Applied Energy 343 (2023) 121205

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Short–term global solar radiation forecasting based on an improved method for sunshine duration prediction and public weather forecasts

Shujing Qin [a], Zhihe Liu [b], Rangjian Qiu [a,*], Yufeng Luo [a], Jingwei Wu [a], Baozhong Zhang [c], Lifeng Wu [d], Evgenios Agathokleous [b]

[a] *State Key Laboratory of Water Resources Engineering and Management, Wuhan University, Wuhan, 430072, China*
[b] *Jiangsu Key Laboratory of Agricultural Meteorology, School of Applied Meteorology, Nanjing University of Information Science and Technology, Nanjing 210044, China*
[c] *State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing 100038, China*
[d] *School of Hydraulic and Ecological Engineering, Nanchang Institute of Technology, Nanchang 330099, China*

HIGHLIGHTS

- We proposed a new sunshine duration converting method ($n_{\_new}$) based on forecast temperature and weather types.
- The $n_{\_new}$ method produced better estimates than the common converting method using only weather types.
- The generalized sunshine–based R_s model incorporating the $n_{\_new}$ method improved the accuracy for forecasting daily R_s.

ARTICLE INFO

Keywords:
Maximum and minimum temperature
Sunshine–based model
Sunshine duration
Temperature–based model
Weather types
Solar radiation prediction

ABSTRACT

Accurate forecasting of daily global solar radiation (R_s) is important for photovoltaic power and other sectors. Numerical models coupled with public weather forecasts information is a feasible method to predict short–term daily R_s. Here, we propose a novel sunshine duration converting method ($n_{\_new}$) based on forecasted air temperature and weather types data, which we validated using measurements from 86 radiation stations. A widely-used, generalized sunshine–based R_s model (R_{s_n}) was then coupled with the $n_{\_new}$ method ($R_{s\_n\ new}$) for forecasting daily R_s. This was further compared to R_{s_n} incorporated with the common sunshine duration converting method ($n_{\_com}$) using only weather types data ($R_{s\_n\ com}$) and a recently developed generalized temperature–based model (R_{s_T}). The results indicated that the $n_{\_new}$ method produced better estimates than the $n_{\_com}$ method, as indicated by increased mean correlation coefficient (R; 13.0%–24.5%) and index of agreement (d_{IA}; 2.9%–9.5%) and decreased mean root mean squared error (RMSE; 12.8%–14.8%) for the 1–7 days lead time over 86 sites. The $R_{s\_n\ new}$ model improved the accuracy for 98% of sites when compared to the $R_{s\_n\ com}$ model, with mean values of R and d_{IA} increasing by 7.7%–11.0% and 2.1%–4.8% and that of RMSE decreasing by 9.7%–12.5% for the 1–7 days lead time. The results suggest that the $R_{s\_n\ new}$ model is advantageous in short–term forecasts. The $R_{s\_n\ new}$ model ranked first for 52.3%–74.4% of sites for the 1–7 days lead time, followed by the R_{s_T} model (25.6%–47.7%). Moreover, there was generally a better performance for the $R_{s\_n\ new}$ model to forecast daily R_s at a longer lead time. Therefore, the $R_{s\_n\ new}$ model using weather forecasts information is highly recommended to forecast short–term daily R_s.

1. Introduction

To achieve carbon neutrality and sustainably support the environment and human activities, investments in research and adoptions of renewable energy from carbon–free sources, such as solar radiation, have been extensively carried out [1–4]. Solar radiation (R_s), as one of the most promising clean, renewable, and sustainable resources in nature, has been extensively utilized worldwide to improve the global energy structure [5–9]. Hence, accurate R_s information is of vital importance in application and design of solar energy systems [8,10–12].

* Corresponding author.
E-mail address: qiurangjian@whu.edu.cn (R. Qiu).

https://doi.org/10.1016/j.apenergy.2023.121205
Received 21 February 2023; Received in revised form 6 April 2023; Accepted 21 April 2023

In addition, R_s is also the critical variable for terrestrial energy balance, meteorological dynamic, hydrological cycle, agricultural management, ecosystem stability, epidemiology spread, and industrial development [13,14]. Since R_s measurements face challenges, great efforts have been carried out to predict R_s based on numerical models using widely available meteorological data [8,15].

A number of models have been proposed to estimate R_s, and empirical models are the most extensively applied as a result of lower requirements for computational costs and programming skills and easily accessible input variables [8,16,17] compared to other types of models, such as machine learning [18,19] and remote sensing methods [20]. The empirical models are generally divided into sunshine– [21–24], temperature– [14,25,26], cloudiness–based [27,28], and comprehensive models [7,9,11] depending upon availability of data and research objectives [7,29]. Among them, sunshine– and temperature–based models have been extensively implemented because of good correlations between R_s and sunshine duration / air temperature (T_a, ℃) [12,17]. The notable sunshine–based Angstrom–Prescott model [30,31], which was constructed on the basis of linear correlation between R_s and the proportion of actual sunshine duration (n, h) to the maximum possible sunshine duration (N, h), has been extensively applied to estimate R_s and was suggested in FAO56 [21]. Subsequently, quadratic, cubic, trigonometric, logarithmic, or exponential relationships between R_s and n/N were proposed for sunshine–based models [7,32]. Liu et al. recently compared the 32 existing sunshine–based R_s models and revealed that the model proposed by Newland [33] provides a compromise between model accuracy and complexity [34]. However, n data are not always widely available. On the contrast, the temperature data are readily measured and recorded worldwide. Hence, the temperature–based models are also widely utilized, which are generally developed on the basis of different combinations of mathematic modes and various input variables of T_a [7,29]. For instance, the notable H–S model considers R_s as a function of diurnal temperature range ($T_d = T_{max}–T_{min}$, ℃) [35]. Bristow and Campbell [36] established an exponential relationship between R_s and T_d to improve model accuracy. Fan et al. [16] enhanced R_s estimation by introducing the mean air temperature (T_{mean}, ℃). Qiu et al. [8] recently proposed four temperature–based R_s models by incorporating maximum air temperature (T_{max}, ℃), minimum air temperature (T_{min}, ℃), T_d, and T_{mean} after reviewing 78 existing temperature–based models.

Short–term forecasting information of daily R_s is important for photovoltaic power, solar energy resource management, and other sectors. Mathematic models combined with public weather forecasting information may be a feasible approach to predict short–term daily R_s because public weather forecasts are easily accessible to public worldwide, providing data of T_a, weather types, and wind speed. The forecasting accuracy depends not only on the quality of the models but also on the quality of the input data. Generally, sunshine–based R_s models produce an extraordinary performance in estimating daily R_s and outperform the temperature–based models using measured meteorological data [17,37,38]. Conversely, public weather forecasts provide more accurate T_a information than weather types (can be converted into n) [39–41]. Hence, it is unclear which type of models is more appropriate for forecasting daily R_s. In addition, the former n converting method is based on only data of weather types, suggesting a great uncertainty. Since weather forecasts can provide reasonable estimates of T_a [39–41], and T_d can indirectly reflect the cloudiness [25,36], we seek to reveal whether the accuracy of converted n can be further improved based on forecasted weather types and T_a data, in turn improving R_s forecasts. Therefore, the main objectives for this study were to: (1) propose a novel n converting method using both forecasted weather types and temperature data; (2) evaluate the performance of daily R_s forecasted from either temperature–based R_s model or sunshine–based R_s model integrated with common and proposed n converting methods, with the goal to recommend the best model to be adopted for forecasting short–term daily R_s.

