❏   691

Prediction of Solar Radiation Intensity using Extreme Learning Machine

Hadi Suyono[1], Hari Santoso[2], Rini Nur Hasanah[3], Unggul Wibawa[4], Ismail Musirin[5]
[1,2,3,4]Electrical Engineering Department, Faculty of Engineering, Universitas Brawijaya
Jalan MT. Haryono 167 Malang 65145 Indonesia
[5]Faculty of Electrical Engineering, Universiti Teknologi MARA, 40450, Shah Alam, Selangor, Malaysia

Article Info	ABSTRACT
	The generated energy capacity at a solar power plant depends on the availability of solar radiation. In some regions, solar radiation is not always available throughout the day, or even week, depending on the weather and climate in the area. To be able to produce energy optimally throughout the year, the availability of solar radiation needs to be predicted based on the weather and climate behavior data. Many methods have been so far used to predict the availability of solar radiation, either by mathematical approach, statistical probability, or even artificial intelligence-based methods. This paper describes a method of predicting the availability of solar radiation using the Extreme Learning Machine (ELM) method. It is based on the artificial intelligence methods and known to have a good prediction accuracy. To measure the performance of the ELM method, a conventional forecasting method using the Multiple Linear Regression (MLR) method has been used as a comparison. The implementation of both the ELM and MLR methods has been tested using the solar radiation data of the Basel City, Switzerland, which are available to public. Five years of data have been divided into training data and testing data for 6 case-studies considered. Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) have been used as the parameters to measure the prediction results based on the actual data analysis. The results show that the obtained average values of RMSE and MAE by using the ELM method respectively are 122.45 W/m2 and 84.04 W/m2, while using the MLR method they are 141.18 W/m2 and 104.87 W/m2 respectively. It means that the ELM method proved to perform better than the MLR method, giving 15.29% better value of RMSE parameter and 24.79% better value of MAE parameter. *Copyright © 2018 Institute of Advanced Engineering and Science.* *All rights reserved.*

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

The increase in electric energy need has a direct relation with the continuous growth of population and human prosperity. To satisfying the energy need, various conservation and diversification of energy programs have been considered. In a country like Indonesia for example, the government has been launching a program to build 35000 MW of power plants during 2009-2020 [1]. Unfortunately, the electrical energy supply of Indonesia is still dominated by coal-fired power plants, which is about 70% of the total generation to be built [2]. Such this kind of fossil energy source has been known to cause serious environmental impacts because of various chemical substances such as carbondioxide (CO_2), sulfurdioxide (SO_2), sulfurtrioxide

(SO$_3$), nitrogenoxides (NO$_x$), particulate matters (PM), condensable PM, mercury (Hg), metals trace and radioactive being caused by the coal burning [3].

Considering the negative impacts of fossil energy source beside its deposits which are continuously depleting, energy diversification using the new and renewable energy sources become favorable. This kind of energy sources produces less environmental impacts [1]. The generation capacity of these new and renewable energy-based generators in general is still relatively small, being around 0.5MW-50MW for each unit. However, if the generating units number is large, the contribution will be significant [4].

One of the renewable energy generation can be realized thorugh the utilization of solar energy using photovoltaic (PV) cells, which convert the solar thermal energy into electrical energy. The injection of some renewable energy-based power generations like the PV, wind turbines, biomass, microhydro, and other plants into the available grid in a power distribution system, being known as dispersed generators (DG), greatly influence the distribution system [5]. Under steady-state condition, it may affect the voltage profile and power losses [6, 7], the number and direction of power flow [8, 9], MVA fault levels [10, 11], reliability system [12-14], and power quality [15, 16]. Under the dynamic-state condition, among the impacts of the PV-based renewable energy injection into grid system are the stability of the frequency and voltage [5, 17-21] and the small signal stability [22] of the electrical system.

The energy generation capacity of a PV-based power system depends on the solar radiation and the weather and climate conditions of the location where it is installed. The weather and climate conditions depend on the geographical and atmospheric features. The geographical features include the latitude, altitude, seasons, whether it is terrain, and even when during the day time of the location. Indonesia for example, geographically it is located along the equator with the available monthly solar radiation of between 4.6 kWh/m^2 and 7.2 kWh/m^2 giving an average of 5.12 kWh/m^2 throughout the year [24]. The atmospheric features lead to the variation in solar radiation, including pressure, humidity, temperature, dust particles content, clouds, aerosols, and snow covering [23]. Therefore, the prediction of solar radiation conditions is very important to harvest the solar energy as effective as possible.

