Solar Irradiance Forecasting Model using Artificial Neural Network (ANN)

Syazwani binti Zahare
Faculty of Electrical Engineering
Universiti Teknologi MARA
40450 Shah Alam Selangor, Malaysia
one_niez5488@yahoo.com

Abstract —**This paper presents the Artificial Neural Network (ANN) model for predicting solar irradiance. The inputs of the ANN are the solar irradiance values for previous five one-minute intervals while the output of the ANN is the solar irradiance of the sixth-minute interval. The solar irradiance data were obtained from a weather monitoring station at Green Energy Research Centre (GERC) at Universiti Teknologi MARA, Malaysia. During testing, the ANN produced a low mean absolute percentage error (MAPE) of 10.5796% and high coefficient of determination, R^2 of 0.8925**.

Keyword- Artificial Neural Network (ANN); Solar Irradiance (SI); Coefficient of Determination (R^2); Mean Absolute Percentage Error (MAPE); prediction

I. INTRODUCTION

Solar irradiance is the portion of the sun's radiation which available at the earth's surface. Solar irradiance also known as a kind of popular renewable energy compared with other energy, like geothermal and tidal [1]. Indeed, this energy useful for many applications, such as increasing the temperature of the water or exciting electrons in a photovoltaic cell. Moreover, it also supplies energy to natural processes like photosynthesis. Solar energy is free, clean and available on the earth throughout the year [3]. In this study, the outputs from a Grid-Connected Photovoltaic (GCPV) system are the data that used as input parameters to predict the solar irradiance.

Artificial Neural Network (ANN) is the one of the popular prediction technique which has the capacity to learn, memorize and create relationships among data [4]. Feed-forward and feedback networks are two major categories of ANN. Feed-forward networks have signals travel one way only which is signaled from input to output [5]. Signals that travel in both directions are known as feedback networks. In this study, the feed-forward network was used.

The ANN model created consists of three layers namely the input layer, hidden layer and output layer [6]. There were several research that has been used the ANN for predicting solar irradiance. In the article of Mubiru et al. has used artificial neural networks to estimate the monthly average of daily global solar irradiation. The input to the network included sunshine hours, cloud cover, maximum temperature, latitude, longitude and altitude. The ANN results compared with the result of empirical model [13].

Fig. 1 shows the proposed ANN model with input parameters of solar irradiance (SI) for every 1 minute, 2 minutes, 3 minutes, 4 minutes and 5 minutes. The output of the prediction is SI for 6 minutes. Number of hidden layer was selected is 1.

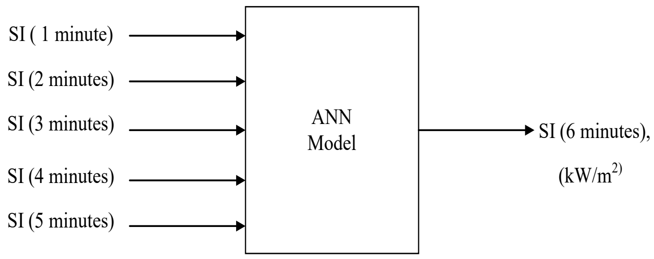

Figure 1: ANN model for the prediction of solar irradiance for 6 minutes-ahead output from grid connected PV system.

II. DATA COLLECTION

The solar irradiance data were obtained from the Grid Connected Photovoltaic System (GCPV) at Universiti Teknologi Mara (UiTM) Green Energy Research Centre. The characteristics of the PV system shown in Table I below.

TABLE I. CHARACTERISTICS OF THE PV SYSTEM

Characteristics	Parameters
Type of PV module	Mono-crystalline
Size of PV array	1.12kWp
Type of Inverter	StecaGrid 2000 + Master
Size of Inverter	1.075kW
Type of Installation	Free standing

The data used in this study were collected from 24 September 2013 to 2 October 2013. Data were recorded at 1 minute interval from the data logger inside the inverter [7]. The data were divided into two groups. 60% from the collected data which is 3000 data were used in the training process while 40% contains 2000 data had been used for testing process [8].

