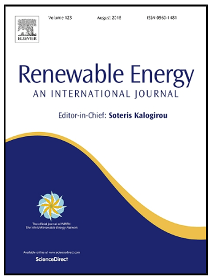
Accepted Manuscript



A spatiotemporal universal model for the prediction of the global solar radiation based on Fourier series and the site altitude

E. Kaplani, S. Kaplanis, S. Mondal

PII: S0960-1481(18)30412-9

DOI: 10.1016/j.renene.2018.04.005

Reference: RENE 9963

To appear in: *Renewable Energy*

Received Date: 24 December 2017

Revised Date: 31 March 2018

Accepted Date: 02 April 2018

Please cite this article as: E. Kaplani, S. Kaplanis, S. Mondal, A spatiotemporal universal model for the prediction of the global solar radiation based on Fourier series and the site altitude, *Renewable Energy* (2018), doi: 10.1016/j.renene.2018.04.005

A spatiotemporal universal model for the prediction of the global solar radiation based on Fourier series and the site altitude

E. Kaplani[a,*], S. Kaplanis[b], S. Mondal[b]

[a]Engineering Division, Faculty of Science, University of East Anglia, Norwich, NR4 7TJ, UK

[b]Renewable Energy Systems Lab, Technological Educational Institute of Western Greece, Meg. Alexandrou 1, Patra 26334, Greece

*Corresponding author. Email address: e.kaplani@uea.ac.uk

Abstract

This paper presents the development, testing and validation of a novel generic type universal model consisting of a set of sine and cosine harmonics in the temporal and spatial domain suitably parameterized for the prediction of the mean expected global solar radiation H(n,φ) on the horizontal for a day, n, at any latitude φ. Its prediction power is further enhanced with the introduction of a correction term for the site altitude taking into account the φ dependent atmospheric height. Solar radiation data from 53 stations around the earth were obtained from GEBA database to train the model. H(n,φ) is expressed by a Fourier series of compact form with the zero frequency component dependent on φ providing the main spatial dependence and two n dependent harmonics in the form of cosine functions giving the time dependence. The φ dependent model parameters follow symmetry rules and are expressed by Fourier series up to the 3rd order harmonic. The 3D spatiotemporal profile of the model is in agreement to the extraterrestrial one. The model was validated using GEBA data from additional 28 sites and compared with NASA, PVGIS and SoDa data, showing the robustness, reliability and prediction accuracy of the proposed model.

Keywords: solar radiation prediction; universal model; Fourier series; site altitude; atmospheric height

1. Introduction

For the sizing of Renewable Energy Systems (RES) configurations it is necessary to provide as input the values of the daily global solar radiation H(n,φ) kWh/m^2/day on the horizontal in any place with latitude φ, for any day n, while in more detailed dynamic simulation models the values of the intensity of the global solar radiation, I(h;n), at a site in any hour h of a day n, are required, [1-4]. Solar radiation is monitored in many stations around the world and data are processed and stored in international databases as in [5-8]. A large number of research studies outline models which provide for H(n) and/or I(h;n) estimates for various sites. Those models are categorized as semi-empirical, ASHRAE [9] and Iqbal [10] models, providing elaborated expressions based on theoretical approaches with regard to the solar light optics such as transmission, reflection and scattering, as well as the atmospheric pressure versus altitude and the ambient temperature for the site and the time period concerned. Both models predict the beam, incident and diffuse components of the global solar radiation in a site enabling the estimation of the mean expected daily global solar radiation based on expressions as below.

$$I_n = C_n A_{ext} e^{-B(P/P_o)\sec(\theta_z)} \quad (1)$$

$$I_n = 0.9751 E_o I_{sc} \tau_r \tau_o \tau_g \tau_w \tau_a \quad (2)$$

where A_{ext} (W/m^2) is the apparent extraterrestrial irradiance given in tables [9]. I_n is the direct normal irradiance (W/m^2), C_n is the ratio of the direct normal irradiance calculated with the local mean clear-day water-vapour over the direct normal irradiance calculated with water-vapour according to the basic atmosphere. P (mbar) is the actual local-air pressure and P_o is the standard pressure (1013.25 mbar). In eq.(2), I_{sc} is the solar constant taken as 1367 W/m^2. E_o (dimensionless) is the eccentricity correction-factor of the Earth's orbit. Finally, τ_r, τ_o, τ_g, τ_w and τ_a are the Rayleigh, ozone, gas, water and aerosols scattering transmittances, (dimensionless), respectively.

A second group comprises of empirical models which provide the daily global solar radiation based on the Angström-Prescott model [11-13] and use various regression based expressions outlined in [14,15]. The parametric values of those expressions are generally valid for the geographical sites or regions they have been determined for. Values for these parameters applicable at any site have been proposed in [12]. The variable in these models is the ratio of the actual sunshine hours, S, over the maximum possible sunshine hours, S_o, in a day n in the site of concern. A third group of empirical solar radiation models correlate further H(n) with the T_{min} and T_{max} air temperature, the relative humidity, RH, and other meteorological parameters, such as the cloud coefficient, C and the precipitation, R, [16-20], as in the general form of eq.(3) with one or more of the above quantities included.

$$\frac{H(n)}{H_{ext}(n)} = f\left(\frac{S}{S_o}, \left(T_{max} - T_{min}\right), RH, R, C\right) \quad (3)$$

where $H_{ext}(n)$ is the daily solar radiation at a site at the top of the earth's atmosphere.

More elaborated models proposed and applied in several projects are the ones in a fourth group which use artificial neural networks (ANN) to provide for H(n), in solar energy systems [21-23]. Finally, there is a group of empirical models which determine H(n) in a site with parameter the number of the day, n, in the year [24-27]. A sub-group uses simple or more complex sine or cosine expressions of Fourier series [28-33],

$$H(n) = A + B\cos\left(\frac{2\pi}{365}n + C\right) \quad (4)$$

$$H(n) = a + bcos(z) + csin(z) + dcos(2z) + esin(2z) \quad (5)$$

where, z=(2π/365)·n, and the parameters A, B, C, and a, b, c, d, e depend on the site and are determined by regression analysis. The predictive performance of those models was shown to be reliable for the region of study. However, the above parameters were determined for the specific region and are not universally applicable. The mean expected hourly global solar radiation on the horizontal may then be determined by the models outlined in [26, 34-36] using the H(n) predicted above as input to satisfy boundary conditions, such as the model analyzed in [28] using the expression,

$$I_{m,exp}(h;n) = A + B\,\exp(-\mu(n)(x(h) - x(12)))\cos\left(\frac{2\pi h}{24}\right) \tag{6}$$

where μ(n) is the solar radiation attenuation coefficient through the atmosphere determined using the predicted H(n) based on eq.(4), x(h) is the solar radiation path in the atmosphere dependent on the hour h in a day n at a site and x(12) is the corresponding path for the solar noon.

Since the parametric expressions of the above models derived through regression analysis were valid only for the regions of concern, a model of universal validity was proposed and tested using 2 cosine functions to predict the global solar radiation in a day at any site [37]. That model was shown to give good H(n) predictions for sites both in the N. and S. Hemispheres. However, the parameters of that model as determined do not guarantee that the H(n) function is continuous when n changes from 365 to 1, i.e. from the end of December to beginning of January the following year.

This paper proposes a reliable and self-consistent generic model of universal applicability composed of a complete set of spatiotemporal terms based on Fourier series satisfying the above requirement. The parameters of the model display symmetries with regard to the N. and S. Hemisphere. Additionally, the proposed model includes a correction for the site altitude and the atmospheric height appropriately parameterized. The model is outlined in Sections 2 and 3 and provides directly the mean expected daily global solar radiation at horizontal H(n,φ) at any site with altitude h_s. The validation of the model is presented and discussed in Sections 4 and 5 where results are given in comparison with the measured data from GEBA and other databases.

2. The Generic Universal Model

The proposed generic model predicts H(n,φ) for any day n and site with latitude φ, and takes also into consideration the site altitude and the atmospheric height. A double harmonic analysis was applied to solar radiation data obtained from GEBA database from a grid of 53 stations around the earth with altitude less than 500m. This process resulted in a set of harmonic spatiotemporal terms whose coefficients are functions of the site latitude.

