

Article

Comparison of Artificial Intelligence and Physical Models for Forecasting Photosynthetically-Active Radiation

Lan Feng ^{1,2}, Wenmin Qin ^{1,†}, Lunche Wang ^{1,*}, Aiwèn Lin ² and Ming Zhang ¹

¹ Laboratory of Critical Zone Evolution, School of Earth Sciences, China University of Geosciences, Wuhan 430074, China; lanfeng@whu.edu.cn (L.F.); 2010202050019@whu.edu.cn (W.Q.); mingzhang@whu.edu.cn (M.Z.)

² School of Resource and Environmental Science, Wuhan University, Wuhan 430079, China; awlin@whu.edu.cn

* Correspondence: wang@cug.edu.cn; Tel.: +86-027-6788-3001

† These authors contributed equally to this work.

Received: 22 September 2018; Accepted: 16 November 2018; Published: 21 November 2018



Abstract: Different kinds of radiative transfer models, including a relative sunshine-based model (BBM), a physical-based model for tropical environment (PBM), an efficient physical-based model (EPP), a look-up-table-based model (LUT), and six artificial intelligence models (AI) were introduced for modeling the daily photosynthetically-active radiation (PAR, solar radiation at 400–700 nm), using ground observations at twenty-nine stations, in different climatic zones and terrain features, over mainland China. The climate and terrain effects on the PAR estimates from the different PAR models have been quantitatively analyzed. The results showed that the Genetic model had overwhelmingly higher accuracy than the other models, with the lowest root mean square error (RMSE = 0.5 MJ m⁻²day⁻¹), lowest mean absolute bias error (MAE = 0.326 MJ m⁻²day⁻¹), and highest correlation coefficient (R = 0.972), respectively. The spatial–temporal variations of the annual mean PAR (APAR), in the different climate zones and terrains over mainland China, were further investigated, using the Genetic model; the PAR values in China were generally higher in summer than those in the other seasons. The Qinghai Tibetan Plateau had always been the area with the highest APAR (8.668 MJ m⁻²day⁻¹), and the Sichuan Basin had always been the area with lowest APAR (4.733 MJ m⁻²day⁻¹). The PAR datasets generated by the Genetic model, in this study, could be used in numerous PAR applications, with high accuracy.

Keywords: photosynthetically-active radiation; physical models; artificial neural network; climate zones; terrain features

1. Introduction

Solar radiation is the most important energy sources, driving the sources and sinks of energy between the earth surface and atmosphere [1]. Ninety-nine percent of the solar energy concentrated in the wavelength of 0.25 μm–4.0 μm, which can be divided into three broadbands—ultraviolet spectrum (0.25–0.40 μm), visible spectrum (0.40–0.70 μm), and near infrared spectrum (0.70–4 μm). Among them, almost 45% of the solar radiation concentrated in the visible spectrum, namely photosynthetically-active radiation (PAR) [2]. PAR is of vital importance in many biological and physical processes, such as chlorophyll synthesis and plant photosynthesis [3–5]. Thus, in a large number of studies accurate observations and estimations of the PAR at the given location, are required [6,7], including concentrating solar power (CPV), meteorology, agriculture, land surface ecosystem, and sustainable development. However, due to the difficulties in calibration, construction,

and maintenance of observation platform, in situ PAR stations are relatively sparse for PAR applications, especially in the mountains and remote areas.

In China, the Chinese Ecosystem Research Network (CERN) was established in 2005, to cover the different terrestrial ecosystems, providing high-quality PAR measurements. However, CERN only consisted of about forty field stations, which cannot meet the requirements of large-scale ecological and meteorological studies. PAR can be estimated from surface solar radiation (SSR) at radiation stations, using empirical conversion factors. However, the number of SSR stations was relatively smaller, in comparison with meteorological stations. A lot of PAR products have been calculated using ground measurements and satellite signals [8], such as PAR products from the International Satellite Cloud Climatology Project (ISCCP) [9,10], long-term PAR records obtained by the Total Ozone Mapping Spectrometer (TOMS) [11], and the PAR records obtained by the Moderate Resolution Imaging Spectroradiometer (MODIS) [12]. However, the accuracy, spatial-temporal continuities, and resolutions of these PAR products cannot meet the requirements for PAR applications.

Many attempts have been made to estimate PAR using meteorological measurements and satellite images. These models could be roughly divided into statistical models and physical models. The statistical models assumed the ratio of PAR to SSR (RPAR) as a constant value, for example, Moon [13] found that RPAR was approximately 0.44. However, McCree [14] considered that RPAR would change under various climate conditions. The RPAR under a clear-sky condition was in the range of 0.47–0.52, while the RPAR under a cloudy-sky condition was from 0.5–0.58. Incorporating the atmospheric attenuators, such as cloud fraction, cloud optical depth, aerosol optical depth, and total ozone amount, a lot of statistical models have been developed for simulating PAR, for example, Janjai et al. [15] developed a statistical model to estimate PAR using the cloud index, aerosol optical depth, precipitable water, total ozone column, and the solar zenith angle. The model performance showed that the R^2 was 0.96, root mean square difference (RMSD) was in the range of 7.3–7.9%, mean bias difference (MBD) ranged from −4.5% to 3.5%. Yu et al. [16] compared ten empirical models for estimating PAR, using ground measurements, the quadratic function model, taking into consideration the solar zenith angle and clearness index, performed better than the other models. Tan and Ismail [17] made a comparative study on the three empirical PAR models in Singapore, the model incorporating SSR, and the clearness index showed better accuracy than the other models. Despite the effectiveness of statistical PAR models, the relatively sparse SSR measurements would limit the application of the above models in remote and mountainous areas.

Taking into consideration the extinction processes in atmosphere, such as ozone absorption, Rayleigh scattering, and aerosol extinction, the physical models could retrieve PAR, with a strict theoretical basis. Physical models could be roughly divided into broadband and spectral models. Dividing the solar spectrum into several small ranges, the spectral models could estimate PAR with high spectral resolutions. Many spectral models have been developed for estimating PAR [18,19]. Despite the high spectral resolutions, these models were subjected to poor computational efficiencies. Compared with the spectral models, broadband models divided the solar spectrum into some relatively wide ranges. A lot of broadband models have been developed for retrieving PAR with a high accuracy and computational efficiency [20,21]. However, these models were limited to sparse meteorological stations. Along with surface meteorological measurements, satellite remote sensing provided an efficient way to retrieve PAR, with high spatial and temporal resolutions, at regional and global scales. Satellite signals derived from numerous polar orbiting satellites (for example, the Moderate Resolution Imaging Spectroradiometer (MODIS), Aura, and MSG2) and geostationary satellites (for example, the Geostationary Operational Environmental Satellites (GOES), Oceansat-1, the Multi-Functional Transport Satellite (MTSAT), and the GMS5) have been applied for PAR estimation. Among them, MODIS products were the most widely used products for retrieving PAR, due to its high spatial resolutions and spectral resolutions [22–25]. Despite the effectiveness of the satellite-based PAR models, they were subjected to some limitations, such as calibration accuracy, cloud screening effects, and the uncertainties of surface reflection [26].

Meanwhile, the artificial intelligence (AI) models, such as support vector machine regression (SVM), artificial bee colony (ABC), and artificial neural network (ANN) have been widely applied for estimating PAR around the world, in recent years. Among them, ANN was the most frequently used method for retrieving PAR, with a satisfied accuracy [27]. López et al. [28] developed a feed-forward multilayered perception neural network (MLP) for PAR estimation, using meteorological measurements. Pankaew et al. [29] proposed a model in Thailand, using the cosine of the solar zenith angle (θ), cloud index, water vapor, and the aerosol optical depth (AOD). The result showed that the RMSE between the estimated and the measured PAR at Chiang Mai, UbonRatchathani, NakhonPathom, and Songkhla stations were 10.9%, 9.3%, 8.7%, and 9.8%, respectively. Yu and Guo [30] conducted a comprehensive comparative study on the performances of the six MLP models and four statistical models proposed by Alados and Alados-Arboledas [31], the result indicates that the MLP models performed better than the statistical models. Wang et al. [27] further compared three ANN models in different climate zones, across China, including the MLP, the Generalized Regression Neural Network (GRNN), and the Radial Basis Neural Network (RBNN). The results revealed that MLP and RBNN showed better performances than GRNN. However, these AI models were limited to inherent drawbacks, such as the intrinsic disadvantages of revealing the explicit physics of the atmospheric radiative transfer processes in the atmosphere, and the slow convergence speed. These AI models should be further improved by combining them with other optimization algorithms.

Great efforts have been made to estimate PAR in China. Zhang et al. [32] calculated the RPAR in the Tibet Plateau (about 0.439). Xie et al. [33] developed an empirical model for predicting PAR, using MODIS products, the results showed that the R^2 between the estimated and the measured PAR in Yuncheng was 0.855. Wang et al. [3] proposed an empirical model for PAR estimation in China, during 1961–2012, the estimates showed good agreements with the PAR measurements at CERN stations. Tang et al. [34] reconstructed the daily PAR over China in the past fifty years, using Qin's PAR model [35], the MBE, RMSE, and R^2 were 2.4 Wm^{-2} , 14 Wm^{-2} , and 0.92, respectively. Tang et al. [8] developed an efficient physical-based parameterization to derive the PAR, using the MODIS atmosphere and land products, based on a clear sky transmittance model [35] and a cloud transmittance parameterization [36]. The RMSE between the estimates and the measurements was 40 Wm^{-2} and 15 Wm^{-2} , for instantaneous and daily scales, respectively. Meanwhile, a lot of PAR models using the LUT method have also been developed in China [37,38], for revealing the spatial and temporal variations of PAR in China. Hu et al. [39] found that the highest PAR in China was in the Qinghai-Tibet Plateau with an RPAR of 1.83 (1.68–1.98). Wang et al. [3] reconstructed the seasonal variations of PAR and RPAR in China, during 1955–2011, the result showed that the annual mean PAR and RPAR were $22.39 \text{ mol m}^{-2}\text{day}^{-1}$ and 1.9, respectively. Further comprehensive studies should be conducted to evaluate the model performances in various ecological zones, climate zones, and terrains, and then the spatial and temporal variation of PAR should be reconstructed with high accuracy and robustness, in China.

This study tried to compare the performances of the ten selected PAR models, including the relative sunshine-based broadband model (BBM), the physical-based model for tropical environment (PBM), the efficient physical-based model (EPP), and the look-up-table method (LUT), back propagation neural network (BP), adaptive neuro fuzzy inference system (ANFIS), least squares support vector machine (LSSVM), genetic algorithm (Genetic), M5 model tree (M5Tree), and the multivariate adaptive regression splines (MARS). Daily PAR measurements at twenty-nine CERN stations across China, during 2005–2014, in different climate zones and terrains, were used to evaluate the model performances. The climate and terrain effects on the PAR estimates from the different PAR models have also been quantitatively analyzed. Daily meteorological measurements at eight hundred and thirty-nine Chinese Meteorological Administration (CMA) stations, aerosol optical depth (0.55 nm) retrievals from the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) datasets, and satellite signals from MODIS and MTSAT were also used as model inputs for

estimating PAR in China. The spatial and temporal variations and related causes of PAR in different climate zones and terrain features, across China, were further analyzed in detail.

2. Materials and Methods

2.1. Sites and Data

2.1.1. Observation Data

Daily PAR measurements at twenty-nine CERN stations, throughout China, were used for the model calibrations and validations. Figure 1 shows the spatial distributions of the CERN and CMA stations used in this study. Table 1 shows the general climatic patterns of these CERN stations, Figure 2 also presents the annual mean air temperature, relative humidity, air pressure at sea-level and the sunshine duration (the hours when solar irradiance is greater than 120 Wm^{-2}) in China. It was clear that these CERN and CMA stations spread across most areas of China, with complicated geomorphology and distinctive climatic features.

Table 1. The statistical indicators representing the meteorological patterns of the Chinese Ecosystem Research Network (CERN) stations.

Statistics	a (m)	pre (mm)	ps (hpa)	rh	sh (h/day)	at (°C)	ws (m/s)	vis (km)	trise (h)	tset (h)
Max	3688	263	1044.3	1	15.1	36.4	20.3	41.25	10.09	22.11
Min	3	0	638	0.05	0	−35.1	0	0.06	3.68	15.60
Std	977.26	8.65	91.46	0.19	4.08	11.99	1.33	7.92	1.13	1.16
Mean	754.79	2.28	939.45	0.65	6.25	13.18	2.09	18.37	6.65	18.70

a is altitude, pre—precipitation, ps—surface pressure, rh—relative humidity, sh—sunshine duration, at—air temperature, ws—wind speed, vis—visibility, trise—sunrise time, tset—sunset time.

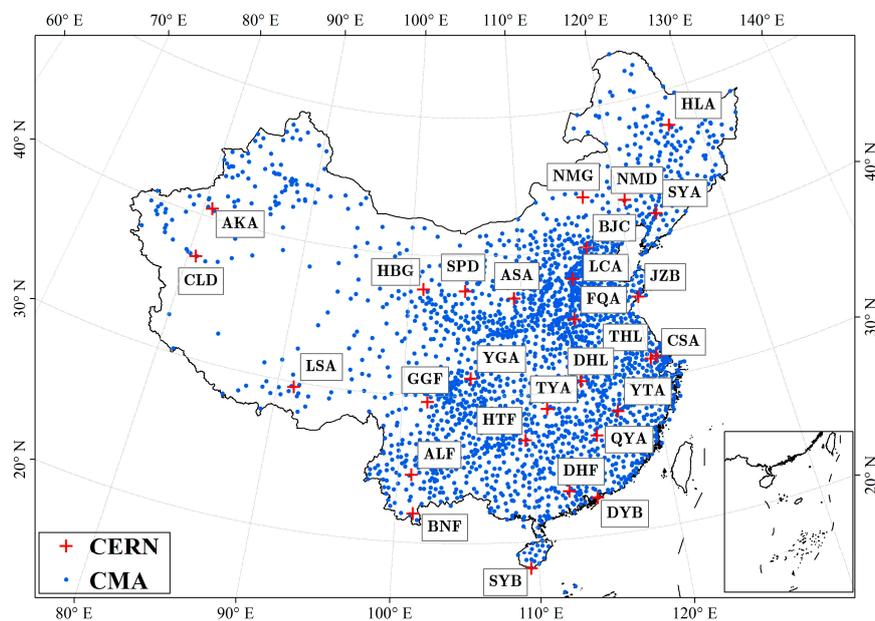


Figure 1. Spatial distribution of the CERN and Chinese Meteorological Administration (CMA) stations used in this study.

