Research Article

A Comparative Study of Some Regression Models to Estimate the Global Solar Radiation on a Horizontal Surface from Sunshine Duration and Meteorological Parameters for Ghardaïa Site, Algeria

Kacem Gairaa and Yahia Bakelli

Unité de Recherche Appliquée en Energies Renouvelables (URAER), Centre de Développement des Energies Renouvelable (CDER), 47133 Ghardaïa, Algeria

Correspondence should be addressed to Kacem Gairaa; gisol47@gmail.com

Received 15 April 2013; Accepted 20 May 2013

Academic Editors: B. Chen, S. S. Kalligeros, S. Li, Z. Oktay, and M. Souliotis

Copyright © 2013 K. Gairaa and Y. Bakelli. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A comparison between some regression correlations for predicting the global solar radiation received on a horizontal plane has been processed. Seven models for estimating the global solar radiation from sunshine duration and two meteorological parameters (air temperature and relative humidity) are presented. The root mean square error (RMSE), mean bias error (MBE), correlation coefficient (CC), and percentage error ($e$) have been also computed to test the accuracy of the proposed models. Comparisons between the measured and the calculated values have been made. The results obtained show that the linear and quadratic models are the most suitable for estimating the global solar radiation from sunshine duration, and for the models based on meteorological parameters, Abdalla and Ojosu’s models give the best performance with a CC of 0.898 and 0.892, respectively.

1. Introduction

Solar energy applications require a complete knowledge and detailed analysis about the potential of the site, so a database at ground level is an important feature of solar energy systems.

Measurements of global solar radiation reaching the surface of the earth and its two components direct and diffuse are essential in the most research fields of solar energy. The daily values as well as the monthly ones are needed to evaluate the performance of existing solar devices and estimate the efficiency of the future installations [1, 2].

When the site under consideration is equipped with a radiometric station operates regularly for several years, it will be easier to exploit solar energy resource. However, in most cases, there are no local measurements and we must resort to approximate methods to predict the characteristics of solar radiation.

In order to estimate the amount of solar energy incident on a horizontal surface, many models have been developed which relate the global solar radiation (H) with the sunshine duration. The first one is that proposed by Angstrom [3–5] which predicts the monthly average daily global solar radiation from sunshine duration and two meteorological parameters (air temperature and relative humidity) are presented. The root mean square error (RMSE), mean bias error (MBE), correlation coefficient (CC), and percentage error ($e$) have been also computed to test the accuracy of the proposed models. Comparisons between the measured and the calculated values have been made. The results obtained show that the linear and quadratic models are the most suitable for estimating the global solar radiation from sunshine duration, and for the models based on meteorological parameters, Abdalla and Ojosu’s models give the best performance with a CC of 0.898 and 0.892, respectively.

In this work, a comparative study between seven models to estimate the global solar radiation on a horizontal surface from sunshine duration and some meteorological parameters
has been conducted, using the local data of Ghardaïa region. The comparison between the estimated and the measured global solar radiation has been illustrated.

2. Database

The data used in this study are taken from the solar radiation laboratory located in the Applied Research Unit for Renewable Energies (URAER). Ghardaïa site is considered an arid and dry area, located in the south of Algeria, about 600 Km south of the capital city; it is framed by the following geographical coordinates: latitude of 32\(^\circ\)36\(\prime\)N, longitude of 3\(^\circ\)48\(\prime\)E, and altitude of 450 m above MSL. The global solar radiation data are measured via an EKO MS-64 Pyranometer with the short wave sensitivity of (7.0 mV/kW/m\(^2\)); more technical specifications of this Pyranometer are described in Table 1. The air temperature and relative humidity measurement data were made by TECNOEL sonde Thermo-Igrometriche; its calibration accuracy is \(\pm 1.5\%\) and temperature sensitivity is \(0.1(\text{C}/\text{mV})\). A data logger and a Campbell Scientific CR10X data acquisition system were used for reading the measurements. From the raw data stored for every five minutes, the mean, maximum, and minimum hourly values were calculated. From the hourly data set, daily and monthly statistics were made for the global solar radiation and meteorological data.

The average hours of sunshine duration are illustrated in Figure 1 for the site under consideration; the highest values were in the summer season with a mean value of 12.23 hours, while the lowest values were in the winter season with a mean value of 7.43 hours. The monthly average daily global solar radiation, air temperature, and relative humidity are shown in Figures 2, 3, and 4.

3. Models Used

A brief description of the mathematical expression of the various models proposed in the present paper is given below.

