Energy Conversion and Management 51 (2010) 881–893

Contents lists available at ScienceDirect

Energy Conversion and Management

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

Proposal of a regressive model for the hourly diffuse solar radiation under all sky conditions

J.A. Ruiz-Arias *, H. Alsamamra, J. Tovar-Pescador, D. Pozo-Vázquez

Department of Physics, Building A3-066, University of Jaén, 23071 Jaén, Spain

ARTICLE INFO

Article history:
Received 14 April 2009
Accepted 21 November 2009
Available online 21 December 2009

Keywords:
Solar radiation
Diffuse fraction
Regressive model
Logistic model

ABSTRACT

In this work, we propose a new regressive model for the estimation of the hourly diffuse solar irradiation under all sky conditions. This new model is based on the sigmoid function and uses the clearness index and the relative optical mass as predictors. The model performance was compared against other five regressive models using radiation data corresponding to 21 stations in the USA and Europe. In a first part, the 21 stations were grouped into seven subregions (corresponding to seven different climatic regions) and all the models were locally-fitted and evaluated using these seven datasets. Results showed that the new proposed model provides slightly better estimates. Particularly, this new model provides a relative root mean square error in the range 25–35% and a relative mean bias error in the range −15% to 15%, depending on the region. In a second part, the potential global character of the new model was evaluated. To this end, the model was fitted using the whole dataset. Results showed that the global fitting model provides overall better estimates that the locally-fitted models, with relative root mean square error values ranging 20–35% and a relative mean bias error ranging −5% to −12%. Additionally, the new proposed model showed some advantages compared to other evaluated models. Particularly, the sigmoid behaviour of this model is able to provide physically reliable estimates for extreme values of the clearness index even though using less parameter than other tested models.

1. Introduction

The amount of solar radiation available on the tilted surface is a key factor in numerous solar energy applications, as thermal and photovoltaic energy systems or self-sustainable buildings. Additionally, it is known that solar radiation rules the life by providing the energy required in many natural process, as the photosynthesis [1,2]. For this reason, this parameter is also important in environmental studies, particularly in those areas, as most part of natural parks, where the usual complex topography may drastically change the incoming solar radiation received on the horizontal surface on the ground [3–5].

To compute the amount of solar radiation striking a tilted surface, the beam and diffuse components of the radiation are needed since the different physical nature of these components determines how they are projected on the surface. Nevertheless, the direct measurement of beam or diffuse components requires complex devices with a high maintenance cost. This explains the scarce availability of this kind of solar radiation data compared to the horizontal solar global radiation. As a consequence, several approaches have been proposed to derive solar radiation components datasets.

As it is known, the interaction between the atmosphere and the solar radiation coming from the Sun, which results in a decomposition into the solar radiation components, is highly complex. The complete modelling of this interaction has been only recently possible, through the use of Numerical Weather Prediction (NWP) models, as for example, the Weather Research and Forecasting Model (WRF) [6]. Particularly, these models simulate the interaction of the solar beam with the atmosphere providing the solar fluxes at the surface as one of their outputs. Spatial resolution of these models can reach up to 1 km in reduced areas, making possible the assessment of the global solar radiation in terms of its beam and diffuse components. The main problem related to the NWP models is their relative complexity and time-demanding computational cost. Another alternative to calculate the solar radiation components is the use of a broadband solar radiation model (many of them integrated within a Geographical Information System) [7–11]. These are a simplified modelization of the atmosphere that only takes into account its radiative properties providing relatively accurate estimates under clear-sky conditions. However, its reliability under other sky conditions is considerable lower.

All these methodologies rely mainly on the physical processes across the atmosphere in order to estimate the solar radiation

* Corresponding author.
E-mail address: jararias@ujaen.es (J.A. Ruiz-Arias).

0196-8904/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.enconman.2009.11.024

components. Nevertheless, another different approach is the use of statistical relationships based on observed databases. The aim is to obtain statistical models that relate measured global, diffuse and/or direct solar radiation data [12,13]. Sometimes, other variables are also taken into account within these statistical models [14]. The statistical approach represents in an easier manner the complex processes that the solar beam suffers in its path trough the atmosphere, although giving up the potential reliability of more sophisticated models as with NWP models, for example. In this context, the hourly interval offers an appropriate agreement between data availability and the inherent solar radiation temporal variability. As a consequence, most of the statistical models are based on the hourly interval of the solar radiation data.

The pioneer study in relating the global radiation and its diffuse counterpart was the Liu and Jordan [15] work, using daily values collected in Massachussets, USA. Orgill and Hollands [16] proposed a regression equation between the hourly clearness index and the hourly diffuse fraction based on data collected during four years in Toronto, Canada. Erbs et al. [17], using data from four stations in the USA, proposed another regression that was validated with three years of data in Australia. Reindl et al. [14] introduced new predictors in the relation. They began studying a set of 28 predictors and, finally, reducing the set to four: clearness index, sine of the solar altitude, monthly mean hourly ambient temperature and monthly mean hourly ambient relative humidity fraction. They trained the model using data collected at five stations (one in the USA and the rest in Europe), and validated the results using data collected at Albany, NY, USA.

The former models are based on piecewise regression equations, which divide the clearness index range into different intervals depending on the study. The diffuse fraction is then fitted for each interval with a polynomial function of a given order. Particularly, Orgill and Hollands [16] and Reindl et al. [14] used a first-order polynomial for the intermediate clearness index range while Erbs et al. [17] used a fourth-order polynomial. More recently, some authors have proposed new regression equations that are not defined as a piecewise function of the clearness index. Particularly, Muneer and Munawwar [18] studied the influence of the synoptic variables, sunshine fraction, cloud cover and optical air mass and their potential ability to improve the estimation of the diffuse fraction. This work had its continuity in Muneer et al. [19]. In the later, the authors proposed to use second-order polynomials for the clearness index and linear or squared relationships for the synoptic variables. This study was carried out using nine stations spread out over India, Japan, Spain and United Kingdom. Clarke et al. [20], using 5 min data collected in Edimburgh during the period 1993–1994, also proposed a quadratic or cubic polynomial correlation for all the range of variation of the clearness index. Nevertheless, the "classical" approach of splitting up the clearness index into intervals has been also recently followed by Jacovides et al. [21], for example. Based on data collected in Cyprus, they proposed to use a third-order polynomial for the intermediate range of the clearness index. A review of some of the existing models for the Mediterranean area is provided by Notton et al. [22]. It is also worth to mention the very interesting works of Perez et al. [23] and Skartveit et al. [24] where the authors also consider the hourly variability of the solar global irradiation in the estimation of the solar radiation components. This is a very convenient approach to somehow include the effect of the clouds. Ineichen [13] has recently published a comparison of these models.

In this work, we present a regressive model based on the sigmoid function to obtain the diffuse fraction with the clearness index and the pressure-corrected optical air mass as predictor variables. The use of more predictors could improve the performance of the model [12,14,18], but most of these predictors (usually synoptic variables) are not easily available everywhere. Therefore, in this work it has been rather preferred to not include any synoptic predictor making the model more general. The model's performance has been compared against that of other models proposed in the bibliography. For a fair comparison, these later models have been re-calibrated based on the dataset used in this study. The work is organized as follows: the dataset used in the study and the quality control procedure applied to the radiation data are described in Section 2. Section 3 presents the here proposed and the other tested models. Section 4 presents the results of the models evaluation and, in Section 5, an additional analysis is presented evaluating the potential global character of the proposed model. Finally, a summary and the conclusions of the work are presented in Section 6.

2. Database

The dataset used in this work consists of global and diffuse solar radiation measurements from 21 locations spread over Europe and the United States (Fig. 1). Particularly, the European part of the dataset is comprised by six stations: three of them in Germany and the rest in Spain. The German stations belong to the European Solar Radiation Atlas database (ESRA) [25], with hourly data covering the period from 1981 to 1990. The measurements have been collected using pyranometers to measure the hemispherical horizontal solar irradiance and solar trackers equipped with shading disks to shield the direct component and record the diffuse irradiance. The uncertainty of the practical pyrheliometers used to register the direct beam irradiance is in the best case the 3% whereas the lowest uncertainty of the pyranometers increases up

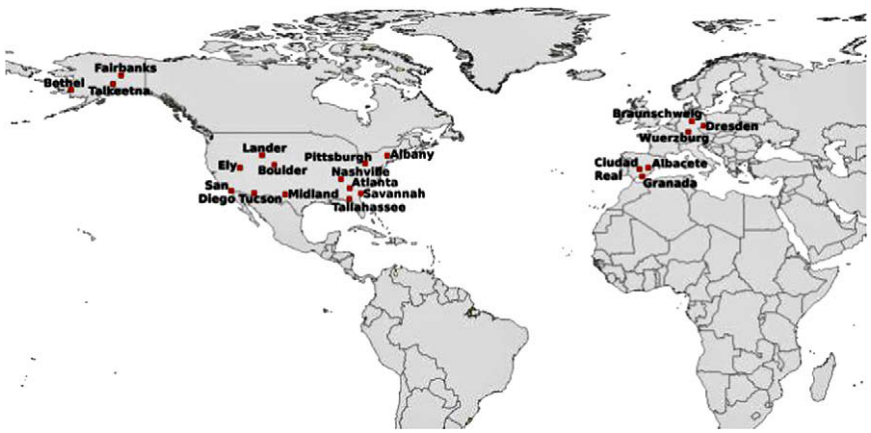

Fig. 1. Geographical location of the stations.

to 5%. Nowadays, in the best case, the uncertainty in the diffuse horizontal irradiance is on the order of 3% ±2 W m^{-2} but it may increase up to 15–20% if the measure has been taken using a shielding shadow band [26].