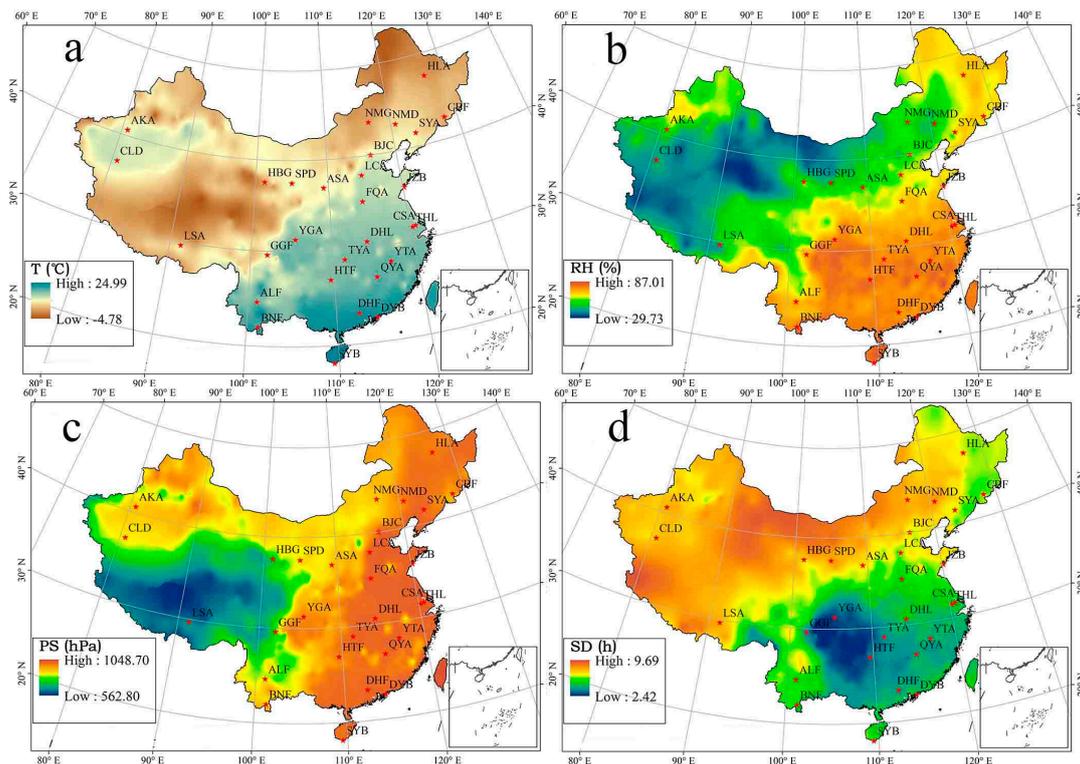


Figure 2. The annual mean values of the meteorological parameters, in China, during 1955–2015; (a) air temperature, (b) relative humidity, (c) air pressure, and (d) sunshine duration.

2.1.2. Satellite Products

The MODIS atmosphere and land products and the MTSAT data were used to derive the input parameters for the PAR models in this study. Atmospheric parameters, including liquid water and ice cloud optical depth (CPO), total column ozone amount (I_{oz}), total column precipitable water (w), liquid water path and ice water path (CWP), effective particle radius for liquid water clouds and ice water clouds (re), cloud fraction (TCP), solar zenith angle (θ), aerosol optical depth (aod) and surface pressure (PS) were derived from the MOD04/MYD04, MOD06/MYD06, and MOD07/MYD07. The ground albedo (α) was derived from the MOD09CMG and MYD09CMG. The top of atmosphere albedo (p_g) was derived from the MTSAT data. Detailed information for MODIS and MTSAT products are presented in Table 2.

Table 2. Basic information about Moderate Resolution Imaging Spectroradiometer (MODIS) and Multi-Functional Transport Satellite (MTSAT) products used in this study.

Data Set Name	Parameters	Spatial Resolution	Temporal Resolution
MTSAT (VIS)	Earth-atmospheric albedo	0.05 degree	Hourly
MOD04/MYD04	Aerosol optical depth (AOD)	5 km	Daily
MOD06/MYD06	Cloud phase optical thickness (CPO), solar zenith angle (θ), cloud water path (CWP), effective particle radius (re), cloud fraction (TCP)	1 km	Daily
MOD07/MYD07	Precipitable water vapor (w), total zone amount (I_{oz})	5 km	Daily
MOD09CMG/MYD09CMG	Surface albedo (ρ_g)	0.05 degree	Daily

2.1.3. Terrain and Climate Division Data

The climate and terrain regionalization data were provided by the Resource and Environment Science Data Center of Chinese Academy of Sciences (<http://www.resdc.cn>). Figures A1 and A2 shows the climate zones and terrain features in China. There were four humidity zones, ten temperate zones, and fifty topographic zones, over mainland China.

2.2. Photosynthetically-Active Radiation Models

2.2.1. Physically-Based Models

BBM

The BBM model is a physical-based broadband model, which was developed by Qin et al. [35], based on the clear-sky spectral transmittance parameterization. Considering the major radiative extinction processes between the surface and the atmosphere, the BBM model has been proved to be an efficient PAR model, with a high accuracy, at seven Surface Radiation Budget Network (SURFRAD) stations and seven hundred and sixteen CMA stations, which is expressed as following equations:

$$R_{all} = \tau_c R_{clr} \quad (1)$$

$$R_{clr} = R_b^{clr} + R_d^{clr} \quad (2)$$

$$\tau_c = 0.2495 + 1.1415r + 0.3910r^2 \quad (3)$$

where R_{all} is the daily PAR under all-sky conditions, and R_{clr} means the daily PAR under clear sky conditions. R_b^{clr} and R_d^{clr} are the beam and diffuse PAR under clear sky conditions, respectively. τ_c means the transmittances due to cloud scattering and absorption; r represents the relative sunshine duration. R_b^{clr} and R_d^{clr} can be calculated using following equations:

$$R_b^{clr} = R_0(d_0/d)^2(\sin\theta)\tau_b \quad (4)$$

$$R_d^{clr} = R_0(d_0/d)^2(\sin\theta)\tau_d \quad (5)$$

where θ is the solar zenith angle (degree); d_0/d is the eccentricity correction factor for the mean sun-earth distance; R_0 is the spectral irradiance (400–700 nm) at the mean distance between the earth and the sun in PAR band. τ_b is the beam transmittance in clear sky conditions; τ_d is the diffuse transmittance in clear sky conditions. τ_b and τ_d can be calculated as follows:

$$\tau_b \approx \tau_g \tau_R \tau_w \tau_o \tau_a \quad (6)$$

$$\tau_d \approx 0.5\tau_g \tau_w \tau_o (1 - \tau_R \tau_a) \quad (7)$$

where τ_g , τ_R , τ_w , τ_o , and τ_a are the transmittances for the mixed gasses absorption, Rayleigh scattering, water vapor absorption, ozone absorption, and the aerosol extinction, respectively.

Where τ_g , τ_R , τ_w , τ_o , and τ_a represent the transmittances for mixed gasses, Rayleigh transmittance, water vapor, ozone, and aerosol, respectively, which can be calculated from:

$$\tau_R = \exp[-0.14057(m')^{0.88384}] \quad (8)$$

$$\tau_w = \exp[-0.00021(mw)^{0.70991}] \quad (9)$$

$$\tau_o = \exp[-0.005218(ml)^{0.96054}] \quad (10)$$

$$\tau_a = \exp[-2.18157(m\beta)^{0.93988}] \quad (11)$$

$$\tau_g \approx 1 \quad (12)$$

where m represents the relative air mass; m' means the pressure-corrected relative air mass; w is the precipitable water vapor (cm); l is the ozone thickness (cm); and β is the Ångström turbidity coefficient.

PBM

This PBM model is a physical-based PAR model taking into consideration the physical relations between PAR and the earth-atmospheric albedo with absorption and scattering atmospheric constituents [22]. The instantaneous PAR at the Earth's surface in PBM, was obtained from:

$$I_{SUR} = \frac{I_{TOA}[1 - \rho_B - (1 - \tau_o) - \tau_o(\alpha_w + \alpha_{aer} + \alpha_g)]}{(1 - \rho_g) + \rho_g(\alpha_w + \alpha_{aer} + \alpha_g) + \rho_g(1 - \alpha_w - \alpha_{aer} - \alpha_g)(1 - \tau_o)} \quad (13)$$

where ρ_B is the earth-atmospheric albedo in the PAR band; ρ_g is the surface albedo; τ_o means the ozone transmittance; α_w , α_{aer} , and α_g denote the absorption coefficients of water vapor, aerosols, and mixed gasses, respectively; I_{TOA} is the extraterrestrial solar irradiation at the top of atmosphere in the PAR band, which could be calculated using following equation:

$$I_{TOA} = R_c [1 + 0.033 \cos(2\pi d_n / 365)] \cos\theta \quad (14)$$

where R_c is the solar constant in the PAR band; d_n means the day number since the first day of the year; θ_z is the solar zenith angle. The ozone transmittance τ_o , absorption of water vapor (α_w), and mixed gasses (α_g) were calculated, using

$$\tau_o = \frac{\int_{0.4\mu m}^{0.7\mu m} I_{o\lambda} \tau_{o\lambda} d\lambda}{\int_{0.4\mu m}^{0.7\mu m} I_{o\lambda} d\lambda} \quad (15)$$

$$a_w = 1 - \frac{\int_{0.4\mu m}^{0.7\mu m} I_{o\lambda} \tau_{w\lambda} d\lambda}{\int_{0.4\mu m}^{0.7\mu m} I_{w\lambda} d\lambda} \quad (16)$$

$$a_g = 1 - \frac{\int_{0.4\mu m}^{0.7\mu m} I_{o\lambda} \tau_{g\lambda} d\lambda}{\int_{0.4\mu m}^{0.7\mu m} I_{g\lambda} d\lambda} \quad (17)$$

where $I_{o\lambda}$ is the extraterrestrial solar irradiance; $\tau_{o\lambda}$, $\tau_{w\lambda}$, and $\tau_{g\lambda}$ are the spectral transmission coefficient for ozone, water vapor, and mixed gasses, respectively.

$$\tau_{o\lambda} = \exp(-k_{o\lambda} l m_r) \quad (18)$$

$$\tau_{g\lambda} = \exp[-1.41 k_{g\lambda} m_a / (1 + 118.93 k_{g\lambda} m_a)] \quad (19)$$

$$\tau_{w\lambda} = \exp[-0.2385 k_{w\lambda} m_r / (1 + 20.07 k_{w\lambda} w m_r)^{0.45}] \quad (20)$$

where $k_{o\lambda}$, $k_{w\lambda}$ and $k_{g\lambda}$ denotes the spectral extinction coefficient for ozone, water vapor, and mixed gasses, respectively; m_a represents the air mass; m_r represents the relative air mass; l means the total ozone amount (cm); and w is the precipitable water vapor (cm).

$$w = 0.8933 \exp(0.1715 \frac{rh p_s}{T}) \quad (21)$$

where rh is the relative humidity; p_s means the surface pressure (mbar); and T means the air temperature (K).

The absorption and scattering ($Daer$) of aerosol was calculated using Equation (23). α_{aer} was partitioned from $Daer$, the detailed descriptions for the estimation of α_{aer} could be found in Reference [40,41].

$$Daer = 1 - \frac{\int_{0.4\mu m}^{0.7\mu m} I_{o\lambda} \tau_{aer\lambda} d\lambda}{\int_{0.4\mu m}^{0.7\mu m} I_{o\lambda} d\lambda} \quad (22)$$

where $\tau_{aer\lambda}$ is aerosol transmission coefficient, calculated as

$$\tau_{aer\lambda} = \exp(-\beta\lambda^{-\alpha}m_a) \quad (23)$$

β is the Ångström turbidity coefficient and α denotes the wavelength exponent. β was calculated as follows:

$$\beta = 0.589 - 0.068VIS + 0.0019VIS^2 \quad (24)$$

where VIS is the visibility (km). It must be noted that Equation (25) was not correct in some situations, for example, when VIS was 15, 16, 17, 18, 19, 20, and 21 km, β was -0.0035 , -0.0126 , -0.0179 , -0.0194 , -0.0171 , -0.011 and -0.0011 , respectively. Thus, we introduced the formula below for the β estimation [42], which was expressed as follows

$$\beta = (0.025 + 0.1 \cos\theta)\exp(-0.7z/1000) \quad (25)$$

where θ is the latitude and z means the surface elevation.

EPP

This physical-based parameterization (EPP) was proposed by Tang et al. [8], based on the BBM and the cloud parameterization developed by Sun et al. [40]. PAR could be calculated using the following equation:

$$R_{all} = \frac{(1 - C_w - C_i)R_{clr} + C_w R_{wc} + C_i R_{ic}}{1 - \rho_{a,all}\rho_g} \quad (26)$$

where C_w and C_i are the cloud fractions for water cloud and ice cloud, respectively. R_{clr} , R_{wc} , and R_{ic} represent PAR in clear sky conditions, water cloudy sky conditions, and ice cloudy sky conditions, respectively; ρ_g is the surface albedo; $\rho_{a,all}$ denotes the atmospheric spherical albedo, which was calculated as follows

$$\rho_{a,all} = (1 - C_w - C_i)\rho_{a,clr} + C_w\rho_{a,wc} + C_i\rho_{a,ic} \quad (27)$$

where $\rho_{a,clr}$, $\rho_{a,wc}$, and $\rho_{a,ic}$ are the atmospheric spherical albedo for clear sky conditions, water cloudy sky conditions, and ice cloudy sky conditions, respectively. R_{clr} , R_{wc} , and R_{ic} could be calculated using following equations:

$$R_{clr} = R(\tau_b^{clr} + \tau_d^{clr}) \quad (28)$$

$$R = R_0(d_0/d)^2\mu \quad (29)$$

$$R_{wc} = R_{clr}\tau_{wc} \quad (30)$$

$$R_{ic} = R_{clr}\tau_{ic} \quad (31)$$

where τ_b and τ_d are beam and diffuse transmittance, respectively; τ_{wc} and τ_{ic} denote the global transmittance for water cloud and ice cloud, respectively.

