Comparison of Solar Radiation Intensity Forecasting Using ANFIS and Multiple Linear Regression Methods

Hadi Suyono[1], Rini Nur Hasanah[2], R. A. Setyawan[3], Panca Mudjirahardjo[4], Anthony Wijoyo[5], Ismail Musirin[6]

[1,2,3,4,5]Electrical Engineering Department, Faculty of Engineering, Universitas Brawijaya, Jalan MT. Haryono 167 Malang 65145, Indonesia
[6]Faculty of Electrical Engineering, Universiti Teknologi MARA, 40450, Shah Alam, Selangor, Malaysia

Article Info	ABSTRACT
	Solar radiation forecasting is important in solar energy power plants (SEPPs) development. The electrical energy generated from the sunlight depends on the weather and climate conditions in the area where the SEPPs are installed. The condition of solar irradiation will indirectly affect the electrical grid system into which the SEPPs are injected, i.e. the amount and direction of the power flow, voltage, frequency, and also the dynamic state of the system. Therefore, the prediction of solar radiation condition is very crucial to identify its impact into the system. There are many methods in determining the prediction of solar radiation, either by mathematical approach or by heuristic approach such as artificial intelligent method. This paper analyzes the comparison of two methods, Adaptive Neuro Fuzzy Inference (ANFIS) method, which belongs into the heuristic methods, and Multiple Linear Regression (MLP) method, which uses a mathematical approach. The performance of both methods is measured using the root mean square error (RMSE) and the mean absolute error (MAE) values. The data of the Swiss Basel city from Meteoblue are used to test the performance of the two methods being compared. The data are divided into four cases, being classified as the training data and the data used as predictions. The solar radiation prediction using the ANFIS method indicates the results which are closer to the real measurement results, being compared to the the use MLP method. The average values of RMSE and MAE achieved are 123.27 W/m^2 and 90.91 W/m^2 using the ANFIS method, being compared to 138.70 W/m^2 and 101.56 W/m^2 respectively using the MLP method. The ANFIS method gives better prediction performance of 12.51% for RMSE and 11.71% for MAE with respect to the use of the MLP method.

Corresponding Author:

Hadi Suyono,
Electrical Engineering Department, Faculty of Engineering, Universitas Brawijaya
Jalan MT. Haryono 167 Malang 65145, Indonesia.
Email: hadis@ub.ac.id

1.   INTRODUCTION

Economic growth and development of a country is always accompanied by an increase in electrical energy consumption. Based on the World Energy Agency projection up to 2040 it is known that the world energy demand increased by 30% with an average increase of 1.6% per year [1]. The projected percentage of electricity demand from 2016 to 2030 has increased significantly as shown in the following countries: China 190%, Africa 29%, Europe 24%, and North America 7% [2]. While in Indonesia, electricity demand growth will reach an average of 8.5% per year, with an average peak load growth of 8.4% per year [3]. With an

estimated average customer growth rate of about 2.7 million per year, the electrification ratio may increase to 94.4% by 2024 [4].

To meet the electrical energy need, new generating plants have been constructed, with the composition of power plants still dominated by thermal power plants, while new and renewable energy (NRE) sources compositions are still relatively limited. Therefore, the use of NRE plants still needs to be improved because it has low environmental impact.

Various renewable energy resources have been continuously explored to alleviate the undesirable burden on the environment engendered by the use of fossil energy resources. Solar energy is one form of renewable energy resources which is abundantly provided by the nature and takes an important role to achieve a sustainable development in a country. It generates electricity from the conversion of photon energy brought by the sunlight to the solar panels or photovoltaic (PV) cells. The development of solar energy power plants (SEPPs) as one of NRE sources is rapidly implemented. However, the SEPPs as a source of solar energy has not been widely used in Indonesia, being compared to the condition in advanced countries such as China, USA, Europe, and Australia, where the SEPPs have been massively implemented [5].

