



Article

Extension of PAR Models under Local All-Sky Conditions to Different Climatic Zones

Ana García-Rodríguez , Sol García-Rodríguez, Diego Granados-López, Montserrat Díez-Mediavilla and Cristina Alonso-Tristán * 

Research Group Solar and Wind Feasibility Technologies (SWIFT), Electromechanical Engineering Department, University of Burgos, 09006 Burgos, Spain; agrodriguez@ubu.es (A.G.-R.); solgr@ubu.es (S.G.-R.); dgranados@ubu.es (D.G.-L.); mdmr@ubu.es (M.D.-M.)

* Correspondence: catristan@ubu.es or cristinaalonso.tristan@gmail.com

Abstract: Four models for predicting Photosynthetically Active Radiation (PAR) were obtained through MultiLinear Regression (MLR) and an Artificial Neural Network (ANN) based on 10 meteorological indices previously selected from a feature selection algorithm. One model was developed for all sky conditions and the other three for clear, partial, and overcast skies, using a sky classification based on the clearness index (k_t). The experimental data were recorded in Burgos (Spain) at ten-minute intervals over 23 months between 2019 and 2021. Fits above 0.97 and Root Mean Square Error (RMSE) values below 7.5% were observed. The models developed for clear and overcast sky conditions yielded better results. Application of the models to the seven experimental ground stations that constitute the Surface Radiation Budget Network (SURFRAD) located in different Köppen climatic zones of the USA yielded fitted values higher than 0.98 and RMSE values less than 11% in all cases regardless of the sky type.



Citation: García-Rodríguez, A.; García-Rodríguez, S.; Granados-López, D.; Díez-Mediavilla, M.; Alonso-Tristán, C. Extension of PAR Models under Local All-Sky Conditions to Different Climatic Zones. *Appl. Sci.* **2022**, *12*, 2372. <https://doi.org/10.3390/app12052372>

Academic Editors: Harry D. Kambezidis and Basil Psiloglou

Received: 12 January 2022
Accepted: 22 February 2022
Published: 24 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: photosynthetically active radiation; k_t sky classification; ANN; multilinear regression models

1. Introduction

Photosynthetically Active Radiation (PAR) is a key factor for photosynthesis, vegetation growth, and climate change. The importance of its measurement and modelling is widely recognized [1–3], as PAR has a major influence on plant canopy growth, agricultural yields, and other environmental variables. Measurement of the PAR band is needed for sky-modelling of biomass growth and forestry production [3–6] and in natural greenhouse illumination [7]. Moreover, PAR affects the relationship between atmospheric systems and plants, so much so that its availability is a regulatory factor in the natural carbon cycle and in CO₂, water, and energy exchanges within the atmosphere [8].

Within the solar radiation spectrum, PAR is the portion with a wavelength between 400 and 700 nm [9]. Unfortunately, PAR sensors are not commonly found at ground meteorological stations [10]. It is therefore usually measured on the basis of other meteorological parameters and variables. The relationship between two different variables and PAR has been analyzed in different works, obtaining mathematical models of greater or lesser accuracy. Many of these models described in the literature for modelling PAR are based on linear regressions of global horizontal radiation [3,11,12], optical air mass [1], and the clearness index [13]. Other authors modelled PAR by simultaneously taking several variables into account: solar radiation, solar zenith angle, columnar perceptible water vapour, and aerosol optical depth [14]. Wang et al. [13] estimated PAR with the clearness index, day length, and zenith angle. Ferrera-Cobos et al. [15] used Global Horizontal Irradiance (R_{aGH}), Global extraterrestrial irradiance (G_0), atmospheric Temperature (T), and Relative Humidity (RH) as input variables for their models. These relationships have been analysed in most studies as a function of sky conditions [16] to estimate PAR for clear skies where the most important

variable is the solar zenith angle. The brightness index indicates whether the model is used in the presence of clouds. In several studies, *PAR* has been treated as a variable that depends on the location where it is measured or estimated [17]. As regards measurement temporality, the ratio of *PAR* and insolation in both tropical and arctic regions has been observed to remain fairly constant on different days and over longer time scales, regardless of cloud cover, atmospheric composition, surface type, season, and day length [18].

Ferrera-Cobos et al. [15] modeled *PAR* in oceanic and Mediterranean climates, testing 22 (11 multi-linear regression and 11 artificial neural network) models, using *RaGH*, G_0 , T , and RH as their input variables. They concluded that areas with different climatic conditions needed different models. The Mediterranean climatic models showed better fits, and the models for oceanic climates needed some corrections depending on their geographical location, as the higher humidity of an oceanic climate, due to higher levels of atmospheric water vapour, means more radiation in the infrared spectrum, increasing the *PAR* to G_0 ratio. Although the combination of using geostationary and polar-orbiting satellites is an optimal solution for estimating *PAR*, a network of *PAR* sensors is necessary for better validation [18]. Sudhakar et al. proposed a *PAR* estimation model for India at six latitudes between 9° to 34° , based on hourly and monthly averages of daily global radiation and a power regression model that accounted for different solar angles, cloud cover, and climatic conditions, which they linked to *RaGH* [3].

Several researchers have used data from different satellites, supplied from CM-SAF, to develop their *PAR* models, although field data had to be used for validation in each study [12,15]—a need that Vindel et al. [19] considered when presenting methodology to determine optimal locations for *PAR* measurement stations. This methodology is based on a clustering process applied to the k_t , *PAR*, calculated by dividing the *PAR* at the Earth's surface by the part of the spectrum corresponding to the *PAR* band at the top of the atmosphere (39.8% of the total).

In recent years, machine learning techniques have been used to develop algorithms for estimating *PAR*. The most decisive variable for *PAR* estimation was *RaGH* when comparing MLR and ANN models, regardless of climate [15]. Even when ANN models were run, their results clearly worsened without *RaGH*. In this sense, Jacovides [20] used the *sunshine fraction* (nN), T , *RaGH*, G_0 , and RH as input variables for the ANN models. They found that sunshine duration plays an important role in obtaining acceptable model predictions and that the model that best predicted *PAR* values combined sunshine duration and *RaGH*.

Lopez et al. [10] presented a model using *PAR* data collected at radiometric stations using a neural network. They estimated *PAR* with *RaGH* as the only measured variable. A second ANN model based on sunshine duration measurements was shown to be an acceptable alternative for calculating *PAR*. In contrast, Pankaew et al. [21] developed an ANN model for estimating hourly *PAR* data using seven atmospheric parameters (cosine of solar zenith angle, cloud index, precipitable water content, and aerosol optical depth) as the input collected from satellite data. They concluded that *PAR* estimation with an ANN model presented a good fit with a root mean square difference of 10.2%. Qin et al. [2] tested eight artificial intelligence models, among which the BackPropagation neural network model yielded the highest accuracy. Wang et al. [22] proposed three improved ANN models: MLP, Generalized Regression Neural Network (GRNN), and Radial Basis Neural Network (RBNN) for *PAR* estimation, from long-term hourly observations of *RaGH* and meteorological variables (air temperature, relative humidity, dew point temperature, water vapour pressure, air pressure). They found that different meteorological parameters influenced *PAR* estimation in accordance with each particular (agricultural land, soil, forest, bay, prairie, desert, and lake) ecosystem.