The model proposed to predict H(n,φ) is expressed through a Fourier series of compact form and is presented in eq.(7).

$$H(n,\varphi) = A(\varphi) + B_1(\varphi)cos\left(l_1(\varphi)\frac{2\pi}{365}n + C_1(\varphi)\right) + B_2(\varphi)cos\left(l_2(\varphi)\frac{2\pi}{365}n + C_2(\varphi)\right) \tag{7}$$

The key requirements and conditions set are:

a. the cyclicity and continuity in the behaviour of the H(n,φ) profiles and especially their rate of change at the end of December and beginning of January to take the same value

b. applicability to both N. and S. Hemispheres

c. the model's coefficients corresponding to the spatial domain [-π/2, π/2] to be expressed with the same order Fourier series

d. the model's coefficients to follow symmetry rules with respect to N. and S. Hemisphere

e. the altitude, h_s, of the site and the variable atmospheric height to be taken into account in the determination of H(n,φ)

The sites chosen were distributed in both S. and N. Hemisphere from East to West as shown in Fig.1. Time series of monthly average global solar irradiance where obtained from GEBA database [5]. The monthly average daily global solar radiation was estimated and averaged over the years data were recorded for each site (in most cases these were more than 10 and in some cases more than 50 years). The estimated monthly averages of daily H(n) values were deployed along 2 consecutive years so that the model coefficients satisfy the requirement (a) above. The monthly averages were mapped to the representative day of each month.

The proposed model based on the compact Fourier series of eq.(7) was fitted on the estimated monthly averages of the daily H(n) values for each of the above sites. A nonlinear regression analysis was applied based on the proposed compact Fourier series model using the nonlinear least squares method. The fundamental frequency is 2π/365.

The model coefficients A, B_1, B_2, and the associated parameters C_1,C_2, l_1, l_2 functions of φ, were derived by nonlinear regression analysis for each one of the 53 sites, satisfying the requirements described above. The coefficient of determination R^2, for any latitude and longitude were between 0.97-0.99 for 96% of the sites, while the NRMSE values were between 0.09-0.34 for all sites.

The frequency parameters l_1, l_2 are φ dependent take integer values [1, 2] corresponding to the 1st and 2nd harmonic. To secure symmetry, parameters C_1, C_2 were normalised based on the well known cosine function properties:

$$\cos(x + (C + 2\lambda\pi)) = \cos(x + C), \quad \cos(x + C) = -\cos(x + (C + \pi)) \tag{8}$$

where, λ is an integer.

In a second stage a Fourier analysis was performed on the model coefficients A, B_1, B_2, and their associated parameters C_1,C_2 as described in Section 3. This resulted in a self-consistent prediction model for H(n,φ).

Fig.1. The 53 sites (drop-shaped) used to train the model and the 28 sites (circles) used for the model validation as illustrated on a Google map.

3. The model coefficients as functions of φ analysed in Fourier series

The Fourier analysis of the coefficients A(φ), B_1(φ), B_2(φ) and the parameters C_1(φ), C_2(φ), determined by nonlinear regression analysis, showed that they may be optimally represented by Fourier series of up to the 3rd order harmonic, providing for the spatial profile of the model expressed through the latitude φ, with the general expression of eq. (9).

$$f(\varphi) = \alpha_0 + \Sigma_{i=1}^{3}\left(\alpha_i \cos\left(i\omega_0\varphi\right) + b_i \sin\left(i\omega_0\varphi\right)\right) \tag{9}$$

The fundamental frequency ω_0 was set equal to 2 to satisfy the condition that φ takes values in [-π/2, π/2].

The Fourier coefficients α_i, b_i and the zero frequency component α_0 for the A(φ), B_1(φ), B_2(φ), C_1(φ) and C_2(φ), obtained through harmonic regression using the nonlinear least squares method taking into account the aforementioned fundamental frequency, are provided in Table 1.

Table 1. The Fourier coefficients of the up to 3rd order harmonics of the H(n,φ) model parameters.

Fourier coefficients	H(n, φ) model parameters				
	A(φ)	B_1(φ)	B_2(φ)	C_1(φ)	C_2(φ)
α_0	4.5180	1.3040	-1.2020	1.9160	1.9160
α_1	0.2055	-0.9208	0.9841	-2.1840	-1.8300
b_1	-0.3439	-1.6650	-1.1560	2.3150	-2.4560
α_2	0.9144	-0.3445	-0.1021	-0.4498	-0.8647
b_2	0.3526	0.3413	0.3068	-1.7250	1.6520
α_3	-0.9101	0.0143	0.2973	0.3847	0.5499
b_3	-0.2346	-0.2715	0.1411	0.1346	0.1396

3.1. On the φ dependence of the model coefficient A

The zero frequency model coefficient A(φ) in eq.(7) is presented in Fig. 2 and exhibits symmetry with respect to the y-axis at φ=0°. It provides for the baseline spatial dependence of the proposed model H(n,φ).

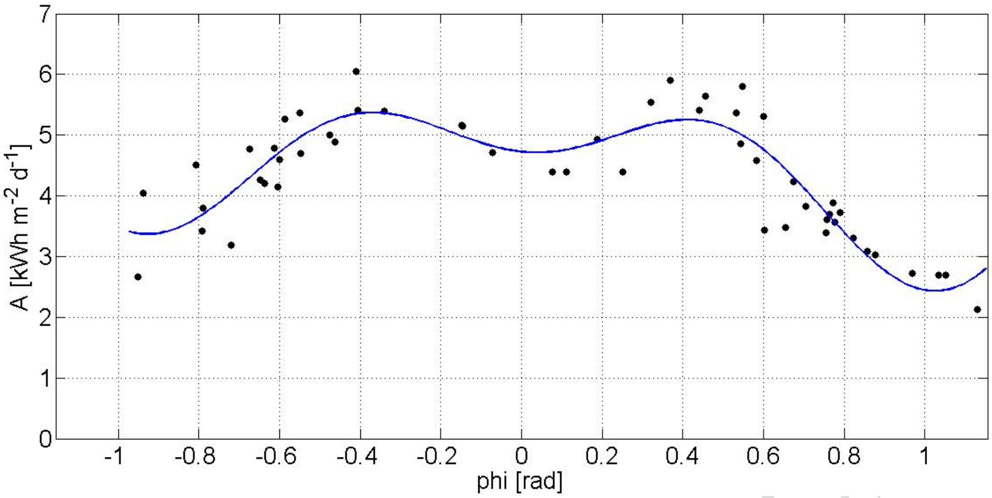
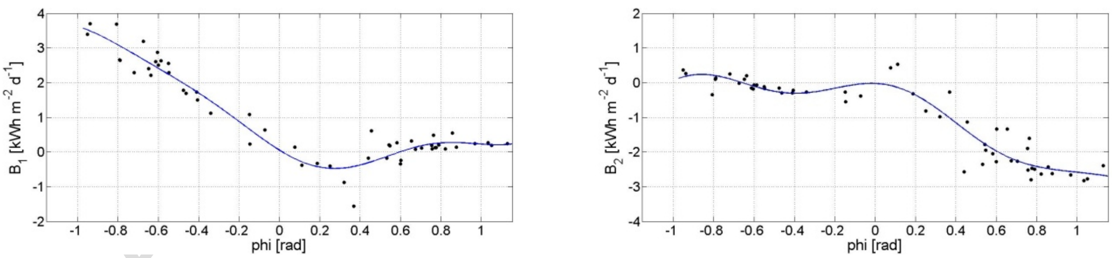


Fig.2 The zero frequency coefficient A of the generic model in kWh/m^{2}day vs φ (rad), as obtained from the nonlinear regression analysis for each of the 53 sites. The fitted curve is a Fourier series of up to 3rd order harmonics whose coefficients are given in Table 1.