Various studies have been undertaken to predict the intensity of solar radiation in a particular place. Many methods can be used to perform the prediction. Each technique has its own properties and precision. The incurred cost is even also to consider in choosing a particular method [25]. Some known prediction methods can be categorized into conventional methods, whether some others are using artificial intelligence methods. Being compared to the conventional methods, the artificial intelligence-based methods offer several advantages, such as relatively easy updates and maintenance, incomplete inputs, and reasoning skills [26].

Several prediction methods of solar radiation are based on the artificial neural network methods [24], linear regression [27], probabilistic methods [28], network monitoring data [29], fuzzy method approach [30, 31], ANFIS, Multiple Linear Regressions (MLR) method [32], and some other methods [33]. The conventional method commonly used for prediction is the multiple regression method. This method can be analyzed by using several independent variables so that the obtained results are more accurate [25]. Another artificial method which can be utilized for prediction is the extreme learning machine (ELM) method, which is based on the of artificial intelligence theory. This method has advantages in terms of accuracy, good generalization performance, and fast learning speed [34].

This paper presents a comparison of methods to obtain the best prediction of solar radiation intensity. The prediction algorithm based on the ELM method is to be compared to that based on the MLR method. The comparison performance parameters to be considered are the root mean square error (RMSE) and the mean absolute error (MAE) obtained on each test performed. It is to be emphasized that a proper modeling is required in implementing the ELM method to result in the good prediction results. In order to obtain the optimal results, several types of modelling variations are considered, including the composition variations of training data and testing data, variations in the number of hidden neurons, and the use of more number of variables and longer data ranges.

The test data used are the Basel region weather data of the Swiss country, being obtained from the Meteoblue website [35] which provides high quality local weather information worldwide for everything on land or sea in the world. The parameters used are temperature, duration of daytime and solar radiation.

2. RESEARCH METHOD

The general process of the research in this paper is presented in Figure 1. The steps to undertake the research on the prediction of solar radiation intensity using ELM method and MLR method can be elaborated as follows:
1) Preparing the parameters data such as temperature, duration of sun exposure, and the solar radiation with the specified time duration in accordance with the existing data.

2) Grouping the available data into two parts, partly as the training data (in the ELM method) or the input data (in MLR method) and some others as the testing data used to compare the performance of both methods in solar radiation intensity prediction.
3) Designing and the development of the prediction system using the EM\LM and MLR methods, and implementing them to perform the solar radiation intensity prediction during certain period of time.
4) The error calculations are performed using two methods, i.e. the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE), as given in the following equation [36]:

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

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

where based on the standard error model (ε) given on each calculation, the error obtained at each iteration is e_i, with $i = 1, 2, 3, …n$. Furthermore both the RMSE and MAE values are compared for both the ELM and MLR methods.

The steps to follow in implementing the ELM method for solar radiation intensity prediction is shown in Figure 2. It can be seen that the steps include three main stages, namely the data preprocessing, training, and testing. The data preprocessing is used to divide the whole data into training data and testing data. The training and testing data division can be done to make a composition with the training data vary between 70% -95% and the testing data vary between 30% -5% of total the data.
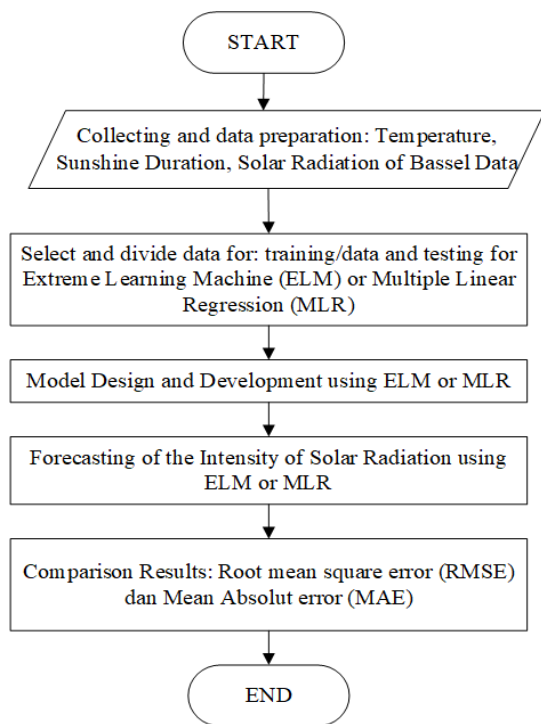

