Hourly global solar radiation prediction based on seasonal and stochastic feature

You Li [a,b], Yafei Wang [c,*], Hui Qian [a], Weijun Gao [c], Hiroatsu Fukuda [c], Weisheng Zhou [d]

[a] *Dual-carbon Research Center, Hangzhou City University, Hangzhou, 310015, China*
[b] *Asia-Japan Research Institute, Ritsumeikan University, Ibaraki 567-8570, Japan*
[c] *Faculty of Environmental Engineering, The University of Kitakyushu, Kitakyushu, 808-0135, Japan*
[d] *College of Policy Science, Ritsumeikan University, Ibaraki, 567-8570, Japan*

ABSTRACT

Accurate and detailed solar radiation data play a crucial role in the simulation of building thermal and photovoltaic systems. However, developing a highly precise and dependable solar radiation model using cost-effective data has proven challenging. This work proposes a new attenuation solar radiation model formed by conducting a comprehensive analysis of existing models and gaining new insights into solar radiation's seasonal and stochastic properties. Meanwhile, the model is constructed using easily obtainable surface meteorological parameters. The results demonstrate that the proposed model exhibits good performance in terms of prediction accuracy. Moreover, the majority of existing hourly solar radiation models have been primarily developed for clear-sky conditions. However, there is a growing demand for solar radiation hourly estimations that can uphold a high level of accuracy and reliability even in different weather state. Conversely, the proposed model is developed and validated by more than twenty year's meteorological data encompassing various weather conditions in Japan. It effectively captures the stochastic nature of solar radiation by utilizing turbidity parameters, even on cloudy and rainy days. Additionally, the inclusion of interaction variables significantly enhances its interpretability.

1. Introduction

1.1. Background

With the population grows and urbanization accelerates, energy has emerged as a crucial factor influencing both the global economy and the environment [1]. As a safe, cost-effective, environmentally friendly, and renewable energy source, solar energy is a viable alternative to enhance the overall energy structure [2,3]. Solar energy has extensive applications across diverse fields such as weather forecasting [4], photovoltaic power utilization [5], zero-energy building design [6,7], climate observation [8], and agricultural usage [9]. Improving the capabilities in hourly solar radiation modeling can directly and significantly impact the quality of applied studies. Moreover, the intermittent, stochastic characteristics of solar energy pose significant challenges when it comes to

---

* Corresponding author. Faculty of Environmental Engineering, The University of Kitakyushu, 808-0135, Japan.
E-mail address: yafeiwangjapan@outlook.com (Y. Wang).

integrating PV power into the grid (e.g., severe blackout [10] or resilience lacking when facing failure and disruption [11,12]). Precise and dependable solar irradiation estimations could enhance the coordination among the overall systems and the public-grid. Thereby enabling smoother penetration of solar power into building's power system [13]. Regrettably, due to financial considerations, it is a significant scarcity of weather observations which monitor hourly solar irradiation compared to these weather sites that monitor general meteorological items such as air temperature. Consequently, in order to conquer those challenges, it becomes imperative to establish models that can estimate solar irradiation from general meteorological items.

1.2. Literature review

Until now, scholars and researchers have devoted significant efforts to solar radiation forecasting modeling and have attained promising outcomes through their research endeavors. Based on the variation in input parameters, hourly solar radiation models can be broadly categorized into two main categories: model according to typical meteorological parameters. It includes empirical formulas and machine-learning model, then model relying on satellite elements.

One approach is to analyze the connections among general meteorological elements and GSR through regression analysis to establish empirical formulas. Because of its simple form of expression, it has been widely utilized throughout its extensive growing. Common parameters in those empirical models include air temperature or difference of air temperature, sunshine duration, and a combination of other weather elements. Angstrom [14] firstly introduced the GSR estimation model in 1924, which is achieved by establishing a linear relationship between global radiation and daylight radiation for a specific location, along with the ratio of the average solar radiation to the maximum radiation. Afterwards, Prescott modified Angstrom's model using extraterrestrial theoretical solar radiation by take the place of the clear sky solar radiation value in 1940. Because it is easier to calculate global radiation; then, the revised model can be known as the Angstrom-Prescott model [15]. After that, many researchers have revised and enhanced the model through their efforts according to the Angstrom-Prescott model, (e.g., Morf H et al., 2013 [16], Paulescu M et al., 2016 [17], Asilevi et al., 2019 [18]). Correspondingly, many researchers have formulated linear, power, exponential, and polynomial equations to establish relationships between the minimum and maximum, or average air temperatures, as well as the differences of air temperature and the clear sky index (e.g., Hargreaves GH et al. [19,20], Liu $M-F$ et al., 2013 [21], Almorox J et al., 2013 [22]). The power function model developed by Hargreaves, which incorporates both the daily temperature variations and extraterrestrial solar radiation, is considered as one of the classic models [19]. The aforementioned models are highly appropriate and accurate for computing GSR daily value. And the GSR daily value holds significant value for concrete application, for example, evaluating the average or max solar generation. However, there is a growing demand for hourly solar radiation data for more accurate application, e.g., buildings' heat utilization.

There have been a lot of literatures on estimating the HGSR by utilizing the monthly mean daily percentage of possible sunshine (e. g., Rietveld, 1978; Yoshida and Shinoki, 1978). Furthermore, there have been proposals to estimate of hourly direct and diffuse solar radiation based on daily global solar radiation (GSR) (e.g., Iqbal, 1979 [23]; Bugler, 1977 [24]; Watanabe et al., 1983 [25]; Flint and Childs, 1987 [26]). Nevertheless, initial studies on hourly solar radiation primarily focused on calculating average values over a one-month period, which limited their ability to capture real-time fluctuations in solar radiation. As a result, these early models were not well-suited for accurately representing the dynamic changes in solar radiation that occur over shorter time intervals. Zhang and Huang developed a real hourly solar model for Beijing and Guangzhou, China [27]. Under all-day conditions, this model incorporates measured data of clouds to figure out the influence of cloud cover on GSR. This model aimed to accurately capture the variations in solar radiation on an hourly basis, considering the unique geographical and climatic characteristics of these two regions. Their work laid the foundation for more accurate and region-specific solar radiation models, many previous studies have verified the accuracy of the model (e.g., Chang K et al., 2019 [28], Kim KH et al. [29,30]).

Recently, it is a growing interest in utilizing artificial intelligence methods, including various random forest model and Artificial Neural Networks (ANN) for GSR estimation. Geetha A et al. [31]investigated different ANN models with three popular algorithms promoted in the literature (Levenberg Marquardt Backpropagation, Resilient Backpropagation and Scaled Conjugate Gradient, respectively). The results showed that the ANN model built using the LM algorithm has the advantage of converging well in a shorter time, providing a suitable solution with minimal error. Miranda E et al. [32] introduced a method based on simple angle calculations and regression models to predict half-hourly diffuse horizontal solar irradiance utilizing state-of-the-art machine learning models from only global horizontal irradiance measurements and geographic coordinates as input items. Their results showed a coefficient of determination ranging from 0.9974 to 0.9983. Nawab F et al. [33] conduct a thorough review of research articles discussing solar irradiance prediction to compare different solar irradiance prediction methods. The review showed that AI methods are more accurate than empirical methods. Among them, ANN and hybrid model have the highest accuracy among artificial intelligence methods, followed by support vector machine and adaptive neuro-fuzzy inference system. Furthermore, many researchers are enthusiastic about integrating multiple deep learning models. Gao, Y. et al. [34] devised a deep LSTM-based generative model for multi-step solar radiation forecasts capable of forecasting one-day in advance. Lai CS et al. [35] presented a hybrid approach for 1-h-ahead global horizontal irradiance prediction based on deep learning. The study showed that the proposed method achieves the smallest solar forecast error compared to Smart Persistence and other state-of-the-art methods. Gao, Y [36]. explored an explain ability in the prediction process's time and space dependence using attentional mechanisms. Their research findings indicated that irradiation is significantly correlated with factors such as air temperature, irradiation time, water vapor pressure and precipitation. But the present outcomes still fall short of achieving full satisfaction.

The GSR prediction models discussed above exhibit high accuracy and broad applicability, primarily as statistical or data-driven models. However, ensuring the forecast's reliability under changing conditions (e.g., geographic climate and weather variations) can

be challenging. Therefore, achieving high-precision solar radiation prediction necessitates a solid theoretical foundation. The first attenuation GSR model was introduced by Roland L. Hulstrom and Richard E. Bird in 1981 [37]. It segregates the attenuation layer of atmosphere into 6 distinct components: ozone layer, mixed gas layer, Rayleigh layer, aerosol layer, water vapor layer and clouds layer [38]. Bird model provides a clear representation in theory, aligning with the physical sensation. Nonetheless, in the calculations of the Bird model, the attenuation effects are expressed through transmittance, necessitating consideration of the transmittance of each layer. This practical implementation can become complicated in practice. Su, G. et al. [39] make modifications to the Richard E. Bird's model by incorporating Air Quality Index (AQI) and relative humidity. As a result, the correlation coefficient of this model is enhanced by 41.3% and 5.7% compared to the existing methods, respectively. However, one drawback is that general attenuation model relies on satellite retrieval data and require inversion using professional software.

Moreover, the majority of those existing models have been established under a clear sky state. However, the demand for high accuracy and reliability in hourly solar radiation models has grown significantly, particularly under complex weather states due to the extreme weather events occur frequently.

According to the preceding research, the work's contribution can be summarized as follows.

- The proposed model founded on a comprehensive analysis of existing models, as well as novel insights into the solar irradiation attenuation effects.
- Low-cost surface meteorological items meticulously chosen as feature variables to the atmospheric transmittance to enhance the reliability and practicality of solar radiation attenuation models.
- The proposed model is meticulously calibrated under diverse weather conditions to account for weather stochasticity inherent in the model.

The remaining sections of this work are structured as: Section 2 presents the methods and materials. Section 3 discusses the results and provides detailed analyses. Lastly, Section 4 presents the conclusions drawn from this work.

2. Materials and methodology

2.1. Formal analysis

Many GSR models have been developed in recent decades, and they generally provide good accuracy regardless of the methods and variables used. And there are many studies reviewing these GSR models [40,41]. They point out that the development of GSR prediction is mainly attributed to several pioneering models. These pioneer models are generally classified as insolation-based, temperature-based, cloud-based, and other climate-parameter-based models, depending on the input variables. Other GSR models are derived from the pioneer model by adding other variables that have been shown to achieve better estimates under validated conditions. Table 1 shows some of the main models, which, although not comprehensive, can show the main formal characteristics of the models.

