NEW MODELS FOR ESTIMATING THE DAILY GLOBAL SOLAR IRRADIANCE

Mihai Răzvan MITROI[1], Valerică NINULESCU[2], Laurenţiu FARA[3]

This study proposes new models for estimating of daily global solar irradiance in the region near the city of Bucharest, Romania, based only on daily data of sunshine duration. Empirical models of the correlation between the clearness index and the relative sunshine duration are used. Two new models and three from the literature are analyzed. The mean bias error, the root mean square error, the slope s of the best-fit line and t-statistic have been calculated to test the performance of the models. The percentage error in estimating the annual global solar irradiance using one of the new models, in the selected region for 2014, is of 3.31 %.

Keywords: solar energy, logistic function, regression analysis

1. Introduction

Solar energy is the most important source of renewable energy and the conversion of this energy into electricity is performed by photovoltaic systems. The design of photovoltaic systems requires knowledge of annual global average irradiance on a horizontal surface as the input variable to predict the amount of electricity produced.

The potential use of solar energy in Romania is quite important. The annual average of global solar irradiance exceeds 1200 kWh/m^2 on almost the entire area of the country, the most favorable solar regions being the Black Sea coast and the south of the country, with an average annual solar irradiance of (1400–1600) kWh/m^2 [1].

In Europe, the leader to the total installed PV capacity at the end of 2017 was Germany with 42 GW, while in Romania this amounted to 1.4 GW [2].

The numerical weather prediction models are very complex nonlinear equations and they are not always available. The correlation of the solar irradiance with various meteorological parameters (sunshine duration, atmospheric pressure,

[1] Associate professor, Dept. of Physics, University POLITEHNICA of Bucharest, Romania, e-mail: rm.mitroi@gmail.com.
[2] Lecturer, Dept. of Physics, University POLITEHNICA of Bucharest, Romania, corresponding author, e-mail: ninulescu.valerica@gmail.com.
[3] Prof., Dept. of Physics, University POLITEHNICA of Bucharest, Romania; Academy of Romanian Scientists, 54 Splaiul Independentei, Bucharest, Romania, e-mail: lfara@renerg.pub.ro.

relative humidity, air temperature and cloudiness) has been proposed using several methods: statistical, neural network, stochastic, etc. [3–17].

This study considers the correlation of the solar global irradiance only with the sunshine duration. The goal is to find the most adequate empirical model of their correlation for the region near Bucharest city, the region which has a typical climate for the southern Romania. Besides three regression correlation models from the literature, we propose two new models: one is based on the logistic function and the other is a generalized exponential model. The performance of the models is analyzed using the following statistical indicators: the mean bias error (*MBE*), the root mean square error (*RMSE*), the slope of the best-fit line (s) and the t-statistic. The global solar irradiance for year 2014 is estimated using these models and compared to the measured value.

2. Correlation between the global solar irradiance and sunshine duration

The maximum possible sunshine duration (hours) in a day is the astronomical day length [18]

$$S_0 = \frac{2}{15}\omega_s \tag{1}$$

where

$$\omega_s = \arccos(-\tan\phi \tan\delta) \tag{2}$$

is the sunset hour angle, ϕ is the latitude (north positive) and

$$\delta = 23.45 \sin[360\,(284+z)/N] \tag{3}$$

is the solar declination angle. In Eq. (3), N is the number of days in a year (365 for an ordinary year and 366 for a leap year) and z is the day of the year, $z = 1,\ldots,N$.

The daily extraterrestrial global solar irradiance (J/m^2) [18] is

$$H_0 = \frac{24\times 3600}{\pi} G_{SC}\left(1 + 0.033\cos\frac{360\,z}{N}\right) \left(\cos\phi\,\cos\delta\,\sin\omega_s + \frac{\pi}{180}\omega_s \sin\phi\,\sin\delta\right) \tag{4}$$

where $G_{SC} = 1367$ W/m^2 is the solar constant.

The measured daily global solar irradiance H_m is correlated to the effective sunshine duration S_m by using the clearness index $k_t = H_m / H_0$ and the relative sunshine duration S_m / S_0 [18]. The clearness index is a parameter that describes the attenuation of the solar irradiance due to atmosphere and this depends on the geographical coordinates of the location for which is calculated. Distinctly overcast or sunny hours can be easily identified by a very high or low clearness index. The distinctly sunny ($k_t \geq 0.75$) and distinctly overcast hours ($k_t \leq 0.2$)

are used for the creation of reference quantities for overcast and sunny hours, respectively.