3.1. Angstrom-Prescott Model (Linear Model) [4, 24]. The Angstrom-Prescott model is widely described by the authors in the study of solar radiation since Prescott has developed in 1940; the general form of this model is given by

\[
\frac{H}{H_0} = a + b \left( \frac{SS}{SS_0} \right),
\]

where \(H_0\) is the monthly average daily extraterrestrial radiation (MJ/m\(^2\)·day) which can be expressed as

\[
H_0 = \frac{24}{\pi} G_{SC} \left( 1 + 0.033 \frac{360 \cdot D_n}{365} \right) \cdot \left( \cos \varphi \cos \delta \sin \omega + \frac{2\pi}{360} \omega \sin \varphi \sin \delta \right),
\]

where \(\omega\) is sunset hour angle in degree defined as

\[
\omega = \cos^{-1} \left( - \tan \varphi \tan \delta \right).
\]
number of the year starting from the first of January, and $\delta$ is the declination angle as given below:

$$\delta = 23.45 \sin \left( \frac{360(284 + D_n)}{365} \right).$$ \hspace{1cm} (4)

SS is the monthly daily sunshine duration; SS$_0$ is the maximum possible monthly average daily sunshine hours or the day length.

3.2. Ögelman et al’s Model (Quadratic Model) [24]. Ögelman et al’s have correlated $(H/H_0)$ with $(SS/SS_0)$ in the form of second-order polynomial equation:

$$\frac{H}{H_0} = a + b \left( \frac{SS}{SS_0} \right) + c \left( \frac{SS}{SS_0} \right)^2. \hspace{1cm} (5)$$

3.3. Ampratwum et al’s Model (Logarithmic Model) [14, 24]. In order to have more precision in the estimation of the global solar radiation, Ampratwum et al. have developed a logarithmic form of linear model as described below:

$$\frac{H}{H_0} = a + b \log \left( \frac{SS}{SS_0} \right). \hspace{1cm} (6)$$

3.4. Almorox et al’s Model (Exponential Model) [8, 24]. The exponential correlation between $(H/H_0)$ and $(SS/SS_0)$ is given by Almorox et al. as

$$\frac{H}{H_0} = a + be^{(SS/SS_0)}. \hspace{1cm} (7)$$

3.5. Abdalla’s Model [25, 26]. Abdalla’s model correlates $(H/H_0)$ with the sunshine duration, maximum air temperature, and average relative humidity, in order to increase the accuracy of the estimating coefficients. This model is based on the Gopinathan model [26], which studied the variation of $(H/H_0)$ as a function of the latitude and the elevation of the site, the fraction of insolation, air temperature, and the maximum average relative humidity. This model has the following expression:

$$\frac{H}{H_0} = a + b\left( \frac{SS}{SS_0} \right) + c(T_{\text{max}}) + d\left( \text{RH} \right), \hspace{1cm} (8)$$

where $T_{\text{max}}$ is the daily mean maximum air temperature and RH is the daily mean relative humidity as a percentage.

3.6. Ojosu and Komolafe’s Model [31]. Ojosu and Komolafe proposed the equation below:

$$\frac{H}{H_0} = a + b\left( \frac{SS}{SS_0} \right) + c\left( \frac{T_{\text{min}}}{T_{\text{max}}} \right) + d\left( \frac{\text{RH}_{\text{min}}}{\text{RH}_{\text{max}}} \right), \hspace{1cm} (9)$$

where $T_{\text{min}}$, $T_{\text{max}}$, $\text{RH}_{\text{min}}$, and $\text{RH}_{\text{max}}$ are the mean minimum air temperature, mean maximum air temperature, mean minimum relative humidity, and mean maximum relative humidity, respectively.

3.7. Hargreaves et al’s Model [29, 30, 32]. Hargreaves et al. were the first to propose a procedure to estimate the global solar radiation by using the difference between daily maximum and daily minimum air temperature and extraterrestrial radiation. The proposed equation has the following form:

$$\frac{H}{H_0} = a + b(T_{\text{max}} - T_{\text{min}})^{0.5}, \hspace{1cm} (10)$$

where $T_{\text{min}}$ and $T_{\text{max}}$ are the mean minimum air temperature and mean maximum air temperature.