The electrical energy generated from the sun depends on the weather and climate in the area where the SEPPs are installed. The solar radiation conditions of the SEPPs will affect the steady-state and dynamic conditions [6], [7] of power systems into which the SEPPs are injected. They influence the number and direction of balanced [8] [9] or unbalanced power flow [10], voltage profile [11], frequency, power quality [12], [13], and the system reliability [14]-[16]. Therefore, the prediction of solar radiation conditions is very important to identify the impact into the power system.

Since the availability of solar sources depends on weather and climate parameters which may change over time, to obtain an optimal power supply system from SEPPs sources it is necessary to predict the availability of solar sources based on previous historical data. There are many methods in determining the prediction of solar radiation either by mathematical approach or by heuristic approach such as artificial intelligent method. Some methods which can be used to predict the availability of solar sources are based on the probabilitic methods [17], network monitoring data [18], artificial neural network and linear regression methods [19], fuzzy logic method [20] [21], mathematical approach using the atmospheric and geometric theory [22], several other approaches [23].

To get a more accurate solar radiation prediction, a combined system of artificial intelligence using the fuzzy and neural networks called Adaptive Neuro Fuzzy Inference System (ANFIS) becomes the main concern in this paper. Adaptive Neuro Fuzzy Inference System (ANFIS) is one of the most commonly used methods for prediction or diagnosis, with fairly good accuracy [24]. The ANFIS method itself is a composite of the fuzzy inference system mechanism introduced in the artificial neural network. The advantage of a fuzzy inference system is that the knowledge of experts can be transformed in the form of rules. However, the establishment of the membership functions may need more extra times [25]. Therefore, with the learning technique improvement of the artificial neural network by reducing the time processing and seaching, the ANFIS method is introduced in this paper to predict the intensity of solar radiation.

To determine the performance of the ANFIS method in predicting the intensity of solar radiation, the Multiple Linear Regression (MLR) method is used as a comparison. The performance of both methods is measured based on the root mean square error (RMSE) and mean absolute error (MAE) values obtained on each test performed. The testing data used in this paper is based on the Swiss Basel City weather parameters taken from the Meteoblue website which provides high quality local weather information worldwide [26]. Many climate parameters can be accessed from Meteoblue, for many places around the world such as temperature, duration of sunlight, wind speed, rainfall data, humadity, solar radiation, and so on.

In this paper, the parameters used to predict the solar radiation are temperature (ºc), humidity (%), precipitation (mm), and sunshine duration (minutes).

2.   ANFIS FORECASTING METHOD

There are several steps which must be done to determine the forecast of solar radiation intensity using ANFIS and MLR methods. In general, the research process conducted by both methods is given in Figure 1. The parameters such as temperature (ºC), humidity (%), precipitation (mm), and sunshine duration (minutes) and solar radiation need to be prepared in advance with specific of time duration. The design and construction of a solution system with ANFIS and MLR are further performed to calculate the desired solar radiation prediction. Existing data need to be grouped into two parts, the first part as training data (on ANFIS method) or input data (on MLR method) and the other part as the testing data, which are the actual data as a comparison of the results obtained from both methods.

```
                        ┌─────────────┐
                        │    START    │
                        └─────────────┘
                               │
                               ▼
              ┌────────────────────────────────────┐
              │  Data Preparation: Temperature,     │
              │  Sunshine Duration, Solar Radiation  │
              │  of Bassel Data                      │
              └────────────────────────────────────┘
                        │              │
                        ▼              ▼
           ┌──────────────────┐  ┌──────────────────┐
           │ Model Design and │  │ Model Design and │
           │ Development using│  │ Development using│
           │     ANFIS        │  │      MLP         │
           └──────────────────┘  └──────────────────┘
                   │                     │
                   ▼                     ▼
           ┌──────────────────┐  ┌──────────────────┐
           │ Forecasting of   │  │ Forecasting of   │
           │ the Intensity of │  │ the Intensity of │
           │ Solar Radiation  │  │ Solar Radiation  │
           │ using ANFIS      │  │ using MLP        │
           └──────────────────┘  └──────────────────┘
                   │                     │
                   └─────────┬───────────┘
                             ▼
                   ┌──────────────────┐
                   │   Comparison     │
                   │ Results: RMSE and│
                   │       MAE        │
                   └──────────────────┘
                             │
                             ▼
                     ┌─────────────┐
                     │     END     │
                     └─────────────┘
```
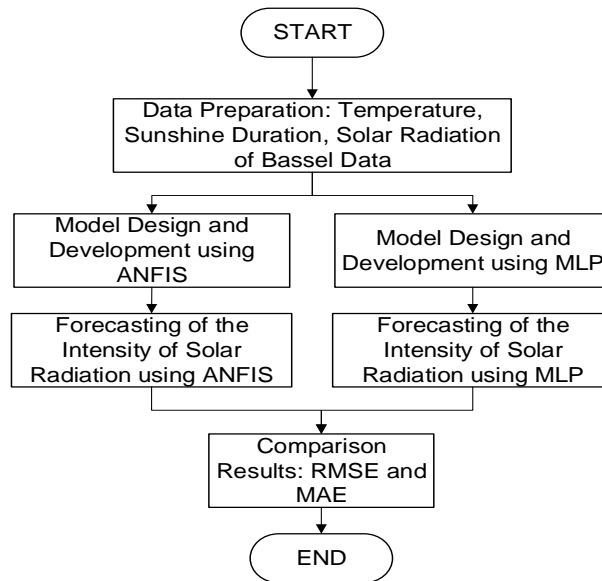
Figure 1. Research methodology