The Spanish stations pertain to the Agencia Estatal de Meteorología (AEMET), which is responsible for their maintenance. The data has been collected using pyranometers Kipp & Zonen CM11 equipped with shadow bands for registering the diffuse irradiance. After quality control, a shadow band correction procedure was applied. The correction algorithm was validated in Southern-Spain by López et al. [27] in a region with pretty similar climatic characteristics to that of the Spanish stations used in this work (Section 2.1). The authors compared four well-known shadow band diffuse irradiance correction algorithms with data collected also with a Kipp & Zonen CM11 pyranometer in the meteorological station at the University of Almeria. They reported a final root mean square error of 13% and a mean bias error of −5% in the diffuse component which might be extrapolated to the three Spanish stations.

The rest of stations of the dataset, located in the USA, are part of the National Solar Radiation Data Base (NSRDB) and cover the period 1961–1990. In the nomenclature of the NSRDB, all these stations are referred as primary stations given that they contain measured solar radiation data for at least a portion of the 30 years record. In the case of Alaska, it has been used two secondary stations (only containing modelled data) at Bethel and Talkeetna. This decision was adopted in order to have three stations in this region: one for calibrating the model and the other two for validation purposes. Ideally, the use of modelled data to derive a new model should be avoided. Nevertheless, the NSRDB offers desirable properties as homogeneity and a long enough period of measurements for climatological studies (the World Meteorological Organization recommends the use of 30-year periods of data). Table 1 shows the proportion of measured and modelled daylight hours as well as the mean estimated uncertainty level for each radiation component for the NSRDB stations here used. This table reveals a mean approximated uncertainty between 9% and 13%, regardless the percentage of modelled data and the solar radiation component. The uncertainty is here defined as the interval around a measured or modelled data value within which the true value will lie 95% of the time [28]. Almost the entire dataset is flagged with uncertainty flags from 4 to 6 which, according to NREL [28] Section 3.4, that roughly correspond to an uncertainty ranging from 6% to 18%. The mean estimated uncertainty in Table 1 has been calculated averaging the flag values, interpreted as continuous values, and then the associated uncertainty estimated by linear interpolation. For example, "mean uncertainty flags" for the station at Boulder are 4.33, 4.87 and 3.92 for the horizontal global, diffuse and direct normal irradiances, respectively (uncertainty flag 3 corresponds to an uncertainty of 4–6%, flag 4 to 6–9% and flag 5 to 9–13%). Linear interpolation yields uncertainties of 10.6%, 12.1% and 9.5%, respectively. The relative low level of uncertainty reported in the NSRDB and the formerly commented properties encourage the use of this dataset in the present work. Obviously, the final uncertainty of the derived models will result of the composition of the initial uncertainty of the data and the intrinsic uncertainty of the models and the methodology.

Table 1
Proportion of measured (M) and modelled (P) daylight hours (in percent) of the NSRDB selected stations and mean approximated uncertainty (U, in percent) based on the quality flags of the data. Measured data include that derived with the closure relation (flag D, see Section 3.4 in NREL [28]). For a detailed explanation on the calculation of the mean estimated uncertainty, see Section 2.

	Global			Direct			Diffuse		
	M	P	U	M	P	U	M	P	U
Tallahassee	15.0	85.0	11.3	12.9	87.1	10.1	12.9	87.1	12.6
Savannah	16.6	83.4	11.4	15.3	84.7	10.8	15.3	84.7	12.5
Atlanta	2.9	97.1	11.0	2.9	97.1	9.7	2.9	97.1	12.5
Midland	10.2	89.8	11.2	8.3	91.7	9.9	8.3	91.7	12.6
Tucson	7.4	92.6	11.2	5.8	94.2	10.0	5.7	94.3	12.6
San Diego	3.2	96.8	11.1	2.9	97.1	10.2	2.9	97.1	12.7
Nashville	50.5	49.5	10.3	14.3	85.7	9.7	14.3	85.7	12.4
Pittsburgh	8.5	91.5	11.3	6.6	93.4	10.0	5.3	94.7	12.7
Albany	9.1	90.9	11.2	9.1	90.9	9.8	9.1	90.9	12.5
Boulder	29.5	70.5	10.6	27.6	72.4	9.5	27.6	72.4	12.1
Ely	35.5	64.5	11.2	17.8	82.2	10.2	17.4	82.6	12.7
Lander	26.6	73.4	11.4	23.3	76.7	10.2	23.3	76.7	12.8
Bethel	0.0	100.0	11.6	0.0	100.0	11.4	0.0	100.0	13.1
Talkeetna	0.0	100.0	11.7	0.0	100.0	11.5	0.0	100.0	13.1
Fairbanks	2.3	97.7	11.8	0.0	100.0	10.8	0.0	100.0	12.7

Given the long time period registered and the wide region covered by the database, several different methodologies and instrumentation have been employed along the data recording process, from modelled data values using other meteorological variables to Eppley pyranometers. A brief history of the solar radiation measurements can be found in the user's manual of the database [28].

Additionally, it is worth to remark that some authors have recommended the use of global solar irradiance calculated from direct and diffuse measured irradiances as Michalsky et al. [29]. Particularly, the model considered here is better derived using only these components. Nevertheless, the length of the available high quality series of simultaneously measured direct and diffuse irradiances is scarce. Then, the measured global solar irradiance with un-shaded pyranometer is an assumable alternative.

Table 2 shows the main geographic characteristics of the stations. Particularly, the stations cover a wide range of latitudes within the northern hemisphere, ranging from 30.38°N (Tallahassee, USA) to 64.82°W (Fairbanks, USA). Stations also span a considerable range of terrain heights: from sea level (e.g., San Diego, USA) to almost 2000 m (Ely, USA). Additionally, Table 2 presents details about the climatic conditions of the stations locations, according to the Koeppen climatic classification. This climatic scheme divides the climate in five main types and some subtypes, based mainly on mean temperature and precipitation values. Each particular climate is symbolized by 2–4 letters. The work of Peel et al. [30] has been used to elucidate the climate of each one of these sites. Within the 21 stations, nine have B type classification, meaning an arid climate. Five stations present a temperate climate, symbolized with the C type, and the rest of stations a cold climate (D type). Therefore, it can be concluded that the stations represent a considerable range of climatic conditions.

2.1. Quality control procedure

With the aim of homogenizing the complete datasets involved in the study, all the data values have been subjected to the quality control procedure described in Younes et al. [31] that, according to these authors, may be used with equal effectiveness for any terrestrial dataset. Note that the SERI QC quality control procedure has been already applied to the NSRDB, and the application of a new quality control should not be inconsistent but rather involving an eventually more restrictive procedure. Following is detailed the four-step quality control here applied.

2.1.1. First test

This test deals with the intrinsic cosine error of the pyranometric sensors. As Younes et al. [31] recommend, all the data values corresponding to a solar altitude α below 7° have been rejected:

$$\alpha \geqslant 7.0^{\circ}. \quad (1)$$

2.1.2. Second test

This is a physical limit imposed to the value of the hemispherical horizontal global solar irradiance, I_G, and the horizontal diffuse

Table 2
Local features of the station locations: geographical situation, elevation, Koeppen's climate index, measurement period and number of data (daylight hours). Station height is given in meters above mean sea level.

Country	Location	Region	Latitude	Longitude	Height	Koeppen's climate	Period of data	Number of records
Spain								
	Granada	Spain	37.14°N	3.63°W	687	BSk	2002–06	10,181
	Ciudad Real	Spain	38.99°N	3.92°W	627	BSk	2002–06	6706
	Albacete	Spain	39.00°N	1.86°W	674	BSk	2002–06	8277
Germany								
	Wuerzburg	Germany	49.77°N	9.97°E	275	Dfb	1981–90	20,087
	Dresden	Germany	51.12°N	13.68°E	246	Dfb	1981–90	24,219
	Braunschweig	Germany	52.30°N	10.45°E	83	Dfb	1981–90	18,323
USA								
	Tallahassee	South-Eastern	30.38°N	84.37°W	21	Cfa	1961–90	79,910
	Savannah	South-Eastern	32.13°N	81.2°W	16	Cfa	1961–90	85,608
	Atlanta	South-Eastern	33.65°N	84.43°W	315	Cfa	1961–90	75,436
	Midland	South-Western	31.93°N	102.20°W	871	BSh	1961–90	68,149
	Tucson	South-Western	32.12°N	110.93°W	779	BWh	1961–90	82,454
	San Diego	South-Western	32.73°N	117.17°W	9	BSk	1961–90	91,482
	Nashville	North-Eastern	36.12°N	86.68°W	180	Cfa	1961–90	85,857
	Pittsburgh	North-Eastern	40.50°N	80.22°W	373	Dfa	1961–90	84,626
	Albany	North-Eastern	42.75°N	73.80°W	89	Dfb	1961–90	87,503
	Boulder	Western	40.02°N	105.25°W	1634	BSk	1961–90	90,411
	Ely	Western	39.28°N	114.85°W	1906	BWk	1961–90	79,674
	Lander	Western	42.82°N	108.73°W	1696	BSk	1961–90	91,086
	Bethel	Alaska	60.78°N	161.80°W	46	Dfc	1961–90	60,182
	Talkeetna	Alaska	62.30°N	150.10°W	105	Dsc	1961–90	60,803
	Fairbanks	Alaska	64.82°N	147.87°W	138	Dwc	1961–90	67,959

solar irradiance, I_D. The limits are based on the clearness index, k_t, and the diffuse fraction, k, defined as:

$$k_t = \frac{I_G}{I_0 \cos Z}, \tag{2}$$

$$k = \frac{I_D}{I_G}, \tag{3}$$

where I_0 is the extraterrestrial direct irradiance and Z is the solar zenith angle. For the sake of completeness, it is also important to define the direct fraction, F_b, as the ratio of the horizontal direct solar irradiance, I_B, to the horizontal global solar irradiance:

$$F_b = \frac{I_B}{I_G}. \tag{4}$$

According to these definitions, both the clearness index (k_t) and the diffuse fraction (k) must verify the following conditions:

$$0 < k_t < 1, \tag{5}$$

$$0 < k < 1. \tag{6}$$

These constraints could have been relaxed since, according to some authors, the clouds albedo could increase the global radiation beyond the extraterrestrial, yielding a clearness index slightly greater than one. However, this is an exceptional situation and it was preferred the data to verify Eqs. (5) and (6).

