Power Generation Forecasting of Dual-Axis Solar Tracked PV System Based on Averaging and Simple Weighting Ensemble Neural Networks

Budiman P.A. Rohman[1], Catur Hilman[2], E. Tridianto[3], T.H. Ariwibowo[3]

[1]Research Center for Electronics and Telecommunication, LIPI, Indonesia
[2]Department of Engineering Physics, ITS, Indonesia
[3]Department of Power Plant Technology, PENS, Indonesia
Corresponding email: budi028@lipi.go.id, hilmanadritya@gmail.com

Abstract

Solar power is a renewable energy interest many researchers around the world to be explored for human life beneficial especially for electric power generation. Photovoltaic (PV) is one of technology developed massively to exploit the solar power for this purpose. However, its performance is very sensitive to environmental condition such as solar irradiance, weather, and climatic behavior. Thus, the hybrid power generation systems are developed to solve this output uncertainty problem. To support this such hybrid system, this paper proposes an ensemble neural network based forecaster of the power output of PV systems which will lead an efficient power management. The object of this research is the PV systems equipped with two axes automated solar tracking with peak power 10Wp. The proposed ensemble forecaster model employs four multi-layer perceptron neural networks with two hidden layers as base forecasters while the input number of historical data is varied in order to exploit the forecaster diversity. The final prediction is calculated both by conventional averaging and simple weighting optimized by the least square fitting technique. According to the research results, the both proposed approaches provide low error rate. Moreover, in term of comparison, the ensemble model with averaging combining technique gives the highest accuracy comparing to the other ensemble and conventional neural network structures.

Keywords: photovoltaic, power output, solar tracking, forecasting, ensemble neural network.

1. INTRODUCTION

Solar power is considered as a prospective renewable energy for supporting largely to the human life. Its availability is covered in huge amount especially in tropical countries. Among many technologies, photovoltaic (PV) technology has taken massively attention of researchers to exploit this solar power. Moreover, the research and development of this technology have been increased significantly in these decades [1]. Its

flexibility and fast progression make this technology became very popular. Nevertheless, this power generation system has very high uncertainty power output because that it is very sensitive to many surrounding aspects such as weather, solar irradiance and climate behavior [2]. The forecasting of the output power of PV systems has been proposed to solve this uncertainty regarding for making the better power management, especially for large hybrid systems both on-grid or off-grid power plants.

Many researchers concern about this issue and they separately proposed some method as follows. Yona et al developed a neural network for predicting the PV power output using the meteorological data in 2007. They compare three different types of neural network: feed-forward neural network, radial basis function neural network (RBFNN) and recurrent neural network (RNN). The experimental results reveal the superiority of RNN as one day ahead forecasting [2]. Then, in 2012, Shi et al proposed the prediction technique to predict the PV system power output based on both weather forecasting data and actual historical power output data using support vector machine [3]. Simultaneously, Su et al proposed the PV power output prediction method using a Gaussian equation which is validated by the measured data of PV systems installed in Macau [4]. Two years later, Li et al, developed ARMAX model based forecaster of PV system power output which is implemented in the on grid-connected PV systems. This study shows that the proposed model outperforms to the ARIMA and RBFNN algorithm proposed by previous research [5]. Yang et al developed a prediction technique which consists of three sequential steps: classification, training, and forecasting. In the classification step, they implement self-organizing mapping and learning vector quantization algorithm, then in the training stage, they use support vector regression and for the last stage, they implement the fuzzy logic [6]. In 2015, Dolora et al, developed the physical hybrid neural network based short-term forecasting of PV system power output [7]. Then, Larson et al developed a forecasting method for day-ahead power output PV systems based on least-squares optimization of numerical weather prediction in 2016 [8].

Meanwhile, the concept of ensemble structure of classifier became very popular and this model has been considered as highly reliable and accurate model facing the inconsistency performance of the classifier, the variance of data and presence of high-level noise. Moreover, the ensemble structure proves its superiority to the single neural network [9-11]. The research studying and exploring the ensemble models are explained as follows. In 2004, the ensemble model of various neural network types was proposed for building a robust weather forecasting. This model consists of multi-layer perceptron neural network (MLPNN), Elman recurrent neural network, RBFNN and Hopfield Model [12]. Then in 2012, the weighting on local learning and diversity in ensemble structure has been investigated[13]. One year later, the ensemble with hierarchical fusion and ten-fold cross-validation were proposed for digital mammogram classification. This

proposed method had a higher performance than the observed single neural network and Adaboost algorithm [14]. Barrow and Crone proposed the use of cross-validation data splitting for model averaging and assessed different forms of cross-validation for creating model diversity [15]. Then, the optimization of ensemble neural network model by using particle swarm optimization integrated Fuzzy Type 1 and 2 for time series prediction was also conducted [16].

