Statistical Comparison Between Empirical Models And Artificial Neural Network Method For Global Solar Radiation At Qena, Egypt

Emad Ali Ahmed
Physics department, faculty of sciences
South Valley University
Qena, Egypt
emad.amer@sci.svu.edu.eg

Abstract—Seven empirical models and the artificial neural network model (ANN) are applied and validated to estimate the daily and monthly global solar radiation on a horizontal surface in Qena, Upper Egypt. Statistical comparison has been done between all these models. The results demonstrated that ANN model with MBE, RMSE, MPE, R, NSE and t values of -0.0692, 0.5338, -0.2647, 0.9946, 0.9890 and 2.4934 respectively, are superior to the other models. This model can be applied as a reliable and accurate global solar radiation estimator for Qena and regions with similar climatic conditions.

Keywords—Global solar radiation; regression models; statistical analysis; artificial neural network; empirical models

I. INTRODUCTION

Egypt has an intention for using further alternative resources of energy due to several economic reasons and more importantly other environmental protection goals. Global solar radiation (GSR) is considered as one of the most important sources of energy that reach our planet. Egypt has great advantageous position, belongs to the global sun-belt, as shown in Figure 1 [1].

Egypt has an intention for using further alternative resources of energy due to several economic reasons and more importantly other environmental protection goals. Global solar radiation (GSR) is considered as one of the most important sources of energy that reach our planet. Egypt has great advantageous position, belongs to the global sun-belt, as shown in Fig. 1 [1].

Solar radiation measurements have effect in several applications such as solar energy systems, agriculture, and architecture; but they are not easily available in all places because of the cost, regular maintenance and tuning requirements. The availability of solar radiation measures is not an easy matter [2]. Thus, suggesting adequate alternatives to estimate such measures based on easily available climatologic data. Assessing global solar radiation is an important requirement for feasibility analysis, design and implementation of solar energy systems. However, because of the unavailability of the instrument in many regions, atmospheric parameters at a particular location are being to predict the global solar radiation in that location.

Fig. 1. Egypt is part of the solar belt.

There are many models to predict long-term daily and monthly average solar radiation in different regions using different combinations of measured weather parameters. The availability of a solar radiation model, in a particular region, is very useful in estimating the amount of power that could be generated from a particular solar energy system [3].

First theoretical model has been introduced by Angstrom [4] for estimating global solar radiation based on sunshine duration. Page [5] and Prescot [6] reconsidered this model to calculate monthly average of the daily global radiation, on a horizontal surface from monthly average daily total insolation on an extraterrestrial horizontal surface. Many models have been developed to estimate the amount of global solar radiation on horizontal surfaces using various climatic parameters, such as sunshine duration, cloud cover, humidity, maximum and minimum ambient temperatures, wind speed, etc. [7-11] Few papers appeared concerning estimation of solar radiation over various places in Egypt [12-14].

An artificial neural network (ANN) provides a computationally efficient way of determining an empirical, possibly nonlinear relationship between a number of inputs and one or more outputs. ANN has been applied for modeling, identification, optimization,
prediction, forecasting and control of complex systems. ANN models are type of solar prediction models and there have been several articles that have used artificial neural networks for predicting solar radiation [15-19].

The aim of the study is the estimation the global solar radiation by using artificial neural network and some empirical models with atmospheric parameters of relative humidity, maximum temperature, sunshine hours and cloud cover. Various regression analyses are performed between the parameters to find the relevant coefficients for establishing the models. Conventional statistical indices such as RMSE, MBE, MABE, R², t-statistic, MPE and MAPE are applied to the estimated values for each model and the actual data to evaluate model performance and to determine the most advantageous.

II. DATA AND METHODOLOGY

A. Data

In this study, the comparison process has been completed in Qena (26.170 N, 32.70 E) which located in the Upper Egypt about 600 Km south of Cairo. The climate of Qena is very hot dry in summer, cold in winter and rarely raining. Also, it receives a large quantity of solar radiation, especially in summer [20]. South valley university station at Qena, which is one of the stations guides of the Egyptian Meteorological authority, provides us with the relevant data. The data collected cover a period of thirteen years from 2001 to 2013, as shown Table 1.