2. Materials and methods

2.1. Datasets

Measured daily meteorological data consisting of R_s (MJ m^{-2} d^{-1}), T_{max} and T_{min} (℃), and n (h) were collected from 86 radiation stations in China during the period 2015–2019. All data applied in this study have been strictly controlled for quality by the China Meteorological Data Service Centre (https://data.cma.cn). In addition, data for a given day were removed when (1) daily R_s higher than corresponding extraterrestrial radiation (R_a, MJ m^{-2} d^{-1}) and (2) $n = 0$, while $R_s > 10$ MJ m^{-2} d^{-1}. To successfully operate the applied logarithmic sunshine–based R_s model, the data of n equal to 0 were set to 1×10^{-7}.

The public weather forecasts data for these 86 sites were obtained from the website of Weather China (https://www.weather.com.cn) for 1–7 days lead time, and from the website of Historic Weather (https://www.tianqihoubao.com) for current–day (i.e. current day weather recorded at the end of day). These data included daily T_{max} and T_{min} and weather types during the period 2015–2019. The daily data for a given day were removed when (1) any of data of T_{max}, T_{min}, and weather types from weather forecast were missing; (2) $T_{max} < T_{min}$; and (3) measured T_{max}, T_{min}, n and R_s from radiation stations were missing.

2.2. R_s estimation models

2.2.1. Generalized temperature–based R_s model

Qiu et al. [8] recently proposed a new temperature–based model, after testing 78 existing models, which was used to estimate daily R_s (noted as R_{s_T}) in this study:

$$R_{s\_T} = a_1 T_{\text{mean}} + \left[a_2 + a_3 T_{\max} + a_4 T_{\min} + a_5 (T_d)^2 + a_6 (T_{\max})^2 + a_7 (T_{\max})^2 T_d + a_8 T_{\max} (T_d)^2\right] R_a \tag{1}$$

where a_1–a_8 are the empirical parameters with generalized values of 0.4397, –0.0363, 0.0603, –0.0944, –0.0024, 0.0008, 0.0003, and –0.0001, respectively [8]; R_a for a given day at any geographic latitude (φ, rad) can be calculated as [21]:

$$R_a = 37.6 d_r (\omega_0 \sin\varphi \sin\delta + \cos\varphi \cos\delta \sin\omega_0) \tag{2}$$

$$d_r = 1 + 0.033\cos\left(\frac{2\pi}{365} J\right) \tag{3}$$

$$\omega_0 = \arccos[-\tan(\varphi)\tan(\delta)] \tag{4}$$

$$\delta = 0.4093\sin\left(\frac{2\pi}{365} J - 1.39\right) \tag{5}$$

where ω_0 and δ are the sunset hour angle and the solar declination (rad), respectively; d_r is the inverse relative distance between the Sun and the Earth; and J is Julian day.

2.2.2. Generalized sunshine–based R_s model

Liu et al. [34] recently evaluated 32 sunshine–based models (noted as R_{s_n}) and suggested that the model proposed by Newland [33] provides a compromise between model accuracy and complexity, which was used here:

$$R_{s\_n} = \left[b_1 + b_2 \frac{n}{N} + b_3 \ln\left(\frac{n}{N}\right)\right] R_a \tag{6}$$

$$N = \frac{24}{\pi} \omega_0 \tag{7}$$

where b_1–b_3 are the empirical parameters, and their generalized values reported are 0.2171, 0.5183, and 0.0022, respectively [34].

2.3. Improved method for estimating sunshine duration using public weather forecasts

The common method to estimate n from weather forecasts is converting from the only weather types data (labeled n_{com}) [42,43], as follows:

$$n_{com} = \alpha N \tag{8}$$

where α is the coefficient of converting from weather types to n. The values of 0.9, 0.7, 0.5, 0.3, and 0.1 for α are recommended corresponding to clear, clear to cloudy, cloudy, overcast, and rainy days [42,43].

We noted that weather forecasts can provide reasonable accuracy for T_{max} and T_{min} [43], and there is a robust relationship between n and T_d based on pooled measured data from 86 stations during 2015 and 2019 (Fig. 1). Hence, we proposed an improved n converting method (noted as n_{new}) incorporating T_d, α, and N, which can be easily obtained from public weather forecasts:

$$n_{new} = \max((c_1\alpha + c_2T_d + c_3)N,\ 0) \tag{9}$$

where c_1–c_3 are empirical coefficients. The n_{new} was calibrated based on pooled current–day weather data (T_d and weather types) from all 86 sites to obtain the generalized c_1–c_3 (0.4914, 0.0283, and –0.0272, respectively). The estimated values of n_{com} and n_{new} for the 1–7 days lead were then used as inputs for the generalized sunshine–based R_s model to predict daily R_s (noted as $R_{s\_n\ com}$, and $R_{s\_n\ new}$, respectively). The flowchart of computational procedure for improving forecasted accuracy of sunshine–based R_s model by integrating the n_{new} converting method is shown in Fig. 2.

2.4. Evaluation indicators

Four statistical indicators were used to assess the model performance, i.e. the regression coefficient (b), the correlation coefficient (R), the root mean squared error (RMSE), and the index of agreement (d_{IA}) [44,45]:

$$b = \sum_{i=1}^{n} M_i\,P_i \Big/ \sum_{i=1}^{n} M_i^2 \tag{10}$$

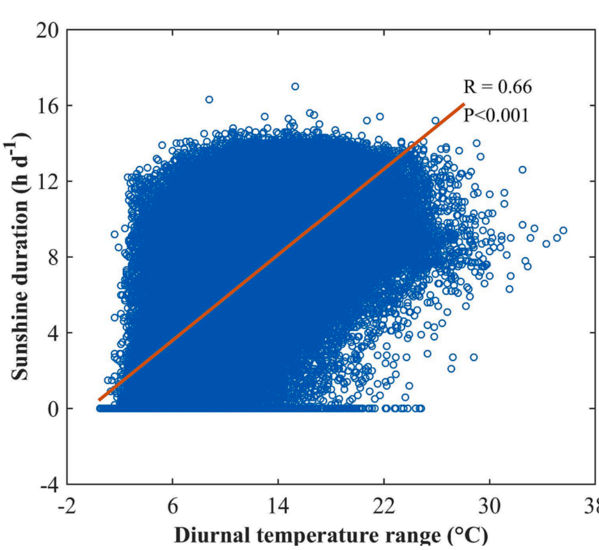

Fig. 1. Correlation between sunshine duration and diurnal temperature range based on pooled data from 86 radiation stations in China during 2015 and 2019.

$$R = \frac{\sum_{i=1}^{n}(M_i - \overline{M})(E_i - \overline{E})}{\sqrt{\sum_{i=1}^{n}(M_i - \overline{M})^2 \sum_{i=1}^{n}(E_i - \overline{E})^2}} \tag{11}$$

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n}(E_i - M_i)^2\right]^{0.5} \tag{12}$$

$$d_{IA} = 1 - \frac{\sum_{i=1}^{N}(M_i - E_i)^2}{\sum_{i=1}^{N}(|E_i - \overline{M}| + |M_i - \overline{M}|)^2} \tag{13}$$

where M_i and $\overline{M}$ are measured values and their averages; and E_i and $\overline{E}$ are the estimated values and their averages. A better performance can be generated when b, R, and $d_{IA} \approx 1.0$ and RMSE ≈ 0 [8,46].

