An Artificial Intelligence Approach to the Prediction of Global Solar Irradiation in India

Sutapa Mukherjee

Abstract Due to a high demand in solar power generation, the predictions and integration of solar energy sources have become an important area of research work. Previous researches have shown that the artificial intelligent-based method for approximation of global solar radiation provides better accuracy and efficiency. The objective of the present work is to find the global solar radiation using multi-layer feedforward neural network (MLFF) for the year 2016 pertaining three cities Kolkata, Roorkee and Chennai based on five years (2011–2015) hourly radiation database collected from Photovoltaic Geographical Information System (PVGIS). Out of 1, 31,325 collected data, 70% were used for training, 15% were used for validation and the rest 15% were used for testing. This neural network model is developed by taking into consideration different parameters like month, day, time, latitude, longitude, elevation, slope, azimuth as inputs and global irradiance on the inclined plane as output. The obtained results indicate that the proposed technique predicts global solar irradiation with a high accuracy (98.74%) which proves the superiority of the applied method over the conventional one.

Keywords Global irradiance · Artificial neural network · Back-propagation

1 Introduction

One of the important challenges we are facing is energy crisis. Out of different renewable sources of energy, as a clean and cost-free, solar energy is very much popular everywhere. One should be conversant with the various components of solar energy for its effective use. There are various components which attribute to the solar energy to different extents like the duration of sunshine, maximum ambient temperature, any day and month, global radiation, latitude and longitude to name a few. The availability of solar energy is predominantly controlled by the global solar

S. Mukherjee (✉)
Electrical Engineering Department, B. P. Poddar Institute of Management & Technology, Kolkata 700052, India
e-mail: sutapa.mukherjee@bppimt.ac.in

S. Bhaumik et al. (eds.), *Proceedings of International Conference on Industrial Instrumentation and Control*, Lecture Notes in Electrical Engineering 815,
https://doi.org/10.1007/978-981-16-7011-4_24

radiation; however, the measurement of same is done only at a couple of locations due to the involvement of relatively expensive equipment and the subsequent maintenance charges. Researches to corelate the solar radiation with various climatic condition and geometrical factors have been done in the past and the major models were based on different inputs, equations and algorithms [1]. Models developed on the basis of the actual solar radiation reaching a given area outperformed the meteorological data model. Attempts were made to overcome the limitation of inaccessibility of insolation data in certain regions, and consequently, temperature-based models and other suitable modifications [2, 3] were proposed to overcome the shortcoming. ANN technique based on resilient propagation (RP) [4] with the physical features and atmospheric conditions as inputs resulted in a much better and consistent output compared to the classical modelling. Subsequently, soft computing technique was deployed to study the effectiveness in estimating the solar radiation and was found to be instrumental in dealing with practical problems where the nature is data is far more complex and dynamic because of the noise and the nonlinearity associated with it. AI-based approach which includes artificial neural network (ANN), fuzzy logic, adaptive neuro-fuzzy interface system (ANFIS) and data mining (DM) is some of the methods which have been effectively utilized to yield better results [5] than conventional modelling. Several models based on ANN and regression techniques [6] were compared and the former yielded in much better results.

Attempt has been made to devise an easier model for evaluating the total solar irradiation taking input parameters like month, day, time, latitude, longitude, elevation, slope, azimuth as inputs and global irradiance on the inclined plane (plane of the array) (W/m^2) as output for three places in India which are Kolkata, Roorkee and Chennai.

1.1 Dataset

It is already mentioned that this study was carried out using the PVGIS database for three Indian cities Kolkata, Roorkee and Chennai [7]. The parameters used in this study are: day, month, latitude, longitude, elevation, slope, azimuth and global solar irradiance which were collected for six consecutive years (2011–2016). From the total dataset, two separate datasets were created among which the first dataset (for 2011–2015) was used for training, testing and validation and the second dataset (for 2016) has been used for prediction. From the first dataset, out of 1, 31,325 collected data, 70% (91,927) were used for training, 15% (19,698) were used for validation and the rest 15% (19,698) were used for testing. From the second dataset, 8755 data points for each city were used for prediction to evaluate the proposed ANN model. Table1 shows the geographical locations of these three cities Kolkata, Roorkee and Chennai.