B. Extraterrestrial radiation.

Solar radiation incident outside the earth's atmosphere is called extraterrestrial solar radiation. The daily extraterrestrial radiation Go was calculated from the following equation [21, 22].

\[
G_o = \frac{2441sc}{\pi} \left[ 1 + 0.033 \cos \left( \frac{360D}{365} \right) \right] \cos \phi \cos \delta \sin \omega_s + \frac{\pi \omega_s}{180} \sin \delta
\]

Where:

\( I_{sc} = 1367 \text{ Wm}^{-2} \) is the solar constant [23].

\( D \) is the Julian day number;

\( \phi \) is the latitude;

\( \delta \) is the declination angle;

\( \omega_s \) is the sunset hour.

\( \delta \) and \( \omega \) are given from these formulae:

\[
\delta = 23.45 \sin \left( \frac{360}{365} \left( 284 + D \right) \right)
\]

\[
\omega_s = \cos^{-1}[-\tan \phi \tan \delta]
\]

The maximum possible sunshine duration \( S_o \) was calculated using the following equation.

\[
S_o = \frac{2}{15} \omega_s
\]

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

C. Models used

The Angstrom correlation (1924) [4] has served as a basic approach to estimate global radiation for a long time. Prescott (1940) [24] put the correlation in a better form known as the Angstrom-Prescott model as;

\[
\frac{G}{G_o} = a + b \frac{S}{S_o}
\]

Where \( G \) is the average daily global radiation, \( G_o \) is average daily extraterrestrial radiation, \( S \) is the day length, \( S_o \) is the maximum possible sunshine duration, and \( a \) and \( b \) are empirical coefficients.

Angstrom’s calculations for Swedish climatic conditions showed that the sum of the empirical regression coefficients \( a \) and \( b \) is equal to unity. Also, investigations for various world geographical latitudes demonstrated that the value of this sum increases as the latitude increases, with some exceptions.

<table>
<thead>
<tr>
<th>years</th>
<th>January</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>July</th>
<th>August</th>
<th>September</th>
<th>October</th>
<th>November</th>
<th>December</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>21</td>
<td>0</td>
<td>31</td>
<td>27</td>
<td>31</td>
<td>31</td>
<td>30</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>2002</td>
<td>28</td>
<td>28</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>31</td>
<td>28</td>
<td>28</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>2003</td>
<td>31</td>
<td>30</td>
<td>30</td>
<td>28</td>
<td>0</td>
<td>31</td>
<td>29</td>
<td>31</td>
<td>31</td>
<td>0</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>2004</td>
<td>29</td>
<td>30</td>
<td>29</td>
<td>28</td>
<td>8</td>
<td>30</td>
<td>29</td>
<td>30</td>
<td>30</td>
<td>0</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>2005</td>
<td>31</td>
<td>31</td>
<td>27</td>
<td>29</td>
<td>30</td>
<td>31</td>
<td>0</td>
<td>31</td>
<td>28</td>
<td>0</td>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td>2006</td>
<td>30</td>
<td>29</td>
<td>28</td>
<td>30</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>30</td>
<td>30</td>
<td>0</td>
<td>30</td>
<td>9</td>
</tr>
<tr>
<td>2007</td>
<td>31</td>
<td>31</td>
<td>20</td>
<td>31</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>31</td>
<td>31</td>
<td>0</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>2008</td>
<td>31</td>
<td>31</td>
<td>11</td>
<td>28</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>31</td>
<td>31</td>
<td>0</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>2009</td>
<td>30</td>
<td>5</td>
<td>30</td>
<td>30</td>
<td>29</td>
<td>0</td>
<td>0</td>
<td>30</td>
<td>30</td>
<td>0</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>2010</td>
<td>30</td>
<td>2</td>
<td>30</td>
<td>30</td>
<td>31</td>
<td>0</td>
<td>31</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>2011</td>
<td>28</td>
<td>29</td>
<td>28</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>30</td>
<td>30</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>2012</td>
<td>31</td>
<td>14</td>
<td>30</td>
<td>0</td>
<td>31</td>
<td>0</td>
<td>31</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>2013</td>
<td>361</td>
<td>319</td>
<td>255</td>
<td>242</td>
<td>235</td>
<td>273</td>
<td>270</td>
<td>0</td>
<td>323</td>
<td>193</td>
<td>361</td>
<td>0</td>
</tr>
</tbody>
</table>
Glover and McCulloch presented coefficient a as a function of the cosine of the latitude, for latitudes under 60 and suggested that [25, 26].