This paper extends the previous research mentioned above in order to improve the accuracy and precision of forecasting of PV power output focusing on the PV systems equipped with dual axis automated solar tracking. This study will be more challenging than the fixed PV systems studied in most previous research since instead of the environmental parameters, the solar tracking employed on the systems will also affect significantly to the power output variance as the black box systems. This research result can be a reliable main reference to make a quick response on the power management of solar tracked PV systems especially if the system is integrated with the big hybrid power generation plant scheme. The method proposed in this research is the ensemble model of a neural network with a variety of input and hidden neuron number. The neural network structure is selected to be employed as the basis predictor in this proposed method besides of other artificial intelligent algorithms such as fuzzy logic or Bayesian classifiers. This method is able to predict without any specific prior knowledge especially in the chaotic time series data.

This paper is an extended version of our paper presented in [17]. We extend our previous work by optimizing the selection of base forecaster and reconfiguring the ensemble structure including the combining method comparison. Then we also evaluate more the performance evaluation criteria accompanied by graphical and several statistical analysis.

The rest of this paper is organized as follows: section 2 describes dual-axis solar tracking PV systems studied in this research while section 3 explains the originality of the proposed ensemble forecaster technique including its base forecaster structure. The methodology of this research is described in section 4. Results and discussion are given in section 5, followed by the conclusion of all conducted studies in section 6.

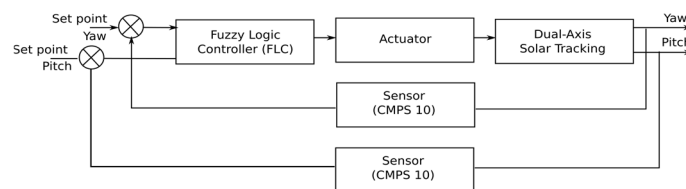## 2. DUAL –AXIS SOLAR TRACKED PV SYSTEM

The system employed in this research is PV system equipped with the dual-axis automated solar tracking system which has been developed and proposed in our previous research [18] showed in Fig. 1. This studied PV system consists of solar cell panel 10Wp, controller box, and 2-axis mechanical parts as an actuator showed by Fig. 1(a). The solar tracking system uses the light dependent resistor and CMPS10 module sensor for tracking the solar position for achieving the highest generated electric power. The Interval Type-2 Fuzzy logic based controller (FLC) is employed in this system as the decision algorithm of the angle of PV depends on the solar

direction and position including azimuth, elevation, and zenith angle. This system uses data acquisition systems to sampling the power output and other variables from the solar panel. The time sampling of measurement is each every five seconds.

The control system of solar tracking works as follows (see Figure1(b)). The solar position is measured by a light-dependent sensor which is fed to FLC as input. This controller computes and sends the control signal to the actuator. The actuator moves the solar panel follow the control signal in both yaw and pitch axis. The current position of solar panel is measured by the CMPS 10 sensor which then fed to the set point for measuring the output error as control feedback.

(a)   (b)

Figure 1. Dual-axis solar tracked PV systems [18]: (a) Prototype (b) Block diagram of control systems

According to our experimental results, this system can result in power output up to 7.7 Watt at the peak time. Besides, with the highest performed configuration (Interval Fuzzy Type-2 with FOU=±0.3), this prototype can achieve the error steady state no more than 1.6% in both pitch and yaw angle regarding to the solar position.

3. PROPOSED ENSEMBLE NEURAL NETWORK
3.1 Multilayer Perceptron Neural Network (MLPNN)

In this study, the ensemble model is built by combining multiple neural networks with MLPNN structure which is very popular and has successfully solved the very complex and diverse problems. MLPNN is a type of neural network which consists of some components as other neural network types including input, weight, layer, neuron, output, and activation function (See Figure2). The layer is a set of neurons. A neuron is a place for multiplying the weight and input then running the activation function. The weight is the multiplier value of input in the neural network.
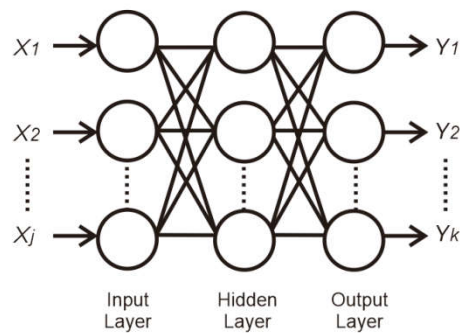
Figure 2. The common structure of MLPNN

