Assessment and Mapping of Solar Energy Potential Using Artificial Neural Network and GIS Technology in the Southern Part of India

Khalid Anwar*‡, Sandip Deshmukh*

*Department of Mechanical engineering, Hyderabad Campus, BITS Pilani, Telangana-500078, India

(khalid@hyderabad.bits-pilani.ac.in, ssd@hyderabad.bits-pilani.ac.in)

‡ Khalid Anwar; BITS Pilani, Hyderabad, Telangana-500078, India,

khalid@hyderabad.bits-pilani.ac.in

Received: 08.03.2018 Accepted: 12.04.2018

Abstract- Prediction and assessment of solar radiation are necessary pre-requisites in developing solar technology. Here, an artificial neural network (ANN) model has been developed to predict potential of solar energy in the Southern part of India: Andhra Pradesh (AP) and Telangana State (TS), lie between 12°41’ and 22°N latitude and 77° and 84°40’E longitude. Generalized feed-forward with back-propagation neural networks were considered using MATLAB. Three layered neural network with different architectures are designed and evaluated. For training and testing the network, geographical and meteorological data of 28 sites over a period of recent 22 years from the NASA geo-satellite database were taken. Geographical parameters (latitude, longitude and altitude), meteorological data (temperature, sunshine duration, relative humidity and precipitation) along with month were used as input data, whereas the mean solar radiation was used as the output of the network. All the parameters taken here are in the form of monthly mean. The ANN model has been evaluated for test locations by calculating mean absolute percentage error (MAPE). The correlation coefficients (R-value) between the output of model and the measured value of solar radiation is calculated. The R-value was more than 0.95, which show high reliability of the model for prediction of solar radiation anywhere within AP and TS. Solar radiation of major cities was predicted using developed model. Predicted solar radiation is analyzed and used to create monthly mean maps using GIS technology. These maps can be useful to estimate solar energy potential at any locations within AP and TS.

Keywords- Solar radiation; Artificial neural network; Renewable energy; GIS technology; Modelling.

1. Introduction

The demand for energy has been tremendously increased since the middle of last century, due to industrial development, increase in global economy and population growth. To meet the energy demand, fossil fuels are being extensively used which lead to continuous increase in greenhouse gas emission. In this way, clean energy is becoming the center of attraction for researchers worldwide to work towards a sustainable future, and reducing the effect of climate change and global warming [1, 2]. Solar and wind are considered renewable alternatives as both are widely available and economically viable for power production at large scale [3–6]. Nowadays, both wind and solar energy is being evaluated and analyzed for efficient utilization across the globe [7–10]. Wind energy has been estimated in in Chad [11] and Techno-Economic Analysis of a Hybrid PV-Wind Turbine System is carried out in Jordan [12]. The potential Solar energy is evaluated in Odisha, India [5]. Among both wind and solar, sun provides heat and light, which can be transformed to any useful forms of energy [13]. In developing countries, exploring and maximum utilization of the solar energy can play vital role to address energy crises [1]. As in India, due to urbanization at an annual rate of 2.67%, demand for electricity increased by 35.19% in the last 5 years and with the generation capacity of 225GW, India faced a peak and energy shortage of 9% and 8.7% in 2012-2013 respectively [14]. The penetration of renewable energy has been also inspired by reasons, for instance, environmental concerns, incentive polices and increased demand of energy across the globe. A key decision for energy production companies is finding the best location to develop renewable energy technology at the same time to choose which one is best energy source (alone or combination of sources) for productivity.
Technology for solar energy has an extraordinarily rapid growth in the last two decades, and now it is a developed, reliable and commercially available technology for electricity production.