LUT

The LUT method introduced in this study was developed by Zhang et al. [24]. First, the input parameters derived from the MODIS products were preprocessed (geometric correction, reprojection, and calibration). Second, the first look-up table was generated to connect atmospheric condition (visibility, cloud optical depth, water vapor amount, ozone amount, aerosol type, and cloud type, etc.)

to the top of the atmosphere radiance (I_{TOA}). Then, the second look-up table was generated to connect the atmospheric condition to PAR. At last, the surface PAR was calculated using the look-up Tables 1 and 2. In this study, GLASS (Global land surface satellite) PAR datasets generated by Zhang et al. [24] were used for the PAR validation, across China (<http://glass-product.bnu.edu.cn>).

2.2.2. The AI models

BP

The BP model is the most widely used AI models for estimating a solar radiation, with a strong learning ability and high accuracy [27]. The basic schematic architecture of the BP neural network was illustrated in Figure 3a. The BP model was formed by the input layer, the hidden layer, and the output layer. Each layer consisted of some neurons connected to each other. The basic idea of BP is to find a function that best maps a set of input parameters to the correct output values, using a gradient descent optimization algorithm, which minimizes the mean square error between the network’s actual output and the expected output values. In this study, six parameters (rh, at, ps, sd, a, d_n) that were closely correlated with the PAR values were set as input parameters for the BP model; daily PAR measurements were set as the model output parameter. The PAR values could be calculated using following equation:

$$F_g = Z \left(\sum_{i=1}^N w_i(t)x_i(t) + nb \right) \tag{32}$$

where F_g is the estimated PAR; $Z(\cdot)$ means the hidden transfer function; $w_i(t)$ means the weight; $x_i(t)$ means the input parameters indiscrete time space; and b means the neuronal bias.

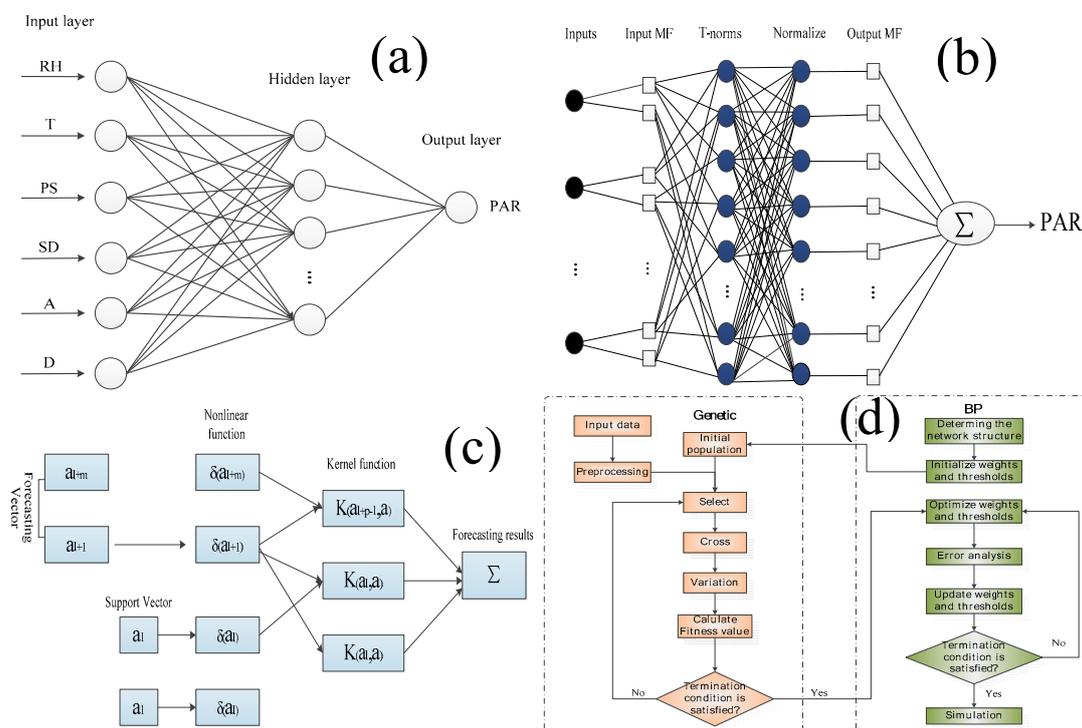


Figure 3. The general structures of the back propagation (BP) network, the adaptive neuro fuzzy inference system (ANFIS), the least squares support vector machine (LSSVM), and the Genetic algorithm model; (a) BP, (b) ANFIS, (c) LSSVM, and (d) Genetic.

ANFIS

The ANFIS is a hybrid intelligence system integrating the self-learning ability of the ANN and the reasoning ability of fuzzy logic [41]. The general structure of the ANFIS model is shown in Figure 3b. The ANFIS model establishes the appropriate membership function of the input and the output variables, by a set of fuzzy If-Then rules and forms output functions [42]. The rh , at , ps , sd , a , and d_n were the input parameters for the ANFIS, the measured PAR value was the output value for ANFIS. There were five layers for ANFIS, including fuzzification, rules, normalization, defuzzification, and summation in this study.

LSSVM

LSSVM is a powerful AI model for solving nonlinear regression evolved from the Support Vector Machine (SVM) [43]. The procedure of LSSVM in this study is shown in Figure 3c. Given a set of inputs a_i (meteorological parameters) and output y_i (PAR values), the LSSVM could reveal the nonlinear relationship between the input and the output values. The nonlinear function of LSSVM could be briefly expressed as:

$$f(a) = \omega^T \delta(a) + bt \quad (33)$$

where ω , δ , and bt are the m -dimensional weight vector, mapping function, and bias term, respectively. In this study, daily rh , at , ps , sd , a , d_n and PAR records at twenty-nine CERN stations were used for training and testing the LSSVM model. More detailed information about the LSSVM model could be found in Kisi [44].

Genetic

The Genetic algorithm is a heuristic algorithm inspired by the process of natural selection that belongs to the larger class of evolutionary algorithms [45]. The Genetic algorithm is commonly used to generate high-quality solutions to optimization and search problems. The Genetic model was used to improve the model accuracy for predicting the PAR values. The Genetic models for estimating PAR were conducted as the following setups (Figure 3d):

- (1) Initialize the random population: The basic structure of the BP neural network in this study was 6–10–1 (Figure 3a) with six input layers, ten hidden layers, and one output layer. Thus, the number of weights was $6 \times 10 + 10 \times 1 = 70$; the number of thresholds was $10 + 1 = 11$. So, the encoding length was $70 + 11 = 81$.
- (2) Selection operation: The new individuals with the high-fitness values would be selected from old individuals using a roulette selection method. The selection probability for individuals were calculated as the following equation:

$$g_i = a/a_i \quad (34)$$

$$P_i = g_i / \sum_{j=1}^n g_i \quad (35)$$

where P_i is the selection probability; g_i is the fitness value, which could be calculated as:

$$g_i = c1 \left(\sum_{j=1}^N \text{abs}(y_i - o_i) \right) \quad (36)$$

where N is the number of input layers of Genetic; y_i is the i -th expected output value; o_i is the i -th predicted output values.

- (3) Crossover operation: The crossover operation was conducted using the arithmetic crossover algorithm:

$$\left. \begin{aligned} a_{cj} &= a_{cj}(1-b) + a_{dj}c2 \\ a_{dj} &= a_{dj}(1-b) + a_{cj}c2 \end{aligned} \right\} \quad (37)$$

where a_{cj} and a_{dj} are the c -th and d -th chromosome at j position, respectively; $c2$ is a constant with the range of 0–1.

- (4) Mutation operation: The mutation operation was conducted using following equations:

$$a_{ij} = \begin{cases} a_{ij} + (a_{ij} - a_{max}) * f(g)r > 0.5 \\ a_{ij} + (a_{min} - a_{ij}) * f(g)r \leq 0.5 \end{cases} \quad (38)$$

$$f(g) = rr(1 - g/G_{max})^2 \quad (39)$$

where a_{max} and a_{min} are the maximum and minimum values for a_{ij} , respectively; r is a random number [0,1]; rr is the random number; g is the number of iterations; and G_{max} is maximum evolution times.

M5Tree

The M5Tree was first developed by Quinlan [46], based on a binary decision tree. The M5Tree could be used to reconstruct the quantitative relationship between the input and the output values. The M5Tree contains three steps [47,48]: (1) Splitting data into subsets to create decision trees; (2) generating the model tree; (3) building the linear regression model. In this study, the rh, at, ps, sd, a , d_n , and the PAR measurements were used for training and testing M5Tree model for predicting the PAR values.

MARS

MARS is a non-parametric regression technique, which could be used to predict the values of a continuous dependent or outcome variable from a set of independent or predictor variables, without any assumption about the underlying functional relationship between the dependent and the independent variables [49,50]. The MARS model for estimating PAR is given as:

$$Y = \alpha + \sum_{m=1}^M \beta_m h_m(X) \quad (40)$$

where Y is the estimated PAR values as a function of the input parameters (rh, at, ps, sd, a , d_n); β_m is the weight; $h_m(X)$ is the basis functions; X is the input parameters. Further details of MARS can be found in Sharda et al. [51].

2.3. Preprocesses for PAR Measurements

The equipment and operation-related errors would degrade the accuracy of the PAR measurements. In this study, the quality control process for the PAR measurements at the CERN stations was conducted following two principles: (1) The ratio between PAR ($\text{mol m}^{-2} \text{s}^{-1}$) and SSR ($\text{MJ m}^{-2} \text{day}^{-1}$) must be in the range of 1.3–2.8 mol MJ^{-1} ; (2) each measured PAR should not exceed the PAR at the top of the atmosphere (G_0), at the same geographical location. Moreover, the instantaneous PAR is typically expressed as the photon flux density ($\text{mol m}^{-2} \text{s}^{-1}$) [3,27]. Dye [52] considered that the ratio of energy flux density to the photosynthetic photon flux density was 1/4.57 (MJ mol^{-1}). Therefore, in our study, the unit of PAR measurement was unified to the energy flux density by multiplying 1/4.57.

2.4. The Statistical Indicators Representing Model Accuracy

In this study, a total of 70% of the database, during the whole study period, were used to train these PAR models, and the remaining datasets were used for testing these models. The model accuracies were validated using the following statistical indicators: The mean absolute bias error (MAE), the mean bias error (MBE), the root mean square error (RMSE), and the correlation coefficient (R):

$$RMSE = \sqrt{\left(\sum_{i=1}^N (P_{est,i} - P_{obs,i})^2\right) / N} \quad (41)$$

$$MAE = \left(\sum_{i=1}^N |P_{est,i} - P_{obs,i}|\right) / N \quad (42)$$

$$MBE = \left(\sum_{i=1}^N (P_{est,i} - P_{obs,i})\right) / N \quad (43)$$

$$R = \frac{(\sum_{i=1}^N (P_{est,i} - \overline{P_{est,i}})(P_{obs,i} - \overline{P_{obs,i}}))}{\sum_{i=1}^N (P_{est,i} - \overline{P_{est,i}}) \sum_{i=1}^N (P_{obs,i} - \overline{P_{obs,i}})} \quad (44)$$

where N is the sample number; P_{est} and P_{obs} represent the estimated and observed PAR, respectively; $\overline{P_{est}}$ and $\overline{P_{obs}}$ represent mean values of the estimated and the observed PAR, respectively.

3. Result and Discussion

3.1. Validation of Daily PAR Estimations at CERN Stations

The model accuracies of these ten PAR models were evaluated at twenty-nine CERN stations over mainland China. More than twenty-five thousand six hundred and sixty-six data samples were used for the training phases of six AI models. Figure 4 shows the model performance of these AI models, in the training phases. All the estimated PAR showed good agreements with the PAR measurements. Among these AI models, the Genetic model performed superior to the other AI models. Then, another eleven thousand data samples were used to validate the model performance of all ten PAR models that were used in this study. Figure 5 illustrates the statistical indicators representing model accuracies of all PAR models. All methods produced PAR estimates that positively correlated with the measurements at the CERN stations. The R values for the BBM, EPP, PBM, LUT, BP, ANFIS, M5Tree, Genetic, MARS, and LSSVM were 0.947, 0.872, 0.900, 0.787, 0.955, 0.970, 0.967, 0.987, 0.955, and 0.961, respectively. All AI models (BP, ANFIS, M5Tree, Genetic, MARS, and LSSVM) showed overwhelming superiority than the BBM, EPP, PBM, LUT, owing to their strong learning. The RMSE for the BBM, EPP, PBM, LUT, BP, ANFIS, M5Tree, Genetic, MARS, and LSSVM were 1.175, 1.565, 2.593, 1.975, 0.912, 0.748, 0.799, 0.5, 0.907, and 0.903, respectively; The MAE for the BBM, EPP, PBM, LUT, BP, ANFIS, M5Tree, Genetic, MARS, and LSSVM were 0.976, 1.310, 2.202, 1.463, 0.731, 0.557, 0.595, 0.326, 0.743, and 0.643, respectively. The Genetic model showed better performances than the other PAR models, with the highest R (0.973) and the lowest RMSE ($0.5 \text{ MJ m}^{-2}\text{day}^{-1}$) and MAE ($0.326 \text{ MJ m}^{-2}\text{day}^{-1}$), due to the optimization of the weight and threshold of the neural network. Affected by the cloud cover and changing weather conditions, parameters derived from MODIS level 2 products and the MTSAT data for the EPP were not always accessible at all CERN stations. Thus, the PBM yielded the largest errors with the highest RMSE ($2.593 \text{ MJ m}^{-2}\text{day}^{-1}$) and MAE ($2.202 \text{ MJ m}^{-2}\text{day}^{-1}$).

Taylor diagrams were introduced to visualize the model accuracies. Figure 6 shows the Taylor diagrams used to visualize the model accuracies for all PAR models in each month of the year, respectively. It was clear that the model accuracies for all models were subject to seasonal climatic characteristics. The physical models (BBM, EPP, PBM and LUT) showed better performances in winter than those in summer, due to the abundant water vapor and large cloud cover in summer.