Table 1 Geographic information of three cities

City	Latitude	Longitude	Elevation (m)	Mounting type	
				Slope	Azimuth
Kolkata	22.57° N	88.37° E	11	45°	120°
Roorkee	29.86° N	77.89° E	274	45°	120°
Chennai	13.08° N	80.28° E	8	45°	120°

1.2 Artificial Neural Network

The ANN model is a mathematical simulation of the way in which the human brain works [8]. It is a highly interconnected network of neurons which is capable of processing the computational data in parallel. The input signal y_i, for $i = 0, 1, 2, \ldots, n$ is processed through the intermediate layers of neurons which assign a weightage w_{ij} to the inputs followed by an application of sigmoid function resulting in an output equivalent to the weighted mean which is given by Eq. 1.

$$u_j = \sum_{i=0}^{n} w_{ij} y_i \tag{1}$$

The multi-layer feedforward mechanism (MLF) along with the back-propagation is the most commonly used ANN method used for estimating the solar radiation [9–11]. This model can be used for problems which are not linearly separable. Apart from input (i) and one output layer (k), it usually contains one or two intermediate/hidden layer (j) which are interconnected by weights W_{ij} and W_{jk} and is shown in Fig. 1a, and the basic structure of a single artificial neuron is shown in Fig. 1b.

A bias is added to the input and a nonlinearity transforms the sum into an output. This is called the activation function of the node. Linear activations are usually present in the output nodes. The logistic sigmoid function given in Eq. 2 is used for hidden node while the linear equation given by Eq. 3 is used for output node [11].

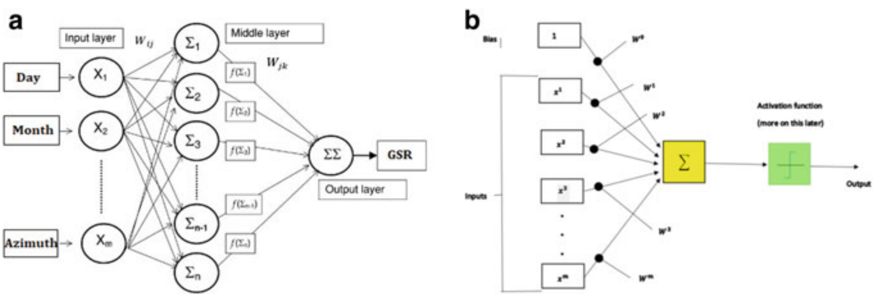

Fig. 1 **a** Typical structure of ANN **b** basic structure of a single artificial neuron

$$f(w) = 1/\left(1 + \mathrm{e}^{-w}\right) \tag{2}$$

$$f(x) = x \tag{3}$$

The synaptic weight is updated on the basis of a procedure called back-propagation in which the error generated at the output after running a cycle is propagated backward through the intermediate layers finally to the input layer [12]. The training cycle is then repeated to minimize the error, and the network gradually moves towards stability. The estimated error used here is the mean squared error (MSE). In addition to the back-propagation technique, other methods like gradient descent, conjugate gradient algorithms, resilient back-propagation, gradient descent with momentum and quasi-Newton algorithms can be used. The number of neurons, the number of intermediate layers and the training algorithm vary from model to model.

1.3 ANN Model Architecture

A three-layer feedforward neural networks (FFNN) simulation has been used for the purpose of investigation. The input variables are fed into the first layer which is basically the input layer followed by the intermediate or the hidden layer which forms the second layer and the third layer is the output layer. Many real-world function problem modelling has been done following this topology [13, 14]. The selection of the neurons in the intermediate layer is very crucial as the complexity of system to be modelled is related to it. For this particular investigation, the selection of optimum number of neurons in the intermediate layer was done by trial-and-error method. Using a range of neurons between 2 and 80 the system was evaluated for minimum error between the predicted and the actual output. For the intermediate layer, the logistic sigmoid function was considered for the transfer function while the output layer was based on the linear transfer function. Back-propagation (BP) with the Levenberg–Marquardt was then applied to train the neural network as it is one of the fastest and more accurate algorithms. The stability of the steepest descent method [15] is combined with the speed of Newton algorithm. This combination is used to compute the Jacobian matrices bypassing the hessian matrix. This approach gives a faster convergence with minimalistic error.