\[
\frac{G}{G_o} = 0.29 \cos(\theta) + 0.52 \frac{S}{S_o} \quad (2)
\]

Akinoglu and Ecevit (1990) [27] suggested a quadratic correlation between the ratio of \( \frac{G}{G_o} \) and \( n \) to estimate the values of global solar radiation

\[
\frac{G}{G_o} = a + b \frac{S}{S_o} + c \left( \frac{S}{S_o} \right)^2 \quad (3)
\]

In order to have more precision in the estimation of the global solar radiation, Ampratwum and Dorvlo (1999) [28] have developed a logarithmic form of linear model as:

\[
\frac{G}{G_o} = a + b \log \frac{S}{S_o} \quad (4)
\]

Almorox [29, 30] presents the exponential correlation between \( \frac{G}{G_o} \) and \( \frac{S}{S_o} \) with the following expression:

\[
\frac{G}{G_o} = a + b S_o \quad (5)
\]

Hargreaves et al. [30-33] 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:

\[
\frac{G}{G_o} = a + b(T_{\text{max}} - T_{\text{min}})^2 \quad (6)
\]

According to the Gopinathan model [10], which studied the variation of \( \frac{G}{G_o} \) as a function of the latitude and the elevation of the site, the fraction of insolation, air temperature, and the maximum average relative humidity, Abdalla (1994) [34] propose model correlates \( \frac{G}{G_o} \) with the sunshine duration, maximum air temperature, and average relative humidity, in order to increase the accuracy of the estimating coefficients, as follow:

\[
\frac{G}{G_o} = a + b \frac{S}{S_o} + c T_{\text{max}} + d \text{RH} \quad (7)
\]

Emad (2013) [15] uses Artificial Neural Network (ANN) method to estimate Global solar radiation in Qena based on the number of sunshine hours, day number and location coordinates. A feed-forward back-propagation neural network was used in this study. A typical neural network consists of an input, a hidden, and output layer. Other components include a neuron, weight, and a transfer function as shown in Fig. 2.

Fig. 2. Typical neuron in a feed forward network.

D. Comparison and Statistical techniques

There are many parameters which deal with the assessment and comparison of daily solar radiation estimation models. Here mathematical expressions and physical meaning of some of these parameters.

1. The Mean Bias Error (MBE)

\[
\text{MBE} = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{G_{i,\text{calc}} - G_{i,\text{meas}}}{G_{i,\text{meas}}} \right)
\]

This test helps to calculate the error or the deviation of the calculated value from the measured value and provides information on long-term performance. A low MBE value is desired. A negative value gives the average amount of underestimation in the calculated value. So, one drawback of these two mentioned tests is that overestimation of an individual observation will cancel underestimation in a separate observation.

2. The Root Mean Square Error (RMSE)

\[
\text{RMSE} = \left[ \frac{1}{n} \sum_{i=1}^{n} \left( \frac{G_{i,\text{calc}} - G_{i,\text{meas}}}{} \right)^2 \right]^{\frac{1}{2}}
\]

The value of RMSE is always positive, representing zero in the ideal case. The normalized root mean square error gives information on the short term performance of the correlations by allowing a term by term comparison of the actual deviation between the predicted and measured values. The smaller the value, the better is the model's performance [35].

3. The Mean Percentage Error (MPE)

\[
\text{MPE} = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{G_{i,\text{calc}} - G_{i,\text{meas}}}{} \right) \times 100
\]

This is one of the measures used to evaluate forecasts using forecast errors. Forecast error is defined as actual observation minus forecast. The mean percentage error is the average or mean of all the percentage errors. A percentage error between −10% and +10% is considered acceptable [36].

4. The Coefficient of Correlation (R)

\[
R = \sqrt{\frac{n \sum G_{\text{calc}} G_{\text{meas}} - \left( \sum G_{\text{calc}} \right) \left( \sum G_{\text{meas}} \right)}{\sqrt{n \sum (G_{\text{calc}})^2 - \left( \sum G_{\text{calc}} \right)^2} \sqrt{n \sum (G_{\text{meas}})^2 - \left( \sum G_{\text{meas}} \right)^2}}
\]
This parameter measures the strength and the direction of a linear relationship between the measured and estimated values. The value of R is between -1 and +1. The + and – signs are used for positive linear correlations and negative linear correlations, respectively. If there is no linear correlation or a weak linear correlation, R is close to 0. A value near zero means that there is a random, nonlinear relationship between the two variables but it should approach to 1 as closely as possible for better modeling.

