Sensitivity analysis based artificial neural network approach for global solar radiation prediction in India

R. Meenal¹, A. Immanuel Selvakumar², Prawin Angel Michael³, Ekambaram Rajasekaran⁴
¹,²,³Department of Electrical and Electronics Engineering, Karunya Institute of Technology and Sciences, India
²Department of Chemistry, VSB Engineering College, India

ABSTRACT

The objective of this paper is to build an artificial neural network model to predict global solar radiation (GSR) with improved accuracy using less number of best input parameters selected using sensitivity analysis. In this work, the input parameters used for training the artificial neural network (ANN) models are bright sunshine duration, maximum and minimum temperature, day length, months, extra terrestrial radiation (H₀), relative humidity and geographical parameters of the locations namely the latitude and longitude. Sensitivity analysis is used to discover how the output data are influenced by the changeability of the input data. Three ANN models namely T-ANN, S-ANN and TS-ANN are proposed with most suitable input parameters selected using sensitivity analysis. The principle of this feature selection using sensitivity analysis is to improve the prediction accuracy of solar radiation models with less number of inputs. The proposed ANN model is also tested under noisy data and proved that ANN is able to perform reasonably good in GSR prediction on practical applications where the data is affected by noise caused by errors on measuring, fault of data acquisition system, recording problems, and so on.

Keywords: ANN, Prediction, Sensitivity analysis, Solar radiation

1. INTRODUCTION

This work presents the application of artificial neural network (ANN) for the prediction of global solar radiation without any significant errors. Since the network of radiation measuring station is very limited due to costly measuring instrument, prediction of GSR is very important in country like India. Several solar radiation models namely clear sky, mathematical, fuzzy, artificial intelligence and various hybrid radiation models are available in the literature for the prediction of GSR [1-4]. GSR can be predicted using usually available meteorological parameters like temperature, sunshine duration, relative humidity, wind speed, atmospheric pressure and so on as input parameters. This GSR prediction can be done with the help of ANN approach without using radiation measuring instrument such as pyranometer. Artificial Neural networks have been successfully applied to various applications such as forecasting, optimization, machine diagnostics, process modelling and control, intelligent searching, voice recognition, fraud detection and so on [5, 6]. The optimum tuning parameters of AA2024-T351 alloy is determined using the ANN computing analysis [7]. S. Karupiah [8] presented an ANN approach for predicting overcurrent relay miscoordination time with the help of MATLAB software. Yadav and Chandel [9] offered a complete review of ANN techniques based solar radiation prediction. Various ANN models [10-15] have been developed for the prediction of daily and monthly global solar radiation. ANN [16-19] is the most commonly used technique for the prediction of GSR and recently machine learning techniques like support vector machine [20] is used for solar radiation prediction.
Identification of the best suitable input variables is the main research area in solar radiation prediction. The authors identified the selection of the most suitable meteorological and geographical input parameters to solar radiation models as a major research gap which requires to be addressed. This paper presents the selection of most suitable input parameters for the estimation of solar radiation using sensitivity analysis of ANN approach. The main goal is to reduce the number of input parameters and to improve the accuracy of the ANN models with less error using sensitivity analysis. Sensitivity analysis has been used in various applications like prediction, optimization and so on [21-25]. Mohammad Hasan Shojaeeefard applied sensitivity analysis approach in Friction Stir Lap Joining of Aluminum to Brass [26]. An ANN approach and sensitivity analysis is used in predicting skeletal muscle forces by MiloslavVilimek [27]. In this paper three ANN models are developed with best input parameters selected using sensitivity analysis.

This paper is organized as follows; Section 2 contains the data set used for training the solar radiation models. Section 3 presents a brief review of the ANN network used in this work. Section 4 presents the sensitivity analysis methodology used to conduct this study. Also, the experimental results are presented and discussed. Finally, the conclusion of this study is given in Section 5.