The smallest RMSE for all physical-based models were in January, the smallest MAE were also in January; the largest RMSE for all physical-based models were also in June, the largest MAE were in June. In contrast, the estimated PAR by the AI models (BP, ANFIS, M5Tree, MARS, Genetic and LSSVM) showed better agreements with the PAR measurements in each month, due to their strong adaptability to the fluctuation of input parameters. The model performances for the Genetic model, throughout the year, were more stable than the other models, the largest RMSE ($0.679 \text{ MJ m}^{-2}\text{day}^{-1}$) and MAE ($0.460 \text{ MJ m}^{-2}\text{day}^{-1}$) for the Genetic were found in April; the smallest RMSE ($0.239 \text{ MJ m}^{-2}\text{day}^{-1}$) and MAE ($0.153 \text{ MJ m}^{-2}\text{day}^{-1}$) were found in July. PBM was not as accurate as other PAR models for estimating PAR values throughout a year, with distinct seasonal variations, and had high RMSE and MAE values, in each month.

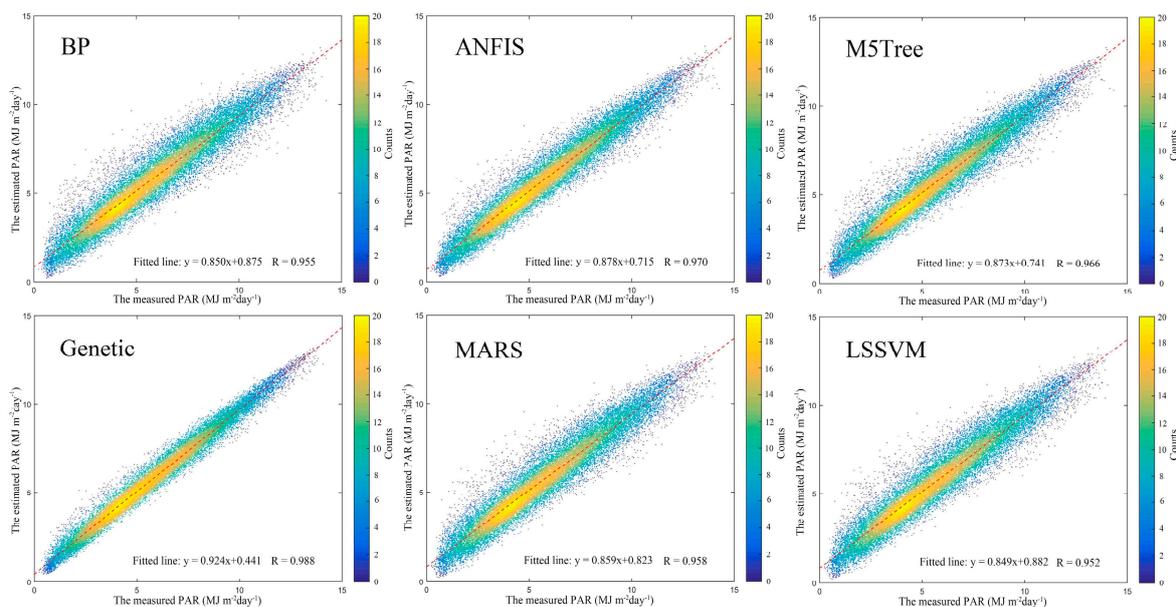


Figure 4. Validation of the estimated photosynthetically-active radiation (PAR) by the Artificial Intelligence (AI) models in the training state, at the CERN stations.

Figure 7 illustrated the spatial distributions of the mean values of the statistical indicators representing model accuracies for all PAR models at CERN stations. Relatively larger errors were mainly observed in the plateau zones, due to the strong heating atmosphere there, for example, the mean RMSE, MAE, R and MBE for all PAR models at Lhasha (LSA), in the Qinghai-Tibetan Plateau, were $1.456 \text{ MJ m}^{-2}\text{day}^{-1}$, $1.042 \text{ MJ m}^{-2}\text{day}^{-1}$, 0.804 , and $-0.528 \text{ MJ m}^{-2}\text{day}^{-1}$, respectively; the mean RMSE, MAE, R and MBE for all PAR models at Gongga (GGF), in the Qinghai-Tibetan Plateau, were $1.730 \text{ MJ m}^{-2}\text{day}^{-1}$, $1.194 \text{ MJ m}^{-2}\text{day}^{-1}$, 0.800 , and $-0.714 \text{ MJ m}^{-2}\text{day}^{-1}$, respectively; the mean RMSE, MAE, R and MBE for all PAR models at Ailao (ALF), in the Yunnan-Guizhou Plateau, were $1.458 \text{ MJ m}^{-2}\text{day}^{-1}$, $0.966 \text{ MJ m}^{-2}\text{day}^{-1}$, 0.851 , and $-0.549 \text{ MJ m}^{-2}\text{day}^{-1}$, respectively. The model deviations were also large in the tropical zones, for example, the mean RMSE, MAE, R and MBE for all PAR models at Xishuangbanna (BNF) (tropical zone) were $1.427 \text{ MJ m}^{-2}\text{day}^{-1}$, $0.954 \text{ MJ m}^{-2}\text{day}^{-1}$, 0.826 , and $-0.465 \text{ MJ m}^{-2}\text{day}^{-1}$, respectively. In contrast, owing to the dry air conditions, relatively lower estimation errors were observed in Northwestern China, for example, the RMSE, MAE, R and MBE for all PAR models at AKA, in the Tarim Desert, were $1.132 \text{ MJ m}^{-2}\text{day}^{-1}$, $0.806 \text{ MJ m}^{-2}\text{day}^{-1}$, 0.937 , and $-0.259 \text{ MJ m}^{-2}\text{day}^{-1}$, respectively; the RMSE, MAE, R and MBE for all PAR models at Cele (CLD), in the Tarim Desert, were $1.206 \text{ MJ m}^{-2}\text{day}^{-1}$, $0.870 \text{ MJ m}^{-2}\text{day}^{-1}$, 0.918 , and $-0.284 \text{ MJ m}^{-2}\text{day}^{-1}$, respectively. The smallest model deviations were observed in Haibei (HBG) and Shapotou (SPD), in the Alashan and the Hexi corridor; the RMSE, MAE, R and MBE, for the HBG, were $1.152 \text{ MJ m}^{-2}\text{day}^{-1}$, $0.881 \text{ MJ m}^{-2}\text{day}^{-1}$, 0.909 and $-0.211 \text{ MJ m}^{-2}\text{day}^{-1}$, respectively;

the RMSE, MAE, R and MBE, for the SPD, were $1.288 \text{ MJ m}^{-2}\text{day}^{-1}$, $0.917 \text{ MJ m}^{-2}\text{day}^{-1}$, 0.924 , and $-0.250 \text{ MJ m}^{-2}\text{day}^{-1}$, respectively.

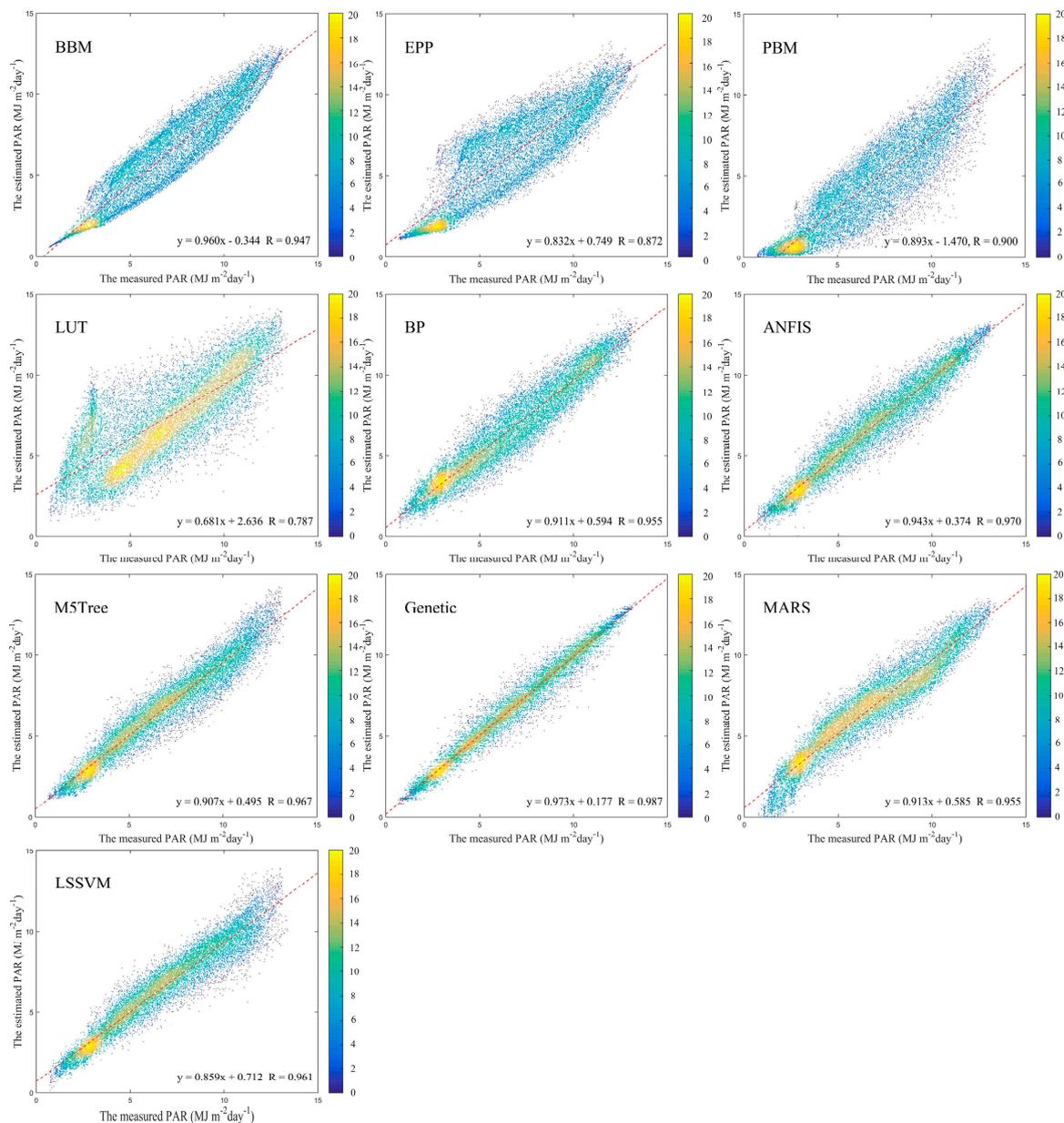


Figure 5. Validation of the PAR models at the CERN stations.

Figures 8 and 9 showed the statistical indicators representing the model accuracies for all PAR models, at different CERN stations. All AI models showed better model accuracies than the BBM, EPP, PBM and LUT, owing to their strong adaptability to the fluctuations of the input parameters. The ranges of RMSE for the BBM, EPP, PBM, LUT, BP, ANFIS, M5Tree, Genetic, MARS, and LSSVM at the CERN stations were $1.027\sim 1.625$, $1.186\sim 2.177$, $1.938\sim 4.074$, $1.168\sim 2.563$, $0.675\sim 1.120$, $0.504\sim 0.978$, $0.515\sim 1.089$, $0.324\sim 0.664$, $0.690\sim 1.091$, and $0.455\sim 1.299 \text{ MJ m}^{-2}\text{day}^{-1}$, respectively. The ranges of MAE for the BBM, EPP, PBM, LUT, BP, ANFIS, M5Tree, Genetic, MARS, and LSSVM at the CERN stations were $0.819\sim 1.434$, $0.992\sim 1.836$, $1.504\sim 3.822$, $0.933\sim 1.996$, $0.540\sim 0.904$, $0.382\sim 0.743$, $0.392\sim 0.831$, $0.215\sim 0.448$, $0.571\sim 0.903$, and $0.339\sim 0.928 \text{ MJ m}^{-2}\text{day}^{-1}$, respectively. The ranges of R for the BBM, EPP, PBM, LUT, BP, ANFIS, M5Tree, Genetic, MARS, and LSSVM at the CERN stations were $0.851\sim 0.978$, $0.662\sim 0.961$, $0.705\sim 0.947$, $0.302\sim 0.919$, $0.924\sim 0.971$, $0.925\sim 0.982$, $0.930\sim 0.981$, $0.969\sim 0.992$,

0.893~0.974, and 0.895~0.985, respectively. The ranges of MBE for the BBM, EPP, PBM, LUT, BP, ANFIS, M5Tree, Genetic, MARS, and SVM at the CERN stations were $-1.434\sim0.220$, $-1.694\sim1.045$, $-3.822\sim-1.382$, $-0.539\sim1.701$, $-0.194\sim0.409$, $-0.309\sim0.259$, $-0.565\sim0.122$, $-0.048\sim0.06$, $-0.4\sim0.288$, and $-0.642\sim-0.02\text{ MJ m}^{-2}\text{ day}^{-1}$, respectively.