III. DESIGNING THE ANN MODEL

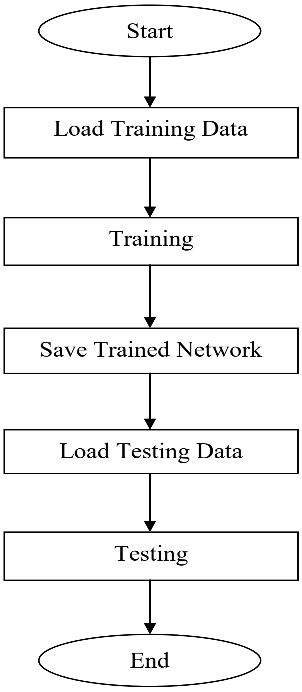

Figure 2. Flow chart for designing artificial neural network model.

Fig. 2 shows the flow chart to design the ANN model used the Matlab R2013a software and its neural network toolbox [9].

The design starts by load the training data from an Excel file. Firstly, load the training data set which consists of SI (from 1 minute to 5 minutes) and SI (6 minutes). These 1 minute to 5 minutes of SI are the inputs to the ANN while 6 minutes of SI as the target output.

The number of neurons in the input layers is 5 while the number of neurons in output layer is 1. In the training set, the inputs and output data will be normalized in a range between -1 to +1 [10]. Next, set number of neurons in the hidden layer, the learning rate, the momentum rate, the activation function and the learning rule. Later, determine the prediction performance involving the coefficient of determination (R^2), the root mean square error (RMSE), mean absolute percentage error (MAPE) and the mean bias error (MBE).

In this study, R^2 [11] and MAPE had been used to evaluate the performance of the ANN model. R^2 was used to find the correlation between parameters in mathematical model [2]. The R^2 and MAPE [12] are calculated as below:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - x_i)^2}{\sum_{i=1}^{n}(x_i - x_{avg})^2} \tag{1}$$

$$\text{MAPE} = \left[\frac{1}{n}\sum_{i=1}^{n}\frac{|x_i - y_i|}{x_i}\right] \times 100 \tag{2}$$

where n is the number of data patterns, y_i is the predicted value obtained from the ANN model, x_i is the actual value and x_{avg} is the average of the actual value.

After that, save the training network obtained. The next step is loading the testing data from the Excel file. Then the data also will be normalized and determine the prediction performance as described in the training procedure.

IV. TRAINING PARAMETERS OF THE ANN MODEL

In the Matlab programming, the x_1, x_2, x_3, x_4 and x_5 refers to the number of neurons in the hidden layer, the learning rate, the momentum rate, the activation function and the learning rule. Table II summarizes the parameters and values that had been used to train the ANN model. The values for the number of hidden layers, number of neurons in the input layer and the number of neurons in the output layer are fixed. As we can see from the Table II, the selected value for mean square error goal and maximum number of iterative updates are 10^{-2} and 1000 respectively.

TABLE II. DESIGN AND TRAINING PARAMETERS OF THE ANN MODEL

Parameters	**Selected value**
Number of hidden layer	1
Number of neurons in hidden layer	1-20
Number of neurons in input layer	5
Number of neurons in output layer	1
Mean square error goal	10^{-2}
Maximum number of iterative updates	1000
Learning rate	0-1
Momentum rate	0-1
Activation function	tansig/logsig/purelin
Learning rule	trainlm/trainscg/trainrp/trainbfg /traingdx

Other than that, the best value for the rest of the training parameters had been selected depends on the highest R^2 and the lowest MAPE obtained. The step by step procedure can be summarized as follows:

Step 1: Verifying the number of neurons in hidden layer from 1 to 20. The best number of neurons had been chosen depends on the lowest number of MAPE.
Step 2: Using the number of neurons that had been verified, choose the activation function. There are three activation function namely tangent-sigmoid (tansig), log-sigmoid (logsig) and purely-linear (purelin). The best activation function based on the lowest MAPE and the highest number of R^2.

Step 3: Choose the learning rule. The learning rule can be changed to Levenberg-Marquardt backpropagation (trainlm), Scaled Conjugate Gradient backpropagation (trainscg), Resilient backpropagation(trainrp), BFGS quasi-Newton backpropagation (trainbfg) and Gradient Descent with Momentum and Adaptive lr backpropagation (traingdx). Similar to activation function the best learning rule also based on the lowest MAPE and the highest number of R^2.