A program written in MATLAB allowing to determine the empirical constants $a$, $b$, $c$, and $d$ for each described model and compared models was thus found with the ground measurements.
4. Statistical Test

The accuracy of the estimated models will be judged by the statistical indicators, such as the correlation coefficient (CC), mean bias error (MBE), root mean square error (RMSE), and the percentage error (e). These indicators are usually applied in the comparison of solar radiation models. The mean square error provides information about the performance of correlations, which allows comparison of the real differences between the estimated values and the measured ones; a low RMSE is desirable. The MBE provides the long-term performance of the model; in general, the positive MBE shows overestimation while the negative MBE indicates underestimation. The correlation coefficient reflects the quality of the model; the more CC close to 1, the more the better quality. The expression of each statistical indicator is given:

\[
CC = 1 - \frac{\sum_{i=1}^{n} (H_{i,m} - H_{i,e})^2}{\sum_{i=1}^{n} (H_{i,m} - \overline{H}_m)^2},
\]

\[
MBE = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{H_{i,e} - H_{i,m}}{H_{i,m}} \right),
\]

\[
RMSE = \left[ \frac{1}{n} \sum_{i=1}^{n} \left( \frac{H_{i,e} - H_{i,m}}{H_{i,m}} \right)^2 \right]^{1/2},
\]

\[
e = \left( \frac{H_{i,m} - H_{i,e}}{H_{i,m}} \right) \times 100,
\]

where \(H_{i,e}\) and \(H_{i,m}\) are the \(i\)th estimated and measured values, \(\overline{H}_m\) is the mean of observed data, and \(n\) is the total number of observations.

5. Results and Discussion

The results of the validation of the models that estimate the global solar radiation from sunshine duration on the monthly basis are presented in Tables 2, 3, 4, and 5. The analysis of the measured and calculated values shows that the maximum of the global solar radiation is observed in June while the minimum values are appearing in December.

For the linear model, the correlation coefficient is significant for all months of the year except in November where its value is relatively low. The quality of the estimate is a slightly worse for
Table 4: Logarithmic model validation.

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>CC</th>
<th>RMSE</th>
<th>MBE</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>0.741</td>
<td>0.079</td>
<td>0.792</td>
<td>0.077</td>
<td>0.0235</td>
</tr>
<tr>
<td>February</td>
<td>0.795</td>
<td>0.314</td>
<td>0.693</td>
<td>0.102</td>
<td>0.0536</td>
</tr>
<tr>
<td>March</td>
<td>0.734</td>
<td>0.079</td>
<td>0.651</td>
<td>0.151</td>
<td>0.0566</td>
</tr>
<tr>
<td>April</td>
<td>0.792</td>
<td>0.218</td>
<td>0.917</td>
<td>0.061</td>
<td>0.0369</td>
</tr>
<tr>
<td>May</td>
<td>0.757</td>
<td>0.121</td>
<td>0.825</td>
<td>0.067</td>
<td>0.0338</td>
</tr>
<tr>
<td>June</td>
<td>0.734</td>
<td>0.112</td>
<td>0.862</td>
<td>0.061</td>
<td>0.0294</td>
</tr>
<tr>
<td>July</td>
<td>0.735</td>
<td>0.164</td>
<td>0.816</td>
<td>0.058</td>
<td>0.0393</td>
</tr>
<tr>
<td>August</td>
<td>0.743</td>
<td>0.176</td>
<td>0.805</td>
<td>0.048</td>
<td>0.0321</td>
</tr>
<tr>
<td>September</td>
<td>0.714</td>
<td>0.097</td>
<td>0.613</td>
<td>0.091</td>
<td>0.0281</td>
</tr>
<tr>
<td>October</td>
<td>0.708</td>
<td>0.121</td>
<td>0.864</td>
<td>0.065</td>
<td>0.0330</td>
</tr>
<tr>
<td>November</td>
<td>0.724</td>
<td>0.09</td>
<td>0.487</td>
<td>0.088</td>
<td>0.0259</td>
</tr>
<tr>
<td>December</td>
<td>0.749</td>
<td>0.176</td>
<td>0.886</td>
<td>0.088</td>
<td>0.0465</td>
</tr>
</tbody>
</table>

Table 5: Exponential model validation.