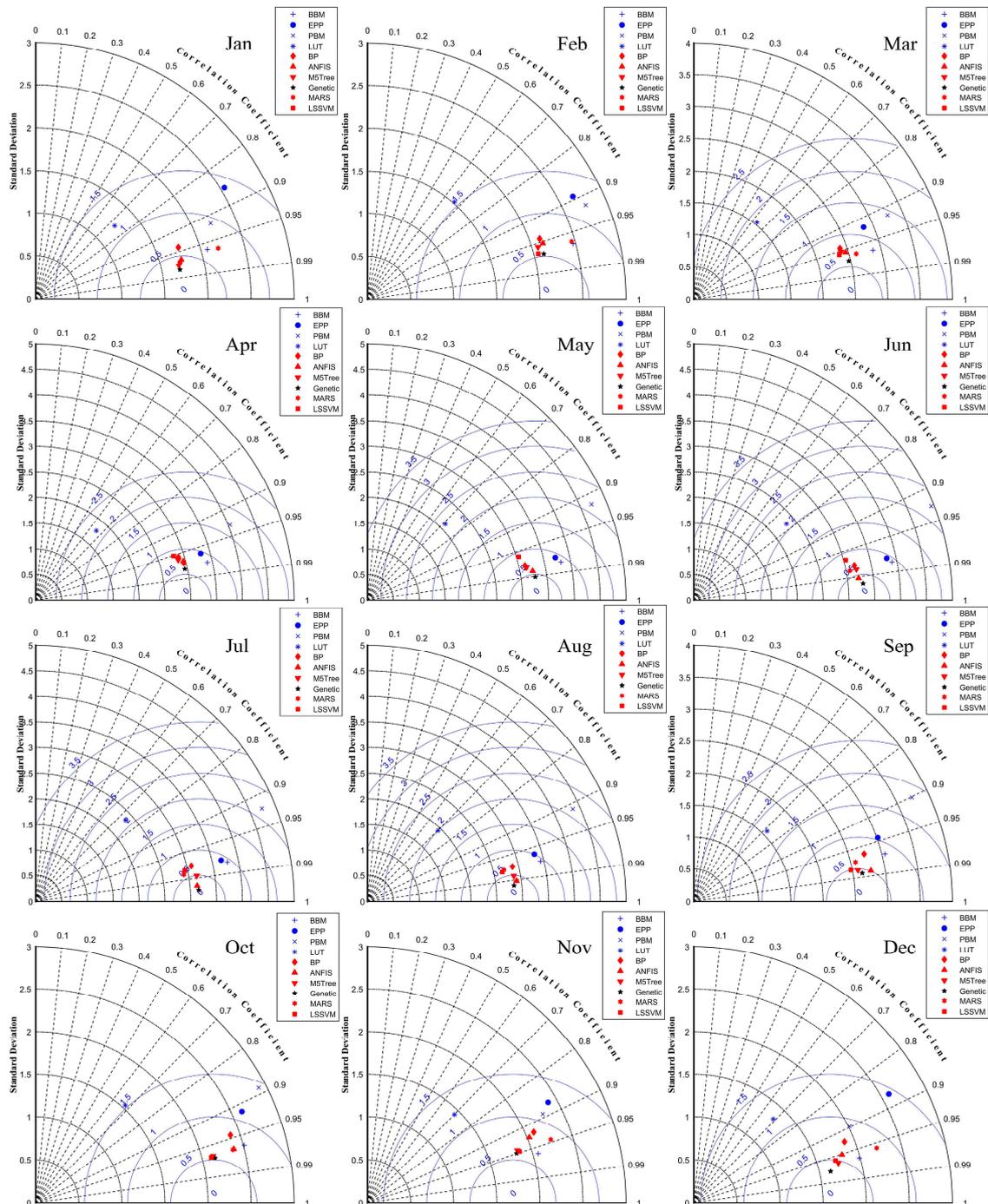


Figure 6. Monthly variations of the statistical indicators representing the model accuracy.

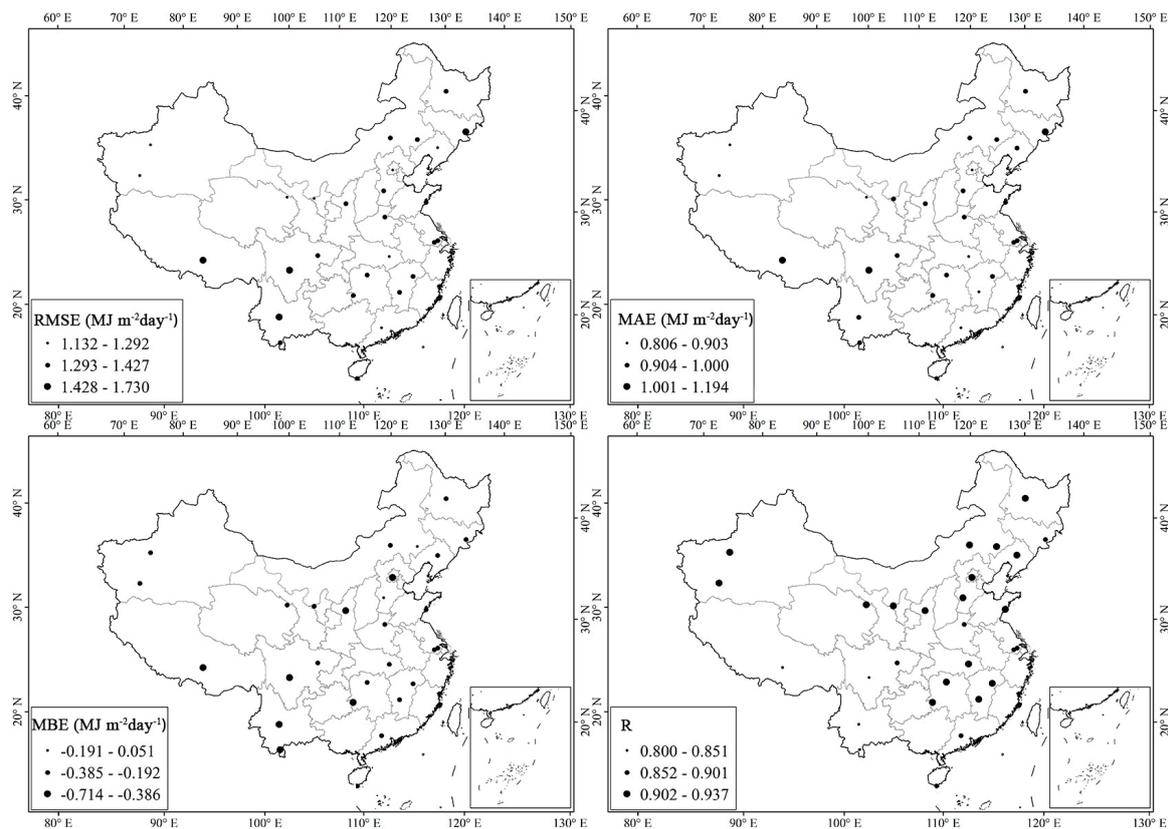


Figure 7. Spatial variations of the statistical indicators representing the model accuracy for ten PAR models.

The Genetic performed superior to other models in all CERN stations, due to its optimized weight and threshold for the neural network. The ranges of the RMSE, MAE, R, and MBE for the Genetic, at all CERN stations, were $0.324\sim 0.664 \text{ MJ m}^{-2}\text{day}^{-1}$, $0.215\sim 0.448 \text{ MJ m}^{-2}\text{day}^{-1}$, $0.969\sim 0.992$ and $-0.048\sim 0.060 \text{ MJ m}^{-2}\text{day}^{-1}$, respectively. The largest model deviations for the Genetic were found in the Changbaishan (CBF) with RMSE, MAE, R and MBE being $0.664 \text{ MJ m}^{-2}\text{day}^{-1}$, $0.448 \text{ MJ m}^{-2}\text{day}^{-1}$, 0.977 and $-0.039 \text{ MJ m}^{-2}\text{day}^{-1}$, respectively; The smallest model deviations for the Genetic were found in the ALF, with RMSE, MAE, R and MBE being $0.324 \text{ MJ m}^{-2}\text{day}^{-1}$, $0.215 \text{ MJ m}^{-2}\text{day}^{-1}$, 0.991 , and $0.020 \text{ MJ m}^{-2}\text{day}^{-1}$, respectively. In contrast, PBM was not accurate as the other models, at all CERN stations. The ranges of the RMSE, MAE, R, and MBE for the PBM, at all CERN stations, were $1.938\sim 4.074 \text{ MJ m}^{-2}\text{day}^{-1}$, $1.504\sim 3.822 \text{ MJ m}^{-2}\text{day}^{-1}$, $0.705\sim 0.947$, and $-3.822\sim -1.382 \text{ MJ m}^{-2}\text{day}^{-1}$, respectively. The largest model deviations for the PBM were found in the GGF with RMSE, MAE, R, and MBE, being $4.074 \text{ MJ m}^{-2}\text{day}^{-1}$, $3.822 \text{ MJ m}^{-2}\text{day}^{-1}$, 0.820 and $-3.822 \text{ MJ m}^{-2}\text{day}^{-1}$, respectively. The smallest model deviations for the PBM were found in the CLD with RMSE, MAE, R and MBE being $1.938 \text{ MJ m}^{-2}\text{day}^{-1}$, $1.527 \text{ MJ m}^{-2}\text{day}^{-1}$, 0.909 , and $-1.479 \text{ MJ m}^{-2}\text{day}^{-1}$, respectively.

In all, the AI models were more accurate and stable than the BBM, EPP, PBM and LUT. BBM, EPP, PBM and LUT were more susceptible to weather conditions than the AI models, which may be attributed to the uncertainties of satellite signals caused by cloud cover and precipitable water vapor. Compared with the AI models, larger spatial and temporal variations of statistical indicators were observed for the BBM, EPP, PBM and LUT. The Genetic showed better accuracies and robustness than the other PAR models, at all selected CERN stations, without significant seasonal variations, due to its strong learning ability and optimized weight and thresholds, for the neural network.

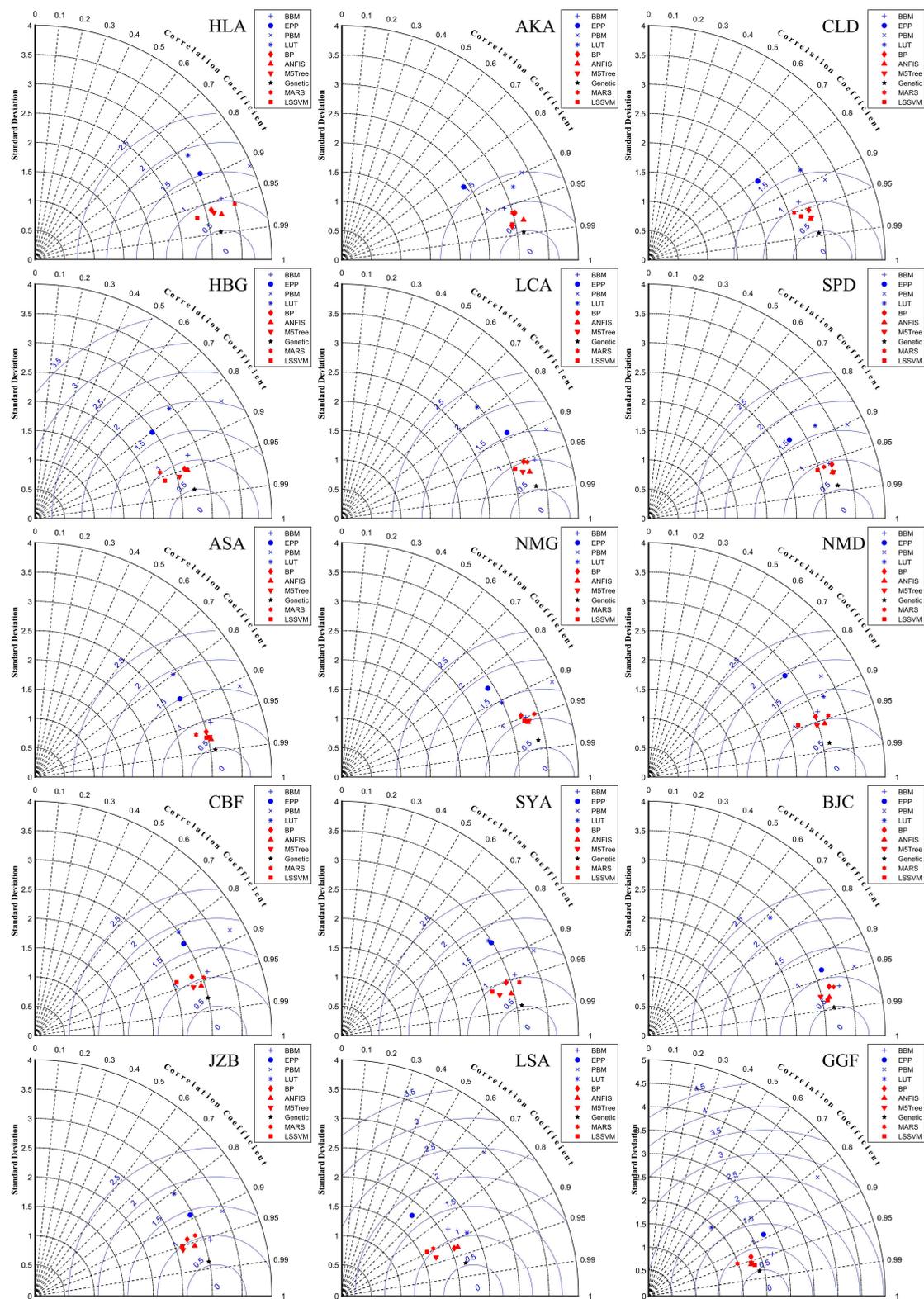


Figure 8. The statistical indicators representing the model accuracy in different stations.

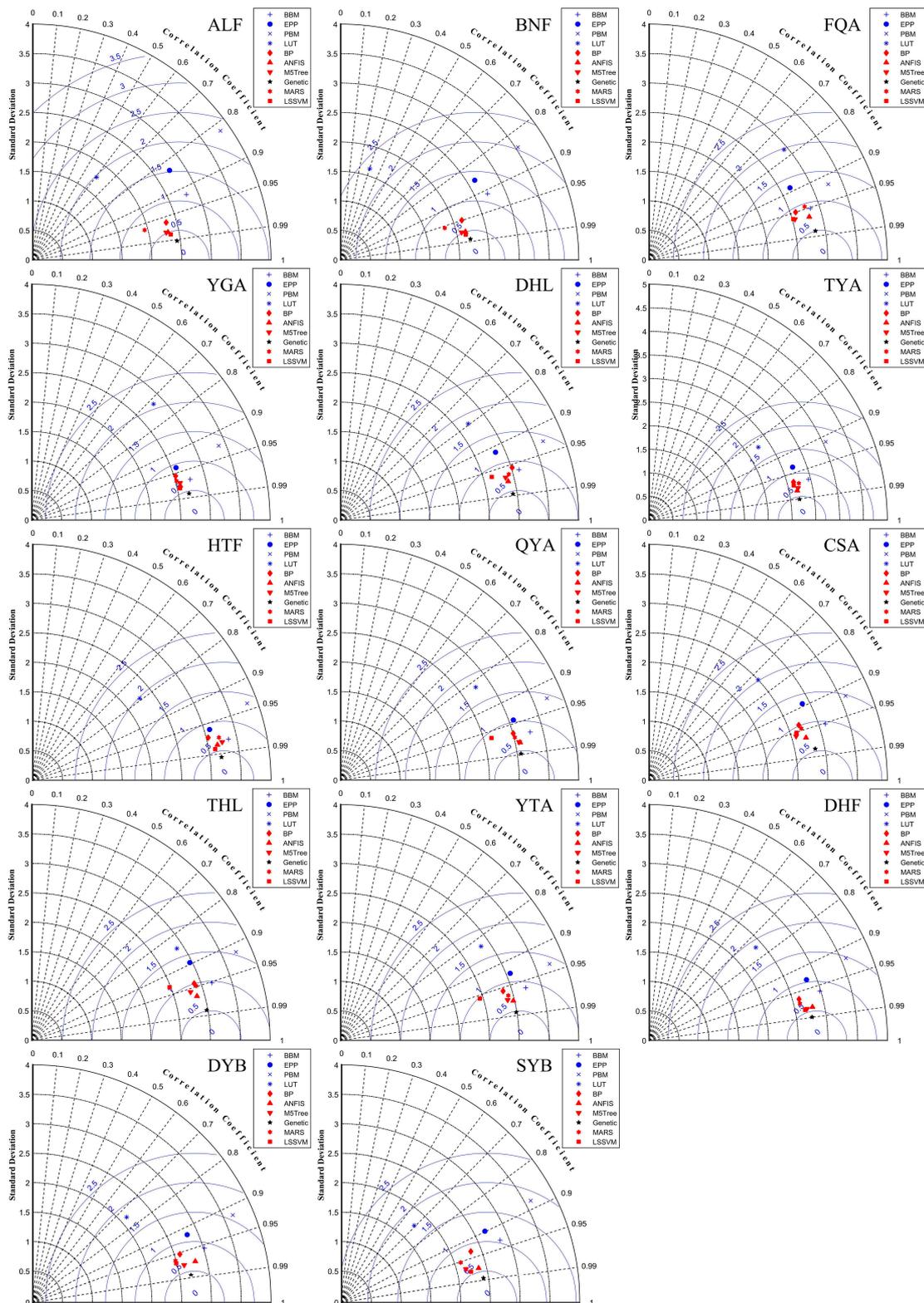


Figure 9. The statistical indicators representing the model accuracy in different stations.