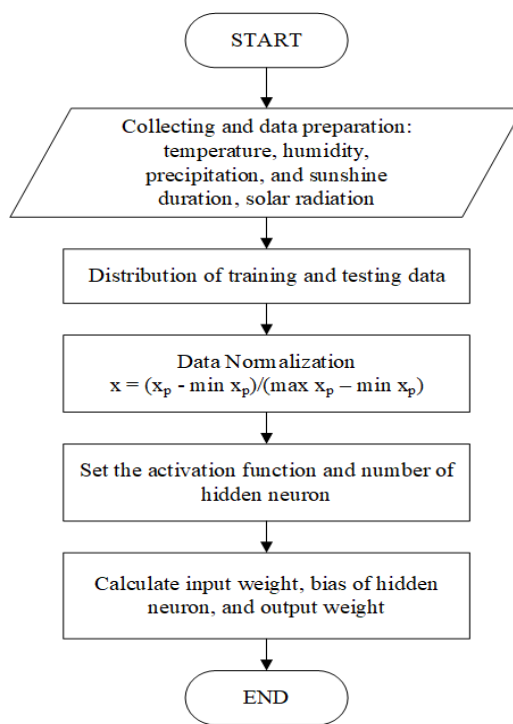Figure 1. The research method in predicting the solar radiation intensity using the ELM method

Figure 2. The training process during the implementation of ELM method

The function of the training process is to develop the model of the ELM method implementation. It is purpose to determine the weights of the input, bias and the output of the implementation system. The results obtained from the weighing process during the training are then implemented to predict the solar radiation intensity using the ELM method. The purpose of this process is to obtain input weight, bias, and output weights. Based on the input weight and the output weights obtained from the training process, the next step is to forecast using ELM (testing). The testing process is used to evaluate the ability of the ELM as a forecasting tool. The flowchart of the testing process in the ELM method is shown in Figure 3.
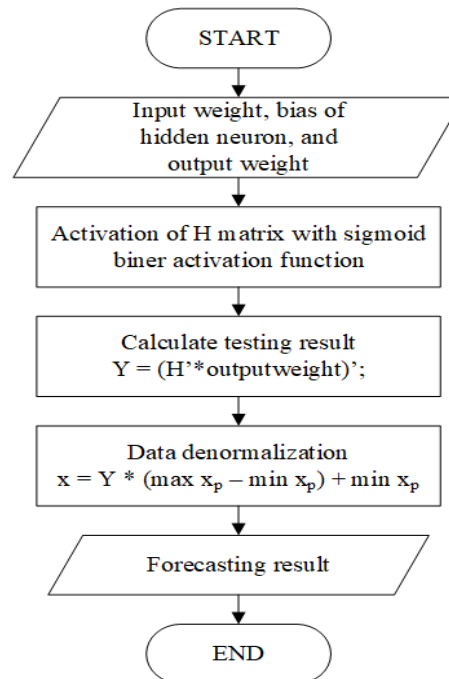
Figure 3. The testing process during the implementation of ELM method

3. RESULTS AND ANALYSIS

3.1. Meteoblue Climatology Data (NOAA)

The data used in this research have been obtained from the NOAA Meteobue Climatology website, namely the data of the Basel City, Switzerland [35]. These data offers longer time span and larger number of independent variables, so that the performance of the ELM method under consideration can be analyzed more comprehensively. The data contain the hourly data of Basel city during the period of January 2012 to March 2018, with a total of 43800 data. The data include the parameters such as the duration of sun radiation, average temperature, humidity, rainfall, and the intensity of solar radiation. Six variations of data composition have been considered during the research, as shown in Table 1. The data composition variations can be formed by the composition of X% of the total data as the training data and (100-X) % of total data as the testing data. The use of MatLab programming has been considered in the ELM method implementation.