The annual global solar irradiance is calculated by the relation

$$W_{\text{annual}} = \sum_{z=1}^{N} H_z \ , \tag{5}$$

where H_z is the global solar irradiance for day z of the year.

3. Database

There is used measured data for the daily sunshine duration (Fig. 1) and the daily global solar irradiance (Fig. 2) on a horizontal surface for years 2008–2014 from the Romania National Meteorological Administration. The monitoring meteo station is located near Bucharest city, at the latitude of 44.50º N, longitude 26.21º E and 90 m altitude. The global solar irradiance was measured with a thermoelectric pyranometer CM 11, manufactured by Kipp & Zonen (Delft-Holland). Sunshine duration is measured with a Campbell-Stokes recorder [19]. Figure 1 shows the maximum possible daily sunshine duration calculated with Eq. (1) and the measured one.

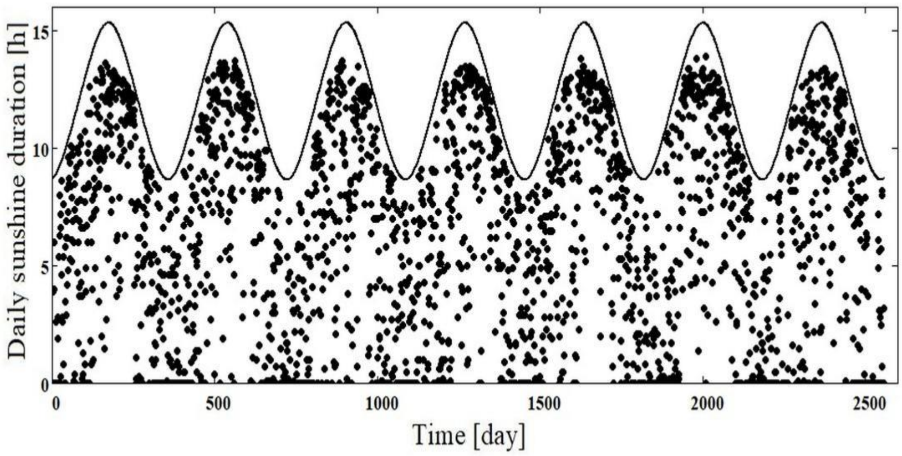

Fig. 1. Daily sunshine duration: maximum possible (—) given by Eq. (1) and measured (•) at Bucharest (2008–2014)

The daily extraterrestrial global solar irradiance given by Eq. (4) and the measured one on a horizontal surface are presented in Fig. 2.

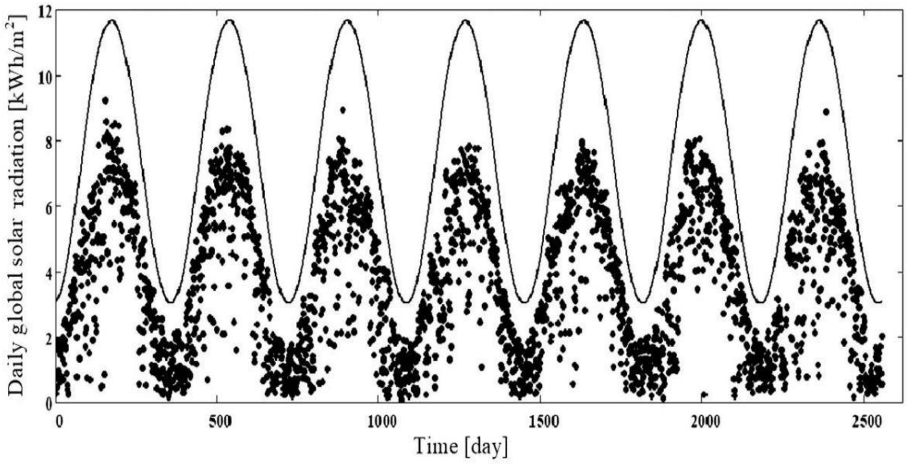

Fig. 2. Daily extraterrestrial global solar irradiance (—) given by Eq. (4) and measured on a horizontal surface (•) for Bucharest (2008–2014)

Annual sunshine duration and annual global solar irradiance—maximum possible and measured—of this region, for years 2008–2014, are given in Table 1 and Table 2. The measured average annual global solar irradiance is about 1333.8 kWh/m^2, its daily average value being 3.65 kWh/m^2; this shows that the analyzed region has a good potential for solar applications.