Yu et al. [23] studied the relationship between hourly *PAR* and *RaGH* from data collected over three years at the Bondville, IL, and Sioux Falls, SD, ground weather stations (United States). From these data, they determined the temporal variability of the *PAR* fraction and its dependence on different sky conditions (defined by the clearness index (k_t)). Furthermore, the results in terms of the normalized Root Mean Square Error ($nRMSE$), the

(R^2) coefficient of determination, the Mean Percentage Error (*MPE*), and Relative Standard Error (*RSE*) from the ANN-based models were compared with the same results from four existing conventional regression models. The authors found that the ANN model could accurately predict hourly *PAR*, especially under cloudy and clear sky conditions.

In general, as the literature review has highlighted, research on the modelling of different components of solar radiation has focused on obtaining models at specific locations, which can rarely be applied directly to other locations and usually requires local recalibration to achieve adequate results. In this study, our aim is to extend locally obtained models for *PAR* estimations as a function of the sky type to other locations. The sky conditions were determined from the clearness index (k_t). Three different *PAR* models, one for each sky type (clear, overcast, and partial), were developed from experimental data collected over ten-minute intervals at Burgos, Spain. The meteorological indices had previously been selected as the model inputs. Multi-Linear Regression (MLR) and ANN were the two modelling procedures. Both models were applied to experimental *PAR* data from all seven ground stations that form the SURFRAD network, corresponding to various Köppen–Geiger climate classification types [24].

The structure of this paper is as follows: after the Introduction Section, the databases, including the meteorological measurement stations, are described in Section 2. This section also includes the definition and description of the meteorological indices used for modelling *PAR*. In Section 3, the feature selection algorithm is described and *PAR* modelling is introduced using MLR and ANN models. In Section 4, the adjustment of the models obtained in Burgos for their application to the seven SURFRAD meteorological stations located in the United States is described. Finally, the main results and conclusions are presented in Section 5.

2. Description of the Experimental Data

Figure 1 shows the location of the weather station in Burgos, Spain (42°21'04" N, 3°41'20" W, 856 metres Above Mean Sea Level) (AMSL) where the data were collected for this study. The prevailing climate in Burgos, both oceanic and Mediterranean, is classified as Csb in the Köppen climate classification system. This weather station, described in previous work [25], is situated on the flat roof of a building at the University of Burgos, with no external obstructions or reflections from other surfaces. In addition, data were collected from all seven stations within the Surface Radiation Budget Network (SURFRAD), dependent on the National Oceanic and Atmospheric Administration (NOAA): Bondville station, Sioux Falls station, Boulder station, Desert Rock station, Fort Peck station, Goodwin Creek Station, and Penn State station. Figure 2 shows the location of the weather station in Burgos and the location of the seven SURFRAD weather stations.

In Burgos, *RaGH* and Diffuse Horizontal Irradiation (*RaDH*) were measured in $W \cdot m^{-2}$ using a Hukseflux pyranometer, model SR11, and a Hukseflux pyrliometer, model DR01, respectively. *PAR* was measured from Photosynthetic Photon Flux Density, Q_p ($\mu mol \cdot s^{-1} \cdot m^{-2}$), data and was then converted into *PAR* data ($W \cdot m^{-2}$) using McCree's conversion factor ($4.57 \mu mol \cdot J^{-1}$) [26] using a EKO quantum sensor, model ML-020P. All meteorological and radiometric data were recorded every ten minutes (averages from 30 s). The experimental campaign took place from April 2019 to February 2021. Experimental data were analyzed and then filtered using conventional quality criteria [27]. If a dataset failed to pass the quality criteria, then all simultaneous datasets were rejected. The original dataset counted 71,600 datums (ten-minute datasets), 36% of which were eliminated after the filtering procedure. In the USA, *PAR*, global, and diffuse irradiance values are given in $W \cdot m^{-2}$. The *PAR* values were measured employing a LI-COR Quantum sensor, while the *RaGH* values were measured using a pyranometer Spectrolab SR-75, and *RaDH* was measured with an Eppley, Model 8–48 pyranometer. The experimental campaign spanned 10 years, from January 2009 to December 2018 at each station, with a temporal resolution of 1 min, moving to ten-minute values by calculating the average of the values within that interval.

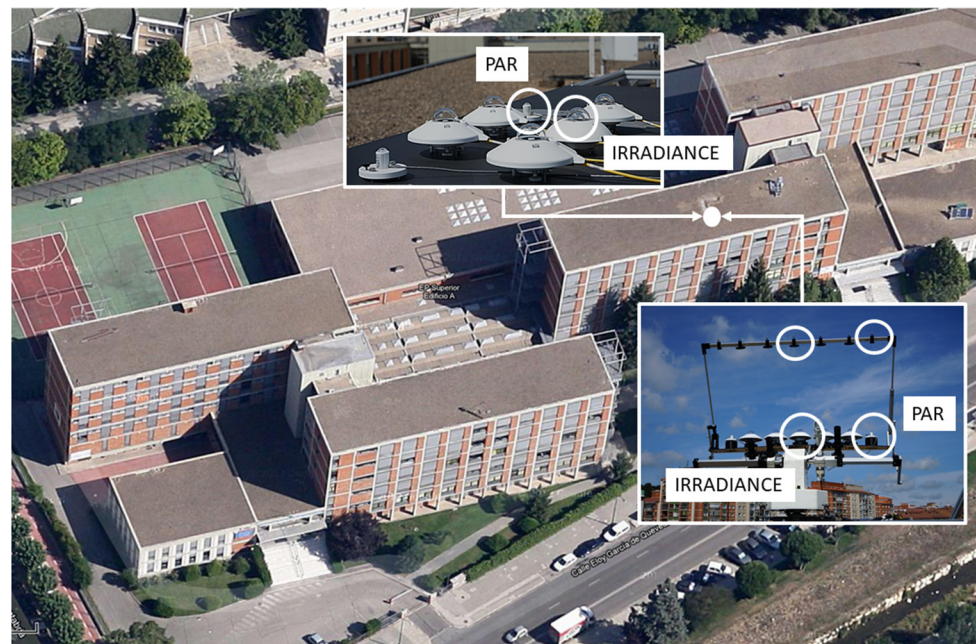


Figure 1. Experimental facility on the roof of the Higher Polytechnic School of.