1.4 Model Performance Evaluation [16]

Using two statistical indicators, mean squared error (MSE) and coefficient of determination (regression value (R)), the performance of the models was evaluated.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(H_\text{p} - H_\text{a})^2 \tag{4}$$

where H_p is the predicted value and H_a is the actual value

$$R^2 = \frac{\left[\sum_{i=1}^{n}\left(H_\text{p} - H_\text{p,avg}\right)\left(H_\text{a} - H_\text{a,avg}\right)\right]^2}{\sum_{i=1}^{n}\left(H_\text{p} - H_\text{p,avg}\right)^2 \sum_{i=1}^{n}\left(H_\text{a} - H_\text{a,avg}\right)^2} \tag{5}$$

2 Results

In the first scenario, the artificial neural network was trained using 2011 to 2015 dataset consisting of the input parameters (day, month, time, latitude, longitude, slope and azimuth) and it was observed that the model achieved MSE value of 0.676 for training and 0.645 for testing which is shown in Table 2.

The regression plot [17] of training dataset has been shown in Fig. 2, from which it can be seen that the overall regression value was obtained of 0.94482 with the coefficients of regression line for slope and offset being 0.89 and 0.18, which indicates a good fit of the data. Also, the best validation performance of 0.665 was achieved after 148 iterations as shown in Fig. 4a. The regression plots of testing dataset are shown in Fig. 3 from which it is seen that the regression value was obtained of 0.93869 with the coefficients of regression line for slope and offset being 1.1 and 0.00013 for Kolkata.

The regression value was obtained of 0.95262 with the coefficients of regression line for slope and offset being 1 and 0.014 for Roorkee. The same was obtained of 0.9076 with the coefficients of regression line for slope and offset being 0.99 and 0.1 for Chennai.

The error histogram demonstrates that the maximum instance of error occurred at 0.2272 signifying the accuracy of the trained ANN model in predicting the GSR value for unknown inputs as shown in Fig. 4b. The error histogram chart during tested data for Kolkata, Roorkee and Chennai is shown in Figs. 5, 6 and 7, respectively.

The trained ANN model was used for the prediction of global solar radiation in three Indian cities (Kolkata, Roorkee and Chennai). The statistical performance measures (MSE and R2 value) for all the three cities are presented in Table 3 for both training and testing datasets.

Table 2 Performance evaluation of the proposed ANN model on the training datasets