2. DATA SET

Global solar radiation is predicted using solar radiation models developed using artificial intelligence based approach. Selection of most suitable input parameter is the most important research area in solar radiation prediction Table 1 shows the input parameters utilized for training the radiation models. The input parameters are maximum temperature \(T_{\text{max}}\), minimum temperature \(T_{\text{min}}\), bright sunshine duration \((S)\), length of the day \((S_0)\), months, extra terrestrial radiation \((H_0)\), Relative humidity (RH) and latitude and longitude. Day length and \(H_0\) are calculated parameters [1]. The measured data for this study are monthly average maximum and minimum temperature, bright sunshine duration and daily GSR in MJ/m\(^2\)/day for various locations of India were collected from India Meteorological Department, Pune. Figure 1 shows the measured monthly average bright sunshine hours for Indian cities Nagpur, Bhubaneswar and New Delhi from IMD, Pune. Hourly global solar radiation data in MJ/m\(^2\)/day is collected for 15 Indian locations. Daily average and monthly average GSR data is calculated from hourly data set. Table 2 shows the measured monthly average GSR in MJ/m\(^2\)/day for the selected eight Indian cities namely Bhubaneswar, Chennai, Hyderabad, Mangalore, Nagpur Patna, New Delhi and Trivandrum. Monthly average data available for few Indian locations is utilized to train the solar radiation models to predict the GSR for other Indian locations where the GSR data is unavailable. Measured monthly average GSR, measured monthly average \(T_{\text{max}}\) and \(T_{\text{min}}\) for the selected three Indian locations are also shown in Figure 1.

The entire dataset is divided into two sections namely the training set (Installation sub-data set) and testing set (Validation sub-data set). The data set is processed further by utilizing solar radiation models with the help of artificial intelligence approach. MATLAB R2010a software is applied for developing the computer codes for solar radiation prediction models. Training data set is applied for training the solar radiation models; further the testing dataset is used for validating the solar radiation models.

<table>
<thead>
<tr>
<th>S. No</th>
<th>Input and Output parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Measured Input Parameter (T_{\text{max}})</td>
</tr>
<tr>
<td>2</td>
<td>Minimum Temperature (\text{°C}(T_{\text{max}}))</td>
</tr>
<tr>
<td>3</td>
<td>(S)-Bright sunshine hour</td>
</tr>
<tr>
<td>4</td>
<td>(RH)-Percentage Relative humidity</td>
</tr>
<tr>
<td>5</td>
<td>Calculated input parameters (H_0)</td>
</tr>
<tr>
<td>6</td>
<td>(S_0)-Day length (\text{MJ/m}^2/\text{day})</td>
</tr>
<tr>
<td>7</td>
<td>Geographical input parameters (\text{°})</td>
</tr>
<tr>
<td>8</td>
<td>Month</td>
</tr>
<tr>
<td>9</td>
<td>Latitude (\text{°}(N))</td>
</tr>
<tr>
<td>10</td>
<td>Longitude (\text{°}(E))</td>
</tr>
<tr>
<td></td>
<td>Output parameter</td>
</tr>
<tr>
<td>11</td>
<td>Global Solar Radiation (\text{MJ/m}^2/\text{day})</td>
</tr>
</tbody>
</table>
Table 2. Measured monthly average GSR in MJ/m²/day for Indian cities collected from IMD, Pune

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td></td>
<td>11.82</td>
<td>11.40</td>
<td>15.36</td>
<td>18.65</td>
<td>17.02</td>
<td>14.80</td>
<td>17.62</td>
<td>17.73</td>
</tr>
<tr>
<td>February</td>
<td></td>
<td>15.18</td>
<td>14.64</td>
<td>18.62</td>
<td>20.70</td>
<td>20.89</td>
<td>16.74</td>
<td>20.20</td>
<td>19.63</td>
</tr>
<tr>
<td>March</td>
<td></td>
<td>18.60</td>
<td>18.18</td>
<td>20.97</td>
<td>22.87</td>
<td>22.46</td>
<td>20.10</td>
<td>19.95</td>
<td>19.30</td>
</tr>
<tr>
<td>April</td>
<td></td>
<td>20.46</td>
<td>21.25</td>
<td>23.37</td>
<td>23.90</td>
<td>22.55</td>
<td>23.18</td>
<td>21.35</td>
<td>18.39</td>
</tr>
<tr>
<td>May</td>
<td></td>
<td>20.58</td>
<td>20.91</td>
<td>23.60</td>
<td>23.97</td>
<td>21.71</td>
<td>22.44</td>
<td>18.83</td>
<td>17.36</td>
</tr>
<tr>
<td>June</td>
<td></td>
<td>17.27</td>
<td>20.01</td>
<td>18.67</td>
<td>18.91</td>
<td>20.00</td>
<td>16.54</td>
<td>17.50</td>
<td>16.62</td>
</tr>
<tr>
<td>July</td>
<td></td>
<td>14.24</td>
<td>17.07</td>
<td>13.81</td>
<td>17.11</td>
<td>18.51</td>
<td>15.31</td>
<td>13.12</td>
<td>16.20</td>
</tr>
<tr>
<td>August</td>
<td></td>
<td>16.17</td>
<td>15.90</td>
<td>14.37</td>
<td>16.66</td>
<td>18.85</td>
<td>15.17</td>
<td>14.41</td>
<td>17.52</td>
</tr>
<tr>
<td>September</td>
<td></td>
<td>14.70</td>
<td>15.51</td>
<td>17.55</td>
<td>17.64</td>
<td>18.61</td>
<td>14.50</td>
<td>17.39</td>
<td>17.84</td>
</tr>
<tr>
<td>October</td>
<td></td>
<td>14.68</td>
<td>15.32</td>
<td>18.23</td>
<td>17.89</td>
<td>16.51</td>
<td>17.51</td>
<td>18.18</td>
<td>16.21</td>
</tr>
<tr>
<td>November</td>
<td></td>
<td>12.86</td>
<td>12.44</td>
<td>15.95</td>
<td>19.34</td>
<td>14.34</td>
<td>15.69</td>
<td>18.09</td>
<td>14.52</td>
</tr>
<tr>
<td>December</td>
<td></td>
<td>11.64</td>
<td>10.60</td>
<td>14.86</td>
<td>18.00</td>
<td>13.32</td>
<td>13.41</td>
<td>17.65</td>
<td>15.73</td>
</tr>
</tbody>
</table>

