Page Proof Instructions and Queries

Journal Title: PPG

Article Number: 823459

Thank you for choosing to publish with us. This is your final opportunity to ensure your article will be accurate at publication. Please review your proof carefully and respond to the queries using the circled tools in the image below, which are available **by clicking "Comment"** from the right-side menu in Adobe Reader DC.*

Please use ***only*** the tools circled in the image, as edits via other tools/methods can be lost during file conversion. For comments, questions, or formatting requests, please use T. Please do ***not*** use comment bubbles/sticky notes.

*If you do not see these tools, please ensure you have opened this file with **Adobe Reader DC**, available for free at https://get.adobe.com/reader or by going to Help > Check for Updates within other versions of Reader. For more detailed instructions, please see https://us.sagepub.com/ReaderXProofs.

No.	Query
	Please note, only ORCID iDs validated prior to acceptance will be authorized for publication; we are unable to add or amend ORCID iDs at this stage.
	Please confirm that all author information, including names, affiliations, sequence, and contact details, is correct.
	Please review the entire document for typographical errors, mathematical errors, and any other necessary corrections; check headings, tables, and figures.
	Please confirm that the Funding and Conflict of Interest statements are accurate.
	Please ensure that you have obtained and enclosed all necessary permissions for the reproduction of artistic works, (e.g. illustrations, photographs, charts, maps, other visual material, etc.) not owned by yourself. Please refer to your publishing agreement for further information.
	Please note that this proof represents your final opportunity to review your article prior to publication, so please do send all of your changes now.
AQ: 1	Please correct the reference entry details for Segura et al. 2012. The DOI does not lead to this article (and so has been removed), the dates do not match (citation date and publication date) and the journal title has not been included (so please add).
AQ: 2	Yang et al. 2006 is listed in the References but there is no citation for this date. Please insert a citation or remove the reference entry.
AQ: 3	In the reference entry for Zeng et al. 2018, please insert the page range.

Progress in Physical Geography
1–15

DOI: 10.1177/0309133318823459
journals.sagepub.com/home/ppg

Improved empirical models for estimating surface direct and diffuse solar radiation at monthly and daily level: A case study in North China

Yao Feng
Queen's University, Canada

Dongmei Chen
Queen's University, Canada

Xinyi Zhao
College of Urban and Environmental Science, Peking University, China

Abstract

Precise knowledge of direct and diffuse solar radiation is important for energy utilization and agricultural activities. However, field measurements in most areas of the world are only for total solar radiation. The satellite-retrieved direct and diffuse solar radiation show poor performance under overcast skies. Therefore, better empirical models are needed to estimate direct and diffuse solar radiation by considering the impact of aerosols over polluted regions. A case study is conducted in North China with the ground-measured solar radiation and satellite-retrieved aerosol optical depth to improve new empirical models at monthly (from 2000 to 2016) and daily (from 2006 to 2009) level. The improved empirical models are validated using the field measurements and compared with the existing models. Results suggest that these models perform well in estimating direct solar radiation at monthly (R^2 = 0.86–0.91, $RMSE$ = 0.76–0.83 MJ/m^2) and daily (R^2 = 0.91–0.94, $RMSE$ = 1.51–1.64 MJ/m^2) level. The accuracy of estimated monthly (R^2 = 0.95–0.96, $RMSE$ = 0.57–0.65 MJ/m^2) and daily (R^2 = 0.91–0.93, $RMSE$ = 1.09–1.15 MJ/m^2) diffuse solar radiation, particularly the maximum diffuse solar radiation value, has been improved compared to the existing models. The models presented in this study can be useful in the improvement and evaluation of solar radiation dataset over polluted regions similar to North China.

Keywords

Direct solar radiation, diffuse solar radiation, aerosols, empirical model, North China

1 Introduction

Solar radiation is the key energy source for many physical and biological processes on the Earth. Total solar radiation (total radiation) passing through the atmosphere comprises the

Corresponding author:

Dongmei Chen, Queen's University Department of French Studies, 99 University Avenue, Kingston, Ontario K7L 3N6, Canada.
Email: chendm@queensu.ca

direct and diffuse components (Qian et al., 2007). Direct solar radiation (direct radiation), the direct beam radiation on the horizontal surface, also known as direct normal irradiance, is important for the heat gain of buildings (Paula et al., 2016) and solar power project (Polo et al., 2016). Diffuse solar radiation (diffuse radiation), radiation that has been scattered by molecules and particles in the atmosphere, plays an active role in thermal, chemical and biological processes on the Earth surface (Mubiru and Banda, 2007; Pedruzo-Bagazgoitia et al., 2017; Yue and Unger, 2017). Therefore, knowledge of total radiation alone is not sufficient, precise information of direct and diffuse radiation is also essential.

Field measurements of direct and diffuse radiation are not available for most areas of the world due to the high technical and financial costs. Satellite products and reanalysis data provide estimates of solar radiation, which have been evaluated by comparing with field measurements (Ineichen, 2014; Urraca et al., 2017, 2018). Despite some good estimates (Amillo et al., 2014; Bakirci and Kirtiloglu, 2018), Alonso-Montesinos et al. (2015) have concluded that the accuracy of satellite-derived direct and diffuse radiation is overall lower than that of total radiation and that it is the best under cloudless skies but the worst under overcast skies. Frank et al. (2018) have also found overestimates of global horizontal radiation by reanalysis data in cloudy skies. Given these facts, empirical models have been widely developed to estimate direct and diffuse radiation.

Liu and Jordan (1960) have presented the relationships of direct, diffuse and total radiation on both clear and cloudy days from hourly, daily and monthly scales. Since then, the estimations of direct and diffuse radiation have been improved by various empirically-derived models (Jiang, 2009; Mubiru and Banda, 2007; Zhang et al., 2017). Clearness index, the ratio of total radiation to extraterrestrial radiation, introduced by Erbs et al. (1982), has been a useful factor in estimating diffuse radiation (Boukelia et al., 2014; Bakirci and Kirtiloglu, 2018). Besides, sunshine duration, the most common type of solar radiation measurements, is also used for estimating diffuse radiation (Jiang, 2009; Sabzpooshani and Mohammadi, 2014). Li et al. (2011, 2012) have improved the temperature-based solar radiation model (Maghrabi, 2009) by incorporating relative humidity in estimating monthly diffuse radiation. Yao et al. (2015) have also developed a new anisotropic model based on the concept of radiation intensity to estimate diffuse radiation. In contrast, very few models have been developed to estimate direct radiation (Safaripour and Mehrabian, 2011; Spitters et al., 1986).

Despite the good general estimates worldwide, these models may not perform well in heavy polluted region where solar radiation would be highly altered by aerosols. In the recent decade, the aerosol-effect on solar radiation has been extensively discussed (Boers et al., 2017; Kim and Ramanathan, 2008; Yang et al., 2009). High aerosol concentration can decrease the duration and intensity of sunshine (Kaiser, 2002), reduce total radiation along with the increasing diffuse radiation (Qian et al., 2007). Zeng et al. (2018) have concluded that the estimation accuracy of solar radiation can be limited by using only sunshine duration without considering aerosols. Fernández-Peruchena et al. (2010) have estimated solar radiation using aerosol optical depth by a clear sky model. Therefore, further efforts are needed to improve the empirical models for estimating direct and diffuse radiation by considering the aerosol-effect over heavily polluted regions under all sky conditions.