Table 1

Annual sunshine duration of Bucharest, for years 2008–2014

Year	Annual sunshine duration [h/year]	
	Maximum possible	Measured
2008	4392.0	2369.2
2009	4380.0	2310.4
2010	4380.0	2153.9
2011	4380.0	2306.6
2012	4392.0	2561.0
2013	4380.0	2275.6
2014	4380.0	2064.2

Table 2

Annual global solar irradiance at Bucharest, for years 2008–2014

	Annual global solar irradiance [kWh/m^2]	
	Maximum possible	Measured
2008	2707.6	1393.1
2009	2700.3	1367.5
2010	2700.3	1291.8
2011	2700.3	1313.9
2012	2707.6	1383.0
2013	2700.3	1352.3
2014	2700.3	1234.8

4. Sunshine-based correlation models

The widely spread method for measuring the global solar irradiance is to measure sunshine duration (period during which direct solar irradiance at the ground surface exceeds the threshold value of 120 W/m^2) by use of a Campbell-Stokes sunshine recorder or a Jordan one, and then use these data to calculate the global horizontal irradiance using a model between the clearness index and the relative sunshine duration. This solution avoids the direct measurement of the global solar irradiance by instruments such as pyranometers and pyrheliometers that require an expensive maintenance. The measured clearness index versus the daily relative sunshine duration for the 2008–2013 years is shown in Fig. 3.

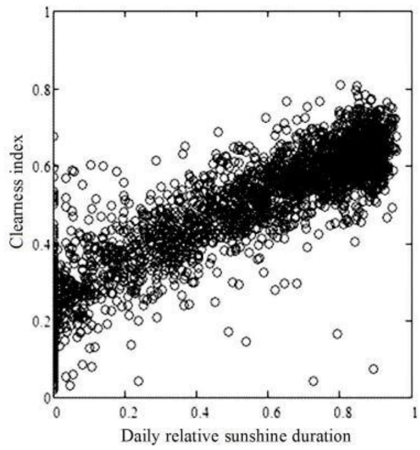

Fig. 3. The measured clearness index versus the daily relative sunshine duration for the 2008–2013 years

Figure 3 shows a fairly rapid growth at the beginning, but due to the restrictions that limit this process (the day has 24 hours, and the annual cycle: spring, summer, autumn or winter) follows a slower growth with limitation. This indicates a process that could be modeled by the logistic function. The logistic function shows an initial exponential increase followed by a period in which the growth slows down and then levels up, approaching (but never reaching) a maximum limit. A new model based on the logistic function is proposed [20]

$$\frac{H_{\mathrm{m}}}{H_0} = \frac{1}{a + b\exp\left(c\,\frac{S_{\mathrm{m}}}{S_0}\right)} \tag{6}$$

We also propose a generalized exponential model (generalized Almorox-Hontoria model [21]) of the form

$$\frac{H_{\mathrm{m}}}{H_0} = a + b\exp\left(c\,\frac{S_{\mathrm{m}}}{S_0}\right) \tag{7}$$

Three models from literature are analyzed for a comparative analysis with the new proposed models:

- *Linear model* (Angström-Prescott) [12, 22−23]:

$$H_{\mathrm{m}}/H_0 = a + b\left(S_{\mathrm{m}}/S_0\right) \tag{8}$$

- *Linear-exponential model* (Bakirci) [24]:

$$H_{\mathrm{m}}/H_0 = a + b\left(S_{\mathrm{m}}/S_0\right) + c\exp\left[\left(S_{\mathrm{m}}/S_0\right)\right] \tag{9}$$

- *Quadratic model* (Ogelman et all.) [25]:

$$H_{\mathrm{m}}/H_0 = a + b\left(S_{\mathrm{m}}/S_0\right) + c\left(S_{\mathrm{m}}/S_0\right)^2 \tag{10}$$

The coefficients *a, b* and *c* in Eqs. (6)−(10) are determined in the paper using a regression method based on the conjugate gradient algorithm; they are generally dependent on the geographical location.