MLPNN works by supervised learning error backpropagation algorithm. Simply, the data will be fed into input layer which will be multiplied by weight in each layer before it is summed and processed by activation function with other nodes in the same layer. This process is continued layer by layer until the output layer. The error of the result will be computed by subtracting the output with the desired value. This error will propagate backward to adjust the weight in accordance with the error correction-rule. Generally, an MLPNN structure consists of more than one hidden layer. Considering that φ is the activation function, the mathematical equation of output of a node in this neural network structure,

$$v_k = \varphi\left(\sum_{j=1}^{p} w_{kj}x_j\right) \tag{1}$$

where x_j is input, w_{kj} is the weight for j-th input in k-th lthe ayer, v_k is the output of a node while y_k is the output of a neuron. During the training, the weight is adjusted which minimize the following error function,

$$e = \sum_{k=1}^{r} \sum_{i=1}^{N} (y \quad y_d) \tag{2}$$

where y is output, y_d is the desired value while r is a number of network and N is number of observed data.

3.2 Proposed Ensemble Neural Network

The ensemble structure is a model combining some base classifiers for minimizing the error of predicted output. By employing this structure, the combined classifiers can enhance the accuracy of output by taking the advantage the diversity result of each base classifier. Some combining methods can be taken such as averaging, weighted averaging and majority voting [19-20]. The averaging method is very popular because of its simplicity and this research will also implement this method.
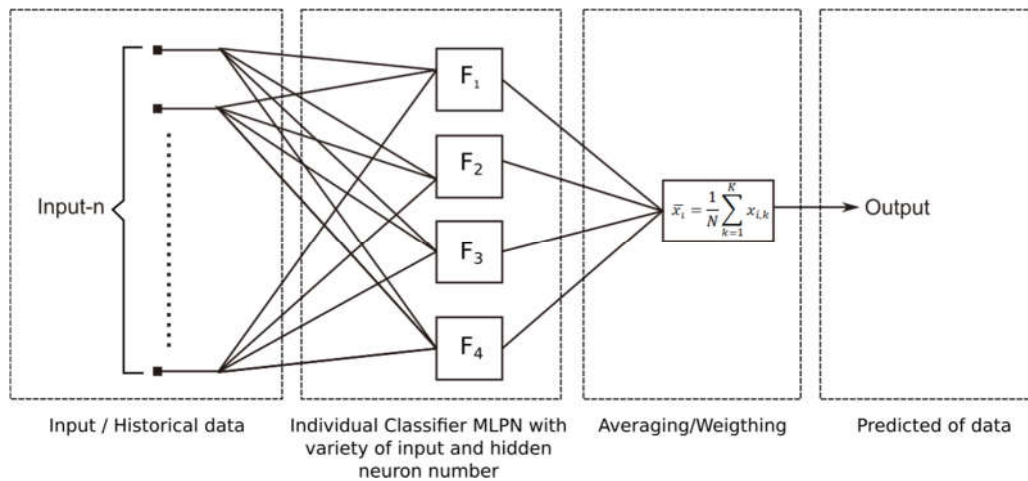
Figure 3. Proposed ensemble neural network with averaging technique

The proposed model of ensemble neural network in this study is shown in Figure 3. The structure consists of four base forecasters which the results of all base forecasters are then computed to take the average or weighted value. This value is then used as the final predicted result. In this scheme, the final forecasting result is five seconds ahead of PV power output in Watt.

The structure of base forecaster employed is MLPNN with two hidden layers. The input of this structure is the normalized power output of PV system with unit Watt. The number of input is varied which will be selected based on the performance of the neural network investigated which represent the historical data of power output of PV systems. Similar to the output of the ensemble model, the output of each base forecaster represents the five seconds ahead of predicted PV power output. The hidden neuron number in the first layer is varied on each base forecaster depend on the number of input (such as 4 neurons when the input number is 4) while the hidden neuron number in the second layer is fixed as many as three neurons.