Fortunately, tropical countries like India are blessed with immense solar energy potential [2]. Major parts of India get about three hundred sunshine days in a year with about eight hours’ sunshine per day and solar radiation in a range of 4 to 7 kWh/m²/day. Currently, India is one of the precursors in the renewable energy markets globally, after JNNSM (Jawaharlal Nehru National Solar Mission) brought it under a policy framework in 2009 within the purview of National Action Plan on Climate Change (NAPCC). The solar power generation in India is expanding fast; it was 3MW in 2008 and reached over 8262MW in 2016. In 2015 government of India increased its solar plans, aiming at huge investment with a target of 20GW by 2020. To realize the growth envisaged by a market study conducted by the World Bank [15], and to comprehend the challenges in the development of solar energy, most of the inventor claims that the time given to complete the target in the JNNSM is impractical. Insufficient data is pointed out as one of the key barriers [15]. In these situations, it is vital that quality information about solar energy potential on local level is obtainable so that optimal design for solar energy technology can be developed. Presently in India, there are only limited location with measuring instruments which have solar energy potential data. Solar radiations modelling and mapping at a specific site are prerequisites in the design and planning process of solar energy system [16, 17]. Hence, measuring or prediction of solar potential is crucial parameter to the development and assessment of solar energy technology. Indian Meteorological Department (IMD), Pune, is the national agent responsible for recording and filling daily meteorological data. Nevertheless, while IMD records other parameters for various sites in India, there are very limited set up pyrometers for recording solar radiation intensities. Therefore, insufficient and imprecise record of solar radiations and other parameters is still parts of the key challenges in developing technology for renewable energy applications. The shortage in meteorological data has led to limited study and finally the utilization of solar potential in India. As India is a big country with diverse climate, a state-wise investigation of solar potential based on geographical condition is needed. This work is intended to overcome some of these limitations by developing a model to estimate solar potential at sites by taking geographical and meteorological information as input parameters.

Artificial neural network (ANN) model is a estimation tool, widely considered as an alternative method which can tackle complex and undefined problems [18, 19]. They are capable of learning from samples and handling random and inadequate data. Non-linear problems can be addressed using ANN. After training, ANN can predict and generalize with high accuracy. ANN has diverse applications in various engineering and science fields, such as, automation, pattern recognition, predicting, computer science, industrial, optimization, and medical sciences. ANN has been considered in the modelling of complex systems and applied for forecast of solar potential around the world. In Turkey, ANN model has been used to forecast solar radiation [16]. Daily solar radiation has been predicted and efficacy of photovoltaic system was improved using ANN [20-22]. Pure ANN and ANN with Computational Fluid Dynamics has been studied to predict wind power in hilly area [23]. An adaptive wavelet network model for predicting daily total solar potential in Algeria was reported by Mellit et al. [24]. Kalogirou presented a detailed overview on application of ANN [18]. Jani et al. applied ANN to study the Performance of solid desiccant - vapor compression and rotary solid desiccant dehumidifier in air conditioning system for hot and humid climatic zone of North India [25, 26]. A comparison between ANN and other model like simple regression has been presented by Mellit and Kalogirou [27]. It is concluded that ANN models are fast and efficient in solving complicated problem. Geographic Information System (GIS) technology was commonly used by researchers to present predicted energy potential from ANN model in the form of maps. That map gives good visual for available solar potential and helps to site selection and designing solar technology. GIS technology was used to plot a solar radiation map in Oman [28]. GIS technology is also useful for mapping all forms of energy potentials. Wind energy potential map has been presented by Ramachandra and Shruthi using GIS in India [29]. In above mentioned studies, parameters used to develop the ANN model are few. Here, eight input parameters, which are different from the other studies, are considered to develop the ANN model. Furthermore, ANN model is optimized, which gives significantly high accuracy.

In this work, ANN model has been developed using feed-forward, back-propagation, multilayer perceptron neural networks. Model is then used to predict the mean solar potential. To develop the model, data of 28 locations is taken within Andhra Pradesh (AP) and Telangana state (TS), India, during the last 22 years. The model has eight input parameters and one output parameter. Primarily, reliability of ANN model has been studied for non-linear relationship between solar radiation and other input parameters, and validated with test location. Thereafter, mean solar radiation was predicted using the developed model for major cities in AP and TS with unavailability of solar radiation data. Predicted solar radiation was presented as solar potential mapping using GIS technology.