The objective of this study is to improve empirical models for estimating direct and diffuse radiation by considering aerosols over polluted regions at monthly (from 2000 to 2016) and daily (from 2006 to 2009) level. A case study is conducted in North China (NC), a very populated and polluted region featured by an

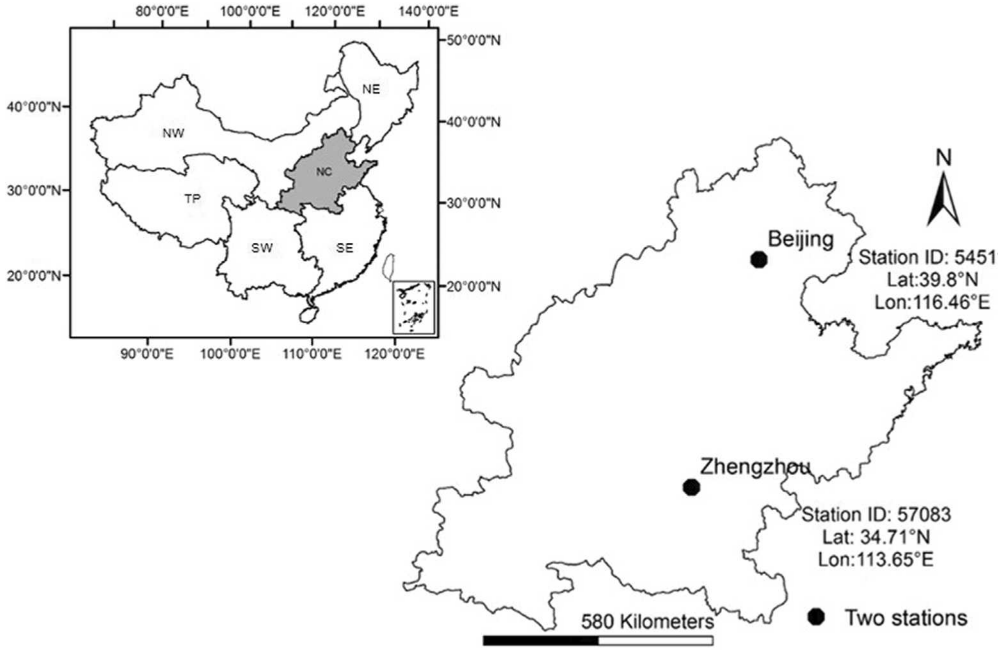

Figure 1. Location of the study area (North China, NC) and spatial distribution of the two stations in North China. The study area is defined by Xu et al. (2015), who have divided China into six regions including North China (NC) Southwest (SW), Southeast (SE), Northeast (NE), Northwest (NW) China and the Tibet Plateau (TP) according to the social-economical and geo-climatic conditions.

increasing proportion of anthropogenic aerosols. The improved empirical models are validated using the field measurements and also compared with the existing empirical models to evlaute the estimation accuracy and performance.

II Data and method

2.1 Data preprocessing

2.1.1 Ground-measured solar radiation. The solar radiation dataset was obtained from the China Meteorological Data Sharing System (CMDSS, http://cdc.cma.gov.cn/) including daily total, direct and diffuse radiation (unit: MJ/m^2). This dataset has been through the preliminary quality checks. Two meteorological and radiometric stations (Beijing 54511 and Zhengzhou 57083) located in NC, with over 99.9% daily total, direct and diffuse radiation from 2000 to 2016, were used to develop and validate the improved empirical models. Daily sunshine duration (unit: hour), the key measurement of solar radiation reaching the Earth's surface, was also collected at the two stations from the CMDSS for the same period. The few missing data in the two radiation datasets were interpolated with a three-day average moving window. Figure 1 presents the location of the study area and the spatial distribution of the two stations.

2.1.2 Satellite-retrieved aerosol concentration. To study the impact of the aerosol on solar radiation at the Earth's surface, satellite-retrieved aerosol optical depth (AOD) was collected from the Terra (MOD) and Aqua (MYD) platforms onboard the Moderate Resolution Imaging Spectroradiometer (MODIS). The spatial resolution is $1° \times 1°$. The MODIS instruments on the Terra and Aqua platforms scan the same area on Earth at approximately 10:30 am and 1:30 pm (local time), respectively. AOD, an effective indicator of aerosol concentration, is the degree to which aerosols prevent the transmission of light by absorption or scattering. Monthly AOD was collected from the Terra platform (MOD08M3) for the 2000–2016 period given few monthly AOD data from the Aqua platform. Daily AOD was collected from both platforms for four normal calendar years from 2006 to 2009 to increase the number of available daily AOD data.

At monthly level, there are generally more than 181 months' AOD data available for 2000 to 2016. The few missing monthly AOD data were interpolated by the seasonal average given the strong seasonality of AOD (Feng, Chen, Ouyang, et al., 2018). At daily level, considering more AOD data with higher accuracy from the Terra platform (Segura et al., 2012), the missing AOD data from the Terra platform were filled with available AOD data from the Aqua platform to increase the number of daily AOD data. Despite the increased number, there are still 42.50–46.80% missing daily AOD data from 2006 to 2009, mainly detected in winter. To spare any biases introduced by different interpolation methods, daily direct and diffuse radiation are separately estimated on days with or without AOD data.

2.2 Model development

2.2.1 Model development at monthly level. Correlations between aerosol concentration (AOD), sunshine duration (SSD), total radiation (R), direct radiation (R_{direct}), and diffuse fraction ($R_{diffuse}/R$, the fraction of diffuse radiation in total radiation) are analyzed using Pearson correlation coefficients at each station (Table 1). Meanwhile, scatterplots of the predictor variables (AOD, SSD, R) against the response variables ($R_{direct}, R_{diffuse}/R$) from the two stations are drawn in Figure 2.

Table 1. Correlation coefficients between $R_{direct}, R_{diffuse}/R$ and $SSD \times R, SSD, AOD$ at monthly level at each station.

	R_{direct}	$R_{diffuse}/R$	
Station ID	$SSD \times R$	SSD	AOD
54511	0.95**	−0.43**	0.74**
57083	0.93**	−0.50**	0.55**

** suggests that correlation coefficients are statistically significant at 0.01 significance level.

Direct radiation, the part of total radiation that directly reaches the Earth's surface (Qian et al., 2007), can be measured by sunshine duration. More direct and total radiation are in reasonable agreement with longer sunshine duration. Therefore, there are significant ($p < 0.01$) positive correlations between direct radiation and sunshine duration modified by total radiation. However, the increase of direct radiation caused by longer sunshine duration may reduce the proportion of diffuse radiation in a given total radiation. Meanwhile, atmospheric aerosols can greatly increase diffuse radiation via scattering and reflection. Hence, diffuse fraction is significantly ($p <$ 0.01) negatively correlated with sunshine duration but positively correlated with aerosol concentration at the two stations.