The combining method in this study will take two type including averaging and weighting. The averaging method is taken as the equation below,

$$\bar{x}_\iota = \frac{1}{N} \sum_{k=1}^{K} x_{i,k} \tag{3}$$

where $x_{i,k}$ is the output of k-th base forecaster while K is the total number of bas forecasters. Then, the weighting method is calculated as below,

$$\bar{x}_\iota = \sum_{k=1}^{K} \alpha_k x_{i,k} \tag{4}$$

where α_k is the k-th weight to be multiplied with the output of k-th forecaster. The weight will be calculated using the least squared method with the objective function F as below,

$$F = \min_{\alpha} \sum_{i=1}^{N} \left(\overline{x_i} \quad \sum_{k=1}^{K} \alpha_k x_{i,k} \right) \tag{5}$$

4. RESEARCH METHODOLOGY

To evaluate the performance of the proposed method we use two days experimental data collected in same weather season, dry, (3 and 7 June 2016) for both training and evaluation phases. This data was collected in Surabaya, Indonesia with coordinate 7o18'44.7804" S and 112o42'53.5716" E in time range around 07.30 WIB to 16.30 WIB, while the data number obtained, is more than 5000 each day with sampling data 5 seconds. Before feeding to the forecasting algorithm, the power output is pre-processed by normalizing it with the global maximum value of data within 15.99 Watt.

The MLPNN parameters used in all of the base forecasters are configured as follows; activation function: tan-sigmoid; learning rate: 0.75; maximum epoch: 5000; and learning method: Levendberg-Marquadt which is recognized as fast and reliable training algorithm for achieving the convergence [21]. The trained of all base MLPNNs build the ensemble structure which then will be tested and evaluated its performance.

To comprehensively evaluate the performance of the proposed method, we take both graphical and several statistical analysis. The graphical analysis is conducted by investigatingvisually the comparison between the forecasted data and actual values including its absolute error. Then, several statistical criteria including Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) also Mean Average Deviation (MAD) and Mean Absolute Percentage Error (MAPE) which is introduced in [22-23]. All of those criteria can be defined mathematically as Eq.(6), Eq.(7), Eq.(8) and Eq.(9) respectively,

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (\overline{x_i} \quad x_i)^2 \tag{6}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\overline{x_i} \quad x_i)^2} \tag{7}$$

$$MAD = \frac{1}{N} \sum_{i=1}^{N} |\overline{x_i} \quad x_i| \tag{8}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{\overline{x_i} \quad x_i}{x_i} \right| \times 100\% \tag{9}$$

where N is the total number of observed data, x_i is i-th desired value while $\overline{x_i}$ is the i-th forecasted value.

