[1]S. I. Sulaiman,
[2]T. K.A. Rahman,
[3]I. Musirin,
[4]S. Shaari

Performance Analysis of Two-variate ANN Models for Predicting the Output Power from Grid-connected Photovoltaic System

IJPEAI
International Journal of Power, Energy and Artificial Intelligence

Abstract—This paper presents the prediction of total AC power output from a grid-connected photovoltaic (PV) system using two-variate artificial neural network (ANN) models. In this study, multi-layer feedforward ANN models for the prediction of total AC power output from a grid-connected PV system has been considered. Three ANN models were developed based on different sets of two-variate ANN inputs. The first ANN model utilizes solar radiation and ambient temperature as its inputs while the second model uses solar radiation and wind speed as its inputs. On the other hand, the third model uses solar radiation and PV module temperature as its inputs. However, all the three models utilize similar type of output which is the total AC power produced from the grid-connected PV system. Data filtering process was introduced to choose quality data patterns to be processed during training. Thus, only informative features were available for the prediction. In addition, the performance of each ANN model was characterized by the correlation coefficient (*R*) and root mean square error (*RMSE*) of the prediction. After training process was completed, testing process was performed to decide whether the training process should be repeated or stopped. Besides selecting the best prediction model, this study also exhibits some of the experimental results which illustrate the effectiveness of the data filtering in predicting the total AC power output from a grid-PV system. Fully trained ANN models are expected to be able to predict the AC power output from a set of un-seen data patterns in the future.

Keywords – artificial neural network (ANN), photovoltaic (PV), correlation coefficient (*R*), root mean square error (*RMSE*), prediction.

I. INTRODUCTION

Photovoltaic involves solar electricity produced from solar energy. As solar energy can be obtained freely in many places most of the time, solar power generation offers significant advantages compared to other sources of renewable energy. While the amount of solar radiation absorbed by PV module is proved to influence the performance of a PV system, the AC PV power output is usually measured as PV system performance.

Besides having the ability to operate as stand-alone systems, PV technology is often coupled with other types of energy technology such as wind generator, fuel generator, micro-hydro power generator as well as the conventional electricity grid network. Thus, the overall power supply performance and reliability can be significantly improved.

One of the most popular type of PV-hybrid power generation is the grid-PV system. This system mainly requires PV arrays to be connected to the conventional grid set up by the power utility. The grid will provide power back-up to the loads if the PV arrays fail to meet the load demand. Grid-PV system can be configured as a centralized system or a decentralized system, depending on the costs and suitability. Nevertheless, many countries have now promoted more decentralized rooftop grid-PV systems to reduce the capital, operating and maintenance costs of their local power utilities.

Similar to other off-grid PV system, the implementation of grid-PV systems can be also undermined by the issue of unpredictability. Despite having all the benefits of solar power generation, the performance of a grid-PV system is controlled by a few limitations. First and foremost, the power generated by PV array may not meet the load demand all the time due to the variations of sun position and climatic conditions [1]. As the apparent motion of the sun is different throughout the year, the amount of light intensity that reaches the earth is diverse in place and time. Secondly, the presence of clouds and rains would also reduce the amount of irradiation delivered to a particular site due to scattering and absorption [2]. Thirdly, the tilt angle and orientation of the PV array would equally restrict the amount of solar energy that can be absorbed by the array [2]. This factor becomes a major problem in rooftop grid-PV system installation because the available area and orientation of rooftop are normally fixed for a building. In addition, the performance of PV array is substantially influenced by surrounding temperature and solar cell temperature [2]-[3].

Due to the unpredictability in PV system, Artificial Neural Network (ANN) is employed to predict PV system parameters. It has been proved to be useful because it requires no prior knowledge of the internal parameters and also involves minimal computations [4].

Firstly, a research on the unpredictability of electrical loads was done using ANN [5]. Later, similar technique was used in predicting electrical energy consumption [6]. The two studies utilized ANN as a tool to predict electrical load variations. On the other hand, the same technique can be employed in predicting the amount of solar radiation

[1]S. I. Sulaiman is with University Teknologi MARA, Shah Alam, Malaysia. He can be reached at shahril_irwan2004@yahoo.com.
[2]T. K. Abdul Rahman is with University Teknologi MARA, Shah Alam, Malaysia. She can be reached at takitik@streamyx.com.
[3]I. Musirin is with University Teknologi MARA, Shah Alam, Malaysia. He can be reached at i_musirin@yahoo.co.uk.
[4]S. Shaari is with University Teknologi MARA, Shah Alam, Malaysia. He can be reached at solarman@salam.uitm.edu.my.

received at a site [7]. Moreover, ANN is also utilized to size PV system parameters with minimum input data [8].

