Daily global solar radiation modelling using multi-layer perceptron neural networks in semi-arid region

Mawloud GUERMOUI*, Abdelaziz RABEHI, Said BENKACIALI, Djelloul DJAFER

Unité de Recherche Appliquée en Energies Renouvelables, URAER, Centre de Développement des Energies Renouvelables, CDER, 47133, Ghardaïa, Algeria
E-mail (*): gue.mouloud@gmail.com

* Corresponding author, phone: 00(213)029 87 01 26, fax: 00(213)029 87 01 46

Received: March 19, 2016 / Accepted: July 14, 2016 / Published: July 31, 2016

Abstract
Accurate estimation of Daily Global Solar Radiation (DGSR) has been a major goal for solar energy application. However, solar radiation measurements are not a simple task for several reasons. In the cases where data are not available, it is very common the use of computational models to estimate the missing data, which are based mainly of the search for relationships between weather variables, such as temperature, humidity, sunshine duration, etc. In this respect, the present study focuses on the development of artificial neural network (ANN) model for estimation of daily global solar radiation on horizontal surface in Ghardaïa city (South Algeria). In this analysis back-propagation algorithm is applied. Daily mean air temperature, relative humidity and sunshine duration was used as climatic inputs parameters, while the daily global solar radiation (DGSR) was the only output of the ANN. We have evaluated Multi-Layer Perceptron (MLP) models to estimate DGSR using three year of measurement (2005-2008). It was found that MLP-model based on sunshine duration and mean air temperature give accurate results in term of Mean Absolute Bias Error, Root Mean Square Error, Relative Square Error and Correlation Coefficient. The obtained values of these indicators are 0.67 MJ/m², 1.28 MJ/m², 6.12%and 98.18%, respectively which shows that MLP is highly qualified for DGSR estimation in semi-arid climates.
Keywords

Global Solar Radiation; Neural Network; Regression; Renewable Energy

Introduction

With increasing demands and interest in energy conservation and environmental protection, solar energy research has witnessed a growing interest and attracted a great deal of attention. Solar radiation is one of the most promising potential renewable energy sources which are believed to be cheap and clean energy in most regions.

The received solar radiation on earth surface depends upon climatic conditions of a location and geographical characteristic of the studied area. An optimal use of solar energy needs an accurate knowledge of solar radiation at a particular geographical location. Nevertheless, these data are not always dispensable particularly in isolated areas, in this respect several approaches have been developed in the literature for modeling and predicting solar radiation components such as: (i) empirical models [1-2], (ii) remote sensing retrievals [3-4] and (iii) soft computing using artificial neural network [5-6]. Many empirical regression models was proposed in the literature for estimating solar radiation, the most popular model was that proposed by Angström-Prescott [7] which establishes a simple relationship between global radiation and sunshine duration according to equation 1:

\[ \frac{R_s}{R_a} = \frac{S}{S_0} + b \] (1)

where \( R_s \) is actual solar radiation, \( R_a \) is extra-terrestrial solar radiation, \( S \) is actual sunshine duration, \( S_0 \) is potential sunshine duration, \( a \) and \( b \) are empirical parameters [8].

In recent years, several studies have been presented by researcher using many meteorological data to estimate the Daily Global Solar Radiation (DGSR). Due to the limited number of solar radiation measuring stations, the estimation of solar radiation are necessary. In this respect new models and improvement are continually proposed which intend to improve the solar radiation values based on available meteorological variables. As mentioned before, Artificial Neural Network (ANN) which used widely to estimate and predict solar radiation is one of the earliest machine learning methods. The main issues of ANN include its instability with short data and its tendency to over fit training data. Dorvio et al [9], used Radial Basis Functions (RBF) and Multi-Layer Perceptron (MLP) networks to estimate the
solar radiation from eight stations in Oman. It was pointed out that, in terms of RMSE between the observed and predicted solar radiation values, both the RBF and MLP models performed well, (0.83 MJ/m²/day) was obtained for the RBF and (1.01 MJ/m²/day) for the MLP. Banghanem et al. [5] used Radial Basis Functions (RBF) neural network to estimate DGSR at Al-Madina (Saudi Arabia). Four RBF models were developed using different meteorological data as input, it was found that the RBF model that uses sunshine duration and air temperature as input gives high performance (r = 98.8%). Senkal et al. [10] used MLP for estimating of the DGSR at 12 cities in Turkey using latitude, longitude, altitude, month, mean diffuse radiation and mean beam radiation as input whereas the output of the network is the solar radiation. According to the authors the MLP model performs better than the physical model, the obtained RMSE between the estimated values using the MLP and the physical model was 91 W/m² and 125 W/m² (testing cities), respectively.

