Artificial Neural Network Performance Analysis for Solar Radiation Prediction, Case Study at Baron Techno Park

Ridwan Budi Prasetyo¹², Haidar Rahman¹, Ikrima Alfi¹, Fredi Prima Sakti¹

¹ Universitas Teknologi Yogyakarta, UTY Jl. Siliwangi (Ringroad Utara), Jombor, Sleman, D.I. Yogyakarta 55285
² Badan Riset dan Inovasi Nasional, BRIN Gedung Energi No. 625, Kawasan PUSPIPTEK, Serpong, 15314 Telp. (021) 7560550

fredipsakti@staff.uty.ac.id, haidarrahman1993@gmail.com, ikrima.alfi@uty.ac.id, ridwan.budi@bppt.go.id

Abstract. As stated in Government Regulation No. 79 of 2014 on National Energy Policy (KEN), the New and Renewable Energy (NRE) mix target is at least 23% by 2025. Now the utilization of solar energy in Indonesia has only reached about 0.05% or 100 MW. Government compiles a roadmap for the use of solar energy that targets the installed PV mini-grid capacity until 2025 is 6.5 GW. Technically, before installing a solar power plant, solar potential data is needed for a certain period of time. This is absolutely necessary considering the potential for solar is intermittent. The solar data are then processed to create a model forecasting so that it can optimize the resulting energy output.

Forecasting using artificial intelligence with artificial neural network algorithms is a good solution because it has higher accuracy. To see the comparison of the performance of the ANN of this research with previous research. The finding shows that Baron 1-7-1 performed better than 2P 1-2-1 and Baron 1-2-1, regardless that the RMSE have a slight difference but still Baron 1-7-1 outperformed the others, with the best value of RMSE 0.15185 and R² of 0.88996.

Keywords: Renewable Energy, Solar Energy, Forecasting, Artificial Neural Network, Accuracy.

1. INTRODUCTION

As stated in Government Regulation No. 79 of 2014 on National Energy Policy (KEN), the New and Renewable Energy (NRE) mix target is at least 23% by 2025. Meanwhile, based on the General Plan for the Provision of Electricity (RUPTL) 2019 to 2028, the number of NRE power plants was 7,790 MW (11.6% of the total national installed capacity of 66,868 WM) until 2019, so it still requires additional number of generators of 13,684 WM in 2020 to 2025 to achieve mix target 23% [1], [2].

Geographically, Indonesia lies on the equator, so it has potential National solar energy reaching 4.8 kWh/m²/day or equivalent to 207,898 MW. Now the utilization of solar energy in Indonesia has only reached about 0.05% or 100 MW. Government compiled a roadmap for the use of solar energy that targets the installed PV mini-grid capacity until 2025 is 6.5 GW [1]. On the other hand, due to its location in the equatorial region, this causes the potential for solar energy in Indonesia to be relatively small. In the initial study of the potential solar power in Indonesia, there is a total potential of 950 MW spread across the Java region and Sulawesi.

The strategy that can be pursued is the development of solar power plants equipped with a battery that can supply electricity 24 hours, this is considering solar power plants intermittent which could potentially damage the existing system. The existence of solar power plants equipped with batteries to become an independent electric power system solution in the 3T archipelago in Indonesia.

Technically, before installing a solar power plant, solar potential data is needed for a certain period of time. This is absolutely necessary considering the potential for solar is intermittent. The solar data are then processed to create a model forecasting so that it can optimize the resulting energy output.
Forecasting using artificial intelligence with artificial neural network algorithms is a good solution because it has higher accuracy.

Past research using support vector regression (SVR), to predict the short-term solar radiation at a hybrid PV power plant in Indonesia. The SVR helps improve the denoising capabilities and simplify computation [3]. Another past research was using Hargreaves model to measure and estimate solar radiation in Kelantan, Malaysia. The Hargreaves model is more of a statistical model that can predict solar radiation only using the maximum and minimum air temperature [4]. The suggested aerosol index for one of the inputs in the ANN model for predicting the solar radiation. Aerosol index is the measurement of attenuation of the sunlight or solar radiation due to haze, dust and other suspended particles. Particles such as smoke, dust, cloud cover etc. block the sunlight by absorbing, diffusing or scattering light. AI defines the amount of sunlight that is prevented from reaching the ground by aerosol particles [5]. Lastly, an assessment of ANN learning method was proposed for predicting solar radiation. 10 different training algorithms and 23 different training datasets were used to model the ANN, for a case study in AUS solar PV, Jordan [6]. This research will focus on the prediction performance for the solar radiation at Baron Techno-park. Using the ANN algorithm for predicting future solar radiation for the next 24 hours.