The root mean square error (RMSE) and mean absolute error (MAE) values are then calculated for each test performed using both methods. In each calculation with the standard error model (ε) given, and the error obtained at each iteration (ei, where i=1, 2, ..., n) then the RMSE and MAE calculations are as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2} \tag{1}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|e_i| \tag{2}$$

Furthermore both RMSE and MAE values are compared for both ANFIS and MLR methods. For a sample size of n≥100 data, the RMSE calculation results indicate that the error distribution is close to "truth" or "right solution", with a standard deviation of 5% against the truth. If more samples are used, the error distribution using RMSE will be more reliable [27].

The calculation procedure using ANFIS is given in the flowchart as shown in Figure 2. In general, this method is divided into several stages: the grouping of the existing data into training and testing data, forming ANFIS network structure, determining the type of membership function to be used, generating fuzzy inference system, determining learning methods, performing ANFIS training, and comparing the forecasting results to the actual data.

The different membership functions have been implemented such as Gaussian, Generalized bell, Gaussian combination, Pi-shaped, and Trapezoidal. Data are grouped into two parts: training data (as input) and the remaining data as result data which will be used as comparative data (actual data) from the calculation result obtained.
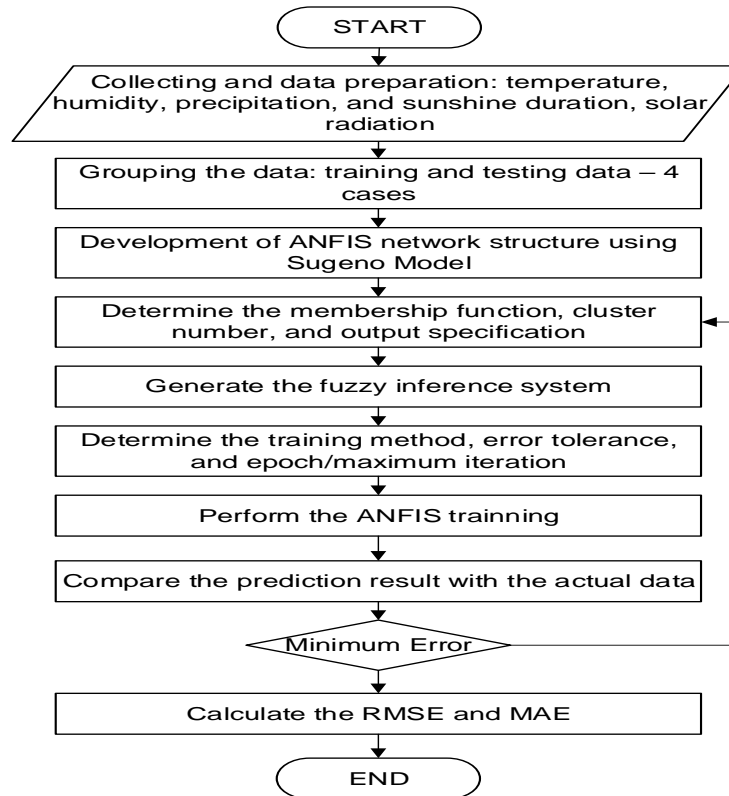
Figure 2. Forecast of solar radiation intensity using the ANFIS method