To better compare the performance of different models, we further applied global performance indicator (GPI) to rank the model performance of more than two models [17,32]:

$$GPI_i = \sum_{j=1}^{4}\alpha_j(g_j - y_{ij}) \tag{14}$$

where j is the number of statistical indicators, α_j equals to 1 for RMSE, and –1 for b, R, and d_{IA}; g_j denotes the median of scaled values of statistical indicator j; y_{ij} denotes the scaled values of the statistical indicators j for model i. Greater GPI value shows better model accuracy.

3. Results and discussion

3.1. Performance of R_s estimated with generalized R_{s_T} and R_{s_n} models using measurements

The performance of generalized R_{s_T} and R_{s_n} models was firstly evaluated using measured data from each radiation station, as shown in Fig. 3. Both types of models can reasonably estimate daily R_s over all stations. The mean values of b, R, RMSE, and d_{IA} over the 86 radiation stations were 0.987, 0.841, 4.315 MJ m^{-2} d^{-1}, and 0.893, respectively, for the generalized R_{s_T} model. Likewise, they were 1.025, 0.939, 2.740 MJ m^{-2} d^{-1}, and 0.956, respectively, for the generalized R_{s_n} model. The above statistical indicators and Fig. 3 also revealed that better performance was generally observed for the generalized R_{s_n} model than the R_{s_T} model for majority stations (98% of all stations), which is in line with previous studies [17,37,38]. Trnka et al. [47] reported that the R_{s_T} models outperformed the cloud–, rainfall–, and temperature–based models in lowlands of the Czech Republic and Austria. Feng et al. [38] also showed that R_{s_n} models have better performance than R_{s_T} models in China. Overall, the results suggest that better model performance can be generated from generalized sunshine–based R_s model than temperature–based model when using measured data. Hence, the predicting accuracy of daily R_s is highly depended on forecasted T_a and weather types, as described in the following sections.

3.2. Comparison of T_{max} and T_{min} between the weather forecasts and measurements

Forecasting accuracy for daily T_{max} and T_{min} is critical when predicting daily R_s. Hence, T_{max} and T_{min} for 1–7 days lead time provided from the weather forecasting system were evaluated with measurements from the radiation stations, as shown in Fig. 4 and Table 1. Results indicated that the weather forecasts can well predict the daily T_{max} and T_{min} for 1–7 days lead time for the 86 sites (Fig. 4), although there was a decreasing trend of accuracy as lead time increased (Fig. 4**,** Table 1). For T_{max} forecasts, values of b, R, RMSE, and d_{IA} varied from 0.898 to 1.044, 0.875–0.994, 1.275–4.303 °C d^{-1}, and 0.933–0.997, respectively, for a 1–day lead time to 0.841–1.021, 0.788–0.974, 2.201–5.290 °C d^{-1}, and 0.871–0.986, respectively, for a 7–day lead time in the 86 sites (Fig. 4).

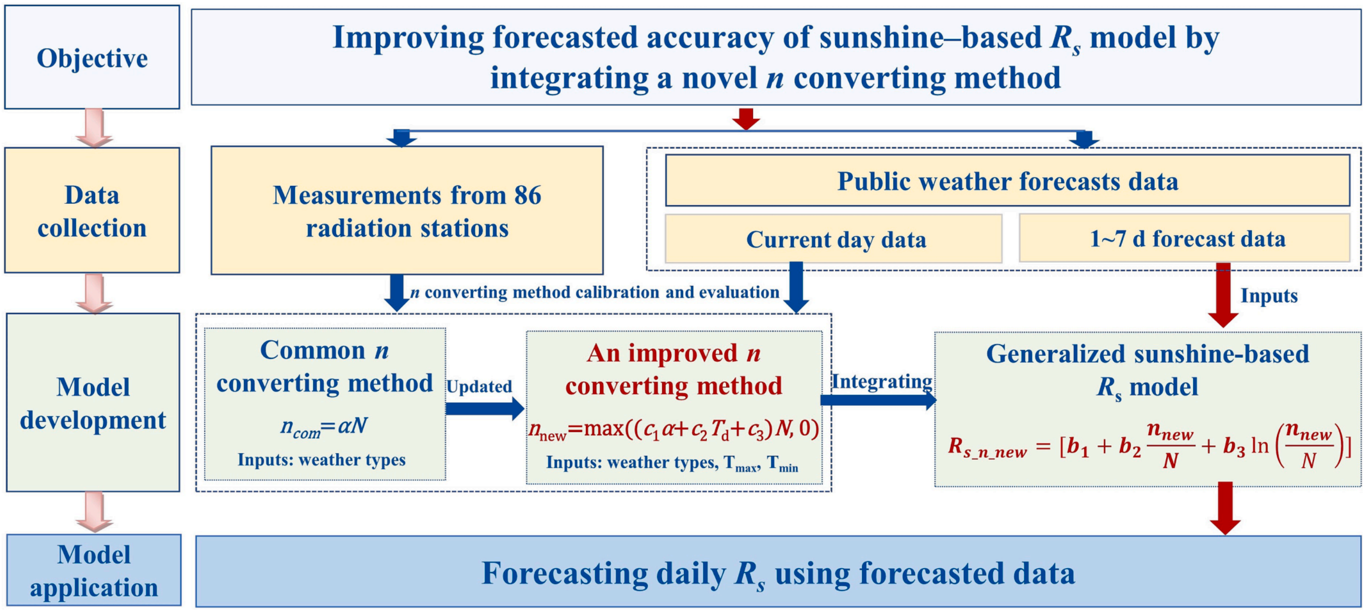

Fig. 2. Flowchart of computational procedure for improving forecasted accuracy of sunshine-based R_s model in this study. R_s is global solar radiation, n is the sunshine duration, N is the maximum possible sunshine duration, α is the coefficient of converting from weather types to n. T_d is diurnal temperature range, T_{max} and T_{min} is the maximum and minimum temperature. $b_1 \sim b_3$ (0.2171, 0.5183, and 0.0022, respectively) and $c_1 \sim c_3$ (0.4914, 0.0283, and –0.0272, respectively) are empirical coefficients.