5. RESULT AND DISCUSSION
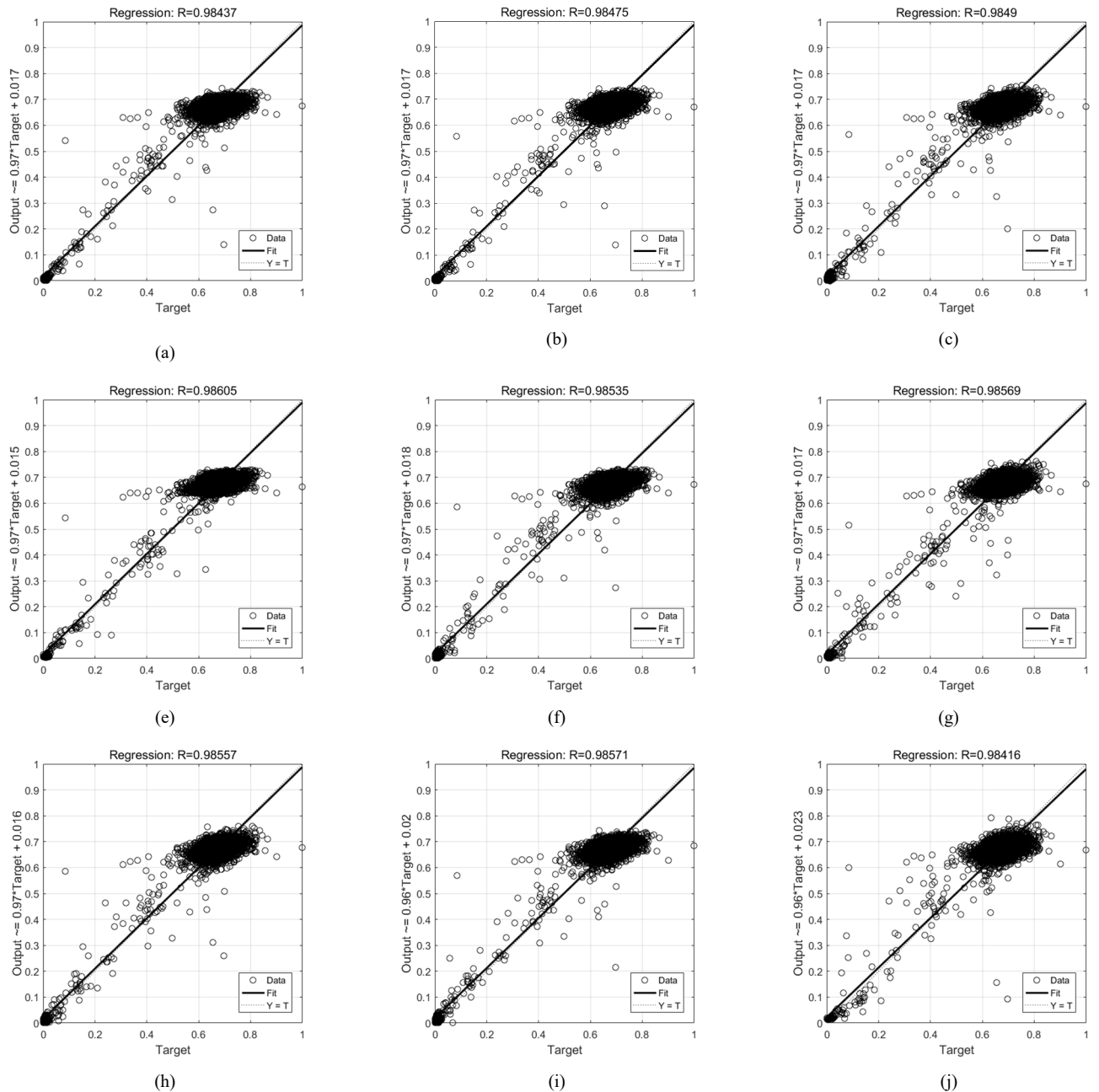## 5.1 Selection of Base Forecasters

Figure 4. Comparison of nine possible base forecaster with model and MSE asfollows: (a) MLPNN2, 0.0019; (b) MLPNN3, 0.0018; (c) MLPNN4, 0.0018; (d) MLPNN5, 0.0017; (e) MLPNN6, 0.0017; (f) MLPNN7, 0.0017; (g) MLPNN8, 0.0017; (h) MLPNN9, 0.0017; (i) MLPNN10, 0.0019

Figure4 shows graphically the regression of training result including the MSE value achieved by nine possible base forecasters which will be selected to build an ensemble structure. The selection will be taken stochastically by

considering those both results. Therefore, considering the lowest MSE and the scattering level of regression of the training result, the four base forecasters are selected with input number: 5, 6, 9, 10 hereinafter these structures are called as MLPNN5, MLPNN6, MLPNN9, and MLPNN10, respectively.

5.2 Performance Evaluation

The simplest way of combining the base classifier in the ensemble structure is taking the mean value so that we investigate this model. Besides, we also develop the weighting combining technique based on the least square method and the selected weights are showed in Table 1 where $W1$, $W2$, $W3$, $W4$ are the weight for the output of MLPNN5, MLPNN6, MLPNN9, and MLPNN10, respectively.

Table 1. The weight of each base forecasters

W1	W2	W3	W4
0.6854	-0.2577	0.6239	-0.0517

Then, Figure5 shows the result of forecasting of the proposed method (both averaging and weighting) and the measured data in the first 150 observed data. It can be seen clearly that the proposed method of averaging technique can predict accurately enough to the PV power output tough some different values in some periods show up. For example, in range time 1 to 40, the method can predict the power almost perfectly. However, in time sampling around 50-60, the predicted data is higher enough than the measured data. The weighting based ensemble also able to predict the value but there are some periods which have very high error values such as in the time sampling around 90 to 100.
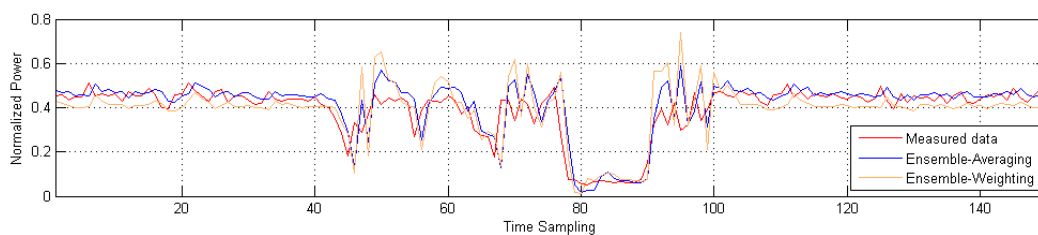

Figure 5. Result of prediction of proposed ensemble neural network

To investigate further about the advantage of ensemble structure to the single MLPNNs, the comparison of predicted power in specific cases was performed which result in Figure 6. According to Fig. 6(a), both proposed methods have the nearest result to the measured data among the observed models. The MLPNN5 which uses 5 input historical data shows the worst results. It means that the selected number of input data in general historical time series based forecasting especially in this case will affect its performance significantly. Then, in the stable case happens in sampling around 1-20, the averaging method can forecast nearly with the measured data while the other forecaster also does same though has a bigger error

