Estimation of Monthly Average Global Solar Radiation using Sunshine Hours

C. Ada Christa¹ C.I. Vijai Kumar² K. Sri Sindhu³

¹²³Prathyusha Institute of Technology

Abstract—In this thesis, the Angstrom-Prescott coefficients are determined by using Artificial Bee Colony algorithm and regression model in order to calculate the monthly average solar radiation for the particular region. In developing countries, the global solar radiation and its components are not available for all locations due to which there are a requirement of using different models for the estimation of global solar radiation that use climatological parameters of the locations. There are various empirical models are used to calculate the solar radiation but in this paper Angstrom-Prescott model is used to calculate the solar radiation of the particular region or area which provide the sunshine hours of that region. With the observation data of global solar radiation and sunshine-hours from one observation station, the Angstrom-Prescott formula is trained to determine all necessary parameters, by which the monthly average solar radiation data of the other weather stations with only sunshine-hour observation data can be derived. Several distribution maps of the global solar radiation are then generated using interpolation models in GIS, based on which spatial and temporal variation pattern of the monthly global solar radiation are discussed.

Key words: Global Solar Radiation, Angstrom Prescott Model, Sunshine hours, GIS

I. INTRODUCTION

Global solar radiation is essential to the earth and human beings. Because of the scarcity of solar radiation observation stations, climatologic calculation methods and interpolation techniques are commonly adopted to derive the monthly global solar radiation and generate the regional distribution map. Solar radiation is the direct form of abundant permanent solar energy resource available on earth, due to nuclear fusion on Sun. Earth surface is receiving about one hundred thousand tera watts (TW) of this renewable energy of solar power at earth’s surface at each moment. Clouds, gases, pollution (including aerosol) and other factor decreases this available power on surface and thus, earth gets about 800 times less solar energy from the Sun at each moment (Schiermeier et.al., 2008)[1]. Despite of the fact is that only 71 minutes of solar energy is good enough to satisfy the demand of solar energy n earth’s population for one year. As a matter of fact about one thousand watts per square meters of solar energy reaches at landmass of the earth.

In addition, the values of the average daily global radiation in the solar energy applications are the major important parameter, measurements of which are not available at every location due to cost, maintenance, and calibration requirements of the measuring equipment. In places where no measured values are available, a common application has been to determine this parameter by appropriate correlations which are empirically established using the measured data [2]. Several empirical models have been used to calculate solar radiation, utilizing available meteorological, geographical and climatological parameters such as sunshine hours (Kadir Bakirci, 2009; Koussa et al., 2009; Bulut and Buyukalaca, 2007; Akinoglu and Ecevit, 1990), air temperature (Fletcher, 2007), latitude (Raja, 1994), precipitation (Rietveld, 1978), relative humidity (Trabea and Shaltout, 2000; Alnaser, 1993), and cloudiness (Kumar and Umanand, 2005). The most commonly used parameter for estimating global solar radiation is sunshine duration.

II. METHODOLOGY & IMPLEMENTATION

A. Angstrom-Prescott Formula:

The Angstrom-Prescott formula is widely adopted in many researches, considering the relationship between monthly global solar radiation and sunshine-hours [3]. The main equation of the Angstrom-Prescott linear regression formula is given by

$$\frac{G}{G_0} = a + b \frac{S}{S_0}$$

Where,

- $G$ is the monthly average global solar radiation which can be directly observed,
- $G_0$ is the monthly average astronomical global solar radiation,
- $S$ represents the monthly average sunshine-hours, which can also be directly observed,
- $S_0$ is the most probably monthly average sunshine-hours, and
- $a$ and $b$ are two regression coefficients to be determined.

$G_0$ can be derived from the given equation once the location and time are given.

$$G_0 = \frac{24}{\pi} I_{SC} E_0 \left[ \frac{\pi}{180} \omega_s (\sin \delta \sin \phi) + (\cos \delta \cos \phi \sin \omega_0) \right]$$

Where,

- $I_{SC}$ - Solar constant with a value of 1367W/m2,
- $E_0$ - Earth’s orbital eccentricity correction factor can be derived from

$$E_0 = 1.00111 + 0.034221 \cos \xi + 0.00128 si n \xi + 0.000719 \cos 2\xi + 0.000077 si n 2\xi$$

$\xi$ is the solar angle and it can be derived through

$$\xi = \frac{2\pi (d_n - 1)}{365}$$

Where $d_n$ is the number of days in one year, ranging between 1 and 365.

$\omega_s$ - Solar angle at sunrise and can be obtained by

$$\omega_s = \cos^{-1}(-\tan \phi \tan \delta)$$

$\delta$ - Solar angle of inclination

$$\delta = (0.006918 - 0.399912 \cos \xi + 0.070257 \sin \xi - 0.006759 \cos 2\xi + 0.000907 \sin 2\xi - 0.002697 \cos 3\xi + 0.001148 \sin 3\xi \frac{180}{\pi})$$

$\phi$ - Latitude of the sample region.