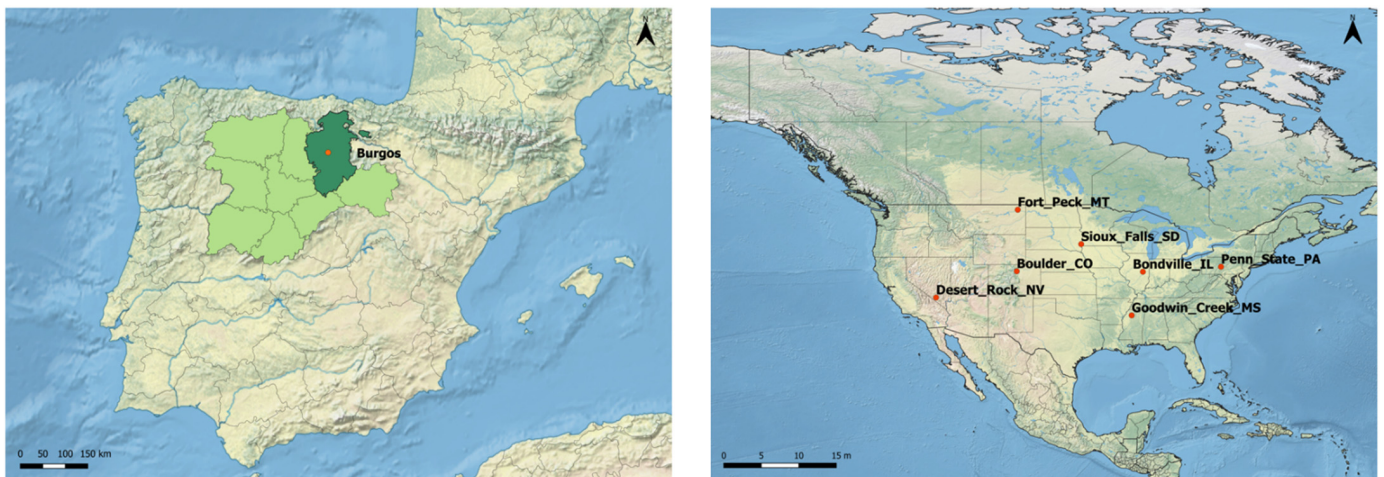


Figure 2. Location of the weather stations (Burgos and USA).

Burgos University, Spain. The Q_p , $RaGH$, and $RaDH$ sensors are shown in the detailed pictures.

Each Meteorological Index (MI) shown in Table 1 was determined at the single Spanish station and the seven meteorological stations in the USA. The following meteorological indices were directly obtained from the experimental measurements: T , P , $RaGH$, and PAR , obtained from Q_p , which were measured with pyranometers and a quantum sensor, respectively. The solar azimuth cosine ($\cos Z$) was calculated from the geometrical data of the location, using well-established mathematical relationships [28]. Finally, the dew point temperature (T_d) was calculated from the vapour water pressure [29] and RH . The other indices, k_f [28], the horizontal diffuse fraction, k_d [30], and Perez's clearness index, ε , and Perez's brightness factor, Δ [31], were calculated using equations described elsewhere [32].

Table 1. Meteorological indices (MIs) measured in Burgos.

MI	MI	Expression	Ref.
$RaGH$	Global Horizontal Irradiance	recorded	
k_d	Horizontal diffuse fraction	$k_d = \frac{RaDH}{RaGH}$	[30]
Q_p	Photosynthetic photon flux density	recorded	
PAR	Photosynthetically active radiation	$PAR = Q_p / 4.57 \mu\text{mol}\cdot\text{J}^{-1}$	[26]
k_t	Clearness index	$k_t = \frac{RaGH}{B_{sc}\cdot\epsilon_0\cdot\cos Z_s}$	[28]
T	Air temperature	recorded	
P	Pressure	recorded	
T_d	Dew point temperature	$T_d = \frac{35.859\cdot\log P_v - 21.48496}{\log P_v - 10.2858}$	[29]
$\cos Z$	Solar azimuth cosine	$\cos Z = \sin\delta\cdot\sin\phi + \cos\delta\cdot\cos\phi\cdot\cos\omega$	[28]
ϵ	Perez’s clearness index	$\epsilon = \frac{\frac{RaDH+RaB}{RaDH} + k\cdot Z_s^3}{1+k\cdot Z_s^3}$	[31]
Δ	Perez’s Brightness factor	$\Delta = \frac{m\cdot RaDH}{B_{sc}\cdot\epsilon_0\cdot\cos Z_s}$	[31]

$\epsilon_0 = 1 + 0.033\cdot\cos[2\cdot\pi\cdot d_n/365]$ is the average value of the orbital eccentricity of the Earth. d_n is the day of the year. B_{sc} is the extraterrestrial irradiance constant ($1361.1 \text{ W}\cdot\text{m}^{-2}$ [33]). $k = 1.04$ (or 5.56×10^{-6} if Z_s is expressed in degrees). Z_s is the angle between the sky zenith and sun. δ, ϕ, ω are the respective declination, hour angle, and geographic latitude of the specific location.

3. Methodology

The dataset was distributed into three categories of sky conditions based on the clearness index, k_t , [28] and the values adapted by Suarez-García [34] considering clear [0.65, 1), partial (0.35, 0.65), and overcast ($0 < 0.35$] skies. Figure 3 shows the Frequency Of Occurrence (FOC) of the different sky types in Burgos during the experimental campaign, which extended from April 2019 to February 2021. As can be seen, in Burgos, the clear sky conditions were prevalent except from November to January, when cloudy sky conditions occur more frequently.

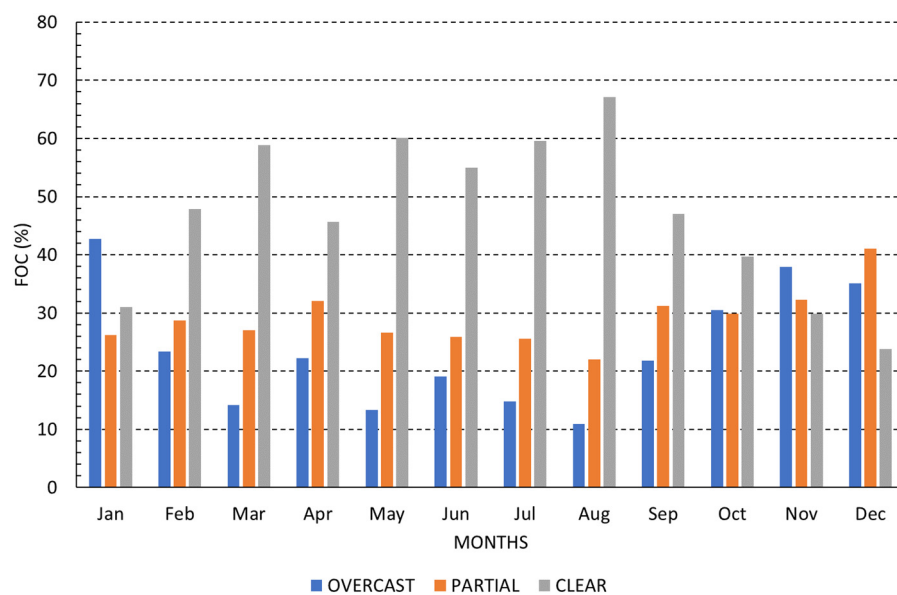


Figure 3. Monthly Frequency Of Occurrence (FOC, %) of clear, partial, and cloudy sky conditions based on the clearness index, k_t , in Burgos, Spain, between April 2019 and February 2021.