2. METHODOLOGY
2.1. Training Algorithms
There are many training methods to choose when using the ANN, these training methods determine the accuracy and learning rate of the prediction [6]. For this research we are only comparing between two methods:

- Levenberg-Marquardt (LM)
- Scaled Conjugate Gradient (SCG)

2.2. Neural Network Structure
After determining the training algorithm, next is to determine the ANN structure, for it is an important part in having good results for prediction. A typical ANN model consist of three layers, i.e., the input layer (X) which accept the environmental variables, the hidden layers (Z) which takes in the input signals and convert them into output signals, and lastly the output layer (Y) which shows the response of the input layers in regards to the hidden layers [7].

![Figure 1. ANN Structure.](image-url)
2.3. ANN Evaluation

Mean absolute percentage error (MAPE) is used as index prediction error and Root Mean Squared Error (RMSE) [3] [8], is being interpretation of MAPE (%) values and RMSE were dimensionless as follows:

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (I_{predicted,i} - I_{measured,i})^2} \tag{1}
\]

\[
MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{I_{measured,i} - I_{predicted,i}}{I_{measured,i}} \right| \times 100% \tag{2}
\]

\[
R^2 = 1 - \frac{\sum_{i=1}^{n}(I_{measured,i} - I_{predicted,i})^2}{\sum_{i=1}^{n}(I_{measured,i} - \overline{I_{measured}})^2} \tag{3}
\]

Coefficient of the determination (R2) is one of the common statistical performance evaluating methods. R2 presents forecasted values and measured values at x and y axis. If the forecasted values and measured values are closer, R2 will be nearer to 1. Meaning that the coefficient of the determination values has the higher prediction, in our case for solar radiation prediction [9]. From equation (5), (6) and (7) the symbols represent, observed data (n), measured data (M), and predicted data (P).

The ANN simulation is run by using the Matlab© program. The size of the data is pre-process from the year 2019. The data contains 28 variables that were taken every hour by the sensors. The data will then be reduced to 11 variables due to only taking the variables that are relevant to the research. The ANN structure that will be implemented is the Feed-Forward Network as a default setting to the Matlab© ANN simulation.

3. RESULT AND DISCUSSION

3.1. Data Correlation

This research is using data from the months of January until May (5 months). The data are taken from Baron Techno-park which will be analyzed for the medium-term day ahead (24 hours) prediction for solar radiation. The reason the data is from January until May is because during the months of January until April is during the rainy season, and April until May is the transition into the fall season. Therefore, using the ANN algorithm to predict the output of the solar radiation, there are some tuning parameters that need to be set in order to have acceptable prediction results. Such tuning parameters are the training algorithms and the ANN structure. Some references from previous research will also be used as comparison in order for better understanding the performance of the ANN we are using in this research.
The first step is to determine the input and output of the ANN, which will be perceived by having to use such tools as RStudio© to find the correlation between the input and output of the data, which can be seen in figure 2. Also, we are using some references from previous data to compare results, and see which better inputs have better correlation.

As seen in figure 2, the point of interest is to find the correlation of the variable output (shortwave radiation) towards the other variables or the inputs (highlighted in orange). The most significant correlated variables that can be determined are the temperature 2, relative humidity, sunshine duration, UV radiation, diffuse short-wave radiation, and temperature 28 (direct shortwave radiation is not included because it is the same as the output variable). However, additional variables can be added into the list as they have a small value of correlation, such as temperature 3, wind speed 5, long wave radiation, and mean sea level pressure. These variables mentioned before will be the inputs to the ANN in order to predict the output which is the shortwave radiation.

3.2. ANN Evaluation
The second step of this research is to evaluate the ANN performance and see whether or not the prediction that was made by the ANN is acceptable. Therefore, for this experiment we will evaluate using coefficient of the determination ($R^2$), root mean squared error (RMSE), and mean absolute percentage error (MAPE). The first experiment will be using significant correlated inputs and also using three different kinds of training algorithms, while having 5-10 hidden layer nodes. This experiment will help determine the best training algorithm and ANN structure.
Tabel 2. ANN Training Table Results.