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

This experiment aims to compare the prediction results obtained using the ELM method and those using the MLR method. It aims to find the smallest error value between the two methods. In the Study Case #1, the data composition is formed by 70% (30660 data) of training data, and 30% (13140 data) of testing data.

Figure 4 shows the comparison of the prediction results using the ELM and MLR methods to the actual data from NOAA. The RMSE and MAE values generated from the ELM method implementation are $132.239W/m^2$ and $91.569W/m^2$, whereas using the MLR method they are $150.547W/m^2$ and $113.405 W/m^2$ respectively. The prediction results using the ELM method are much closer to the actual data being compared to the results of MLR prediction method.

3.3. Study Case #2: Data Composition 75% - 25%

In the Study Case #2, the data composition is formed by 75% (32850 data) of training data, and 25% (10950 data) of testing data. The comparison results are given in Figure 5. As seen, the RMSE and MAE values generated from the ELM method implementation are $126.254W/m^2$ and $86.345W/m^2$, whereas using the MLR method they are $142.521W/m^2$ and $103.805W/m^2$ respectively. The prediction results using the ELM method are much closer to the actual data being compared to the results of MLR prediction method.

3.4. Study Case #3: Data Composition 80% - 20%

In the Study Case #3, the data composition is formed by 80% (35040 data) of training data, and 20% (8760 data) of testing data. The comparison results are given in Figure 6. As indicated, the RMSE and MAE

values generated from the ELM method implementation are 131.265W/m^2 and 90.637W/m^2, whereas using the MLR method they are 148.272W/m^2 and 108.645W /m^2 respectively. The prediction results using the ELM method are much closer to the actual data being compared to the results of MLR prediction method.

Table 1. The variation of training data and testing data compositions used in the research

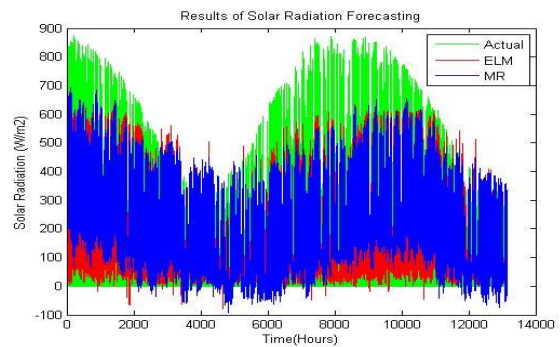| Experiment number	% training data - % testing data	Training data amount	Testing data amount
1	70% - 30%	30660	13140
2	75% - 25%	32850	10950
3	80% - 20%	35040	8760
4	85% - 15%	37230	6570
5	90% - 10%	39420	4380
6	95% - 5%	41610	2190

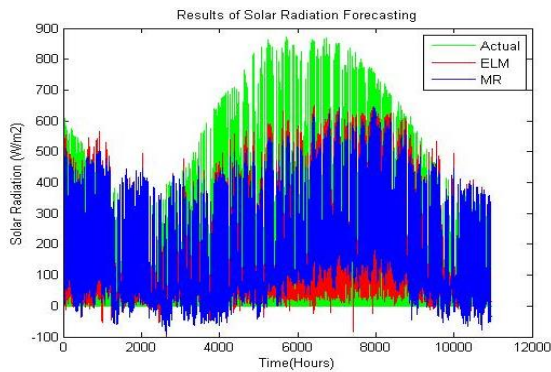Figure 4. The Comparison of ELM and MLR for Case Study #1

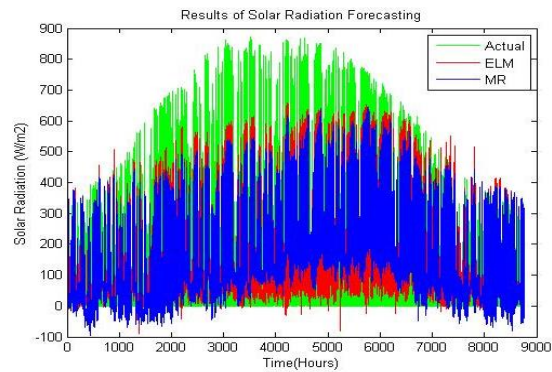Figure 5. The Comparison of ELM and MLR for Case Study #2

