (2007). Estimation of daylight downward longwave atmospheric irradiance under clear-sky and all-sky conditions. *Journal of Applied Meteorology and Climatology*, 46(6), 878–889. https://doi.org/10.1175/jamc2503.1
Brunt, D. (1932). Notes on radiation in the atmosphere. *Quarterly Journal of the Royal Meteorological Society*, 58(247), 389–420. https://doi.org/10.1002/qj.49705824704
Brutsaert, W. (1975). Derivable formula for long-wave radiation from clear skies. *Water Resources Research*, 11(5), 742–744. https://doi.org/10.1029/wr011i005p00742