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>CC</th>
<th>RMSE</th>
<th>MBE</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>0.14</td>
<td>0.261</td>
<td>0.861</td>
<td>0.04</td>
<td>−0.00071</td>
</tr>
<tr>
<td>February</td>
<td>0.128</td>
<td>0.26</td>
<td>0.713</td>
<td>0.063</td>
<td>−0.00047</td>
</tr>
<tr>
<td>March</td>
<td>0.342</td>
<td>0.173</td>
<td>0.878</td>
<td>0.032</td>
<td>0.00003</td>
</tr>
<tr>
<td>April</td>
<td>0.295</td>
<td>0.195</td>
<td>0.901</td>
<td>0.025</td>
<td>−0.00026</td>
</tr>
<tr>
<td>May</td>
<td>0.321</td>
<td>0.181</td>
<td>0.868</td>
<td>0.032</td>
<td>−0.00077</td>
</tr>
<tr>
<td>June</td>
<td>0.295</td>
<td>0.183</td>
<td>0.916</td>
<td>0.024</td>
<td>−0.00040</td>
</tr>
<tr>
<td>July</td>
<td>0.335</td>
<td>0.163</td>
<td>0.846</td>
<td>0.025</td>
<td>0.00078</td>
</tr>
<tr>
<td>August</td>
<td>0.403</td>
<td>0.132</td>
<td>0.67</td>
<td>0.032</td>
<td>−0.00022</td>
</tr>
<tr>
<td>September</td>
<td>0.192</td>
<td>0.224</td>
<td>0.774</td>
<td>0.055</td>
<td>−0.00044</td>
</tr>
<tr>
<td>October</td>
<td>0.236</td>
<td>0.199</td>
<td>0.901</td>
<td>0.028</td>
<td>−0.0015</td>
</tr>
<tr>
<td>November</td>
<td>0.283</td>
<td>0.187</td>
<td>0.621</td>
<td>0.045</td>
<td>0.104</td>
</tr>
<tr>
<td>December</td>
<td>0.188</td>
<td>0.225</td>
<td>0.822</td>
<td>0.073</td>
<td>−0.0096</td>
</tr>
</tbody>
</table>

the months of the February and November when the RMSE is between 0.064 and 0.069.

For the quadratic correlation, the result shows a slight improvement in the accuracy of the model for the month of August, where the coefficient of determination is equal to 0.83 (compared to the linear model, CC = 0.73); the quality of estimation is also improved in the months of June and December with an RMSE of 0.018 and 0.039, respectively. The regression based on the logarithmic and exponential models did not improve significantly the accuracy of estimation of the global solar radiation than the two previous models, as seen from the values of the CC, MBE, and RMSE. (Correlation coefficient reached his minimum values in November (CC = 0.48) for the logarithmic model.)

Another manner of the comparison between the measured and the predicted values was illustrated by the scatter diagram plotted in Figures 5, 6, 7, and 8. A strong scattering of the points in the superior or in the inferior half of the regression line indicates, respectively, an overestimate or an underestimate.

For the linear model, the dispersion of estimated values around the regression line is fairly low in the months of January, March-August, October, and December, where the correlation coefficient ranges between 0.85 and 0.94; while this dispersion is small in the quadratic model, this indicates an excellent fitting between the global radiation and the sunshine duration.

For the logarithmic model, the dispersion of the estimated values is strong especially in the months of January, March, and August, and it is low in the months of February, September and November, while this dispersion remains low for the exponential model in the months of February, September, and November and it is strong too for the other months of the year.

The confrontation between estimated and measured monthly average global solar radiation by the different models is shown in Table 6, which indicated that the percentage error never exceeds ±10% for all models. It was 1.903% as maximum value for the linear model, 1.974% for the quadratic model, −8.489% for the logarithmic model, and 3.217% for the exponential model. Among the four proposed models, we found in the first rank, the linear and quadratic correlations which give results that agree well with the measured values, compared with the other two models, which underestimate or overestimate measurements.

The validation of the various models on the annual basis is shown in Table 7. For the correlation based on the sunshine duration, we can note that the linear and quadratic
Figure 5: Scatter diagram of linear model.
Figure 6: Scatter diagram of quadratic model.
Figure 7: Scatter diagram of logarithmic model.
Figure 8: Scatter diagram of exponential model.
Several empirical models have been tested to estimate the global solar radiation on a horizontal surface, using the sunshine duration and the meteorological parameters. The goal of this paper is the validation of seven models to predict the monthly and the annual global solar radiation within a dataset of Ghardaia area and, thus, select the suitable model. The models were compared on the basis of statistical tests.

Depending on the obtained results, we can conclude that the linear and the quadratic models are the most accurate for estimating the monthly global solar radiation on a horizontal surface according to the statistical tests (CC: 0.956, RMSE: 0.018, and MBE: 43.10^{-3}); the percentage error never exceeds ±5%, which indicate an excellent fitting between the global solar radiation and sunshine duration. On the basis of annual estimation of global solar radiation, and for the models based on the sunshine duration, always the linear and the quadratic models provide the best accuracy (CC: 0.787, RMSE: 0.0451, and MBE: −0.00073); whereas for models using the meteorological parameters in addition to sunshine duration, we found that Abdalla and Ojosu’s models are the most suitable with a CC of 0.898 and 0.892, respectively, and the relative error ranges between −0.289% and 0.639%.

### References