In relation with PV system output prediction, feedforward back-propagation ANN models were developed to predict the energy (in kWh) output of a grid-PV system using different combination of solar radiation, module temperature and clearness index [9]. Likewise, the output performance of 150W monocrystalline module had specifically modeled using the same architecture but with a two-layer configuration and different inputs [4]. The ANN utilizes solar irradiance, ambient temperature and module temperature as its inputs while voltage and current are identified at its outputs. These studies have proven that ANN is capable of predicting the output of PV systems. Although these ANN-related studies have produced satisfactory results in terms of accuracy and convergence speed, none of these efforts had demonstrated the capability of predicting total AC power output from the inverters of a grid-PV system, justifying this work.

This study presents the implementation of two-variate multilayer feedforward ANN models in predicting the AC power output from a grid-connected photovoltaic (PV) system. Three models have been developed for this investigation

II. DEVELOPMENT OF TWO-VARIATE ANN MODELS

ANN is a generalization process for mathematical models based on biological nervous system [10]. The fundamental processing element of an ANN is called neuron. In basic computational model, the neuron collect input signals from other neurons or sources and merge them. It will then perform necessary computation before mapping them to an output.

Although a few ANN architectures and training algorithms have been introduced for predicting purposes, the multi-layer feedforward neural network with Levenberg-Marquardt backpropagation training algorithm has been widely used in solving many engineering problems due to its good generalization capability and simplicity. Therefore, this type of ANN has been employed in this study. Generally, it consists of one input layer, one hidden layer and an output layer. However, it can also be realized with more than one hidden layer. In this work, the different input combinations are tested with the feedforward ANN model of two hidden layers.

The main objective of this study is to develop ANN models that can accurately predict the total AC power output from a rooftop grid-connected PV system using multiple inverters. Secondly, this research is also aimed at improving the training process through data filtering.

The ANN data consists of solar radiation, *SR* (in kW/m^2) falling on horizontal plane, ambient temperature, *AT* (in ºC), wind speed, *WS* (in m/s), module temperature, *MT* (in ºC) and total AC power output (in kW) that have been collected from the 42kW$_p$ grid-PV system mounted on the roof of Quadrangle Building, University of New South Wales, Australia. The data patterns obtained are based on 15-minute interval.

In general, three two-variate ANN models have been developed based on different types of input configuration in order to predict the total AC power output from the grid-PV system. However, all models have a single output which is the total AC power in kilowatts. Apart form that, *SR* forms the core type of input in each ANN model since the PV system is heavily dependent on the amount of sunlight falling onto the PV modules.

A. ANN Model 1

The first ANN model utilizes *SR* and *AT* as its inputs and total AC power as its output. Since the ambient temperature could also cool down the PV modules and hence influences the performance of PV arrays, *AT* is selected to be the second input to the ANN besides *SR*. The proposed model is illustrated in Fig. 1.

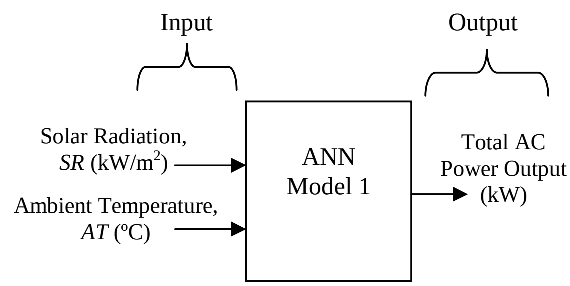

Fig. 1. ANN Model 1 based on *SR* and *AT* as inputs

B. ANN Model 2

The second model utilizes *SR* and *WS* as its input and total AC power as its output. As flowing of wind could significantly cool down the PV modules and hence influences the performance of solar cells inside the modules, *WS* is chosen to be one of the input in the second ANN model. The second ANN model is illustrated in Fig. 2.

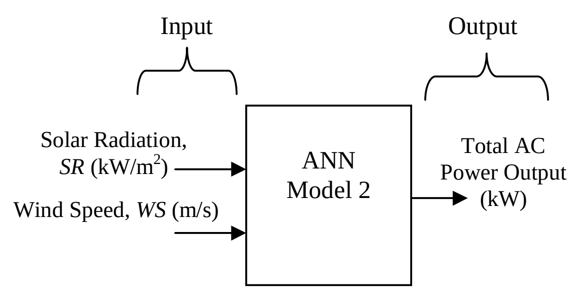

Fig. 2. ANN Model 2 based on *SR* and *WS* as inputs

C. ANN Model 3

The third model utilizes *SR* and *MT* as its input and total AC power as its output.

