Estimation of Solar Panel Output based on Weather Parameters using Machine Learning Algorithms

Raheel Siddiqui[1], Sulaiman Umer[1], Asif Iqbal[2], Farman Ullah[1], Ajmal Khan[1], Kyung Sup Kwak[2]
[1]*Department of Electrical and Computer Engineering, COMSATS University Islamabad, Attock Campus*
[2]*Department of Information and Communication Engineering, Inha University, Incheon, South Korea*

ABSTRACT

Solar energy is one of the most extensively used renewable energy sources. However, it is highly variable and needs accurate estimation for its wide range of integration into the electricity grid. Solar voltage and current are estimated in areas where only sunlight is considered as a primary solar parameter, and information about their weather conditions are unknown. Weather plays a vital role in the prediction of solar panel output. In this paper, we propose solar panel output prediction considering the solar panel and weather parameters using machine learning algorithms. We estimate the solar panel voltage and current consider the weather parameters such as temperature, humidity, rain rate, wind speed, and wind direction. For estimating the output voltage and current, Linear Regression (LR) and Artificial Neural Network (ANN) are applied on weather and solar data. The datasets are extracted from Bancroft close 49KW substation, which is placed in the UK, for three months. The performance of the given model is evaluated using two matrices Root Mean Square Error (RMSE) and Absolute Error (AE). The Neural Network shows better accuracy compared to the linear regression.

Keywords—Solar Energy; Renewable energy; Linear Regression; Artificial Neural Network (ANN); Solar Panel Output

1-INTRODUCTION

Conventionally, energy is extracted from non-renewable resources like fossil fuels such as oil, coal and natural gas. However, these resources have a limit and are meant to end one day. These resources also have drawbacks, such as polluting the environment and have a risk of human diseases. Hence the world needs to move towards renewable resources to cater to these issues [1] and has focused on renewable energy up to a greater extent. However, it has not met the climate goals set internationally by the Paris agreement or the goals set by sustainable development [2-3]. As solar power obtained from Photovoltaic (PV) is one of the most widely used renewable energy process, experts have been discussing methods to obtain accurate estimations of solar power [5]. According to a study in 2016, solar panel systems have increased by 100 times since 2000, and the total generation is expected to be 540 GW in 2019 and is expected to reach 609 GW by 2025. It means that the increase in solar power usage has increased exponentially in the last decade. Hence, the need to accurately estimate has increased as well [6].

Internet-of-Things plays an important role in data collection and providing an analysis platform [7]. Harvested solar energy is mostly relying on solar irradiance and other climatological factors such as temperature, wind speed and humidity [4]. Machine learning techniques such as Linear Regression, Neural network (NN), Deep Learning (DL) are used in training and testing data sets of weather parameters humidity, temperature, wind direction and irradiance. Furthermore, electrical parameters voltage, current and power are also predicted by using the techniques mentioned above [8-9]. However, in this article prime goal to estimate the solar power by using datasets of weather parameters and compare the root mean square (RMSE) and absolute error (AE) of available datasets using the machine learning algorithms

2-THE PROPOSED ARCHITECTURE FOR SOLAR PANEL OUTPUT PREDICATION

Figure. 1 shows the proposed architecture of solar panel output prediction based on weather parameter based on the

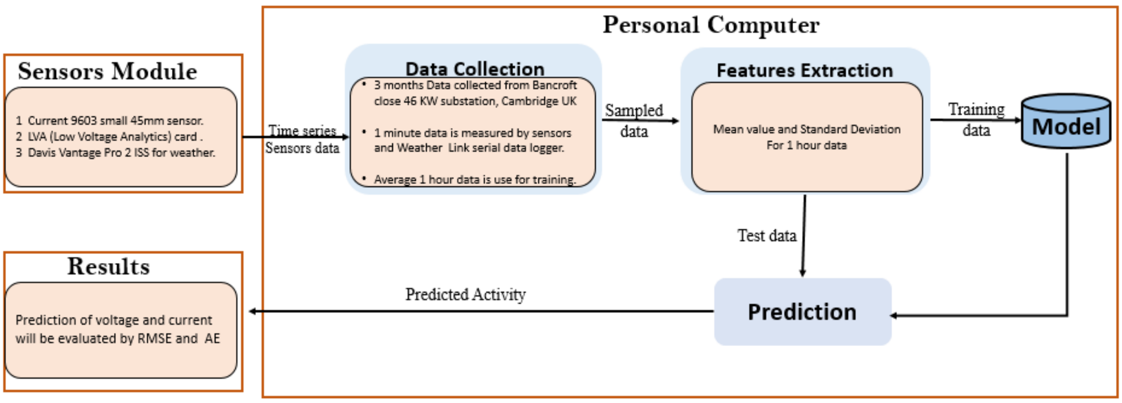

Figure 1 The Propose Architecture for Estiamation of Solar Panel Output based on Weather Data using Machiner Learning algorithms