2.1.3. Third test

In this step, a maximum value is imposed to the horizontal global solar irradiance and maximum and minimum values are imposed to the diffuse solar irradiance [32,33]. These boundary values are calculated using the model of Page [7,25], which parameterizes the sky with the air mass 2 Linke turbidity factor. The maximum value of the horizontal global solar irradiance and the minimum value of the diffuse solar irradiance are calculated using an air mass 2 Linke turbidity equals to 2.5 (extremely clear sky) whereas the maximum value of the diffuse solar irradiance is calculated with an air mass 2 Linke turbidity 572 times the solar altitude (heavily overcast sky) expressed in radians. Therefore:

$$I_G \leqslant I_{G,C}, \tag{7}$$

$$I_{D,C} \leqslant I_D \leqslant I_{D,OC}, \tag{8}$$

where $I_{G,C}$ and $I_{D,OC}$ are the maximum horizontal global and diffuse solar irradiances, respectively, and $I_{D,C}$ is the minimum diffuse solar irradiance estimated with the model of Page. This test has been successfully evaluated by Younes et al. [31] in 11 locations in the northern hemisphere from England to Japan.

2.1.4. Fourth test

This is essentially a statistical outlier analysis. The whole k_t range (from 0 to 1) was split into ten equal-size intervals. For every interval, the mean and standard deviation of the diffuse fraction were calculated. Values below and above twice standard deviations from the mean were removed. Occasionally, this limit was slightly modified to take into account local features of some stations.

Fig. 2 shows the k–k_t scatter plot for each quality control step for the station located in Boulder, USA. The filtering process in this station reduced the number of records up to 90,411 (63.3% of available daylight records). Note that test 3 rejects several outliers for low clearness index values, albeit at the same time, it seems to also eliminate some, a priori, good points for intermediate clearness indices and low diffuse fractions. The amount of these points varies from one station to another. Overall, given the high amount of data points, the application of test 3 has been considered statistically positive.

Finally, the shadow band correction factor procedure proposed by Muneer and Zhang [34] was applied to the diffuse solar irradiance measured in the Spanish stations. It was selected among other available methodologies because it has been successfully tested by López et al. [27] in Almería, close to the Spanish stations. The authors used data measured with Kipp & Zonen CM11 pyranometers, one of them equipped with a shadow band. The RMSE of the measured and corrected data was reduced in a 10% up to 12.9% and the bias was reduced (in module) in a 16% up to −5%. These results can be extrapolated to the data in the Spanish stations here used.

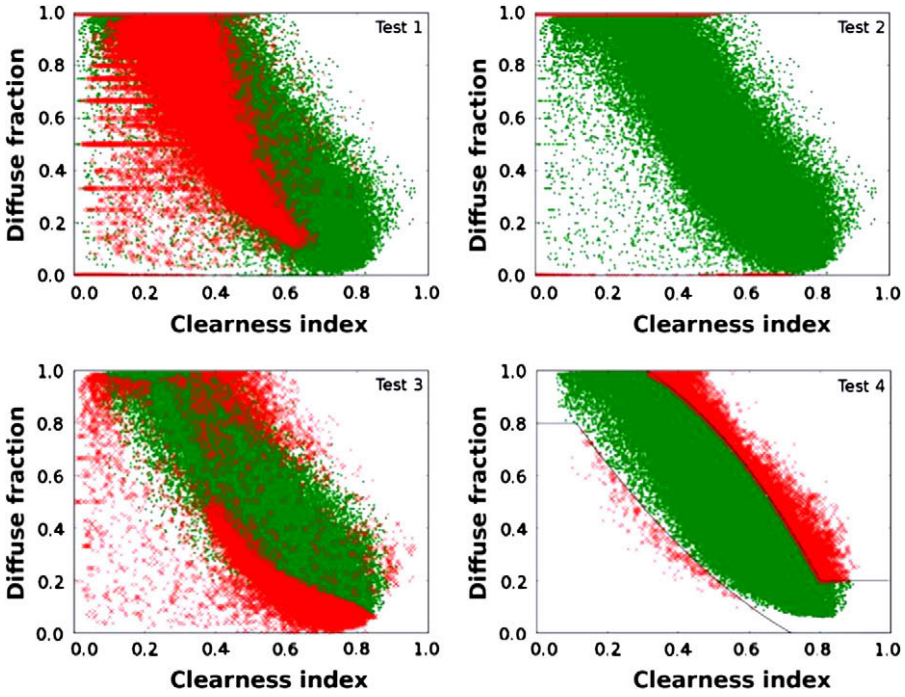

Fig. 2. Quality control procedure for the station located in Boulder (USA). The rejected data points are marked with red-cross symbols and the green dots are the points that pass the tests. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. Models and methodology

3.1. The proposed model

Traditionally the regression equation has related the diffuse fraction with the clearness index, instead of relating the global solar radiation with its diffuse component. The reason is that, on the one hand, seasonal and diurnal variations of solar irradiance are driven by well-established astronomical relationships. But, on the other hand, the solar irradiance shows considerable stochastic short-time variability, ruled by less predictable variables as frequency and height of the clouds and their optical properties, aerosols, ground albedo, water vapour or atmospheric turbidity [35]. As a consequence, the solar irradiance can be considered as the sum of two components: one deterministic and one stochastic. The stochastic component can be isolated using the clearness index and the diffuse fraction.

Fig. 3 shows the scatter plot of k against k_t for the station located in Boulder (USA) along with three curves resulting from fitting the data using a second-order polynomial (P2), a third-order polynomial (P3) and a sigmoid (or logistic) curve (G0). The latter is a real-valued and differentiable curve, with either a non-negative or non-positive first derivative and one inflection point. In this case, the curve has the functional form $1 - a_0 \exp[a_1 \exp(a_2 t)]$ which is based on the so-called Gompertz curve.

Note that, for strong overcast conditions, it would be physically expected that $k \to 1$ as $k_t \to 0$. Among the curves in Fig. 3, this characteristic is only partially attained by the sigmoid curve, since the second-order polynomial takes values greater than one and the third-order polynomial decreases for small clearness indices. On the other hand, for clear days, the diffuse fraction is expected to tend to small values, albeit strictly never equals to zero (note that even a completely clear atmosphere will scatter some amount of solar radiation). Again, among the curves in Fig. 3, this condition is only partially fulfilled by the sigmoid curve, since the second-order polynomial predicts a negative diffuse fraction and the third-order polynomial predicts an increase. On the intermediate range of clearness index values, the three curves behave very similarly.

A special characteristic of the sigmoid curve is that there is no need to break the hourly diffuse fraction down into intervals as function of the clearness index, the "classical" approach [14–17,21]. This is a desirable feature given that the introduction of breaking points in the regression definition may increase the local dependency of the model.

In spite of these interesting properties, there is scarce bibliography using the sigmoid curve in the solar radiation modelling field. Particularly, a sigmoid curve was proposed by Boland and Ridley

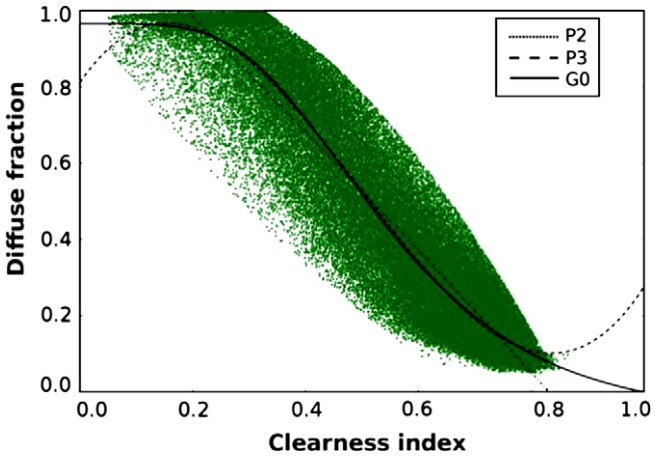

Fig. 3. Scatter plot for the station located in Boulder (USA) (green dots) and fitting curves using a second-order polynomial (P2), a third-order polynomial (P3) and a sigmoid curve (G0). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

[36], who used a logistic model to fit hourly and fifteen minutes data collected in Geelong, Australia. Recently, the same authors [37] have published a work where they evaluate the performance of the sigmoid model and they also study its possible applicability for any place.

Additionally, a study has been carried out in order to elucidate the type of sigmoid curve most suitable to regress the hourly diffuse fraction and the clearness index. The solar beam suffers several attenuation processes on its way through the atmosphere, as Rayleigh scattering or absorption by uniformly mixed gases, water vapour or aerosols. Assuming that all extinction processes occur independently of each other within narrow spectral regions, the incident beam spectral irradiance at normal incidence, $E_{bn,\lambda}$, is then obtained as:

$$E_{bn,\lambda} = \prod_i \tau_{i\lambda} E_{0n,\lambda}, \tag{9}$$

where $\tau_{i\lambda}$ is the spectral transmittance of the ith attenuation process and $E_{0n,\lambda}$ is the extraterrestrial spectral irradiance at wavelength λ at the actual Sun–Earth distance [38].

By similarity with Eq. (9), at broadband scale, the whole attenuation process can be expressed as:

$$E_{bn} = \prod_i \tau_i E_{0n}, \tag{10}$$

where E_{bn} and E_{0n} are obtained by integrating $E_{bn,\lambda}$ and $E_{0n,\lambda}$ over all the wavelength spectrum, and τ_i is the broadband transmittance for the ith attenuation process. Again, by similarity with the Bouguer's law (applied strictly only to monochromatic radiation), the broadband transmittance can be expressed as:

$$\tau_i = \exp(-m_i\delta_i), \tag{11}$$

where m_i is the optical mass and δ_i the broadband optical depth for the ith process [39].