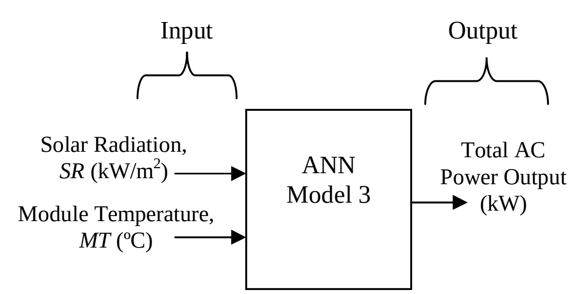

Fig. 3. ANN Model 3 based on *SR* and *MT* as inputs

Since the module temperature directly affects the performance of PV modules, *MT* has become one of the inputs in the third model. The third model is illustrated in Fig. 3.

III. DEVELOPMENT OF TRAINING AND TESTING PROGRAM

The neural network program is written using MATLAB software package (Version R2006a). The training and testing are implemented in a single program so that the performance from each process can be easily monitored and compared. Besides that, data filtering is introduced to improve the regression coefficient and root mean square error, *RMSE* performance in training process. This new feature could improve the performance of ANN by discarding poorly related input and output data during training. In addition, number of poor data to be eliminated could be changed depending on the assigned maximum absolute error performance allowed from each data pattern during training. In this study, the absolute error performance obtained should be less than 2kW. This value is initially chosen based on the maximum of 5% error in the prediction of output power from the 42kW grid-PV system. The absolute error, *E* and root mean square error, *RMSE* are computed as

$$E = |a - t| \qquad (1)$$

$$RMSE = \sqrt{\frac{\sum (a - t)^2}{n}} \qquad (2)$$

where *a* is the actual output of the training data and *t* is the target output of the training data. In addition, *n* is the number of data patterns for training.

The training process is performed using a three-layer feedforward ANN trained by Levenberg-Marquardt backpropagation algorithm, *trainlm*. With supervised learning capability, the Levenberg-Marquardt algorithm has the combination of the advantage found in gradient descent backpropagation technique and the advantage of Newton method. As it only operates with sum squared error function, it is popularly utilized in many prediction tasks [11]. Therefore, Levenberg-Marquardt training algorithm is selected in this study due to its proven track record in many predicting tasks based on regression model [9].

Before the training algorithm is applied, the data set is normalized such that all input and output parameter values lie within a range from -1 to 1. The normalization of a particular data value is calculated using the formula

$$x_{norm} = (Max_{norm} - Min_{norm}) \times \left[\frac{x_{actual} - D_{\min}}{D_{\max} - D_{\min}}\right] + Min_{norm} \qquad (3)$$

where x_{norm} is the normalized data value and x_{actual} is the actual data value to be normalized. D_{min} is the minimum actual data value while D_{max} is the maximum actual data value. In addition, Max_{norm} is the predetermined maximum normalized data value which is equal to 1 whereas Min_{norm} is the predetermined minimum normalized data value which is equal to -1.

On the other hand, the testing process is conducted consecutively after obtaining the best trained ANN model. The overall procedure for the prediction program using ANN can be summarized as follows:

i. Start training process by loading training data.
ii. Adjust ANN architecture and training parameters.
iii. Perform training process.
iv. If the training converges, proceed to the next step. Otherwise, return to step ii.
v. Determine absolute error, E_{abs} for each training pattern, regression coefficient, *R* and Root Mean Square Error, *RMSE*.
vi. If *R* of training is greater or equal to 0.99, proceed to the next step. If not, return to step ii.
vii. If the prediction produces patterns with E_{abs} greater or equal to 2kW, proceed to the next step. Otherwise, go to step ix.
viii. Perform data filtering. Remove data patterns which produce E_{abs} greater or equal to 2kW. Then, return to step iii.
ix. Save trained ANN.
x. Start testing process by loading testing data and recalling the trained ANN in training process.
xi. Perform testing process using the successfully trained ANN.
xii. If the testing converges, proceed to the next step. Otherwise, return to step ii.
xiii. Determine the regression coefficient, *R* and Root Mean Square Error, *RMSE* for testing.
xiv. If *R* is greater than or equal to 0.99, the testing process is stopped. Otherwise, return to step ii.