weather parameters using machine learning algorithms. The weather data is collected using various sensors which is pre-processed and various kind of features are extracted from the data. In this paper, we use the PV power and weather data mined from UK Power Networks [10]; large scale data collected from domestic sites and substations in the UK. This dataset comprises of solar panel current, voltage, power, and weather parameters such as humidity, solar irradiance, dew point, temperature, wind speed, wind direction, biometric pressure, rain and many others. In this research article, only three months of data is extracted from the available dataset for analysis. PV and weather data of Croftclose 49KW substation is used for training the model. Duration of data is 1st January 2014 to 1st March 2014. For training purpose, 1-minute interval record is converted into a 1-hour interval record by taking the mean value of all 60 minutes data. The total number of records is 2136.

2.3 FEATURE EXTRACTION

There are fourteen main features taken into consideration in available data set in order to use for training the model. Six features out the fourteen are related to the solar panel which is minimum voltage (V_min), Maximum voltage (V_max), minimum current (I_min), Maximum current (I_max), minimum power (P_min) and Maximum power (P_max). Eight features of outdoor weather such as Humidity, Temperature, Wind speed, Wind direction, Dew point, Rain, Rain rate, Atmospheric pressure and Solar irradiance. In this process, the maximum voltage (V_max) and maximum current (I_max) values of the solar panel are predicted through applying linear regression (LR) models, neural network (NN) algorithms. Performance is evaluated by comparing the results of root means square error (RMSE) and absolute error (AE).

2.4 The MACHINE LEARNING ALGORITHMS FOR SOLAR OUTPUT PREDICATION

2.4.1 Linear Regression Model

Linear regression technique is applied on supervised learning. It predicts the dependent variable (y-axis) based on the independent variable (x-axis), which is why it gives a linear relation. Then it is applied to real data type problems. In this study, V_max and I_max are predicted by using two different ratios of training and testing data.

Mathematical equation of Linear Regression [11].

$$h(x) = \sum_{i=0}^{n} \theta_i x_i \qquad (1)$$

$$J(\theta) = \frac{1}{2} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2 \qquad (2)$$

h(x) is the hypothesis value in which thirteen weather and solar parameters are used for estimation. J (theta) is the cost function or it can use as an error function. It should be minimum for maximum accuracy. $y^{(i)}$ is the actual value

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta). \qquad (3)$$

The equation mentioned above shows the gradient descent function. α denotes the Learning rate it should be 0 ~ 0.5.

$$\theta_j := \theta_j + \alpha \sum_{i=1}^{m} \left(y^{(i)} - h_\theta(x^{(i)})\right) x_j^{(i)} \qquad (4)$$

For every value of J repeat the above mentioned equation until convergence is obtained.

2.4.2 Neural Network (NN) Model

Artificial Neural Networks or ANN is an information processing model influenced by the way the human nervous system works, such as knowledge about the brain process, it operates together with a vast number of highly integrated computing components (neurons) to solve a particular problem. Figure. 2 shows the ANN model used for solar output parameters prediction.

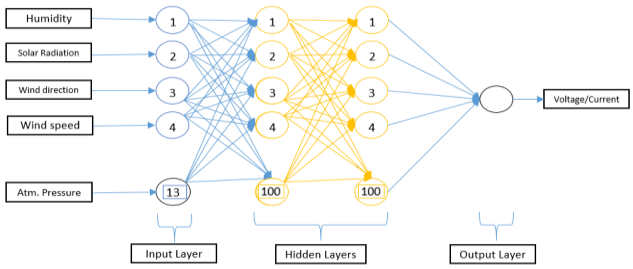

Figure 2 Structure of Artificial Neural Network (ANN)

In the above figure, we are predicting voltage and current. There are 13 input parameters; in this scenario, the ANN has two hidden layers. Each layer comprises of 100 neurons. Learning rate (α) is 0.01 and Epsilon error is 1.00e-4. The different number of training cycles are analyzed in this study. The output is predicted by using two ratios of training and testing data, which are 8:2 and 9:1.

3-RESULTS AND DISCUSSION

We used the RMSE and AE as a performance metrics to predict the voltage and current as solar panel output. Figure. 3-6 shows the curve fit by the linear regression model using the 80-20 and 90-10 percent training and testing data. In Table 1, RMSE value is less in 8:2 ratio for both V_max and I_max, and it is better as compared to 9:1 training and testing data ratio.