Results	MSE	*R*
Training	0.676	0.944
Validation	0.665	0.945
Testing	0.645	0.948

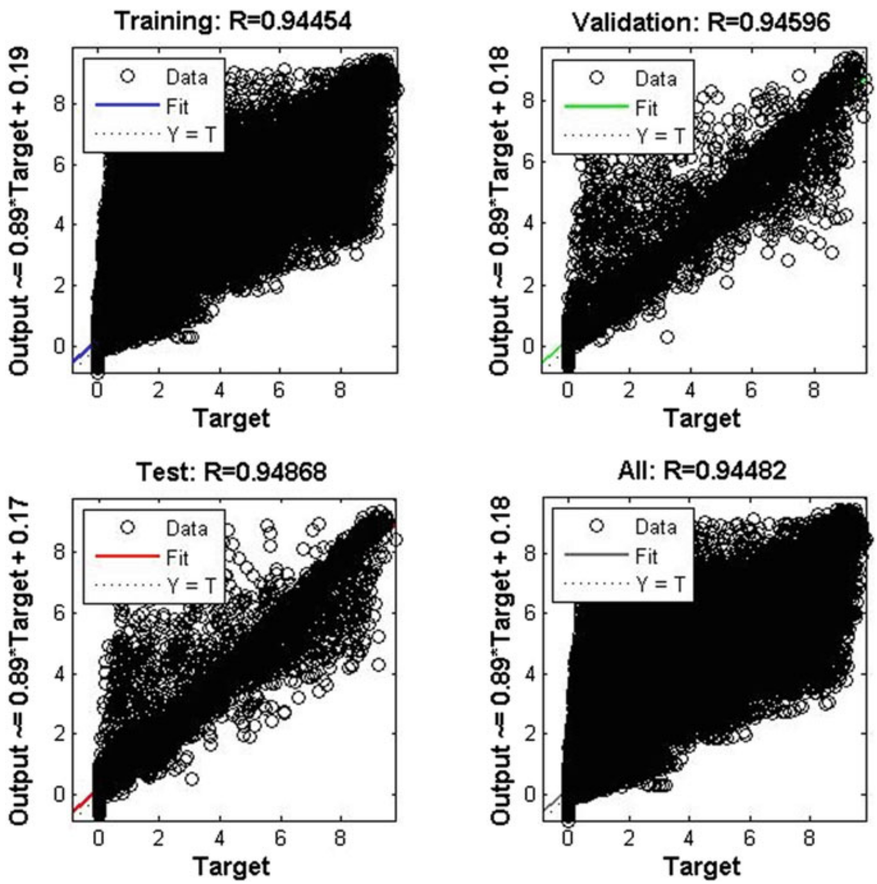

Fig. 2 Regression plot of training dataset

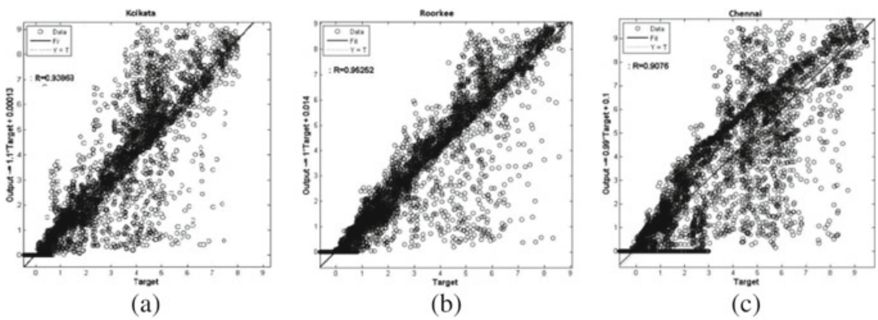

Fig. 3 Regression plot of testing dataset for **a** Kolkata **b** Roorkee and **c** Chennai