After determining the best ANN architecture and training parameters for each ANN model, the best prediction model for this study is chosen based on the model that produces lowest *RMSE* and highest *R* during training and testing.

IV. RESULTS AND DISCUSSIONS

The results of this work can be categorized into three sections. The first section describes the best architecture and training parameters for all the three different models, while the second section illustrates the prediction performance of each model. The final section eventually reveals the best model for the prediction of total AC power output from a grid-connected PV system.

After extensive investigation, the best ANN architecture and optimum training parameters for each model are illustrated in Table I. In Table I, all the three models obtain similar best configuration of transfer function where purely linear transfer function has been found to be the most suitable transfer function in each layer. Despite having the same configuration of transfer function, further investigation shows that the three models require different neural configuration. Model 1 and Model 2 have neural configuration of (3,5,1) and (2,3,1) respectively. On the other hand, Model 2 has a configuration of (2,9,1). Therefore, Model 2 can be implemented using a smaller neural configuration compared to Model 1 and Model 3 with a 5 number of nodes in total. In contrast, Model 3 has the highest neural configuration with total of 11 nodes. The number of nodes signifies the complexity of an ANN if it is going to be realized as hardware.

TABLE I: ANN ARCHITECTURE AND TRAINING PARAMETERS OF TWO-VARIATE MODELS

Parameters / Results	ANN Model 1	ANN Model 2	ANN Model 3
Types of elements in input layer	*SR* & *AT*	*SR* & *WS*	*SR* & *MT*
Number of training patterns before filtering	1000	1000	1000
Number of training patterns after filtering	838	832	838
Number of testing patterns	1000	1000	1000
Number of neurons in first hidden layer	3	2	2
Number of neurons in second hidden layer	5	3	9
Number of elements in output layer	1	1	1
Type of transfer function	purelin-purelin-purelin	purelin-purelin-purelin	purelin-purelin-purelin

TABLE II: REGRESSION PERFORMANCE OF TWO-VARIATE MODELS

Performance Type	ANN Model 1	ANN Model 2	ANN Model 3
R before filtering for training	0.99288	0.99266	0.99271
R after filtering for training	0.99859	0.99852	0.99846
% Difference of *R* for training	0.58	0.59	0.58
R for testing	0.99362	0.99350	0.99332

In Table II, during training process, Model 2 experienced the highest improvement of 0.59% in *R* value from 0.99266 to 0.99852 after filtering. Apart from that, after filtering process, Model 1 has experienced an increase of 0.58% of *R* value from 0.99288 to 0.99859 while Model 3 has received an improvement of 0.58% in *R* value from 0.99271 to 0.99846. However, Model 1 has the highest *R* value before and after filtering during training process. The model yields *R* value of 0.99288 before filtering and 0.99859 after filtering. Similarly, in testing process, the highest *R* value of 0.99362 is achieved in Model 1. In short, Model 1 exhibits the best regression performance among the three models despite having a slightly lower percentage of regression improvement after filtering compared to other models.

In terms of *RMSE* performance as illustrated in Table III, besides producing the lowest *RMSE* before filtering, Model 1 obtains the highest reduction of *RMSE* after filtering. The *RMSE* is reduced from 1.2722kW to 0.5708kW after filtering. On the other hand, Model 2 and Model 3 exhibit slightly lower *RMSE* reduction of 54.69% and 54.13% respectively. The *RMSE* of Model 2 is reduced from 1.2919kW to 0.5853kW whereas *RMSE* of Model 3 is reduced from 1.2873kW to 0.5972kW. Although Model 2 has the highest *RMSE* value before filtering, the *RMSE* value has been significantly reduced after filtering. Instead, Model 3 exhibits the highest *RMSE* value after filtering. In testing process, the lowest *RMSE* is still achieved in Model 1 with a value of 1.3085kW. The highest *RMSE* of 1.4206kW is attained by Model 3.

TABLE III: ROOT MEAN SQUARE ERROR PERFORMANCE OF TWO-VARIATE MODELS

Performance Type [in kW]	ANN Model 1	ANN Model 2	ANN Model 3
RMSE before filtering for training	1.2722	1.2919	1.2873
RMSE after filtering for training	0.5708	0.5853	0.5972
% Difference of *RMSE* for training	55.13	54.69	53.61
RMSE for testing	1.3085	1.3768	1.4206

The prediction results for each model in testing process are depicted in Fig. 4 to Fig. 6. All three models shows satisfactory prediction with most actual forecasted output data match the targeted data.