Additionally, and similarly to the diffuse fraction, the direct beam fraction (F_b) is defined as the ratio of horizontal direct beam irradiance to the horizontal global irradiance. Therefore, considering the global irradiance as the contribution of the direct and diffuse components, F_b and k are related by means of the following expression:

$$F_b + k = 1. \tag{12}$$

From the definition of the direct beam fraction, Eq. (2), and the clearness index, Eq. (4), the direct beam fraction can be re-written as:

$$F_b = \frac{1}{k_t}\frac{E_{bn}}{E_{0n}} = \frac{1}{k_t}\exp\left(-\sum_i m_i\delta_i\right). \tag{13}$$

Removing F_b in Eq. (12):

$$k = 1 - \frac{1}{k_t}\exp\left(-\sum_i m_i\delta_i\right). \tag{14}$$

If we consider the Taylor series expansion of the factor k_t^{-1},

$$k_t^{-1} = \sum_{n=0}^{\infty}(1-k_t)^n, \quad |1-k_t| < 1, \tag{15}$$

and we only take the first term of the expansion ($n = 0$), Eq. (14) can be written, in first approximation of the Taylor expansion, as:

$$k = 1 - \exp\left(-\sum_i m_i\delta_i\right). \tag{16}$$

Let consider now the atmosphere as a background Rayleigh atmosphere and the superimposed effects of the rest of atmosphere constituents (water vapour and aerosols, principally). Then Eq. (16) can be given as:

$$\begin{aligned} k &= 1 - \exp\left[-m_R\delta_R - \sum_j m_j\delta_j\right] = 1 - \exp\left[-m_R\delta_R(1+\varepsilon)\right] \\ &= 1 - \exp\{-\exp[\ln(m_R\delta_R) + \ln(1+\varepsilon)]\}, \end{aligned} \tag{17}$$

where ε is the ratio of the attenuation coefficient by atmospheric constituents not included in the Rayleigh atmosphere to that of the Rayleigh atmosphere. It represents the relative noise produced over the Rayleigh atmosphere by the rest of atmospheric constituents. If the noise is assumed to be on the order of the Rayleigh background, the natural logarithm in Eq. (17) can be expanded in polynomials such as

$$k = 1 - \exp\left\{-\exp\left[const. + \varepsilon - \frac{1}{2}\varepsilon^2 + \frac{1}{3}\varepsilon^3 - \cdots\right]\right\}. \tag{18}$$

The functional form of Eq. (18) resembles a regression equation of the type:

$$K(x_i) = a - b\exp\{-\exp[F(x_i)]\}, \tag{19}$$

where $F(x_i)$ is a polynomial of the predictors x_i. This functional form also resembles the sigmoid Gompertz's curve and x_i represents the predictors of the diffuse fraction k.

In the present work, we evaluate the use of different versions of the sigmoid function in Eq. (19) to fit the hourly diffuse fraction.

3.2. Brief discussion on the number of predictors

As some authors suggest [14,19], for a given clearness index and, especially, for intermediate values, the range of possible values of the diffuse fraction is too wide to use a regression equation with only the clearness index as predictor. For instance, for a clearness index of 0.5 and, as can be observed in Fig. 3, the interval of diffuse fraction ranges approximately from 0.3 to 0.8. Therefore, it would be probably useful to include other predictors as temperature, humidity, sunshine fraction, cloud cover or optical air mass [12,14,18,36]. The main problem arises because these predictors (usually synoptic variables) are not always available. Given that the main aim of the statistical models is to easily estimate the diffuse fraction from measurements available in radiometric stations, in this work it has been rather preferred to not include any synoptic predictor, making the model more general. It is then assumed a certain reduction of the possible accuracy in the results as toll for an easily applicable model. Consequently, to evaluate the proposed model performance, we have only tested models that use the clearness index alone or combined with the relative optical air mass. Other authors, as Reindl et al. [14], use the sine of the solar height, which is related to the optical air mass. In this work, we have rather preferred to use just the later variable, because its closer relationship to the attenuation processes in the atmosphere.

Another very interesting approach adopted by some authors [23,24] has been the inclusion of the short-term hourly variability of the irradiance as an estimator of the cloudiness. This approach has proven to improve the performance of the model without include new synoptic variables at the expense of increasing the complexity of the model.

3.3. The tested models

Four models recently appeared in the bibliography of diffuse fraction regression models, two of them using the clearness index alone and two using the optical air mass as well, together with the "classical" model of Reindl et al. [14], have been tested against three different versions of the sigmoid-function-based model. This totalizes eight tested models, described below.

3.3.1. Models using only the clearness index as predictor

- Second-order polynomial, used in Clarke et al. [20], hereinafter referred as P2:

$$k(k_t) = a_0 + a_1 k_t + a_2 k_t^2. \quad (20)$$

- Third-order polynomial, used in Clarke et al. [20], hereinafter referred as P3:

$$k(k_t) = a_0 + a_1 k_t + a_2 k_t^2 + a_3 k_t^3. \quad (21)$$

- The model proposed by Reindl et al. [14], hereinafter referred as R:

$$k(k_t) = \begin{cases} 1.020 - 0.248 k_t, & 0.0 \leqslant k_t \leqslant 0.3 \\ 1.450 - 1.670 k_t, & 0.3 \leqslant k_t \leqslant 0.78. \\ 0.147, & 0.78 \leqslant k_t \end{cases} \quad (22)$$

- New model here proposed, based on a linear dependency with k_t in the sigmoid function, hereinafter referred as G0:

$$k(k_t) = a_0 - a_1 \exp[-\exp(a_2 + a_3 k_t)]. \quad (23)$$

3.3.2. Models using the clearness index and height-corrected optical air mass as predictors

- New model here proposed, based on a linear dependency with k_t and m in the sigmoid function, hereinafter referred as G1:

$$k(k_t, m) = a_0 - a_1 \exp[-\exp(a_2 + a_3 k_t + a_4 m)]. \quad (24)$$

- New model here proposed, based on a quadratic dependency with k_t and m in the sigmoid function, hereinafter referred as G2:

$$k(k_t, m) = a_0 - a_1 \exp[-\exp(a_2 + a_3 k_t + a_4 k_t^2 + a_5 m + a_6 m^2)]. \quad (25)$$

- Regression equation proposed in Clarke et al. [20], hereinafter referred as M1:

$$k(k_t, m) = (a_0 + a_1 m) + (a_2 + a_3 m) k_t + (a_4 + a_5 m) k_t^2. \quad (26)$$

- Regression equation proposed in Clarke et al. [20], hereinafter referred as M2:

$$k(k_t, m) = (a_0 + a_1 m + a_2 m^2) + (a_3 + a_4 m + a_5 m^2) k_t + (a_6 + a_7 m + a_8 m^2) k_t^2. \quad (27)$$

3.4. Analysis of the models performance

In order to assess the performance of the different models, a number of statistical scores have been computed.

- the squared coefficient of correlation (r^2) between modelled and measured diffuse fraction values, which represents the proportion of the linear variability "explained" by the model. It has been assessed as

$$r^2 = \frac{\left[\sum_t (p_i - \bar{p})(m_i - \bar{m})\right]^2}{\sum_i (p_i - \bar{p})^2 \sum_i (m_i - \bar{m})^2}, \quad (28)$$

  where p_i is the ith predicted diffuse fraction data point, m_i is the ith measured diffuse fraction data point, $\bar{p}$ is the predicted mean value and $\bar{m}$ is the measured mean value. It ranges from 0 to, ideally, 1 for a perfect linear relationship.

- The mean bias error (MBE), which measures the systematic error of the model. It evaluates the tendency of the model to under- or over-estimate the measured values. Here, we have used the relative MBE (rMBE) to the measured mean value, obtained as follows:

$$\mathrm{rMBE} = 100 \frac{\sum_i (p_i - m_i)}{\sum_i m_1}. \quad (29)$$

- The root mean squared error (RMSE), that estimates the level of scattering of the predicted values, was also computed. Again, we have used the relative RMSE (rRMSE) to the measured mean value:

$$\mathrm{rRMSE} = 100 \frac{\sqrt{N \sum_i (p - i - m_i)^2}}{\sum_i m_i}. \quad (30)$$

- Also the skewness and the kurtosis of the error distribution have been computed. They measure, respectively, the level of asymmetry and the peakedness of the error distribution with respect to a normal distribution.
- Finally, we have also computed and accuracy score (AS), to easily compare the overall model's performance. The accuracy score allows elucidating the best behaved model attending to the statistics used in its definition. In this case, we have calculated the score as:

$$\begin{aligned} AS = {} & 0.24 \frac{r_i^2 - r_{\min}^2}{r_{\max}^2 - r_{\min}^2} + 0.24 \left[1 - \frac{|\mathrm{rMBE}|_i - |\mathrm{rMBE}|_{\min}}{|\mathrm{rMBE}|_{\max} - |\mathrm{rMBE}|_{\min}}\right] \\ & + 0.24 \left[1 - \frac{\mathrm{rRMSE}_i - \mathrm{rRMSE}_{\min}}{\mathrm{rRMSE}_{\max} - \mathrm{rRMSE}_{\min}}\right] \\ & + 0.14 \left[1 - \frac{|\mathrm{Skewness}|_i - |\mathrm{Skewness}|_{\min}}{|\mathrm{Skewness}|_{\max} - |\mathrm{Skewness}|_{\min}}\right] \\ & + 0.14 \frac{Kurtosis_i - Kurtosis_{\min}}{Kurtosis_{\max} - Kurtosis_{\min}}. \end{aligned} \quad (31)$$

  Note that different weights were applied to each addend, being the sum of them equals to 1. Particularly, the same weight (0.24) has been selected for the correlation coefficient, the rMBE and the rRMSE, whereas a lower value (0.14) was used for the skewness and the kurtosis. The rationale behind this is to assign a greater relative importance to the firsts over the skewness and the kurtosis. According to this definition, the maximum AS value is 1 and the greater the score the better the model. It is worth to remark that AS is only useful for the inter-comparison of the involved set of models. It must be re-calculated if the set of models changes or some of them is modified.

