Proceedings of 2012 4th International Conference on Machine Learning and Computing
IPCSIT vol. 25 (2012) © (2012) IACSIT Press, Singapore

A Genetic Algorithm-Based Hybrid Multi-Layer Feedforward Neural Network for Predicting Grid-Connected Photovoltaic System Output

Shahril Irwan Sulaiman [+], Titik Khawa Abdul Rahman and Ismail Musirin

Faculty of Electrical Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

Abstract. This paper presents a Hybrid Multi-Layer Feedforward Neural Network (HMLFNN) technique for predicting the output from a grid-connected photovoltaic (GCPV) system. In the proposed HMLFNN, Genetic Algorithm (GA) was selected as the optimizer for the training process of the Multi-Layer Feedforward Neural Network (MLFNN). GA was used to optimize the number of neurons in the hidden layer, the learning rate, the momentum rate, the type of activation function and the learning algorithm. In addition, the MLFNN utilized solar irradiance (SI) and module temperature (MT) as its inputs and kWh energy as its output. When compared with the classically trained MLFNN, the proposed HMLFNN was found to be superior in terms of having shortest computation time and lower prediction error.

Keywords: Genetic Algorithm (GA), Multi-Layer Feedforward Neural Network (MLFNN), Grid-connected Photovoltaic (GCPV), Solar Irradiance (SI), Module Temperature (MT)

1. Introduction

Photovoltaic (PV) technology has been a crucial type of renewable energy used in many countries as the sunlight is freely available throughout the year. In urban areas where the utility-grid is highly available, grid-connected photovoltaiv systems are often used as an alternative mode of electricity generation. In GCPV system, the DC power generated by the PV array is channelled via an inverter for power conditioning before directing the resulting AC power to the grid or the load. Nevertheless, the expected energy output from the GCPV system is frequently inconsistent due to the fluctuating weather conditions throughout the day. As a result, there is a need for predicting the the output from the system such that the performance of the system could be justified.

A few studies had been conducted to predict the output from GCPV systems. An MLFNN model for predicting the output power was initially developed using solar irradiance, ambient temperature and wind speed [1]. However, another study found out that a two-variate MLFNN using SI and MT or SI and ambient temperature as the inputs had presented the best output power prediciton model [2]. However, the training parameters of these MLFNN models were chosen using classical approach through a heuristic process. As the parameters were selected based on trial and error process, the overall training became very tedious and time consuming. Therefore, an evolutionary MLFNN were developed to optimize the training parameter selection during MLFNN training [3-5]. Nevertheless, only the number of neurons in hidden layer, the learning rate and the momentum rate were evolved during the training process, thus limiting the accuracy of the training performance. In this, study, a hybrid MLFNN was proposed to predict the energy output from a GCPV system. However, besides the training parameters mentioned earlier, the type of activation function and the type of learning algorithm were also allowed to evolve during the training process. In addition, GA was used as an optimizer for the selection of the best training parameters for the MLFNN, thus providing a hybrid approach for the prediction technique. After the evolutionary training of the MLFNN, testing process was conducted to validate the training process.

[+] Corresponding author. Tel.: +603-55211996; fax: +603-55211990.
E-mail address: shahril@salam.uitm.edu.my.

2. System Output Prediction Model

The GCPV system output prediction was implemented using an MLFNN with single hidden layer. The MLFNN contains SI, in Wm^{-2} and MT, in °C as the inputs and the kWh energy of the system as the sole output. The input and output data were obtained from a GCPV system located at the rooftop of Malaysian Green Technology Corporation (MGTC), Bandar Baru Bangi, Malaysia. The system comprises 45.36 kWp poly-crystalline PV array and a 40 kW inverter (IG2). The irradiance and temperature sensors were connected to a built-in data logger inside the inverter while all data were recorded at fifteen minute interval. 80% of the collected data or 800 data patterns were assigned for the training process whereas the remaining 20% of the collected data or 200 data patterns had been used for the testing process. The implementation of the MLFNN was conducted in two stages, i.e. the training and testing processes, which were performed in Matlab.