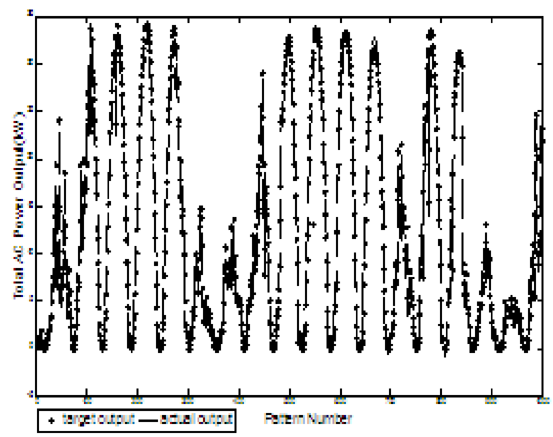

Fig. 4. Model 1- Results of prediction for testing

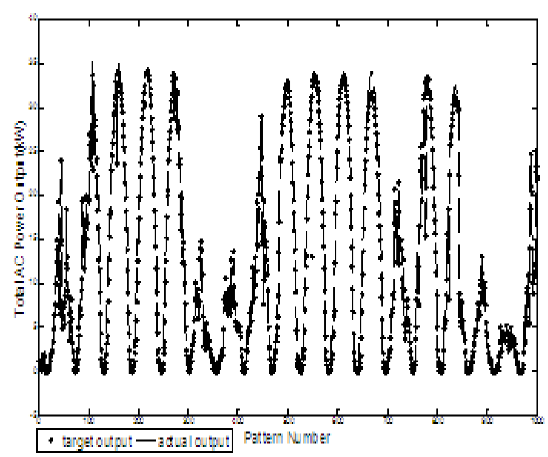

Fig. 5. Model 2- Results of prediction for testing

After the correlation coefficient and *RMSE* performance are considered, it is obviously found that Model 1 produces the most accurate ANN for predicting the total AC power output from grid-connected PV system. The second best model is Model 2 where the *R* and *RMSE* performance are significantly improved after data filtering process. The worst performing model is Model 3 as the model yields lowest *R* value in training and testing. It also produces highest *RMSE* during training and testing. In short, ANN Model 1 which utilizes *SR* and *AT* as its input is found to be the best performing two-variate ANN models in this study.

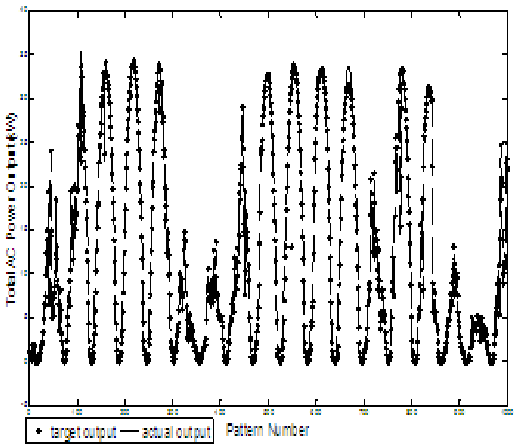

Fig. 6. Model 3- Results of prediction for testing

V. Conclusion

In this study, two-variate ANN models have been successfully developed for predicting total AC power output from a grid-PV system. The ambient temperature has found to be the most important factor in affecting the performance of the grid-connected PV system besides the solar radiation. Apart from that, while regression coefficient and root mean square error are validated to be useful criteria in measuring predicting performance, embedded data filtering employed in these ANN models has actually increased the predicting performance. Besides that, this study also proves that training and testing can be conducted consecutively such that training parameter selection and performance comparison become simpler and quicker. In conclusion, an ANN based technique has been successfully developed to satisfactorily predict the total AC power output from a grid-PV system.

VI. ACKNOWLEDGEMENT

The authors would like to thank the School of Photovoltaic Engineering, University of New South Wales, Australia for making the training data and testing data available for study.

VII. REFERENCES

[1] H. Asano, K. Yajima, and Y. Kaya, “Influence of Photovoltaic Power Generation on Required Capacity for Load Frequency Control,” *IEEE Trans. on Energy Conversion*, vol. 11, no. 1, pp.188-193, 1996

[2] S.R. Wenham, M.A. Green, and M.E. Watt. *Applied Photovoltaics*. Centre for Photovoltaic Devices and Systems, Sydney: The University of New South Wales, 1995, ch. 1.