3.1. Feature Selection

The meteorological indices with the greatest influence on the *PAR* estimation models were selected to improve the precision and reduce the complexity of obtaining the models in Burgos, regardless of whether all sky types that can appear in Burgos were considered or whether the data were classified into clear, partial, and overcast skies. The first step determines the relationship of the different MIs to the *PAR*. The Pearson criterion, based on the Pearson correlation coefficient (*r*), was used to determine the influence or weight that each of index has on the *PAR* component. The criteria are as follows: if *r* is close to 0, the corresponding MI has a very weak relationship with *PAR*, whereas if it is close to 1, or -1 , the relationship is very strong. The Thumb rule [35] established five *r* intervals for the correlation: direct ($1 \geq |r(PAR, MI_i)| \geq 0.9$), strong ($0.9 > |r(PAR, MI_i)| \geq 0.7$), moderate ($0.7 > |r(PAR, MI_i)| \geq 0.5$), weak ($0.5 > |r(PAR, MI_i)| \geq 0.3$), and negligible ($|r(PAR, MI_i)| < 0.3$). Table 2 shows the different intervals of Pearson’s coefficients for the different MIs according to the classification of skies over Burgos with the *k_t* sky classification (clear, partial, and overcast).

Table 2. Pearson’s *r* (*PAR*, *MI_i*) based on the sky conditions according to the *k_t* sky classification (clear, partial, and overcast).

<i>k_t</i> Sky Type	$ r(PAR, MI_i) $				
	[1–0.9]	(0.9–0.7]	(0.7–0.5]	(0.5–0.3]	(0.3,0]
All sky conditions	<i>RaGH</i>	<i>cosZ</i> , <i>k_t</i>	<i>k_d</i> , ϵ	<i>T</i>	Δ , <i>P</i> , <i>T_d</i>
Clear	<i>RaGH</i> , <i>cosZ</i>		<i>k_t</i> , ϵ	<i>T</i>	<i>k_d</i> , Δ , <i>P</i> , <i>T_d</i>
Partial	<i>RaGH</i> , <i>cosZ</i>			<i>T</i>	<i>k_t</i> , <i>k_d</i> , Δ , ϵ , <i>P</i> , <i>T_d</i>
Overcast	<i>RaGH</i>	<i>cosZ</i>	<i>k_t</i> , Δ		<i>k_d</i> , ϵ , <i>P</i> , <i>T</i> , <i>T_d</i>

From the results shown in Table 2, it can be observed that *RaGH* is the MI that has a very strong and direct influence on *PAR* for all types of skies, coinciding with the results obtained by Ferrera-Cobos et al. [15]. Likewise, for all-sky types, *cosZ* and *k_t* also have a strong relationship, and *k_d* and ϵ have a moderate relationship with *PAR*. In the case of analyzing clear skies, there is a direct and strong relation with *RaGH* and a moderate relation with *k_t*, *cosZ*, and ϵ . For partial skies, there is also a direct and strong relation, in addition to *RaGH*, with *cosZ*. For overcast skies, there is also a direct and strong relation with *cosZ* and a moderate relation with *k_t*, and Δ . The rest of the indices (*T*, *k_d*, *P*, *T_d*) presented a weak or negligible relation with *PAR*, so these MIs were discarded as inputs for modelling *PAR*. These results agree with the literature, insofar as *RaGH* and *cosZ* are two of the variables with the strongest influence on *PAR* measurements [17,36].

3.2. Multilinear Regression Models

Once the meteorological indices with high influence on the *PAR* component were determined, taking into account all the sky types and the three sky types (clear, partial, and overcast), four MLR models were developed with which to estimate *PAR*: one MLR model for all sky types and one for each sky type. The indices selected with Pearson’s correlation coefficient (*r*) were taken as input variables in all models.

The working method to obtain the MLR models was as follows: the experimental dataset was divided into two groups. The first group contained 85% of the data and was used to fit the models. The other group, with the remaining 15% of the data, was used to validate the models using the conventional statistics corresponding to the coefficient of determination (*R*²), the normalized Mean Bias Error (*nMBE*), and the normalized Root Mean Square Error (*nRMSE*), calculated from Equations (1) and (2):

$$nMBE = \frac{1}{PAR_{exp}} \frac{\sum_{i=1}^n (PAR_{mod} - PAR_{exp})}{n} \times 100 (\%) \tag{1}$$

$$nRMSE = \frac{1}{PAR_{exp}} \sqrt{\frac{\sum_{i=1}^n (PAR_{mod} - PAR_{exp})^2}{n}} \times 100 (\%) \tag{2}$$

where n represents the amount of experimental data used for fitting the models; PAR_{exp} is the experimental value of PAR , and PAR_{mod} is the modelled PAR value. Table 3 shows the mathematical expressions of the four regression models and the fit of each one with their mean square error. As Table 3 shows, the model fitting results presented good correlations with the experimental data, given $R^2 > 0.97$ and $nRMSE$ lower than 8%. The small and negative values of $nMBE$ indicated that the models presented a good fit although they tended to underestimate the PAR values. It can be seen that the model for clear skies showed the highest accuracy with an R^2 of 0.99 and had the lowest errors, for $nRMSE$ and for $nMBE$.

Table 3. Multilinear regression models of PAR .

Sky Conditions	Multilinear Regression Model	R^2	$nRMSE$ (%)	$nMBE$ (%)
All skies (MLR1)	$PAR = 12.12 + 0.40 \cdot RaGH + 15.74 \cdot \cos Z - 11.44 \cdot kt - 10.64 \cdot kd - 0.47 \cdot \epsilon$	0.994	4.37	-2.74×10^{-13}
Clear skies (MLR2)	$PAR = -18.12 + 0.33 \cdot RaGH + 83.15 \cdot \cos Z + 24.19 \cdot kt + 0.71 \cdot \epsilon$	0.990	3.27	-1.45×10^{-14}
Partial skies (MLR3)	$PAR = -1.81 + 0.40 \cdot RaGH + 13.75 \cdot \cos Z$	0.977	6.80	-5.32×10^{-13}
Overcast skies (MLR4)	$PAR = -0.03 + 0.42 \cdot RaGH + 6.88 \cdot \cos Z + 1.58 \cdot kt - 6.12 \cdot \Delta$	0.978	7.33	1.87×10^{-13}

3.3. Artificial Neural Network Model

Another PAR estimation method used an ANN trained with the Levenberg–Marquardt Back-Propagation (LMBP) algorithm [37]. In this case, a 3-layer configuration was chosen: the input layer, with each MI for each model, the hidden layer with the information processing centres (neurons), and the output layer, with the results. The explanation of the iterative process and the fitting are explained elsewhere [32].

For this purpose, four networks were developed and tested in this work, one for all skies and another three for each sky type, based on each selected MI with its previously determined Pearson’s correlation coefficient shown in Table 2. The R^2 , the $nMBE$, and the $nRMSE$ statistics, previously defined above, were used for the validation of the models. Table 4 summarizes the statistical results of each ANN. It can be seen that a very good fit was obtained for each ANN with an $R^2 > 0.97$ and $nRMSE < 8\%$.