Scatterplots reveal first-order polynomial correlations between direct radiation and sunshine duration modified by total radiation (Figure 2a), between diffuse fraction and sunshine duration (Figure 2b) and aerosol concentration (Figure 2c). Thereby, direct radiation can be estimated using total radiation and sunshine duration, whereas diffuse fraction is influenced by sunshine duration and aerosol concentration. Based on the analysis above, two empirical models are developed to estimate direct and diffuse radiation at monthly level, as shown in equation (1)

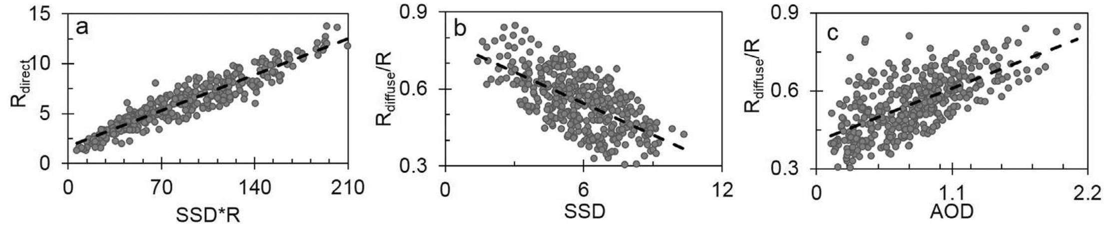

Figure 2. Scatterplots of $R_{direct}, R_{diffuse}/R$ against $SSD \times R, SSD, AOD$ at monthly level from the two stations.

Table 2. Correlation coefficients between $R_{direct}, R_{direct}/R, R_{diffuse}/R$ and SSD, AOD at daily level from the two stations.

	R_{direct}	R_{direct}/R	$R_{diffuse}/R$	
Year	SSD	AOD	SSD	AOD
Y2006	0.86**	−0.82**	−0.83**	0.82**
Y2007	0.87**	−0.75**	−0.81**	0.75**
Y2008	0.86**	−0.79**	−0.80**	0.79**
Y2009	0.89**	−0.84**	−0.74**	0.84**

** suggests that correlation coefficients are statistically significant at 0.01 significance level.

$$R_{direct} = a_1 \times R \times SSD + b_1$$
$$\frac{R_{diffuse}}{R} = a_2 \times SSD + b_2 \times AOD + c_2 \quad (1)$$

where a, b, c are coefficients to be determined.

2.2.2 Model development at daily level. Similar analysis is conducted at daily level with the Pearson correlation coefficients and scatterplots presented in Table 2 and Figure 3, respectively.

At daily level, similar to that at monthly level, there are significant ($p < 0.01$) positive correlations between direct radiation and sunshine duration from 2006 to 2009. Since the sum of direct and diffuse radiation is total radiation, the increased diffuse fraction by aerosols can reduce direct fraction. Therefore, there are negative correlations between direct fraction and aerosol concentration. Similar significant ($p <$ 0.01) correlations between diffuse fraction and sunshine duration, aerosol concentration at monthly level are also detected at daily level.

Scatterplots suggest second-order polynomial correlations between direct radiation and sunshine duration (Figure 3a), and between direct fraction and aerosol concentration (Figure 3b) but they are opposite in signs. Meanwhile, diffuse fraction suggests a first-order polynomial correlation with sunshine duration (Figure 3c) but a logarithmic correlation with aerosol concentration (Figure 3d). Thereby, at daily level, direct radiation increases in line with sunshine duration, whereas diffuse fraction cannot always increase with aerosol concentration as diffuse radiation cannot exceed total radiation and total radiation can be reduced by aerosols (Chou et al., 2006; Qian et al., 2007).

Based on the analysis above, two models incorporating three scenarios (S1–S3) are developed to estimate daily direct and diffuse radiation. S1 refers to the scenario where sunshine duration is zero. S2 and S3 refer to the scenarios with or without daily AOD data, respectively.

In S1 where sunshine duration (SSD) is zero, direct and diffuse radiation are estimated as in equation (2)

$$R_{direct} \approx 0$$
$$R_{diffuse} \approx R \quad (2)$$

Two weather conditions are considered in S1; one is the overcast or heavily polluted days when diffuse radiation is almost equal to total radiation whereas direct radiation is close to zero (and $R_{direct} \approx 0$); the other is rainy days without any solar radiation ($R_{diffuse} \approx R_{direct} \approx R = 0$).

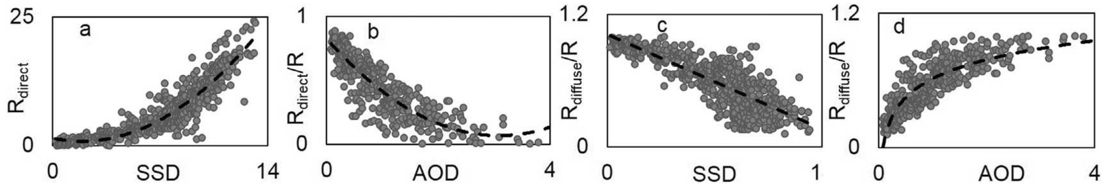

Figure 3. Scatterplots of R_{direct}, R_{direct}, $/R$, $R_{diffuse}/R$ against SSD, AOD at daily level from the two stations. Take the year 2006 as an example.

In S2 where aerosol optical depth (AOD) and sunshine duration (SSD) are both available, based on Table 2 and Figure 3, direct and diffuse radiation are estimated as in equation (3)

$$\frac{R_{direct}}{R} = a_3 \times AOD^2 + b_3 \times AOD + c_3$$
$$\frac{R_{diffuse}}{R} = a_4 \times SSD + b_4 \times \log(AOD) + c_4 \tag{3}$$

In S3 where aerosol optical depth (AOD) is not available, alternative methods are applied. Direct radiation is estimated as in equation (4)

$$R_{direct} = a_5 \times SSD^2 + b_5 \times SSD + c_5 \tag{4}$$

where a, b, c are coefficients to be determined.

Meanwhile, the method from Erbs et al. (1982) in equation (5) is used to estimate daily diffuse radiation. So far, all the empirical models have been developed.

$$\begin{aligned}
&\text{For } w_s < 1.4208\\
&\frac{R_{diffuse}}{R} = 1.0 - 0.2727K_T + 2.4495 \times {K_T}^2\\
&\qquad - 11.9514 \times {K_T}^3 + 9.3879\\
&\qquad \times {K_T}^4 \text{ for } K_T < 0.715\\
&\frac{R_{diffuse}}{R} = 0.143 \text{ for } K_T \geq 0.715\\
&\text{For } w_s \geq 1.4208\\
&\frac{R_{diffuse}}{R} = 1.0 + 0.2832 \times K_T - 2.5557 \times {K_T}^2\\
&\qquad + 0.8448 \times {K_T}^3 \text{ for } K_T < 0.722\\
&\frac{R_{diffuse}}{R} = 0.175 \text{ for } K_T \geq 0.722\\
&K_T = \frac{R}{R_0}
\end{aligned} \tag{5}$$

where K_T is the clearness index, w_s is sunset hour angle and R_0 is the extraterrestrial radiation.