3.2. Validation of PAR Models in Various Climate Zones and Terrains

Many radiation extinction processes occur when solar radiation passes through the atmosphere and is eventually reflected back to space. These extinction processes would vary with time and locations. Temperature was directly proportional to surface solar radiation, without radiation-damping

processes in the atmosphere. Table 3 showed the statistical errors for all PAR models in different temperate zones over China, the largest errors were found in temperature zone in the plateau (HII), due to the strong heating atmosphere. The mean RMSE and MAE for all models in the HII were 1.25 and 0.994 MJ m⁻²day⁻¹, respectively. It was clear that the Genetic performed superior to other PAR models, in different temperate zones, the largest RMSE (0.540 MJ m⁻²day⁻¹) and MAE (0.368 MJ m⁻²day⁻¹) for the Genetic were observed in the HII, while the smallest RMSE (0.324 MJ m⁻²day⁻¹) and MAE (0.215 MJ m⁻²day⁻¹) were in VII. PBM was not as accurate as the other models, in most temperate zones.

Table 3. The root mean square error (RMSE) and mean absolute bias error (MAE) for the PAR models in different temperate zones.

Statistics	Models	HII	II	IIE	III	IV	V	VI	VII
RMSE	BBM	1.178	1.112	1.186	1.116	1.189	1.201	1.266	1.453
	EPP	1.666	1.466	1.369	1.627	1.500	1.568	1.555	1.832
	PBM	2.776	2.454	2.814	2.263	2.728	2.748	2.780	3.141
	LUT	1.820	2.095	1.857	1.854	2.013	1.922	2.195	2.007
	BP	0.918	0.963	0.876	0.973	0.823	0.873	0.902	0.675
	ANFIS	0.830	0.769	0.567	0.806	0.711	0.708	0.719	0.507
	M5Tree	0.817	0.819	0.663	0.858	0.828	0.743	0.754	0.515
	Genetic	0.540	0.520	0.377	0.519	0.488	0.479	0.513	0.324
	MARS	0.988	0.923	0.780	0.959	0.861	0.855	0.897	0.770
	SVM	0.969	0.861	0.579	0.998	0.958	0.878	0.849	0.458
	Mean	1.250	1.198	1.107	1.197	1.210	1.197	1.243	1.168
MAE	BBM	0.946	0.928	1.041	0.935	0.992	0.989	1.056	1.226
	EPP	1.460	1.220	1.198	1.348	1.266	1.308	1.284	1.556
	PBM	2.319	2.100	2.544	1.868	2.365	2.334	2.447	2.729
	LUT	1.319	1.533	1.430	1.336	1.495	1.430	1.740	1.605
	BP	0.724	0.769	0.735	0.789	0.653	0.710	0.710	0.540
	ANFIS	0.642	0.558	0.435	0.609	0.524	0.538	0.526	0.394
	M5Tree	0.617	0.602	0.518	0.648	0.604	0.564	0.557	0.392
	Genetic	0.368	0.333	0.251	0.339	0.317	0.320	0.330	0.215
	MARS	0.823	0.749	0.654	0.787	0.706	0.703	0.727	0.663
	SVM	0.723	0.639	0.435	0.716	0.669	0.630	0.569	0.343
	Mean	0.994	0.943	0.924	0.938	0.959	0.953	0.994	0.966

II for mid temperate; III for warm temperate; IV for north subtropical zone; V for the mid-subtropics; VI for the south subtropics; VII for the edge of tropical zone; HII for temperature zone in plateau; IIE for mid tropical zone with humid weather. The unit for RMSE and MAE: MJ m⁻²day⁻¹.

The underlying surface properties would have significant effects on the accuracy of the PAR estimations. In this study, seven types of underlying surfaces, including grassland, city, lake, desert, farmland, forest, and water were considered to reveal the effects of the underlying surfaces on the model accuracies. Table 4 shows the RMSE and MAE, in different underlying surfaces, for all PAR models. Wetland was a land area that permanently or seasonally saturated with water, thus, the radiation processes in wetlands were more complicated than the other ecosystems, which made it more difficult to estimate PAR [43]. The mean RMSE and MAE in wetlands were 1.306 and 1.043 MJ m⁻²day⁻¹, respectively. The surfaces in the city had considerable influences on PAR balances, thus, modeling PAR in the city was also complicated. The mean RMSE and MAE in the city areas were 1.229 and 0.996 MJ m⁻²day⁻¹, respectively. It was obvious that the Genetic model showed much higher accuracy than the other models in all underlying surfaces, the RMSE for the Genetic in wetland, desert, lake, forest, farmland, city, and grassland were 0.571, 0.491, 0.485, 0.546, 0.500, 0.461 and 0.474 MJ m⁻²day⁻¹, respectively. The MAE for the Genetic in the wetland, desert, lake, forest, farmland, city, and grassland were 0.384, 0.329, 0.312, 0.369, 0.324, 0.297 and 0.305 MJ m⁻²day⁻¹, respectively. The PBM showed the poorest model performances in all underlying surfaces, the RMSE for the Genetic in the wetland, desert, lake, forest, farmland, city, and grassland were 2.868, 2.133,

2.433, 2.171, 2.487, 3.032 and 2.554 MJ m⁻²day⁻¹, respectively; the MAE for the Genetic in the wetland, desert, lake, forest, farmland, city, and grassland were 2.437, 1.866, 2.131, 1.668, 2.111, 2.674 and 2.243 MJ m⁻²day⁻¹, respectively. In all, the underlying surface properties were indispensable factors influencing the model accuracies, especially for the grassland and desert.

Table 4. The RMSE and MAE for the PAR models in different underlying surface.

Statistics	Models	Wetland	Desert	Lake	Forest	Farmland	City	Grassland
RMSE	BBM	1.239	1.058	1.063	1.148	1.124	1.341	1.116
	EPP	1.757	1.226	1.352	1.852	1.492	1.647	1.378
	PBM	2.868	2.133	2.433	2.171	2.487	3.032	2.554
	LUT	1.752	2.354	2.027	1.562	2.038	2.100	1.937
	BP	1.002	0.875	0.952	0.981	0.900	0.830	0.937
	ANFIS	0.912	0.743	0.735	0.839	0.734	0.655	0.739
	M5Tree	0.908	0.860	0.835	0.842	0.813	0.677	0.786
	Genetic	0.571	0.491	0.485	0.546	0.500	0.461	0.474
	MARS	1.062	0.884	0.882	0.988	0.894	0.835	0.897
	SVM	0.993	0.639	1.115	1.000	0.933	0.711	0.820
	Mean	1.306	1.126	1.188	1.193	1.191	1.229	1.164
MAE	BBM	1.011	0.872	0.879	0.943	0.935	1.131	0.950
	EPP	1.535	1.004	1.122	1.582	1.251	1.365	1.166
	PBM	2.437	1.866	2.131	1.668	2.111	2.674	2.243
	LUT	1.234	1.822	1.437	1.090	1.486	1.683	1.489
	BP	0.789	0.721	0.780	0.773	0.721	0.661	0.769
	ANFIS	0.706	0.556	0.541	0.639	0.548	0.477	0.553
	M5Tree	0.684	0.651	0.630	0.645	0.604	0.497	0.587
	Genetic	0.384	0.329	0.312	0.369	0.324	0.297	0.305
	MARS	0.878	0.724	0.726	0.803	0.733	0.688	0.742
	SVM	0.770	0.487	0.786	0.723	0.672	0.488	0.589
	Mean	1.043	0.903	0.934	0.923	0.939	0.996	0.939

The unit for RMSE and MAE: MJ m⁻²day⁻¹.

3.3. Spatial and Temporal Variations of PAR in China

The annual and monthly mean PAR, during 1955–2015, were calculated to reveal the spatial and temporal variations of PAR across China, based on the Genetic model, using meteorological measurements at eight hundred and thirty-nine CMA stations. Figure 10 illustrates the mean PAR values during 1955–2015, the annual PAR values presents a clear decreasing trend at the rate of -0.003 MJ m⁻²day⁻¹/year during 1955–2015. Figure 11 shows the spatial distributions of the annual mean estimated PAR (APAR) over mainland China. Generally, the PAR was higher in Western China than that in the Southern and Northeastern China, the Tibetan Plateau has always been an area with the highest PAR, over mainland China, due to the small atmospheric extinction effects, the maximum APAR was about 8.668 MJ m⁻²day⁻¹ in the Tibetan Plateau. In contrast, the Sichuan Basin in Southern China had always been an area with the lowest PAR, due to the perennial cloudy weather and strong atmospheric extinctions [34], the annual mean PAR was about 4.733 MJ m⁻²day⁻¹ in the Sichuan Basin. The Northeastern China was also an area with low PAR, owing to the relatively short sunshine durations and humid weather there. Figure 12 illustrates the monthly variation of PAR over mainland China, PAR values were generally higher in the summer than that in other seasons, because of higher solar zenith angle and longer sunshine duration, in the summer than that in other seasons. The monthly mean PAR values from January to December were 3.676, 4.697, 6.112, 7.610, 8.570, 8.754, 8.795, 8.281, 6.965, 5.523, 4.202, and 3.468 MJ m⁻²day⁻¹, respectively. The Qinghai Tibetan plateau has always been an area with the highest monthly mean PAR values throughout the year. The largest monthly mean PAR values for the Qinghai Tibetan plateau from January to December were 6.718, 7.975, 9.482, 10.893, 11.917, 12.550, 11.580, 10.727, 9.877, 8.762, 7.261, and 6.373 MJ m⁻²day⁻¹, respectively. In contrast,

the Sichuan Basin has always been an area with the lowest monthly mean PAR values. The largest monthly mean PAR values in the Sichuan Basin from January to December were 1.359, 2.619, 3.616, 4.645, 5.495, 5.943, 5.516, 6.368, 4.058, 3.129, 1.656, and 0.991 MJ m⁻²day⁻¹, respectively.

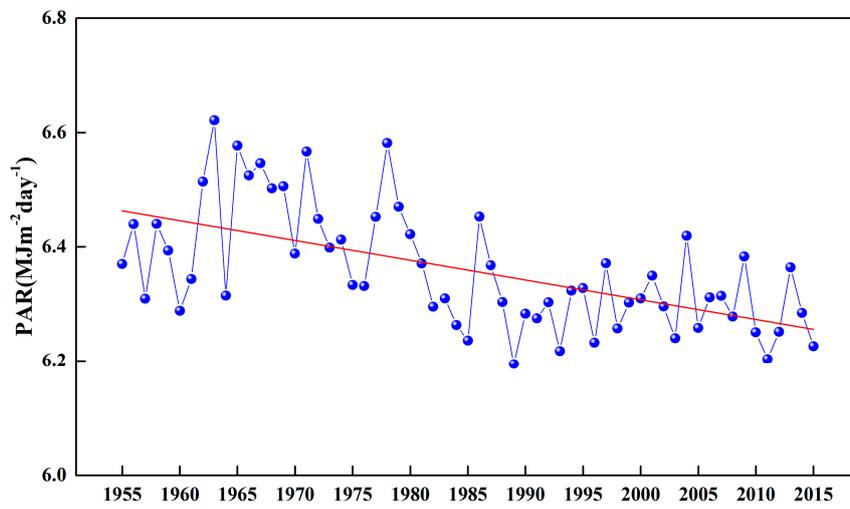


Figure 10. Annual mean PAR over China, during 1955–2015 (red line is the trend line).

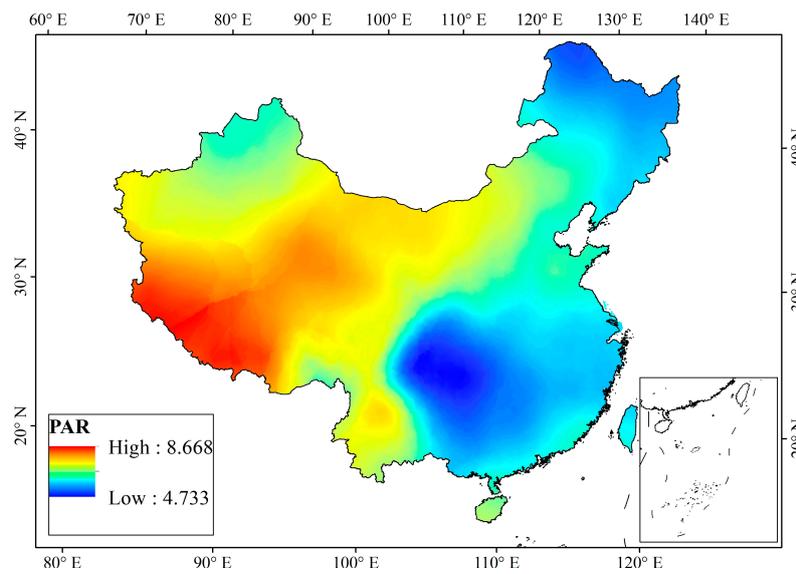


Figure 11. Annual mean PAR (MJ m⁻²day⁻¹) over mainland China.