Table 4. Statistical results of the ANN models.

Sky Conditions	R^2	$nRMSE$ (%)	$nMBE$ (%)
All skies (ANN1)	0.994	4.22	2.84×10^{-3}
Clear skies (ANN2)	0.992	3.01	-4.68×10^{-2}
Partial skies (ANN3)	0.977	6.80	4.06×10^{-3}
Overcast skies (ANN4)	0.978	7.28	-3.50×10^{-2}

A check was performed with 15% of the values not used in the generation of the models to validate the results of both the MLR and ANN models, which confirmed the good fit of the models for all sky types, both for the regressions and the neural networks. The results, presented below in Table 5, show the MLR results on the left and ANN results on the right.

As can be seen from the data in the table, both “data treatment” methods are optimal, as the R^2 is greater than 0.97 and the errors are less than 7%. The best-fitting model in both cases is the one applied to clear skies with fit values of 0.99 and mean square errors of less than 3%.

Table 5. Validation of MLR and ANN models.

Sky Conditions	R^2	$nRMSE$ (%)	$nMBE$ (%)	Sky Conditions	R^2	$nRMSE$ (%)	$nMBE$ (%)
All skies (MLR1)	0.994	4.48	-5.20×10^{-3}	All skies (ANN1)	0.994	4.35	-4.00×10^{-3}
Clear skies (MLR2)	0.992	3.00	1.21×10^{-1}	Clear skies (ANN2)	0.993	2.782	4.89×10^{-2}
Partial skies (MLR3)	0.978	6.72	-6.02×10^{-2}	Partial skies (ANN3)	0.978	6.71	-5.26×10^{-2}
Overcast skies (MLR4)	0.982	6.63	-4.84×10^{-2}	Overcast skies (ANN4)	0.982	6.59	-5.10×10^{-1}

4. Extension of the Models to Other Locations

The final objective of this work was to verify whether the models obtained with the experimental values obtained in Burgos (Spain) were valid for other climatic zones, for which purpose all seven SURFRAD stations in the USA were selected and the multilinear regression and neural network models were applied with the coefficients obtained for Burgos to check their fit.

The seven American stations are at the geographical locations shown in Table 6. The type of climate at each station was also included in this table, according to the Köppen climatic classification [38]: hot desert climate (Bwh); hot-summer humid continental climate (Dfa); warm-summer humid continental climate (Dfb); humid subtropical climate (Cfa); cold semi-arid climate (BSk).

Table 6. Geographical data of SURFRAD weather stations.

	Latitude (°N)	Latitude (°W)	Altitude (m.a.s.l.)	Climate
Bondville, Illinois	40.05192	88.37309	230	Dfa
Table Mountain, Boulder, Colorado	40.12498	105.2368	1689	Bsk
Desert Rock, Nevada	36.62373	116.01947	1007	Bwh
Fort Peck, Montana	48.30783	105.1017	634	Bsk
Goodwin Creek, Mississippi	34.2547	89.8729	6	Cfa
Penn State, Univ. Pennsylvania	40.72012	77.93085	376	Dfb
Sioux Falls, South Dakota	43.73403	96.62328	473	Dfa

(m.a.s.l.: meters above sea level).

The models based on the data collected at Burgos were applied to the seven American meteorological stations and compared with their own primary data. For this purpose, the R^2 coefficient and the $nRSME$ and $nMBE$ errors were obtained for all stations, both with the MLR and the ANN models for all sky types and for each particular sky type, as shown in Table 7.

Observing Table 7, it can be seen that the fit values were also very good with an R^2 greater than 0.99 for all stations and both the MLR and ANN models. The $nRMSE$ was always less than 7%, and all $nMBE$ values showed that all the models overestimated the PAR values, albeit with minimally higher values. The station with the lowest $nRMSE$ and the best fit was Penn State, for both the MLR and ANN results.

If the results for clear skies are analyzed (Table 8), it can be observed that the fit was also very good with an R^2 higher than 0.98 (although slightly lower than the study carried out for all types of sky conditions). In this case, the $nRMSE$ was lower, and at no station did it exceed 6%.

When the study was centred upon partial skies (Table 9), the R^2 values were greater than 0.99 and the errors were less than 9.5%. Finally, when the study was performed with overcast skies (Table 10), the R^2 value was greater than 0.98 and the errors with MLR were less than 8%. In this case, the $nRMSE$ at the Desert Rock station where the climate is hot and arid was slightly higher than the rest (11.07%) in the ANN study.

If the results are analysed by climatic zones, it can be seen that there were hardly any differences in the fit (R^2); all values were higher than 0.98 regardless of the sky type analysed. Figures 4 and 5 show the $nRMSE$ values for the four sky categories (all-sky type, clear, partial, and overcast) at all stations, for the MLR and the ANN models. When analysing the results for all sky types, the highest error occurred in the cold semi-arid zone (Bsk) (far from the climatic zone for which the regression and ANN models were obtained) and the lowest RMSE in continental zones with greater similarity to the climatic zone in which the models were obtained in Burgos. For clear skies in both the MLR and ANN models, the errors were slightly lower at all stations; the lowest RMSE was also obtained in continental climate zones (Dfb and Dfa) and the highest error in desert zones (Bwh).

Table 7. Model fit values for MLR models (left) and for ANN models (right) for all-sky types.

USA Stations	MLR			ANN		
	R^2	$nRMSE$ (%)	$nMBE$ (%)	R^2	$nRMSE$ (%)	$nMBE$ (%)
Bondville, Illinois	0.994	4.62	1.89	0.994	4.90	2.54
Table Mountain, Boulder, Colorado	0.996	4.93	3.48	0.996	5.48	4.18
Desert Rock, Nevada	0.997	6.10	5.36	0.997	6.75	6.04
Fort Peck, Montana	0.994	6.46	4.94	0.994	6.87	5.45
Goodwin Creek, Mississippi	0.995	4.66	2.48	0.995	5.15	3.33
Penn State, Univ. Pennsylvania	0.995	4.35	1.34	0.995	4.62	2.07
Sioux Falls, South Dakota	0.995	6.34	5.00	0.995	6.84	5.59

Table 8. Model fit values for MLR (left) and ANN (right) models under clear sky conditions.

USA Stations	MLR			ANN		
	R^2	$nRMSE$ (%)	$nMBE$ (%)	R^2	$nRMSE$ (%)	$nMBE$ (%)
Bondville, Illinois	0.985	4.05	1.61	0.985	4.13	1.28
Table Mountain, Boulder, Colorado	0.993	4.50	3.30	0.992	4.67	3.13
Desert Rock, Nevada	0.994	5.50	4.87	0.994	5.77	4.77
Fort Peck, Montana	0.988	5.41	4.16	0.987	5.31	3.74
Goodwin Creek, Mississippi	0.985	4.25	2.37	0.984	4.46	2.06
Penn State, Univ. Pennsylvania	0.987	3.77	1.20	0.986	4.02	0.95
Sioux Falls, South Dakota	0.991	5.44	4.49	0.990	5.36	4.03

Table 9. Model fit values for MLR (left) and ANN (right) under partial sky conditions.