2.3 Model validation

Measurements at each station are used to calibrate and validate these models for the initial model validation. Once these models perform well at each station, measurements from the two stations are used to calibrate and validate those models for regional application. Apart from the initial model validation, another two types of model validations are conducted. One is the validation of the empirical models between the two stations (the Type I validation), and the other is the comparisons between the empirical models presented in this study and the existing models in previous studies at the two stations (the Type II validation). The Type I validation ensures that the empirical models developed at site are applicable on a large regional scale. The Type II validation highlights the significance of aerosols in estimating diffuse radiation over heavy polluted regions.

In the Type I validation, the empirical models are calibrated using measurements from one station. Based on the calibrated models and measurements from the other station, the estimated direct and diffuse radiation are compared with

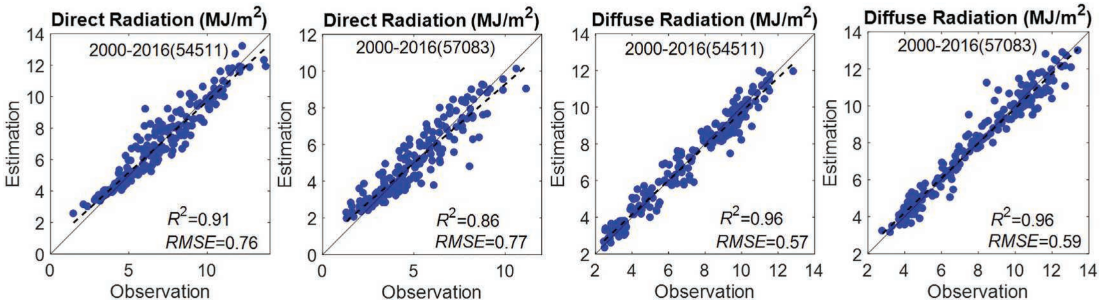

Figure 4. Comparisons of the observed against the estimated monthly direct and diffuse radiation at each station using equation (1).

the field measurements to validate the models. The Type I validation is applied to all empirical models presented in this study.

In the Type II validation, only the empirical models for estimating diffuse radiation at monthly and daily level are compared with those existing models in previous studies. At monthly level, according to Li et al. (2011, 2012), diffuse radiation ($R_{diffuse}$) is estimated using extraterrestrial radiation (R_0), actual and maximum sunshine duration (S, S_0), surface temperature (T_a) and relative humidity (R_h) as in equations (6-1) and (6-2). At daily level, Erbs et al. (1982) have proposed to estimate daily diffuse radiation using clearness index as in equation (5). These existing models mentioned above are compared with the empirical models in this study.

$$\frac{R_{diffuse}}{R} = a + b\left(\frac{S}{S_0}\right) + c\left(\frac{S}{S_0}\right)^2 + d\left(\frac{S}{S_0}\right)^3 \tag{6 − 1}$$

$$\frac{R_{diffuse}}{R_0} = a + b\left(\frac{S}{S_0}\right) + c\log\left(\frac{S}{S_0}\right) + dT_a + eR_h \tag{6 − 2}$$

In all the model validations, the coefficient of determination (R^2) and root mean square error ($RMSE$) are used to indicate the estimation accuracy.

III Results and discussions

3.1 Model calibration at monthly level

Sunshine duration, total radiation and aerosol concentration at each station are used to calibrate two models in equation (1) with the measured direct and diffuse radiation for model validation. Both empirical models show good performance in estimating monthly direct ($R^2 = 0.86$–0.91) and diffuse ($R^2 = 0.96$) radiation at each station (Figure 4).

Then, measurements from both stations are used to calibrate the two generic models in equation (1) as suggested in equations (1-1) and (1-2), respectively. Figure 5 presents the validations of the calibrated models. Both models perform well in estimating monthly direct ($R^2 = 0.89$) and diffuse ($R^2 = 0.95$) radiation. In combination with the initial validations at each station in Figure 4, two empirical models at monthly level can explain 86–91% of direct radiation variations and 95–96% of diffuse radiation variations, which confirms the reliability and application of both models.

$$R_{direct} = 0.0516 \times SSD \times R + 1.6999 \tag{1 − 1}$$

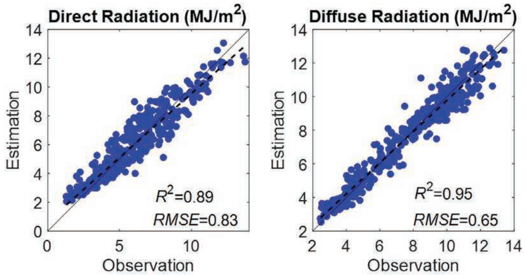

Figure 5. Comparisons of the observed against the estimated monthly direct and diffuse radiation using equations (1-1) and (1-2).

$$\frac{R_{diffuse}}{R} = -0.0387 \times SSD + 0.1778 \times AOD + 0.6362 \quad (1-2)$$

3.2 Model calibration at daily level

3.2.1 Daily direct radiation. The empirical model for daily direct radiation is calibrated and initially validated at each station in S2 and S3 (Table 3). The estimated direct radiation can reach a good agreement with field measurements at both stations under two scenarios.

Then, measurements from two stations are used to calibrate and validate the empirical models for daily direct radiation in S2 and S3 (Tables 4 and 5). Despite the slight differences, similar coefficients and good performance are suggested in S2 ($R^2 \geq 0.88$) and S3 ($R^2 \geq 0.89$) in four years.

Then, three scenarios (S1–S3) are all considered to estimate daily direct radiation at two stations in each year (Figure 6). Results suggest that the empirical model incorporating three scenarios performs well in estimating daily direct radiation ($R^2 = 0.91$–0.94, RMSE $= 1.51$–1.64 MJ/m^2) from 2006 to 2009.

3.2.2 Daily diffuse radiation. Similarly, the empirical model for daily diffuse radiation is calibrated and initially validated at each station in S2 (Table 6). The estimated diffuse radiation can reach a good agreement with field measurements on both stations in S2 in four years.

Then, measurements from two stations are used to calibrate and validate the empirical model for daily diffuse radiation in S2 (Table 7). Despite the differences in coefficients, the estimated daily diffuse radiation suggests a good agreement with field measurements in S2 ($R^2 \geq 0.85$) from 2006 to 2009.

Diffuse radiation in S3 (where daily AOD is not available) is estimated using the method from Erbs et al. (equation (5)), which has been proved as a reliable method in estimating daily diffuse radiation (Feng, Chen and Zhao, 2018). Difference between the improved empirical model and the Erbs et al. method is discussed in Section 3.3.2.