5. The Nash–Sutcliffe Equation (NSE)

\[
NSE = 1 - \frac{1}{n} \sum_{i=1}^{n} \left( \frac{G_{\text{calc},i} - G_{\text{meas},i}}{G_{\text{meas},i}} \right)^2
\]

Where, \( G_{\text{meas}} \) is the mean measured global radiation.

To improve the results and better comparison this parameter is also selected as an evaluation criterion. A model is more efficient when NSE is closer to 1 [37]. The errors that have been estimated help to compare the models but they do not make the model statistically significant.

6. t-Statistic Test

As defined by Student [38] in one of the tests for mean values, the random variable t with \( n - 1 \) degrees of freedom may be written here as follows:

\[
t = \sqrt{\frac{(n - 1) \times (\text{MBE})^2}{\text{(RMSE)}^2 - (\text{MBE})^2}}
\]

The t-statistic allows models to be compared and at the same time it is carried out to determine statistical significance of the predicted values by the models. The smaller the value of t the better is the performance. To determine whether a model’s estimates are statistically significant, one simply has to determine, from standard statistical tables, the critical t value. For the model’s estimates to be judged statistically significant at the calculated t value must be less than the critical value.

III. RESULTS AND DISCUSSIONS

In the study, the daily average of global solar radiation, the day length, temperature and relative humidity were taken of the average of the values in the corresponding period from 2001 to 2013 for the site under consideration as shown in Figs. 3-6.

Regression analysis is performed by using Microsoft Excel 2010 and SPSS Statistics 22 to find values of the regression coefficients that minimize the sum of the squared residual values. The obtained regression coefficients after analysis are given in Table 2 for each model.
Table 2. Obtained regression coefficients.

<table>
<thead>
<tr>
<th>Model No.</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eq. 1</td>
<td>0.4061</td>
<td>0.2850</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Eq. 2</td>
<td>0.2603</td>
<td>0.5200</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Eq. 3</td>
<td>-0.7275</td>
<td>2.8960</td>
<td>-1.4986</td>
<td>-</td>
</tr>
<tr>
<td>Eq. 4</td>
<td>0.6891</td>
<td>0.5762</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Eq. 5</td>
<td>0.3726</td>
<td>0.1178</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Eq. 6</td>
<td>0.5736</td>
<td>0.0004</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Eq. 7</td>
<td>0.5165</td>
<td>0.3080</td>
<td>-0.0013</td>
<td>-0.0027</td>
</tr>
</tbody>
</table>

By using the developed models, the total daily global solar radiation values are estimated throughout the study period. Also, the monthly average of daily global solar radiation is calculated as shown in Table 3.

Table 3. Measured and calculated monthly averages of daily global solar radiation (Mj/m\(^2\).day) at Qena (2001-2013).