2. Data Collection

To carry out this work, 28 locations were identified in AP and TS, India with an interval of 1 degree in latitude and 1 degree in longitude. Data were obtained for all specified locations during the period of 22 years (1983 - 2005). The source of data is NASA geo-satellite database. The data used here are monthly mean meteorological parameters for a given geographical location. Temperature, sunshine duration, relative humidity, precipitation and
solar radiation were taken as meteorological parameters. Three sets of data were made out of 28 locations consisting of data of 20, 4 and 4 locations respectively. The first set with 20 locations, second set with 4 locations and third set with 4 locations are applied to train the model, validate the efficiency of model and test the model for accuracy respectively. The test location is selected randomly at different location. Every location has meteorological data for 12 months. Each input data set for the model has combination of 3 geographical coordinates (latitude, longitude and altitude), and 4 meteorological data (mean temperature, mean sunshine duration, mean relative humidity, and mean precipitation) along with month of the year (1 to 12). In total there are 8 data in each input data set which corresponds to one output as solar radiation.

**Table 1.** Geographical location of data stations.

<table>
<thead>
<tr>
<th></th>
<th></th>
<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>1</td>
<td>19.5</td>
<td>78.5</td>
<td>8</td>
<td>18.5</td>
<td>83.5</td>
<td>15</td>
<td>17.5</td>
<td>82.5</td>
<td>22</td>
<td>15.5</td>
<td>78.5</td>
</tr>
<tr>
<td>2</td>
<td>19.5</td>
<td>79.5</td>
<td>9</td>
<td>18.5</td>
<td>84.5</td>
<td>16</td>
<td>16.5</td>
<td>77.5</td>
<td>23</td>
<td>15.5</td>
<td>79.5</td>
</tr>
<tr>
<td>3</td>
<td>18.5</td>
<td>77.5</td>
<td>10</td>
<td>17.5</td>
<td>77.5</td>
<td>17</td>
<td>16.5</td>
<td>78.5</td>
<td>24</td>
<td>14.5</td>
<td>77.5</td>
</tr>
<tr>
<td>4</td>
<td>18.5</td>
<td>78.5</td>
<td>11</td>
<td>17.5</td>
<td>78.5</td>
<td>18</td>
<td>16.5</td>
<td>79.5</td>
<td>25</td>
<td>14.5</td>
<td>78.5</td>
</tr>
<tr>
<td>5</td>
<td>18.5</td>
<td>79.5</td>
<td>12</td>
<td>17.5</td>
<td>79.5</td>
<td>19</td>
<td>16.5</td>
<td>80.5</td>
<td>26</td>
<td>14.5</td>
<td>79.5</td>
</tr>
<tr>
<td>6</td>
<td>18.5</td>
<td>80.5</td>
<td>13</td>
<td>17.5</td>
<td>80.5</td>
<td>20</td>
<td>16.5</td>
<td>81.5</td>
<td>27</td>
<td>13.5</td>
<td>78.5</td>
</tr>
<tr>
<td>7</td>
<td>18.5</td>
<td>82.5</td>
<td>14</td>
<td>17.5</td>
<td>81.5</td>
<td>21</td>
<td>15.5</td>
<td>77.5</td>
<td>28</td>
<td>13.5</td>
<td>79.5</td>
</tr>
</tbody>
</table>

![Fig. 1. Location of sites on AP and TS Map used in ANN model](image)

**Table 2.** Mean solar radiation (in KW h/m²) of 28 locations

<table>
<thead>
<tr>
<th>No.</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>1</td>
<td>4.88</td>
<td>5.73</td>
<td>6.34</td>
<td>6.68</td>
<td>6.59</td>
<td>4.83</td>
<td>4.04</td>
<td>3.84</td>
<td>4.67</td>
<td>5.07</td>
<td>4.93</td>
<td>4.71</td>
</tr>
</tbody>
</table>
3. ANN Approaches

ANN is a computational model, branch of artificial intelligence. ANN is based on the structure and functions of biological neural networks. This have the ability to train themselves, store for future calculation and recall the data for prediction based on input dataset. One of the unique feature of ANN models is, it establishes a typical kind of relation among distinct input parameters. Moreover, it provides weighting to each parameters based on that relation. The main components of a standard neural network are input units, junctions or nodes, output units. All three components are arranged in a series of layers, starting from input to hidden to output layers. The connections between components are established by some number, which is known as weight. This weight could be either positive or negative. The component got higher weight number is considered as having high influence in the relation.