It can be found that all GSR models endeavor to calculate the proportion among the irradiation reaching the horizontal plane and an irradiation threshold, typically representing extraterrestrial GSR. Initially, it possesses a non-dimensional nature with a straightforward interpretation, representing the transmittance of solar irradiation in the air under clear and overcast days. Consequently, the

Table 1
Monthly or Daily solar radiation methods.

No.	Author	Formulation	Explanatory item
1	Angstrom (1924) [14]-Prescott (1940) [15]	$\frac{I}{I_0} = c_0 + c_1 \frac{S_d}{S_0}$	Sunshine duration (S_d)
2	Ogelman et al. (1984)	$\frac{I}{I_0} = c_0 + c_1 \frac{S_d}{S_0} + c_2 \left(\frac{S_d}{S_0}\right)^2$	Sunshine duration (S_d)
3	Ampratwum and Dorvlo (1999)	$\frac{I}{I_0} = c_0 + c_1 \ln\left(\frac{S_d}{S_0}\right)$	Sunshine duration (S_d)
2	Teke and Yildirim (2014)	$\frac{I}{I_0} = c_0 \sin\left(\frac{S_d}{S_0} + c_1\right)$	Sunshine duration (S_d)
3	Hargreaves and Samani (1982) [19]	$\frac{I}{I_0} = c_1 \Delta T^{0.5}$	Air temperature (T)
4	Hargreaves and Samani (1985)	$\frac{I}{I_0} = c_0 + c_1 \Delta T^{0.5}$	Air temperature (T)
5	Falayi (2008) [42]	$\frac{I}{I_0} = c_0 + c_1 T_{\min}$	Air temperature (T)
6	Li et al. (2010) [43]	$\frac{I}{I_0} = c_0 + c_1 T_{\min} + c_2 T_{\max}$	Air temperature (T)
6	Black (1956) [44]	$\frac{I}{I_0} = c_0 + c_1 C + c_2 C^2$	Daily average cloud cover (C)
7	Glover and McCulloch (1958) [45]	$\frac{I}{I_0} = c_1 \cos\varphi + c_2 \frac{S_d}{S_0}$	Latitude, Sunshine duration (S_d)
8	Swartman and Ogunlade (1967) [46]	$\frac{I}{I_0} = c_0 + c_1 RH + c_2 \frac{S_d}{S_0}$	Relative humidity (RH), Sunshine duration (S_d)

aforementioned ratio may be regarded as the clearness index. Like detailed in Section one, numerous works have sought to develop an equational correlation among the clearness index and typical weather items. This category of those equations can be represented as follows:

$$\frac{I}{I_0} = f(\sigma) \tag{1}$$

In this equation, I represents the GSR. I_0 denotes the extraterrestrial GSR. The function expression $f(\sigma)$ encompasses various meteorological parameters. These models demonstrate higher accuracy when solving monthly and daily-scale predictions due to the meteorological parameters utilized, which offer insights into the seasonal changes of GSR through cumulative values. However, the applicability of most of those models diminish as the resolution becomes hourly or instantaneous values. This limitation arises because the equational features in the fractions will cause denominator close to zero, resulting in significant calculation error in the forecasts outcomes. Typically, the daily cumulative values of extraterrestrial irradiation of Japanese stations ranges from 10 MJ/m^2 to 30 MJ/m^2. However, during sunset or sunrise, the irradiation approaches 0, and those factors could not be disregarded.

Table 2 shows the main hourly solar radiation models. Watanabe model (1983) is an early developed average-hour model whose main function is to split the GSR into direct and diffuse radiation. In the equation, the GSR is obtained by adding the first half (direct radiation) and the second half (diffuse radiation). The model is a statistical model driven by the atmospheric transmittance (P). Ueyama (2005) model is obtained by modifying the Watanabe model, and its main contribution is to relate the atmospheric transmittance to the surface parameters (S and r), making Make the model more convenient. Meanwhile, Ueyama (2018) simplifies the expression for atmospheric transmittance by removing precipitation. It can be found that early models of hourly solar radiation are usually complex, considering the actual relationship between atmospheric transmittance and direct and diffuse radiation, respectively. However, observing these equations, we notice that the expressions for direct and diffuse radiation are partly the same ($I_{sc} \sin \theta$). Zhang and Huang model (2002) summarizes and derives this partial contribution. It builds a new model with extraterrestrial radiation ($I_{sc} \sin \theta$) as the interaction term and selects surface meteorological parameters instead of the original atmospheric transmittance. Theoretically, it is a realistic hourly solar radiation model because the driving parameters can be obtained from the hour. From there, some scholars have optimized the Zhang and Huang model by changing the variables (e.g., Chang Kai (2020)).

2.2. Formal innovations

In endeavors to achieve high-precision solar radiation prediction, stochastic variations of the sky' state can significantly augment the randomness of the solar radiation distribution. As a result, the form of the model is also evolving from a statistically driven model to a surface meteorological parameter driven model. Meanwhile, individual seasonal meteorological parameters such as temperature exhibit poor correlation with instantaneous GSR. Hence, opting for a connection between multiple weather items, particularly representation variables representing the sky' state, proves to be a more favorable selection.

Furthermore, the complexity of the HGSR prediction primarily lies in two aspects, namely, seasonality and weather stochasticity. The weather stochasticity diminishes the seasonal distribution patterns of GSR, rendering some classical daily value models less in regions with complex weather conditions [50]. Extraterrestrial solar radiation contains all the seasonal features of GSR, while atmospheric transmittance (or its characterizing variable) represents the degree of sky clarity and thus can be considered to represent the stochastic characteristics of weather. Formally expressed as they mainly build models with extraterrestrial radiation ($I_{sc} \sin \theta$) as the interaction term and with the product of extraterrestrial solar radiation and atmospheric transmittance. Although they have proven to have very good performance in many cases, they also have some drawbacks. For example, the accuracy of the Zhang and Huang models is low in cloudy and rainy days, and it is not clear how we can interpret and modify them. This is because the model contains complex physical implications despite its formal simplicity. To address the above knowledge gaps, we propose the following methodological process (Fig. 1), and the conceptual model can be like Eq. (2):

Table 2
Hourly solar radiation methods.

No.	Author	Formulation	Explanatory item
9	Watanabe et al. (1983) [25]	$I = I_{sc} \sin \theta \cdot P^{\csc \theta} + I_{sc} \sin \theta \cdot Q / (1 + Q)$ $Q = (0.9013 + 1.123 \sin \theta) P^{0.489 \csc \theta} (1 - P^{\csc \theta})^{2.525}$	Atmospheric transmittance(P) (Statistical Model)
10	Ueyama (2005) [47]	$I = I_{sc} \sin \theta \cdot P^{\csc \theta} + I_{sc} \sin \theta \cdot (1 - Kd / Kds)$ $P = c_1 S + c_2 r + c_3 I_0 + c_0$ $Kds = Kd + c_1 Kd^{c_2} \cdot (1 - Kd)^{c_3}$ $Kd = P^{\csc \theta}$	Sunshine duration(S), Precipitation(r)
11	Ueyama (2018) [48]	$I = I_{sc} \sin \theta \cdot P^{\csc \theta} + I_{sc} \sin \theta \cdot (1 - Kd / Kds)$ $P = c_1 S + c_2 I_0 + c_0$ $Kds = Kd + c_1 Kd^{c_2} \cdot (1 - Kd)^{c_3}$ $Kd = P^{\csc \theta}$	Sunshine duration(S)
12	Zhang and Huang (2002) [49]	$I = [I_{sc} \cdot \sin \theta \cdot \{c_0 + c_1 CC + c_2 CC^2 + c_3 (T_n - T_{n-3}) + c_4 RH\} + d] / k$	Cloud cover (CC), Air temperature (T_n-T_{n-3}), Relative humidity (RH)
13	Chang Kai et al. (2020) [27]	$I = I_{sc} \cdot \sin \theta \cdot \{c_0 + c_1 S + c_2 (T_n - T_{n-3}) + c_3 RH\} + c_4$	Sunshine duration(S), Air temperature (T_n-T_{n-3}), Relative humidity (RH)

$$I = I_0 - \Delta I \tag{2}$$

where ΔI represents loss value of GSR in the air. As mentioned earlier, the interaction of seasonality and weather stochasticity is the main reason for the difficulty in predicting solar radiation. Therefore, we make a formal innovation based on the existing models and construct a new weakening model by subtraction. As shown in Fig. 1 (a), the main seasonal variation of solar radiation is reflected by extraterrestrial solar radiation (Periodic variations in solar radiation are determined by the path and angle of radiation through the atmosphere. Fig. 1 (b)). We can consider that the extraterrestrial solar radiation varies periodically on a yearly basis. Also, the decrease in solar radiation as it traverses through the air (Fig. 1 (c)) is a possible variable that responds to the state of the sky. Thus, the proposed conceptualized model can predict solar radiation while by describing the potential of different features of solar radiation.

2.3. The attenuation effect of solar radiation

The attenuation of irradiation could be segmented into the layers in the atmosphere, including cloud layer, water vapor layer, ozone layer, Rayleigh layer, aerosol layer and mixed gas layer, (which plotted in Fig. 1 (c) and (d)). The surface meteorological parameters are selected as characterization variables to the atmospheric transmittance, thus the reduction of solar radiation in the atmosphere can be described as:

$$\Delta I = c_4 \cdot (c_1 \cdot L_a + c_2 \cdot L_w + c_3 \cdot L_c + c_0) \tag{3}$$

In the equation, the loss values of solar radiation in the aerosol layer (L_a, measured in W/m^2), water vapor layer (L_w measured in (W/m^2), and clouds layer (L_c, measured in W/m^2) are presented. The coefficients C_0 to C_4 (dimensionless) are also presented. Because the uniformity in composition of the Rayleigh layer ozone layer, and mixed gas layer, it easily becomes to identify a constant variable (C_0) in those layers.