Sozen et al. [11] proposed a neural network model for modeling solar radiation based on geographical coordinates (latitude, longitude and altitude), meteorological data (sunshine duration and mean temperature) and the corresponding month from seventeen stations in Turkey as inputs of the network. According to the others, the maximum mean absolute percentage error was found to be less than 6.7% and the correlation coefficient achieves 99.89% for the testing stations. Another interesting work proposed in [12] using hybrid model that combines Markov chain and neural network called ANN-MTM (Markov Transition Matrix). The geographic coordinates was used as input of the proposed model while the outputs are the DGSR. The obtained results between the measured and the predicted data generated by the proposed model in term of root mean square error (RMSE) is below 8% and the correlation coefficient between 90 and 92%.

Artificial neural network

Artificial neural network models employ artificial intelligence techniques and are data driven; they learn and memorize a data structure and subsequently simulate the structure. ANN can be trained to be used to approximate a non-linear function [12], or to classify outputs [13]. There are several algorithms available to train a neural network but this depends on the type of the topology of the neural networks. The aim of the present study is represented
by developing artificial neural network (ANN) model for estimation of daily global solar radiation on horizontal surface in Ghardaia city (South Algeria).

The most prominent topology of ANN is the feed-forward networks. An ANN is composed of nodes organized into layers and connected through weight elements. At each node, the weighted inputs are aggregated, threshold and inputted to an activation function to generate an output of that nod. Mathematically, this can be represented by equation 2:

\[ y(t) = f \left( \sum_{i=1}^{n} w_i x_i - w_0 \right) \]

where \( x_i \) are neuron inputs, \( w_i \) are the weights and \( f[.] \) is the activation function.

Multilayer Feed-forward is a kind of neural network, which consist of a number of layers: the first has neurons directly connected to the input data, and they are linked to one or more neurons in a hidden layer, or directly connected to the neurons in the output layer. In this kind of network, all neurons in one layer are full connected to all neurons of the next layer, and there are no feedbacks or recurrent connections.

In this work, was used a Multilayer Feed-forward Neural Network with different inputs parameters, as show in Figure 1. Hyperbolic tangent sigmoid transfer function is used in the hidden layer and linear transfer function for neurons in the output layer. The neural network was trained with the Levenberg-Marquardt Back-propagation algorithm [14], due to its high efficiency and fast convergence, although their computational requirements are high.

![Figure 1. Architecture of MLP neural network](image)

Materials and methodology

Site location and solar radiation data

The experimental data used in this work (Global Solar Radiation and sunshine duration, air temperature, relative humidity) , have been collected at the Applied Research Unit for Renewable Energies, (URAER) situated in the south of Algeria far from Ghardaïa.
The climate of Ghardaia region is semi-arid with a minimum and maximum air temperatures ranging from 14°C to 47°C and from 2°C to 37°C during summer and winter months respectively. The daily GSR varies between a minimum of 607Wh/m²/day to a maximum of 7574Wh/m²/day and the annual-mean-daily GSR is about 5656Wh/m²/day. The data are recorded every 5 min with a high precision by a radiometric station installed at the rooftop of the URAER (Unité de Recherche Appliquée en Energie Renouvelable) as shown on Figure 3.

The prediction model uses the data collected between 2005 and 2008 (for some days, however, there is no information due to electrical outages). The daily evolution of global solar
Daily global solar radiation modelling using multi-layer perceptron neural networks in semi-arid region
Mawloud GUERMOUI, Abdelaziz RABEHI, Said BENKACIALI, Djelloul DJAFER

radiation on horizontal surfaces and sunshine duration is illustrated in Figure 4.

**Figure 4. Daily evolution of the global horizontal radiation and sunshine duration**

**Model validation**

The performance of MLP models for estimating DGSR on horizontal surface is evaluated by comparing the estimated values with the measured values using different statistical indexes such as Mean Absolute Bias Error (MABE), Root Mean Square Error (RMSE), Relative Square Error (RRMSE), Determination Coefficient ($R^2$) and Correlation Coefficient($r$). The MABE, give the mean absolute value of bias error. Its expression is given by equation 3:

$$MABS = \frac{1}{n} \sum_{i=1}^{n} |H' - H|$$  \hspace{1cm} (3)

where $H'$ is the estimated value and $H$ is the measured value. The RMSE represents the difference between the predicted values and the measured values. In fact, RMSE identifies the model's accuracy. It is calculated by:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (H' - H)^2}$$  \hspace{1cm} (4)

The RRMSE is calculated by dividing the RRMSE to the average of measured data as:

$$RRMSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^{n} (H' - H)^2}}{\frac{1}{N} \sum_{i=1}^{N} H} \times 100$$  \hspace{1cm} (5)

The performance of the model is defined by the RRMSE range as follows:
• Excellent if: $\text{RRMSE} \leq 10$
• Good if: $10\% < \text{RMSE} < 20\%, \ 10 \leq \text{RRMSE} \leq 20$
• Fair if: $20\% < \text{RMSE} < 30\%, \ 20\% \leq \text{RRMSE} \leq 30\%$
• Poor if: $\text{RMSE} > 30\%, \ \text{RMSE} \geq 30\%$