5. Model performance assessment

To compare the accuracy and validation of the above correlation models, the following set of statistical indicators are used in the present study [26–28]:

- *MBE* – it is a measure of systematic errors and provides an indication of the tendency to underpredict or overpredict the modeled values:

$$MBE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - x_i\right) \tag{11}$$

where x_i is a measured value, y_i is a calculated value and n is the total number of observations.

● *RMSE* – it gives information about the short-term performance of the models; *RMSE* takes on a positive value and it is zero in the ideal case:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - x_i\right)^2} \tag{12}$$

● *s* – it is desired to be equal to one; a value exceeding one indicates overestimation, while a value below one indicates underestimation of the calculated variable:

$$s = \frac{\sum_{i=1}^{N}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2} \tag{13}$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ are the mean values of *x* and *y*, respectively.

● *t*-statistic – it indicates if the prediction of a model is statistically meaningful; a smaller value of *t*-statistic shows a better performance of the model:

$$t - \text{statistic} = \left[\frac{(n-1)MBE^2}{RMSE^2 - MBE^2}\right]^{1/2} \tag{14}$$

6. Methodology

In analysis, the relative daily sunshine duration data from period 2008 to 2014 ($7 \times 365 + 2$ days), for the chosen region are employed. The comparative analysis of the newly proposed models [Eqs. (6) and (7)] with the three models from literature [Eqs. (8)–(10)] is performed. The method contains the following steps:

1) The coefficients *a, b* and *c* from models are obtained by a regression method–conjugate gradient algorithm. The coefficients are calculated using 2008−2013 data ($6 \times 365 + 2$ days);

2) The performance of each model is evaluated by use of the statistical indicators and the models are compared;

3) The daily and annual global solar irradiance are estimated for the year 2014 with the use of each model.

All the programs used are written in Mathcad environment.

7. Results and discussion

The regression coefficients *a, b* and *c* from all regression models are presented in Table 3.

Table 3

The regression coefficients *a*, *b* and *c* in Eqs. (6)–(10)

Model	*a*	*b*	*c*
Logistic	1.460	3.164	3.571
Generalized exponential	0.838	-0.631	-1.256
Linear	0.232	0.474	—
Linear-exponential	0.543	1.006	-0.331
Quadratic	0.209	0.718	-0.274

Table 4 shows the values of the statistical indicators for the new models [Eq. (6)–(7)] and the models from literature [Eqs. (6)–(10)]. The values of *MBE* and *RMSE* indicate a good performance for all models, and comparable to the values obtained for this region [26]. The value of the slope of the best fit line is less than one for all models which indicates an underestimation of the computed values. The value of *t*-statistic for all models does not exceed *t*-critic whose value is 2.578 for $\alpha = 0.01$ and $n = 2192$ [29]. As one can see, no model has all five of the best indicators. Each indicator has similar values for all models. Therefore, the proposed new models, the logistic model and generalized exponential model [Eqs. (6), (7)], are suitable to be used for the estimation of the clearness index.

Table 4

The statistical indicators for models described by Eqs. (6)–(10)

Model	$MBE\times100$	*RMSE*	*s*	*t*-statistic
Logistic	0.555	0.550	0.940	0.472
Generalized exponential	0.640	0.545	0.942	0.550
Linear	0.275	0.562	0.946	0.229
Linear-exponential	0.579	0.546	0.943	0.496
Quadratic	0.605	0.545	0.943	0.519

Table 5 presents the statistical indicators for estimating the clearness index k_t for 2014, using the models given by Eqs. (6)–(10). The value of *t*-statistic for all models does not exceed *t*-critic whose value is 2.589 for $\alpha = 0.01$ and *n*=365 [29]. The values in the table indicate that there is a good agreement between the measured and estimated data sets. However, the generalized exponential model [Eq. (7)] singles out with the best values of *MBE*, *RMSE* and *s*.

Table 5

The statistical indicators for the estimating k_t for 2014 year

Model	*MBE*	*RMSE*	*s*	*t*-statistic
Logistic	0.121	0.557	0.964	0.193
Generalized exponential	0.112	0.548	0.968	0.224
Linear	0.116	0.565	0.954	0.193
Linear-exponential	0.114	0.550	0.968	0.212
Quadratic	0.116	0.552	0.968	0.212

The clearness index, estimated by use of Model 1–5 [Eq. (6)–(10)] and the measured one, versus the daily relative sunshine duration for the 2014 year is shown in Fig. 4. The measured clearness index data is close and symmetrically scattered from the theoretical curve obtained with the generalized exponential model.