3.2 On the φ dependence of the model coefficients B_1 and B_2

The model coefficients $B_1(\varphi)$, $B_2(\varphi)$ presented in Figs. 3(a)-(b) appear to be anti-symmetric to one another, with $B_1(\varphi)$ corresponding mainly to the S. Hemisphere and $B_2(\varphi)$ to the N. Hemisphere. $B_1(\varphi)$ and $B_2(\varphi)$ take values close to zero for φ>0.5 rad in Fig.3(a) and φ<-0.5 rad in Fig.3(b), respectively. This implies that the two time domain harmonics of the model with amplitude $B_1(\varphi)$ and $B_2(\varphi)$, eq.(7), converge to the one cosine model, in sites satisfying the above latitude range in either of the Hemispheres. In the region -0.5 rad≤φ≤0.5 rad there is contribution from both $B_1(\varphi)$ and $B_2(\varphi)$ in the model with the two time domain harmonics differing in frequency (see Section 3.4) and in phase (see Section 3.3).

$B_1(\varphi)$ and $B_2(\varphi)$ are expressed through a Fourier series of up to 3rd order harmonics, whose coefficients are provided in Table 1.



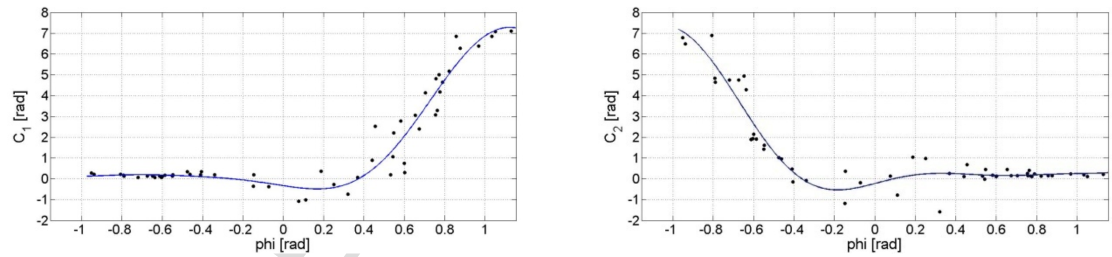
(a) (b)

Fig.3 (a) B_1 (kWh/m^{2}day) and (b) B_2 (kWh/m^{2}day) vs φ (rad). The fitted curves are Fourier series of up to 3rd order harmonics.

3.3 On the φ dependence of the model parameters C_1 and C_2

Parameters $C_1(\varphi)$ and $C_2(\varphi)$ correspond to the phase shift in the two time domain harmonics of the model. Eqs.(8) were applied in the values of C_2 obtained from the nonlinear regression analysis of the H(n) values from the 53 sites, in order to secure symmetry of the function with respect to C_1. In this process the sign of B_2 was adjusted accordingly. The $C_1(\varphi)$ and $C_2(\varphi)$ shown in Figs.4(a)-(b) are expressed through a Fourier series of up to 3rd order harmonics, whose coefficients are provided in Table 1. $C_1(\varphi)$ and $C_2(\varphi)$ appear symmetric to one another with respect to the y-axis at $\varphi=0^o$. This symmetry is also reflected in the Fourier coefficients of $C_1(\varphi)$ and $C_2(\varphi)$ shown in Table 1, where generally the respective α_i coefficients which correspond to the cosine (even function) have the same sign, whereas the b_i coefficients which correspond to the sine (odd function) have the opposite sign, reinforcing a mirror symmetry between $C_1(\varphi)$ and $C_2(\varphi)$ on the y axis at $\varphi=0^o$.

It may be observed that $C_1(\varphi)$ and $C_2(\varphi)$ take values close to zero for $\varphi<-0.5$ rad in Fig.4(a) and $\varphi>0.5$ rad in Fig.4(b) respectively. This indicates that when the two time domain harmonics of the model converge to one cosine model, in either of the Hemispheres, with $B_1(\varphi)$ or $B_2(\varphi)$ prevailing, the corresponding phase shift $C_1(\varphi)$ or $C_2(\varphi)$ respectively is zero. This reduces the model to one cosine model with zero phase shift. The values of $C_1(\varphi)$ or $C_2(\varphi)$ are larger mainly when the contribution of $B_1(\varphi)$ or $B_2(\varphi)$ respectively is small, in which case this emphasizes the effects of seasonality in these regions with $|\varphi|>0.5$ rad. In the tropical and extra-tropical regions with -0.5 rad $\leq\varphi\leq$0.5 rad, where both the two time domain harmonics of the model contribute, the values of $C_1(\varphi)$ and $C_2(\varphi)$ reveal larger seasonal effects, leading for example in the N. Hemisphere the daily solar radiation to be slightly higher in Spring than in Autumn. This effect is illustrated in the 2D and 3D representation of the model in Section 3.5.



(a) (b)

Fig.4 (a) C_1(rad) and (b) C_2 (rad) vs φ (rad). The fitted curves are Fourier series of up to 3rd order harmonics.

3.4 On the φ dependence of the model parameters l_1 and l_2

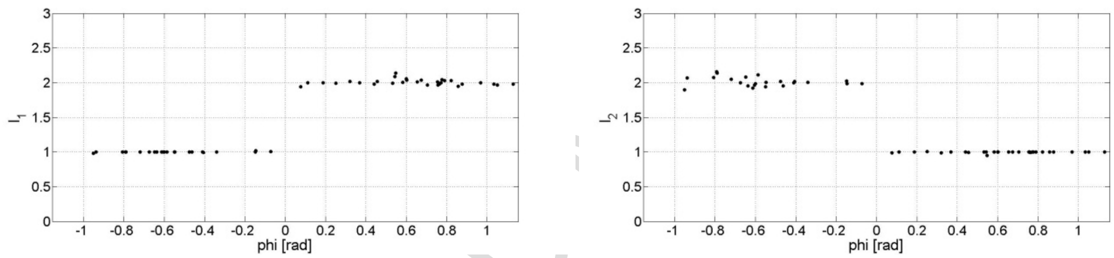
The nonlinear regression analysis of the measured H(n) data for the 53 sites showed $l_1(\varphi)$ and $l_2(\varphi)$ to take values 1 or 2 as presented in Figs. 5(a)-(b), exhibiting a mirror symmetry with respect to the y axis at $\varphi=0^o$. The values of $l_1(\varphi)$ and $l_2(\varphi)$ reflect the condition that these are multipliers of the fundamental frequency $2\pi/365$ in the two cosine day-dependent terms of eq.(7). Therefore, in the proposed model $l_1(\varphi)$ and $l_2(\varphi)$ are provided by the following equations.

$$l_1(\varphi) = \begin{cases} 1, & \varphi < 0^o \\ 2\,, & \varphi \geq 0^o \end{cases} \tag{10a}$$

$$l_2(\varphi) = \begin{cases} 1, & \varphi \geq 0^o \\ 2\,, & \varphi < 0^o \end{cases} \tag{10b}$$

This indicates that when the two time domain harmonics of the model converge to one cosine model, as for example in latitudes with φ>0.5 rad in the N. Hemisphere then the amplitude of the cosine $B_1(\varphi)$ tends to zero and $B_2(\varphi)$ prevails with an $l_2(\varphi)$ frequency multiplier equal to 1, i.e. the frequency is the fundamental $2\pi/365$. This agrees with the one cosine model of eq.(4). Similar analysis holds for latitudes with φ<-0.5 rad in the S. Hemisphere, where $B_2(\varphi)$ tends to zero and $B_1(\varphi)$ prevails with $l_1(\varphi)$ equal to 1.

In regions with latitudes -0.5rad $\leq\varphi\leq$ 0.5rad, the amplitudes $B_1(\varphi)$ and $B_2(\varphi)$ are comparable with a contribution from both cosines of eq.(7) where one of the $l_1(\varphi)$ or $l_2(\varphi)$ is 1 and the other 2 as shown in Fig.5. This is reflected in the two peaks of the $H(n,\varphi)$ profile whose time distance depends on the phase shift $C_1(\varphi)$ and $C_2(\varphi)$, as shown in Section 4.