USA Stations	MLR			ANN		
	R^2	$nRMSE$ (%)	$nMBE$ (%)	R^2	$nRMSE$ (%)	$nMBE$ (%)
Bondville, Illinois	0.990	5.66	3.15	0.990	5.67	3.19
Table Mountain, Boulder, Colorado	0.994	6.21	4.67	0.994	6.24	4.72
Desert Rock, Nevada	0.996	9.29	8.23	0.996	9.33	8.31
Fort Peck, Montana	0.991	8.06	6.51	0.991	8.08	6.56
Goodwin Creek, Mississippi	0.992	5.87	4.00	0.992	5.90	4.04
Penn State, Univ. Pennsylvania	0.991	5.00	2.27	0.991	5.00	2.30
Sioux Falls, South Dakota	0.992	7.62	6.22	0.992	7.64	6.26

Table 10. Model fit values for MLR (left) and ANN (right) under overcast sky conditions.

USA Stations	MLR			ANN		
	R^2	$nRMSE$ (%)	$nMBE$ (%)	R^2	$nRMSE$ (%)	$nMBE$ (%)
Bondville, Illinois	0.984	6.29	−0.23	0.984	7.13	3.32
Table Mountain, Boulder, Colorado	0.992	4.60	0.70	0.992	6.23	4.18
Desert Rock, Nevada	0.993	7.91	5.81	0.993	11.07	9.40
Fort Peck, Montana	0.989	6.11	2.61	0.988	8.48	6.37
Goodwin Creek, Mississippi	0.987	5.75	0.19	0.986	6.82	3.48
Penn State, Univ. Pennsylvania	0.987	6.11	−1.67	0.987	6.15	1.91
Sioux Falls, South Dakota	0.989	6.07	1.88	0.988	8.14	5.54

The errors at all stations were somewhat higher for overcast skies ranging from 4.6 to 11.07, although the ANN model still yielded the highest RMSE for the hot desert zone (Bwh). The overcast skies can present very different conditions depending on parameters such as the degree of cloud cover, its thickness, and the presence of aerosols, which have a great influence on the phenomena of absorption and dispersion of solar radiation.

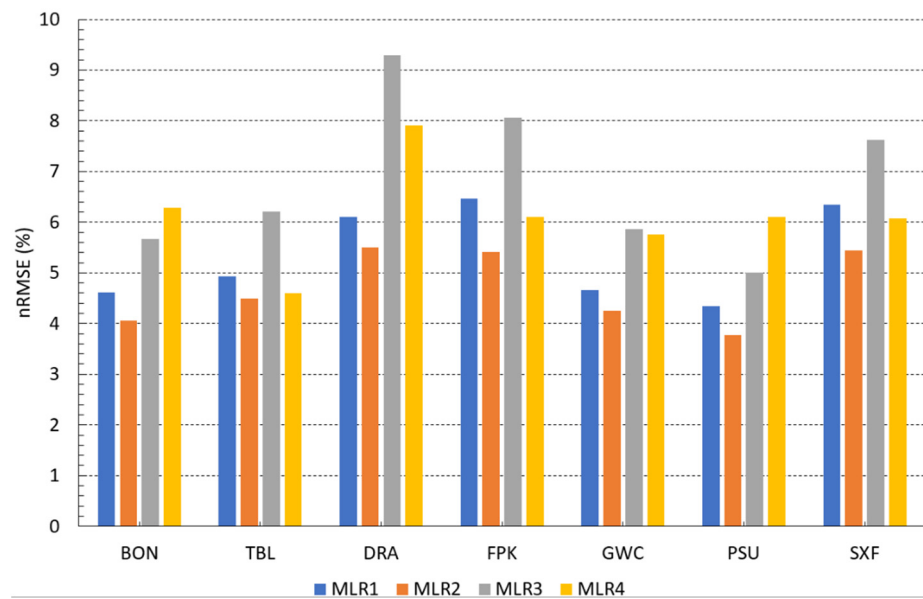


Figure 4. nRMSE (MLR) values (%) for all sky type categories at the SURFRAD stations.

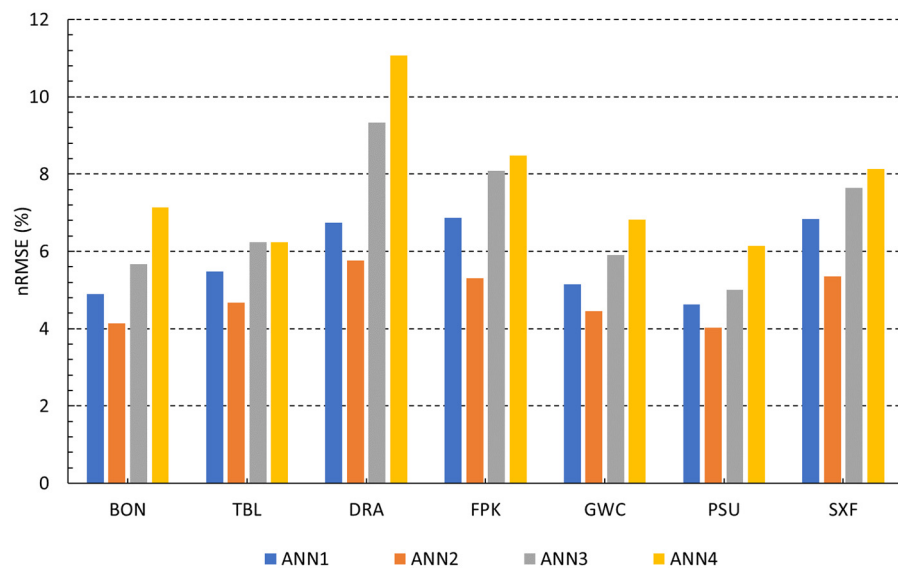


Figure 5. nRMSE (ANN) values (%) for all sky type categories at the SURFRAD stations.

5. Conclusions

A set of MLR and ANN models were obtained to model *PAR* from a set of meteorological indices in Burgos. It was determined that the most influential variable when modelling *PAR*, both when analysing all-sky types and for each particular type (clear, partial, and, overcast skies) was *RaGH*. For clear and partial skies, *cosZ* also had a very strong influence, while this same MI also had a strong relationship with *PAR* for all-sky types and for overcast skies. The MI k_t had a strong relationship with *PAR* for all-sky types, while the relationship was moderate for clear, partial, and overcast skies. Perez’s clearness index, ϵ , showed a moderate relationship with *PAR* for all-skies and clear skies, and Δ was the MI with a moderate relationship for overcast skies.

Four *PAR* estimation models were obtained with MLR models that yielded an $R^2 > 0.97$, reaching an $R^2 > 0.99$ in the cases of all-skies types and clear skies; the mean square errors were lower than 8%, although this error was 3.27% for clear skies. The statistical coefficients of the ANN model were also very well adjusted because the R^2 value was in all cases

greater than 0.97 and the $nRMSE$ was also less than 8%. In this case, the model fitted clear skies better than any other, obtaining an $R^2 > 0.99$ and an $nRMSE$ of 3.01%.