Step 4: Determine the best number of learning rate. Using the results for x1, x4 and x5 obtained, choose the learning rate from 0.05 to 1.00 with an interval of 0.05. The lowest number of MAPE had been chosen as a result for learning rate.

Step 5: Continue with the results of x1, x2, x4 and x5, identify the momentum rate from 0.10 to 1.00 with increment of 0.05. Like the previous results, the best number of momentum rate also selected from the lowest number of MAPE.

V. RESULT AND DISCUSSIONS

The ANN model was trained using 3000 of data patterns and the remaining which is 2000 data patterns was used for testing. The performance results for ANN training and testing shown in Table III. This performance shows the accuracy of the ANN model that had been used as a prediction model in this study.

The result shows that number of neurons in hidden layer are 10. Other than that, learning rate and momentum rate selected values are 0.15 and 0.30 respectively. The mean square error goal is fixed with 10^{-2} and the maximum number of iterative updates of 1000. The activation function had been set to logsig and the best learning rule had been chosen was trainscg.

Fig. 3 presents the result for number of neurons in the hidden layer. From the number of neurons that had been verified, the result shows that 10 are the numbers of neurons in hidden layer which was having the lowest percentage of MAPE during the testing process.

TABLE III. TRAINING PARAMETERS AND RESULTS FOR ANN TRAINING AND TESTING

Parameters	Selected value
Number of training data patterns	3000
Number of testing data patterns	2000
Number of hidden layer	1
Number of neurons in hidden layer	10
Number of neurons in input layer	5
Number of neurons in output layer	1
Mean square error goal	10^{-2}
Maximum number of iterative updates	1000
Learning rate	0.15
Momentum rate	0.30
Activation function	logsig
Learning rule	trainscg
Coefficient of determination (R^2) for training	0.8796
Coefficient of determination (R^2) for testing	0.8925
Mean absolute percentage error (MAPE) for training	10.5796
Mean absolute percentage error (MAPE) for testing	12.8279

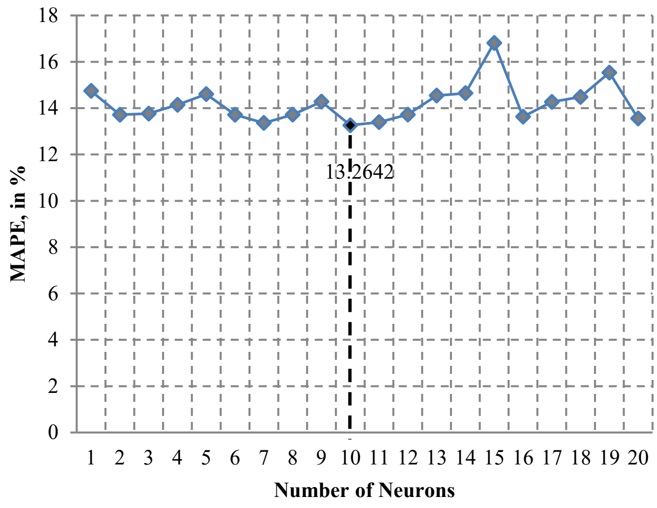

Figure 3. MAPE testing performance for number of neurons in hidden layer.

Fig. 4 shows the comparison between three activation function i.e. tangent-sigmoid (tansig), log-sigmoid (logsig) and purely-linear (purelin). We can see from the bar chart that logsig are the best activation function that presents the lowest MAPE and the highest number of R^2.

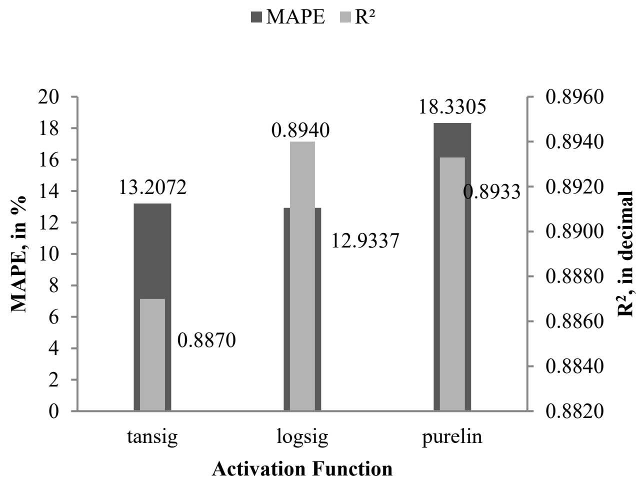

Figure 4. Testing performance for three different types of activation function.