3. RESULTS AND DISCUSSIONS
3.1. Meteoblue Climatology Data System

For long-term forecasting, the data from an open source weather model, Meteoblue, with the regional nonhydrostatic weather forecasting model based on a model of the US National Weather Service NOAA, have been used for validation. It covers the data of solar radiation intensity, temperature, humidity, rainfall, and duration of solar irradiation. The data used to measure the performance of both ANFIS and MLR methods is data obtained from the NASA Meteoblue Climatology website, i.e. Basel City, Switzerland [28].

The climatology data of Basel City - Switzerland are quite comprehensive with the amount of data as much as 43800 hours with the time period from January 2012 to March 2018. The use of this large data aims to test the performance of both methods used, i.e. ANFIS and MLR, with more independent variables.

To determine the best performance of ANFIS method, the different membership functions such as Gaussian, Generalized bell, Gaussian combination, Pi-shaped, and Trapezoidal have been implemented. Each membership function is tested and compared the result based on the RMSE and MAE values. In this analysis, the Gaussian combination membership function (GCMF) is selected to be used on the overall testing since the GCMF has given the best result of RMSE and MAE values.

There are four cases for the test to be performed as given in Table 1. Training data and testing data can be divided by the composition of training data as much as X% of total data and testing data as much as (100-X)% of the total data. The first three cases used data per hour over five years with 70%, 80%, and 90% for training data, whereas the fourth case is for short-term prediction, i.e. 24 hours prediction, with one year data for the training. The results for the above four cases by comparison of the two methods used are given in the following.

Table 1. Case Study and Data Composition for Training and Testing

Case #	Data Bassel	Data Composition	Training Data	Testing Data
1	43800	70% - 30%	30660	13140
2	43800	80% - 20%	35040	8760
3	43800	90% - 10%	39420	4380
4	8640	99.72% - 0.28%	8616	24

3.2. Case #1: Data Composition of 70% - 30%

This testing aims to compare the results of solar radiation forecasting obtained using the ANFIS and MLR methods. Based on the comparison of both methods, the smallest error value (RMSE and MAE) are calculated. The composition data used in Case #1 is 70% of the total data (30660 data) as the training data while 30% of the total data (13140 data) being used as the testing data.