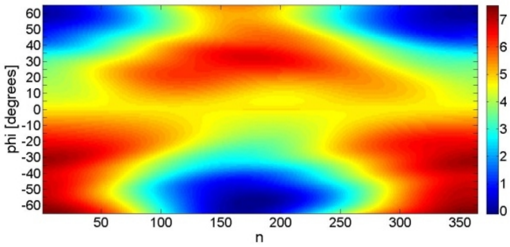
(a) (b)

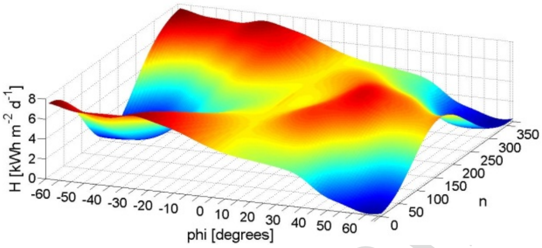
Fig.5(a) Parameter l_1 and (b) l_2 vs φ (rad).

3.5 Model representation

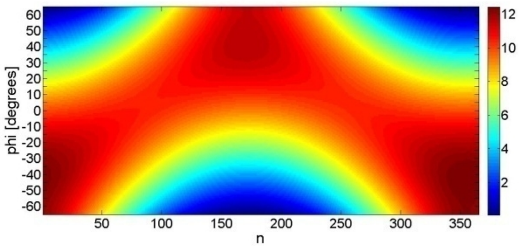
The proposed universal model given by eq.(7) with the coefficients $A(\varphi)$, $B_1(\varphi)$, $B_2(\varphi)$ and the parameters $C_1(\varphi)$, $C_2(\varphi)$ determined by eq.(9) and $l_1(\varphi)$, $l_2(\varphi)$ by eq.(10) was executed for latitudes from -65^o to +65^o and all days of the year ($1\leq n\leq 365$). The resulting 2D and 3D image representations are shown in Figs.6(a)-(b). For comparison reasons, the 2D and 3D image representation of the extraterrestrial solar radiation H_{ext} are shown in Figs.6(c)-(d). The general spatiotemporal profile of the proposed model is in agreement to the extraterrestrial one. Features such as the higher solar radiation received in the S. Hemisphere in December compared to the solar radiation received in the N. Hemisphere in June as a result of the Earth's orbit, are preserved as shown in Fig.6, where additionally these are also higher to that in the equator.

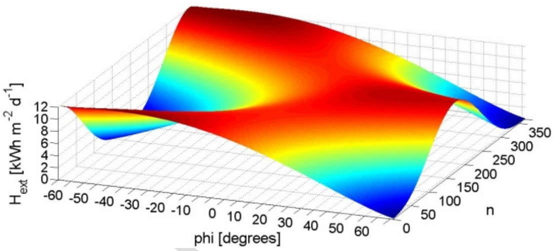
Seasonal variations of the global solar radiations may be observed in the 2D and 3D image representations of the model. In the N. Hemisphere at the tropical region, where the daily solar radiation profile is expressed by two peaks, i.e. both $B_1(\varphi)$ and $B_2(\varphi)$ are contributing, a higher peak is observed during the spring months than during the autumn months. This is a result of the difference in phase between the two harmonic terms and is generally in agreement with solar radiation from databases.





(a) (b)





(c) (d)

Fig.6 (a) 2D and (b) 3D image representation of the proposed universal model, (c) 2D and (d) 3D image representation of the extraterrestrial solar radiation respectively, for latitudes from -65° to +65° and the number of day in the year n. The color map displays the solar radiation in kWh m^{-2} d^{-1}.

3.6 Correction for the site altitude

The impact of the site altitude to the H(n,φ) prediction was investigated. For sites with altitude h_s, the predicted H(n,φ) values corresponding to the sea level were corrected according to eq.(11). The term $\exp(h_s/h_{atm}(\varphi))$ is in conformity to other correlated atmospheric quantities which affect the solar radiation transmission, such as pressure and air density versus altitude [38]. The $H(n,\varphi,h_s)$ taking into account the site altitude is provided by:

$$H(n,\varphi,h_s) = H(n,\varphi) \cdot e^{h_s/h_{atm}(\varphi)} \quad (11)$$

where h_{atm} is the height of the atmospheric layer for a site, φ. In this paper h_{atm} is taken to be that of the Tropopause, which includes more than 80% of the air mass. To estimate the Tropopause height the vertical profile of the atmospheric temperature is required to determine the rate of temperature decrease versus altitude, i.e. the Lapse Rate Tropopause (LRT). The tropopause altitude is the lowest level at which the LRT decreases to 2°C/km or less provided that the LRT in the upper levels does not exceed 2°C/km. The LRT is experimentally proven to be dependent on the latitude and temperature [39-41], as there is a tropopause warming and this causes increase in the tropopause height seasonally [42]. This phenomenon brings a very important issue into the proposed model which is the seasonal variations most important in regions outside the tropic zones. In general, the altitude of the first LRT decreases from 16.2

km in the equator to 8.5 km near the polar regions. Between 20^o -50^o either N or S there is a strong gradient of Tropopause Layer (TL) with φ. The input of the atmospheric height, $h_{atm}(\varphi)$ into this proposed model was analyzed in Fourier series with up to 4th order harmonics shown in eq.(12), which fitted very well to the LRT data provided in [41] and agree with the profiles in [40-43]. This is shown in Fig.7.

$$h_{atm}(\varphi) = \alpha_0 + \Sigma_{i=1}^{4}\left(\alpha_i \cos\left(i\omega_0\varphi\right) + b_i \sin\left(i\omega_0\varphi\right)\right) \quad (12)$$

where φ is given in radians and the fundamental frequency ω_0=2 due to the period in φ equal to π. The Fourier coefficients α_i, b_i and the zero frequency component α_0 are shown in Table 2. For more accurate predictions it is important to introduce to the model the seasonal variations to $h_{atm}(\varphi)$ which depend on the latitude and the month.

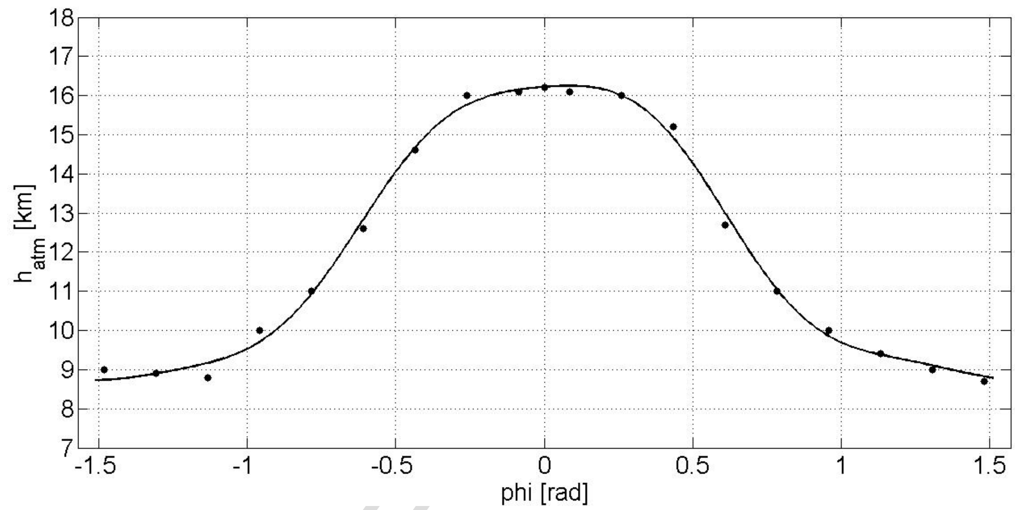


Fig. 7 The curve of the altitude (km) of the first LRT.

Table 2: The Fourier coefficients of the h_{atm} Fourier series up to the 4th order harmonics.