In present study, the multi-layer, feed-forward with back propagation neural networks (FBPNN) model is used to predict solar radiations through geographical parameters and mean monthly meteorological data as inputs. This ANN model is considered as the most efficient and popular among the recent researchers for undefined complex problem. The FBPNN model evidenced to be the best for diverse data condition, giving considerably highest correlation coefficient values. In figure 3 general schematic of ANN model is given. In present study there are eight input parameters to model. Output of the model is solar radiation. Five steps have been followed to develop the ANN model, as shown in flow chart Figure 4.

4. Design of ANN Model

ANN model based on FBPNN was developed in MATLAB. The model network is composed of 3 layers which is; i) input ii) hidden and iii) output layer. As shown in figure 3, the basic architecture of network has 8 neurons in input layer and one neuron in output layer with varying number of neurons in hidden layer. Each neuron in input layer takes one input parameters. One neuron in output layer is to give estimated value of solar radiation based on relation established in hidden layer among input parameters.
Technically, in any ANN model, information flows in both the directions. First information goes to model as input to train the model. Once training is done, the model will give us information as output with accepted errors range. Critical part in ANN model development is training with input data, because of knowledge received that includes of updating or adjusting the baises in networks through various algorithms. These weights are updated until the ANN model restores the target output. There is a feedback mechanism in the model called backpropagation. Backpropagation matches the output generated by model with actual or measured value. Initially model was designed with multiple combinations of parameters, and different structures of network with single hidden layers. Double hidden layers were also tested where the number of neurons in each hidden layer varied from 1 to 15. The structure of ANN model with one hidden layer gives high accuracy. Therefore, network with one hidden layer is considered for study and results are presented for the same.

Fig. 2. Sunshine durations for 10 stations for representation.

Fig. 3. Architecture of ANN model in current study.
5. Results and Discussion

5.1. ANN Model

Training and testing of data yield predicted values and then these values have to be compared with the actual output value. This comparison is done using mean average percentage error (MAPE), given by the following formula:

$$ MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{H_{ai} - H_{ei}}{H_{ai}} \right| $$

Here, $H_{ai}$ represents the actual solar radiation, $H_{ei}$ represents the estimated solar radiation and $n$ stands for the number of testing examples.

![Fig. 5. Variation of MAPE with number of neurons in hidden layer.](image)

MAPE predicts the accuracy of a given forecasting method and expresses accuracy as percentage and neglects the signs of errors. MAPE values were calculated in all different architecture of model. Model has been evaluated for number of neurons in hidden layer, start with one and then increased to 15. It was observed that the MAPE first decreases with increase in number of neurons in hidden layer and then increases after a certain number. Optimum number of neurons is found for minimum MAPE and then considered for further calculations. The MAPE was calculated for each month of each location. The MAPE for training location will be minimum for obvious reason, which is not reported here. The average MAPE over 12 months for four test location is studied and presented. Fig. 5 shows the variation of MAPE with number of neurons in hidden layer and it is found that 9 neurons in one hidden layer are giving better results with minimum MAPE as 2.16, which is considered as optimum model for prediction of solar radiation. Optimum model is also simulated with varying number of iteration. The minimum MAPE has come with 25 iterations for the training of network. It can be noted from figure 5 that prediction of solar radiation from the optimum model is close and similar in their trend to the measured data. MAPE calculation shows that the error in predicted value is less than 2.5%. R-value has calculated to check the overall efficiency of the optimum model. It can be seen from figure 6, the optimum model gives R-value 0.999 for training, 0.989 for validation, 0.987 for testing and 0.996 for all data, the corresponding $R^2$ values (coefficient of determination) are 0.998, 0.979, 0.975 and 0.992 respectively. With all these statistics it can be clearly stated that the model estimation is acceptable with permissible error.
Fig. 6. Performance of developed ANN model. R-value for training, validation, testing, and whole datasets for network with optimal configuration.