Figure 1. Measured input parameters for Indian cities provided by IMD, Pune,
(a) measured bright sunshine hours, (b) measured GSR in MJ/m²/day,
(c) measured maximum temperature, (d) measured minimum temperature

3. METHODOLOGY

Three artificial neural network (ANN) models are developed to estimate the monthly average daily GSR values in different locations of India. Sensitivity analysis approach is utilized to build the ANN models to select the best input parameters. The aim of the sensitivity analysis is to decrease the number of inputs required for the prediction of GSR values without raising the network error. ANN models provide a better solution without requiring the knowledge of mathematical calculations among the parameters. The design, training parameters, architecture, and sensitivity and performance analysis of the proposed ANN models is discussed here in detail. Through the proposed methodology, the prediction accuracy is improved with less number of inputs. The proposed ANN models for GSR prediction is developed using the Lavenberg-Marquardt back propagation algorithm (LM algorithm).
3.1. Global solar radiation prediction using artificial neural network

ANN techniques have been applied for a wide range of applications such as prediction, image processing, pattern recognition, optimization, forecasting and simulation and so on. Various researchers have predicted GSR by using ANN models with appropriate meteorological data as input parameters [11]. Artificial neural networks replicate the biological neural networks. A simple model of an artificial neuron is shown in Figure 2. The neuron is the basic element of an ANN. The inputs to the ANN are \(X_1, X_2, ..., X_n\) and the weights linked to the input are \(w_1, w_2, ..., w_n\). The sum function is used to process the captured input information. The new signal developed by the neuron is set by the activation function. Therefore, the neurons are based on uncomplicated mathematical function which is given by the following equation.

\[
y = f\left[\sum_{i=1}^{n} I_i w_i\right]
\]

where \(I_i\) the \(i^{th}\) input of ANN, \(w_i\) represents the weight and \(f\) represents the activation function.

The multiple input single output ANN is given by the following equation:

\[
y = f_1\left[\sum_{i=1}^{n_{th}} w_{1i} \sum_{k=1}^{n_{th}} I_k w_k\right]
\]

In this study, tasksig activation function is applied for hidden layer and linear function in output layer. In this study, three artificial neural network (ANN) models are developed using the sensitivity analysis technique. These three ANN models are utilized to estimate the monthly average daily solar radiation values in India.

![Figure 2. Model of an artificial neuron](image-url)

4. SENSITIVITY ANALYSIS

A sensitivity analysis is a technique which is used to study the behavior of a model. This technique is used to assess the importance of each input variables on the values of the output parameter of the radiation model. In a more general sense, the sensitivity analysis determines how different values of the input parameters affect the the output variable. With this technique, it is feasible to determine which input variables should be considered as the most important and least important ones for the prediction of solar radiation. In this study, a correlation analysis is carried out to determine the most influencing input variables. Table 3 shows the ranking of input parameters based on the value of correlation coefficient. The prediction results will be better if the value of \(R\) is high which is nearer to one. It is observed that relative humidity scored the least \(R\) value of 0.44; maximum temperature got the highest value of 0.8 followed by bright sunshine duration. The calculated correlation coefficient values ranges between 0.44 and 0.80.
Table 3. Ranking of input parameters based on the value of correlation coefficient (R)