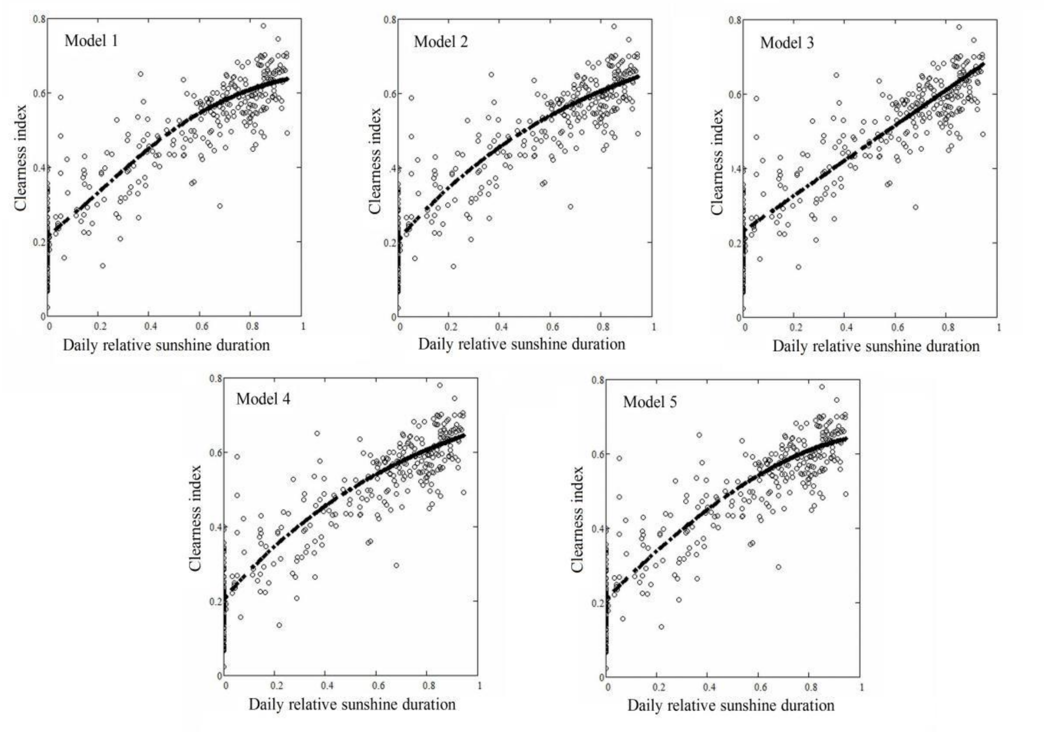

Fig. 4. Clearness index versus the daily relative sunshine duration for the 2014 year: estimated by use of the five models [Eqs. (6)–(10)] (—) and measured (◦)

Figure 5 presents the estimated and calculated daily global solar irradiance with Models 1−5 versus the measured value for the 2014 year. The arrangement of the points, in all cases, is practically symmetrical with respect to the angle bisector of first quadrant thus proving a consistent modeling.