2.4. Solar geometry

2.4.1. The extraterrestrial irradiation

The GSR is dependent on the sun's incident altitude and the incidence direction, which received on the horizontal plane. In the previous works, the solar altitude angle (θ) is consistently regarded as a crucial variable for describing the Sun's position. It can be derived from the geographical latitude (φ), the solar hour angle (ω), and the solar declination angle (δ) through Eq. (4):
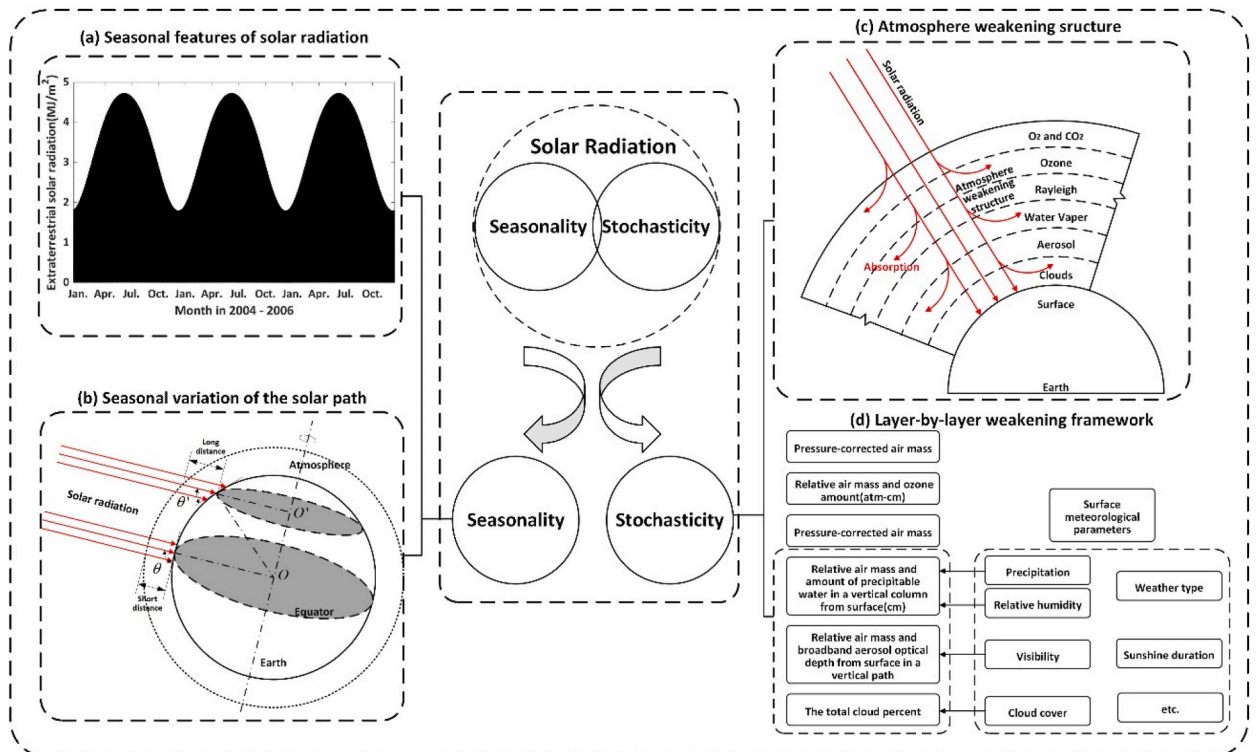

Fig. 1. The schematic diagram of seasonal and random characteristics of solar radiation.

$$\sin \theta = \sin \varphi \sin \delta + \cos \varphi \cos \delta \cos \omega \tag{4}$$

The extraterrestrial irradiation is envisioned as the max value of GSR on the horizontal plane. To calculate I_0, we employ Eq. (5):

$$I_0 = I_{sc} \cdot E_0 \cdot \sin \theta \tag{5}$$

here, I_{sc} represents the solar constant, equivalent to 1361 W/m^2. E_0 stands for the eccentricity factor. Fig. 1 (a) portrays an accurate depiction of the seasonal and temporal variations of extraterrestrial irradiation. As a result, the I_0 can serve as the initial parameter for the proposed attenuation model.

2.4.2. Solar radiation pathway

Solar radiation must travel across the air when reaching the ground. So, the positions between the Earth and Sun are in constant flux, resulting in continuous changes in the length of GSR through the atmosphere. Thus, for a comprehensive analysis of the internal connection among the attenuation effects of GSR and the representation variables, it becomes imperative to take into account the temporal and seasonal tendency in the pathway across the air. And it can be considered as the length of the pathway. The solar altitude angle has piqued our interest due to its straightforward characterization of the positions between the Earth and Sun. As depicted in Fig. 1 (c), when θ is small, those irradiation travel a greater distance across the air, leading to a more substantial attenuation of the irradiation is more significant. Consequently, Eq. (3) is enhanced like Eq. (6):

$$\Delta I = \sin \theta \cdot (c_1 \cdot L_a + c_2 \cdot L_w + c_3 \cdot L_c + c_0) \tag{6}$$

2.5. Data sources

This paper utilizes a publicly available dataset containing meteorological items. The data were collected from 1st Jan of 2000 to 31st Dec. Of 2021 from 7 different stations in Japan. The selection of these 7 stations ensures adequate coverage of Japan's geographical features thereby preventing longitude and latitude variations from impacting modelling. The geographical sites and information of the observation sites are depicted in Fig. 2. The dataset used in this study is sourced from the Japan Meteorological Agency and is comprehensively described in Table 3, which outlines the meteorological items. Thorough inspection and quality control have been conducted on all the data to ensure reliability.

2.6. Method

In this study, an attenuation solar radiation approach is employed, utilizing readily accessible surface meteorological parameters. The process involves establishing the attenuated layers as an equation of GSR through these steps as follows.

- Data pre-processing. The raw data undergoes an initial screening to identify and handle outliers and missing data. Because technological limitations in weather stations, they only recorded 3-h intervals values of visibility, weather type, and cloud cover from January 2000 until December 2019 (In the selected experimental period, 2000 to 2021). Thus, for data consistency, we select
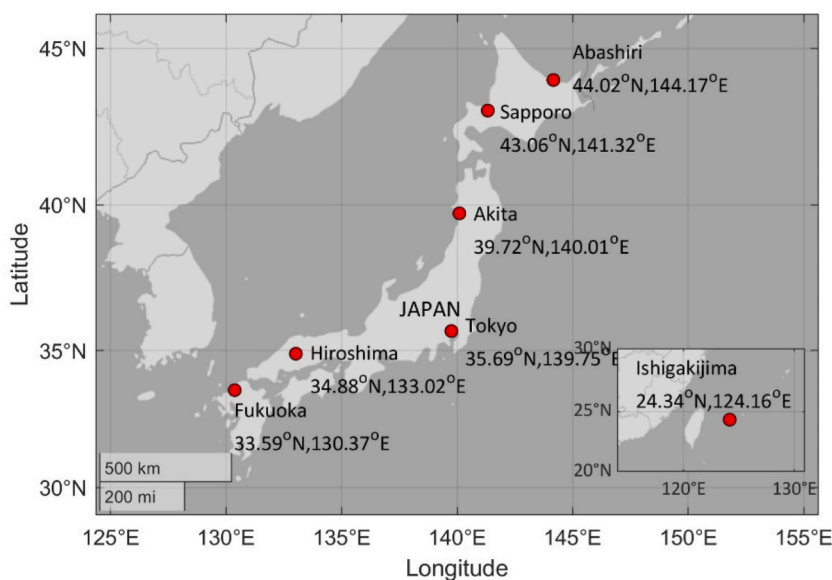

Fig. 2. The geographical information of the observation sites selected in the work.

Table 3
Weather items collected from the stations and their derivation.

No.	Surface Meteorological Parameter	Symbol	Unit	Resolution(h)	Description
1	Solar radiation	I	MJ/m^2	1	–
2	Air temperature	T	°C	1	–
3	Sunshine duration	S	hour	1	–
4	Relative humidity	RH	%	1	–
5	Cloud cover	CC	dimensionless	3	Appendix Table 1
6	Precipitation	r	mm	1	–
7	Weather type	W	dimensionless	3 or 1	Appendix Table 2
8	Wind speed	WS	m/s	1	–
9	Wind Direction	WD	dimensionless	1	–
10	Visibility	V	km	3 or 1	–
11	Dewpoint temperature	Td	°C	1	–
12	Water vapor pressure	WVP	hPa	1	–
13	Atmospheric pressure	Pa	hPa	1	–
14	Enthalpy	H	kJ/kg	1	–
15	Three-hourly temperature difference	$\Delta T3$	°C	1	–
16	Two-hourly temperature difference	$\Delta T2$	°C	1	–
17	Three-hourly enthalpy difference	$\Delta H3$	kJ/kg	1	–
18	Two-hourly enthalpy difference	$\Delta H2$	kJ/kg	1	–
19	Day Number	N	dimensionless	1	–
20	Solar altitude angle	θ	degree	1	–
21	Solar declination angle	δ	Degree	1	–

datasets that include all the mentioned items for subsequent modeling as original resources. After the screening process, the left data sets comprise 205,722 sets of measured data. Among all data sets, 184,382 data sets are adopted to establish the model then, 21,340 data sets are reserved for model accuracy verification. The data details of all weather stations are detailed in Table 4.

- Model establishments. The attenuation factor of GSR in the air is employed as the intermediary to build an equational connection among the assumed representation variables and the actual measured GSR reaching on the horizontal plane. By utilizing all data sets of weather parameters spanning from 1st Jan. Of 2000 to 31st Dec. Of 2018, various general equational form, including, quadratic polynomial, cubic polynomial equations and exponential, linear, are compared as the most suitable equation form.
- Validation. To conclude, the proposed model undergoes validation using solar radiation measurements obtained from seven locations in Japan between January 1, 2019, and December 31, 2021. The validation includes 15 distinct weather conditions, categorized to 3 different sky conditions. Then, the accuracy of the assumed model is thoroughly validated during this validation process.

2.7. Statistical validation

According to the previous studies, various statistical validation indexes are commonly employed to assess the accuracy of the HGSR. This study adopts the following evaluation metrics: Correlation coefficient (R), the relative standard error (RSE), the efficient Nash-Sutcliffe Equation (NSE), the root mean square error (RMSE) and the relative root mean square error (rRMSE).

(1) RSE

The calculation for the relative standard error is as Eq. (7):

$$RSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{e_i - m_i}{m_i}\right)^2} \tag{7}$$

where e_i denotes the calculated data and m_i represents the actual data.

Table 4
Dataset accessibility for the selected weather stations.

Sites	All datasets	Modeling datasets	Validation datasets
Abashiri	28,629	27,125	1504
Sapporo	31,417	27,170	4247
Akita	25,376	23,868	1508
Tokyo	31,199	26,947	4252
Hiroshima	30,697	26,539	4158
Fukuoka	31,048	26,845	4203
Ishigakijima	27,356	25,888	1468
Total	205,722	184,382	21,340

(2) RMSE

The calculation for root mean square error is defined like Eq. (8):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(e_i - m_i)^2} \tag{8}$$

(3) rRMSE

The calculation for relative root mean square error is described like Eq. (9):

$$rRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(e_i - m_i)^2}}{m_a} \times 100\% \tag{9}$$

where m_a is the mean - value of the actual data.