Fourier coefficients								
α_0	α_1	b_1	α_2	b_2	α_3	b_3	α_4	b_4
11.95	3.971	0.1123	0.7537	0.00892	-0.2332	0.05556	-0.2204	0.00086

4. Results and Model Validation

The proposed spatiotemporal model, as expressed by eqs.(7)-(12), was validated with 28 extra sites with GEBA stations widespread from tropical and extra-tropical, to temperate and cold climates. Those sites shown in Fig.1 are independent from the set of sites used for the training of the proposed model. The validation was performed against the estimated monthly average daily global solar radiation averaged over the years data were recorded in GEBA for each of these sites. A comparison between the predicted H(n,φ) monthly mean daily values and the GEBA data is given in Figs.8(a)-(f) and Figs.9(a)-(f), where the corresponding data profiles from NASA, SoDa and PVGIS databases are also shown, to provide for a complete picture of the inherent deviations between the various databases and the predicted H(n,φ) values by this model. The statistical analysis between predicted and measured (GEBA) values for the 28 sites is given in Table 3, where the correlation coefficient R, the Normalized Mean Bias Error (NMBE), the Normalized Root Mean Square Error (NRMSE) and t-statistic results are provided. In general, the values predicted by this model follow the profile of the GEBA values and in most cases the abs(NMBE) is smaller than 0.2, the NRMSE is smaller than 0.25, the correlation coefficient is higher than 0.90, and the t-statistic is below or close to the t critical value 3.106 for a=0.01. However there are cases where one or more of the statistics fall outside this range and these are discussed below to disclose any factors of deviation. In these cases it is very important to examine how the corresponding data from the other databases behave and get a complete picture of the proposed model.

The predicted H(n,φ) profiles shown in Figs.8(a)-(d) for the regions with latitude 40.67°S, 34.95°S, 26.57°S, and 19.12°S compared to GEBA data have a very good correlation coefficient generally higher than 0.99 and low NMBE and NRMSE but relatively high t-statistic especially for the site in Fig.8(c) where this model provides lower than the GEBA values, however similar to SoDa. In Fig.8(d), the predicted H(n,φ) profile for September-December-April is very close to GEBA compared to the other database results. In general, the model performs very well and the corrected to the site altitude $H(n,\varphi,h_s)$ values are closer to GEBA in most of the periods of the year, see Fig.8(d).

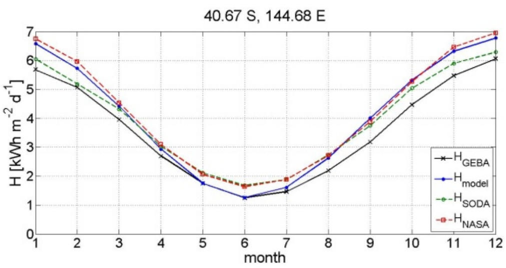
Fig.8(e) presents the comparison of the model performance for a site with latitude 5.08°S in the tropic zone in Tanzania with satisfactory statistic results as far as it concerns the NMBE and NRMSE, while the correlation coefficient has a low value, and the t statistic is higher than the critical value (site 7 in Table 3). The model results are shown along with the other 5 databases, GEBA, NASA, SoDa, PVGIS-CMSAF and PVGIS-Helioclim. In this case, it is very important to discuss over the low correlation between the predicted and GEBA data and investigate on the deviations observed. It is underlined that considerable deviations also exist between PVGIS-CMSAF and PVGIS-Helioclim and GEBA. The investigation on the poor correlation coefficient, R for the sites 6, 7, 8 in Table 3, focused on the H(n,φ) profiles of several sites in Tanzania such as Iringa 7.67°S 35.75°E (site 6 in Table 3), Arusha 3.33°S 36.62°E, Morogoro 6.83°S 37.65°E, Tabora airport 5.08°S 32.83°E (site 7 in Table 3) and Kilimanzaro airport 3.42°S 37.07°E. Although these sites differ by 1°-3° in latitude they experience largely different profiles. The research study in [44] mapping the Tanzania solar resources shows these different profiles which for the case of Morogoro, Arusha and Kilimanzaro are similar to the ones predicted by this model as the profile displayed in Fig.8(e). The deviations are attributed to the different micro-climatic conditions which prevail in the regions in Tanzania with plain and mountainous areas.

For the case of Momote 2.1°S, 147.72°E (site 8 in Table 3) shown in Fig.8(f), a large deviation is observed in spring and autumn months, and a low correlation coefficient but good results for the statistical criteria NRMSE and NMBE and the t-statistic. It is observed that large deviation is also exhibited between the SoDa and GEBA profile for this site. A similar performance appears for the site with latitude 10.33°N in the N. Hemisphere, shown in Fig.9(a).

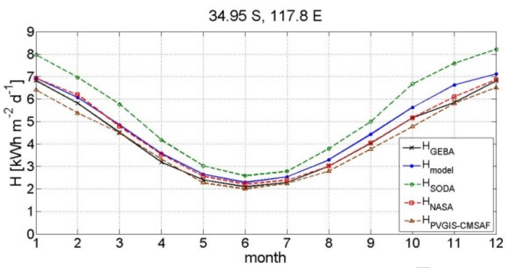
A substantial deviation between the predicted and measured H(n,φ) values is shown in Fig.9(b). This is the case of Guangzhou in China, with latitude 23.13°N, and longitude 113.32°E. All four statistics in this case give poor results. An investigation of the large deviation in this case gave that the clearness index K_T calculated for this site is significantly low and correspondingly the fraction of the diffuse to the global solar radiation is considerably high compared to the other sites of the same latitude, as presented in Table 4. The reason of the high deviation between predicted and measured is attributed mainly to the high anthropogenic pollution which prevails in that region rather than to the climatic conditions. The other databases, and particularly SoDa and PVGIS show also substantial deviations providing higher values compared to GEBA as shown in Fig.9(b).

Fig.9(c) provides the comparison between the results of this model, corrected for the site altitude, for Lhasa, Tibet (29.40°N, 91.80°E and altitude 3.649 km) with the GEBA data and the other databases. In this case, the correction to the site's altitude was introduced to provide $H(n,\varphi,h_s)$. The statistical criteria NMBE, NRMSE and R take very good values, but the t statistic exceeds the critical value. It is noteworthy that the Tibetan Plateau (TP) due to its large volume and height perturbs the tropopause height especially during the short boreal Summer when the TP behaves as a heat sink and boosts the Tropopause to higher altitudes by about 2 km than in the Plain (region in China with the same latitude as TP but different longitude) while during the boreal Winter the Tropopause altitude in the TP drops even lower than that in the Plain [45]. The height of the first LRT exhibits considerable seasonal variations ranging from about 13km during Winter up to about 19km during Summer as discussed in [45]. These seasonal variations in the first LRT height for the case of Tibet were introduced directly in the height correction term in the model eq.(11), which resulted in a considerable improvement of the $H(n,\varphi,h_s)$ prediction. This prediction profile both with the seasonal h_{atm} and with the h_{atm} determined by eq.(12) is presented in Fig.9(c). The statistical results shown in Table 3 correspond to the latter profile. It is interesting to note that the predicted profile $H(n,\varphi,h_s)$ is in a very good agreement with the results given in [38] for the TP.

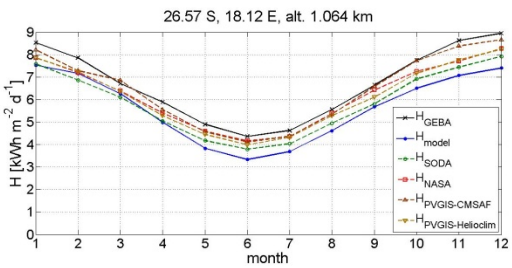
Finally, for the sites 35.05°N 106.62°W and altitude 1.631km, 41.7°N 87.98°W and 55.35°N 131.57°W the predicted profiles shown in Figs.9(d)-8(f) are in good agreement with the measured data expressed also through the statistical criteria where the correlation coefficient is higher than 0.99 and the NMBE, NRMSE take low values as shown in Table 3. Additional cases for the model performance are presented for the other sites in Table 3. The agreement of the predicted solar radiation by the proposed model with the corresponding measured data from GEBA database for the 28 validation sites is shown in Fig.10 along with the dichotomy. The resulting correlation coefficient is 0.881 and RMSE 0.806 $kWhm^{-2}d^{-1}$. The case of Guangzhou, China and Lhasa, Tibet are shown to have larger deviation.