Additionally, the Akaike's Information Criterion (AIC) has been also provided in the local inter-comparison step of the models. The AIC is a model selection score based on the Kullback–Leibler information loss and closely related to the concept of entropy. It describes the trade-off between precision and complexity of the model. In the special case of least squares estimation with normally distributed errors and the number of experimental points far larger than the number of predictors, AIC can be calculated as

$$\mathrm{AIC} = n \log\left(\frac{\sum_i (p_i - m_i)^2}{n}\right) + 2K, \quad (32)$$

where n is the number of experimental points and K the number of predictor variables [40]. According to this score, the smaller the AIC the better the model.

4. Evaluation of the models

The evaluation process was carried out in three steps. In a first step, the 21 stations were grouped into seven regions, namely Spain, Germany, South-Western USA, Western USA, North-Eastern USA, South-Eastern USA and Alaska (Table 3). These regions were selected to contain three stations and represent different climatic conditions. At each region, one station was used to train the models and the other two for an independent validation process. Another common method of validation consists on the use of a certain portion of the experimental time series (previously re-

moved from the original data). Given that the record length of the Spanish stations is only 5 years long we have used the validation procedure based on independent stations.

In a second step, seven local regression analyses (one per region) for each model (Eqs. (20), (21), (23)–(27)) were carried out. This local treatment allowed the parameters of the models to account for the local climatic and geographic features of the seven regions under study, making the evaluation straightforward and fair. The different models parameters are presented in Table 4 (except for model R which has fixed parameters) whereas the models performance evaluation, in terms of the scores defined in Eqs. (28)–(32), is presented in Table 5. Overall, it is concluded that those models which use k_t and m in the regression generally present a slightly better fit than those that use k_t alone. But, according to the AIC, this enhancement of the model's performance shouldn't be enough as to resolutely conclude that the use of the optical mass overall improves the model. In terms of the AIC, the increase of the model's complexity after including the optical mass as a second predictor is "greater" than the enhancement of the model's performance. However, in the case of the optical mass, its inclusion into the regression equation does not imply any extra input information or effort, provided that it can be readily calculated with the same information needed to assess the clearness index given the horizontal global solar irradiance. Therefore, in spite of the AIC values, the use of the models with k_t and m as predictors would be justified. Note that, since the data has been fitted using the least squares method, the rMBE is zero in all the cases. The stations of Tucson and Boulder (USA) and the station of Albacete (Spain) present the highest rRMSE values (23–25%). The rest of stations present values in the range 14–16%. For the European stations the explained variability is over 80% while for the stations in USA is about 90%. This is probably related to the greater length of the USA time series. Focussing on models P2, P3 and G0, which only use the clearness index as predictor, it can be concluded that the model G0 provides, overall, considerable better estimates than model P2 and slightly better than P3. However, note that the statistics presented in Table 5 refers to the entire range of measured values. At this regard, Fig. 3 allows a preliminary comparison of the behaviour for low, intermediate and high clearness index. It can be seen that P3 and G0 have a similar behaviour for intermediate clearness indices. On the contrary, for clearness indices close to zero, P3 predicts a decreasing diffuse fraction which, although might eventually be found, is not the expected behaviour. The same applies to the increasing diffuse fraction predicted by P3 for high clearness indices. Therefore, although the performance scores of model P3 may be equivalent to that of model G0, the later is able to provide estimates statistically more consistent in the entire range of the clearness index. Similar conclusions can be derived from the evaluation of the models that use both the clearness index and the optical mass as predictors. Likewise, the model G2, based on the sigmoid curve, performs slightly better than the others. In addition, model G2 has fewer parameters than M2.

Table 3
Subregions into which the 21 stations were divided for the local evaluation study and training and validation stations used in the study.

Region	Training station	Validation stations
Spain	Albacete	Granada Ciudad Real
Germany (Germ.)	Dresden	Braunschweig Wuerzburg
South-Western (SW) USA	Tucson	San Diego Midland
South-Eastern (SE) USA	Savannah	Atlanta Tallahassee
Western USA	Boulder	Ely Lander
North-Eastern (NE) USA	Pittsburgh	Albany Nashville
Alaska	Talkeetna	Bethel Fairbanks

Finally, in a third step, the models trained using one of the stations at each region were used to estimated the rest of the stations (two at each region) values. In this independent validation procedure, the squared correlation coefficient, the rRMSE and the rMBE were used. At this stage, the R model was incorporated to the comparative evaluation process. Table 6 presents the results of this validation. As could be expected, the performance of the models is lower than for the training dataset (Table 5). Particularly, the explained variability decreases and the rRMSE and rMBE increase. Interestingly, for Europe, the more simple models P2, P3, G0 and R provide better estimates than the rest. Particularly, the R model provides the best results at the Granada and Ciudad Real stations (Spain subregion), where rRMSE values range from 34% to 40% and rMBE values range from 10% to 17%. Additionally, the stations of Braunschweig and Wuerzburg (Germany) present better relative scattering (27–30%) and mean relative error (11–15%) values when using the simpler models P2, P3, G0 and R.

For the stations located in the USA regions, the simpler models show lower relative scattering and mean relative error than the stations located in Europe. Contrarily to the European stations the use of the optical mass as additional predictor does improve the estimates. An interesting feature of these results is that the stations of the North-Eastern and South-Eastern USA regions show considerable better results (rRMSE ranges from 22% to 26%) than Western and South-Western regions stations (rRMSE ranges from 22% to 35%). Similar results are found in terms of the rMBE: North-Eastern and South-Eastern regions stations range from −8% to 4% and Western and South-Western regions stations range from −15% to 15%. Note in Table 2 that the Western and South-Western USA regions and the Spanish stations correspond to a B climate in the Koeppen's classification, that is, an arid climate. This result might point that this kind of climate have associated a particular frequency distribution of hourly irradiance that makes difficult its characterization by simple statistics models as the evaluated here. In Alaska, the rRMSE vary between 19% and 27% and the rMBE between −13% and 12%.

To sum up, and regarding the simpler models, the G0 and R models provide the best estimates. Particularly, the R model provided the worst results in San Diego and Bethel but, contrary, the best in Fairbanks and Spain. For the rest of stations, the performance is fair. On the other hand, the G0 model provides fair estimates for all the locations, with no large differences in the performance with respect to the other models and stations.

Regarding the more complex models, those using the clearness index and the air mass as predictors, the best estimates are probably found using the G2 and M2 models. Overall, both models provide similar estimates, although G2 gives slightly better results in Europe. Additionally, the behaviour for extreme clearness indices is better in the case of the G2 model. To illustrate this end, we have obtained the hourly diffuse irradiation residuals for all the models for the San Diego station (Fig. 4). This station was chosen because its rMBE is very similar for both G2 and M2 models. As can be observed, the residuals for low and high clearness indices are smaller using the G2 model. On the other hand, for intermediate values, the G2 is better in the lower half range and the M2 model in the higher range. An additional and important advantage of the G2 model is that it has seven coefficients, while M2 needs to fit nine.

Table 4
Fitting coefficients of the models evaluated in this study; in parentheses the corresponding subregion.

Location	Model	a0	a1	a2	a3	a4	a5	a6	a7	a8
Albacete (Spain)	P2	0.962	0.088	−1.482						
	P3	0.718	1.981	−5.741	2.903					
	G0	0.086	−0.880	−3.877	6.138					
	G1	0.096	−0.853	−4.816	7.153	0.178				
	G2	0.108	−0.871	−3.898	3.701	2.769	0.377	−0.038		
	M1	0.917	−0.020	0.384	0.132	−1.558	−0.337			
	M2	0.952	−0.036	−0.001	0.429	0.023	0.038	−1.735	−0.096	−0.067
Dresden (Germ.)	P2	1.014	−0.753	−0.608						
	P3	0.913	0.324	−3.781	2.735					
	G0	0.140	−0.962	−1.976	4.067					
	G1	0.119	−0.991	−1.815	3.889	−0.065				
	G2	−1.618	−2.617	−4.031	7.484	−4.497	−0.034	−0.006		
	M1	1.044	−0.013	−0.920	0.058	−0.601	0.062			
	M2	0.932	0.094	−0.020	−0.077	−0.754	0.150	−1.796	1.225	−0.217
Tucson (SW USA)	P2	1.404	−1.936	0.358						
	P3	0.877	1.688	−7.103	4.745					
	G0	0.988	1.073	2.298	−4.909					
	G1	0.970	1.037	2.948	−5.628	−0.134				
	G2	0.962	1.088	3.382	−5.999	0.608	−0.420	0.051		
	M1	1.405	−0.045	−1.797	0.217	0.428	−0.418			
	M2	1.270	0.097	−0.024	−0.877	−0.660	0.137	−0.515	0.430	−0.112
Savannah (SE USA)	P2	1.252	−1.117	−0.442						
	P3	0.907	1.493	−6.321	4.066					
	G0	0.988	1.000	2.456	−5.172					
	G1	0.980	1.000	2.909	−5.541	−0.122				
	G2	0.973	1.000	3.352	−5.528	−0.136	−0.455	0.055		
	M1	1.248	−0.045	−1.126	0.310	−0.014	−0.642			
	M2	1.082	0.089	−0.015	−0.189	−0.416	0.070	−0.892	−0.043	−0.018
Pittsburgh (NE USA)	P2	1.197	−0.779	−0.743						
	P3	0.770	2.572	−8.557	5.557					
	G0	1.001	1.000	2.450	−5.048					
	G1	0.994	1.000	2.936	−5.440	−0.130				
	G2	0.984	1.000	3.531	−6.342	0.740	−0.385	0.041		
	M1	1.192	−0.047	−0.737	0.302	−0.377	−0.632			
	M2	1.119	0.007	−0.005	−0.260	−0.035	0.025	−0.805	−0.391	0.012
Boulder (Western USA)	P2	1.278	−1.447	−0.107						
	P3	0.812	2.142	−8.168	5.488					
	G0	0.967	1.024	2.473	−5.324					
	G1	0.961	1.048	2.847	−5.472	−0.116				
	G2	0.956	1.268	3.202	−6.712	2.228	−0.213	0.021		
	M1	1.225	−0.015	−1.122	0.096	−0.226	−0.278			
	M2	1.061	0.139	−0.026	−0.316	−0.655	0.127	−1.003	0.438	−0.117
Talkeetna (Alaska)	P2	1.280	−1.297	−0.369						
	P3	0.721	3.171	−11.05	7.793					
	G0	0.985	0.962	2.655	−6.003					
	G1	0.989	1.000	2.760	−5.862	−0.048				
	G2	0.976	1.000	3.221	−7.145	1.280	−0.125	0.010		
	M1	1.401	−0.082	−2.000	0.502	0.780	−0.778			
	M2	1.403	−0.063	−0.007	−1.837	0.257	0.065	0.324	−0.264	−0.120
Global	G0	0.952	1.041	2.300	−4.702					
	G2	0.944	1.538	2.808	−5.759	2.276	−0.125	0.013		