[3] S. Roberts, *Solar Electricity*, Hertfordshire: Prentice Hall International (UK), 1991, p.25.

[4] M. Balzani and A. Reatti, “Neural Network Based Model of a PV Array for the Optimum Performance of PV System”, in *Proc. 2005 PhD Research in Microelectronics and Electronics Conf.*, vol. 2, pp. 123-126.

[5] D.C. Park, R.J. El-Sharkawi, and R.J. Marks II, “Electric Load Forecasting Using an Artificial Neural Network,” *IEEE Trans. on Power Syst*., 6(2), pp. 442-449, 1999.

[6] G.E. Nasr, E.A. Badr, and M.R. Younes, “Neural Network in Forecasting Electrical Energy Consumption: Univariate and Multivariate Approaches,” *Int. Journal of Energy Research*, vol. 26, pp. 67-68, 2002.

[7] A.S.S. Dorvlo, J.A. Jervase, and A. Al-Lawati, “Solar Radiation Estimation Using Artificial Neural Networks,” *Applied Energy*, vol. 71, pp. 307-319, 2002.

[8] A. Mellit, M. Benghanem, A. Hadj Arab, and A. Guessoum, “Modelling of Sizing the Photovoltaic System Parameters Using Artificial Neural Network”, in *Proc. IEEE Control Applications Conf*., vol. 1, pp.353-357, 2003.

[9] I. Ashraf and A. Chandra, “Artificial Neural Network Based Models for Forecasting Electricity Generation of Grid Connected Solar PV Power Plant”, *Int. Journal of Global Energy Issues*, vol. 21, no. 1/2, pp. 119-130, 2004.

[10] S. Kumar, *Neural Networks A Classroom Approach*, Singapore: McGraw-Hill, 2005, p. 61.K.L Priddy and P.E. Keller, *Artificial Neural Networks: An Introduction*, New Delhi: Prentice Hall of India, 2007, p. 121.

VIII. BIBLIOGRAPHIES

Shahril Irwan Sulaiman received his B.Eng in Electrical & Electronics Engineering from Universiti Tenaga Nasional (UNITEN), Malaysia in 2002 and M.Eng.Sc in Photovoltaic Engineering from University of New South Wales, Australia in 2003. Since 2004, he has been a lecturer in Department of Electronics, Faculty of Electrical Engineering, Universiti Teknologi MARA (UiTM), Malaysia. His research interests are in the areas of photovoltaics (PV), renewable energy technologies and artificial intelligence. He is currently working towards his PhD degree at Universiti Teknologi MARA (UiTM), Malaysia.

Dr. Titik Khawa Abdul Rahman received BSc E.E. (Hons) and PhD on 1988 from Loughborough University of Technology and University of Malaya, Malaysia on 1996 respectively. She is currently a professor at the Faculty of Electrical Engineering and currently the Head of Quality Academic Assurance, Academic Affair Division of Universiti Teknologi MARA Malaysia. She has written more 80 technical papers and also supervising post-graduate students. Her research interest includes voltage profile studies, artificial neural network, Evolutionary Computation such as Evolutionary Programming and Genetic Algorithm, Artifical Immune System (AIS), Economic Dispatch and Loss Minimization. She is currently the secretary for the Power and Energy Society (PES) Chapter, IEEE Malaysia Section.

Dr. Ismail Musirin obtained Diploma of Electrical Power Engineering in 1987, Bachelor of Electrical Engineering (Hons) in 1990; both from Universiti Teknologi Malaysia, MSc in Pulsed Power Technology in 1992 from University of Strathclyde, United Kingdom and PhD in Electrical Engineering from Universiti Teknologi MARA (UiTM), Malaysia in 2005. He holds the Head, Electrical Power Engineering Department, Faculty of Electrical Engineering (UiTM). He has authored and co-authored more than 100 technical papers in the international and national conferences and also in the international journals. He is reviewer for IEEE, IET and WSEAS journals and conferences. His research interest includes power system stability, optimization techniques, distributed generator and artificial intelligence.

Dr. Sulaiman Shaari received his PhD, MSc and BSc from DeMontfort University, United Kingdom, Univ of Missouri, United States of America and Kansas State Univ, United States of America, respectively. He teaches at the Faculty of Applied Science, Universiti Teknologi MARA Malaysia as an Associate Professor. He also heads the Solar Energy Applications Research Laboratory at the university. His main research interests are in the design, sizing, monitoring and evaluation of photovoltaic systems.