Figure 3 shows the comparison between the predicted solar radiation using ANFIS method (red line), MLR method (blue line), and actual data (green line) for case # 1. The RMSE and MAE calculations obtained were 139.34 W/m^2 and 101.63 W/m^2 for the ANFIS method, while 150.55 W/m^2 and 114.40 W/m^2 for the MLR method. The ANFIS method gives results that are closer to the actual value than the MLR method.
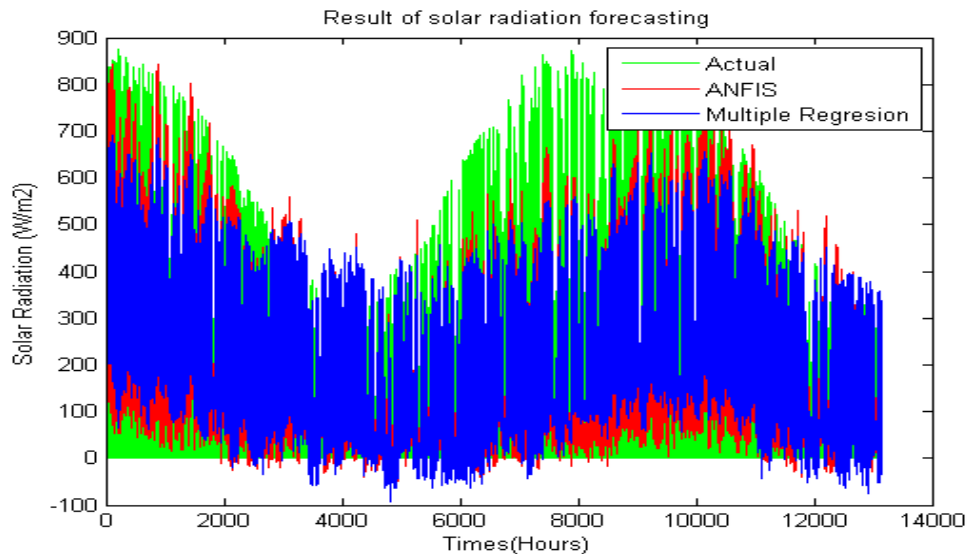
Figure 3. Comparison result ANFIS and MLR for Case #1: 70%-30%

3.3. Case #2: Data Composition of 80% - 20%

The composition data used in Case #2 is 80% of the total data (35040 data) as the training data while 20% of the total data (8760 data) being used as testing data. Figure 4 shows the comparison between the predicted solar radiation using the ANFIS method (red line), MLR method (blue line), and the actual data (green line) for Case #2. The RMSE and MAE calculations obtained were 138.35 W/m^2 and 98.46 W/m^2 for the ANFIS method, while 148.27 W/m^2 and 108.65 W/m^2 for the MLR method. The ANFIS method gives results that are closer to the actual value than the MLR method.
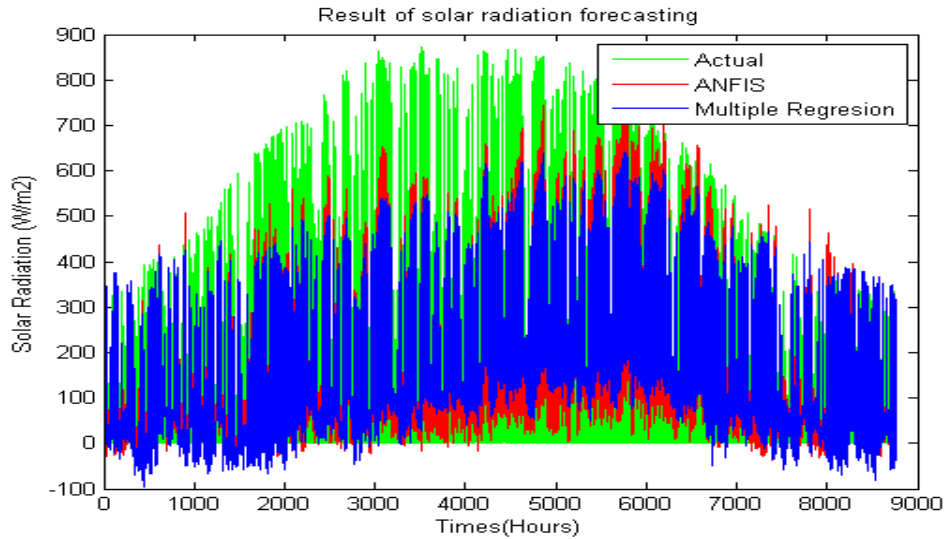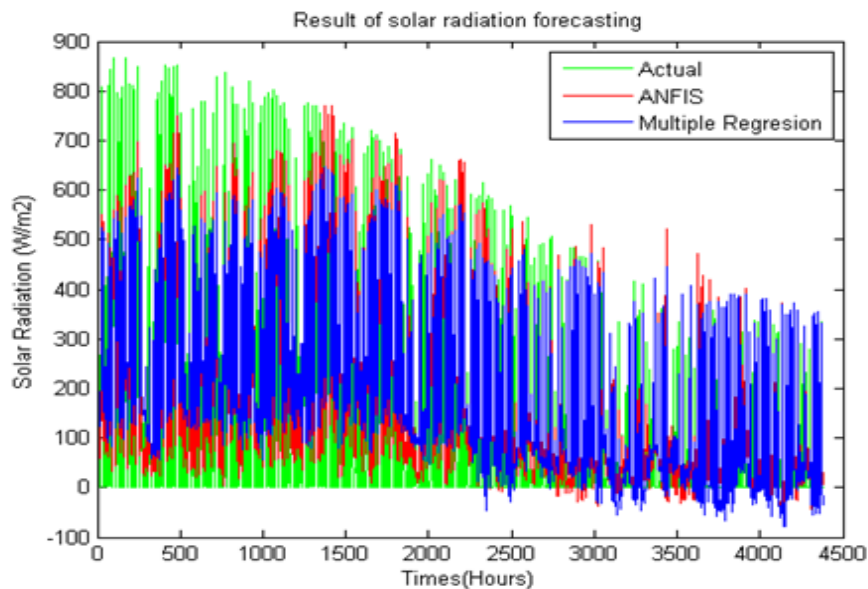
Figure 4. Comparison result ANFIS and MLR for Case #2: 80%-20%