Figure 6. The Comparison of ELM and MLR for Case Study #3

3.5. Study Case #4: Data Composition 85% - 15%

In the Study Case #4, the data composition is formed by 85% (37230 data) of training data, and 15% (6570 data) of testing data. The comparison results are given in Figure 7. As seen, the RMSE and MAE values generated from the ELM method implementation are 137.932W/m^2 and 96.272W/m^2, whereas using the MLR method they are 156.741W/m^2 and 119.219W /m^2 respectively. The prediction results using the ELM method are much closer to the actual data being compared to the results of MLR prediction method.

3.6. Study Case #5: Data Composition 90% - 10%

In the Study Case #5, the data composition is formed by 90% (39420 data) of training data, and 10% (4380 data) of testing data. The comparison results are given in Figure 8. As seen, the RMSE and MAE values generated from the ELM method implementation are 121.968W/m^2 and 82.657W/m^2, whereas using the MLR method they are 147.015W/m^2 and 113.072W /m^2 respectively. The prediction results using the ELM method are much closer to the actual data being compared to the results of MLR prediction method.

3.7. Study Case #6: Data Composition 95% - 5%

In the Study Case #6, the data composition is formed by 95% (41610 data) of training data, and 5% (2190 data) of testing data. The comparison results are given in Figure 9. As seen, the RMSE and MAE values generated from the ELM method implementation are 85.064W/m^2 and 56.749W/m^2, whereas using the MLR method they are 101.978W/m^2 and 71.088W /m^2 respectively. The prediction results using the ELM method are much closer to the actual data being compared to the results of MLR prediction method.
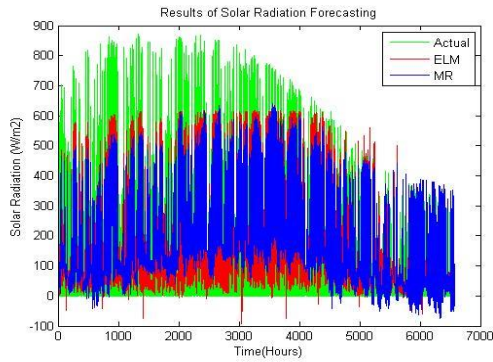
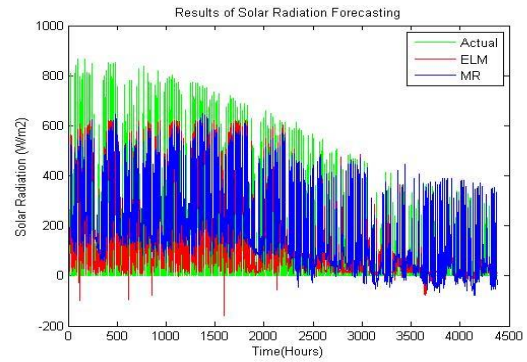Figure 7. The Comparison of ELM and MLR for Case Study #4

Figure 8. The Comparison of ELM and MLR for Case Study #5

3.8. General Comparison of ELM and MLR Method Results

Summary comparison of the RMSE and MAE error calculation results from solar radiation intensity prediction using both the ELM and MLR methods with 6 case studies considered is given in Figure 10. The ELM method implementation resulted in the average RMSE error of 122.45W/m^2 and the average MAE error of 84.04W/m^2, whereas the MLR method yielded an average RMSE error of 141.18W/m^2 and an average MAE error of 104.87 W/m^2. The ELM method showed a better performance than the MLR method did, being indicated with the RMSE and the MAE error values respectively 15.29% and 24.79% better that those of the MLR method.