Then, three scenarios (S1–S3) are all considered to estimate daily diffuse radiation over two stations in each year (Figure 7). Results suggest that the empirical model incorporating three scenarios shows a good performance in estimating daily diffuse radiation ($R^2 = 0.91$–0.93, $RMSE = 1.09$–1.15 MJ/m^2) from 2006 to 2009.

3.3 Model validation

3.3.1 Type I validation at monthly and daily level. In the Type I validation, measurements from one station (Beijing, 54511) are used for model calibration and measurements from the other station (Zhengzhou, 57083) are used for model validation (Figure 8). R^2 and *RMSE* suggest high estimation accuracy of four models for estimating direct and diffuse radiation at monthly and daily level. This confirms that the empirical models developed at site (one or two stations) in this study can be applied to a larger region (North China) with reliable estimates of direct and diffuse radiation.

3.3.2 Type II validation at monthly level. Monthly diffuse radiation estimated by the methods from Li et al. (2011, 2012) (equations (6-1) and (6-2)) can reach a good agreement with field

Table 3. Comparisons of the observed and the estimated daily direct radiation in S2 (where daily AOD is available) and S3 (where daily AOD is not available) at each station in each year.

		Y2006		Y2007		Y2008		Y2009	
	Station ID	R^2	*RMSE*	R^2	*RMSE*	R^2	*RMSE*	R^2	*RMSE*
S2	54511	0.90	1.70	0.91	1.88	0.89	1.87	0.93	1.90
	57083	0.89	1.85	0.86	1.82	0.86	1.75	0.89	1.87
S3	54511	0.92	1.19	0.89	1.24	0.92	1.09	0.92	0.75
	57083	0.79	0.93	0.84	0.81	0.83	1.91	0.89	1.23

Table 4. Coefficients in equation (3) to estimate daily direct radiation and the comparisons of the observed and the estimated daily direct radiation by R^2 and *RMSE* in S2 (where daily AOD is available) in four years.

S2	$a(AOD^2)$	$b(AOD)$	c	R^2	*RMSE*
Y2006	0.079	−0.491	0.823	0.89	0.79
Y2007	0.072	−0.464	0.803	0.88	1.99
Y2008	0.071	−0.453	0.800	0.88	1.83
Y2009	0.075	−0.491	0.845	0.91	1.95

Table 5. Coefficients in equation (4) to estimate daily direct radiation and the comparisons of the observed and the estimated daily direct radiation by R^2 and *RMSE* in S3 (where daily AOD is not available) in four years.

S3	$a(SSD^2)$	$b(SSD)$	c	R^2	*RMSE*
Y2006	0.126	−0.198	0.243	0.90	1.11
Y2007	0.104	−0.037	0.178	0.89	1.05
Y2008	0.105	−0.043	0.260	0.89	1.10
Y2009	0.113	−0.097	0.175	0.89	1.11

measurements (Figure 9), whereas the improved model in equation (1) obtains better estimates ($R^2 = 0.96$, $RMSE = 0.57–0.59$ MJ/m^2). Besides, comparisons of the estimates and the measurements suggest that our model can improve the estimation accuracy of monthly maximum diffuse radiation (Figure 10). Additionally, differences of diffuse radiation ($\Delta R_{diffuse}$) between the estimates by equations (6-1) and (6-2) and the measurements are significantly ($p < 0.01$) correlated with monthly AOD (Table 8), which highlights the significance of aerosols in improving the estimation accuracy of diffuse radiation.

According to the previous study at the two stations in North China (Feng, Chen and Zhao, 2018), the Erbs et al. method (1982) can have a good estimate of monthly diffuse radiation ($R^2 = 0.96$, $RMSE = 0.59–0.92$ MJ/m^2) using the clearness index. The method from Chen et al. (1999) can significantly underestimate diffuse radiation ($R^2 = 0.95–0.96$, $RMSE = 1.78–2.55$ MJ/m^2) by not considering aerosols. In addition, another model in the work of Li et al. (2011) also fails to reach a good estimate of the maximum monthly diffuse radiation value ($R^2 = 0.91–0.93$, $RMSE = 0.80–0.93$ MJ/m^2). These comparisons reveal that the estimates of monthly diffuse radiation, particularly the maximum diffuse radiation, can be significantly improved by our empirical model ($R^2 = 0.96$, $RMSE = 0.57–0.59$ MJ/m^2).

3.3.3 Type II validation at daily level. The Erbs et al. (1982) method performs well in estimating daily diffuse radiation ($R^2 = 0.89–0.94$, $RMSE = 1.20–1.28$ MJ/m^2) and therefore can be an optimum method when daily AOD is not available (Figure 11). However, clearness index (K_T), the ratio of total radiation to extraterrestrial radiation, is directly modified by total radiation instead of atmospheric components like aerosols. Besides, extraterrestrial radiation

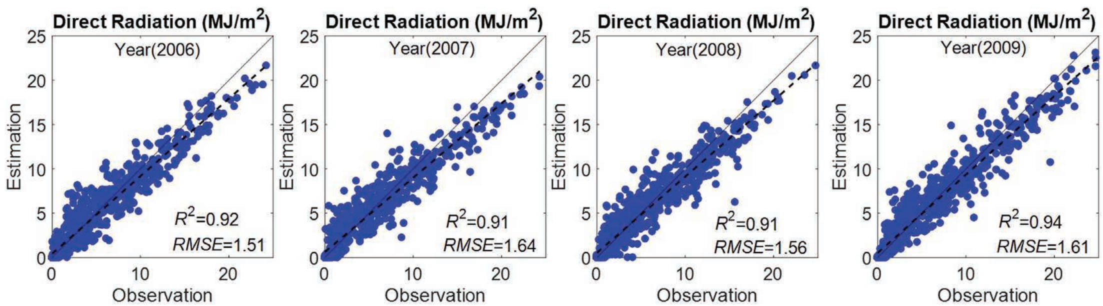

Figure 6. Scatterplots of the observed against the estimated daily direct radiation using the improved empirical model.

Table 6. Comparisons of the observed and the estimated daily diffuse radiation in S2 (where daily AOD is available) at each station in each year.

S2	Y2006		Y2007		Y2008		Y2009	
Station ID	R^2	*RMSE*	R^2	*RMSE*	R^2	*RMSE*	R^2	*RMSE*
54511	0.87	1.39	0.86	1.34	0.89	1.43	0.83	1.09
57083	0.86	1.26	0.86	1.36	0.92	1.18	0.86	1.41

Table 7. Coefficients in equation (3) to estimate daily diffuse radiation in S2 (where daily AOD is available) from 2006 to 2009 and the comparisons of the observed and the estimated daily diffuse radiation indicated by R^2 and *RMSE*.