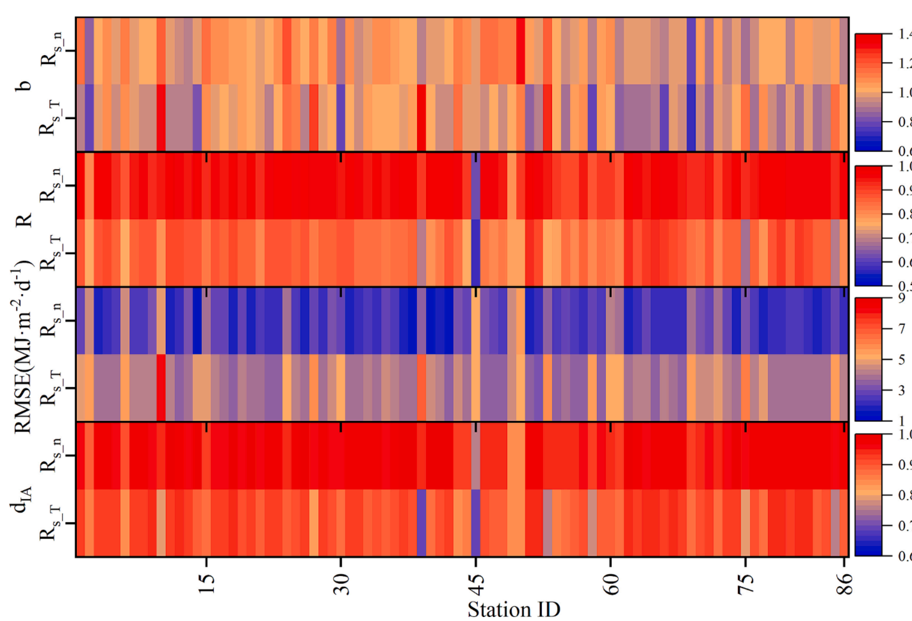

Fig. 3. Performance of solar radiation estimated by the generalized temperature– (R_{s_T}) and sunshine–based models (R_{s_n}), respectively, driving by measured data from 86 radiation stations in China during the period 2015–2019. b is the slope of the regression function; R is the correlation coefficient; RMSE is the root mean squared error; and d_{IA} is the index of agreement.

For T_{min} estimation, they were varied from 0.862 to 1.056, 0.856–0.991, 1.173–5.314 °C, and 0.917–0.995, respectively, for the 1–day lead time to 0.850–1.070, 0.833–0.973, 1.778–5.755 °C, and 0.900–0.986, respectively, for the 7–day lead time in the 86 sites (Fig. 4).

The results here supported the finding that the weather forecasting system can accurately forecast both daily T_{max} and T_{min}, which is in line with previous studies [8,40,48,49]. For instance, Yang et al. [40] reported that the values of RMSE and R ranged 2.52–5.03 °C and 0.88–0.97 for T_{max} and 1.24–4.76 °C and 0.95–0.99 for T_{min} for the lead times of 1–7 days at eight cities during 2012–2014. The forecasting accuracy of T_{max} and T_{min} was also reasonable for the 1–7 days lead time during 2012–2016 in an assessment of forecasting performance for 61 cities distributed in varying climate zones of China [40]. These studies show that T_{max} and T_{min} with an 1–7 days lead time can be widely well forecasted, which will benefit forecasting daily R_s for both temperature– and sunshine–based models.

3.3. *Performance of common and novel sunshine duration converting methods*

Performance of sunshine–based R_s models highly depends on the accuracy of converted n from weather forecasts. Fig. 5 and Table 2 show

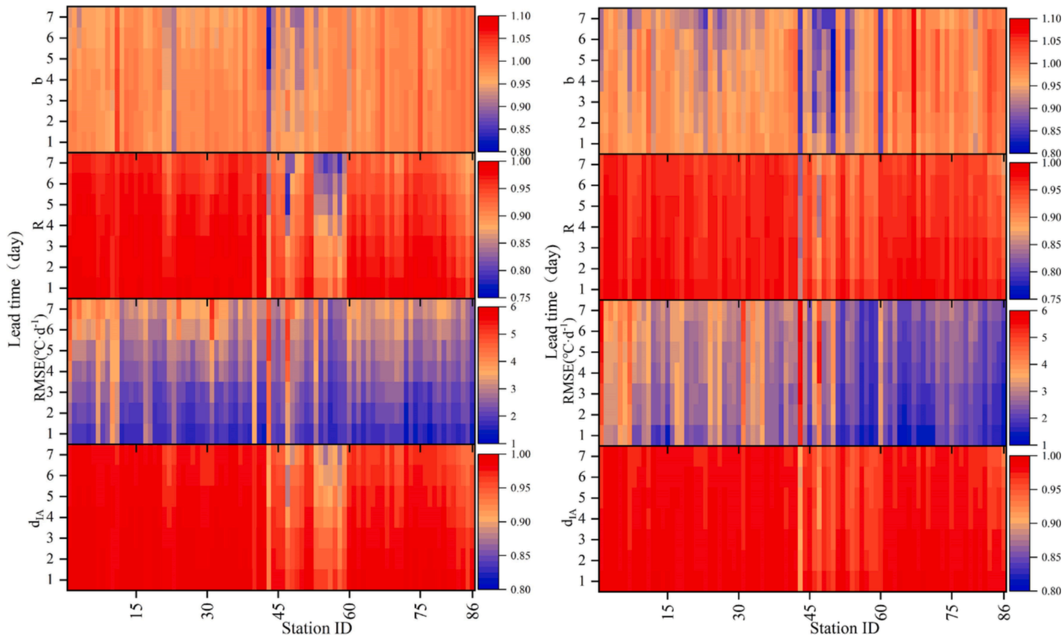

Fig. 4. Statistical indicators for forecasting performance of maximum (left) and minimum (right) temperatures for 1–7 days lead time during the period 2015–2019 for 86 sites in China. b is the slope of the regression function; R is the correlation coefficient; RMSE is the root mean squared error; and d_{IA} is the index of agreement.

Table 1
The overall goodness of fit statistical indicators when comparing the daily values of maximum (T_{max}) and minimum (T_{min}) air temperatures from weather forecasts and radiation stations measurements for 1–7 days lead time during the period 2015–2019. Values are the means ± SD (n = 86). b is the slope of the regression function; R is the correlation coefficient; RMSE is the root mean squared error; d_{IA} is the index of agreement.

	Indicator	Lead days (day)						
		1	2	3	4	5	6	7
T_{max}	b	0.981 ± 0.020	0.978 ± 0.023	0.978 ± 0.022	0.975 ± 0.025	0.972 ± 0.028	0.968 ± 0.029	0.964 ± 0.029
	R	0.974 ± 0.024	0.968 ± 0.027	0.963 ± 0.030	0.954 ± 0.035	0.946 ± 0.041	0.938 ± 0.043	0.928 ± 0.046
	RMSE (°C d^{-1})	1.978 ± 0.542	2.215 ± 0.498	2.386 ± 0.488	2.675 ± 0.482	2.886 ± 0.523	3.164 ± 0.572	3.424 ± 0.578
	d_{IA}	0.986 ± 0.013	0.982 ± 0.015	0.980 ± 0.017	0.975 ± 0.020	0.971 ± 0.024	0.966 ± 0.025	0.960 ± 0.028
T_{min}	b	0.969 ± 0.030	0.965 ± 0.036	0.965 ± 0.040	0.963 ± 0.042	0.961 ± 0.044	0.958 ± 0.045	0.945 ± 0.043
	R	0.969 ± 0.018	0.962 ± 0.020	0.960 ± 0.021	0.955 ± 0.023	0.954 ± 0.023	0.951 ± 0.024	0.947 ± 0.024
	RMSE (°C d^{-1})	2.423 ± 0.799	2.641 ± 0.784	2.687 ± 0.792	2.842 ± 0.802	2.879 ± 0.779	2.962 ± 0.809	3.049 ± 0.768
	d_{IA}	0.983 ± 0.010	0.979 ± 0.011	0.979 ± 0.012	0.976 ± 0.013	0.975 ± 0.013	0.973 ± 0.014	0.971 ± 0.014