(Figure6(b)). In the fluctuation case (sampling 40-55) there are some performance differences among observed forecasters as shown in Figure 6(c). It is seen that the proposed method face difficulties to predict the power output especially in the sampling 47 where another forecaster performs the same. It shows that the proposed method is not reliable enough in the very fluctuate data with significant changes. However, in the case of sharp decrease and climbing (Figure 6(d) and Figure 6(e)), the proposed method shows its superiority over all of the observed model.

The specific case results mentioned above are confirmed by the comparison of the absolute error shown by Figure 7. According to this figure, the highest error of all forecaster including the proposed method happens in the fluctuation condition/period. Nevertheless, the proposed method provides better result among all methods in that case since it takes the diversity value from the base forecasters.

Table 1 shows the numerical comparisons of the proposed method with its base of forecaster including MSE, RMSE, MAD, and MAPE aspects. According to this result, it can be seen clearly that the ensemble structure with the averaging technique is able to improve the performance of its base forecasters with single MLPNN structures. The proposed ensemble neural network has the lowest value in all studied criteria. Specifically, in MSE value, the ensemble model performs within 0.0016 or error rate around 4%. Comparing with the lowest performed forecaster, MLPNN5, this proposed method can suppress the MSE up to 0.0031 or 5.5%. In other words, the proposed ensemble structure can take optimally the advantages of diversity of each basis MLPNN model results.

6. CONCLUSION

We have described the proposed ensemble neural network based forecasting method for the power output of PV system equipped with the dual-axis automated solar tracker. The ensemble structure consists of four base models which are MLPNN with varied input and hidden neuron number. For combining the results of base classifiers, we use the simple averaging method and least square based weighting technique. According to the simulation results, the most optimum structure is ensemble of 5, 6, 9, 10 input and averaging technique as combining method. Using this structure, the MSE of prediction is around 0.0016 or error rate around 4%. In term of comparison, this method shows its superiority to the single forecasters based on MLPNN which is commonly used for forecasting studies.
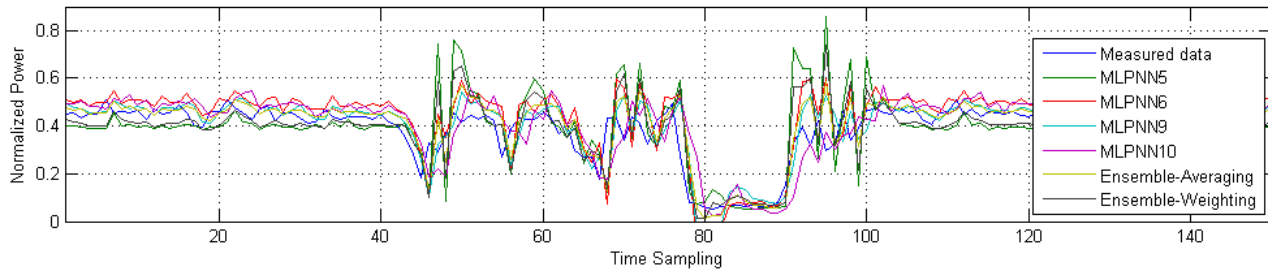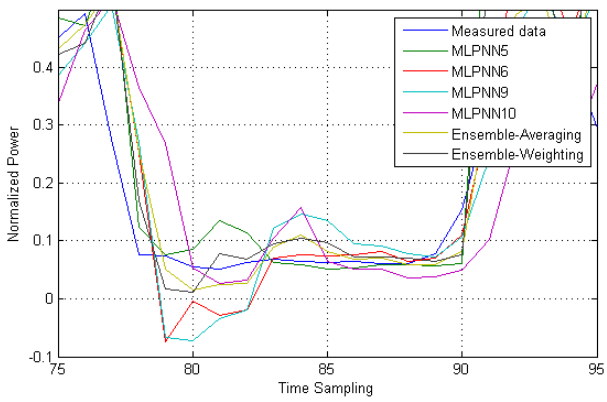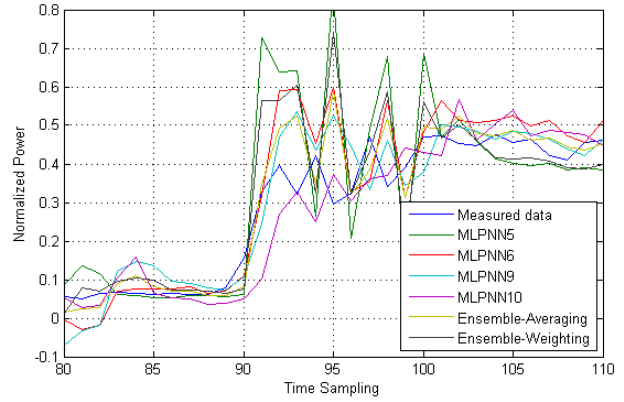
(a)
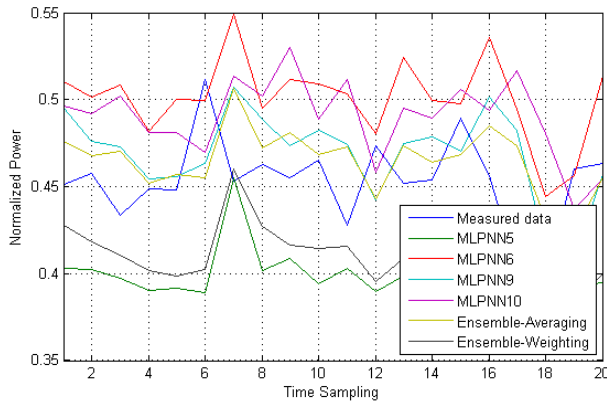
(b)
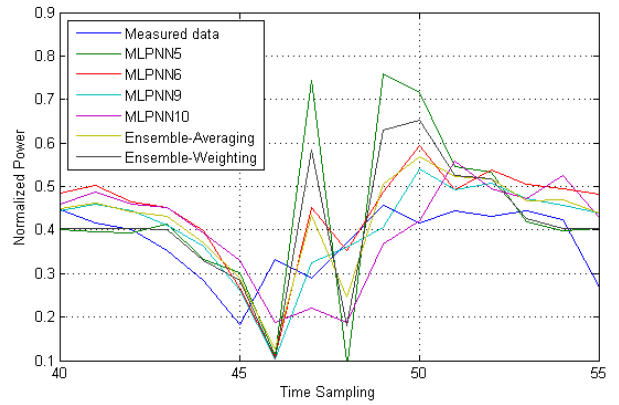
(c)