(4) R

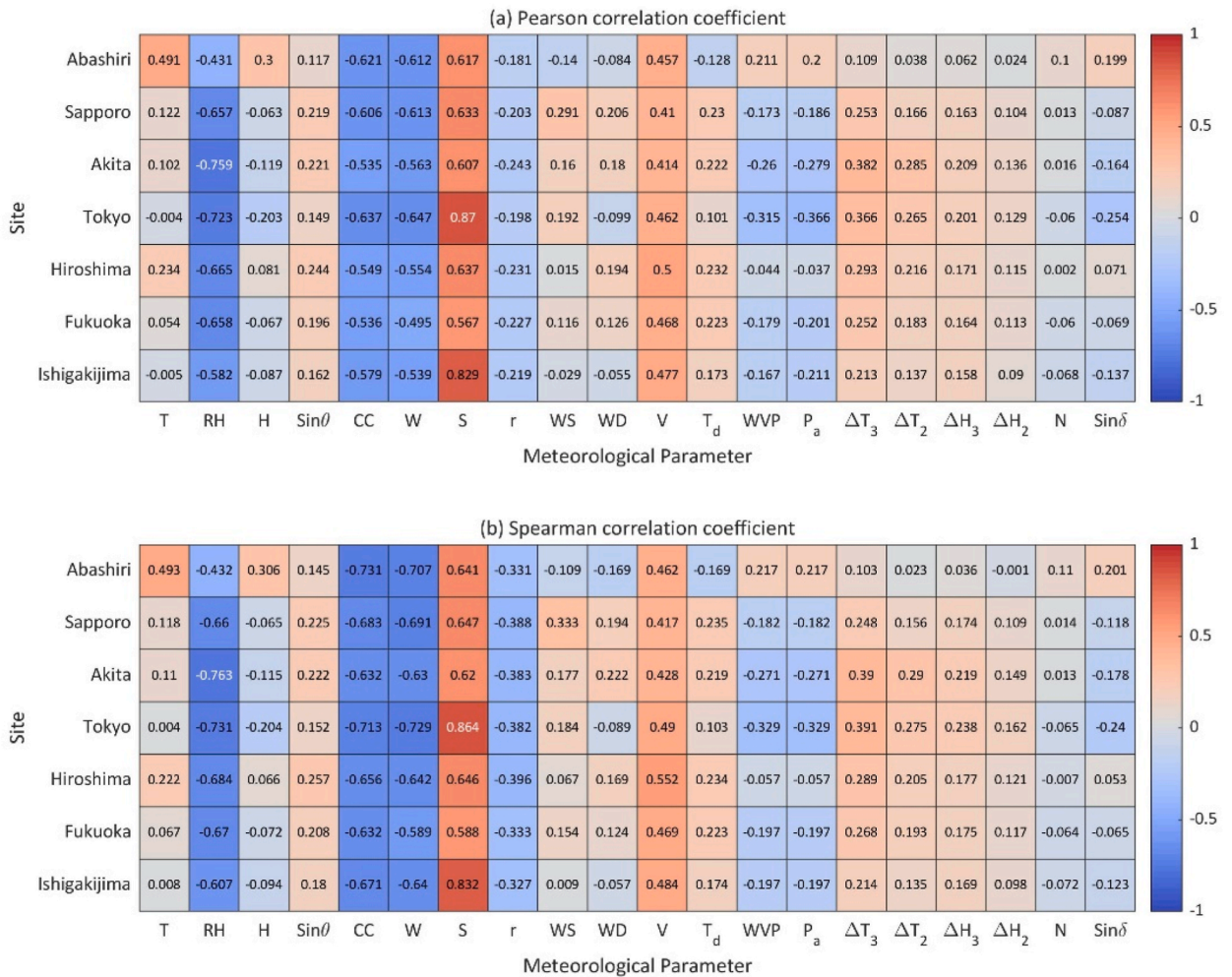The calculation for correlation coefficient is as Eq. (10):

Fig. 3. Correlation coefficient between meteorological parameters with the clearness index. (Note: the absolute value of the coefficient is used to demonstrate the magnitude of the correlation in a concise manner.)

$$R = \frac{\sum ((m_i - m_a)(e_i - e_a))}{\sqrt{\sum (m_i - m_a)^2 \sum (e_i - e_a)^2}} \tag{10}$$

where e_a represents the mean - value of the calculated data and m_a represents the mean - value of the actual data.

(5) NSE

The calculation for Nash-Sutcliffe equation can be described like Eq. (11):

$$NSE = 1 - \frac{\sum (m_i - e_i)^2}{\sum (m_i - m_a)^2} \tag{11}$$

3. Result and discussion

3.1. Correlation coefficient

A Pearson correlation coefficient and a Spearman correlation coefficient (bivariate correlation) are obtained by calculating the correlation between the parameters and the GSR, which are demonstrated in Fig. 3 (a, b). Pearson correlation coefficient shows the linear relations while Spearman correlation coefficient shows the non-linear relations between the parameters. There are 20 meteorological observations and variants can be obtained from weather stations. Fig. 3 illustrates the correlation coefficient values of these elements with (the color represents the strength of the correlation). Typically, correlation coefficients (absolute value) greater than 0.5 are considered to have a relatively strong correlation, while those greater than 0.3 and less than 0.5 are considered to have a weaker correlation. Therefore, the weather type (W), sunshine duration (S), relative humidity (RH), and cloud cover (CC) show a strong linear correlation with GSR. The visibility (V) shows a weak linear correlation. The Spearman coefficient of precipitation (r) is significantly
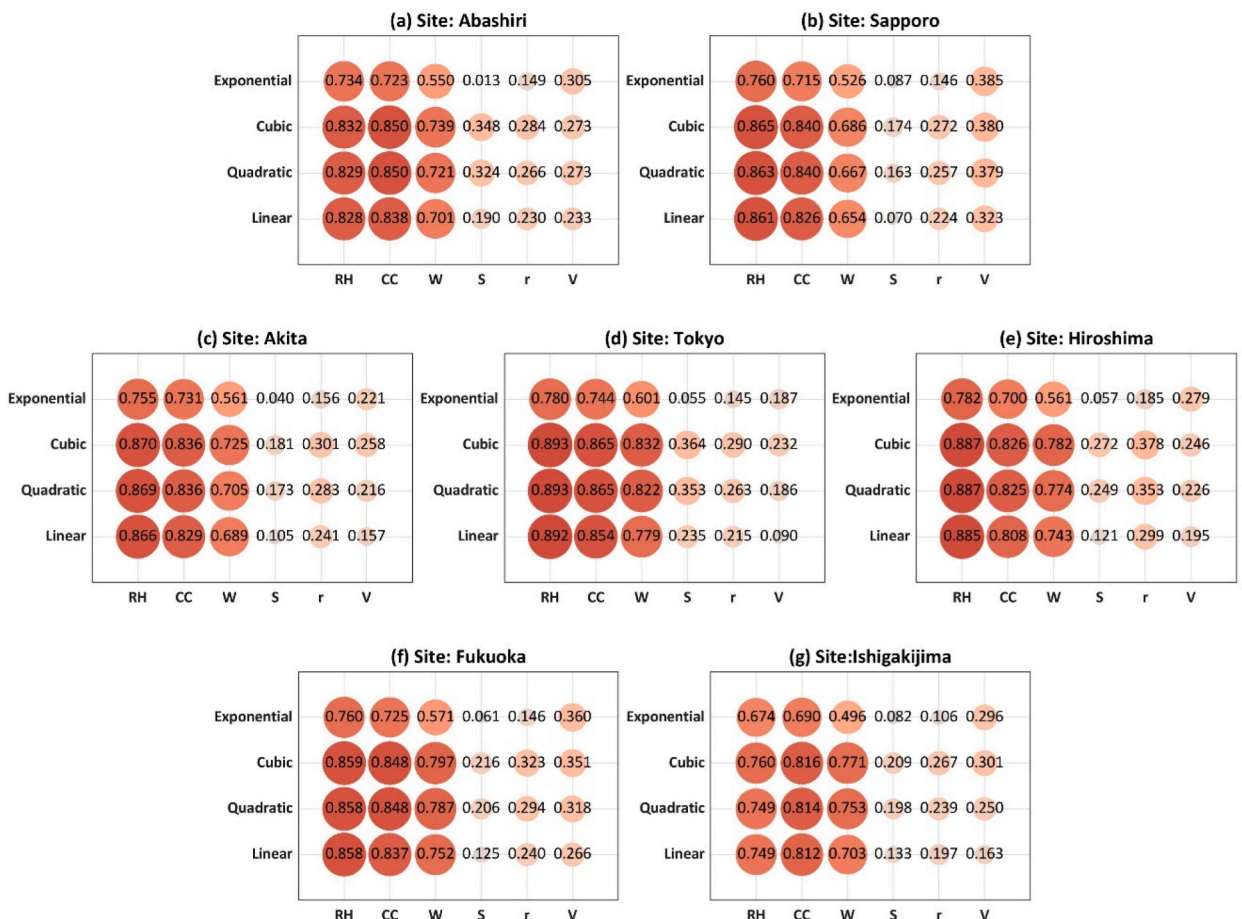

Fig. 4. Pearson coefficient (R) of models built by various equations (linear, Quadratic, Cubic, and Exponential).

higher comparing with the Pearson coefficients and weakly correlated to GSR. To simplify the computation, the visibility (V), precipitation (r), weather type (W), relative humidity (RH), sunshine duration (S), and cloud cover (CC)are selected to further feature analysis.

3.2. Model formation

This paper considers *RH, CC, W, S, r, V* as potential feature variables for the three weakened layers. The basis for establishing an equational form as Eq. (6) is proposed. This section discusses the most suitable equational expressions for them, including the quadratic polynomial function, linear function, exponential function, and cubic polynomial function, which are adopted for comparison. Fig. 4 (a–g) presents the results, where circle's the size and the color the indicate the Pearson coefficient with a higher value indicating a stronger correlation (closer to 1). The correlation between ΔI and *W, CC, RH* appears to be higher. At the same time, the cubic polynomial form of six representative variables exhibits a highest R-value among all seven locations. Consequently, the cubic polynomial equation is employed to build the model for GSR weakening effects. Additionally, there is an importance to declare despite showcasing only 4 equation functions in the main text for fitting, we have actually employed over 10 diverse functions, encompassing logarithmic form, power form, logistic form, and more. Due to their complexity or weak relevance to functions, we have opted not to include the remaining equations in the main text, in order to maintain brevity.