For all case studies considered, the ELM method gives a smaller error value being compared to the MLR method implementation, based on the error calculations using the RMSE and MAE. In addition, calculations using the RMSE errors tend to result in higher values than using the MAE error parameter.
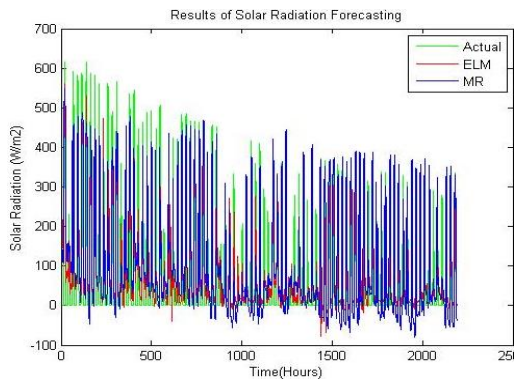

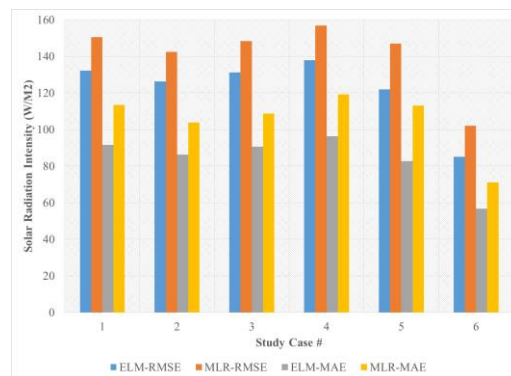Figure 9. The Comparison of ELM and MLR for Case Study #6

Figure 10. The Comparison of ELM and MLR for All Cases Study

4. CONCLUSION

The research on the comparison of solar radiation intensity prediction using the artificial intelligence-base ELM method and the conventional MLR method brings to some conclusions that the optimum data composition giving the smallest RMSE and MAE errors among the whole prediction experiments conducted using the ELM method is formed by 85% training data and 15% testing data.

The implementation of the ELM method on the data of Basel city, Switzerland resulted in the prediction of solar radiation intensity with the smallest RMSE value of 85.064W/m^2 and the smallest MAE value of 56.749W/m^2, while using the MLR method the smallest RMSE value is 101.978W/m^2 and the smallest MAE value of 71.088 W/m^2, which were obtained in Case Study #6 with the composition of 95% training data and 5% testing data.

The ELM method showed a better performance than the MLR method did, being indicated with the RMSE and the MAE error values respectively 15.29% and 24.79% better that those of the MLR method.

REFERENCES

[1]   Secretariat General of National Energy Council of Indonesia (NEC), Indonesia Energy Outlook 2016 (Indonesia Version), ISSN 2527-3000.

[2]   PLN. 2015 Annual Report, PT Pembangkitan Jawa Bali. First Edition, Jakarta, 2015

[3]   Shahzad BK, Yousaf M. Coal Fired Power Plants: Emission Problems and Controlling Techniques. *Journal of Earth Science and Climate Change*. 2017; 8(7): 1-9.

[4]   Jinkens N, Allan R, Cossley P, Kirschen D, Strbac G. Embedded Generation. London: The Institution of Electrical Engineers. 2000.

[5]   Suyono H, Hasanah RN, Mudjirahardjo P, Purnomo MFE. *Steady-State and Dynamic Impact of Hybrid Embedded Generation in Distribution System*. International Seminar on Intelligent Technology and Its Applications (ISITIA) and Regional Conference on Electrical and Electronic Engineering (RCEEE) 2017 (ISITIA – RCEEE 2017). 2017: 142-147.

[6]   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.

[7]   Suyono H, Hasanah RN. Analysis of Power Losses due to Distributed Generation Increase on Distribution System. *Jurnal Teknologi*. 2016; 78(6-3): 23-28.

[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]   Murdoch N, Berry J, Kazerooni A. Distributed generation connections under a fault-level active network management scheme. *CIRED - Open Access Proceedings Journal IET Journals & Magazines*. 2017; 2017(1): 1707–1710.

[11]   Boljevic S, Conlon MF. *Fault current level issues for urban distribution network with high penetration of distributed generation*. IEEE Conferences 2009 6th International Conference on the European Energy Market. 2009: 1–6.

[12]   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.

[13]   Argüello, Lara JD, Rojas JD, Valverde G. Impact of Rooftop PV Integration in Distribution Systems Considering Socioeconomic Factors. *IEEE Systems Journal*. 2017; pp (99): 1–12.

[14]   Suyono H, Wijono, Hasanah RN, Dhuha S. *Power distribution system reliability improvement due to injection of distributed generation*. 2017 IEEE 10th International Conference on Electrical and Electronics Engineering (ELECO). 2017: 1485–1490.