5. Proposal of a global regression equation

In the previous section, the different versions of the here proposed radiation model based on the sigmoid function, along with other representative models of the current literature, were fitted and evaluated for the seven different regions. In this section, we present an additional analysis aiming to evaluate the potential global applicability of these sigmoid models for any location around the world. Particularly, the G0 model was selected among the models using only k_t as predictor and the G2 model was selected among the models using also the optical mass. Then, they were fitted using the seven training stations detailed in Table 3. Hereinafter, we will refer these fitted models as global models, meaning that they have been adjusted with the training stations in USA and Europe altogether. Given the different record length of the German and American stations datasets, only the period 1981–1990 was used. However, the whole dataset of the Spanish stations was considered. The resultant models for Europe and USA are the following:

$$k_0(k_t) = 0.952 - 1.041e^{-\exp(2.300-4.702k_t)}, \quad (33)$$

$$k_2(k_t, m) = 0.944 - 1.538e^{-\exp(2.808-4.759k_t+2.276k_t^2+0.125m+0.013m^2)}. \quad (34)$$

Again, the AIC is greater for k_2 (7.89) than for k_0 (5.86). Table 7 shows the results of the validation process of these two global models. For comparison purposes, the validation results of the models G0 and G2 fitted locally (previously showed in Table 6) are also provided. In addition, and for the sake of clearness, Fig. 5 shows the rRMSE and rMBE of the global models validation for the seven regions under study.

Table 5
Evaluation scores for the training stations of the different regression models locally-fitted. rRMSE and rMBE are given in %. The location column shows the name of the stations used for training the model; in parentheses, the corresponding subregion. The accuracy score (AS) is a weighted mean of the five previous columns. The Akaike's information criterion (AIC) is a selection model score which describes the trade-off between the precision and the complexity of the model. Bold-faced figures in the AS column mean the highest AS at each region for the models with only clearness index as predictor and clearness index and relative optical air mass as predictors.

Location	Model	r^2	rRMSE	rMBE	Kurtosis	Skewness	AS	AIC
Albacete (Spain)	P2	0.8240	25.34	0.00	−0.100	−0.074	0.32	5.72
	P3	0.8269	25.13	0.00	0.062	−0.090	0.43	5.74
	G0	0.8279	25.06	0.00	0.102	−0.059	**0.59**	5.74
	G1	0.8365	24.42	0.00	0.238	−0.060	0.94	7.79
	G2	0.8381	24.31	0.00	0.246	−0.066	**0.97**	7.78
	M1	0.8340	24.61	0.00	0.057	−0.090	0.66	7.76
	M2	0.8344	24.59	0.00	0.062	−0.095	0.67	7.78
Dresden (Germ.)	P2	0.8026	16.79	0.00	0.237	0.449	0.35	5.91
	P3	0.8050	16.69	0.00	0.333	0.429	**0.58**	5.92
	G0	0.8043	16.72	0.00	0.310	0.431	0.52	5.91
	G1	0.8095	16.50	0.00	0.371	0.520	0.70	7.94
	G2	0.8119	16.39	0.00	0.407	0.501	**0.89**	7.95
	M1	0.8106	16.45	0.00	0.316	0.513	0.73	7.94
	M2	0.8117	16.40	0.00	0.299	0.497	0.81	7.95
Tucson (SW USA)	P2	0.8877	26.15	0.00	0.365	0.029	0.36	6.18
	P3	0.8947	25.32	0.00	0.692	0.132	0.44	6.24
	G0	0.8950	25.28	0.00	0.713	0.097	**0.49**	6.24
	G1	0.9053	24.01	0.00	0.995	0.076	0.79	8.33
	G2	0.9103	23.37	0.00	1.188	0.038	**0.97**	8.38
	M1	0.8998	24.71	0.00	0.496	0.046	0.61	8.28
	M2	0.9016	24.48	0.00	0.573	0.013	0.71	8.30
Savannah (SE USA)	P2	0.8828	18.26	0.00	−0.288	0.315	0.24	6.06
	P3	0.8871	17.92	0.00	−0.091	0.308	0.42	6.09
	G0	0.8867	17.96	0.00	−0.111	0.250	**0.44**	6.09
	G1	0.8973	17.10	0.00	−0.033	0.225	0.75	8.18
	G2	0.9037	16.56	0.00	0.115	0.175	**1.00**	8.23
	M1	0.9008	16.80	0.00	−0.124	0.278	0.76	8.21
	M2	0.9026	16.65	0.00	−0.126	0.231	0.84	8.22
Pittsburgh (NE USA)	P2	0.8809	15.76	0.00	−0.233	0.380	0.24	6.02
	P3	0.8897	15.17	0.00	0.089	0.357	0.54	6.09
	G0	0.8899	15.15	0.00	0.073	0.330	**0.55**	6.09
	G1	0.9026	14.25	0.00	0.114	0.242	0.87	8.12
	G2	0.9070	13.93	0.00	0.151	0.181	**1.00**	8.24
	M1	0.9008	14.39	0.00	−0.041	0.300	0.73	8.18
	M2	0.9018	14.31	0.00	−0.072	0.258	0.77	8.19
Boulder (Western USA)	P2	0.8732	24.70	0.00	−0.073	0.203	0.31	5.92
	P3	0.8850	23.53	0.00	0.410	0.254	0.52	6.01
	G0	0.8857	23.45	0.00	0.457	0.230	**0.56**	6.01
	G1	0.8973	22.24	0.00	0.853	0.203	0.88	8.11
	G2	0.8996	21.98	0.00	0.927	0.211	**0.92**	8.13
	M1	0.8885	23.17	0.00	0.224	0.167	0.67	8.03
	M2	0.8893	23.08	0.00	0.228	0.156	0.71	8.04
Talkeetna (Alaska)	P2	0.8864	16.03	0.00	−0.144	0.072	0.31	5.84
	P3	0.8994	15.08	0.00	0.554	0.072	0.77	5.84
	G0	0.8997	15.06	0.00	0.570	0.044	**0.82**	5.86
	G1	0.9031	14.81	0.00	0.741	−0.029	**0.97**	7.93
	G2	0.9038	14.75	0.00	0.845	−0.060	0.95	7.98
	M1	0.8956	15.37	0.00	0.400	−0.116	0.57	8.08
	M2	0.8963	15.32	0.00	0.418	−0.103	0.61	8.02

Overall, results in Table 7 show that the global models strongly improve the estimates of the local version of the models for the two European regions, the Alaska region and the Western and South-Western USA regions. Particularly, for the stations located in these regions, the most important improvements are obtained for the rMBE: in most of the cases the rMBE values obtained using the global models are less than one third of the locally-fitted models. Improvements in terms of the rRMSE are also important but to a lower extend, while the explained variability remains similar.

On the other hand, for the North-Eastern and South-Eastern USA regions the performances of the global models are lower than the locally-fitted models. Particularly, rRMSE values slightly increase and rMBE show notable increment. All the stations located in these regions have a Cfa climate, except the Albany station. This station has a Dfa climate, but is located close to a region of Cfa climate [30] and shares important characteristics with this type of climate, as the existence of significant precipitation in all the seasons. The different behaviour of the stations located in the North-Eastern and South-Eastern USA regions might be related, therefore, to the Cfa climate characteristics.

The main differential characteristic of this kind of climate compared to the other analysed region climates is the nonexistence of a dry season. This probably yields that the Cfa regions have a greater proportion of cloudy conditions than climates with dry season. As a consequence, for a given clearness index, the attenuation of the solar radiation will be more frequently caused by clouds in comparison with the rest of the stations climates. Additionally, the beam fraction will be smaller than in the case of climates having dry seasons and, thus, the diffuse fraction will be greater. Since most of the stations used in the global fitting procedure are located in climates having a dry season, the global model fit will be biased by the data corresponding to these stations. This, finally, results in

Table 6
Evaluation scores for the validation stations of the different regression models locally-fitted. rRMSE and rMBE are given in %. The location column shows the name of the stations used for validation; in parentheses, the corresponding subregion.