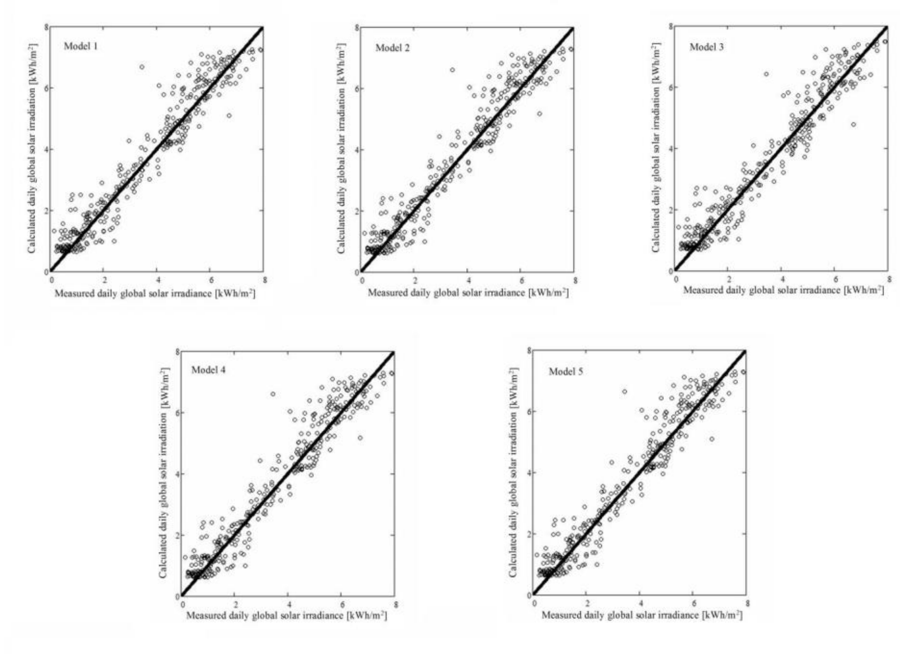

Fig. 5. Estimated daily global solar irradiance by use of Models 1−5 versus measured value, for 2014 year

Table 6

The global solar irradiance for 2014 year: estimated and the percentage error

Model	Estimated annual global solar irradiance [kWh/m^2]	ε_r [%]
Logistic	1278.8	3.56
Generalized exponential	1275.7	3.31
Linear	1277.3	3.44
Linear-exponential	1277.2	3.38
Quadratic	1276.6	3.43

The maximum possible and measured global annual solar irradiance, Table 2, and that estimated with Models 1−5 using only the daily duration of the sun Table 6, of regions selected for 2014 are presented. The obtained values show that the analyzed region has a good potential for solar applications. The percentage error $\varepsilon_r = (y - x)/x$ obtained with the generalized exponential model is 3.31 %, the best value compared to all the other models (see Table 6). Therefore, at least one of the two new models, the generalized exponential model, is a suitable model to estimate the global solar irradiance for the chosen region.

8. Conclusions

In this paper two new model based on the correlation between the clearness index and the relative sunshine duration are proposed. The new models make use Eq. (6) and Eq. (7) for the first time in the estimation of the clearness index. The statistical indicators for estimating the clearness index for 2014 used in the evaluation show that the new model, generally exponential, has the best accuracy compared to the other models analyzed. The percentage error in the estimation of the annual global solar irradiance by use of the new model is less than 3.56 % for the year 2014. This result indicates a good estimation of the annual global solar irradiance for analyzed region near Bucharest city, Romania.

The results obtained for the chosen region indicate that the new models can also be used with good results for regions with similar climatic conditions in determining the global solar irradiance using only the daily sunshine duration.

R E F E R E N C E S

[1] Solar irradiance data. https://solargis.com/maps-and-gis-data/download/romania.

[2] https://iea-pvps.org/wp-content/uploads/2020/11/IEA _PVPS_ Trends_ Report_2020-1.pdf.

[3] *K. Bakirci*, "Prediction of global solar radiation and comparison with satellite data", J. Atmospheric Sol.-Terr. Phys., **vol.** 152, pp. 41–49, 2017.

[4] *V. Badescu* and *A. Dumitrescu*, "New types of simple non-linear models to compute solar global irradiance from cloud cover amount", J. Atmospheric Sol.-Terr. Phys., **vol.** 117, pp. 54–70, 2014.

[5] *S. Bhardwaj, V. Sharma, S. Srivastava, O.S. Sastry, B. Bandyopadhyay, S.S. Chandel* and *J.R.P. Gupta*, "Estimation of solar radiation using a combination of hidden Markov model and generalized fuzzy model", Solar Energy, **vol.** 93, pp. 43–54, 2013.

[6] *S.A.M. Maleki, H. Hizam* and *C. Gomes*, "Estimation of hourly, daily and monthly global solar radiation on inclined surfaces: models re-visited", Energies, **vol.** 10, no. 1, 134, 2017.

[7] *S. Hongkong* and *S. Pattanasethanon*, "Observation modeling for continuous predicting global and diffuse solar radiation on horizontal surface", American Journal of Environmental Science, **vol.** 9, no. 3, pp. 201–209, 2013.

[8] *S.E. Berrizbeitia, E.J. Gago* and *T. Mune*er, "Empirical models for the estimation of solar sky-diffuse radiation. A review and experimental analysis", Energies, **vol.** 13, 701, 2020.

[9] *C. Voyant, P. Haurant, M. Muselli, C. Paoli and Marie-Laure Nivet*, "Time series modeling and large scale global solar radiation forecasting from geostationary satellites data", Solar Energy, **vol.** 102, pp. 131–142, 2014.