The $R^2$ and $r$ indicate the strength of a linear relationship between the measured and predicted values. They are calculated as follows:

$$r = f \left( \frac{\sum_{i=1}^{n} (H' - \bar{H}) \times (H' - \bar{H})}{\sqrt{\sum_{i=1}^{n} (H_p - \bar{H'}) \times \sum_{i=1}^{n} (H - \bar{H})}} \right)$$

(6)

**MLP-based implementation of solar radiation models**

As mentioned before, the main objective of this letter is to assess the potential of MLP for modeling the global solar radiation from other parameters (e.g. Air temperature, Sunshine Ratio, Relative Humidity) in semi arid area (Ghardaia). In this respect several MLP-models are developed based on different input parameters and one output (which is the Clearness Index $K_T$).

Several combinations of the input parameters are used; we take Air Temperature, Relative Humidity and sunshine ration (SS) as inputs and clearness index, $K_T = H/H_0$ as output.

The collected data are divided into three subsets, first set for training (600 samples), the second (100 sample) for model validation and the rest for testing the performance of the studied MLP-models. The obtained statistical results are presented in Table1 for the testing data set. From Table 1, we can infer several observations. Firstly, we can point out that using different inputs influence the precision of the models. Secondly, MLP models that use air temperature achieve high performance compared with those based on relative humidity. Thirdly, combining air temperature and relative humidity does not improve significantly the performance of the models.

An important observation is that using sunshine ratio alone as input achieves high performance due to its high correlation with clearness index. Combining sunshine duration with air temperature and relative humidity boosts the performance and the model with higher accuracy are obtained using $\{T_{Mean}, SS\}$ as input parameters.
Table 1. The studied MLP-models with different input attributes

<table>
<thead>
<tr>
<th>Models</th>
<th>Input Parameters</th>
<th>MABE (MJ/m²)</th>
<th>RMSE (MJ/m²)</th>
<th>RRMSE</th>
<th>r(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLP-1</td>
<td>TMean, HMean</td>
<td>2.12</td>
<td>3.05</td>
<td>14.6</td>
<td>89.15</td>
</tr>
<tr>
<td>MLP-2</td>
<td>TMean, HMean, SS</td>
<td>0.76</td>
<td>1.38</td>
<td>6.59</td>
<td>97.98</td>
</tr>
<tr>
<td>MLP-3</td>
<td>TMean, SS</td>
<td>0.75</td>
<td>1.28</td>
<td>6.12</td>
<td>98.18</td>
</tr>
<tr>
<td>MLP-4</td>
<td>HMean, SS</td>
<td>0.84</td>
<td>1.33</td>
<td>6.36</td>
<td>98.06</td>
</tr>
<tr>
<td>MLP-5</td>
<td>SS</td>
<td>0.91</td>
<td>1.55</td>
<td>7.42</td>
<td>97.27</td>
</tr>
</tbody>
</table>

The graphical prediction output of the all five MLP- models against the measured data is plotted in the Figures (4-8) as follow:

Figure 4. MLP-1 model output prediction

Figure 5. MLP-2 model output prediction
Figure 6. MLP-3 model output prediction

Figure 7. MLP-4 model output prediction
Conclusion

The conducted examination showed the appreciable effect of input parameters on the precision of MLP-models. It has been demonstrated that \( H = \{SS, T_{\text{Mean}}\} \) provides better accurate precision than the other proposed MLP-models.

It should be noted that the fraction of sunshine duration SS play an important role for obtaining high precision accuracy which has been seen in the case when we have eliminated the fraction of sunshine duration from the input of the MLP-models.

Additionally, more cases studies with various climate conditions should also be considered in the future investigations.

Acknowledgements

We would like to thank and acknowledge the support of the team at the Research Unit on Renewable Energies of Ghardaïa who collected the data used in the present study.
Appendix

Clearness index $K_T$, which is the ratio of global solar radiation on horizontal surface (H) to the extraterrestrial solar radiation on horizontal surface ($H_0$). The extra-terrestrial solar radiation ($H_0$) is given as:

$$H_0 = \frac{24 \times 3600}{\pi} I_0 f \left( \cos \lambda \cos \delta \sin w_s + \frac{\pi}{180} w_s + \lambda \sin \delta \right)$$

$$f = 1 + 0.33 \left( \cos \frac{360 \lambda}{365} \right)$$

$$\lambda = 23.45 \sin \left( \frac{360 (284 + n)}{365} \right)$$

$$w_s = \arccos \left( -\frac{\tan \lambda \tan \delta}{1} \right)$$

where $I_0$ is the solar constant equal to 1360 W/m², $f$ is the eccentricity correction factor, $\lambda$ is the latitude of the region, $\delta$ is the solar declination and $w_s$ is the mean sunrise hour angle, $n$ represents the number of days.

References


5. Benghanem M., Mellit A., Radial basis function network-based prediction of global solar