The proposed function can be characterized as Eq. (12):

$$\Delta I = \sin\theta \cdot \begin{bmatrix} c_1 RH + c_2 RH^2 + c_3 RH^3 + c_4 CC + c_5 CC^2 + c_6 CC^3 + \\ c_7 W + c_8 W^2 + c_9 W^3 + c_{10} S + c_{11} S^2 + c_{12} S^3 + \\ c_{13} r + c_{14} r^2 + c_{15} r^3 + c_{16} V + c_{17} V^2 + c_{18} V^3 + c_0 \end{bmatrix} \tag{12}$$

where the cloud cover and weather type are represented by *CC* and *W* in (dimensionless). *r* represents the precipitation (mm). The sunshine duration is denoted by *S* in (hour). *V* denotes visibility (km). The relative humidity is represented by *RH* in (%). c_0 to c_{18} represent the constants in (dimensionless). Table 5 provides the coefficient tables for the seven locations.

3.3. Comprehensive forecast result

Taking into account environmental conditions and ensuring the reliability of the data, the validation data to calculate the estimated irradiation data using the assumed model has been selected from 1st Jan. Of 2019 to 31st Dec. Of 2021. Fig. 5(a–g) illustrates the regression analysis for validation data. It showcases seven representative observation sites in a north-to-south order: Abashiri, Sapporo, Akita, Tokyo, Hiroshima, Fukuoka, and Ishigakijima. The coefficient of determination (R^2) for the proposed weakening model is as follows for the observation sites: Abashiri (0.93), Sapporo (0.86), Akita (0.88), Tokyo (0.96), Hiroshima (0.9), Fukuoka (0.88), and Ishigakijima (0.84), (plotted in Fig. 5). The average of R^2 approaches to 0.89. The regression analysis closely aligns with the fit line, indicating consistent trends across all locations. Due to the challenges in obtaining the satellite data for accurate comparison, Zhang and Huang model has been selected as the comparison model to evaluate the difference in accuracy. The performance of the Zhang and Huang model is proven in numerous existing researches, as it is established based on the connection among GSR, cloud cover, relative humidity, and 3-h difference of air temperature. The function equation of Zhang and Huang model can be referred in Table 2. Following the coefficient alignment using the weather data measured in Japan, the Zhang and Huang model exhibits a mean R^2 value of 0.83 and the best R^2 value of 0.86 (Fig. 6(a–g)). Consequently, the proposed model demonstrates an average improvement in

Table 5
Table of coefficients.

| Coefficient	Ishigakijima	Fukuoka	Hiroshima	Tokyo	Akita	Sapporo	Abashiri
$c0$	0.395	0.474	0.485	0.649	0.495	0.414	0.569
$c1$	0.35	−0.236	−0.68	0.27	−0.202	0.051	0.663
$c2$	−0.574	1.104	2.226	0.062	0.678	−0.013	−1.601
$c3$	0.493	−0.648	−1.262	−0.085	−0.151	0.418	1.218
$c4$	0.69	−0.139	0.018	0.16	−0.193	−0.254	0.32
$c5$	−1.77	0.07	−0.347	−0.404	0.023	0.223	−0.758
$c6$	1.37	0.143	0.428	0.333	0.224	0.082	0.579
$c7$	−0.458	0.53	0.135	−0.504	1.084	0.607	−0.67
$c8$	1.229	−0.478	0.129	1.119	−2.23	−0.462	1.809
$c9$	−0.771	0.051	−0.253	−0.588	0.501	−0.23	−1.336
$c10$	−0.653	−0.664	−0.591	−0.87	−0.677	−0.617	−0.936
$c11$	1.048	1.148	0.944	1.07	1.065	0.904	1.116
$c12$	−0.609	−0.7	−0.552	−0.602	−0.62	−0.498	−0.605
$c13$	0.857	0.714	0.342	0.262	0.238	0.392	1.233
$c14$	−2.229	−1.502	−0.6	−0.498	−0.245	−0.785	−3.083
$c15$	1.493	0.946	0.359	0.278	0.112	0.638	2.168
$c16$	0.108	0.141	0.074	−0.23	0.286	0.436	0.873
$c17$	−0.446	−0.458	−0.123	0.371	−0.92	−0.787	−2.546
$c18$	0.362	0.357	0.06	−0.156	0.759	0.416	1.94
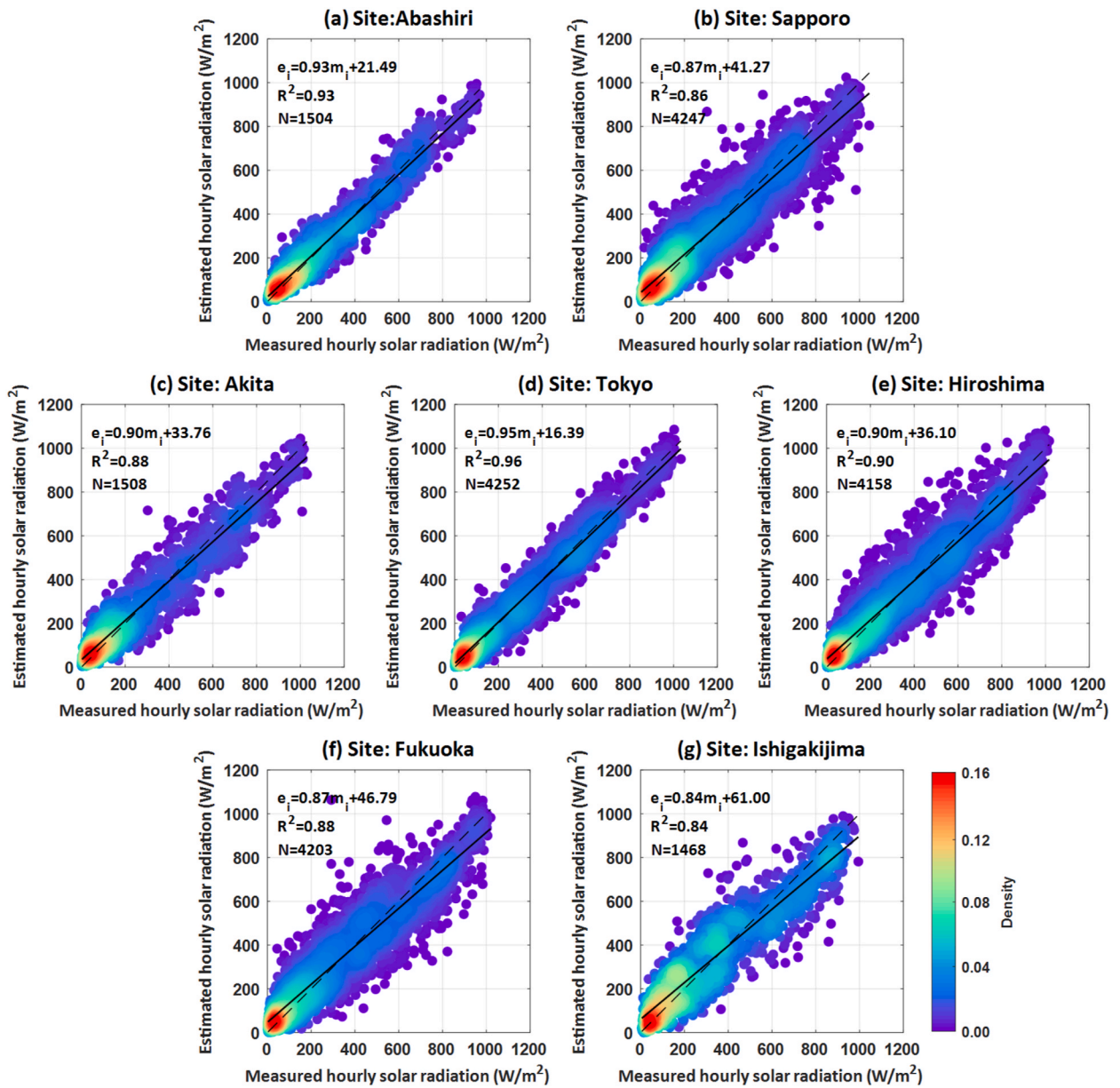
Fig. 5. Scatterplot of predicted and actual hourly irradiation data in Japan, spanning from 1st Jan. Of 2019 to 31st Dec. Of 2021.

accuracy of 7.59% and a maximum improvement of 11.63%.

A summary of others tested statistical indicators, including R, RSE, RMSE, rRMSE and NSE, can be found in Table 6. The more the values of NSE and R approach one, and the values of RMSE and RSE approach zero, the greater the reliability and accuracy. The result indicates that the optimal and average for NSE, RMSE, RSE, and R are as follows: R exhibits a best value of 0.979 and a mean value of 0.945, RSE displays a best value of 0.552 and a mean value of 0.756, RMSE demonstrates a best value of 59.024 and a mean value of 88.2, and NSE shows a best value of 0.946 and a mean value of 0.884. Yao et al. [51] conducted a comparison of 15 hourly solar radiation models, which consisted of 4 proposed solar radiation models and 11 previous decomposition models. They conducted that the RMSE values predominantly fell within the range of 131.4–142.2 W/m². In a previous work by J. Zhang et al. [52], a comprehensive review and comparison of solar radiation prediction models was conducted. The findings revealed that HGSR models predominantly exhibited RMSE values ranging from 88.3 to 142.2 W/m². Li [21] developed 6 models based on weather items including parameters maximum and minimum air temperatures, relative humidity, sunshine duration. These models are compared, then the analysis revealed that a significant portion of rRMSE fell within the range of 24.9% and 41.1% across 83 locations. Upon comparison with the aforementioned conclusions, there is evident that the proposed model exhibits enhanced forecast performance when compared to measured data.
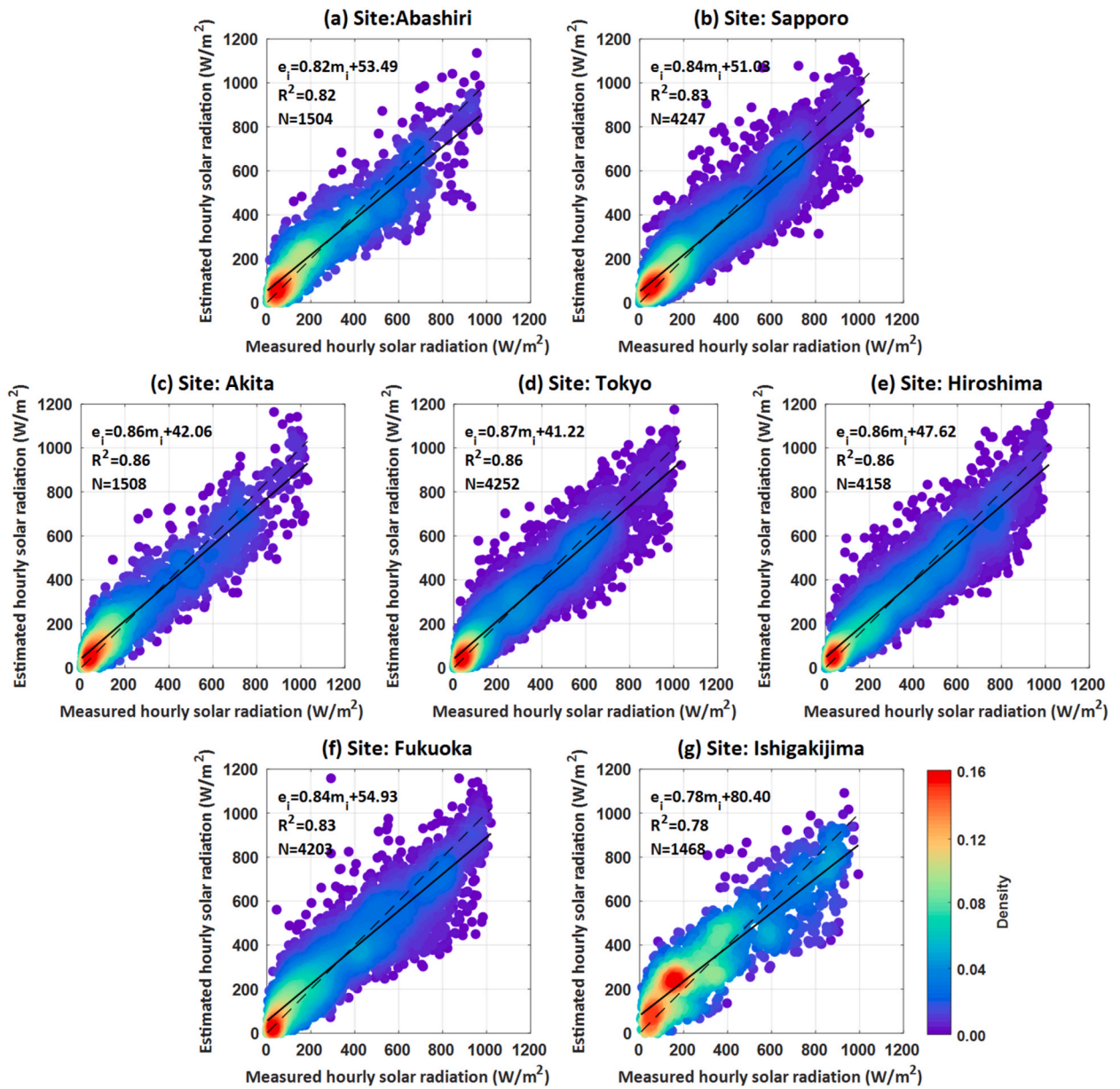
Fig. 6. Scatterplot of predicted and actual hourly irradiation data in Japan of Zhang and Huang model (from1st Jan. Of 2019 to 31st Dec. Of 2021).