| Training Algorithm | ANN Structure | $R^2$ (0-1) | RMSE | MAPE (%) |
|--------------------|---------------|------------|------|----------|
| LM                 | 1-5-1         | 0.99362    | 33.2302 | 19.6930 |
|                    | 1-6-1         | 0.99324    | 34.2034 | 24.0609 |
|                    | 1-7-1         | 0.99334    | 33.9710 | 18.2846 |
|                    | 1-8-1         | 0.99349    | 33.9042 | 24.0487 |
|                    | 1-9-1         | 0.99379    | 32.7822 | 22.8084 |
|                    | 1-10-1        | 0.99350    | 34.0372 | 25.5497 |
| SCG                | 1-5-1         | 0.99102    | 39.4148 | 46.5516 |
|                    | 1-6-1         | 0.99243    | 36.2073 | 35.4605 |
|                    | 1-7-1         | 0.99212    | 36.9858 | 39.2477 |
|                    | 1-8-1         | 0.99049    | 40.5735 | 74.3327 |
|                    | 1-9-1         | 0.99151    | 38.3287 | 27.8087 |
|                    | 1-10-1        | 0.99157    | 38.1818 | 48.7482 |

From the results in table 2 and figure 3, comparing between the two methods of training, it can easily be seen which training algorithm and the neural network structure is more superior. Looking at the $R^2$, it is to look for a high result meaning it has a high correlation between its predictions and moreover, looking for a lower MAPE which tells us a higher accuracy in prediction. However, for the RMSE test run it has a relative value between 30 to 40, meaning that it has no significant impact in determining the best training and structure. Therefore, the results show that the Levenberg-Marquardt training algorithm with the ANN structure of 1-7-1 was a better ANN model.

Figure 3. Graphic Results ANN Training.
Figure 4. Real and Predicted Data Comparison Results.

For figure 4 it shows the results of the ANN and see the comparison in predicting the medium day ahead (24 hours), which will conclude the satisfactory results as the graphic shows only a slight error in comparison between the real data and the predicted data.

Another evaluation for this research will compare with other researches to see whether the method used in this research can perform the same or better with the given references [10]. Note that for each comparison the data will be different but with the same variables, therefore for each experiment determining the same structure model and training methods with the referred research [10] and also comparing the results from table 1 and figure 3.

**Table 3. ANN Model 2P 1-2-1.**

| Model | RMSE  | $R^2$  | No. of Iteration |
|-------|-------|--------|-----------------|
| 2P    | 0.18536 | 0.52223 | 8900            |
|       | 0.18532 | 0.52261 | 20900           |
|       | 0.18514 | 0.52479 | 34100           |
|       | 0.18500 | 0.52494 | 50000           |
|       | 0.18500 | 0.52494 | 78000           |
|       | 0.18500 | 0.52494 | 10000           |

**Table 4. ANN Model Baron 1-2-1.**

| Model  | RMSE  | $R^2$  | No. of Iteration |
|--------|-------|--------|-----------------|
| Baron 1-2-1 | 0.20772 | 0.88672 | 8900            |
|        | 0.19973 | 0.88676 | 20900           |
|        | 0.19611 | 0.88658 | 34100           |
|        | 0.19413 | 0.88661 | 50000           |
|        | 0.16693 | 0.88663 | 78000           |
|        | 0.15377 | 0.88712 | 10000           |
Tabel 5. ANN Model Baron 1-7-1.

| Model   | RMSE   | $R^2$  | No. of Iteration |
|---------|--------|--------|------------------|
| Baron 1-7-1 | 0.20384 | 0.89170 | 8900             |
|          | 0.17754 | 0.89265 | 20900            |
|          | 0.17687 | 0.88983 | 34100            |
|          | 0.17537 | 0.89325 | 50000            |
|          | 0.17508 | 0.88480 | 78000            |
|          | 0.15185 | 0.88996 | 10000            |

Looking at figure 5 and figure 6 which graph from the results of table 3, table 4 and table 5. It is determined that the ANN model from this research performed slightly better in terms of error than that of the previous research [10]. The correlation is also much higher, but even though that’s the case still the previous research with far lower correlation still has a lower error in its experiment. This might be the cause of the quality of its data and also the algorithm of the ANN. Having said that, the ANN have performed well and will be better with fine tuning to help predict the medium day ahead for Baron Techno-Park. Its application in knowing the day ahead will benefit in maintenance scheduling and also solar power potential of the area for further additional installation of solar panels in-the-near-future.