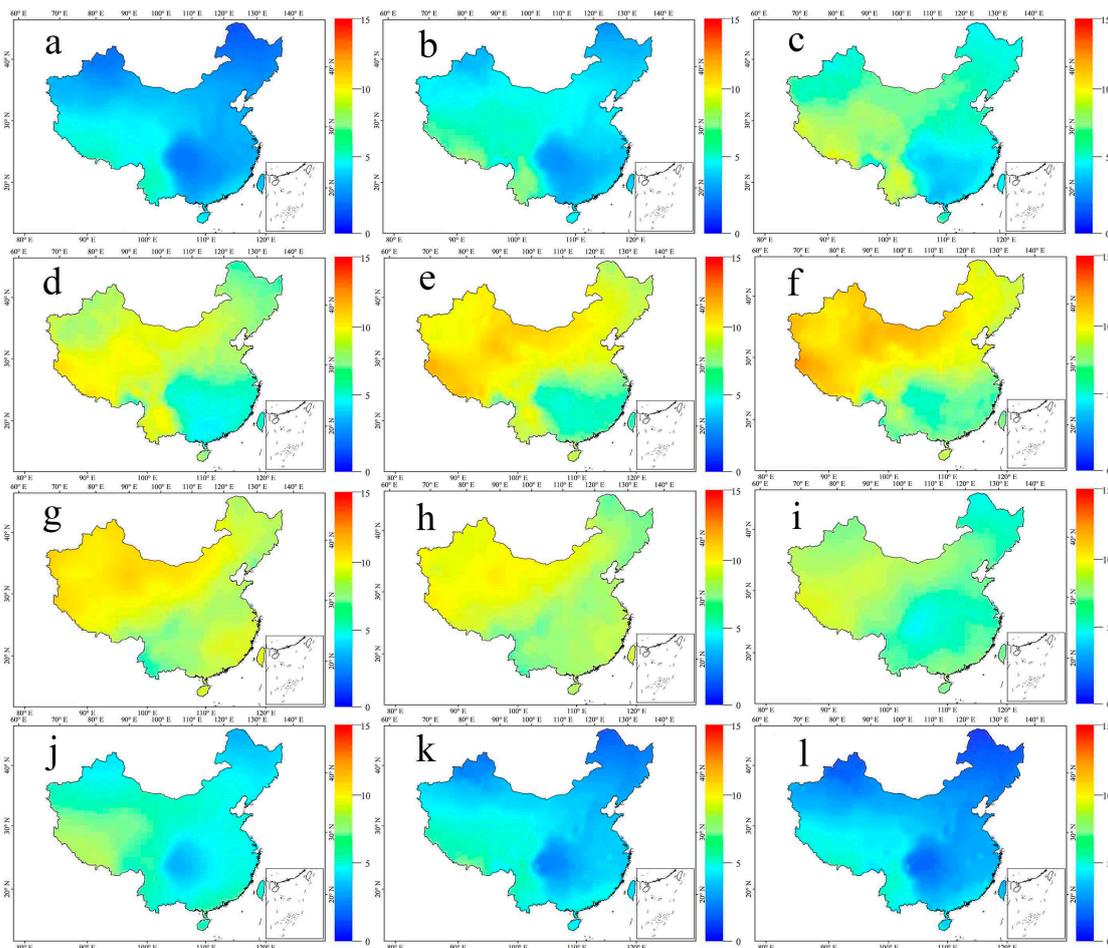


Figure 12. Spatial and temporal variations of PAR ($\text{MJ m}^{-2}\text{day}^{-1}$) over mainland China; (a) January, (b) February, (c) March, (d) April, (e) May, (f) June, (g) July, (h) August, (i) September, (j) October, (k) November, and (l) December.

4. Conclusions

The aim of this research was to make a comparative study on the model accuracies of ten models for estimating PAR over mainland China. The performances of the Genetic model, together with other nine PAR models, were evaluated in different climate zones and terrain features, using long-term continuous meteorological and radiation measurements at twenty-nine CERN stations and satellite signals. The spatial and temporal variations of PAR values, during 1955–2015, over mainland China, were further investigated.

Generally, the AI models showed better performances than the BBM, EPP, PBM and LUT. Among all PAR models, the Genetic performed superior to the other PAR models at all CERN stations in terms of RMSE, MAE and R. The model performances for Genetic were more stable than the other PAR models, throughout the year, without large monthly variations. The largest RMSE ($0.679 \text{ MJ m}^{-2}\text{day}^{-1}$) and MAE ($0.460 \text{ MJ m}^{-2}\text{day}^{-1}$) for the Genetic were found in April; the smallest RMSE ($0.0239 \text{ MJ m}^{-2}\text{day}^{-1}$) and MAE ($0.153 \text{ MJ m}^{-2}\text{day}^{-1}$) were found in July.

Meanwhile, the climate and terrain effects on the PAR estimation for all PAR models were investigated. PAR models showed different performances in different ecosystems, the largest mean RMSE ($1.306 \text{ MJ m}^{-2}\text{day}^{-1}$) and MAE ($1.043 \text{ MJ m}^{-2}\text{day}^{-1}$) for all PAR models were found in wetlands, due to the complicated radiation processes. The model deviations were also larger in city ecosystems with RMSE and MAE of 1.229 and $0.996 \text{ MJ m}^{-2}\text{day}^{-1}$, respectively. In all, the Genetic model performed better than the other PAR models, with strong robustness in different climate zones and terrain features.

The spatial and temporal variations of the annual mean PAR (APAR) value over mainland China were further analyzed. The Tibetan Plateau had always been an area with the highest APAR ($8.668 \text{ MJ m}^{-2}\text{day}^{-1}$). In contrast, the Sichuan Basin had always been an area with the lowest APAR values ($4.733 \text{ MJ m}^{-2}\text{day}^{-1}$), across China. The PAR values in China were generally higher in the summer than the other seasons.

Certainly, these PAR models should be tested and applied in other climate zones, terrain features, and ecosystems, around the world. Additionally, the relationships between the PAR and natural characteristics (such as climate patterns and terrain features) and social economic conditions (economic development and population density), should be quantitatively analyzed in future.

Author Contributions: L.F., W.Q. and L.W. designed the research; W.Q., M.Z. and L.F. performed the experiments and analyzed the data; L.F. and W.Q. wrote the manuscript; L.W. and A.L. revised the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (No.41601044), the Special Fund for Basic Scientific Research of Central Colleges, China University of Geosciences, Wuhan (CUGL170401 and CUGCJ1704), the Opening Foundation of Key Laboratory of Middle Atmosphere and Global Environment Observation (LAGEO), Institute of Atmospheric Physics, Chinese Academy of Sciences.

Acknowledgments: We would like to thank the China Meteorological Administration (CMA) and Chinese Ecosystem Research Network (CERN) for providing the meteorological and radiation data.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

R_{all}	the daily PAR under all-sky conditions	nb	the neuronal bias
R_{clr}	the daily PAR under clear sky conditions	a	altitude
R_b^{clr}	the beam PAR under clear sky conditions	st	surface temperature
R_d^{clr}	the diffuse PAR under clear sky conditions	pre	precipitation
τ_c	the transmittances due to cloud scattering and absorption	ps	surface pressure
r	the relative sunshine duration	rh	relative humidity
d_0/d	eccentricity correction factor for the mean sun-earth distance	sd	sunshine duration
R_0	the spectral irradiance (400–700 nm) at the mean distance between earth and sun in PAR band	at	air temperature
τ_b	the beam transmittance in clear sky conditions	ws	wind speed
τ_d	the diffuse transmittance in clear sky conditions	vis	visibility
τ_{wc}	global transmittance for water cloud	trise	sunrise time
τ_{ic}	global transmittance for ice cloud	AOD	Aerosol optical depth
τ_g	transmittances for mixed gasses absorption	CWP	cloud water path
τ_R	transmittances for Rayleigh scattering	w	Precipitable water vapor
τ_w	transmittances for water vapor absorption	Ioz	total zone amount
τ_o	transmittances for ozone absorption	g	Surface albedo
τ_a	transmittances for aerosol extinction	re	effective particle radius
α_w	the absorption coefficients of water vapor	TCP	cloud fraction
α_{aer}	the absorption coefficients of aerosols	θ	solar zenith angle
α_g	the absorption coefficients of mixed gasses	ρ_B	Earth-atmospheric albedo
I_{TOA}	extraterrestrial solar irradiation at the top of atmosphere in PAR band	ω	m-dimensional weight vector
$\rho_{a,all}$	atmospheric spherical albedo	δ	mapping function
$\rho_{a,clr}$	atmospheric spherical albedo for clear sky conditions	b	bias term
$\rho_{a,wc}$	atmospheric spherical albedo for water cloudy sky conditions	P_i	selection probability
$\rho_{a,ic}$	atmospheric spherical albedo for ice cloudy sky conditions	g_i	fitness value

d_n	day number since the first day of the year	N	is the number of input layers of Genetic
C_w	cloud fractions for water cloud	y_i	i -th expected output value
C_i	cloud fractions for ice cloud	o_i	i -th predicted output values
R_{clr}	PAR in clear sky conditions	$Z(\cdot)$	hidden transfer function
R_{wc}	PAR in water cloudy sky conditions	$w_i(t)$	the weight
R_{ic}	PAR in ice cloudy sky conditions	$x_i(t)$	the input parameters indiscrete time space
Fg	the estimated PAR	R_c	solar constant

Appendix A

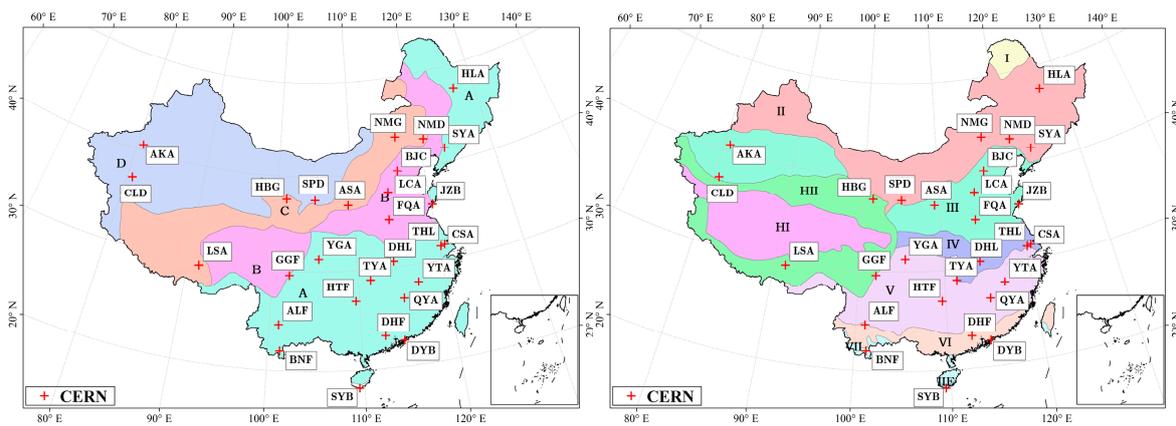


Figure A1. The humidity zones and temperate zones in China (A—humid, B—semi-humid, C—semi-arid, D—arid; I—cold temperate, II—mid temperate, III—warm temperate, IV—north subtropical zone, V—the mid-subtropics, VI—the south subtropics, VII—the edge of tropical zone, HI—sub-frigid zone in plateau, HII—temperature zone in plateau, and HIE—mid tropical zone with humid weather).

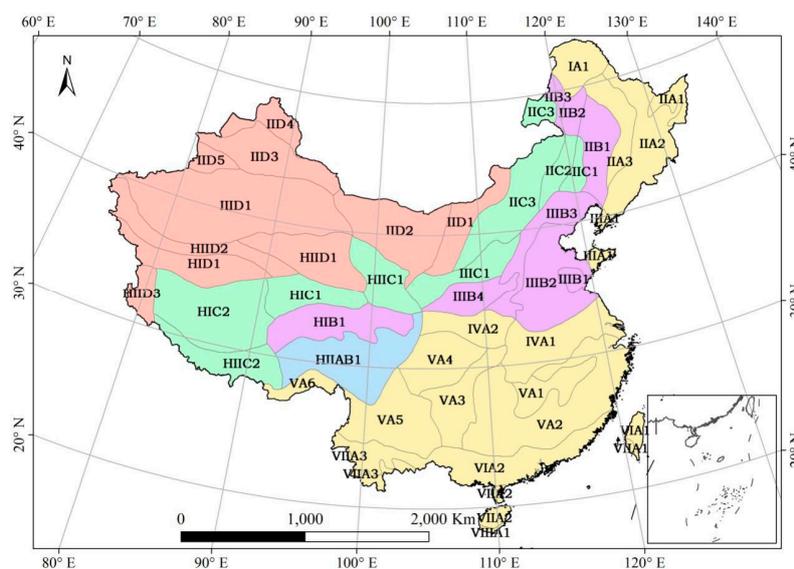


Figure A2. The topographic zones over mainland China. (IA1—Greater Khingan Range, IIA2—mountain area in the eastern part of Northeast China, IIB3—the foothills of the piedmont of the Sanhe mountain, IIA3—the piedmont plain in the eastern part of the Northeast China, IIA1—Sanjiang plain, IIB2—southern part of the Greater Khingan Range, IIB1—Central Songliao Plain, IIC3—Eastern Inner Mongolia high plain, IID4—Altai Mountains and the Tacheng Basin, IID3—Junggar Basin, IIC3—Eastern Inner Mongolia high plain, IIC2—southern part of the Greater Khingan Range, IIC1—Southwestern

Songliao Plain, **IID5**—Ili Basin, **IID1**—the Western and Hetao region of the Inner Mongolia high plain, **IID1**—the Tarim and Turpan Basin, **IID2**—the Alashan and the Hexi Corridor, **HIID2**—the North Wing of the Kunlun mountains, **IIIB3**—mountain and hills in North China, **IIIA1**—hills in Jiaodong and Liaodong, **IIIB2**—North China Plain, **HID1**—Alpine plateau in Kunlun, **HIID1**—Qaidam Basin, **IIIC1**—Jinzhong-Shaanxi-Gandong Plateau, **HIIC1**—Qilian mountain area, **IIIA1**—hills in Jiaodong and Liaodong, **IIIB1**—Shandong hills, **IIIB4**—Shanxi-Guanzhong Basin, **HIID3**—Alishan Mountain, **HIC1**—Southern Qinghai Plateau Gully, **HIC2**—Qiangtang Plateau Lake Basin, **IVA1**—Huainan and the middle and lower reaches of the Yangtze River, **HIB1**—GologNagqu hilly plateau, **IVA2**—the Hanzhoung Basin, **HIAB1**—the deep Alpine valley in Tibet, Sichuan Province, **VA4**—the Sichuan Basin, **VA2**—Jiangnan and Nanling Mountains, **VA3**—the Guizhou Plateau, **VA1**—the Chiang-nan Hilly Region, **VA5**—the Yunnan Plateau, **VA6**—South East Himalaya, **VIA2**—Fujian and Guangdong Guangxi hilly plain, **VIA1**—mountains and plains in North Central Taiwan, **VIIA1**—the lowlands in Southern Taiwan, **VIIA3**—the hilly Valley in Southern Yunnan, **VIIA2**—Hills in the Qiong Lei, **VIIIA1**—Qiong Lei lowlands and the Dongsha-Xisha-Nansha, **HIIC2**—the Zangnan mountain area).