3.4. Case #3: Data Composition of 90% - 10%

The composition data used in Case #3 is 90% of the total data (39420 data) as training data while 10% of the total data (4380 data) being used as the testing data. Figure 4 shows the comparison between the predicted solar radiation using the ANFIS method (red line), MLR method (blue line), and the actual data (green line) for Case #3. The RMSE and MAE calculations obtained were 138.35 W/m^2 and 98.46 W/m^2 for the ANFIS method, while 148.27 W/m^2 and 108.65 W/m^2 for the MLR method. The ANFIS method gives results that are closer to the actual value than the MLR method.

Figure 1. Comparison result ANFIS and MLR for Case #3: 90%-10%

3.5. Case #4: Short Term (24 hours) Prediction

In this testing, the short-term forecasting is done within one day (24 hours) with input/training data for one year. The predicted RMSE and MAE results obtained are 81.22 W/m^2 and 63.67 W/m^2 for the ANFIS method, while 108.95 W/m^2 and 70.13 W/m^2 for the MLR method. The MLR method gives results that are

closer to the actual value than the ANFIS method. However, ANFIS provides a more realistic value than MLR for the calculation between 16:00 to 17:00 where the sunlight starts to disappear.

A summary comparison of the RMSE and MAE error values for the ANFIS and MLR methods is shown in Table 2. The average RMSE and MAE are 131.68 W/m^2 and 95.60 W/m^2 for the ANFIS method, while 141.18 W/m^2 and 105.04 W/m^2 for the MLR method. ANFIS showed better prediction performance of 7.22% for RMSE and 9.87% for MAE compared with MLR method. Figure 6 shown comparison result ANFIS and MLR for case #4: short term prediction