For better understanding and detail visualization of efficiency of developed model, predicted solar radiation was compared and plotted with actual value for all 12 months. This comparison has been done for the test locations. Figure 7 presents the comparison plot for 4 test location (station No. 25 to 28 from table 1). The result indicates great fitting of value and prediction for each month is very near to actual value for all test location. This assured that the model is reliable and can be used to estimate solar radiation any location within AP and TS where the solar radiation data is not available. The average MAPE for four test location is given in table 3 shows accuracy of predicted value is significantly high. The developed model is capable to predict solar radiation with acceptable accuracy within AP and TS. This model can predict solar radiation taking input parameters at any location within the boundary where instrument is not installed for measurement of solar radiation. Solar radiation can be predicted in the form of monthly mean value for each month individually at any locations. The optimum model was then used to predict the solar radiation for 15 different major towns in AP and TS.

Table 3. MAPE of test locations.

<table>
<thead>
<tr>
<th>Test location</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Station No.</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
</tr>
<tr>
<td>MAPE</td>
<td>2.02</td>
<td>2.19</td>
<td>2.41</td>
<td>2.00</td>
</tr>
</tbody>
</table>
Fig. 7. Testing the model efficiency by comparing predicted with measured values for four test locations, (a) station No. 25, (b) station No. 26, (c) station No. 27, and (d) station No. 28.

Table 4. Statistical presentation of monthly mean solar radiation in AP and TS.

<table>
<thead>
<tr>
<th>Month</th>
<th>Solar radiation potential kWh m$^2$ day</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Max.</td>
</tr>
<tr>
<td>1</td>
<td>5.26</td>
</tr>
<tr>
<td>2</td>
<td>6.11</td>
</tr>
<tr>
<td>3</td>
<td>6.55</td>
</tr>
<tr>
<td>4</td>
<td>6.68</td>
</tr>
<tr>
<td>5</td>
<td>6.59</td>
</tr>
<tr>
<td>6</td>
<td>5.18</td>
</tr>
<tr>
<td>7</td>
<td>4.67</td>
</tr>
<tr>
<td>8</td>
<td>4.75</td>
</tr>
<tr>
<td>9</td>
<td>5.00</td>
</tr>
<tr>
<td>10</td>
<td>5.09</td>
</tr>
<tr>
<td>11</td>
<td>5.01</td>
</tr>
<tr>
<td>12</td>
<td>4.84</td>
</tr>
</tbody>
</table>

5.2. Mapping of solar potential

The output of the model, solar radiation in monthly mean format, for January–December was analyzed using GIS technology. ArcGIS 10.0 is used in present work for mapping of solar radiation. ArcGIS 10.0 is a geographic information system which can be analyzed and develop map. To develop map there are various technique available such as, one is presented by Famoso et. al. to study Ozone Concentration in Catania, Italy [30]. We used Kriging method to generate surface plot as it is well established and accepted technique. In this method 3D surface plot is generated using scattered Z-value with X-Y co-ordinate. Here, solar radiation is taken as Z-value and geographical co-ordinates (lat. And long.) is considered as X-Y co-ordinate. The solar radiation at all available location with their geographical co-ordinates is imported to ArcGIS and converted into shape file. Then the kriging method is applied to interpolate and to create the surface plot. Kriging is based on geo-statistical analysis which involves auto-correlation. It considered both distance and direction and fits a relation to each surrounding point. The Surface plot is then analyze using spatial analyst to form a contour plot of solar radiation. The steps used to create map in this study are summarized in Figure 8. This contour map is generated for each month, which is given in figure 9-20. The interval of contour is kept as 0.1 kW h/m$^2$ day, the contours are iso-lines and solar radiation value along the line is constant. The maps can provide better visualization and understanding of solar potential within the concern region in all month. It also helps to find the hot spot for solar potential and make better decision. It is found from the maps (Figs. 8-19) that solar potential varied widely across region and month to month. It is high in west region and getting low as we go to the east. As it can be seen in the figures, there is sufficient solar energy available in AP and TS, especially in March, April and May. Table 4 is the statistical presentation of solar potential in AP and TS. It can be
noted that the solar potential in AP and TS varies from 6.68 to 3.31 kW h/m²/day. April has the maximum solar potential 6.68 kW h/m²/day and minimum solar potential is available in August as 3.13 kW h/m²/day (Table 4). The solar radiation value is stable in April as it has least standard deviation as 0.09 and unstable in August with high standard deviation as 0.37.