Location		P2	P3	G0	R	G1	G2	M1	M2
Granada (Spain)	r^2	0.612	0.634	0.640	0.642	0.648	0.649	0.622	0.623
	rRMSE	35.65	34.92	34.65	34.34	34.64	34.85	36.27	36.23
	rMBE	11.82	11.86	11.73	10.35	13.05	13.52	13.82	13.82
Ciudad Real (Spain)	r^2	0.579	0.595	0.604	0.600	0.601	0.594	0.563	0.565
	rRMSE	37.6	36.67	36.18	35.87	37.34	38.13	39.82	39.80
	rMBE	14.81	14.03	13.65	12.05	15.56	16.10	16.90	17.01
Braunschweig (Germ.)	r^2	0.685	0.676	0.676	0.674	0.680	0.679	0.688	0.689
	rRMSE	26.98	27.32	27.31	28.07	27.91	28.11	27.83	27.74
	rMBE	−11.33	−11.22	−11.22	12.80	−12.05	−12.21	−12.47	−12.49
Wuerzburg (Germ.)	r^2	0.680	0.679	0.678	0.676	0.678	0.678	0.671	0.664
	rRMSE	27.97	27.92	27.92	28.68	28.66	28.84	29.20	29.39
	rMBE	−11.79	−11.16	−11.12	13.03	−12.01	−12.07	−13.00	−13.16
San Diego (SW USA)	r^2	0.831	0.829	0.830	0.815	0.838	0.842	0.849	0.847
	rRMSE	24.32	24.37	24.32	27.49	22.39	21.94	21.75	21.67
	rMBE	−3.83	−5.35	−5.01	14.62	−5.43	−5.60	−6.11	−5.18
Midland (SW USA)	r^2	0.840	0.836	0.836	0.852	0.836	0.839	0.844	0.840
	rRMSE	27.47	27.83	27.88	24.65	25.40	24.12	24.49	24.15
	rMBE	−10.13	−10.67	−10.53	8.21	−7.31	−5.05	−5.44	−4.56
Atlanta (SE USA)	r^2	0.806	0.816	0.816	0.828	0.818	0.823	0.834	0.831
	rRMSE	25.23	25.05	25.10	22.43	23.60	22.47	21.84	21.76
	rMBE	−7.91	−7.24	−6.48	−0.88	−5.03	−3.75	−4.41	−4.17
Tallahasse (SE USA)	r^2	0.803	0.800	0.796	0.812	0.799	0.802	0.819	0.815
	rRMSE	22.94	23.34	23.69	22.07	22.58	22.00	21.08	21.15
	rMBE	−3.43	−3.69	−3.52	3.29	−1.68	−0.26	−0.05	0.35
Albany (NE USA)	r^2	0.835	0.830	0.830	0.844	0.837	0.834	0.843	0.843
	rRMSE	22.14	22.64	22.69	21.55	21.61	21.85	21.08	21.17
	rMBE	3.43	3.66	3.70	3.05	3.89	4.13	2.78	2.96
Nashville (NE USA)	r^2	0.861	0.848	0.850	0.866	0.860	0.860	0.884	0.884
	rRMSE	20.50	21.85	21.80	20.39	19.74	19.36	17.42	17.42
	rMBE	−0.77	0.22	0.22	−1.04	0.88	1.78	−0.08	0.25
Ely (Western USA)	r^2	0.752	0.782	0.785	0.802	0.797	0.805	0.791	0.787
	rRMSE	33.03	30.61	30.65	29.86	28.06	27.16	29.21	29.58
	rMBE	−14.50	−9.35	−10.28	12.72	−7.78	−6.73	−10.16	−10.11
Lander (Western USA)	r^2	0.744	0.775	0.774	0.797	0.781	0.786	0.768	0.766
	rRMSE	35.41	33.93	34.06	30.34	31.82	31.14	32.57	32.67
	rMBE	−14.32	−12.48	−12.76	9.43	−11.04	−10.59	−11.99	−11.93
Fairbanks (Alaska)	r^2	0.779	0.774	0.781	0.819	0.784	0.788	0.807	0.806
	rRMSE	26.75	27.29	27.10	21.07	26.19	25.78	24.71	24.81
	rMBE	−12.67	−12.59	−12.58	3.10	−12.13	−11.83	−11.81	−11.85
Bethel (Alaska)	r^2	0.838	0.849	0.851	0.809	0.854	0.854	0.839	0.841
	rRMSE	19.88	19.33	19.24	24.45	18.72	18.70	19.66	19.45
	rMBE	−2.41	−2.03	−2.09	11.44	−1.19	−0.86	−1.10	−1.24

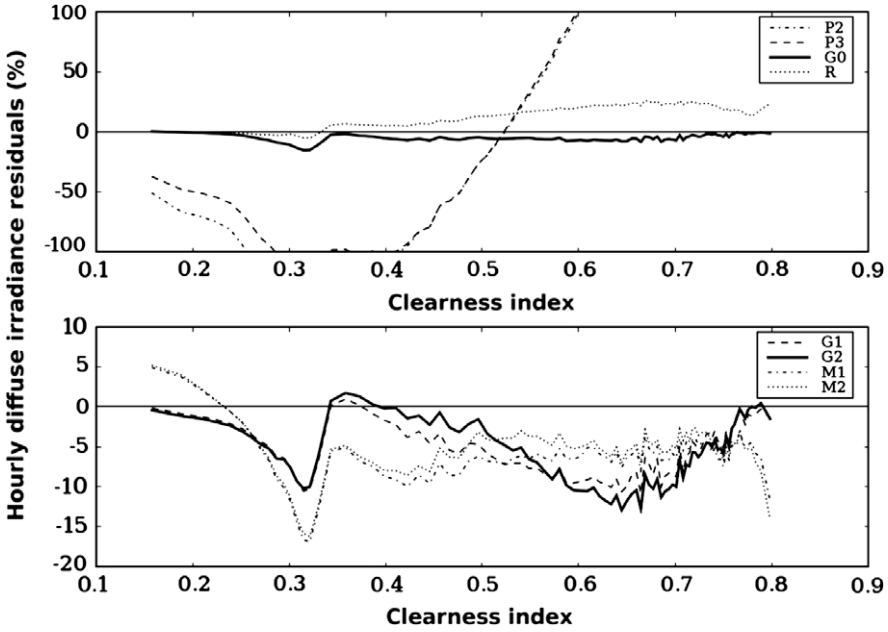

Fig. 4. Residuals of the hourly diffuse irradiation (%) for the station located in San Diego (USA) for the seven evaluated models. A filtering has been applied for a clearer visualization.

Table 7
Evaluation scores for the validation stations of the global models (G0 and G2). For the sake of clearness, the scores for the same models but locally-fitted are also provided (G0 local and G2 local). Relative RMSE and MBE are given in %. The location shows the name of the stations and in parentheses, the corresponding subregion. The Koeppen climate index is also displayed.

Location	Climate		G0 local	G0 global	G2 local	G2 global
Granada (Spain)	BSk	r^2	0.640	0.633	0.649	0.637
		rRMSE	34.65	34.33	34.85	33.10
		rMBE	11.73	−3.78	13.52	−2.31
Ciudad Real (Spain)	BSk	r^2	0.604	0.605	0.594	0.593
		rRMSE	36.18	34.47	38.13	34.11
		rMBE	13.65	−2.64	16.10	−0.99
Braunschweig (Germ.)	Dfb	r^2	0.676	0.682	0.679	0.680
		rRMSE	27.31	24.21	28.11	24.91
		rMBE	−11.22	1.37	−12.21	4.25
Wuerzburg (Germ.)	Dfb	r^2	0.678	0.679	0.678	0.682
		rRMSE	27.92	24.99	28.84	25.46
		rMBE	−11.12	0.88	−12.07	3.74
San Diego (SW USA)	BSk	r^2	0.830	0.832	0.842	0.843
		rRMSE	24.32	23.92	21.94	21.91
		rMBE	−5.01	1.04	−5.60	0.46
Midland (SW USA)	BSh	r^2	0.836	0.849	0.839	0.852
		rRMSE	27.88	25.76	24.12	23.96
		rMBE	−10.53	−4.58	−5.05	−1.72
Atlanta (SE USA)	Cfa	r^2	0.816	0.808	0.823	0.823
		rRMSE	25.10	27.99	22.47	25.66
		rMBE	−6.48	−12.51	−3.75	−10.79
Tallahassee (SE USA)	Cfa	r^2	0.796	0.793	0.802	0.806
		rRMSE	23.69	25.67	22.00	23.63
		rMBE	−3.52	−9.17	−0.26	−7.15
Albany (NE USA)	Dfb	r^2	0.830	0.843	0.834	0.852
		rRMSE	22.69	24.47	21.85	22.86
		rMBE	3.70	−8.52	4.13	−8.47
Nashville (NE USA)	Cfa	r^2	0.850	0.846	0.860	0.866
		rRMSE	21.80	26.41	19.36	23.83
		rMBE	0.22	−12.49	1.78	−11.90
Ely (Western USA)	BWk	r^2	0.785	0.805	0.805	0.807
		rRMSE	30.65	28.09	27.16	26.57
		rMBE	−10.28	−1.45	−6.73	1.00
Lander (Western USA)	BSk	r^2	0.774	0.795	0.786	0.796
		rRMSE	34.06	30.67	31.14	29.11
		rMBE	−12.76	−4.04	−10.59	−2.69
Fairbanks (Alaska)	Dwc	r^2	0.781	0.820	0.788	0.823
		rRMSE	27.10	23.47	25.78	23.27
		rMBE	−12.58	−8.31	−11.83	−9.91
Bethel (Alaska)	Dfc	r^2	0.851	0.840	0.854	0.845
		rRMSE	19.24	19.64	18.70	19.21
		rMBE	−2.09	0.55	−0.86	−0.24

an underestimation of the values for the stations with Cfa climate index. Particularly, the rMBE value considerable increases (at least by a factor of two) for the North and South-Eastern region stations when using the global models compared to the locally-fitted models. Relative RMSE values also increase, but to a lower extend, while small changes are found for the explained variability.

Overall, it could be concluded that both the G0 and G2 global models provide fair estimates for the entire dataset. Particularly, model G2 provides the best estimates (Fig. 5), with rRMSE ranging from around 20% in the Alaska region to 35% in Spain, and with rMBE ranging from less than −5% in Spain to −12% in the eastern USA region.

6. Summary and conclusions

In this work, we propose a new regressive model for the estimation of the hourly diffuse solar irradiation under all sky conditions.