$S_0$ is the most probably monthly average sunshine-hours, can be derived from
$S_0 = \frac{2}{15} \cos^{-1}(-\tan \phi \tan \delta)$

B. Determination of the Regression Parameters: A, B and C:
The monthly mean daily global solar radiation and sunshine hours on horizontal surfaces were obtained from the archives of National Aeronautics and Space Administration (NASA) for a year while the sunshine duration were obtained from the archives of the Department of Meteorology for the particular region. In this work, four types of regression models are used to calculate the regression coefficients [5]. They are
1) Linear Regression Equation (Model 1):
This model is obtained from the source of Angstrom, (1924) and Prescott (1940) equation takes the form as
\[ \frac{G}{G_0} = a + b \frac{S}{S_0} \]
\[ y = a + b \times \]
where a and b are the regression coefficients can be determined.
2) Exponential Regression Equation (Model 2):
This model is obtained from the source of Almorox et. al. (2005) equation takes the form as
\[ \frac{G}{G_0} = a + b \exp \left( \frac{S}{S_0} \right) \]
\[ y = a + b \exp \left( x \right) \]
3) Logarithmic Regression Equation (Model 3):
This model is obtained from the source of Ampratwum and Dorvlo (1999) and in the form as follows
\[ \frac{G}{G_0} = a + b \log \left( \frac{S}{S_0} \right) \]
\[ y = a + b \log \left( x \right) \]
4) Polynomial Regression Equation (Model 4):
This model is obtained from the source of Akinoglu and Ecevit (1990) and in the form as
\[ \frac{G}{G_0} = a + b \left( \frac{S}{S_0} \right) + c \left( \frac{S}{S_0} \right)^2 \]
\[ y = a + b \left( x \right) + c \left( x \right)^2 \]
a, b and c are the regression coefficients can be determined.

C. Artificial Bee Colony (ABC) Algorithm:
Artificial Bee Colony Algorithm (ABC) is nature-inspired metaheuristic, which imitates the foraging behaviour of bees. ABC as a stochastic technique is easy to implement, has fewer control parameters, and could easily be modify and hybridized with other metaheuristic algorithms. Due to its successful implementation, several researchers in the optimization and artificial intelligence domains have adopted it to be the main focus of their research work. The ABC algorithm is a swarm based, meta-heuristic algorithm based on the model first proposed by [6] on the foraging behaviour of honey bee colonies. The model is composed of three important elements: employed and unemployed foragers, and food sources. The employed and unemployed foragers are the first two elements, while the third element is the rich food sources close to their hive.

The main steps of the algorithm are as below:
Step 1: Initialize Population
Step 2: repeat
Step 3: Place the employed bees on their food sources
Step 4: Place the onlooker bees on the food sources depending on their nectar amounts
Step 5: Send the scouts to the search area for discovering new food sources
Step 6: Memorize the best food source found so far
Step 7: until requirements are met

In ABC algorithm, each cycle of the search consists of three steps: sending the employed bees onto their food sources and evaluating their nectar amounts; after sharing the nectar information of food sources, the selection of food source regions by the onlookers and evaluating the nectar amount of the food sources; determining the scout bees and then sending them randomly onto possible new food sources [5]. At the initialization stage, a set of food sources is randomly selected by the bees and their nectar amounts are determined. At the first step of the cycle, these bees come into the hive and share the nectar information of the sources with the bees waiting on the dance area. A bee waiting on the dance area for making decision to choose a food source is called onlooker and the bee going to the food source visited by herself just before is named as employed bee. After sharing their information with onlookers, every employed bee goes to the food source area visited by her at the previous cycle since that food source exists in her memory, and then chooses a new food source by means of visual information in the neighborhood of the one in her memory and evaluates its nectar amount. At the second step, an onlooker prefers a food source area depending on the nectar information distributed by the employed bees on the dance area. As the nectar amount of a food source increases, the probability of that food source chosen also increases. After arriving at the selected area, she chooses a new food source in the neighborhood of the one in the memory depending on visual information as in the case of employed bees. The determination of the new food source is carried out by the bees based on the comparison process of food source positions visually [7]. At the third step of the cycle, when the nectar of a food source is abandoned by the bees, a new food source is randomly determined by a scout bee and replaced with the abandoned one. In this model, at each cycle at most one scout goes outside for searching a new food source and the number of employed and an onlooker bee is selected to be equal to each other. These three steps are repeated through a predetermined number of cycles called Maximum Cycle Number (MCN) or until a termination criterion is satisfied.