Table 6
The other statistical indexes for validation.

Site	The Zhang and Huang model					The proposed model				
	RSE	R	rRMSE	RMSE	NSE	RSE	R	rRMSE	RMSE	NSE
Ishigakijima	1.281	0.883	34.733	125.980	0.715	1.415	0.917	28.912	104.880	0.815
Fukuoka	1.539	0.910	33.266	110.930	0.802	0.612	0.938	28.531	95.139	0.859
Hiroshima	1.437	0.927	28.956	99.808	0.839	0.828	0.949	24.935	85.946	0.886
Tokyo	1.118	0.926	30.241	94.445	0.847	0.552	0.979	18.898	59.024	0.946
Akita	1.398	0.927	33.362	101.096	0.839	0.666	0.938	27.195	82.407	0.896
Sapporo	1.216	0.911	33.187	103.201	0.803	0.821	0.930	28.768	89.462	0.858
Abashiri	1.516	0.905	34.801	103.751	0.781	0.404	0.965	21.047	62.748	0.929
Mean	1.358	0.912	32.649	105.602	0.803	0.756	0.945	25.271	82.800	0.884

3.4. Interaction effects

Indeed, physical interactions such as clouds and particles have the potential to generate feedback that influences transmission. Therefore, it is essential to take into account the interaction effects among meteorological items. In statistical analysis, when one explanatory variable influences the relationship between the other variables to the dependent items, that is typically represented as a product value between those two distinct parameters [53]. The performance the proposed expansion model upon the inclusion of interaction variables is compared in Table 7. Model 1 is selected as the reference model for comparison. For the sake of calculation, all parameters are included as linear functions in the expansion model. The findings indicate that incorporating the interactive variables of weather type and cloud cover can enhance the model's accuracy, as well as the interactive variables of cloud cover with sunshine duration, relative humidity and weather type, respectively. Specifically, the inclusion of the interaction variable between cloud cover and sunshine duration (model 5) contributes to an enhancement of 0.08 in R^2 of the model. This is enhancement is attributed to the fact that sunshine duration serves as a precise indicator of whether the sun is obstructed by cloudiness [54]. Moreover, variations in clouds typically coincide with shifts in relative position of the clouds in relation to the sun. Then model 6 is comparison model without interaction items. And the exhibition of the model 6 also demonstrates that R^2 will climbs up and cannot solely attributed to the addition of variables. Hence, the incorporation of those interaction terms can be considered, and it holds high physical significance.

3.5. New feature of solar radiation model

Based on the previous research, numerous models of HGSR predominantly rely on clear sky conditions. Such models exhibit high performance when weather extreme events as well as variation in clouds is not taken into account. Additionally, it is possible to calculate the theoretical irradiation on the horizontal plane assuming clear-sky states. Nevertheless, given the yearly climate change and the escalating frequency of weather extreme conditions, there is an urgent need for solar radiation models that can account for various types of weather conditions. At the same time, it is not clear how to modify existing hourly solar radiation models (e.g., Zhang and Huang model) for different weather adaptations.

Consequently, the validation of proposed attenuation model is conducted based on the classified weather types presented in Table 8. Where the weather types are divided to three distinct groups: rainy (which involve moisture-mixing turbid conditions), cloudy, and sunny (encompassing mostly clear and sunny, as well as slightly cloudy).

In theory, the prediction of GSR becomes more challenging as weather conditions and becomes more complex and atmospheric turbidity increase. This is because the seasonal character of solar radiation is attenuated by the stochastic character of weather. This is the main reason why solar radiation is difficult to interpret and apply on overcast days. However, we found that there is an opportunity to explain the randomness of solar radiation separately from the seasonal characteristics through formal innovation. The findings indicate that the weakening value of solar radiation demonstrates better adaptability in cloudy or rainy situations. For all weather stations, the average values of R^2 are 0.96, 0.90 and 0.71 for rainy, cloudy, and sunny states, respectively, when using the attenuation value of GSR in the air as the forecast targets (plotted in Fig. 7(a–u)). This can be attributed to the fact that the assumed model is grounded on the solar radiation attenuation effects. In situations where the turbidity of atmosphere is getting higher, both the diffusion or absorption effect of irradiation are more pronounced in the atmosphere, thus influencing a performance of the model. The weakening process holds significant importance, which makes it easier to observe the connection among the feature variables (e.g., weather type, sunshine duration) of those attenuation layers.

3.6. Interpretability

Indeed, interpretability plays a crucial role in solar radiation models as it enables a better understanding of the model's functioning and provides insights for model improvement. The contribution of this work is to provide a new way of thinking to investigate the factors influencing solar radiation. In contrast to commonly used models, we try to isolate the seasonal and stochastic characteristics of solar radiation and look for characterizing variables in simply obtained surface meteorological parameters.

As depicted in Fig. 4, the most substantial influences on the stochastic characteristics of GSR are observed from *RH*, *CC*, and *W*. These factors exhibit the propensity of GSR to be attenuated in those attenuation layers, which signifies the turbidity of the air. Certainly, the distance of irradiation travels across the air must be taken into account, which aligns with our physical perception. At the same time, the distribution trends between these parameters and their combinations can provide good feedback for turbid skies (Fig. 8 (a–f)).

Table 7
The models with interaction items.

No.	Formulation	R^2 (hourly)
1	$\Delta I = \sin\theta \cdot (c_1 RH + c_2 CC + c_3 W)$	0.81
2	$\Delta I = \sin\theta \cdot (c_1 RH + c_2 CC + c_3 W + c_4 CC \cdot W)$	0.82
3	$\Delta I = \sin\theta \cdot (c_1 RH + c_2 CC + c_3 W + c_4 S \cdot W)$	0.87
4	$\Delta I = \sin\theta \cdot (c_1 RH + c_2 CC + c_3 W + c_4 S \cdot RH)$	0.87
5	$\Delta I = \sin\theta \cdot (c_1 RH + c_2 CC + c_3 W + c_4 S \cdot CC)$	0.89
6	$\Delta I = \sin\theta \cdot (c_1 RH + c_2 CC + c_3 W + c_4 S)$	0.84
7	$\Delta I = \sin\theta \cdot (c_1 RH + c_2 CC + c_3 W + c_4 S + c_5 S \cdot CC)$	0.91

Table 8
Classified weather types.

Symbol	Descriptions	Definitions
1	Mostly clear and sunny	Sunny
2	Sunny	
3	Slightly cloudy	
4	Cloudy	Cloudy
5	Mist	
6	Dust storm	
7	Drifting snow	
8	Fog	
9	Drizzle	Rainy
10	Rain	
11	Sleet	
12	Snow	
13	Hail	
14	Hailstorm	
15	Thunderstorm	
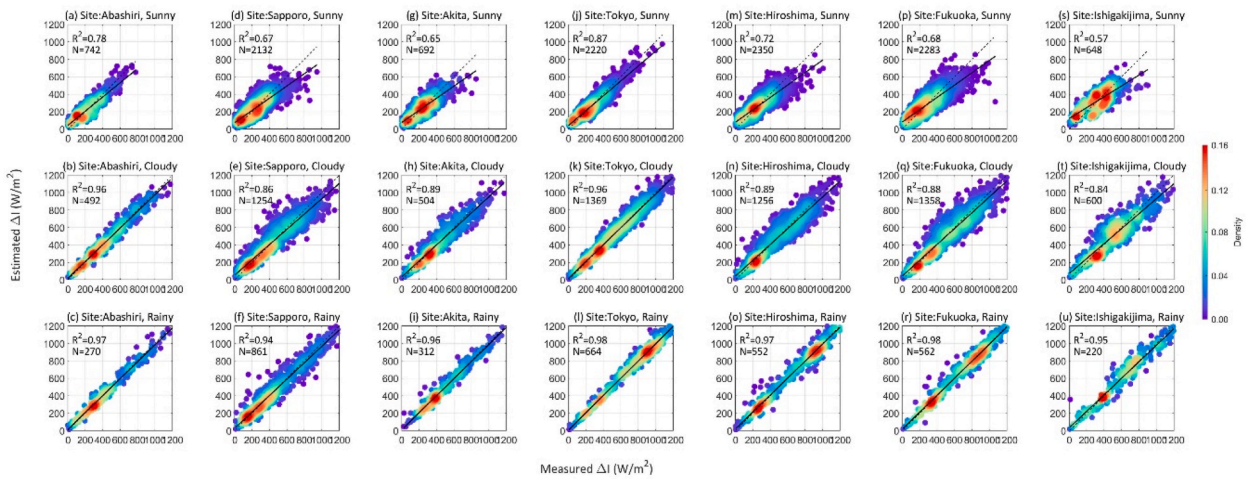
Fig. 7. Scatterplot of predicted and actual hourly irradiation data in Japan under rainy cloudy and sunny conditions.