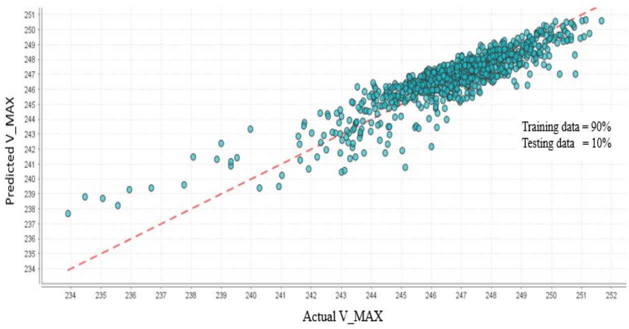

Figure 3 V_max predicted by using training data is 90% training data is 10%

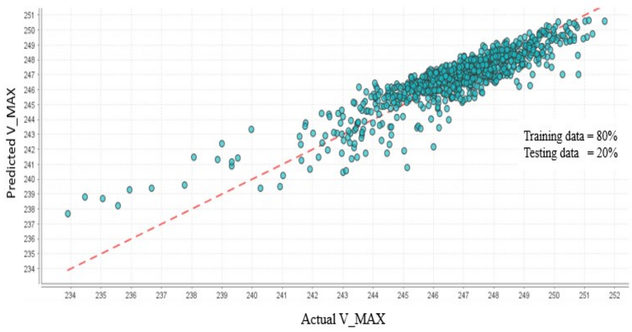

Figure 4 (V_max predicted by using training data is 80% and testing data is 20%

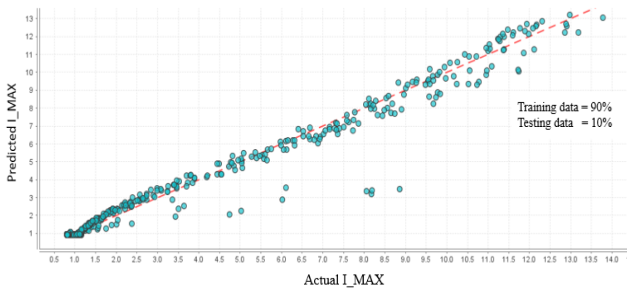

Figure 5 (I_max predicted by using training data is 90% and testing data is 10%

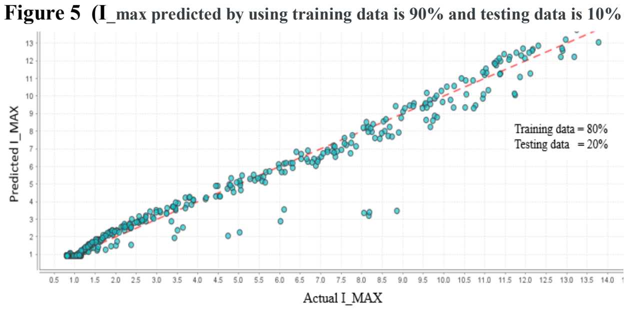

Figure 6 I_max predicted by using training data is 80% and testing data is 20%

Table 1 Linear Regression for RMSE and AE using different training data

Training & Testing Data	Min Tolerance	Ridge	Root Mean Square Error(RMSE)	Absolute Error (AE)
8:2 of V_max	0.05	1.00E-08	0.977	0.755
9:1 of V_max	0.05	1.00E-08	1.03	0.779
8:2 of I_max	0.05	1.00E-08	1.027	0.882
9:1 of I_max	0.05	1.00E-08	1.093	0.806

Voltage is predicted using ANN, by this prediction RMSE values are analyzed by using training cycles (T.C) from 5 to 500, as shown in figure 7. The two ratios of training and testing data have been used. Firstly, we have chosen 90% training data and 10% testing data. This ratio shows minimum RMSE value 1.056 at 200 training cycles. In ratio 8:2 of training and testing data, the minimum RMSE value is 0.956 at 300 training cycles.