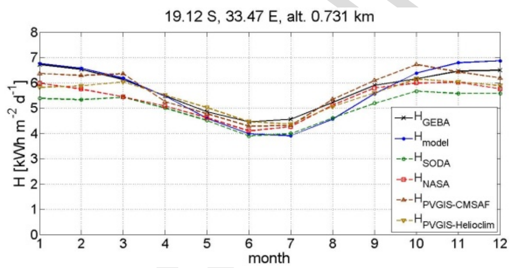
(a)



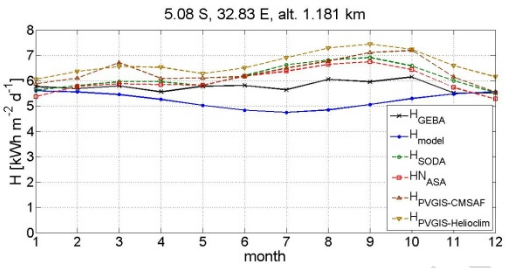
(b)



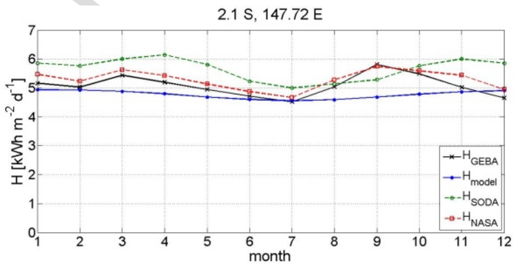
(c)



(d)

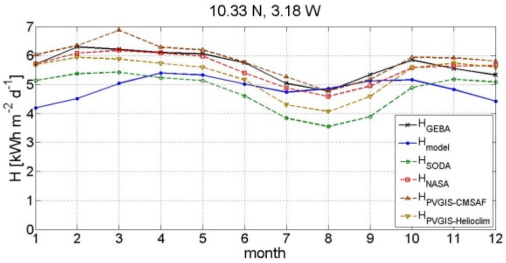


(e)

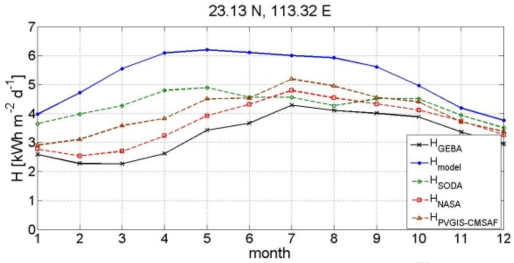


(f)

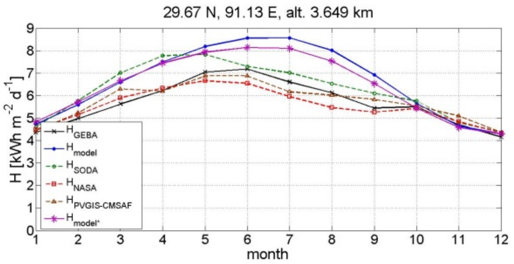
Fig.8(a)-(f). Predicted H(n,φ) values, corrected to the site's altitude where appropriate, versus GEBA, SoDa, NASA, PVGIS-CMSAF and PVGIS-Helioclim available data for different sites in the S. Hemisphere. The latitude, longitude and altitude (where appropriate) of the sites are shown in the figures' title.



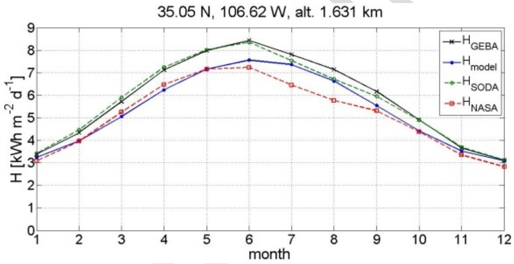
(a)



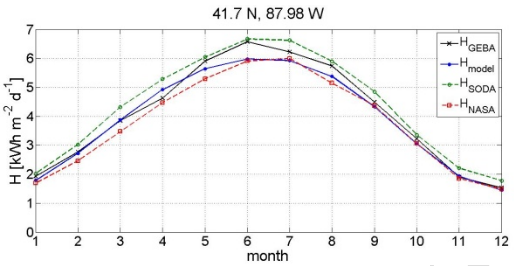
(b)



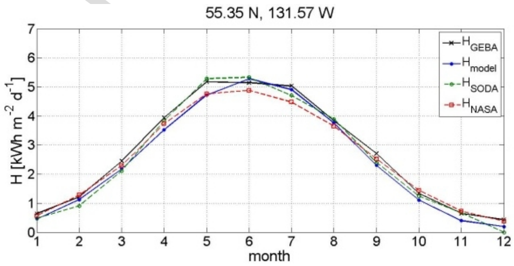
(c)



(d)



(e)



(f)

Fig. 9(a)-(f). Same as in Figure 8 but for sites in the N. Hemisphere. Note, in 9(c) an additional curve H_{model*} is presented in which the height correction uses h_{atm} values directly from the seasonal altitude of the first LRT for the Tibetan Plateau.

Table 3: Statistics of the comparison between the predicted by this model $H(n,\varphi)$ and the measured values from GEBA database. For the sites with considerable altitude the correction for the height based on eqs.(11) -(12) was applied.

Site #	Latitude (deg), Longitude (deg), Altitude (km)	NMBE	NRMSE	R	t-statistic
1	-40.67, 144.68	0.140	0.167	0.996	5.165
2	-34.95, 177.8	0.078	0.086	0.996	6.906
3	-26.57, 18.12, 1.064[a]	-0.153	0.160	0.983	-11.416
4	-19.12, 33.47, 0.731[a]	-0.019	0.063	0.983	-1.040
5	-17.95, 122.23	-0.150	0.168	0.876	-6.614
6	-7.67, 35.75, 1.426[a]	-0.148	0.182	-0.240	-4.627
7	-5.08, 32.83, 1.181[a]	-0.094	0.118	-0.357	-4.326
8	-2.1, 147.72	-0.062	0.093	0.225	-3.014
9	4.4, 18.52	-0.035	0.080	0.711	-1.635
10	10.33, -3.18	-0.138	0.164	0.253	-5.177
11	10.62, -61.35	0.089	0.127	0.425	3.261
12	19.53, 41.05	-0.140	0.147	0.986	-10.553
13	22.65, 88.45	0.154	0.205	0.699	3.760
14	23.07, 72.63	-0.035	0.163	0.568	-0.738
15	23.13, 113.32	0.601	0.657	0.392	7.522
16	23.17, -82.35	0.030	0.062	0.979	1.817
17	29.67, 91.13, 3.649[a]	0.166	0.204	0.956	4.688
18	32.27, -64.33	0.085	0.103	0.990	4.805
19	35.05, -106.62, 1.631[a]	-0.086	0.099	0.996	-5.869
20	35.67, 138.62	0.212	0.266	0.942	4.388
21	37.92, 12.52	-0.062	0.114	0.996	-2.141
22	41.7, -87.98	-0.036	0.064	0.994	-2.265
23	49.63, 100.17	-0.196	0.228	0.971	-5.580
24	50.35, 80.25	-0.248	0.265	0.996	-8.779
25	51.32, -108.4	-0.271	0.281	0.996	-12.242
26	51.52, -0.12	0.067	0.137	0.991	1.868
27	55.35, -131.57	-0.080	0.099	0.996	-4.478
28	58.75, -94.07	-0.214	0.257	0.984	-4.959

Note: t critical (a=0.01): 3.106

The latitude in the Northern Hemisphere is taken positive and in the Southern Hemisphere negative. The longitude towards East from Greenwich is taken positive and towards West negative.

[a] indicates the corrected to height solar radiation $H(n,\varphi,h_s)$ for the marked sites with significant altitude

Table 4. The ratio of the diffuse over the global solar radiation and the clearness index on an annual basis in sites with the same latitude as with Guangzhou and different longitudes.