![Figure 5. RMSE comparison.](image-url)
4. CONCLUSION AND OUTLOOK
In conclusion, the findings for the performance analysis of the ANN for the case study in Baron Techno-park in predicting the medium day ahead solar radiation, have been satisfactory. Looking at figure 2 where to determine the best correlation input to the output and running an experiment to find the best ANN structure and training algorithm, which was determine from table 2 that the Levenberg-Marquardt training algorithm with the ANN structure of 1-7-1 was the best results that was obtained. Nonetheless, another experiment was tested to see the comparison of the performance of the ANN of this research with a previous research [10]. The finding shows from table 3, table 4 and table 5, that despite Baron 1-7-1 performed better than 2P 1-2-1 and Baron 1-2-1, regardless that the RMSE have a slight difference but still Baron 1-7-1 outperformed the others, with the best value of RMSE 0.15185 and $R^2$ of 0.88996.

ACKNOWLEDGEMENT
We would like to acknowledge our gratitude to DPRM Kemenristek/BRIN that has given us their full support for this research through their funding in order to help us publish this research in a prestigious journal. We hope in the success of this research being published we can continue to work together to have better research publicises in the near future. We also would like to acknowledge the collaboration team from Baron Techno-park and Universitas Teknologi Yogyakarta for their hard work and dedication for this research.

REFERENCES
[1] M. G. Pae, R. B. Prasetyo, S. Suharyanto, and T. Haryono, “Performance Analysis of Hybrid Power Plant (Wind and Solar) in Baron Techno Park Gunung Kidul Yogyakarta,” J. Ilm. Tek. Elektro Komput. dan Inform., vol. 3, no. 2, p. 110, 2017, doi: 10.26555/jiteki.v3i2.7830.
[2] M. G. Pae, T. Prasetyo, S. Suharyanto, T. Haryono, and R. B. Prasetyo, “Effect of Load Growth on PLTH Baron Techno Park Performance,” IJITEE (International J. Inf. Technol. Electr. Eng.), vol. 1, no. 4, 2018, doi: 10.22146/ijitee.35028.
[3] P. Aji, K. Wakamori, and H. Mineno, Short-Term Solar Power Forecasting Using SVR on Hybrid PV Power Plant in Indonesia, vol. 1035. Springer International Publishing, 2020.
[4] M. I. Yusoff, M. Irvanto, S. Ibrahim, G. Nair, S. I. S. Hassan, and M. Fitra, “Estimating Daily Solar Radiation Using Hargreaves Model in Eastern Malaysia,” Appl. Mech. Mater., vol. 699, no.
April, pp. 564–569, 2014, doi: 10.4028/www.scientific.net/amm.699.564.

[5] A. Kumar, M. Rizwan, and U. Nangia, “Artificial Neural Network Based Model for Short Term Solar Radiation Forecasting Considering Aerosol Index,” 2018 2nd IEEE Int. Conf. Power Electron. Intell. Control Energy Syst. ICPEICES 2018, pp. 212–217, 2018, doi: 10.1109/ICPEICES.2018.8897290.

[6] S. Al-Dahidi, O. Ayadi, J. Adeeb, and M. Louzazni, “Assessment of Artificial Neural Networks Learning Algorithms and Training Datasets for Solar Photovoltaic Power Production Prediction,” Front. Energy Res., vol. 7, no. November, pp. 1–18, 2019, doi: 10.3389/fenrg.2019.00130.

[7] Z. Pang, F. Niu, and Z. O’Neill, “Solar Radiation Prediction Using Recurrent Neural Network and Artificial Neural Network: A Case Study With Comparisons,” Renew. Energy, vol. 156, pp. 279–289, 2020, doi: 10.1016/j.renene.2020.04.042.

[8] M. F. Li, X. P. Tang, W. Wu, and H. Bin Liu, “General Models for Estimating Daily Global Solar Radiation for Different Solar Radiation Zones in Mainland China,” Energy Convers. Manag., vol. 70, pp. 139–148, 2013, doi: 10.1016/j.enconman.2013.03.004.

[9] R. K. B. Navas, S. Prakash, and T. Sasipraba, “Artificial Neural Network Based Computing Model for Wind Speed Prediction: A Case Study of Coimbatore, Tamil Nadu, India,” Phys. A Stat. Mech. its Appl., vol. 542, 2020, doi: 10.1016/j.physa.2019.123383.

[10] V. N. Bapat, “Global Solar Radiation Estimation Model with Two Parameters and its ANN validation,” IEEE Spons. 9th Int. Conf. Intell. Syst. Control, no. 1, pp. 2–4, 2015.