Prediction of Solar Radiation Intensity using Extreme Learning Machine

Hadi Suyono¹, Hari Santoso², Rini Nur Hasanah³, Unggul Wibawa⁴, Ismail Musirin⁵
¹,²,³,⁴Electrical Engineering Department, Faculty of Engineering, Universitas Brawijaya
Jalan MT. Haryono 167 Malang 65145 Indonesia
⁵Faculty of Electrical Engineering, Universiti Teknologi MARA, 40450, Shah Alam, Selangor, Malaysia

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.

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 carbon dioxide (CO₂), sulfur dioxide (SO₂), sulfur trioxide...
(SO$_2$), nitrogen oxides (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 thorough 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 ELM 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]:

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} e_i^2}
\]

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} |e_i|
\]

where based on the standard error model \(e\) given on each calculation, the error obtained at each iteration is \(e_i\), with \(i = 1, 2, 3, \ldots n\). 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](image1.png)

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

![Figure 2](image2.png)

**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.
3. RESULTS AND ANALYSIS

3.1. Meteoblue Climatology Data (NOAA)

The data used in this research have been obtained from the NOAA Meteoblue 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.

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/m² and 91.569W/m², whereas using the MLR method they are 150.547W/m² and 113.405 W/m² 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 132.239W/m² and 91.569W/m², whereas using the MLR method they are 142.521W/m² and 103.805 W/m² 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 131.265 W/m² and 90.637 W/m², whereas using the MLR method they are 148.272 W/m² and 108.645 W/m² 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>
<thead>
<tr>
<th>Experiment number</th>
<th>% training data</th>
<th>% testing data</th>
<th>Training data amount</th>
<th>Testing data amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>70% - 30%</td>
<td></td>
<td>30660</td>
<td>13140</td>
</tr>
<tr>
<td>2</td>
<td>75% - 25%</td>
<td></td>
<td>32850</td>
<td>10950</td>
</tr>
<tr>
<td>3</td>
<td>80% - 20%</td>
<td></td>
<td>35040</td>
<td>8760</td>
</tr>
<tr>
<td>4</td>
<td>85% - 15%</td>
<td></td>
<td>37230</td>
<td>6570</td>
</tr>
<tr>
<td>5</td>
<td>90% - 10%</td>
<td></td>
<td>39420</td>
<td>4380</td>
</tr>
<tr>
<td>6</td>
<td>95% - 5%</td>
<td></td>
<td>41610</td>
<td>2190</td>
</tr>
</tbody>
</table>

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 137.932 W/m² and 96.272 W/m², whereas using the MLR method they are 156.741 W/m² and 119.219 W/m² 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 121.968 W/m² and 82.657 W/m², whereas using the MLR method they are 147.015 W/m² and 113.072 W/m² 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 85.064 W/m² and 56.749 W/m², whereas using the MLR method they are 101.978 W/m² and 71.088 W/m² respectively. The prediction results using the ELM method are much closer to the actual data being compared to the results of MLR prediction method.
3.8. 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 122.45 W/m² and the average MAE error of 84.04 W/m², whereas the MLR method yielded an average RMSE error of 141.18 W/m² and an average MAE error of 104.87 W/m². 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 than 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.

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 85.064 W/m² and the smallest MAE value of 56.749 W/m², while using the MLR method the smallest RMSE value is 101.978 W/m² and the smallest MAE value of 71.088 W/m², 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 than those of the MLR method.
REFERENCES


Prediction of Solar Radiation Intensity using Extreme Learning Machine (Hadi Suyono)