S2	$a(SSD)$	$b(\log(AOD))$	c	R^2	*RMSE*
Y2006	−0.033	0.169	0.844	0.87	1.33
Y2007	−0.036	0.161	0.866	0.87	1.32
Y2008	−0.033	0.161	0.843	0.90	1.24
Y2009	−0.037	0.159	0.871	0.85	1.38

varies only spatially rather than temporally, which is fixed for a given location. In this case, the Erbs et al. (1982) method relating diffuse fraction ($R_{diffuse}/R$) as a function of K_T limits the estimation accuracy of daily maximum diffuse radiation. Therefore, it is important to consider aerosols when estimating daily diffuse radiation.

Aerosol concentration is also referred when comparing the two methods. Both stations and

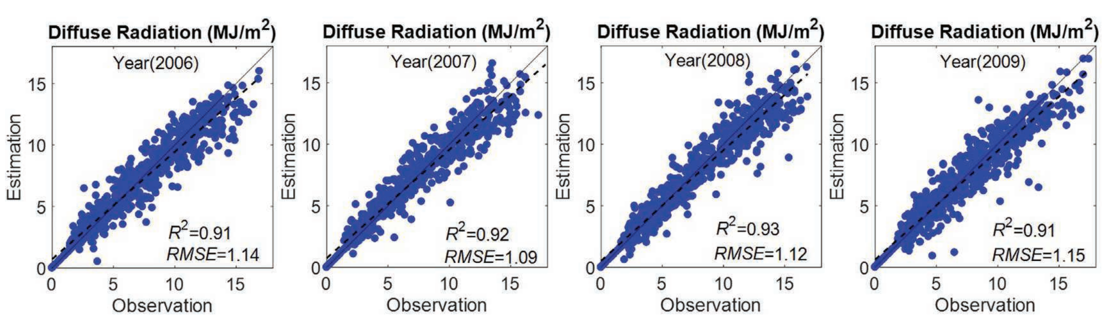

Figure 7. Scatterplots of the observed against the estimated daily diffuse radiation using the improved empirical model.

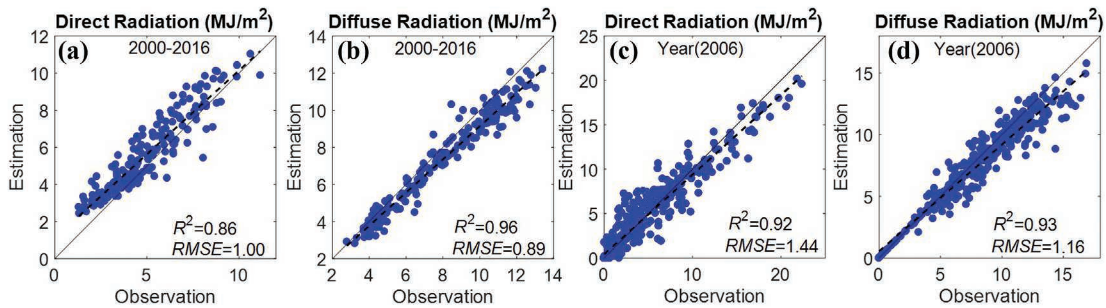

Figure 8. Type I validations of four improved empirical models at monthly (a, b) and daily (c, d) level. Only results in the year 2006 are shown at daily level.

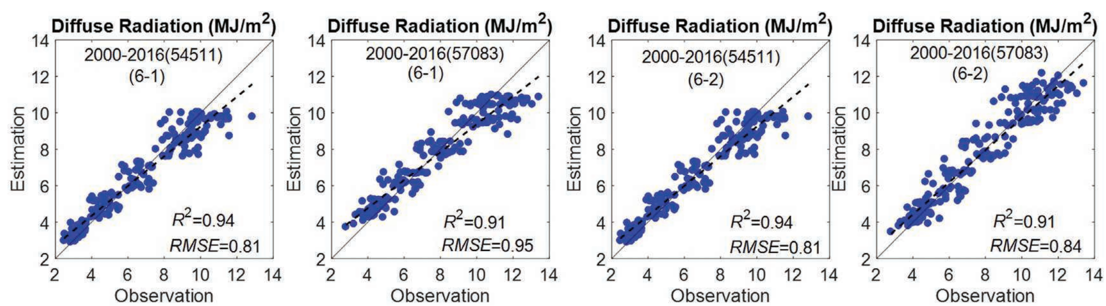

Figure 9. Scatterplots of the observed against the estimated monthly diffuse radiation at each station using equations (6-1) and (6-2).

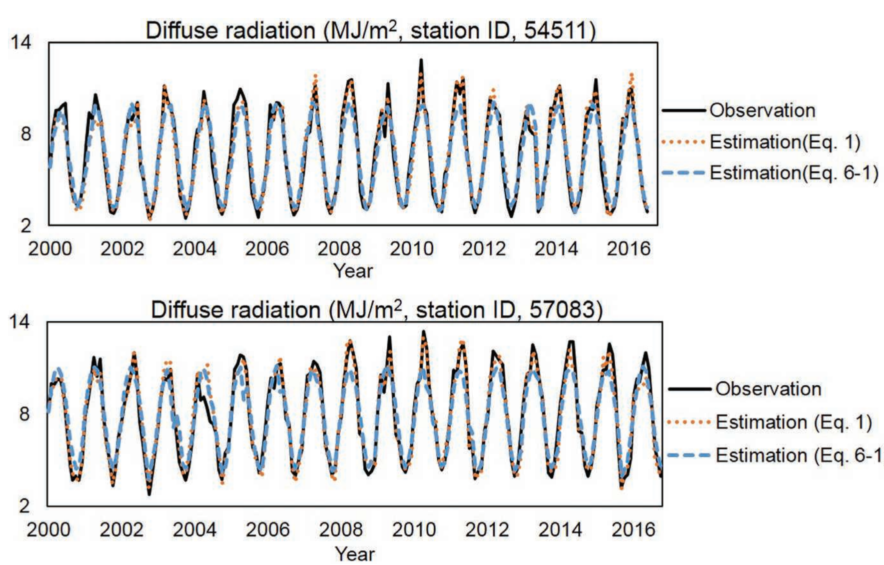

Figure 10. Comparisons of the observed and the estimated monthly diffuse radiation by the improved model in this study equation (1) and the model in previous study at each station. Here, only the model in equation (6-1) is taken for an example.

NC suggest that aerosol concentration is higher in 2006 but lower in 2009 (Figure 12). Compared to the estimates from Erbs et al. (1982) ($R^2 = 0.89$, $RMSE = 1.28$ MJ/m^2), our method shows great improvement in estimating daily diffuse radiation ($R^2 = 0.91$, $RMSE = 1.14$ MJ/m^2) in 2006. In 2009, despite slightly lower R^2 (0.91), still, our method obtains smaller $RMSE$ (1.15 MJ/m^2). Particularly, different from the Erbs et al. (1982) method, our model can improve the estimation accuracy of daily diffuse radiation, particularly the daily maximum diffuse radiation value, by considering different sky conditions. Therefore, it is more applicable in regions with high aerosol concentration.

Table 8. Correlation between Δdiffuse radiation and AOD at monthly level at the two stations.