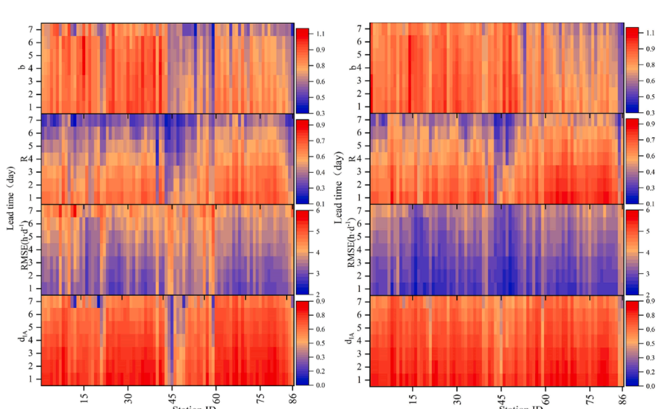

Fig. 5. Forecasting performance of converted sunshine duration from weather forecasts based on common (left) and improved sunshine duration models (right), respectively, for 1–7 days lead time during the period 2015–2019 for the 86 sites in China. b is the slope of the regression function; R is the correlation coefficient; RMSE is the root mean squared error; and d_{IA} is the index of agreement.

the comparison between the n_{com} and n_{new} methods for the 86 sites. Overall, the n_{com} method underestimates n by 20%–34% (Table 2), and the mean values of R, RMSE, and d_{IA} for the n_{com} method were within the ranges of 0.326–0.609, 3.256–4.204 h d^{-1}, and 0.514–0.732, respectively, for the 1–7 days lead time over 86 sites (Table 2). In addition, the predicted accuracy was decreased as lead time increased. Generally, the n_{new} method provided an acceptable performance and improved the accuracy relative to the n_{com} method for the 1–7 days lead as indicated by b closer to 1.00, higher R and d_{IA}, and lower RMSE (Table 2**,** Fig. 5). Overall, the n_{new} method increased mean R and d_{IA} by 13.0%–24.5% and 2.9%–9.5%, and decreased mean RMSE by 12.8%–14.8%, for the 1–7 days lead time compared to the n_{com} model for the 1–7 days lead time (Table 2).

The n_{new} method improves the n prediction mainly by introducing T_d into the calculation, which consists with the pre–analyzed close relationship between daily n and T_d based on measurement from the 86 stations during 2015 and 2019 (Fig. 1). This is because T_d is directly correlated with atmospheric transmittance [36] and can serve as an indicator of cloudiness by assuming that the clear skies will increase T_{max} due to higher shortwave radiation while decrease T_{min} due to higher transmissivity [25,50]. Therefore, T_d is also commonly introduced in many temperature–based R_s models. For instance, Qiu et al. [8]

Table 2
Overall forecasting performance for sunshine duration based on the common (n_{com}) and improved (n_{new}) methods for 1–7 days lead time during the period 2015–2019. Values are the means ± SD (n = 86). b is the slope of the regression function; R is the correlation coefficient; RMSE is the root mean squared error; d_{IA} is the index of agreement.

Models	Indicator	Lead days (day)						
		1	2	3	4	5	6	7
n_{com}	b	0.798 ± 0.103	0.783 ± 0.105	0.773 ± 0.108	0.763 ± 0.119	0.749 ± 0.122	0.738 ± 0.124	0.664 ± 0.127
	R	0.609 ± 0.094	0.572 ± 0.092	0.546 ± 0.088	0.488 ± 0.084	0.446 ± 0.086	0.400 ± 0.081	0.326 ± 0.085
	RMSE (h d^{-1})	3.256 ± 0.401	3.383 ± 0.381	3.470 ± 0.364	3.660 ± 0.359	3.789 ± 0.351	3.927 ± 0.341	4.204 ± 0.410
	d_{IA}	0.732 ± 0.104	0.707 ± 0.106	0.689 ± 0.108	0.651 ± 0.112	0.623 ± 0.114	0.594 ± 0.108	0.514 ± 0.141
n_{new}	b	0.823 ± 0.096	0.807 ± 0.096	0.797 ± 0.099	0.781 ± 0.103	0.770 ± 0.103	0.760 ± 0.107	0.726 ± 0.108
	R	0.688 ± 0.093	0.647 ± 0.087	0.617 ± 0.087	0.557 ± 0.087	0.516 ± 0.091	0.470 ± 0.091	0.406 ± 0.091
	RMSE (h d^{-1})	2.811 ± 0.295	2.940 ± 0.276	3.025 ± 0.268	3.185 ± 0.271	3.288 ± 0.276	3.405 ± 0.285	3.581 ± 0.312
	d_{IA}	0.759 ± 0.068	0.730 ± 0.070	0.709 ± 0.071	0.670 ± 0.073	0.642 ± 0.076	0.613 ± 0.073	0.563 ± 0.084

found that 63% of collected temperature–based R_s models (totally 78) employed T_d as an input variable. Almorox et al. [25] indicated that the temperature–based R_s models including only T_{max}, T_{min}, or T_{mean} as single input variable of temperature generated greater deviation than those using T_d. Besides, another convincing reason to incorporate T_d into forecasting n is that T_a is generally accurately forecasted in the weather forecast systems, as shown in this study (Fig. 3) and elsewhere [8,41,48]. Thus, the potential advantage of sunshine–based R_s model using newly developed n estimation method involving T_d was demonstrated.

3.4. Comparison of R_s forecasted from generalized temperature– and sunshine–based models using public weather forecasts

Performance of daily R_s forecasted from R_{s_T}, $R_{s\_n\ com}$, and $R_{s\_n\ new}$ was evaluated using measurements from the 86 radiation stations, as shown in Table 3 and Fig. 6. The $R_{s\_n\ com}$ model overestimated daily R_s by 9.0%–13.5%, and mean values of R, RMSE, and d_{IA} were within the ranges of 0.588–0.741, 5.264–6.572 MJ m^{-2} d^{-1}, and 0.705–0.798, respectively, for the 1–7 days lead time over the 86 sites (Table 3). Compared to the $R_{s\_n\ com}$ model, the $R_{s\_n\ new}$ model improved the accuracy for 98% of the stations for 1–7 days lead time and produced an acceptable accuracy for forecasting daily R_s with an overestimation of only 3.3%–4.5% (Table 3, Fig. 6). Overall, the $R_{s\_n\ new}$ model increased mean R and d_{IA} by 7.7%–11.0% and 2.1%–4.8%, respectively, and decreased mean RMSE by 9.7%–12.5%, for 1–7 days lead time compared to the $R_{s\_n\ com}$ model (Table 3). The R_{s_T} model slightly underestimated daily R_s by 2.1%–4.3% (Table 3), and mean values of R, RMSE, and d_{IA} were 0.613–0.790, 4.744–6.086 MJ m^{-2} d^{-1}, and 0.752–0.855, respectively, for the 1–7 days lead time over the 86 sites. The overall accuracy of the R_{s_T} model was superior to that of the $R_{s\_n\ com}$ model, as indicated by b closer to 1.00, higher R and d_{IA}, and lower RMSE. However, the R_{s_T} model produced higher mean RMSE and lower mean R, but higher mean d_{IA} and closer b relative to the $R_{s\_n\ new}$ model. In addition, similar to forecasted T_{max}, T_{min}, and converted n by the n_{com} and n_{new} models, the forecasted accuracy of daily R_s from all the three models decreased gradually as forecasting lead time increased in all sites (Fig. 6, Table 3). For instance, mean values of R and d_{IA} for the $R_{s\_n\ new}$ model decreased from 0.520 to 0.918 and 0.565–0.949, in a 1–day lead time to 0.364–0.830, and 0.458–0.895, respectively, in a 7–day lead time, while RMSE increased from 3.250 to 7.863 to 4.294–7.820 MJ m^{-2} d^{-1} (Fig. 6).