Table 2. Comparison of ANFIS and MLR Results

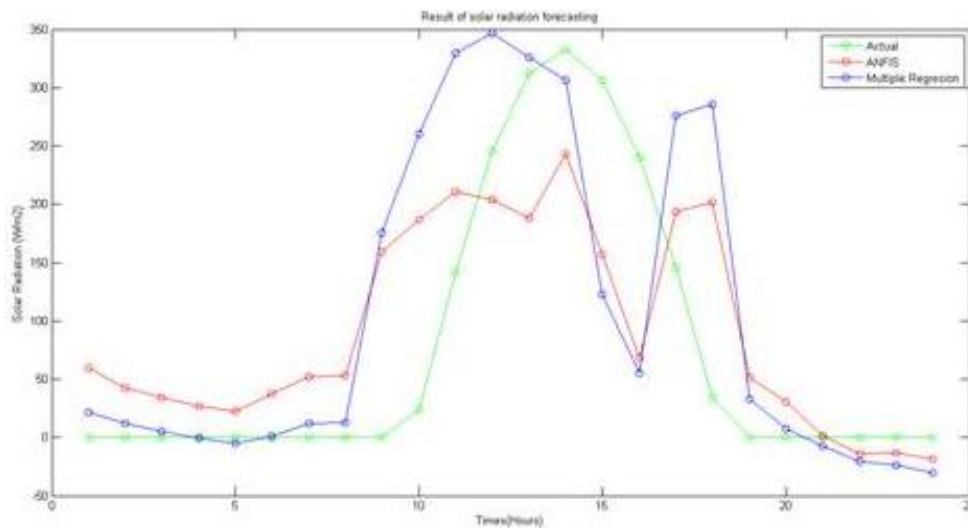| Case #	ANFIS (W/m^2)		MLR (W/m^2)	
	RMSE	MAE	RMSE	MAE
1	139.34	101.63	150.55	114.40
2	138.35	98.46	148.27	108.65
3	134.19	99.91	147.02	113.07
4	81.22	63.67	108.95	70.13
Average	123.27	90.91	138.70	101.56

Figure 6. Comparison result ANFIS and MLR for case #4: short term prediction

4. CONCLUSION

Based on the results obtained from the simulation and analysis undertaken in this paper, it can be concluded that the forecasting of optimal solar radiation intensity was obtained in Case #3 using the ANFIS method with 90% training data and 10% testing data. This optimal results can be seen from the smallest RMSE and MAE values among the other experiments. Based on the simulated type of membership function curve, the membership function with Gaussian combination is the best result.

The use of ANFIS method gives better results than MLR method when implemented with huge data. It can be known from the calculation results which give smaller RMSE and MAE values using the ANFIS method. Contrarily, the results of the calculation using the MLR method give better results for relatively fewer of data, as shown in Case # 4.

The average values of RMSE and MAE were 131.68 W/m^2 and 95.60 W/m^2 using the ANFIS method, whereas using the MLR method they were 141.18 W/m^2 and 105.04 W/m^2 respectively. The ANFIS method showed the better prediction performance of 7.22% for RMSE and 9.87% for MAE with respect to the MLR method.