Station ID	Equation 6-1	Equation 6-2
54511	0.51**	0.33**
57083	0.65**	0.44**

** suggests the correlation coefficients are statistically significant at 0.01 significance level.

IV Conclusions

Empirical models are improved to estimate direct and diffuse radiation at monthly and daily level in heavily polluted regions using the ground-measured total radiation, sunshine duration and satellite-retrieved aerosol concentration. Results of the case study in North China reveal that the improved empirical models perform well in estimating monthly ($R^2 = 0.86$–0.91, $RMSE = 0.76$–0.83 MJ/m^2) and daily ($R^2 = 0.91$–0.94, $RMSE = 1.51$–1.64 MJ/m^2) direct radiation, and the estimation accuracy of monthly ($R^2 = 0.95$–0.96, $RMSE = 0.57$–0.65 MJ/m^2) and daily ($R^2 = 0.91$–0.93, $RMSE$

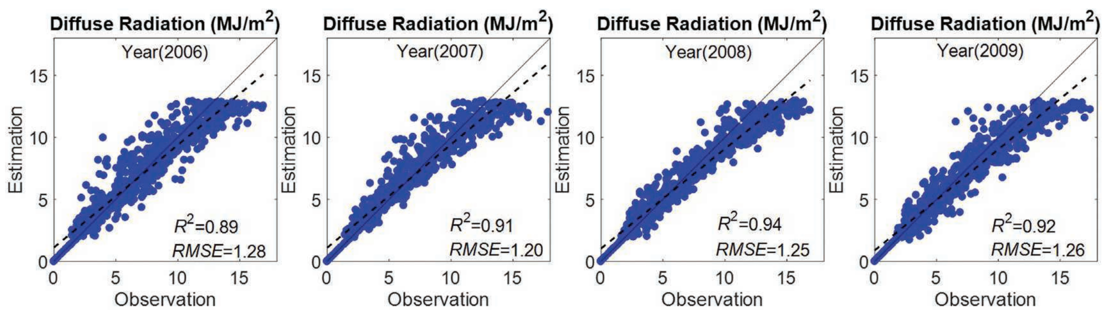

Figure 11. Scatterplots of the observed against the estimated daily diffuse radiation from two stations using equation (5).

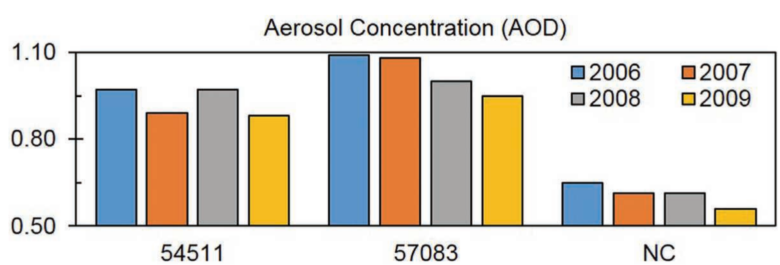

Figure 12. Annual average aerosol optical depth at each station and NC in four years.

$= 1.09–1.15$ MJ/m^2) diffuse radiation, particularly the maximum diffuse radiation value, have been improved compared to the existing methods.

Since the ground-measured total radiation and sunshine duration are more readily available and the satellite-based aerosol concentration is open for access, models presented in this study can be applicable in polluted regions featured by high anthropogenic aerosol concentration similar to North China. However, certain limitations of these empirical models should be mentioned. First, when these empirical models based on two stations are applied at the regional level, the estimates of direct and diffuse radiation are subjected to some uncertainties due to the varying stationary conditions such as the meteorology and air pollution. Despite these uncertainties, these models still provide a reliable improvement in the estimation accuracy of direct and diffuse radiation over the highly-polluted region. Besides, different from the fine-mode anthropogenic aerosols which can increase diffuse fraction, the coarse-mode natural aerosols, such as sands or dust in Northwest China, may reduce diffuse fraction rather than increase it (Feng and Li, 2018). Therefore, the empirical models for diffuse radiation will need further calibration with more ground-measured data over different regions featured by varied aerosol types and levels. Additionally, in regions with relatively good air quality, where the aerosol-effect on solar radiation may be negligible, these models may have no significant improvement in the estimation accuracy of direct and diffuse radiation. Future efforts are still needed to improve those models based on more in situ measurements with specific calibrations for different aerosol types or under different aerosol concentration levels.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Dongmei Chen https://orcid.org/0000-0002-4308-6743

References

Alonso-Montesinos J, Batlles FJ and Bosch JL (2015) Beam, diffuse and global solar irradiance estimation with satellite imagery. *Energy Conversion and Management* 105: 1205–1212. DOI: 10.1016/j.enconman.2015.08.037.

Amillo AG, Huld T and Müller R (2014) A new database of global and direct solar radiation using the eastern meteosat satellite, models and validation. *Remote Sensing* 6(9): 8165–8189. DOI: 10.3390/rs6098165.

Bakirci K and Kirtiloglu Y (2018) Prediction of diffuse solar radiation using satellite data. *International Journal of Green Energy* 15(2): 76–79. DOI: 10.1080/15435075.2018.1423976.

Boers R, Brandsma T and Pier Siebesma A (2017) Impact of aerosols and clouds on decadal trends in all-sky solar radiation over the Netherlands (1966–2015). *Atmospheric Chemistry and Physics* 17(13): 8081–8100. DOI: 10.5194/acp-17-8081-2017.

Boukelia TE, Mecibah MS and Meriche IE (2014) General models for estimation of the monthly mean daily diffuse solar radiation (case study: Algeria). *Energy Conversion and Management* 81: 211–219. DOI: 10.1016/j.enconman.2014.02.035.

Chen J, Liu J, Cihlar J, et al. (1999) Daily canopy photosynthesis model through temporal and spatial scaling for remote sensing applications. *Ecological Modelling* 124(2–3): 99–119. DOI: 10.1016/S0304-3800(99)00156-8.

Chou MD, Lin PH, Ma PL, et al. (2006) Effects of aerosols on the surface solar radiation in a tropical urban area. *Journal of Geophysical Research Atmospheres* 111(15): 1–9. DOI: 10.1029/2005JD006910.

Erbs DG, Klein SA and Duffie JA (1982) Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation. *Solar Energy* 28(4): 293–302. DOI: 10.1016/0038-092X(82)90302-4.

Feng Y and Li Y (2018) Estimated spatiotemporal variability of total, direct and diffuse solar radiation across

China during 1958–2016. *International Journal of Climatology* (April): 1–10. *DOI: 10.1002/joc.5676*.

Feng Y, Chen D and Zhao X (2018) Estimated long-term variability of direct and diffuse solar radiation in North China during 1959–2016. *Theoretical and Applied Climatology*: 1–11. DOI: doi.org/10.1007/s00704-018-2579 -1.

Feng Y, Chen D, Ouyang X, et al. (2018) Variability of satellite-based total aerosols and the relationship with emission, meteorology and landscape in North China during 2000–2016. *Environmental Earth Sciences* 77(499): 1–11. DOI: 10.1007/s12665-018-7685-y.