The GPI was further employed to rank the model performance for forecasting daily R_s. Overall, the $R_{s\_n\ new}$ model was advantageous and outperformed the other two models in a short–term forecast for as many as 52.3%–74.4% of the stations for the 1–7 days lead time. The R_{s_T} model ranked second in terms of dominance, ranking top for 25.6%–47.7% of all stations for the 1–7 days lead time. Only in few cases the $R_{s\_n\ com}$ model outperformed the other two models (Fig. 7). In addition, the percentage of stations ranking first generally showed an increased trend as lead time increased for the $R_{s\_n\ new}$ model but a decreased trend for the R_{s_T} model, indicating a better performance for the $R_{s\_n\ new}$ model to forecast R_s at a longer lead time. Although the $R_{s\_n\ new}$ model is not superior to the other two models for all the sites reported here, the potential of sunshine–based R_s model pronounced in the context of continuous improvement of weather forecasting accuracy and n

Table 3
Overall performance of forecasted solar radiation by using generalized sunshine–based model coupled with common ($R_{s\_n\ com}$) and improved sunshine duration converted methods ($R_{s\_n\ new}$) and generalized temperature–based model (R_{s_T}), respectively, for 1–7 days lead time during the period 2015–2019 for 86 sites in China. Values are the means ± SD (n = 86). b is the slope of the regression function; R is the correlation coefficient; RMSE is the root mean squared error; and d_{IA} is the index of agreement.

Model	Indicators	Lead time (day)						
		1	2	3	4	5	6	7
$R_{s\_n\ com}$	b	1.099 ± 0.137	1.103 ± 0.140	1.101 ± 0.140	1.096 ± 0.146	1.095 ± 0.147	1.090 ± 0.148	1.135 ± 0.156
	R	0.741 ± 0.119	0.720 ± 0.122	0.705 ± 0.125	0.677 ± 0.127	0.651 ± 0.129	0.627 ± 0.130	0.588 ± 0.130
	RMSE (MJ m^{-2} d^{-1})	5.264 ± 0.997	5.442 ± 0.965	5.565 ± 0.950	5.788 ± 0.922	5.966 ± 0.899	6.157 ± 0.871	6.572 ± 0.857
	d_{IA}	0.812 ± 0.100	0.798 ± 0.103	0.789 ± 0.104	0.773 ± 0.107	0.758 ± 0.110	0.744 ± 0.109	0.705 ± 0.110
$R_{s\_n\ new}$	b	1.035 ± 0.098	1.038 ± 0.100	1.038 ± 0.099	1.035 ± 0.102	1.035 ± 0.103	1.033 ± 0.103	1.045 ± 0.105
	R	0.801 ± 0.086	0.779 ± 0.087	0.762 ± 0.093	0.729 ± 0.100	0.708 ± 0.105	0.684 ± 0.109	0.653 ± 0.114
	RMSE (MJ m^{-2} d^{-1})	4.700 ± 0.776	4.878 ± 0.754	4.998 ± 0.741	5.228 ± 0.732	5.373 ± 0.722	5.529 ± 0.710	5.753 ± 0.700
	d_{IA}	0.835 ± 0.080	0.820 ± 0.084	0.809 ± 0.087	0.789 ± 0.094	0.775 ± 0.098	0.762 ± 0.100	0.739 ± 0.104
R_{s_T}	b	0.979 ± 0.112	0.979 ± 0.115	0.979 ± 0.115	0.977 ± 0.118	0.975 ± 0.120	0.973 ± 0.122	0.957 ± 0.116
	R	0.790 ± 0.086	0.762 ± 0.085	0.741 ± 0.092	0.695 ± 0.105	0.673 ± 0.112	0.646 ± 0.118	0.613 ± 0.127
	RMSE (MJ m^{-2} d^{-1})	4.744 ± 0.804	4.995 ± 0.778	5.149 ± 0.760	5.509 ± 0.806	5.664 ± 0.821	5.849 ± 0.841	6.086 ± 0.874
	d_{IA}	0.855 ± 0.072	0.837 ± 0.074	0.824 ± 0.078	0.798 ± 0.085	0.784 ± 0.089	0.768 ± 0.093	0.752 ± 0.095

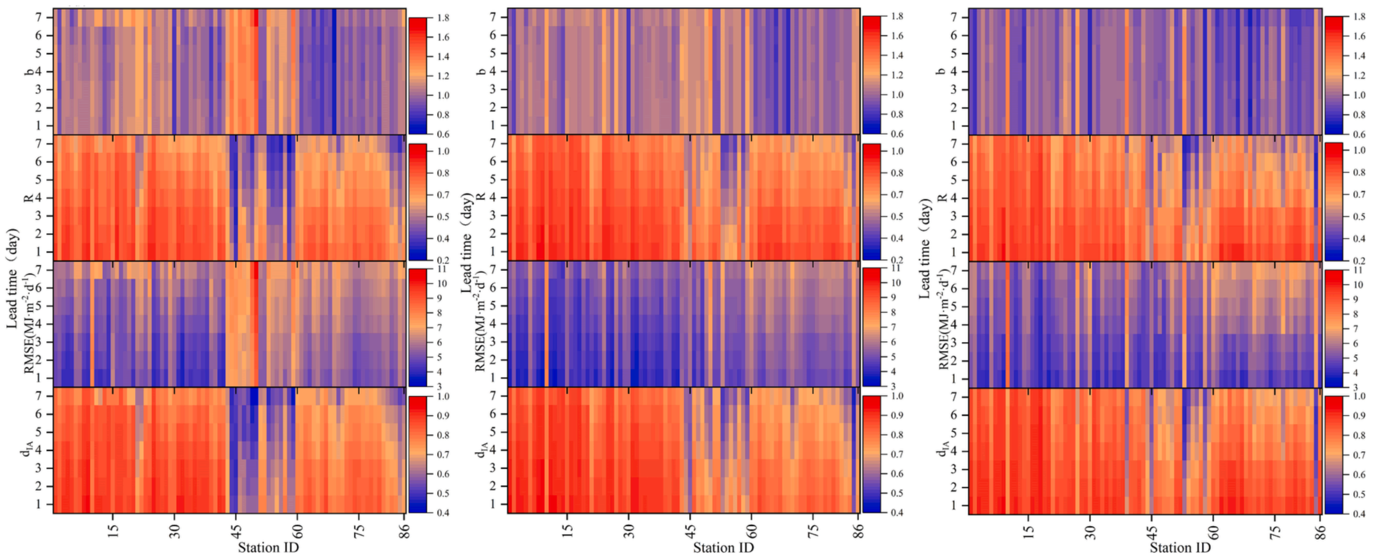

Fig. 6. Forecasting performance of solar radiation by using the generalized sunshine–based model coupled with common (left) and improved sunshine duration converted methods (middle), and generalized temperature–based model (right), respectively, for 1–7 days lead time during the period 2015–2019 for 86 sites in China. b is the slope of the regression function; R is the correlation coefficient; RMSE is the root mean squared error; and d_{IA} is the index of agreement.