Site	Latitude	Longitude	H_d/H	K_T^*
Guangzhou, China	23.13°N	113.32°E	0.66[a], 0.55[b]	0.33
Macau, China	22.20°N	113.54°E	0.48[a], 0.51[b]	0.42
Ahmedabad, India	23.07°N	72.63°E	0.37[a], 0.36[b]	0.61
Casa Blanca, Cuba	23.17°N	82.35°W	0.47[a]	0.52
Tamanrasset, South Algeria	22.47°N	5.31°E	0.25[a], 0.34[b], 0.28[c]	0.70

*calculated from the ratio H/H_{ext} with the annually average H obtained from METEONORM and the annually average H_{ext} calculated for the specific site latitude
[a] METEONORM
[b] PVGIS-CMSAF
[c] PVGIS-Helioclim

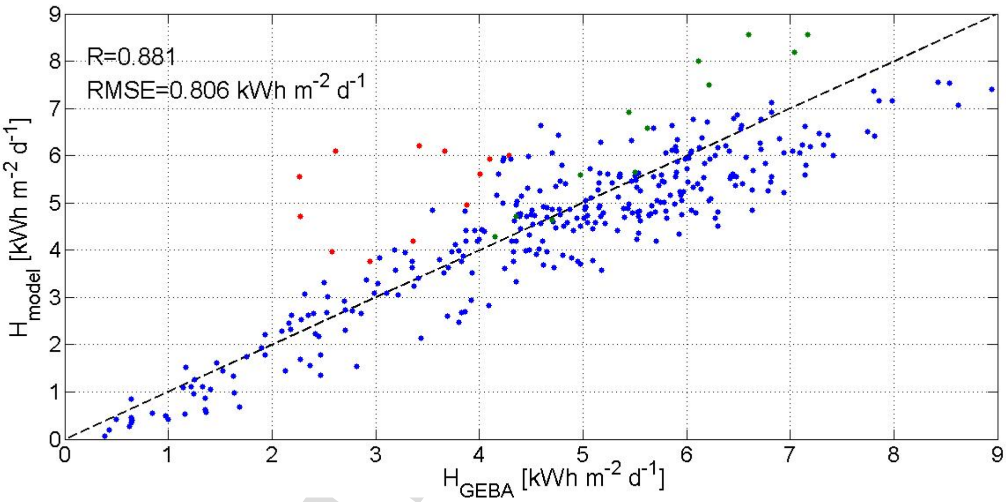


Fig. 10. H(n,φ) predicted by the model vs the measured values from GEBA database for the 28 sites and the 12 months. Correction for the site altitude was applied were appropriate. The special cases of Guangzhou, China (red dots) and Lhasa, Tibet (greed dots) are highlighted.

5. Discussion

The proposed universal model was developed by analyzing H(n) data for a number of years obtained from GEBA database for 53 sites uniformly distributed around the world. The data analysis process follows a double spatiotemporal harmonic analysis. The analysis showed that the model parameter $A(\varphi)$ is symmetric with respect to the y axis at $\varphi=0^o$, and similarly the parameters $C_1(\varphi)$ and $C_2(\varphi)$, $l_1(\varphi)$ and $l_2(\varphi)$ take symmetric forms in the φ space mirrored on the y axis at $\varphi=0^o$, while $B_1(\varphi)$ and $B_2(\varphi)$ appear anti-symmetric inverted at the origin $\varphi=0^o$. The latter is also realized through the 3D image representation of the proposed model (Fig.6(b)). Obviously, due to the Earth's orbit, an absolute symmetry between N. and S. Hemispheres does not exist and therefore an absolute symmetry in the model parameters was not expected.

The model was validated in 28 additional sites, randomly selected and covering a large geographical space extended within abs(φ)<65^{o} and any longitude. For the majority of the cases examined at least 3 out of the 4 statistical criteria used had values which displayed a successful prediction as compared to GEBA measured data. Predicted profiles were also compared with the corresponding profiles from NASA, PVGIS and SoDa databases. It is underlined that the validation included some abnormal cases such as the region of Guangzhou in China where the anthropogenic environmental pollution reaches 52 μg/m^{3}compared to the national standard 35 μg/m^3 and is the major factor in deviations between the predicted and measured data (Figs.9(b),10). Another such case was sites in Tanzania differing in latitude by 1^{o}-2^{o}, where the microclimate pattern in that region caused the deviations between the predicted H(n,φ) and the measured profiles.

An important feature of this model is that it converges to the one cosine model for sites with abs(φ)>0.5rad where the model coefficients B_1(φ) for the S. Hemisphere and B_2(φ) for the N. Hemisphere become almost zero. The investigation of the impact of the site's altitude to H(n,φ) resulted in an effective correction term dependent on φ and incorporated into the model. This was shown in several cases with site altitudes ranging from 0.73 to 3.65 km the latter corresponding to Lhasa, Tibet. Additionally, it was shown that the variation of the h_{atm} with latitude plays a significant role in the prediction of solar radiation and the incorporation of the seasonal variations of the h_{atm} in the $H(n,φ,h_s)$ improves the predicted profile as compared to the measured data, shown for Lhasa in Fig.9(c).

The monthly average daily global solar radiation data used to train the model were averaged over multiple years that data were recorded in the GEBA database for each site, which makes the proposed model resilient to annual fluctuations in the solar radiation profile and promotes the long-term applicability of the model. Nevertheless, long-term trends with decadal changes in the global solar radiation have been analysed in [5,46] and attributed among other causes to changes at the tropopause, aerosol characteristics and pollution. This highlights the need for consideration of the influence of these parameters in a local but also temporal level.

6. Conclusions

The development of a spatiotemporal universal model to predict the expected mean daily global solar radiation, H(n,φ), and its validation results were described and argued in this paper. The model is based on a Fourier series of compact form with variable the day of the year, n, while its φ-dependent parameters, A, B_1, B_2, C_1, C_2 are given by Fourier series of up to 3rd order harmonics. It is applicable as a generic model which through a set of mathematical expressions may predict the mean expected daily solar radiation at the horizontal, H(n,φ), for any site at any day. Further, it may be used in the prediction of the solar irradiance at any hour of the day, I(h;n;φ), with the least required data. The impact of the site's altitude was incorporated into the model using an exponential correction term and Fourier series up to the 4th order harmonic for the estimation of the φ-dependent atmospheric height necessary for the correction. The results obtained using the altitude correction and the seasonal variations of the atmospheric height were impressive. The validation process showed that the model is reliable and self-consistent. The predicted H(n,φ) values for a very large spectrum of latitudes and longitudes show that the model predicts H(n,φ) very close to the measured global solar radiation. The model predicts with a good accuracy the cases where

H(n,φ) exhibits 2 peaks during a year within and near to the tropic zones. For abs(φ)>0.5rad the model converges to the one cosine model. Finally, the proposed model can be easily incorporated into any sizing software for solar energy applications.

Further work will focus on integrating in this model the hourly prediction of the global solar irradiance I(h,n,φ) at the site altitude and the seasonal variation of the atmospheric height, which are not yet considered in such models, providing a complete generic universal model for both H(n,φ) and I(h,n,φ).

Acknowledgements

The authors express their appreciation to Prof. M. Wild from the Institute for Atmospheric and Climate Science, ETH Zurich, for providing access into GEBA database and for communication about GEBA data collection and management.

References

[1] A. Mellit, S.A. Kalogirou, H. Shaari S,Salhi, A. Hadj Arab. Methodology for predicting sequences of mean monthly clearness index and daily solar radiation data in remote areas: Application of sizing a stand-alone PV system. Renewable Energy Vol.33 (2008) 1570-1590.

[2] E. Kaplani, S. Kaplanis. A stochastic simulation model for reliable PV system sizing providing for solar radiation fluctuations. Applied Energy, Vol.97 (2012) 970-981.

[3] Ravinder Kumar, L. Umanand. Estimation of global radiation using clearness index model for sizing photovoltaic system. Renewable Energy, Vol.30 (2005) 2221–2233.

[4] T. Markvart, A. Fragaki, J.N. Ross. PV system sizing using observed time series of solar radiation. Solar Energy, Vol.80 (2006) 46–50.

[5] M. Wild, A. Ohmura, C. Schär, G. Müller, D. Folini, M. Schwarz, M.Z. Hakuba, and A. Sanchez-Lorenzo. The Global Energy Balance Archive (GEBA) version 2017: a database for worldwide measured surface energy fluxes, Earth Syst. Sci. Data, Vol.9 (2017) 601-613, https://doi.org/10.5194/essd-9-601-2017.