The best performances for learning rule are 12.9337% for MAPE and 0.8940 for R^2. Trainscg shows the lowest MAPE and the highest number of R^2 compared to others i.e. Levenberg-Marquardt backpropagation (trainlm), Scaled Conjugate Gradient backpropagation (trainscg), Resilient backpropagation (trainrp), BFGS quasi-Newton backpropagation (trainbfg) and Gradient Descent with Momentum and Adaptive lr backpropagation (traingdx). The result presented in Fig. 5.

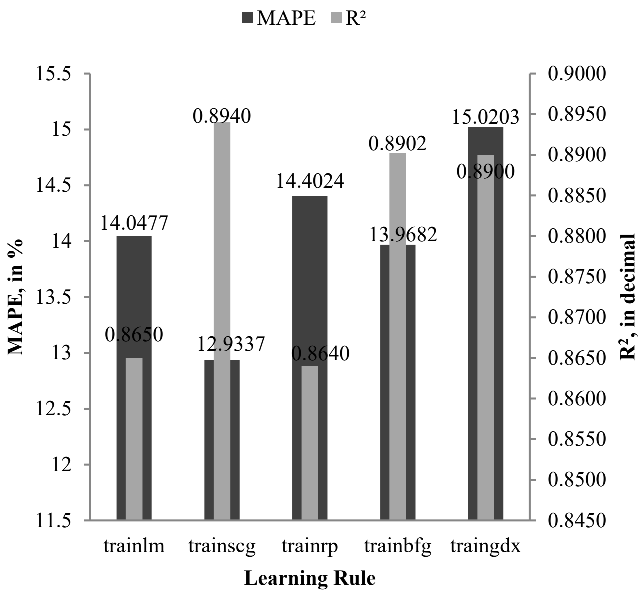

Figure 5. Testing performance for five different types of learning rule.

As we can see from the result illustrated in Fig. 6, 12.8279% was the lowest percentage in the MAPE performance. Based on that result, 0.15 had chosen as the best number of learning rate.

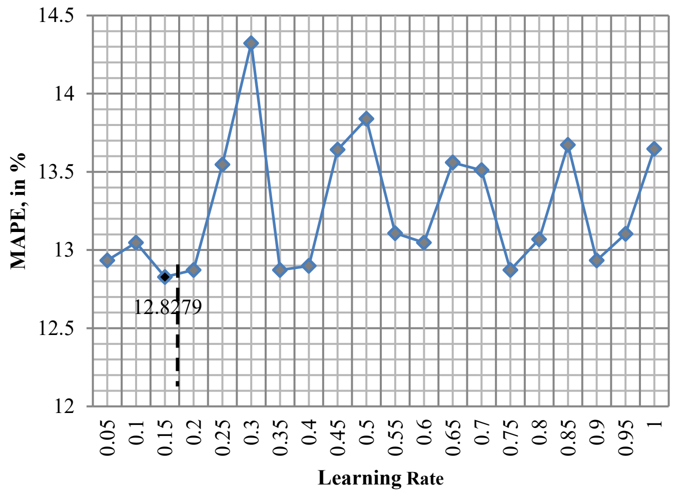

Figure 6. MAPE testing performance for learning rate.

Fig. 7 shows the momentum rate was 0.3 when the MAPE was having the lowest percentage. This result obtained after verified the momentum rate from 0.10 to 1.00 with increment of 0.05.

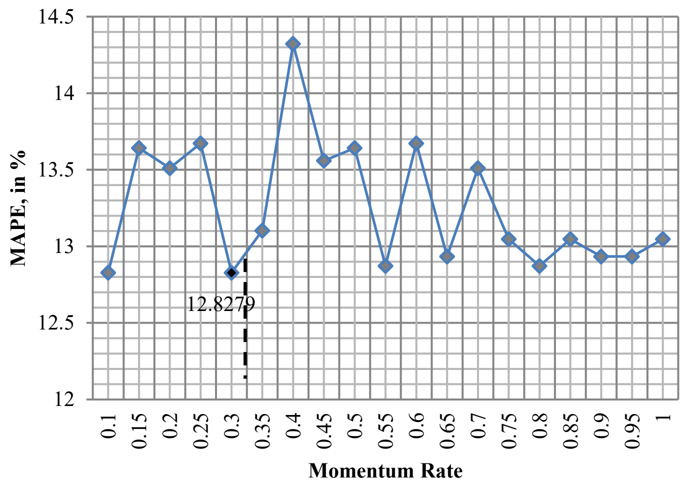

Figure 7. MAPE testing performance for momentum rate.