4. Conclusions

In this work, we introduce a novel HGSR model that incorporates the layer-by-layer attenuation effects. A dataset comprising 205,722 datasets of weather data that measured from 7 stations in Japan was utilized for both model development and validation processes. Taking into account the simplicity of data acquisition, we chose to include only readily available six meteorological parameters from weather stations as input variables. Furthermore, the model's accuracy was assessed in this work across diverse weather states. The specific conclusions can be summarized as follows.

1) An attenuation hourly solar radiation model is introduced according to the weakening effects of irradiation travels through the air. The maximum and average values of R^2 are found to be 0.96 and 0.89, respectively. The proposed model exhibited an average improvement of 7.59% in accuracy and a best improvement of 11.63% in comparison to the existing model.
2) The characteristics of solar radiation are determined by a combination of seasonality and stochasticity. we try to isolate the seasonal and stochastic characteristics of solar radiation and look for characterizing variables in simply obtained surface meteorological parameters. Stochasticity can be characterized by a variable representing weather turbidity, while seasonality can be represented by the interaction term between the seasonality variable and the weather turbidity variable. This will serve as a valuable comment on many solar radiation models that are primarily developed in clear-sky conditions, enabling them to more intricate weather conditions.
3) The attenuation hourly solar radiation model is a versatile framework allowing for the enhancement or replacement of the characterization factors of the weakening layers based on the attenuation effects. The incorporation of interactive items will notably enhance the model's accuracy, particularly through the inclusion of the interactive variable involving cloud cover and sunshine duration.
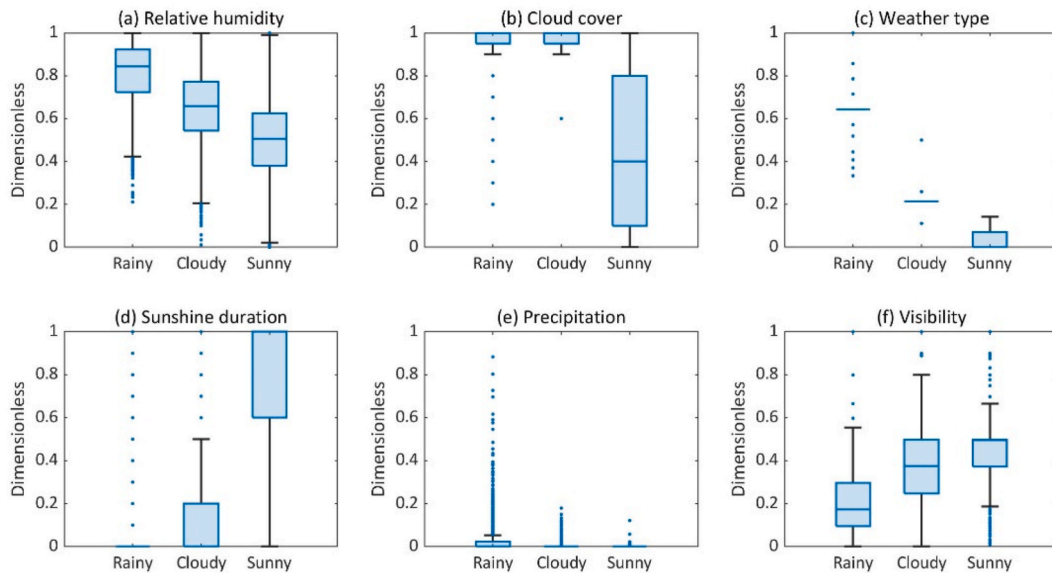
Fig. 8. Boxplot of meteorological parameters under different weather conditions.

4) The attenuation hourly solar radiation model is built and validated based on more than twenty years of actual weather data collected from seven Japanese distinct stations, encompassing diverse weather states. Consequently, our approach enables the generation of a typical meteorological year (TMY) specifically tailored for Japan. This holds significant value for various applications and techniques, e.g., the photovoltaic assessment throughout the year.

Author contribution statement

You Li: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper. Yafei Wang: Performed the experiments; Analyzed and interpreted the data; Wrote the paper. Hui Qian: Analyzed and interpreted the data. Weijun Gao: Analyzed and interpreted the data. Hiroatsu Fukuda: Conceived and designed the experiments. Weisheng Zhou: Analyzed and interpreted the data.

Funding

Data availability statement

The authors do not have permission to share data.

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.

Acknowledgements

Appendix A

A.1. Description of cloud cover and weather type in weather forecast

Appendix Table 1
Cloud cover symbol description

Symbol	Description	Numericalization
blank	No observations	blank
–	No clouds are observed	0
0.0	If clouds are present but cloud cover is less than 0.1	0
0	If clouds are present but cloud cover is less than 1	0
0+	Cloud cover is more than 0.1 but less than 1	0.5
1	One-tenth of the total cloud cover	1
2	Two-tenths of the total cloud cover	2
3	Three-tenths of the total cloud cover	3
4	Four-tenths of the total cloud cover	4
5	Five-tenths of the total cloud cover	5
6	Six-tenths of the total cloud cover	6
7	Seven-tenths of the total cloud cover	7
8	Eight-tenths of the total cloud cover	8
9	Nine-tenths of the total cloud cover	9
10-	If cloud cover is 10 but there are areas with no clouds	9.5
10	The total cloud cover	10

Appendix Table 2
Weather type description

Symbol	Description
1	Mostly clear and sunny
2	Sunny
3	Slightly cloudy
4	Cloudy
5	Mist
6	Dust storm
7	Drifting snow
8	Fog
9	Drizzle
10	Rain
11	Sleet
12	Snow
13	Hail
14	Hailstorm
15	Thunderstorm

1.2. Correlation coefficients based on Eq. (1) and Eq. (2)

Appendix Fig. 1 shows the correlation coefficients based on Eq. (1) and Eq. (2), respectively. It can be found that the sensitivity of the two methods to various meteorological parameters is diverse. In the method based on actual solar radiation on the horizontal surface, sunshine duration, visibility, and 3-h temperature difference have the highest correlation to solar radiation values. This is because on a large scale, solar radiation reflects certain seasonal characteristics and therefore correlates better with meteorological parameters that can represent date or seasonal characteristics, such as sunshine duration and temperature differences. While in the proposed model, relative humidity, weather type, and cloudiness have the highest correlation on the solar radiation loss values. This is because the effect of these parameters on the reduction of solar radiation as it passes through the atmosphere is the most significant.
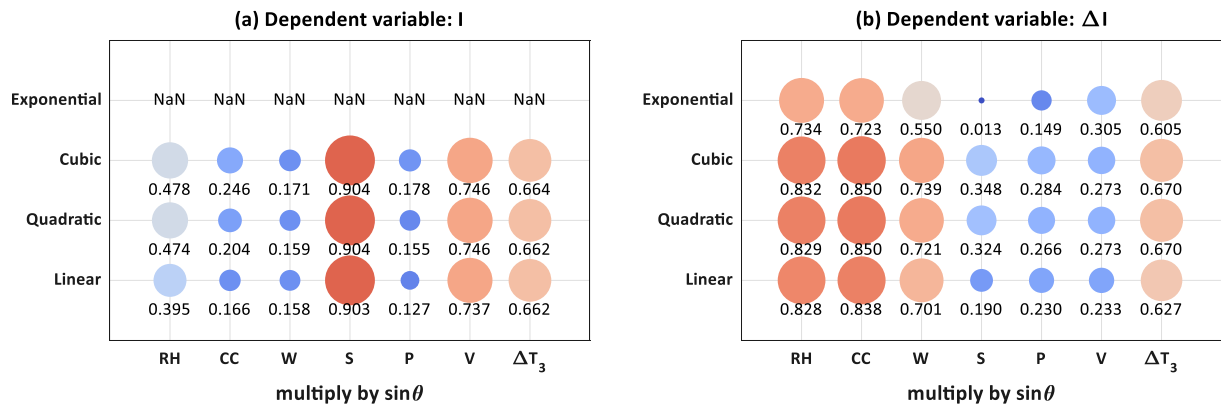
Appendix Fig. 1. Correlation coefficients (R) of the two methods for different meteorological parameters in Abashiri, (NaN means that the function form is not applicable to the method), ($\Delta I_3 = I_n - I_{n-3}$).

References

[1] A.E. Gurel, U. Agbulut, H. Bakir, A. Ergun, G. Yildiz, A state of art review on estimation of solar radiation with various models, Heliyon 9 (2023), e13167.

[2] D.W. Van der Meer, J. Widén, J. Munkhammar, Review on probabilistic forecasting of photovoltaic power production and electricity consumption, Renew. Sustain. Energy Rev. 81 (2018) 1484–1512.

[3] J. Yan, Y. Yang, P. Elia Campana, J. He, City-level analysis of subsidy-free solar photovoltaic electricity price, profits and grid parity in China, Nat. Energy 4 (2019) 709–717.

[4] J. Sward, T. Ault, K. Zhang, Genetic algorithm selection of the weather research and forecasting model physics to support wind and solar energy integration, Energy (2022), 124367.

[5] Y. Dai, Y. Wang, M. Leng, X. Yang, Q. Zhou, LOWESS smoothing and Random Forest based GRU model: a short-term photovoltaic power generation forecasting method, Energy 256 (2022), 124661.

[6] L. Xiao, L.-L. Qin, S.-Y. Wu, Proposal and application of comprehensive thermal comfort evaluation model in heating seasons for buildings with solar Trombe wall, Appl. Therm. Eng. (2022), 118774.

[7] Y. Yin, H. Chen, X. Zhao, W. Yu, H. Su, Y. Chen, et al., Solar-absorbing energy storage materials demonstrating superior solar-thermal conversion and solar-persistent luminescence conversion towards building thermal management and passive illumination, Energy Convers. Manag. 266 (2022), 115804.

[8] L. Ji, Y. Wu, L. Sun, X. Zhao, X. Wang, Y. Xie, et al., Solar photovoltaics can help China fulfill a net-zero electricity system by 2050 even facing climate change risks, Resour. Conserv. Recycl. 186 (2022), 106596.