<table>
<thead>
<tr>
<th>Month</th>
<th>Measured</th>
<th>Eq. 1</th>
<th>Eq. 2</th>
<th>Eq. 3</th>
<th>Eq. 4</th>
<th>Eq. 5</th>
<th>Eq. 6</th>
<th>Eq. 7</th>
<th>ANN model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feb.</td>
<td>18.44</td>
<td>18.28</td>
<td>20.08</td>
<td>18.31</td>
<td>18.29</td>
<td>18.29</td>
<td>16.03</td>
<td>18.00</td>
<td>18.20</td>
</tr>
<tr>
<td>Apr.</td>
<td>25.05</td>
<td>24.22</td>
<td>26.27</td>
<td>24.29</td>
<td>24.23</td>
<td>24.23</td>
<td>21.66</td>
<td>24.93</td>
<td>24.87</td>
</tr>
<tr>
<td>May</td>
<td>27.11</td>
<td>26.17</td>
<td>28.64</td>
<td>26.23</td>
<td>26.19</td>
<td>26.19</td>
<td>23.10</td>
<td>27.18</td>
<td>27.07</td>
</tr>
<tr>
<td>Jun.</td>
<td>28.07</td>
<td>27.17</td>
<td>30.08</td>
<td>27.16</td>
<td>27.18</td>
<td>27.18</td>
<td>23.55</td>
<td>27.83</td>
<td>27.86</td>
</tr>
<tr>
<td>Jul.</td>
<td>27.36</td>
<td>26.84</td>
<td>29.73</td>
<td>26.83</td>
<td>26.85</td>
<td>26.86</td>
<td>23.24</td>
<td>27.35</td>
<td>27.35</td>
</tr>
<tr>
<td>Aug.</td>
<td>25.73</td>
<td>25.55</td>
<td>28.34</td>
<td>25.50</td>
<td>25.56</td>
<td>25.57</td>
<td>22.08</td>
<td>25.74</td>
<td>25.91</td>
</tr>
<tr>
<td>Sept.</td>
<td>23.11</td>
<td>22.83</td>
<td>25.20</td>
<td>22.88</td>
<td>22.84</td>
<td>22.84</td>
<td>19.87</td>
<td>22.87</td>
<td>22.99</td>
</tr>
<tr>
<td>Nov.</td>
<td>15.61</td>
<td>16.37</td>
<td>18.17</td>
<td>16.31</td>
<td>16.38</td>
<td>16.39</td>
<td>14.14</td>
<td>15.72</td>
<td>15.52</td>
</tr>
</tbody>
</table>

Table 4. The calculated values of the statistical indices for each model at Qena

<table>
<thead>
<tr>
<th>Model No.</th>
<th>(\text{MBE})</th>
<th>(\text{RMSE})</th>
<th>(\text{MPE})</th>
<th>(R)</th>
<th>(\text{NSE})</th>
<th>(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eq. 1</td>
<td>-0.1528</td>
<td>0.7978</td>
<td>0.1199</td>
<td>0.9951</td>
<td>0.9755</td>
<td>3.7232</td>
</tr>
<tr>
<td>Eq. 2</td>
<td>2.0097</td>
<td>2.1064</td>
<td>10.1296</td>
<td>0.9925</td>
<td>0.8295</td>
<td>60.7638</td>
</tr>
<tr>
<td>Eq. 3</td>
<td>-0.1400</td>
<td>0.7851</td>
<td>0.1685</td>
<td>0.9953</td>
<td>0.9763</td>
<td>3.4579</td>
</tr>
<tr>
<td>Eq. 4</td>
<td>-0.1424</td>
<td>0.7927</td>
<td>0.1664</td>
<td>0.9952</td>
<td>0.9759</td>
<td>3.4845</td>
</tr>
<tr>
<td>Eq. 5</td>
<td>-0.1390</td>
<td>0.7947</td>
<td>0.1855</td>
<td>0.9951</td>
<td>0.9757</td>
<td>3.3905</td>
</tr>
<tr>
<td>Eq. 6</td>
<td>-2.8479</td>
<td>3.1364</td>
<td>-12.3615</td>
<td>0.9937</td>
<td>0.6220</td>
<td>41.3525</td>
</tr>
<tr>
<td>Eq. 7</td>
<td>-0.0999</td>
<td>0.5573</td>
<td>-0.2931</td>
<td>0.9943</td>
<td>0.9881</td>
<td>3.4780</td>
</tr>
<tr>
<td>ANN model</td>
<td>-0.0692</td>
<td>0.5338</td>
<td>-0.2647</td>
<td>0.9946</td>
<td>0.9890</td>
<td>2.4934</td>
</tr>
</tbody>
</table>
Fig. 7. Correlation of calculated and measured global solar radiation for each model.
IV. CONCLUSIONS

Assessing global solar radiation is an important requirement for feasibility analysis, design and implementation of solar energy systems. However, because of the unavailability of the instrument in many locations, atmospheric parameters at a particular location are being to predict the global solar radiation in that location. In this study, seven empirical models in addition to artificial neural network model (ANN) are developed and validated to estimate the daily and monthly global solar radiation on a horizontal surface in Qena, Upper Egypt. These models were studied based on the available climatic parameters of sunshine hour, maximum and minimum temperature, and relative humidity. According to the statistical comparison between all models, the ANN model is superior to the other models; it could be applied for estimation of global solar radiation in Qena and other similar climatic conditions regions.

REFERENCES

of the Royal Society of South Australia, Vol. 64, pp. 114-118.