D. Performance Statistics:
1) Mean Bias Error:
The Mean Bias Error gives an idea of the divergence between the monthly average daily radiation values estimated by the model used and the measured value [8]. A positive value shows over estimation and a negative value is under estimation.
\[ MBE = \frac{1}{n} \sum_{i=1}^{n} \left( G_{calc} - G_{meas} \right) \]
2) Mean Percentage Error:
The MPE can be defined as the percentage deviation of the monthly average daily radiation values estimated by the model used from the measured values. A percentage error between −10% and +10% is considered acceptable.

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

Where

- \( G_{\text{calc}} \) is calculated value
- \( G_{\text{meas}} \) is the measured value of the average daily global solar radiation
- \( n \) is the number of observations.

3) Root Mean Square Error:
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.

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

4) The Nash–Sutcliffe Equation:
A model is more efficient when NSE is closer to 1 (Chen et al., 2004).

\[ NSE = 1 - \frac{\sum_{i=1}^{n} (G_{\text{meas}} - G_{\text{calc}})^2}{\sum_{i=1}^{n} (G_{\text{meas}} - \bar{G}_{\text{meas}})^2} \]

where \( \bar{G}_{\text{meas}} \) is the mean measured global radiation.

5) T-Statistics Test:
As defined by Student (Bevington, 1969) in one of the tests for mean values, the random variable \( t \) with \( n-1 \) degrees of freedom (Daniel and Terrell) as follows:

\[ t = \left[ \frac{(n-1)(MSE)^2}{(RMSE)^2 - (MBE)^2} \right]^\frac{1}{2} \]

The smaller the value of \( t \) the better is the performance.

III. SIMULATION RESULTS AND DISCUSSION

A. Input Data Table:

<table>
<thead>
<tr>
<th>MONT HS</th>
<th>G measur ed</th>
<th>G0</th>
<th>S</th>
<th>S0</th>
<th>G/G 0</th>
<th>S/S0</th>
</tr>
</thead>
<tbody>
<tr>
<td>JAN</td>
<td>23.004</td>
<td>34.71</td>
<td>8.91</td>
<td>11.74</td>
<td>0.66</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>94</td>
<td>42</td>
<td>27</td>
<td>95</td>
</tr>
<tr>
<td>FEB</td>
<td>23.616</td>
<td>36.67</td>
<td>9.46</td>
<td>11.87</td>
<td>0.64</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>43</td>
<td>84</td>
<td>4</td>
<td>68</td>
</tr>
<tr>
<td>MAR</td>
<td>21.348</td>
<td>37.73</td>
<td>8.53</td>
<td>12.05</td>
<td>0.56</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21</td>
<td>23</td>
<td>12</td>
<td>58</td>
<td>81</td>
</tr>
<tr>
<td>APR</td>
<td>19.008</td>
<td>37.15</td>
<td>7.70</td>
<td>12.20</td>
<td>0.51</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>26</td>
<td>67</td>
<td>76</td>
<td>16</td>
<td>13</td>
</tr>
<tr>
<td>MAY</td>
<td>18</td>
<td>35.96</td>
<td>7.94</td>
<td>12.32</td>
<td>0.50</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>82</td>
<td>07</td>
<td>05</td>
<td>51</td>
</tr>
<tr>
<td>JUNE</td>
<td>16.452</td>
<td>35.60</td>
<td>6.7</td>
<td>12.34</td>
<td>0.46</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>68</td>
<td>17</td>
<td>17</td>
<td>2</td>
<td>29</td>
</tr>
<tr>
<td>JULY</td>
<td>15.192</td>
<td>36.38</td>
<td>5.88</td>
<td>12.26</td>
<td>0.41</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td>03</td>
<td>39</td>
<td>18</td>
<td>76</td>
<td>99</td>
</tr>
<tr>
<td>AUG</td>
<td>15.012</td>
<td>37.25</td>
<td>5.28</td>
<td>12.11</td>
<td>0.40</td>
<td>0.43</td>
</tr>
</tbody>
</table>

B. Output For Regression Model:
1) Linear Regression Model 1:

Fig. 1: Linear regression model

Fig. 2: Values of regression coefficients a and b

2) Exponential Regression Model:

Fig. 3: Exponential regression model
Energy is the primary requirement for development of any country. The world is heading towards energy crisis as fossil fuel reserves are depleting rapidly and demand for energy is increasing exponentially. Therefore, we need to exploit the widely available non-conventional energy resource “The Sun.” Designing a solar system is crucial for fulfilling the energy requirement of the people. In this direction the knowledge of the amount of irradiance reaching any point on the Earth’s surface is very essential. Since the regression coefficients vary from time to time and place to place hence it is necessary to determine these coefficients for the particular region, then the obtained values will be taken as reference value for the nearby region in order to calculate the solar radiation the nearby area or region.

REFERENCES