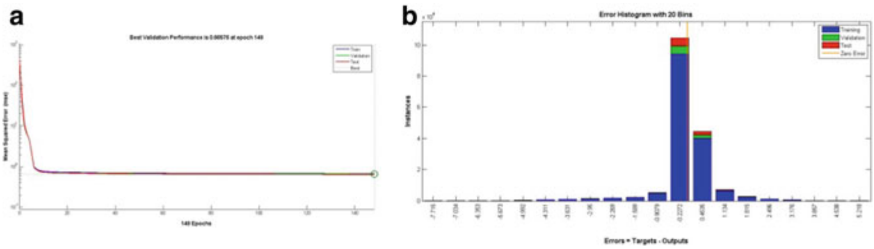

Fig. 4 **a** Best validation performance **b** error histogram of trained ANN

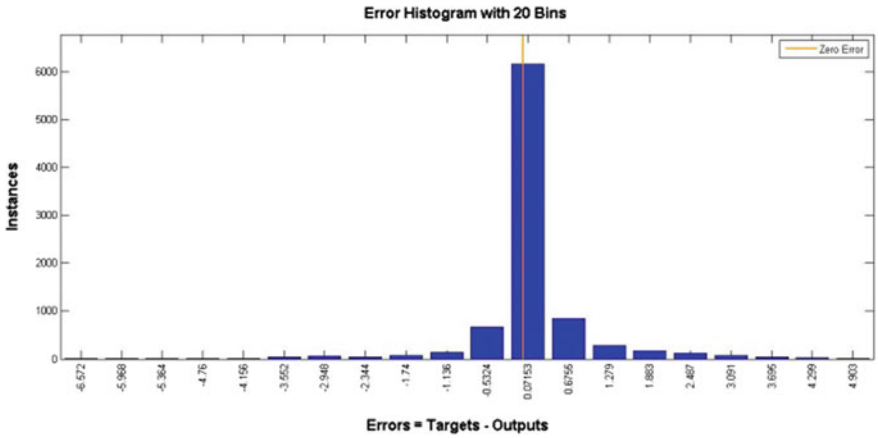

Fig. 5 Error histogram for tested data of Kolkata

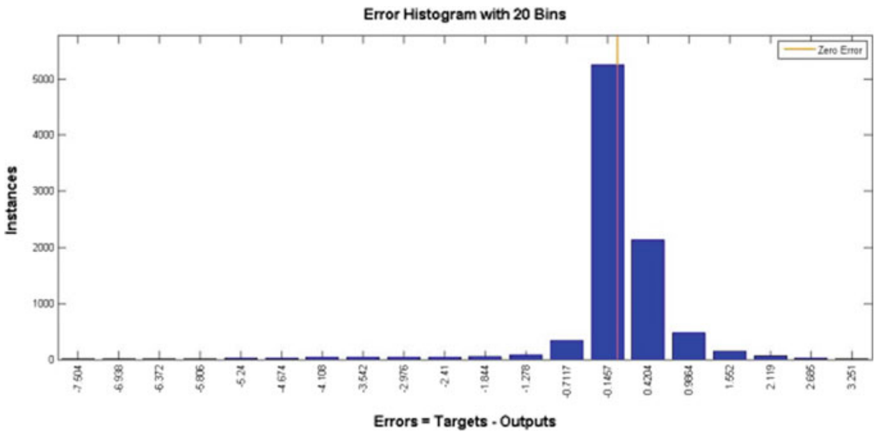

Fig. 6 Error histogram for tested data of Roorkee

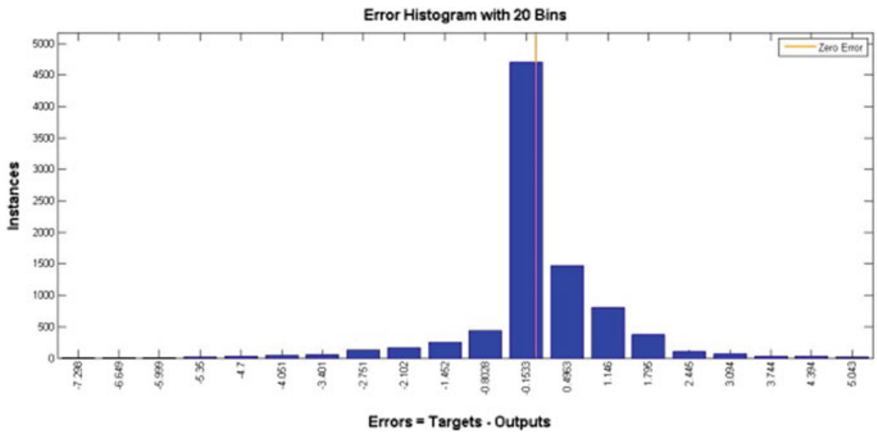

Fig. 7 Error histogram for tested data of Chennai

Table 3 Performance evaluation of the proposed ANN model on the prediction dataset

Cities	MSE	R^2
Kolkata	0.731	0.939
Roorkee	0.544	0.953
Chennai	0.129	0.908

From Table 3, it can be observed that the proposed ANN model can efficiently predict the GSR with low mean squared error and high R^2 value for all three cities.