These models obtained in Burgos were directly applied to seven SURFRAD stations in the USA to verify their utility, obtaining a very good fit for all of them for all sky conditions, with both MLR and ANN models, as the R^2 values were greater than 0.98 and the mean square errors were less than 11%. It is worth noting that clear skies were the sky type condition where the $nRMSEs$ were the lowest, between 3 and 5%. Models for overcast sky conditions presented worse results, mainly due to the variability of the phenomena of absorption and dispersion of solar radiation that can appear under cloudy sky conditions and that cannot be characterized by a single parameter such as k_t . It was also noteworthy that the models showed a worse fit at Desert Rock station under all sky conditions, each with an $nRMSE$ at around 9–10% for partial and overcast skies. There were no significant differences when the data were studied by climatic zones, which in all cases showed a good fit with an R^2 higher than 0.98, a lower $nRMSE$ in continental zones (Dfa and Dfb), and higher errors in semi-arid and desert zones (Bsk and Bwh).

Both the MLR and ANN models that were used to estimate PAR on the basis of experimental data collected at Burgos may be used in other locations independently of the climatology, insofar as they yielded very good statistical coefficients across the seven meteorological stations representative of very different climates within the United States. They also fitted well for different climates, although in the case of a hot arid climate, the models presented the largest quadratic errors for each type of sky.

Author Contributions: Conceptualization, M.D.-M. and C.A.-T.; methodology, A.G.-R. and S.G.-R.; software, D.G.-L.; validation, D.G.-L., A.G.-R., M.D.-M.; formal analysis, A.G.-R. and S.G.-R.; investigation, A.G.-R. and S.G.-R.; original draft preparation, M.D.-M.; writing—review and editing, C.A.-T.; visualization, A.G.-R. and S.G.-R.; supervision, M.D.-M. and C.A.-T.; project administration, M.D.-M.; funding acquisition, C.A.-T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Spanish Ministry of Science and Innovation, grant number RTI2018-098900-B-I00, and Consejería de Empleo e Industria, Junta de Castilla y León, grant number INVESTUN/19/BU-0004.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to thank the SURFRAD network for placing complete and quality weather data in the public domain. Diego Granados López gratefully acknowledges economic support from the Junta de Castilla-León (ORDEN EDU/556/2019).

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

Abbreviation

Acronyms

ANN	Artificial Neural Network
MPE	Mean Percentage Error
MI	Meteorological Index
MLR	MultiLinear Regression
m.a.s.l.	meters above sea level
NOAA	National Oceanic and Atmospheric Administration
$nRMSE$	Normalized Root Mean Square Error
$nMBE$	Normalized Mean Bias Error
RSE	Relative Standard Error
SURFRAD	Surface Radiation Budget Network

Meteorological variables

$\cos Z$	Solar azimuth cosine ($^\circ$)
Δ	Perez's brightness factor (dim)
ϵ	Perez's clearness index (dim)
$RaGH$	Global Horizontal Irradiance ($W \cdot m^{-2}$)

$RaDH$	Diffuse Horizontal Irradiance ($W \cdot m^{-2}$)
RaB	Beam Irradiance ($W \cdot m^{-2}$)
k_d	Horizontal diffuse fraction (dim)
k_t	Clearness Index (dim)
P	Atmospheric Pressure (kPa)
PAR	Photosynthetically Active Radiation ($W \cdot m^{-2}$)
Q_p	Photosynthetic photon flux density ($\mu mol \cdot s^{-1} \cdot m^{-2}$)
r	Pearson correlation coefficient (dim)
RH	Relative Humidity (%)
T	Atmospheric Temperature ($^{\circ}C$)
T_d	Dew point temperature ($^{\circ}C$)

References

- Hu, B.; Yu, Y.; Liu, Z.; Wang, Y. Analysis of Photosynthetically Active Radiation and Applied Parameterization Model for Estimating of PAR in the North China Plain. *J. Atmos. Chem.* **2016**, *73*, 345–362. [\[CrossRef\]](#)
- Qin, W.; Wang, L.; Zhang, M.; Niu, Z.; Luo, M.; Lin, A.; Hu, B. First Effort at Constructing a High-Density Photosynthetically Active Radiation Dataset during 1961–2014 in China. *J. Clim.* **2019**, *32*, 2761–2780. [\[CrossRef\]](#)
- Sudhakar, K.; Srivastava, T.; Satpathy, G.; Premalatha, M. Modelling and Estimation of Photosynthetically Active Incident Radiation Based on Global Irradiance in Indian Latitudes. *Int. J. Energy Environ. Eng.* **2013**, *4*, 21. [\[CrossRef\]](#)
- Aguiar, L.J.G.; Fischer, G.R.; Ladle, R.J.; Malhado, A.C.M.; Justino, F.B.; Aguiar, R.G.; da Costa, J.M.N. Modeling the Photosynthetically Active Radiation in South West Amazonia under All Sky Conditions. *Theor. Appl. Climatol.* **2012**, *108*, 631–640. [\[CrossRef\]](#)
- Landsberg, J.J.; Waring, R.H. A Generalised Model of Forest Productivity Using Simplified Concepts of Radiation-Use Efficiency, Carbon Balance and Partitioning. *For. Ecol. Manag.* **1997**, *95*, 209–228. [\[CrossRef\]](#)
- Rubio, M.A.; López, G.; Tovar, J.; Pozo, D.; Batlles, F.J. The Use of Satellite Measurements to Estimate Photosynthetically Active Radiation. *Phys. Chem. Earth* **2005**, *30*, 159–164. [\[CrossRef\]](#)
- Alados, I.; Foyo-Moreno, I.; Alados-Arboledas, L. Photosynthetically Active Radiation: Measurements and Modelling. *Agric. For. Meteorol.* **1996**, *78*, 121–131. [\[CrossRef\]](#)
- Deo, R.C.; Downs, N.J.; Adamowski, J.F.; Parisi, A.V. Adaptive Neuro-Fuzzy Inference System Integrated with Solar Zenith Angle for Forecasting Sub-Tropical Photosynthetically Active Radiation. *Food Energy Secur.* **2019**, *8*, e00151. [\[CrossRef\]](#)
- Alados-Arboledas, L.; Olmo, F.J.; Alados, I.; Pérez, M. Parametric Models to Estimate Photosynthetically Active Radiation in Spain. *Agric. For. Meteorol.* **2000**, *101*, 187–201. [\[CrossRef\]](#)
- López, G.; Rubio, M.; Martínez, M.; Batlles, F. Estimation of Hourly Global Solar Radiation Using Artificial Neural Network. *Agric. For. Meteorol.* **2001**, *107*, 279–291. [\[CrossRef\]](#)
- Majnooni-heris, A. Estimating Photosynthetically Active Radiation (PAR) Using Air Temperature and Sunshine Durations. *J. Biodivers. Environ. Sci.* **2014**, *5*, 371–377.
- Vindel, J.M.; Valenzuela, R.X.; Navarro, A.A.; Zorzalejo, L.F.; Paz-Gallardo, A.; Souto, J.A.; Méndez-Gómez, R.; Cartelle, D.; Casares, J.J. Modeling Photosynthetically Active Radiation from Satellite-Derived Estimations over Mainland Spain. *Remote Sens.* **2018**, *10*, 849. [\[CrossRef\]](#)
- Wang, L.; Gong, W.; Li, C.; Lin, A.; Hu, B.; Ma, Y. Measurement and Estimation of Photosynthetically Active Radiation from 1961 to 2011 in Central China. *Appl. Energy* **2013**, *111*, 1010–1017. [\[CrossRef\]](#)
- Zempila, M.M.; Taylor, M.; Bais, A.; Kazadzis, S. Modeling the Relationship between Photosynthetically Active Radiation and Global Horizontal Irradiance Using Singular Spectrum Analysis. *J. Quant. Spectrosc. Radiat. Transf.* **2016**, *182*, 240–263. [\[CrossRef\]](#)
- Ferrera-Cobos, F.; Vindel, J.M.; Valenzuela, R.X.; González, J.A. Models for Estimating Daily Photosynthetically Active Radiation in Oceanic and Mediterranean Climates and Their Improvement by Site Adaptation Techniques. *Adv. Space Res.* **2020**, *65*, 1894–1909. [\[CrossRef\]](#)
- Möttus, M.; Ross, J.; Sulev, M. Experimental Study of Ratio of PAR to Direct Integral Solar Radiation under Cloudless Conditions. *Agric. For. Meteorol.* **2001**, *109*, 161–170. [\[CrossRef\]](#)
- Nwokolo, S.; Amadi, S.O. A Global Review of Empirical Models for Estimating Photosynthetically Active Radiation. *Trends Renew. Energy* **2018**, *4*, 236–327. [\[CrossRef\]](#)
- Frouin, R.; Pinker, R.T. Estimating Photosynthetically Active Radiation (PAR) at the Earth's Surface from Satellite Observations. *Remote Sens. Environ.* **1995**, *51*, 98–107. [\[CrossRef\]](#)
- Vindel, J.M.; Valenzuela, R.X.; Navarro, A.A.; Zorzalejo, L.F. Methodology for Optimizing a Photosynthetically Active Radiation Monitoring Network from Satellite-Derived Estimations: A Case Study over Mainland Spain. *Atmos. Res.* **2018**, *212*, 227–239. [\[CrossRef\]](#)
- Jacovides, C.P.; Tymvios, F.S.; Boland, J.; Tsitouri, M. Artificial Neural Network Models for Estimating Daily Solar Global UV, PAR and Broadband Radiant Fluxes in an Eastern Mediterranean Site. *Atmos. Res.* **2015**, *152*, 138–145. [\[CrossRef\]](#)
- Pankaew, P.; Pattarapanitchai, S.; Buntoung, S.; Wattan, R.; Masiri, I.; Sripradit, A.; Janjai, S. Estimating Photosynthetically Active Radiation Using an Artificial Neural Network. In Proceedings of the 2014 International Conference and Utility Exhibition on Green Energy for Sustainable Development (ICUE), Pattaya, Thailand, 19–21 March 2014.