3. Hybrid Multi-Layer Feedforward Neural Network

The GCPV system output prediction was implemented using an MLFNN with single hidden layer. The MLFNN contains SI, in Wm^{-2} and MT, in °C as the inputs and the kWh energy of the system as the sole output. The input and output data were obtained from a GCPV system located at the rooftop of Malaysian Green Technology Corporation (MGTC), Bandar Baru Bangi, Malaysia. The system comprises 45.36 kWp poly-crystalline PV array and a 40 kW inverter (IG2). The irradiance and temperature sensors were connected to a built-in data logger inside the inverter while all data were recorded at fifteen minute interval. 80% of the collected data or 800 data patterns were assigned for the training process whereas the remaining 20% of the collected data or 200 data patterns had been used for the testing process. The implementation of the MLFNN was conducted in two stages, i.e. the training and testing processes, which were performed in Matlab.

The GA-based HMLFNN was mainly developed to search for the optimal training parameters, i.e. the number of neurons in the hidden layer, the learning rate, the momentum rate, the transfer function in the hidden layer and the learning algorithm. These training parameters, also known as the decision variables for the optimization task, were transcribed as x_1, x_2, x_3, x_4 and x_5 respectively for the evolution process. On the other hand, the objective function for the optimization process was to minimize the RMSE during training. The proposed algorithm was also written in Matlab using the following steps:

Step 1: Generate M population of sets of random numbers, x_1, x_2, x_3, x_4 and x_5. x_1 was set to have integer values from 1 to 20. On the other hand, both x_2 and x_3 were set to have continuous decimal values between 0 and 1. In addition, x_4 was transcribed to have an integer value selected from 1 to 3 where these integers represent logistic-sigmoid (LOGSIG), hyperbolic tangent-sigmoid (TANSIG) and purely linear (PURELIN) activation function respectively. Besides that, x_5 was set to have an integer value selected from 1 to 3 where these integers represent Levenberg-Marquardt (LM), Scaled Conjugate Gradient (SCG) and BFGS quasi-Newton backpropagation (BFG) algorithm respectively. Eventually, each set of random numbers forms the initial candidates of the optimal solutions known as parent.

Step 2: Perform fitness evaluation of each set of random number by training the MLFNN to determine the Root Mean Squared Error (RMSE) value of the prediction. The RMSE was calculated using

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(z_{p,i} - z_{t,i}\right)^2}{n}} \tag{1}$$

where n is the number of data patterns. $z_{p,i}$ and $z_{t,i}$ are the predicted and the target value of the MLFNN output respectively. Besides RMSE, the coefficient of determination, R^2 was also calculated using

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(z_{p,i} - z_{t,i}\right)^2}{\sum_{i=1}^{n}\left(z_{t,i} - z_{t\_avg}\right)^2} \tag{2}$$

where z_{t_avg} is the average value of the target MLFNN outputs. R^2 is a model fit commonly used in estimating the correlation between parameters in mathematical model. Nevertheless, RMSE had been selected as the sole fitness value and key performance indicator for the prediction while the R^2 was used to verify the validity of the RMSE of the network, i.e. the R^2 of the prediction must be reasonably high, i.e. greater than 0.9000 although low RMSE was achieved by the network.

***Step 3*:** Select M parents for mating based on the fitness value. The top M parents with the best fitness value were selected for the generation of offspring.

***Step 4*:** Create offspring by crossover and mutation. In this study, Gaussian mutation scheme was used as the mutation operator.

Step 5: Perform fitness evaluation of each offspring by repeating step 2.

***Step 6*:** If the difference between the maximum fitness and minimum fitness in the offspring population was less than 0.01, the program was stopped and the current population were transcribed as the final solutions. Otherwise, step 4 was repeated.

Upon completion of the training process, the trained network was saved. Later, the trained network was used for testing using the testing data. The performance of the testing process was also quantified using RMSE and R.