References

- Purohit, I.; Purohit, P. Inter-comparability of solar radiation databases in Indian context. *Renew. Sustain. Energy Rev.* **2015**, *50*, 735–747. [[CrossRef](#)]
- Feng, L.; Lin, A.; Wang, L.; Qin, W.; Gong, W. Evaluation of sunshine-based models for predicting diffuse solar radiation in China. *Renew. Sustain. Energy Rev.* **2018**, *94*, 168–182. [[CrossRef](#)]
- Wang, L.; Gong, W.; Feng, L.; Lin, A.; Hu, B.; Zhou, M. Estimation of hourly and daily photosynthetically active radiation in Inner Mongolia, China, from 1990 to 2012. *Int. J. Climatol.* **2015**, *35*, 3120–3131. [[CrossRef](#)]
- Zhang, B.; Li, W.; Xie, G. Ecosystem services research in China: Progress and perspective. *Ecol. Econ.* **2010**, *69*, 1389–1395. [[CrossRef](#)]
- Beer, C.; Reichstein, M.; Tomelleri, E.; Ciais, P.; Jung, M.; Carvalhais, N.; Rödenbeck, C.; Arain, M.A.; Baldocchi, D.; Bonan, G.B.; et al. Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate. *Science* **2010**, *329*, 834. [[CrossRef](#)] [[PubMed](#)]
- Loutzenhiser, P.G.; Manz, H.; Felsmann, C.; Strachan, P.A.; Frank, T.; Maxwell, G.M. Empirical validation of models to compute solar irradiance on inclined surfaces for building energy simulation. *Sol. Energy* **2007**, *81*, 254–267. [[CrossRef](#)]
- Majasalmi, T.; Rautiainen, M.; Stenberg, P. Modeled and measured fPAR in a boreal forest: Validation and application of a new model. *Agric. For. Meteorol.* **2014**, *189–190*, 118–124. [[CrossRef](#)]
- Tang, W.; Qin, J.; Yang, K.; Niu, X.; Min, M.; Liang, S. An efficient algorithm for calculating photosynthetically active radiation with MODIS products. *Remote Sens. Environ.* **2017**, *194*, 146–154. [[CrossRef](#)]
- Rossow, W.B.; Schiffer, R.A. Advances in Understanding Clouds from ISCCP. *Bull. Am. Meteorol. Soc.* **1999**, *80*, 2261–2287. [[CrossRef](#)]
- Pinker, R.T.; Laszlo, I. Global distribution of photosynthetically active radiation as observed from satellites. *J. Clim.* **1992**, *5*, 56–65. [[CrossRef](#)]
- Eck, T.F.; Dye, D.G. Satellite estimation of incident photosynthetically active radiation using ultraviolet reflectance. *Remote Sens. Environ.* **1991**, *38*, 135–146. [[CrossRef](#)]
- King, M.D.; Kaufman, Y.J.; Menzel, W.P.; Tanré, D. Remote sensing of cloud, aerosol, and water vapor properties from the moderate resolution imaging spectrometer (MODIS). *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 2–27. [[CrossRef](#)]
- Moon, P. Proposed standard solar-radiation curves for engineering use. *J. Franklin Inst.* **1940**, *230*, 583–617. [[CrossRef](#)]
- McCree, K.J. A solarimeter for measuring photosynthetically active radiation. *Agric. Meteorol.* **1966**, *3*, 353–366. [[CrossRef](#)]
- Janjai, S.; Kumharn, W.; Laksanaboonsong, J. Determination of Angstrom's turbidity coefficient over Thailand. *Renew. Energy* **2003**, *28*, 1685–1700. [[CrossRef](#)]
- Yu, X.; Wu, Z.; Jiang, W.; Guo, X. Predicting daily photosynthetically active radiation from global solar radiation in the Contiguous United States. *Energy Convers. Manag.* **2015**, *89*, 71–82. [[CrossRef](#)]

17. Tan, P.Y.; Ismail, M.R.B. Photosynthetically active radiation and comparison of methods for its estimation in equatorial Singapore. *Theor. Appl. Climatol.* **2016**, *123*, 873–883. [[CrossRef](#)]
18. Gueymard, C. *SMARTS2: A simple Model of the Atmospheric Radiative Transfer of Sunshine: Algorithms and Performance Assessment*; Florida Solar Energy Center: Cocoa, FL, USA, 1995.
19. Bosch, J.L.; López, G.; Battles, F.J. Global and direct photosynthetically active radiation parameterizations for clear-sky conditions. *Agric. For. Meteorol.* **2009**, *149*, 146–158. [[CrossRef](#)]
20. Alados, I.; Olmo, F.J.; Foyo-Moreno, I.; Alados-Arboledas, L. Estimation of photosynthetically active radiation under cloudy conditions. *Agric. For. Meteorol.* **2000**, *102*, 39–50. [[CrossRef](#)]
21. Gueymard, C.A. REST2: High-performance solar radiation model for cloudless-sky irradiance, illuminance, and photosynthetically active radiation—Validation with a benchmark dataset. *Sol. Energy* **2008**, *82*, 272–285. [[CrossRef](#)]
22. Janjai, S.; Wattan, R. Development of a model for the estimation of photosynthetically active radiation from geostationary satellite data in a tropical environment. *Remote Sens. Environ.* **2011**, *115*, 1680–1693. [[CrossRef](#)]
23. Singh, R.K.; Liu, S.; Tieszen, L.L.; Suyker, A.E.; Verma, S.B. Novel approach for computing photosynthetically active radiation for productivity modeling using remotely sensed images in the Great Plains, United States. *J. Appl. Remote Sens.* **2012**, *6*, 63521–63522.
24. Zhang, X.; Liang, S.; Zhou, G.; Wu, H.; Zhao, X. Generating Global Land Surface Satellite incident shortwave radiation and photosynthetically active radiation products from multiple satellite data. *Remote Sens. Environ.* **2014**, *152*, 318–332. [[CrossRef](#)]
25. Li, L.; Xin, X.; Zhang, H.; Yu, J.; Liu, Q.; Yu, S.; Wen, J. A method for estimating hourly photosynthetically active radiation (PAR) in China by combining geostationary and polar-orbiting satellite data. *Remote Sens. Environ.* **2015**, *165*, 14–26. [[CrossRef](#)]
26. Li, Z.; Zhao, X.; Kahn, R.; Mishchenko, M.; Remer, L.; Lee, K.H.; Wang, M.; Ilaszlo, S.; Nakajima, T.; Maring, H. Uncertainties in satellite remote sensing of aerosols and impact on monitoring its long-term trend: A review and perspective. *Ann. Geophys.* **2009**, *27*, 2755–2770. [[CrossRef](#)]
27. Wang, L.; Kisi, O.; Zounemat-Kermani, M.; Hu, B.; Gong, W. Modeling and comparison of hourly photosynthetically active radiation in different ecosystems. *Renew. Sustain. Energy Rev.* **2016**, *56*, 436–453. [[CrossRef](#)]
28. López, G.; Rubio, M.A.; Mart, N.M.; Battles, F.J. Estimation of hourly global photosynthetically active radiation using artificial neural network models. *Agric. For. Meteorol.* **2001**, *107*, 279–291. [[CrossRef](#)]
29. Pankaew, P.; Milton, E.J.; Dash, J. Estimating hourly variation in photosynthetically active radiation across the UK using MSG SEVIRI data. In Proceedings of the IOP Conference Series-Earth and Environmental Science, Jakarta, Indonesia, 23–24 January 2014.
30. Yu, X.; Guo, X. Hourly photosynthetically active radiation estimation in Midwestern United States from artificial neural networks and conventional regressions models. *Int. J. Biometeorol.* **2016**, *60*, 1247–1259. [[CrossRef](#)] [[PubMed](#)]
31. Alados, I.; Alados-Arboledas, L. Direct and diffuse photosynthetically active radiation: Measurements and modelling. *Agric. For. Meteorol.* **1999**, *93*, 27–38.
32. Zhang, X.Z.; Zhang, Y.G.; Zhou, Y.H. Measuring and modelling photosynthetically active radiation in Tibet Plateau during April–October. *Agric. For. Meteorol.* **2000**, *102*, 207–212. [[CrossRef](#)]
33. Xie, X.; Gao, W.; Gao, Z. Estimation of land photosynthetically active radiation in clear sky using MODIS atmosphere and land products. *Proc. SPIE Int. Soc. Opt. Eng.* **2008**, 7083. [[CrossRef](#)]
34. Tang, W.; Qin, J.; Yang, K.; Niu, X.; Zhang, X.; Yu, Y.; Zhu, X. Reconstruction of daily photosynthetically active radiation and its trends over China. *J. Geophys. Res. Atmos.* **2013**, *118*, 13292–13302. [[CrossRef](#)]
35. Qin, J.; Yang, K.; Liang, S.; Tang, W. Estimation of Daily Mean Photosynthetically Active Radiation under All-Sky Conditions Based on Relative Sunshine Data. *J. Appl. Meteorol. Climatol.* **2012**, *51*, 150–160. [[CrossRef](#)]
36. Sun, Z. Improving transmission calculations for the Edwards-Slingo radiation scheme using a correlated-k distribution method. *Q. J. R. Meteorol. Soc.* **2011**, *137*, 2138–2148. [[CrossRef](#)]
37. Liu, R.; Liang, S.; He, H.; Liu, J.; Zheng, T. Mapping incident photosynthetically active radiation from MODIS data over China. *Remote Sens. Environ.* **2008**, *112*, 998–1009.
38. Zheng, T.; Liang, S.; Wang, K. Estimation of incident photosynthetically active radiation from GOES visible imagery. *J. Appl. Meteorol. Climatol.* **2008**, *47*, 853–868. [[CrossRef](#)]

39. Hu, B.; Wang, Y.; Liu, G. Measurements and estimations of photosynthetically active radiation in Beijing. *Atmos. Res.* **2007**, *85*, 361–371. [[CrossRef](#)]
40. Sun, Z.; Liu, J.; Zeng, X.; Liang, H. Parameterization of instantaneous global horizontal irradiance: Cloudy-sky component. *J. Geophys. Res. Atmos.* **2012**, *117*, D14. [[CrossRef](#)]
41. Jang, J.S.R. Fuzzy Modeling Using Generalized Neural Networks and Kalman Filter Algorithm. In Proceedings of the National Conference on Artificial Intelligence, Anaheim, CA, USA, 14–19 July 2012; pp. 762–767.
42. Zou, L.; Wang, L.; Xia, L.; Lin, A.; Hu, B.; Zhu, H. Prediction and comparison of solar radiation using improved empirical models and Adaptive Neuro-Fuzzy Inference Systems. *Renew. Energy* **2017**, *106*, 343–353. [[CrossRef](#)]
43. Kumar, M.; Kar, I.N. Non-linear HVAC computations using least square support vector machines. *Energy Convers. Manag.* **2009**, *50*, 1411–1418. [[CrossRef](#)]
44. Kisi, O. Pan evaporation modeling using least square support vector machine, multivariate adaptive regression splines and M5 model tree. *J. Hydrol.* **2015**, *528*, 312–320. [[CrossRef](#)]
45. Holland, J.H. Adaptation in natural and artificial systems. *Q. Rev. Biol.* **1975**, *6*, 126–137.
46. Quinlan, J.R. Learning with Continuous Classes. In Proceedings of the Australian Joint Conference on Artificial Intelligence, Hobart, Tasmania, Australia, 16–18 November 1992; pp. 343–348.
47. Solomatine, D.P. M5 Model Trees and Neural Networks: Application to Flood Forecasting in the Upper Reach of the Huai River in China. *J. Hydrol. Eng.* **2004**, *9*, 491–501. [[CrossRef](#)]
48. Rahimikhoob, A.; Asadi, M.; Mashal, M. A Comparison Between Conventional and M5 Model Tree Methods for Converting Pan Evaporation to Reference Evapotranspiration for Semi-Arid Region. *Water Resour. Manag.* **2013**, *27*, 4815–4826. [[CrossRef](#)]
49. Friedman, J.H. Multivariate Adaptive Regression Splines. *Ann. Stat.* **1991**, *19*, 1–67. [[CrossRef](#)]
50. Samui, P. Determination of ultimate capacity of driven piles in cohesionless soil: A Multivariate Adaptive Regression Spline approach. *Int. J. Numer. Anal. Met.* **2012**, *36*, 1434–1439. [[CrossRef](#)]
51. Sharda, V.N.; Patel, R.M.; Prasher, S.O.; Ojasvi, P.R.; Prakash, C. Modeling runoff from middle Himalayan watersheds employing artificial intelligence techniques. *Agric. Water Manag.* **2006**, *83*, 233–242. [[CrossRef](#)]
52. Dye, D.G. Spectral composition and quanta-to-energy ratio of diffuse photosynthetically active radiation under diverse cloud conditions. *J. Geophys. Res. Atmos.* **2004**, *109*. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).