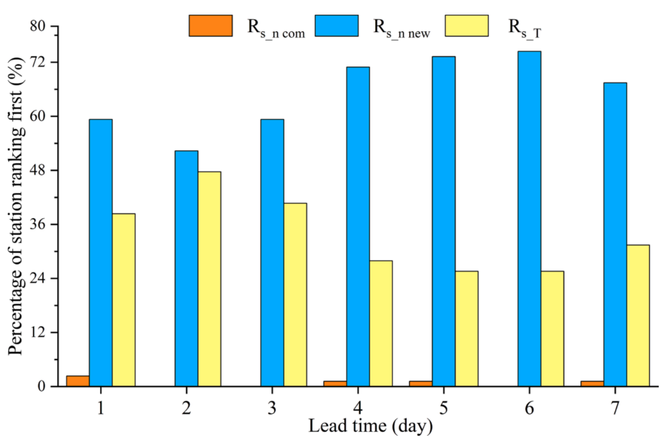

Fig. 7. The percentage of GPI values ranking first for the 86 sites in China per model for 1–7 days lead time during the period 2015–2019. $R_{s\_n\ com}$ and $R_{s\_n\ new}$ are the generalized sunshine–based model coupled with common and improved sunshine duration converting methods. R_{s_T} is the generalized temperature–based model.

converting method.

The results indicated that facilitating sunshine–based model with the proposed n_{new} method that involves T_d can improve R_s forecast due to the close relationship between n and T_d (Fig. 1), which has also been demonstrated in previous studies [11,16,51], since T_d can greatly affect the estimation of daily n and R_s [6]. Overall, the improved model can be implemented in the majority of the locations in China, which can provide reliable information on solar energy–related industries. To make accurate R_s information easily accessible to the public, this study can be potentially applied to provide valuable information to display forecast daily R_s in public weather forecasting system on both websites and smartphone apps.

4. Conclusions

The forecasted daily R_s can offer important information for sectors of energy industry, agriculture production, ecology stability, and climate change. To make reasonable R_s information easily accessible to the public, we used the strength of sunshine–based R_s model and dedicate to solve the weakness of the common n converting method. We developed a novel n converting method by combining weather types and T_d that are easily available from weather forecasts, and then integrated this method into the generalized sunshine–based R_s model. Here, we verified that the proposed novel n converting method was able to improve forecasting accuracy of n compared to the common converting method. Moreover, the generalized R_{s_n} model integrating the n_{new} converting method outperformed the model coupled with the n_{com} converting method and the generalized R_{s_T} model. Hence the $R_{s\_n\ new}$ model is recommended for predicting daily R_s using weather types, T_{max}, and T_{min} data from weather forecast systems.

CRediT authorship contribution statement

Shujing Qin: Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Zhihe Liu:** Methodology, Formal analysis, Visualization, Writing – review & editing. **Rangjian Qiu:** Supervision, Conceptualization, Data curation, Funding acquisition, Methodology, Writing – review & editing. **Yufeng Luo:** Data curation, Writing – review & editing. **Jingwei Wu:** Methodology, Writing – review & editing. **Baozhong Zhang:** Data curation, Writing – review & editing. **Lifeng Wu:** Methodology, Software, Writing – review & editing. **Evgenios Agathokleous:** Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We acknowledge support from the National Natural Science Foundation of China (52179036, 52130906, and 51822907), and the Startup Foundation for Introducing Talent of Wuhan University (2042022rc0013), and the China Meteorological Data Service Centre for providing the meteorological data.

References

[1] Wang F, Harindintwali JD, Yuan Z, Wang M, Wang F, Li S, et al. Technologies and perspectives for achieving carbon neutrality. The Innovation 2021;2(4):100180.

[2] Ye J, Wang C, Gao C, Fu T, Yang C, Ren G, et al. Solar-driven methanogenesis with ultrahigh selectivity by turning down H2 production at biotic-abiotic interface. Nat Commun 2022;13.

[3] Li L, Zhang Y, Zhou T, Wang K, Wang C, Wang T, et al. Mitigation of China's carbon neutrality to global warming. Nat Commun 2022;13.

[4] Cherp A, Vinichenko V, Tosun J, Gordon JA, Jewell J. National growth dynamics of wind and solar power compared to the growth required for global climate targets. Nat Energy 2021;6:742–54.

[5] Park J, Das A, Park J. A new approach to estimate the spatial distribution of solar radiation using topographic factor and sunshine duration in South Korea. Energ Conver Manage 2015;101:30–9.

[6] Liu Y, Zhou Y, Wang D, Wang Y, Li Y, Zhu Y. Classification of solar radiation zones and general models for estimating the daily global solar radiation on horizontal surfaces in China. Energ Conver Manage 2017;154:168–79.

[7] Chen J, He L, Yang H, Ma M, Chen Q, Wu S, et al. Empirical models for estimating monthly global solar radiation: A most comprehensive review and comparative case study in China. Renew Sustain Energy Rev 2019;108:91–111.

[8] Qiu R, Li L, Wu L, Agathokleous E, Liu C, Zhang B, et al. Modeling daily global solar radiation using only temperature data: Past, development, and future. Renew Sustain Energy Rev 2022;163:112511.

[9] Besharat F, Dehghan AA, Faghih AR. Empirical models for estimating global solar radiation: a review and case study. Renew Sustain Energy Rev 2013;21:798–821.

[10] Sun H, Gui D, Yan B, Liu Yi, Liao W, Zhu Y, et al. Assessing the potential of random forest method for estimating solar radiation using air pollution index. Energ Conver Manage 2016;119:121–9.

[11] Fan J, Wang X, Wu L, Zhang F, Bai H, Lu X, et al. New combined models for estimating daily global solar radiation based on sunshine duration in humid regions: a case study in South China. Energ Conver Manage 2018;156:618–25.

[12] Qiu R, Liu C, Cui N, Gao Y, Li L, Wu Z, et al. Generalized extreme gradient boosting model for predicting daily global solar radiation for locations without historical data. Energ Conver Manage 2022;258:115488.

[13] Urraca R, Martinez-de-Pison E, Sanz-Garcia A, Antonanzas J, Antonanzas-Torres F. Estimation methods for global solar radiation: case study evaluation of five different approaches in central Spain. Renew Sustain Energy Rev 2017;77:1098–113.

[14] Fan J, Chen B, Wu L, Zhang F, Lu X, Xiang Y. Evaluation and development of temperature-based empirical models for estimating daily global solar radiation in humid regions. Energy 2018;144:903–14.

[15] Hussain S, AlAlili A. A hybrid solar radiation modeling approach using wavelet multiresolution analysis and artificial neural networks. Appl Energy 2017;208:540–50.