(d)

(e)

Figure 6. Comparison of proposed ensemble neural network with single MLPNN with specific cases: (a) all cases, (b) stable, (c) fluctuation, (d) after decrease sharply, (e) after climbing sharply
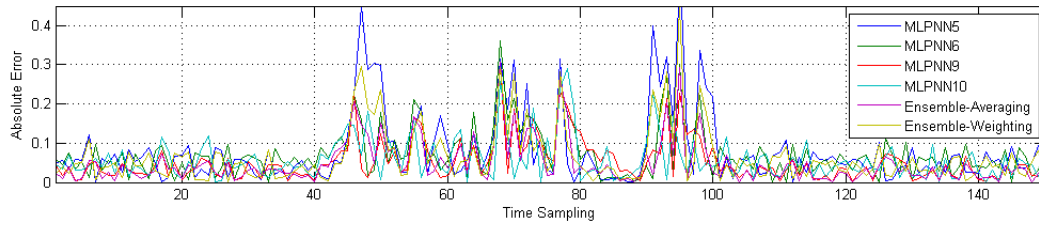
Figure 7. Comparison of absolute error

Table 2. Comparison of statistical performance

Model	MSE	RMSE	MAD	MAPE (%)
MLPNN5	0.0047	0.0686	0.0551	13.4342
MLPNN6	0.0042	0.0648	0.0535	13.7297
MLPNN9	0.0019	0.0434	0.0294	8.1686
MLPNN10	0.0030	0.0543	0.0423	11.2734
Ensemble-Averaging	**0.0016**	**0.0394**	**0.0259**	**7.2250**
Ensemble-Weighting	0.0032	0.0563	0.0439	10.9126

However, according to the investigation of prediction in some specific cases, the proposed method has lacked in the fluctuation data especially if the data changes happen very significantly during the very short period. Tough has very high error, comparing with the single forecaster, the ensemble model with averaging technique is generally provide a better prediction. Thus, the future works can take this drawback as a problem to be solved. There are many options for further development before the real implementation such by employing the optimization algorithm or adaptive corrector value.

REFERENCES

[1] Häberlin, Heinrich, **Photovoltaics system design and practice**,*John Wiley & Sons*, 2012.