Tab.1 Training parameters and performance using different MLFNN techniques

Training parameters & performance	Classical MLFNN	GA-based HMLFNN
Number of neurons in hidden layer	18	16
Learning rate	0.1000	0.4100
Momentum rate	0.7500	0.3500
Type of activation function	LOGSIG	LOGSIG
Type of learning algorithm	LM	BFG
RMSE, in kWh	377.06	376.11
R^2	0.98133	0.97816
Computation time, in sec	18012	4487

4. Results and Discussions

The performance of the proposed GA-based HMLFNN is tabulated in Tab. 1. Although similar type of activation function was produced in both classical and proposed training method, the GA-based MLFNN was able to be implemented using smaller number of neurons with different set of learning rate and momentum rate when compared to the classical MLFNN. In addition, unlike classical MLFNN which utilized LM as the learning algortihm, the GA-based HMLFNN had yielded BFG as the optimal type of learning algorithm. In terms of prediction performance, GA-based HMLFNN was had outperformed the classical MLFNN by producing lower RMSE with slightly lower R^2. Besides that, the GA-based HMLFNN was also found to be superior in terms of computation efficiency since the MLFNN could be trained at approximately 4 times faster compared to the classical training approach.

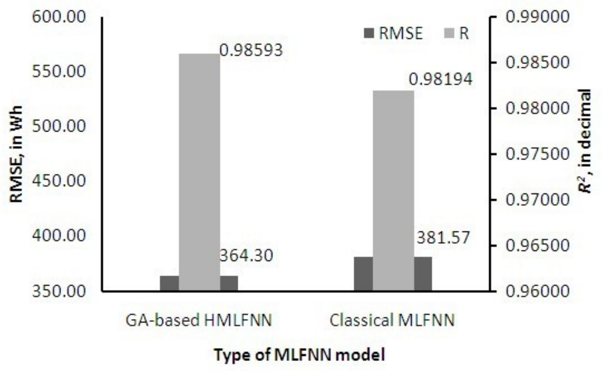

Fig. 1: Testing performance of different MLFNN techniques

On the other hand, the testing performance of the proposed HMLFNN is shown in Fig. 1. The GA-based HMLFNN had produced lower RMSE of almost 4.5% when compared to the classical MLFNN. In addition, the resulting R^2 of the GA-based HMLFNN was also slightly higher than the R^2 of the classical MLFNN.

Since the GA-based HMLFNN had produced smaller RMSE when compared to the classical MLFNN during both training and testing, the proposed HMLFNN method is justified. In fact, the proposed HMLFNN also has faster computation time compared to the classical MLFNN.

5. Conclusions

This paper presents a hybrid MLFNN approach of predicting the energy output of a GCPV system. GA was used to optimize the training of the MLFNN by selecting the optimal number of neurons in hidden layer, the learning rate, the momentum rate, the type of activation function and the type of learning algorithm. Results showed that the proposed GA-based HMFLNN had outperformed the classical MLFNN by producing lower RMSE in both training and testing, thus indicating better prediction accuracy. Apart from that, the proposed HMLFNN was also found to have faster computation time compared to the classical MLFNN.

6. Acknowledgements

This work was supported in part by the Ministry of Higher Education, Malaysia under the Fundamental Research Grant Scheme, Ref: 600-RMI/ST/FRGS 5/3/Fst (81/2010), and by the Excellence Fund, Universiti Teknologi MARA, Malaysia, Ref: 600-RMI/ST/DANA 5/3/Dst (283/2009).

7. References

[1] S. I. Sulaiman, I. Musirin, and T. K. A. Rahman. Prediction of Grid-Photovoltaic System Output Using Three-variate ANN Models. *WSEAS Transactions on Information Science and Applications* 2009, **6**(8):1339-1348.

[2] S. I. Sulaiman, T. K. A. Rahman, I. Musirin, and S. Shaari. Performance analysis of two-variate ANN models for predicting the output power from grid-connected photovoltaic system. *International Journal of Power, Energy and Artificial Intelligence* 2009, **2**(1): 73-77.

[3] X. Yao and Y. Liu. Towards designing artificial neural networks by evolution. *Applied Mathematics and Computation* 1998, **91**: 83-90.

[4] S. I. Sulaiman, T. K. A. Rahman, and I. Musirin. Partial Evolutionary ANN for Output Prediction of A Grid-connected Photovoltaic System. *International Journal of Computer and Electrical Engineering* 2009, **1**(1): 41-46.

[5] X. Fu, D. Per, and S. Zhang. Evolving neural network using variable string genetic algorithm for color infrared aerial image classification. *Chinese Geographical* Science 2008, **18**(2): 162-170.