[15]   Esparza M, Segundo J, Núñez C, Wang X, Blaabjerg F. A Comprehensive Design Approach of Power Electronic-Based Distributed Generation Units Focused on Power-Quality Improvement. *IEEE Transactions on Power Delivery*. 2017; 32(2): 942–950.

[16]   Zeng Z, Yang H, Guerrero JM, Zhao R. Multi-functional distributed generation unit for power quality enhancement. *IET Power Electronics*. 2015; 8(3): 467–476.

[17]   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; pp (99): 1–1.

[18]   Glover E, Chang C-C, Gorinevsky D, Lall S. *Frequency stability for distributed generation connected through grid-tie inverter*. 2012 IEEE International Conference on Power System Technology (POWERCON). 2012: 1–6.

[19]   Dong D, Li J, Boroyevich D, Mattavelli, Cvetkovic I, Xue Y. *Frequency behavior and its stability of grid-interface converter in distributed generation systems*. 2012 Twenty-Seventh Annual IEEE Applied Power Electronics Conference and Exposition (APEC). 2012: 1887–1893.

[20]   Angelim JH, Affonso CM. Impact of distributed generation technology and location on power system voltage stability. *IEEE Latin America Transactions*. 2016; 14(4): 1758–1765.

[21]   Karlsson P, Bjornstedt J, Strom M. *Stability of voltage and frequency control in distributed generation based on parallel-connected converters feeding constant power loads*. 2005 European Conference on Power Electronics and Applications. 2005.

[22]   Krismanto AU, Mithulananthan N. Identification of modal interaction and small signal stability in autonomous microgrid operation. *IET Generation, Transmission & Distribution*. 2018; 12(1): 247– 257.

[23]   Rathod APS, Mittal P, Kumar B. *Analysis of factors affecting the solar radiation received by any region*. 2016 International Conference on Emerging Trends in Communication Technologies (ETCT). 2016: 1–4.

[24]   Rumbayan M, Abudureyimu A, Nagasaka K. Mapping of solar energy potential in Indonesia using artificial neural network and geographical information system. *Renewable and Sustainable Energy Reviews*. 2012; 16: 1437–1449.

[25]   Makridakis S, Wheelwright SC, McGee VE. Forecasting, 2nd Edition, Vol. I. Printed V. Translation Jakarta: Erlangga. 1995

[26]   Kusumadewi S. Artificial Intelligence (Technique and Its Application) (Indonesian Version). Yogyakarta: Graha Ilmu. 2003.

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

[28]   Khallat MA, Rahman S. A Probabilistic Approach to Photovoltaic Generator Performance Prediction. *IEEE Power Engineering Review*. 1986; PER-6(9): 24–24.

[29] 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.

[30] Mellit A, Arab AH, Khorissi N, Salhi H. *An ANFIS-based Forecasting for Solar Radiation Data from Sunshine Duration and Ambient Temperature*. IEEE 2007 Power Engineering Society General Meeting. 2007: 1–6.

[31] Mellit, Kalogirou SA. *Neuro-Fuzzy Based Modeling for Photovoltaic Power Supply System*. 2006 IEEE International Power and Energy Conference. 2006: 88–93.

[32] Suyono H, Hasanah RN, Setyawan RA, Mudjirahardjo P, Wijoyo A, Musirin I. Comparison of Solar Radiation Intensity Forecasting Using ANFIS and Multiple Linear Regression Methods. *Bulletin of Electrical Engineering and Informatics*. 2018; 7(2).

[33] Mori H, Takahashi M. *A prediction method for photovoltaic power generation with advanced Radial Basis Function Network*. IEEE PES Innovative Smart Grid Technologies. 2012: 1–6.

[34] Cheng G, Cai L, Pan H. *Comparison of Extreme Learning Machine with Support Vector Regression for Reservoir Permeability Prediction*. IEEE 2009 International Conference on Computational Intelligence and Security. 2009; 2: 173-176.

[35] Meteoblue, "Weather history download Basel", can be accessed in: https://www.meteoblue.com/en/weather/archive/export/basel_switzerland_2661604

[36] 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.