[6] Solar radiation data (SoDa). http://www.soda-is.com

[7] NASA. Surface meteorology and Solar Energy. A renewable energy resource web site (release 6.0). https://eosweb.larc.nasa.gov/sse/

[8] Joint Research Centre, Institute for Energy and Transport. Photovoltaic Geographical Information System (PVGIS). http://re.jrc.ec.europa.eu/pvgis/

[9] ASHRAE handbook: HVAC applications, Atlanta (GA): ASHRAE; 1999

[10] M. Iqbal. An introduction to solar radiation. Toronto: Academic press, 1983.

[11] A. Dumas, A. Andrisani, M. Bonnici, M. Madonia, M. Trancossi. A new correlation between solar energy radiation and some atmospheric parameters. Cornell University Library, arXiv:1401.7837, (2014).

[12] M.R. Rietveld. A new method for estimating the regression coefficients in the formula relating solar radiation to sunshine. Agricultural Meteorology, Vol.19 (1978) 243–252.

[13] C. Gueymard, P. Jindra, V. Estrada-Cajigal. A critical look at recent interpretations of the Angstrom approach and its future in global solar radiation prediction. Solar Energy Vol.54(5) (1995) 357–363.

[14] L.T. Wong, W.K. Chow. Solar radiation model, Applied Energy, Vol.69 (2001) 191–224

[15] D.T. Cotfas, P.A. Cotfas, E. Kaplani, C. Samoila. Monthly average daily global solar and diffuse solar radiation based on sunshine duration and clearness index for Brasov, Romania. Journal of Renewable and Sustainable Energy, Vol.6, 053106 (2014); doi: http://dx.doi.org/10.1063/1.4896596.

[16] V. Badescu. Correlations to estimate monthly mean daily global solar irradiation: application to Romania. Energy, Vol. 24 (1999) 883-93.

[17] F. Besharat, A.A. Dehghan, A.R. Faghih. Empirical models for estimating global solar radiation: A review and case study. Renewable and Sustainable Energy Reviews, Vol. 21 (2013) 798-821.

[18] H. Li, F. Cao, X. Bu, L. Zhao. Models for calculating daily global solar radiation from air temperature in humid regions—A case study. Journal of Environ. Prog.& Sustain. Energy, 34(2) (2015) 595-599.

[19] O.O. Ajayi, O.D. Ohijeagbon, C.E. Nwadialo, O. Olasope. New model to estimate daily global solar radiation over Nigeria. Sustain. Energy Technol. Assess. Vol. 5(2014) 28-36

[20] I. Korachagaon, V.N. Bapat. General formula for the estimation of global solar radiation on earth's surface around the globe. Renewable Energy, Vol. 41, (2012) 394–400

[21] S.A. Kalogirou. Artificial neural networks and genetic algorithms for the modeling, simulation and performance prediction of solar energy systems. Chapter in: Assessment and Simulation Tools for Sustainable Energy Systems (2013), Springer, pp. 225–245.

[22] P. Neelamegam, V.A. Amirtham. Prediction of solar radiation for solar systems by using ANN models with different back propagation algorithms. Journal of Applied Research and Technology, Vol.14 (2016) 206-214

[23] S. Reddy, M. Ranjan. Solar resource estimation using artificial neural networks and comparison with other correlation models. Energy Conversion and Management, Vol.44(2003) 2519–2530.

[24] H. Khorasanizadeh, K.Mohammadi, M. Jalilvand. A statistical comparative study to demonstrate the merit of day of the year-based models for estimation of horizontal global solar radiation. Energy Conversion and Management, Vol.87 (2014) 37–47.

[25] H. Li, W. Ma, Y. Lian, X. Wang. Estimating daily global solar radiation by day of year in China. Applied Energy, Vol.87(10) (2010) 3011–3017

[26] C. Gueymard."Prediction and Performance Assessment of Mean Hourly Global Radiation. Solar Energy, Vol.68(3) (2000) 285–303.

[27] R. Mullen, L. Marshall, B. McGlynn. A Beta Regression Model for Improved Solar Radiation Predictions, Journal of Applied Meteorology and Climatology, Vol.52 (2013), 1923-1938.

[28] S. Kaplanis, E. Kaplani. A model to predict expected mean and stochastic hourly global solar radiation $I(h;n_j)$ values. Renewable Energy, Vol.32 (8) (2007) 1414–1425.

[29] Y.-C. Sun, R. Kok. A solar radiation model with a Fourier transform approach. Canadian Biosystems Engineering/Le génie des biosystèmes au Canada, Vol. 49 (2007) 7.17-7.24

[30] U. Ali Rahoma, A.H. Hassan, Fourier transforms investigation of global solar radiation at true noon: in the Desert climatology. American Journal of Applied Sciences, Vol.4(11) (2007), 902-907.

[31] H. Zang, Q. Xu, H. Bian. Generation of typical solar radiation data for different climates of China. Energy, Vol. 38 (2012) 236-248.

[32] A.S.S. Dorvlo. Fourier analysis of meteorological data for Seeb. Energy Conversion and Management, Vol.41(12) (2000) 1283–1291.

[33] N. Fodor, J. Mika. Using analogies from soil science for estimating solar radiation. Agric. For. Meteor., 151, (2011) 78–86.

[34] M. Collares-Pereira, A. Rabl. The average distribution of solar radiation-correlations between diffuse and hemispherical and between daily and hourly insolation. Solar Energy, Vol.22 (1979) 155-164

[35] A. Baig, P. Akhter, A. Mufti. A novel approach to estimate the clear day global radiation. Renewable Energy, Vol.1(1) (1991) 119-123.

[36] S. Kaplanis. New methodologies to estimate the hourly global solar radiation; Comparisons with existing models. Renewable Energy, Vol.31(6) (2006) 781–790.

[37] S. Kaplanis, J. Kumar, E. Kaplani. On a universal model for the prediction of the daily global solar radiation. Renewable Energy, Vol.91 (2016) 178-188.

[38] J.M. Wallace, P.V. Hobbs. "Atmospheric Science: An Introductory Survey. Elsevier, 2nd Edition (2006).

[39] T. Reichler, M. Dameris, R. Sausen. Determining the tropopause height from gridded data. Geophysical Research Letters, Vol. 30(20), 2042, (2003) doi:10.1029/2003GL018240.

[40] J. Thuburn, G.C. Craig. GCM Tests of Theories for the Height of the Tropopause. Journal of the Atmospheric Sciences, Vol. 54(1996) 869-882.

[41] S. Feng, Y. Fu, Q. Xiao. Trends in the global tropopause thickness revealed by radiosondes. Geophysical Research Letters, Vol.39, L20706, (2012) doi:10.1029/ 2012GL053460

[42] B.D. Santer et al. "Behavior of tropopause height and atmospheric temperature in models, reanalyses, and observations: Decadal changes. Journal of Geophysical Research, Vol. 108(D1), 4002, (2003) doi:10.1029/2002JD002258

[43] T. Rieckh, et al. "Characteristics of tropopause parameters as observed with GPS radio occultation. Atmos. Meas. Tech., Vol.7, 3947–3958, (2014) doi:10.5194/amt-7-3947-2014.

[44] Solar Resource Mapping in Tanzania: Solar Modeling Report; World bank Group. (2015), https://www.esmap.org/

[45] S. Feng, Y. Fu, Q. Xiao. Is the tropopause higher over the Tibetan Plateau? Observational evidence from Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC) data. Journal of Geophysical Research, Vol. 116, D21121, (2011) doi:10.1029/2011JD016140

[46] M. Wild. Global dimming and brightening: A review. Journal of Geophysical Research, Vol. 114, D00D16 (2009) doi: 10.1029/2008JD011470.

Highlights

- A spatiotemporal universal model to predict the mean daily global solar radiation
- Generic model based on Fourier series with symmetries in the N. and S. Hemisphere
- Model incorporates the site altitude and the atmospheric height as a function of φ
- Model trained using GEBA data from 53 sites and validated at extra 28 random sites
- Model predictions compared with GEBA, NASA, PVGIS and SoDA data