[10] *C. Arbizu-Barrena, J.A. Ruiz-Arias, F.J. Rodriguez-Benitez, D. Pozo-Vazquez* and *J. Tovar-Pescador*, "Short-term solar radiation forecasting by advecting and diffusing MSG cloud index", Solar Energy, **vol.** 155 (Supplement C), pp. 1092–1103, 2014.

[11] *K. Bakker, K. Whan, W. Knap* and *M. Schmeits*, "Comparison of statistical post-processing methods for probabilistic NWP forecasts of solar radiation", Solar Energy, Elsevier, **vol.** 191, pp. 138–150, 2019.

[12] *M. Yorukoglu* and *A.N. Celik*, "A critical review on the estimation of daily global solar radiation from sunshine duration", Energy Convers. Manag., **vol.** 47, pp. 2441–2450, 2006.

[13] *H. Morf*, "A stochastic solar irradiance model adjusted on the Ångström-Prescott regression", Solar Energy, **vol.** 87, pp. 1–21, 2013.

[14] *K. Bakirci*, "Models for the estimation of diffuse solar radiation for typical cities in Turkey", Energy, Elsevier, **vol.** 82(C), pp. 827–838, 2015.

[15] *Y.A. Kaplan*, "A new model for predicting the global solar radiation", Environ. Prog. Sustainable Energy, **vol.** 37, pp. 870–880, 2017.

[16] *C. Paoli*, *C. Voyant, M. Muselli* and *M.L. Nivet*, "Forecasting of preprocessed daily solar radiation time series using neural networks", Solar Energy, Elsevier, **vol.** 84, pp. 2146–2160, 2010.

[17] *W. Zhu, B. Wu, N. Yan, Z. Ma, L. Wang, W. Liu, Q. Xing* and *J. Xu*, "Estimating sunshine duration using hourly total cloud amount data from a geostationary meteorological satellite, Atmosphere, **vol.** 11, no. 1, 26, 2020.

[18]. *J. Duffie* and *W. Beckman*, "Solar engineering of thermal processes", John Wiley & Sons, New York, 2013.

[19] *V. Badescu, C. A. Gueymard, S. Cheval, C. Oprea, M. Baciu, A. Dumitrescu, F. Iacobescu, I. Milos* and *C. Rada*, "Accuracy analysis for fifty-four clear-sky solar radiation models using routine hourly global irradiance measurements in Romania", Renewable Energy, **vol.** 55, pp. 85–103, 2013.

[20] *G. Ranferi, M.A. Reyes* and *H.C. Rosu*, "A note on Verhulst's logistic equation and related logistic maps", J. Phys., A 43, 205204, 2010.

[21] *J. Almorox* and *C. Hontoria*, "Global solar radiation estimation using sunshine duration in Spain", Energy Convers. Manag., **vol.** 45 (9-10), pp. 1529–1535, 2004.

[22] *M. Iqbal*, "An introduction to solar radiation", Academic Press, Don Mills, Ontario, 1983.

[23] *A. Angström*, "Solar and terrestrial radiation", Q.J.R. Meteorol. Soc., **vol.** 50, pp. 121–126, 1924.

[24] *K. Bakirci*, "Correlations for estimation of daily global solar radiation with hours of bright sunshine in Turkey", Energy, Elsevier, **vol.** 34 (4), pp. 485–501, 2009.

[25] *H. Ogelman, A. Ecevit* and *E. Tasdemiroğlu*, "A new method for estimating solar radiation from bright sunshine data", Solar Energy, **vol.** 33, pp. 619–625, 1984.

[26] V. Badescu, "Assessing the performance of solar radiation computing models and model selection procedures", J. Atmospheric Sol.-Terr. Phys., **vol.** 105–106, pp. 119–134, 2013.

[27] *C.C.Y. Ma* and *M. Iqbal*, "Statistical comparison of models for estimating solar radiation on inclined surfaces", Solar Energy, Elsevier, **vol.** 31 (3), 313–317, 1983.

[28] *R.J. Stone*, "Improved statistical procedure for the evaluation of solar radiation estimation models", Solar Energy, **vol.** 51, pp. 289–291, 1993.

[29] *D. Zwillinger* and *S. Kokoska*, "CRC standard probability and statistics tables and formulae", Chapman & Hall/CRC, Boca Raton, London, New York, Washington, D.C., 2000.