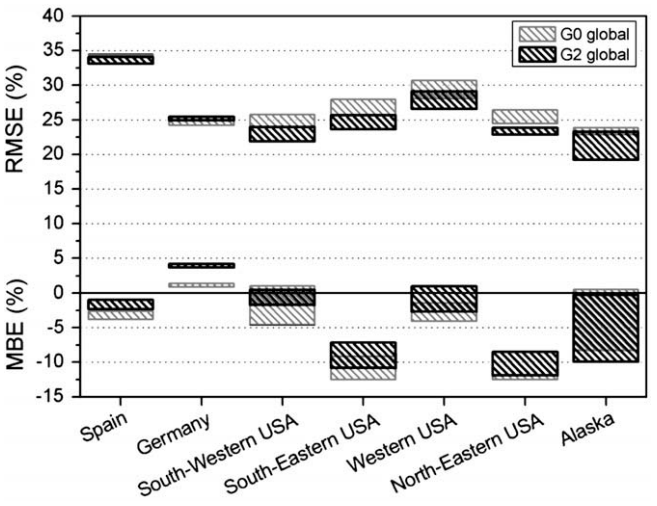

Fig. 5. Ranges of the relative RMSE and MBE values (for the seven analysis regions) obtained based on the proposed global models.

The model is based on a sigmoid function and uses the clearness index and the relative optical mass as predictors. The model's performance was compared against other four regressive models recently proposed in the bibliography and the model of Reindl et al. [14]. For the evaluation, a set of radiation data corresponding to 21 stations in the USA and Europe was used.

In a first part, the 21 stations were grouped into seven subregions (three at each region, namely Spain, Germany, South-Western USA, Western USA, North-Eastern USA, South-Eastern USA and Alaska), corresponding to seven different climatic regions. Both the new model (in three different versions) and the five models taken from the bibliography were locally-fitted and validated using these seven sub-datasets. Particularly, one station at each region was used to train the models and the other two for an independent validation process. Results showed that the new proposed model offers slightly better estimates, in terms of rRMSE, rMBE and explained variability. Particularly, the new model provides relative RMSE in the range 25–35% and the relative MBE in the range −15% to 15%, depending on the considered region. Additionally, the new proposed model shows some important advantages compared to other evaluated models. Particularly, the logistic behaviour of this model is able to provide more reliable estimates (statistically speaking) for extreme values of the clearness index. This avoids the use of piecewise regressive models, that usually introduce extra local dependencies. Moreover, the new model needs less parameters than most of the other analysed models.

In a second part, the potential global spatial applicability of the new model was evaluated. To this end, the seven training stations, one per region, were merged in a same dataset and the model was fitted using this new set of data. The other fourteen stations were used for an independent validation process. Results showed that the global fitting model, based on the sigmoid function, provides overall better estimates than the locally-fitted models. Particularly, the new model provides relative RMSE values between 20% and 35% and a relative MBE between −5% and −12%.

Acknowledgements

The Spanish Ministry of Science and Technology (Project ENE2007-67849-C02-01) and the Andalusian Ministry of Science and Technology (Project P07-RNM-02872) financed this study. Part of the data were provided by Agencia Estatal de Meteorología de España (AEMET). H. Alsamamra is supported by a grant from the

Spanish International Cooperation Agency (Spanish Foreign Office Ministry). The NSRDB was kindly provided by the NREL.

References

[1] Fu P, Rich PM. A geometric solar radiation model with applications in agriculture and forestry. Comput Electron Agric 2002;37:25–35.
[2] Pierce Jr KB, Lookingbill T, Urban D. A simple method for estimating potential relative radiation (PRR) for landscape-scale vegetation analysis. Landscape Ecol 2005;20:137–47.
[3] Tovar J, Olmo FJ, Alados-Arboledas L. Local-scale variability of solar radiation in a mountainous region. J Appl Meteorol 1995;34:2316–22.
[4] Oliphant AJ, Spronken-Smith RA, Sturman AP, Owens IF. Spatial variability of surface radiation fluxes in mountainous region. J Appl Meteorol 2003;42:113–28.
[5] Tovar-Pescador J, Pozo-Vázquez D, Ruiz-Arias JA, Batlles J, López G, Bosch JL. On the use of the digital elevation model to estimate the solar radiation in areas of complex topography. Meteorol Appl 2006;13:279–87.
[6] Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Wang W, et al. A description of the advanced research WRF version 2. NCAR Technical Note, NCAR/TN-468+STR; 2005.
[7] Rigollier C, Bauer O, Wald L. On the clear sky model of the ESRA – European Solar Radiation Atlas – with respect to the Heliosat method. Sol Energy 2000;68:33–48.
[8] Wilson JP, Gallant JC. Secondary topographic attributes. In: Wilson JP, Gallant JC, editors. Terrain analysis: principles and applications. John Wiley and Sons; 2000. p. 91–105.
[9] Mészáros I, Miklánek P, Parajka J. Solar energy income modelling in mountainous areas, ERB and NEFRIEND Proj. 5 Conference on interdisciplinary approaches in small catchment hydrology. Slovak NC IHP UNESCO/UH SAV; 2002. p. 127–35.
[10] Súri M, Hofierka J. A new GIS-based solar radiation model and its application to photovoltaic assessments. Trans GIS 2004;2:175–90.
[11] Ruiz-Arias JA, Alsamamra H, Tovar-Pescador J, Pozo-Vázquez D. A comparative analysis of DEM-based models to estimate the solar radiation in mountainous terrain. Int J Geogr Inform Sci 2009;23:1049–76.
[12] López G, Rubio MA, Batlles FJ. Estimation of hourly direct normal from measured global solar irradiance in Spain. Renew Energy 2000;21:175–86.
[13] Ineichen P. Comparison and validation of three global-to-beam irradiance models against ground measurements. Sol Energy 2008;82:501–12.
[14] Reindl DT, Beckman WA, Duffie JA. Diffuse fraction correlations. Sol Energy 1990;45:1–7.
[15] Liu BYH, Jordan RC. The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. Sol Energy 1960;4:1–19.
[16] Orgill JF, Hollands KGT. Correlation equation for hourly diffuse radiation on a horizontal surface. Sol Energy 1976;19:357–9.
[17] Erbs DG, Klein SA, Duffie JA. Estimation of the diffuse radiation fraction for hourly, daily and monthly average global radiation. Sol Energy 1982;28:293–302.
[18] Muneer T, Munawwar S. Potential for improvement in estimation of solar diffuse irradiance. Energy Convers Manage 2006;47:68–86.
[19] Muneer T, Younes S, Munawwar S. Discourses on solar radiation modelling. Renew Sust Energy Rev 2007;11:551–602.
[20] Clarke P, Munawwar S, Davidson A, Muneer T, Kubie J. Technical note: an investigation of possible improvements in accuracy of regressions between diffuse and global solar radiation. Build Serv Eng Res Technol 2007;28:189–1997.
[21] Jacovides CP, Tymvios FS, Assimakopoulos VD, Kaltsounides NA. Comparative study of various correlations in estimating hourly diffuse fraction of global solar radiation. Renew Energy 2006;31:2492–504.
[22] Notton G, Cristofari C, Muselli M, Poggi P. Calculation on an hourly basis of solar diffuse irradiations from global data for horizontal surfaces in Ajaccio. Energy Convers Manage 2004;45:2849–66.
[23] Perez R, Ineichen P, Maxwell E, Seals R, Zelenka A. Dynamic global to direct irradiance conversion models. ASHRAE Trans Res Ser 1992:354–69.
[24] Skartveit A, Olseth JA, Tuft ME. An hourly diffuse fraction model with correction for variability and surface albedo. Sol Energy 1998;63:173–83.
[25] Scharmer K, Greif J. European Solar Radiation Atlas, 4th ed. Paris: Presses de l'Ecole, Ecole des Mines de Paris; 2000.
[26] Gueymard CA, Myers DR. Solar radiation measurement: progress in radiometry for improved modelling. In: Badescu V, editor. Modeling solar radiation at the Earth surface: recent advances. Berlin, Heidelberg: Springer-Verlag; 2008. p. 1–27.
[27] López G, Muneer T, Claywell R. Comparative study of four shadow band diffuse irradiance correction algorithms for Almería, Spain. J Sol Energy – Trans ASME 2004;126:696–701.
[28] National Renewable Energy Laboratory. National Solar Radiation Data Base User's Manual (1961–1990), vol. 1.0. Ashville, NC, USA; National Climatic Data Center; 1992.
[29] Michalsky J, Dutton E, Rubes E, Nelson D, Stoffel T, Wesley M, et al. Optimal measurement of surface shortwave irradiance using current instrumentation. J Atmos Ocean Technol 1999;16:55–69.
[30] Peel MC, Finlayson BL, McMahon TA. Updated world map of the Köppen–Geiger climate classification. Hydrol Earth Syst Sci Discuss 2007;4:439–73.
[31] Younes S, Claywell R, Muneer T. Quality control of solar radiation data: present status and proposed new approaches. Energy 2005;30:1533–49.
[32] Munner T, Fairooz F. Quality control of solar radiation and sunshine measurements – lessons learnt from processing worldwide databases. Build Serv Eng Res Technol 2002;23:151–66.
[33] Muneer T. Solar radiation and daylight models. 2nd ed. Oxford: Elsevier Butterworth-Heinemann; 2004.
[34] Muneer T, Zhang X. A new method for correcting shadow band diffuse irradiance data. J Sol Energy – Trans ASME 2002;124:34–43.
[35] Woyte A, Belmans R, Nijs J. Fluctuations in instantaneous clearness index: analysis and statistics. Sol Energy 2007;81:195–206.
[36] Boland J, Ridley B. Models of diffuse solar fraction. Renew Energy 2008;3:575–84.
[37] Boland J, Scott L, Luther M. Modelling the diffuse fraction of global solar radiation on a horizontal surface. Environmetrics 2001;12:103–16.
[38] Iqbal M. An introduction to solar radiation. New York: Academic Press; 1983.
[39] Gueymard CA. Turbidity determination from broadband irradiance measurements: a detailed multicoefficient approach. J Appl Meteorol 1998;37:414–35.
[40] Burnham KP, Anderson DR. Multimodel inference: understanding AIC and BIC in model selection. Sociol Methods Res 2004;33:261–304. doi:10.1177/0049124104268644.