Fernández-Peruchena C, Gaston M, Guisado M V, et al. (2010) Solar radiation estimation and prediction using measured and predicted aerosol optical depth. *International Conference on Solar Heating, Cooling and Buildings*: 1–7. Available at: http://hal.archives-ouvertes.fr/hal-00919963/.

Frank CW, Wahl S, Keller JD, et al. (2018) Bias correction of a novel European reanalysis data set for solar energy applications. *Solar Energy* 164(September 2017): 12–24. DOI: 10.1016/j.solener.2018.02.012.

Ineichen P (2014) Long term satellite global, beam and diffuse irradiance validation. *Energy Procedia* 48: 1586–1596. DOI: 10.1016/j.egypro.2014.02.179.

Jiang Y (2009) Estimation of monthly mean daily diffuse radiation in China. *Applied Energy* 86(9): 1458–1464. DOI: 10.1016/j.apenergy.2009.01.002.

Kaiser DP (2002) Decreasing trends in sunshine duration over China for 1954–1998: Indication of increased haze pollution? *Geophysical Research Letters* 29(21): 1–4. DOI: 10.1029/2002GL016057.

Kim D and Ramanathan V (2008) Solar radiation budget and radiative forcing due to aerosols and clouds. *Journal of Geophysical Research Atmospheres* 113(2): 1–34. DOI: 10.1029/2007JD008434.

Li H, Bu X, Lian Y, et al. (2012) Further investigation of empirically derived models with multiple predictors in estimating monthly average daily diffuse solar radiation over China. *Renewable Energy* 44: 469–473. DOI: 10.1016/j.renene.2012.01.104.

Li H, Ma W, Wang X, et al. (2011) Estimating monthly average daily diffuse solar radiation with multiple predictors: A case study. *Renewable Energy* 36(7): 1944–1948. DOI: 10.1016/j.renene.2011.01.006.

Liu BYH and Jordan RC (1960) The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. *Solar Energy* 4(3): 1–19. DOI: 10.1016/0038-092X(60)90062-1.

Maghrabi AH (2009) Parameterization of a simple model to estimate monthly global solar radiation based on meteorological variables, and evaluation of existing solar radiation models for Tabouk, Saudi Arabia. *Energy Conversion and Management* 50(11): 2754–2760. DOI: 10.1016/j.enconman.2009.06.024.

Mubiru J and Banda EJKB (2007) Performance of empirical correlations for predicting monthly mean daily diffuse solar radiation values at Kampala, Uganda. *Theoretical and Applied Climatology* 88(1–2): 127–131. DOI: 10.1007/s00704-006-0249 -1.

Paula A, Rocha DA, Goffart J, et al. (2016) On the uncertainty assessment of incident direct solar radiation on building facades due to shading devices. *Energy & Buildings* 133: 295–304. DOI: 10.1016/j.enbuild.2016.09.058.

Pedruzo-Bagazgoitia X, Ouwersloot HG, Sikma M, et al. (2017) Direct and diffuse radiation in the shallow cumulus–vegetation system: Enhanced and decreased evapotranspiration regimes. *Journal of Hydrometeorology* 18(6): 1731–1748. DOI: 10.1175/JHM-D-16-0279.1.

Polo J, Wilbert S, Ruiz-Arias JA, et al. (2016) Preliminary survey on site-adaptation techniques for satellite-derived and reanalysis solar radiation datasets. *Solar Energy* 132: 25–37. DOI: 10.1016/j.solener.2016.03.001.

Qian Y, Wang W, Leung LR, et al. (2007) Variability of solar radiation under cloud-free skies in China: The role of aerosols. *Geophysical Research Letters* 34(12): 1–5. DOI: 10.1029/2006GL028800.

Sabzpooshani M and Mohammadi K (2014) Establishing new empirical models for predicting monthly mean horizontal diffuse solar radiation in city of Isfahan, Iran. *Energy* 69: 571–577. DOI: 10.1016/j.energy.2014.03.051.

Safaripour MH and Mehrabian MA (2011) Predicting the direct, diffuse, and global solar radiation on a horizontal surface and comparing with real data. *Heat and Mass Transfer/Waerme- und Stoffuebertragung* 47(12): 1537–1551. DOI: 10.1007/s00231-011-0814-8.

Segura S, Estelles V and Utrillas MP (2012) Validation of MODIS aerosol optical depth over a Mediterranean and coastal urban environment (Valencia, Spain). 14(February 2015): 13206.**[AQ1]**

Spitters CJT, Toussaint HAJM and Goudriaan J (1986) Separating the diffuse and direct component of global radiation and its implications for modeling canopy photosynthesis: Part I. Components of incoming radiation. *Agricultural and Forest Meteorology* 38(1–3): 217–229. DOI: 10.1016/0168-1923(86)90060-2.

Urraca R, Gracia-Amillo AM, Huld T, et al. (2017) Quality control of global solar radiation data with satellite-based products. *Solar Energy* 158(September): 49–62. DOI: 10.1016/j.solener.2017.09.032.

Urraca R, Huld T, Gracia-Amillo A, et al. (2018) Evaluation of global horizontal irradiance estimates from ERA5 and COSMO-REA6 reanalyses using ground and satellite-based data. *Solar Energy* 164(October 2017): 339–354. DOI: 10.1016/j.solener.2018.02.059.

Xu W, Luo XS, Pan YP, et al. (2015) Quantifying atmospheric nitrogen deposition through a nationwide monitoring network across China. *Atmospheric Chemistry and Physics* 15(21): 12 345–12 360. DOI: 10.5194/acp-15-12345-2015.

Yang K, Koike T and Ye B (2006) Improving estimation of hourly, daily, and monthly solar radiation by importing global data sets. *Agricultural and Forest Meteorology* 137(1–2): 43–55. DOI: 10.1016/j.agrformet.2006.02.001.[AQ2]

Yang Y, Zhao N, Hao X H, et al. (2009) Decreasing trend of sunshine hours and related driving forces in North China. *Theoretical and Applied Climatology* 97(1–2): 91–98. DOI: 10.1007/s00704-012-0583-4.

Yao W, Li Z, Zhao Q, et al. (2015) A new anisotropic diffuse radiation model. *Energy Conversion and Management* 95: 304–313. DOI: 10.1016/j.enconman.2015.01.016.

Yue X and Unger N (2017) Aerosol optical depth thresholds as a tool to assess diffuse radiation fertilization of the land carbon uptake in China. *Atmospheric Chemistry and Physics* 17(2): 1329–1342. DOI: 10.5194/acp-17-1329-2017.

Zeng X, Zhao N and Ma Y (2018) Distributed modeling of surface solar radiation based on aerosol optical depth and sunshine duration in China. *Earth and Environmental Science* 121.[AQ3]

Zhang J, Zhao L, Deng S, et al. (2017) A critical review of the models used to estimate solar radiation. *Renewable and Sustainable Energy Reviews* 70(July 2016): 314–329. DOI: 10.1016/j.rser.2016.11.124.