VI. CONCLUSIONS

This paper has presented the prediction of solar irradiance. This prediction was performed using the ANN. The best performance of this ANN model has 10 neurons in the hidden layer, logsig as activation function and also trainscg as learning rule. The coefficient of determination (R^2) of ANN model is 0.8925 while mean absolute percentage error (MAPE) was found to be 12.8279% during the testing process. The low MAPE value and high R^2 value had shown that the ANN model has a good predictive performance and was useful in predicting solar irradiance. In order to improve this research, this work can be extended by change the predicting into 1 hour or 1 day-ahead.

ACKNOWLEDGMENT

The author would like to thank Green Energy Research Centre (GERC) at Universiti Teknologi MARA, for providing the set of data used for training and testing for this study.

REFERENCES

[1] M. Cococcioni, E. D'Andrea, and B. Lazzerini, "24-hour-ahead forecasting of energy production in solar PV systems," *2011 11th International Conference on Intelligent Systems Design and Applications*, pp. 1276–1281, Nov. 2011.

[2] S. I. Sulaiman, T. K. A. Rahman, S. Member, I. Musirin, and S. Shaari, "Optimizing Three-layer Neural Network Model for Grid-connected Photovoltaic System Output Prediction 1," no. July, pp. 338–343, 2009.

[3] T. Khatib, A. Mohamed, K. Sopian, and M. Mahmoud, "Assessment of Artificial Neural Networks for Hourly Solar Radiation Prediction," *International Journal of Photoenergy*, vol. 2012, pp. 1–7, 2012.

[4] J. Mubiru, "Using Artificial Neural Networks to Predict Direct Solar Irradiation," *Advances in Artificial Neural Systems*, vol. 2011, pp. 1–6, 2011.

[5] O. S. Eluyode and D. T. Akomolafe, "Comparative study of biological and artificial neural networks," vol. 2, no. 1, pp. 36–46, 2013.

[6] A. Ouammi, D. Zejli, H. Dagdougui, and R. Sacile, "Analysis of moroccan wind and solar potential using artificial neural network approach," *2012 7th International Conference on System of Systems Engineering (SoSE)*, pp. 43–47, Jul. 2012.

[7] S. I. Sulaiman, T. K. A. Rahman, I. Musirin, and S. Shaari, "Fast Evolutionary Programming-based Hybrid Multi-Layer Feedforward Neural Network for predicting Grid-Connected Photovoltaic system output," *2012 IEEE 8th International Colloquium on Signal Processing and its Applications*, pp. 44–47, Mar. 2012.

[8] F. Wang, Z. Mi, S. Su, and H. Zhao, "Short-Term Solar Irradiance Forecasting Model Based on Artificial Neural Network Using Statistical Feature Parameters," *Energies*, vol. 5, no. 12, pp. 1355–1370, May 2012.

[9] V. Sivamadhavi and R. S. Selvaraj, "Prediction of monthly mean daily global solar," no. 6, pp. 1501–1510, 2012.

[10] S. I. Sulaiman, T. K. Abdul Rahman, I. Musirin, and S. Shaari, "Artificial neural network versus linear regression for predicting Grid-Connected Photovoltaic system output," *2012 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*, pp. 170–174, May 2012.

[11] A. Azadeh and A. Maghsoudi, "Using an Integrated Artificial Neural Networks Model for Predicting Global Radiation: The Case Study of Iran."

[12] K. Angela, S. Taddeo, and M. James, "Predicting Global Solar Radiation Using an Artificial Neural Network Single-Parameter Model," *Advances in Artificial Neural Systems*, vol. 2011, pp. 1–7, 2011.

[13] J. Mubiru and E. J. K. B. Banda, "Estimation of monthly average daily global solar irradiation using artificial neural networks," *Solar Energy*, vol. 82, no. 2, pp. 181–187, Feb. 2008.