22. 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](#)]
23. 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](#)]
24. Peel, M.; Finlayson, B.; McMahon, T. Updated World Map of the Köppen-Geiger Climate Classification. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 1633–1644. [[CrossRef](#)]
25. García-Rodríguez, A.; García-Rodríguez, S.; Díez-Mediavilla, M.; Alonso-Tristán, C. Photosynthetic Active Radiation, Solar Irradiance and the Cie Standard Sky Classification. *Appl. Sci.* **2020**, *10*, 8007. [[CrossRef](#)]
26. Akitsu, T.; Kume, A.; Hirose, Y.; Ijima, O.; Nasahara, K. On the Stability of Radiometric Ratios of Photosynthetically Active Radiation to Global Solar Radiation in Tsukuba, Japan. *Agric. For. Meteorol.* **2015**, *209–210*, 59–68. [[CrossRef](#)]
27. Gueymard, C.A.; Ruiz-Arias, J.A. Extensive Worldwide Validation and Climate Sensitivity Analysis of Direct Irradiance Predictions from 1-Min Global Irradiance. *Sol. Energy* **2016**, *128*, 1–30. [[CrossRef](#)]
28. Iqbal, M. *An Introduction to Solar Radiation*; Academic Press: New York, NY, USA, 1983.
29. Heating, Refrigerating American Society of y Air Conditing Engineers. American Society of Heating, Refrigerating y Air conditioning Engineers. In *ASHRAE Handbook*; Ashrae: Atlanta, GA, USA, 2015; ISBN 9781936504459.
30. Erbs, D.G.; Klein, S.A.; Duffie, J.A. Estimation of the Diffuse Radiation Fraction for Hourly, Daily and Monthly-Average Global Radiation. *Sol. Energy* **1982**, *28*, 293–302. [[CrossRef](#)]
31. Perez, R.; Ineichen, P.; Seals, R.; Michalsky, J.; Stewart, R. Modeling Daylight Availability and Irradiance Components from Direct and Global Irradiance. *Sol. Energy* **1990**, *44*, 271–289. [[CrossRef](#)]
32. García-Rodríguez, A.; Granados-López, D.; García-Rodríguez, S.; Díez-Mediavilla, M.; Alonso-Tristán, C. Modelling Photosynthetic Active Radiation (PAR) through Meteorological Indices under All Sky Conditions. *Agric. For. Meteorol.* **2021**, *310*, 108627. [[CrossRef](#)]
33. Gueymard, C.A. A Reevaluation of the Solar Constant Based on a 42-Year Total Solar Irradiance Time Series and a Reconciliation of Spaceborne Observations. *Sol. Energy* **2018**, *168*, 2–9. [[CrossRef](#)]
34. Suárez-garcía, A.; Granados-lópez, D.; González-peña, D.; Alonso-tristán, C. Benchmarking of Meteorological Indices for Sky Cloudiness Classification. *Sol. Energy* **2020**, *195*, 499–513. [[CrossRef](#)]
35. Mukaka, M.M. Statistics Corner: A Guide to Appropriate Use of Correlation Coefficient in Medical Research. *Malawi Med. J.* **2012**, *24*, 69–71. [[PubMed](#)]
36. Foyo-Moreno, I.; Alados, I.; Alados-Arboledas, L. A New Conventional Regression Model to Estimate Hourly Photosynthetic Photon Flux Density under All Sky Conditions. *Int. J. Climatol.* **2017**, *37*, 1067–1075. [[CrossRef](#)]
37. Du, Y.C.; Stephanus, A. Levenberg-Marquardt Neural Network Algorithm for Degree of Arteriovenous Fistula Stenosis Classification Using a Dual Optical Photoplethysmography Sensor. *Sensors* **2018**, *18*, 2322. [[CrossRef](#)] [[PubMed](#)]
38. Kottek, M.; Grieser, J.; Beck, C.; Rudolf, B.; Rubel, F. World Map of the Köppen-Geiger Climate Classification Updated. *Meteorol. Z.* **2006**, *15*, 259–263. [[CrossRef](#)]