REFERENCES

[1] International Energy Agency (IEA). World Energy Outlook 2017. IEA Publications. 2017.
[2] European Energy Agency (EEA). 2017.
[3] Secretariat General of National Energy Council of Indonesia (NEC). Indonesia Energy Outlook 2016. 2016.
[4] PLN PT Pembangkitan Jawa Bali. 2015 Annual Report. Jakarta: PLN, 2015.
[5] World Energy Council (WCE). World Energy Resources Solar 2016. 2016
[6] Suyono H, Hasanah RN, Mudjirahardjo P, Edy Purnomo MF. *Steady-State and Dynamic Impact of Hybrid Embedded Generation in Distribution System.* Proceedings of International Seminar on Intelligent Technology and Its Applications (ISITIA) and Regional Conference on Electrical and Electronic Engineering (RCEEE) 2017 (ISITIA – RCEEE 2017). Surabaya. 2017; 1: 142-147.
[7] Islam M, Mithulananthan N, Hossain MJ. Dynamic voltage support by TL-PV systems to mitigate short-term voltage instability in residential DN. *IEEE Transactions on Power Systems.* 2017; (99): 1 – 1.
[8] Thomson M, Infield DG. Network power-flow analysis for a high penetration of distributed generation. *IEEE Transaction on Power System.* 2007; 22(3): 1157–1162.
[9] Widén J, Shepero M, Munkhammar J. Probabilistic load flow for power grids with high PV penetrations using copula-based modeling of spatially correlated solar irradiance. *IEEE Journal of Photovoltaics.* 2017; 7(6): 1740-1745.
[10] Syafii Syafii, K. M. Nor. Unbalanced active distribution analysis with renewable distributed energy resources, *TELKOMNIKA ((Telecommunication Computing Electronics and Control).* 2015; 13(1): 21-31
[11] Katiraei F, Mauch K, Dignard-Bailey L. Integration of photovoltaic power systems in high-penetration clusters for distribution networks and mini-grids. Int. J. Distrib. Energy Res. 2007; 3(3): 207–223.
[12] Hu H, Shi Q, He Z, He J, Gao S. Potential Harmonic Resonance Impacts of PV Inverter Filters on Distribution Systems. *IEEE Transactions on Sustainable Energy.* 2015; 6(1): 151–161.
[13] Amirullah Amirullah, Ontoseno Penangsang, Adi Soeprijanto; Power quality analysis of integration photovoltaic generator to three phase grid under variable solar irradiance level, *TELKOMNIKA (Telecommunication Computing Electronics and Control).* 2016; 14(1): 29-38
[14] Liu N, Wu T, Xu T, Ma Y. Reliability evaluation method for distribution network. *IET Journals & Magazines, the Journal of Engineering.* 2017; 2017(13): 1771–1776.
[15] Argüello, Lara JD, Rojas JD, Valverde G. Impact of rooftop PV integration in distribution systems considering socioeconomic factors. *IEEE Systems Journal.* 2017; 99: 1-12.
[16] Suyono H, Wijono, Hasanah RN, Dhuha S. *Power Distribution System Reliability Improvement due to Injection of Distributed Generation.* Proceedings of 10th International Conference on Electrical and Electronics Engineering (ELECO). 2017(1): 1485–1490.
[17] Khallat MA, Rahman S. A Probabilistic Approach to Photovoltaic Generator Performance Prediction. *IEEE Power Engineering Review.* 1986; PER-6(9): 24–24.
[18] Zhong Z, Tan J, Zhang T, Zhu L. PV power short-term forecasting model based on the data gathered from monitoring network. *IEEE Journals & Magazines China Communications.* 2014; 11(14):61– 69.
[19] Sulaiman SI, Abdul Rahman TK, Musirin I, Shaari S. Artificial neural network versus linear regression for predicting Grid-Connected Photovoltaic system output. Proceedings of IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER). 2012; 170–174.
[20] Mellit A, Arab AH, Khorissi N, Salhi H. An ANFIS-based Forecasting for Solar Radiation Data from Sunshine Duration and Ambient Temperature. *IEEE Power Engineerng Society General Meeting.* 2007; 1–6.
[21] Mellit A, Kalogirou SA. *Neuro-Fuzzy Based Modeling for Photovoltaic Power Supply System.* Proceedings of IEEE International Power and Energy Conference. 2006; 88–93.
[22] Mustapha Elyaqouti, Lahoussine Bouhouch, Ahmed Ihlal. Modelling and predicting of the characteristics of a photovoltaic generator on a horizontal and tilted surface. *International Journal of Electrical and Computer Engineering (IJECE)*, 2016; 6(6): pp. 2557-2576
[23] Mori H, Takahashi M. *A prediction method for photovoltaic power generation with advanced Radial Basis Function Network.* Proceedings of IEEE PES Innovative Smart Grid Technologies. 2012; 1–6.
[24] Suparta W, Alhasa KM. Modeling of Tropospheric Delays Using ANFIS. Springer, 2016.
[25] Mathur N, Gleska I, Buis A. Comparison of adaptive neuro-fuzzy inference system (ANFIS) and Gaussian processes for machine learning (GPML) algorithms for the prediction of skin temperature in lower limb prostheses. *Medical Engineering and Physics.* 2016; 38: 1083–1089.
[26] Makridakis S. Forecasting: Methods and Applications. John Wiley and Sons, 2002.
[27] Chai T, Draxler RR. Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature. *Geoscientific Model Development.* 2014; 7: 1247–1250.
[28] Meteoblue, ”Weather history download Basel”, can be accessed in: https://www.meteoblue.com/en/weather/archive/export/basel_switzerland_2661604