[9] R. Qiu, X. Li, G. Han, J. Xiao, X. Ma, W. Gong, Monitoring drought impacts on crop productivity of the US Midwest with solar-induced fluorescence: GOSIF outperforms GOME-2 SIF and MODIS NDVI, EVI, and NIRv, Agric. For. Meteorol. 323 (2022), 109038.

[10] Y. Li, F. Qian, W. Gao, H. Fukuda, Y. Wang, Techno-economic performance of battery energy storage system in an energy sharing community, J. Energy Storage 50 (2022), 104247.

[11] F. Creutzig, P. Agoston, J.C. Goldschmidt, G. Luderer, G. Nemet, R.C. Pietzcker, The underestimated potential of solar energy to mitigate climate change, Nat. Energy 2 (2017) 1–9.

[12] M. Shabani, E. Dahlquist, F. Wallin, J. Yan, Techno-economic impacts of battery performance models and control strategies on optimal design of a grid-connected PV system, Energy Convers. Manag. 245 (2021), 114617.

[13] J. Dong, M.M. Olama, T. Kuruganti, A.M. Melin, S.M. Djouadi, Y. Zhang, et al., Novel stochastic methods to predict short-term solar radiation and photovoltaic power, Renew. Energy 145 (2020) 333–346.

[14] A. Angstrom, Solar and terrestrial radiation. Report to the international commission for solar research on actinometric investigations of solar and atmospheric radiation, Q. J. R. Meteorol. Soc. 50 (1924) 121–126.

[15] J. Prescott, Evaporation from a water surface in relation to solar radiation, Trans. Roy. Soc. S. Aust. 46 (1940) 114–118.

[16] H. Morf, A stochastic solar irradiance model adjusted on the Ångström–Prescott regression, Sol. Energy 87 (2013) 1–21.

[17] M. Paulescu, N. Stefu, D. Calinoiu, E. Paulescu, N. Pop, R. Boata, et al., Ångström–Prescott equation: physical basis, empirical models and sensitivity analysis, Renew. Sustain. Energy Rev. 62 (2016) 495–506.

[18] P.J. Asilevi, E. Quansah, L.K. Amekudzi, T. Annor, N.A.B. Klutse, Modeling the spatial distribution of Global Solar Radiation (GSR) over Ghana using the Ångström-Prescott sunshine duration model, Scientific African 4 (2019), e00094.

[19] G.H. Hargreaves, Z.A. Samani, Estimating potential evapotranspiration, J. Irrigat. Drain. Div. 108 (1982) 225–230.

[20] G.H. Hargreaves, Z.A. Samani, Reference crop evapotranspiration from temperature, Appl. Eng. Agric. 1 (1985) 96–99.

[21] M.-F. Li, X.-P. Tang, W. Wu, H.-B. Liu, General models for estimating daily global solar radiation for different solar radiation zones in mainland China, Energy Convers. Manag. 70 (2013) 139–148.

[22] J. Almorox, M. Bocco, E. Willington, Estimation of daily global solar radiation from measured temperatures at Cañada de Luque, Córdoba, Argentina, Renew. Energy 60 (2013) 382–387.

[23] M. Iqbal, A study of Canadian diffuse and total solar radiation data—I monthly average daily horizontal radiation, Sol. Energy 22 (1979) 81–86.

[24] J. Bugler, The determination of hourly insolation on an inclined plane using a diffuse irradiance model based on hourly measured global horizontal insolation, Sol. Energy 19 (1977) 477–491.

[25] T. Watanabe, Procedures for separating direct and diffuse insolation on a horizontal surface and prediction of insolation on tilted surface, Transactions of the Architectural Institute of Japan 330 (1983) 96–108.

[26] A.L. Flint, S.W. Childs, Calculation of solar radiation in mountainous terrain, Agric. For. Meteorol. 40 (1987) 233–249.

[27] K. Chang, Q. Zhang, Development of a Solar Radiation Model Considering the Hourly Sunshine Duration for All-Sky Conditions – A Case Study for Beijing, 234, Atmospheric Environment, China, 2020.

[28] K. Chang, Q. Zhang, Improvement of the hourly global solar model and solar radiation for air-conditioning design in China, Renew. Energy 138 (2019) 1232–1238.

[29] K.H. Kim, J.-C. Baltazar, J.S. Haberl, Evaluation of meteorological base models for estimating hourly global solar radiation in Texas, Energy Proc. 57 (2014) 1189–1198.

[30] K.H. Kim, J.K.-W. Oh, W. Jeong, Study on solar radiation models in South Korea for improving office building energy performance analysis, Sustainability 8 (2016) 589.

[31] A. Geetha, J. Santhakumar, K.M. Sundaram, S. Usha, T.M.T. Thentral, C.S. Boopathi, et al., Prediction of hourly solar radiation in Tamil Nadu using ANN model with different learning algorithms, Energy Rep. 8 (2022) 664–671.

[32] E. Miranda, J.F.G. Fierro, G. Narvaez, L.F. Giraldo, M. Bressan, Prediction of site-specific solar diffuse horizontal irradiance from two input variables in Colombia, Heliyon 7 (2021), e08602.

[33] F. Nawab, A.S. Abd Hamid, A. Ibrahim, K. Sopian, A. Fazlizan, M.F. Fauzan, Solar irradiation prediction using empirical and artificial intelligence methods: a comparative review, Heliyon 9 (2023), e17038.

[34] Y. Gao, S. Miyata, Y. Akashi, Multi-step solar irradiation prediction based on weather forecast and generative deep learning model, Renew. Energy 188 (2022) 637–650.

[35] C.S. Lai, C. Zhong, K. Pan, W.W.Y. Ng, L.L. Lai, A deep learning based hybrid method for hourly solar radiation forecasting, Expert Syst. Appl. (2021) 177.

[36] Y. Gao, S. Miyata, Y. Akashi, Interpretable deep learning models for hourly solar radiation prediction based on graph neural network and attention, Appl. Energy 321 (2022).

[37] R.E. Bird, R.L. Hulstrom. Simplified Clear Sky Model for Direct and Diffuse Insolation on Horizontal Surfaces. Solar Energy Research Inst., Golden, CO (USA) 1981.

[38] Y. Li. Study On Energy Sharing In Pv Communities Considering Solar Radiation Prediction. The University of Kitakyushu2023.

[39] G. Su, S. Zhang, M. Hu, W. Yao, Z. Li, Y. Xi, The modified layer-by-layer weakening solar radiation models based on relative humidity and air quality index, Energy 239 (2022).

[40] J.-I. Prieto, D. García, Global solar radiation models: a critical review from the point of view of homogeneity and case study, Renew. Sustain. Energy Rev. (2021), 111856.

[41] W. Yao, Z. Li, Y. Wang, F. Jiang, L. Hu, Evaluation of global solar radiation models for Shanghai, China, Energy Convers. Manag. 84 (2014) 597–612.

[42] E. Falayi, J. Adepitan, A. Rabiu, Empirical models for the correlation of global solar radiation with meteorological data for Iseyin, Nigeria, Int. J. Phys. Sci. 3 (2008) 210–216.

[43] M.-F. Li, H.-B. Liu, P.-T. Guo, W. Wu, Estimation of daily solar radiation from routinely observed meteorological data in Chongqing, China, Energy Convers. Manag. 51 (2010) 2575–2579.

[44] J. Black, The distribution of solar radiation over the earth's surface, Archiv für Meteorologie, Geophysik und Bioklimatologie, Serie B. 7 (1956) 165–189.

[45] J. Glover, J. McCulloch, The empirical relation between solar radiation and hours of sunshine, Q. J. R. Meteorol. Soc. 84 (1958) 172–175.

[46] R. Swartman, O. Ogunlade, Solar radiation estimates from common parameters, Sol. Energy 11 (1967) 170–172.

[47] H. Ueyama, Estimating hourly direct and diffuse solar radiation for the compilation of solar radiation distribution maps, J. Agric. Meteorol. 61 (2005) 207–216.

[48] H. Ueyama, Development of statistical methods for estimating hourly direct and diffuse solar radiation using public data for precise cultivation management, J. Agric. Meteorol. 74 (2018) 29–39.

[49] Q. Zhang, J. Huang, S. Lang, SYMPOSIUM PAPERS-HI-02-16 recent developments in availability of international weather data-development of typical year weather data for Chinese locations, ASHRAE Transactions-American Society of Heating Refrigerating Airconditioning Engin 108 (2002) 1063–1078.

[50] Y. Li, Y. Wang, W. Yao, W. Gao, H. Fukuda, W. Zhou, Graphical decomposition model to estimate hourly global solar radiation considering weather stochasticity, Energy Convers. Manag. 286 (2023), 116719.

[51] W. Yao, Z. Li, T. Xiu, Y. Lu, X. Li, New decomposition models to estimate hourly global solar radiation from the daily value, Sol. Energy 120 (2015) 87–99.

[52] J. Zhang, L. Zhao, S. Deng, W. Xu, Y. Zhang, A critical review of the models used to estimate solar radiation, Renew. Sustain. Energy Rev. 70 (2017) 314–329.

[53] J. Jaccard, C.K. Wan. LISREL Approaches to Interaction Effects in Multiple Regression. sage1996.

[54] J. Chakchak, N.S. Cetin, Investigating the impact of weather parameters selection on the prediction of solar radiation under different genera of cloud cover: a case-study in a subtropical location, Measurement 176 (2021), 109159.

Nomenclature

CC: Cloud cover
e_i: Estimated value
E_0: Eccentricity factor
GSR: Global solar radiation (W/m^2)
HGSR: Hourly global solar radiation (W/m^2)
I_0: Extraterrestrial solar radiation (W/m^2)
I_{sc}: Solar constant (equal to 1361 W/m^2)
ΔI: Loss value of solar radiation through the atmosphere (W/m^2)
ΔT: Daily maximum temperature difference (°C)
ΔT_3: Three-hour temperature difference (°C)
L_a: Loss value of solar radiation in the aerosol layer
L_w: Loss value of solar radiation in the water vapor layer
L_c: Loss value of solar radiation in the clouds layer
m_i: Measured value
n: The number of estimated values of measured data
NSE: Nash-Sutcliffe equation
P: Precipitation (mm)
R: Correlation coefficient
RH: Relative humidity (%)
RMSE: Root mean square error
RSE: Relative standard error
rRMSE: Relative root mean square error
S: Sunshine duration (hour)
V: visibility (km)
W: Weather type
δ: Solar declination angle (degrees)
φ: Geographical latitude (degrees)
θ: Solar altitude angle (degrees)
ω: Solar hour angle (degrees)
c_0 to c_{18}: Cubic polynomial factor