1. Data adding to ArcGIS: latitude, longitude and solar radiation as X and Y and Z value respectively. Defining the geographical coordinate system.
2. Converting the data to Shape file.
3. Creating raster from shape file from solar radiation (Z value) of each month using interpolation in spatial analyst tool.
4. Developing contour of solar radiation for each month by analyzing the raster file.
5. Labeling the map. Adding the boundary for AP and TS, coordinates, direction and scale bar.

Fig. 8. Steps evolve in generation of contour map of solar radiation using GIS technology.

Fig. 9. Contour map of solar radiation (kW h/m² day) for AP and TS (January).

Fig. 10. Contour map of solar radiation (kW h/m² day) for AP and TS (February).

Fig. 11. Contour map of solar radiation (kW h/m² day) for AP and TS (March).

Fig. 12. Contour map of solar radiation (kW h/m² day) for AP and TS (April).

Fig. 13. Contour map of solar radiation (kW h/m² day) for AP and TS (May).
Fig. 14. Contour map of solar radiation (kW h/m² day) for AP and TS (June).

Fig. 15. Contour map of solar radiation (kW h/m² day) for AP and TS (July).

Fig. 16. Contour map of solar radiation (kW h/m² day) for AP and TS (August).

Fig. 17. Contour map of solar radiation (kW h/m² day) for AP and TS (September).

Fig. 18. Contour map of solar radiation (kW h/m² day) for AP and TS (October).

Fig. 19. Contour map of solar radiation (kW h/m² day) for AP and TS (November).
6. Conclusion

A FBPN model was developed to predict the solar radiation and estimation was mapped using GIS approach in southern part of India. The developed ANN model is optimized for number of neurons in hidden layer. The input to the model are meteorological data along with location (lat., long., and alt.) and month. Model was trained to give solar radiation as output. Meteorological data at 28 different locations is used to develop model. Model was trained with data of 20 locations and validated with 4 locations. Remaining data of 4 locations was kept aside and not used for model development, which was used to test the model. This was to ensure the model efficiency.

To evaluate the model performance, MAPE was calculated. Nine neurons in one hidden layer has been found giving least MAPE for all four test location in this study. The average MAPE for test location is 2.16. The developed model gives an error less the 2.5%, which can be considered as quite good accuracy. The R-value showed by model was 0.98745 for test data. The model can be considered as capable to estimate solar radiation at any locations all over the AP and TS with high accuracy. It is found that the highest solar radiation available is 6.68 kW h/m²·day in the month of April and lowest is 3.31 kW h/m²·day in the month August. It has been seen that the solar radiation is high in west region and decrease towards east of AP and TS.

Predicted solar radiation was analyzed and mapped using GIS technology. Kriging method was used to interpolate between the segregated points and generate the surface for mean solar radiation value. Surface plot is then used to create contour plot of solar radiation with an interval of 0.1 kW h/m²·day. The contour map was plotted for each month separately. These maps will give better visualization of solar potential in AP and TS, which will help the decision maker to find hot spot for solar technology.

Acknowledgments

We thank NASA website. These data were acquired from the NASA Langley Research Center and Solar Energy web portal supported by the NASA LaRC POWER Project.

References