[16] Fan J, Wu L, Zhang F, Cai H, Ma X, Bai H. Evaluation and development of empirical models for estimating daily and monthly mean daily diffuse horizontal solar radiation for different climatic regions of China. Renew Sustain Energy Rev 2019;105:168–86.

[17] Feng Yu, Gong D, Zhang Q, Jiang S, Zhao Lu, Cui N. Evaluation of temperature-based machine learning and empirical models for predicting daily global solar radiation. Energ Conver Manage 2019;198:111780.

[18] Meenal R, Selvakumar AI. Assessment of SVM, empirical and ANN based solar radiation prediction models with most influencing input parameters. Renew Energy 2018;121:324–43.

[19] Wu L, Huang G, Fan J, Zhang F, Wang X, Zeng W. Potential of kernel-based nonlinear extension of Arps decline model and gradient boosting with categorical features support for predicting daily global solar radiation in humid regions. Energ Conver Manage 2019;183:280–95.

[20] Bakirci K. Prediction of global solar radiation and comparison with satellite data. J Atmos Sol Terr Phys 2017;152-153:41–9.

[21] Allen RG, Pereira LS, Raes D, Smith M. Crop evapotranspiration: Guidelines for computing crop requirements-FAO Irrigation and Drainage Paper 56. Rome, Italy: Food and Agriculture Organization of the United Nations; 1998.

[22] Almorox J, Hontoria C. Global solar radiation estimation using sunshine duration in Spain. Energ Conver Manage 2004;45(9-10):1529–35.

[23] Duzen H, Aydin H. Sunshine-based estimation of global solar radiation on horizontal surface at Lake Van region (Turkey). Energ Conver Manage 2012;58:35–46.

[24] Chelbi M, Gagnon Y, Waewsak J. Solar radiation mapping using sunshine duration-based models and interpolation techniques: application to Tunisia. Energ Conver Manage 2015;101:203–15.

[25] Almorox J, Bocco M, Willington E. Estimation of daily global solar radiation from measured temperatures at Cañada de Luque, Córdoba. Argentina Renew Energy 2013;60:382–7.

[26] Hassan GE, Youssef ME, Mohamed ZE, Ali MA, Hanafy AA. New temperature-based models for predicting global solar radiation. Appl Energy 2016;179:437–50.

[27] Gu L, Fuentes JD, Garstang M, Silva JTd, Heitz R, Sigler J, et al. Cloud modulation of surface solar irradiance at a pasture site in southern Brazil. Agric For Meteorol 2001;106(2):117–29.

[28] Kostić R, Mikulović J. The empirical models for estimating solar insolation in Serbia by using meteorological data on cloudiness. Renew Energy 2017;114:1281–93.

[29] Prieto J-I, García D. Global solar radiation models: A critical review from the point of view of homogeneity and case study. Renew Sustain Energy Rev 2022;155:111856.

[30] Angstrom A. Report to the international commission for solar research on actinometric investigations of solar and atmospheric radiation. Q J R Meteorolog Soc 1924;50:121–6.

[31] Prescott J. Evaporation from a water surface in relation to solar radiation. Trans Roy Soc S Aust 1940;46:114–8.

[32] Despotovic M, Nedic V, Despotovic D, Cvetanovic S. Review and statistical analysis of different global solar radiation sunshine models. Renew Sustain Energy Rev 2015;52:1869–80.

[33] Newland FJ. A study of solar radiation models for the coastal region of South China. Sol Energy 1989;43(4):227–35.

[34] Liu ZH, Qiu RJ, Liu CW. Determination of generalized solar radiation model based on sunshine percentage for improving reference crop evapotranspiration estimation accuracy. Water Saving Irrigation 2023;95:80–8.

[35] Hargreaves G, Samani Z. Estimating potential evapotranspiration. J Irrig Drain Eng 1982;108:225–30.

[36] Bristow KL, Campbell GS. On the relationship between incoming solar radiation and daily maximum and minimum temperature. Agric For Meteorol 1984;31(2):159–66.

[37] Zhang J, Zhao L, Deng S, Xu W, Zhang Y. A critical review of the models used to estimate solar radiation. Renew Sustain Energy Rev 2017;70:314–29.

[38] Feng Yu, Gong D, Jiang S, Zhao Lu, Cui N. National-scale development and calibration of empirical models for predicting daily global solar radiation in China. Energ Conver Manage 2020;203:112236.

[39] Luo YF, Chang XM, Peng SZ, Khan S, Wang WG, Zheng Q, et al. Short-term forecasting of daily reference evapotranspiration using the Hargreaves-Samani model and temperature forecasts. Agric Water Manage 2014;136:42–51.

[40] Yang Y, Cui Y, Bai K, Luo T, Dai J, Wang W, et al. Short-term forecasting of daily reference evapotranspiration using the reduced-set Penman-Monteith model and public weather forecasts. Agric Water Manage 2019;211:70–80.

[41] Yang Y, Luo Y, Wu C, Zheng H, Zhang L, Cui Y, et al. Evaluation of six equations for daily reference evapotranspiration estimating using public weather forecast message for different climate regions across China. Agric Water Manage 2019;222:386–99.

[42] Cai J, Liu Yu, Lei T, Pereira LS. Estimating reference evapotranspiration with the FAO Penman-Monteith equation using daily weather forecast messages. Agric For Meteorol 2007;145(1-2):22–35.

[43] Yang Y, Cui Y, Luo Y, Lyu X, Traore S, Khan S, et al. Short-term forecasting of daily reference evapotranspiration using the Penman-Monteith model and public weather forecasts. Agric Water Manage 2016;177:329–39.

[44] Gong X, Qiu R, Ge J, Bo G, Ping Y, Xin Q, et al. Evapotranspiration partitioning of greenhouse grown tomato using a modified Priestley-Taylor model. Agric Water Manage 2021;247:106709.

[45] Qiu R, Katul GG, Wang J, Xu J, Kang S, Liu C, et al. Differential response of rice evapotranspiration to varying patterns of warming. Agric For Meteorol 2021;298-299:108293.

[46] Qiu R, Li L, Liu C, Wang Z, Zhang B, Liu Z. Evapotranspiration estimation using a modified crop coefficient model in a rotated rice-winter wheat system. Agric Water Manage 2022;264:107501.

[47] Trnka M, Žalud Z, Eitzinger J, Dubrovský M. Global solar radiation in Central European lowlands estimated by various empirical formulae. Agric For Meteorol 2005;131(1-2):54–76.

[48] Luo Y, Traore S, Lyu X, Wang W, Wang Y, Xie Y, et al. Medium range daily reference evapotranspiration forecasting by using ANN and public weather forecasts. Water Resour Manage 2015;29(10):3863–76.

[49] Yan X, Mohammadian A. Forecasting daily reference evapotranspiration for Canada using the Penman-Monteith model and statistically downscaled global climate model projections. Alex Eng J 2020;59(2):883–91.

[50] Allen RG. Self-calibrating method for estimating solar radiation from air temperature. J Hydrol Eng 1997;2(2):56–67.

[51] Yıldırım HB, Teke A, Antonanzas-Torres F. Evaluation of classical parametric models for estimating solar radiation in the Eastern Mediterranean region of Turkey. Renew Sustain Energy Rev 2018;82:2053–65.