<table>
<thead>
<tr>
<th>S. No</th>
<th>Input parameters</th>
<th>Correlation Coefficient (R)</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Relative humidity (RH)</td>
<td>0.44</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>Minimum temperature (T_{min})</td>
<td>0.54</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>Length of the day</td>
<td>0.60</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>H_0</td>
<td>0.63</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Bright sunshine duration</td>
<td>0.68</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>Maximum temperature (T_{max})</td>
<td>0.80</td>
<td>1</td>
</tr>
</tbody>
</table>

4.1. ANN models

Three artificial neural network (ANN) models namely S-ANN, T-ANN and ST-ANNN are developed using sensitivity analysis technique. The S-ANN model has bright sunshine hour (S) and length of the day (S_0) as its input. As the meteorological parameter bright sunshine hour is not available in all the measuring stations of India, the temperature-based ANN model namely T-ANN is build with the help of maximum and minimum temperature as its input variables. The last model ST-ANN is developed using maximum and minimum temperature, bright sunshine hour values, day length are used as inputs parameters. For all the above mentioned three ANN models, latitude, longitude and month of the selected Indian cities are used as common input parameters and the output is the monthly average daily GSR data. Figure 3 shows the architecture of the ST-ANN model. Lavenberg-Marquardt back propagation algorithm (LM algorithm) is applied to estimate the solar radiation in selected Indian locations.

The main aim of the sensitivity analysis technique is to decrease the number of inputs variables required for estimation of solar radiation without raising the error of the network. To accomplish this, ANN model is developed by different combination of inputs and then trained respectively. Next, the accomplished performance is reported. Table 4 presents the results of training error namely RMSE for single input parameter and combination of two inputs. The developed ANN model with least RMSE has the the majority influence on the global solar radiation. From Table 4, it is found that model with the combination of bright sunshine duration and day length inputs performs excellent in the estimation of solar radiation with least error followed by the model using maximum and minimum temperature as inputs. Thus, for radiation prediction the first two models with less RMSE value is selected. Best performing models are available in Table 5. From Table 5, it is verified that the inclusion of the geographical parameters results in minimizing the regression error of the ANN models. In addition, the model with the combination of temperature and sunshine inputs gives excellent prediction results.

![Figure 3. Architecture of the ST-ANN model](image-url)
4.2. Sensitivity analysis under noise

In the following analysis, a simulated small random error has been introduced in input variables to see the influence on desired output. The sensitivity analysis is performed on temperature based ANN model (T-ANN). In T-ANN model, 5 inputs are used namely, month, latitude longitude, maximum and minimum temperature. In this analysis, the geographical parameters are kept fixed. A small random error has been introduced in input variables namely maximum and minimum temperature to see the influence on the output radiation. The following three cases is considered to demonstrate the stability of the ANN:

Case a) 5% of error added to single input and two inputs
Case b) 10% of error added to single input and two inputs
Case c) 20% of error added to single input and two inputs

In all the three cases shown in Table 6, R value is above 0.94 in training results. If small random error is introduced in any one of the inputs, there is no much influence on the output. The prediction results are good. ANN is able to perform well up to 10% of error introduced in the input parameters. From Table 6, it is observed that under noise condition also, all the statistical indicators namely MBE and RMSE are within the acceptable range.

As the noise level is increasing to 20% there is a small degradation in ANN performance which is shown in Table 7. It is found that as the error introduced in the input increases, the RMSE value is also increasing. From Table 8, it is found that the selected ANN model with the best input variables identified using sensitivity analysis performs better than the other ANN models with less RMSE value and R value closer to one. As conclusion, ANN can perform reasonably good in GSR prediction on practical applications.
5. CONCLUSION

Three categories of ANN models are developed to predict the global solar radiation. The best models namely T-ANN, S-ANN and ST-ANN are proposed based on sensitivity analysis. The stability of the model is checked by conducting sensitivity analysis under noise condition by introducing some error in the inputs given to the ANN models. From the results reported in this work, it is concluded that correlation coefficient and prediction accuracy of the ANN models build in this work with most suitable input variables are better than the other ANN models available in the literature (R=0.99). The effectiveness of the ANN models developed in this work with best input variables has been verified by conducting comparative study with other ANN models. It has been observed that the selected ANN models have performed superior over the other monthly ANN models. Therefore it is obviously verified that the input variables namely the maximum temperature and bright sunshine duration identified using sensitivity analysis plays an important role in obtaining the superior prediction accuracy. In addition, ANN is able to perform reasonably good in GSR prediction on practical applications where the data is affected by noise caused by errors on measuring, fault of data acquisition system, recording problems, and so on.

ACKNOWLEDGEMENTS

India Meteorological Department is acknowledged for the data support.

REFERENCES