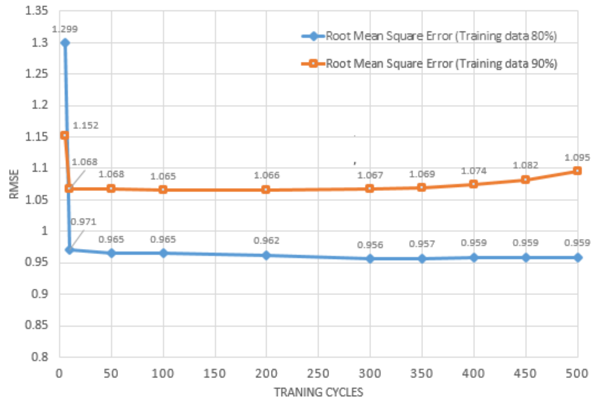

Figure 7, **RMSE is analyzed on different training cycles while predicting V_max**

Table 2 Neural Network table for RMSE and AE using different training data

Training and testing data ratio	Minimum Root Mean Square Error (RMSE)	Minimum Absolute Error (AE)
V_Max at 8:2	0.956	0.730
V_Max at 9:1	1.056	0.605
I_Max at 8:2	1.013	0.769
I_Max at 9:1	1.165	0.868

In table 2, Using ANN algorithm minimum values of RMSE and AE are realized. While predicting V_max and I_max, using different training cycles (T.C) as given.

CONCULSION

In this article, voltage and current are predicted using solar and weather parameters, i.e. voltage, current, power, humidity, atmospheric pressure, wind direction, wind speed, rain, dew point, temperature and rain rate. Data is collected from Bancroft close 49KW substation. Artificial Neural Networks (ANN) and Linear Regression (LR) are used to predict the solar parameters V_MAX and I_MAX using two sets of training and testing data ratio 9:1 and 8:2. The performance is analyzed through the root mean square RMSE and Absolute error (AE).

When ANN and LR techniques are applied for predicting V_MAX using Training data 90%, minimum AE value is achieved by the ANN algorithm. In the case of training data 80%, minimum RMSE value obtained is 0.956 by ANN algorithm, which is lowest compared to the LR algorithm. On the other hand, while predicting I_MAX, the ANN and LR are used. The minimum value of RMSE is obtained by the ANN algorithm.

After analyzing all the parameters and performance matrices, ANN is more accurate as compared to LR for this specific location data set.

References

[1] Owusu, Phebe Asantewaa, and Samuel Asumadu Sarkodie. "A review of renewable energy sources, sustainability issues

and climate change mitigation." *Cogent Engineering* 3, no. 1 (2016): 1167990.

[2] Md Bengir Ahmed, Mohammad Asaduzzaman Chowdhury, Shameem Ahmed, and Mohammod Abul Kashem. "Prediction of solar irradiation and performance evaluation of grid connected solar 80KWp PV plant in Bangladesh." *Energy Reports* 5 (2019): 714-722.

[3] Shaw, Subham, and M. Prakash. "Forecasting Solar Potential Using Support Vector Regression." In *2019 Devices for Integrated Circuit (DevIC)*, pp. 5-8. IEEE, 2019.

[4] Wang, Zheng, Irena Koprinska, and Mashud Rana. "Solar power prediction using weather type pair patterns." In *2017 International Joint Conference on Neural Networks (IJCNN)*, pp. 4259-4266. IEEE, 2017.

[5] Sharma, Navin, Pranshu Sharma, David Irwin, and Prashant Shenoy. "Predicting solar generation from weather forecasts using machine learning." In *2011 IEEE international conference on smart grid communications (SmartGridComm)*, pp. 528-533. IEEE, 2011.

[6] Rana, Mashud, Irena Koprinska, and Vassilios G. Agelidis. "Solar power forecasting using weather type clustering and ensembles of neural networks." In *2016 International Joint Conference on Neural Networks (IJCNN)*, pp. 4962-4969. IEEE, 2016.

[7] Iqbal, Asif, Farman Ullah, Hafeez Anwar, Kyung Sup Kwak, Muhammad Imran, Waseef Jamal, and Atta ur Rahman. "Interoperable Internet-of-Things platform for smart home system using Web-of-Objects and cloud." Sustainable Cities and Society 38 (2018): 636-646.

[8] Kim, Seul-Gi, Jae-Yoon Jung, and Min Kyu Sim. "A Two-Step Approach to Solar Power Generation Prediction Based on Weather Data Using Machine Learning." *Sustainability* 11, no. 5 (2019): 1501.

[9] Sharma, Navin, Pranshu Sharma, David Irwin, and Prashant Shenoy. "Predicting solar generation from weather forecasts using machine learning." In *2011 IEEE international conference on smart grid communications (SmartGridComm)*, pp. 528-533. IEEE, 2011.

[10] Link for Dataset: http://Ups-solar.co.uk/?msclkid=aa74b4ab0c5f16290005ebb2faf11c41

[11] Ng, Andrew. "CS229 Lecture notes." CS229 Lecture notes 1, no. 1 (2000): 1-3.