[2] Yona, Atsushi, et al., **Application of neural network to one-day-ahead 24 hours generating power forecasting for photovoltaic system,***Intelligent Systems Applications to Power Systems*, 2007. ISAP 2007. International Conference on. IEEE, 2007.

[3] Shi, Jie, et al.,**Forecasting power output of photovoltaic systems based on weather classification and support vector machines**,*IEEE Transactions on Industry Applications* 48.3, pp 1064-1069, 2012.

[4] Su, Yan, et al.,**Real-time prediction models for output power and efficiency of grid-connected solar photovoltaic systems**. *Applied Energy* 93,pp 319-326, 2012.

[5] Li, Yanting, Yan Su, and Lianjie Shu, **An ARMAX model for forecasting the power output of a grid connected photovoltaic system,***Renewable Energy* 66, pp 78-89, 2014.

[6] Yang, Hong-Tzer, et al, **A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output,***IEEE transactions on sustainable energy* 5.3,pp 917-926, 2014.

[7] Dolara, Alberto, et al,**A physical hybrid artificial neural network for short term forecasting of PV plant power output**,*Energies* 8.2,pp 1138-1153, 2015.

[8] Larson, David P., Lukas Nonnenmacher, and Carlos FM Coimbra,**Day-ahead forecasting of solar power output from photovoltaic plants in the American Southwest**,*Renewable Energy* 91, pp 11-20, 2016.

[9] Rokach, Lior,**Ensemble-based classifiers**, *Artificial Intelligence Review*, 33.1-2, pp 1-39, 2010.

[10] L. K. Hansen and P. Salamon, **Neural Network Ensembles**, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 993-1001, 1990.

[11] Rahman, Akhlaqur, and Sumaira Tasnim,**Ensemble Classifiers and Their Applications: A Review**, arXiv preprint arXiv:1404.4088, 2014.

[12] Maqsood, Imran, Muhammad Riaz Khan, and Ajith Abraham, **An ensemble of neural networks for weather forecasting**. Neural Computing and Applications, 13.2, pp 112-122.13, 2004.

[13] Y. Li, S. Gao, and S. Chen, **Ensemble feature weighting based on local learning and diversity**, **26th AAAI Conf. Artificial Intelligence**, pp. 1019–1025. 14, 2012.

[14] Mc Leod, Peter, and Brijesh Verma,**Variable Hidden Neuron Ensemble for Mass Classification in Digital Mammograms**. *IEEE Computational Intelligence Magazine*, 8.1, pp. 68-76. 15, 2013.

[15] Barrow, Devon K., and Sven F. Crone, **Crogging (cross-validation aggregation) for forecasting - A novel algorithm of neural network ensembles on time series subsamples**. *Neural Networks (IJCNN), The 2013 International Joint Conference on. IEEE*, 2013.

[16] Pulido, Martha, and Patricia Melin, **Ensemble Neural Network Optimization Using the Particle Swarm Algorithm with Type-1 and Type-2 Fuzzy Integration for Time Series Prediction**. *Recent Advances on Hybrid Approaches for Designing Intelligent Systems. Springer International Publishing*, 99-112 17, 2014.

[17] Hilman, Catur, E. Tridianto, TH Ariwibowo, Budiman PA Rohman, **Forecasting of Power Output of 2-Axis Solar Tracked PV Systems using Ensemble Neural Network**, *International Electronics Symposium (IES)* IEEE, 2017.

[18] Hilman, Catur and Ali Musyafa, **Rancang Bangun Dual-Axis PV Solar Tracker System MenggunakanInterval Type-2 Fuzzy Logic Controller**, *Seminar Nasional Pascasarjana XIV – ITS*, 2015.

[19] Naftaly, Ury, Nathan Intrator, and David Horn, **Optimal ensemble averaging of neural networks**. *Network: Computation in Neural Systems.* 8.3, pp. 283-296. 12, 1997.

[20] Rahman, Akhlaqur, and Sumaira Tasnim, **Ensemble Classifiers and Their Applications: A Review,** *arXiv preprint arXiv:1404.4088*, 2014.

[21] Hagan, Martin T., and Mohammad B. Menhaj, **Training feedforward networks with the Marquardt algorithm**, *Neural Networks, IEEE Transactions on* 5.6, pp. 989-993, 1994.

[22] Jiang, Yingni, **Prediction of monthly mean daily diffuse solar radiation using artificial neural networks and comparison with other empirical models**, *Energy policy,*Vol.36,n10,pp.3833-3837, 2008.

[23] Simon Haykin, **Neural Networks. A Comprehensive Foundation**, *2ndEdition, Prentice Hall*, 1999.