3 Discussion and Conclusion

In this study, multilayer feedforward (MLFF) ANN technique [18] based on back-propagation algorithm is developed, trained and tested to predict hourly global solar radiation for 2016 for three cities Kolkata, Roorkee and Chennai. The model is developed and proposed by including input parameters like month, day, time, latitude, longitude, elevation, slope, azimuth and global solar irradiation as output. Using the prepared ANN model created with the PVGIS 5 years solar radiation dataset for three cities and testing this one for the prediction of 2016 data revealed that the predictions of solar radiation using our proposed model shows a better result with the obtained PVGIS dataset for 2016. This developed model is suitable for predicting solar radiation for any locations in India for which solar radiation is required for site-specific solar energy applications especially for solar power generation and production.

References

1. Besharat, F., Dehghan, A.A., Faghih, A.R.: Empirical models for estimating global solar radiation: a review and case study. Renew. Sustain. Energy Rev. **21**, 798–821 (2013)
2. Chen, J.-L., Liu, H.-B., Wu, W., Xie, D.-T.: Estimation of monthly solar radiation from measured temperatures using support vector machines—a case study. Renew. Energy **36**, 413–420 (2011)
3. Olatomiwa, L., Mekhilef, S., Shamshirband, S., Mohammadi, K., Petkovic, D., Sudheer, C.: A support vector machine–firefly algorithm-based model for global solar radiation prediction. Sol. Energy **115**, 632–644 (2015)
4. Senkal, O.: Solar radiation and precipitable water modeling for Turkey using artificial neural networks. Meteorol. Atmos. Phys. **127**, 481–488 (2015)
5. Gopalakrishnan, K., Khaitan, S.K. Kalogirou, S.: Soft Computing in Green and Renewable Energy Systems. Springer, Berlin (2011)
6. Paliwal, M., Kumar, U.A.: Neural networks and statistical techniques: a review of applications. Expert Syst. Appl. **36**, 2–17 (2009)
7. PVGISHomepage. https://ec.europa.eu/jrc/en/pvgis
8. Haykin, S., Lippmann, R.: Neural networks, a comprehensive foundation. Int. J. Neural Syst. **5**, 363–364 (1994)
9. Yadav, A.K., Chandel, S.S.: Solar radiation prediction using artificial neural network techniques: a review. Renew. Sustain. Energy Rev. **33**, 772–781 (2014)
10. Qazi, A., Fayaz, H., Wadi, A., Raj, R.G., Rahim, N.A., Khan, W.A.: The artificial neural network for solar radiation prediction and designing solar systems: a systemic literature review. J. Cleaner Prod. **114**, 1–12 (2015)
11. Rezrazi, A., Hanini, S., Laidi, M.: An optimisation methodology of artificial neural network models for predicting solar radiation: a case study. Theor. Appl. Climatol. **123**, 1–15 (2015)
12. Esmaeelzadeh, S.R., Adib, A., Alahdin, S.: Long-term stream flow forecasts by adaptive neuro-fuzzy inference system using satellite images and K-fold cross validation (case study: Dez, Iran). KSCE J. Civ. Eng. 1–9 (2014)
13. Ata, R.: Artificial neural networks applications in wind energy systems: a review. Renew. Sustain. Energy Rev. **49**, 534–562 (2015)
14. Antonopoulos, V.Z., Antonopoulos, A.V.: Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate variables. Comput. Electron. Agric. **132**, 86–96 (2017)
15. Wilamowski, B.M., Chen, Y., Malinowski, A.: Efficient algorithm for training neural networks with one hidden layer. In: IJCNN. Proceedings of the 1999 International Joint Conference, pp. 1725–1728. Citeseer (1999)
16. Quej, V.H., Almorox, J., Arnaldo, J.A., Saito, L.: ANFIS, SVM and ANN soft-computing techniques to estimate daily global solar radiation in a warm sub-humid environment. J. Atmos. Solar-Terr. Phys. **155**, 62–70 (2017)
17. Jain, R., Goel, B.: Prediction ofglobal solar radiation using artificial neural network. Int. J. Adv. Res. Electr. Electron. Instrum. Eng. **2**, 5170–5175 (2013)
18. Aghmadi, A., El Hani, S., Elbouchikhi, E., Naseri, N., Elharouri, K.: Prediction global solar radiation using ANN model based on meteorological data in Morocco. In: 8th International Conference on Power Science and Engineering, pp. 50–54 (2019)