2020 IEEE REGION 10 CONFERENCE (TENCON)
Osaka, Japan, November 16-19, 2020

Forecasting PV Panel Output Using Prophet Time Series Machine Learning Model

Md. Mehedi Hasan Shawon
Department of EEE
BRAC University
Dhaka, Bangladesh
shawon.eee.bracu@gmail.com

Sumaiya Akter
Department of EEE
BRAC University
Dhaka, Bangladesh
sumaiya.eee.bracu@gmail.com

Md. Kamrul Islam
Department of EEE
BRAC University
Dhaka, Bangladesh
md.kamrul.islam@g.bracu.ac.bd

Sabbir Ahmed
Department of EEE
BRAC University
Dhaka, Bangladesh
sabbir.ahmed2@g.bracu.ac.bd

Md. Mosaddequr Rahman
Department of EEE
BRAC University
Dhaka, Bangladesh
mosaddeq@bracu.ac.bd

***Abstract*—Due to climate change effects, the demand for renewable energy is growing immensely around the world. Photovoltaic (PV) panels are widely popular as a vital source of renewable energy all over the world as well as in Bangladesh. However, besides solar irradiance, the panel output is greatly affected by some of the weather parameters like temperature, humidity, wind, etc. Reliable forecasting of PV panel output is essential for capacity planning in advance to efficiently manage the energy distribution. This paper presents a method to forecast the PV panel output energy using a machine learning model, known as the Prophet Model used for a univariate time series forecasting. For this study, the PV panel generated data are collected from an outdoor experimental set-up throughout the full winter season in Bangladesh. Based on the data, forecasting of one-day-ahead PV panel short circuit current is done, and then the estimation of PV panel output energy is made. The results show the proposed forecasting method to be quite encouraging and reliable one while providing a higher coefficient of determination value with an average 0.9772 for one-day-ahead PV panel output energy forecasting.**

Keywords—Prophet Model, Forecasting, PV panel, Short circuit current, Renewable energy.

I. Introduction

Global warming is a growing concern around the world, and Bangladesh is in a vulnerable position to face this problem. However, renewable energy is one of the most effective tools to fight against global warming. The Photovoltaic (PV) panel is broadly accepted around the world for its trouble-free and easily accessible clean energy. Bangladesh has an immense number of solar home systems, which makes the largest off-grid electrification in the world, covering 14 percent of its total population. In addition, a number of projects are currently going on to expand the contribution of solar energy in Bangladesh. However, the PV panel generation is sensible to weather parameters like solar irradiance, temperature, humidity, wind, and its production is affected by weather conditions such as rain, cloud, fog, etc. Hence, there is always an uncertainty of PV panel generation. Therefore, reliable and more accurate forecasting of one-day-ahead PV panel output energy becomes essential for the commercial or non-commercial planners to plan in advance and efficiently manage the energy distribution.

Numerous works of literature highlight solar PV generation forecasting using a variety of methods. The main classes of these methods include Numerical Weather Prediction (NWP), Satellite/Sky image-based forecasting, Regression, Hybrid System, Machine Learning, Artificial Neural Network (ANN), etc. NWP is a traditional technique for several hour solar irradiance forecasting in advance [1]. A combination of two or more methods is characterized by Hybrid System to utilize each model's unique strengths. The combination of processed sky images with cloud motion information and machine learning is performed for 10 minutes ahead of solar irradiance forecasting with high accuracy [2]. Hybrid methodology like Mycielski-Markov is in practice for short term solar power forecasting, which has a 0.87 correlation of determination value [3]. Besides, an innovative method like weighted Gaussian process regression is also used for short term solar power forecasting with higher accuracy than other typical machine learning approaches [4].

The support vector machine (SVM) proposed by Vapnik [5] is a machine learning algorithm that is widely practiced in past years for solar PV generation forecasting [6-7]. The accuracy of SVM depends on the proper kernel functions and parameters [8]. SVM is also assembled with NWP to improve forecasting accuracy, and this approach shows promising results to forecast solar power [9]. Support vector regression (SVR), an extended form of SVM, is also a popular technique for short-term PV power output forecasting [10].

However, the latest methods are introduced each year to increase forecasting accuracy. In this study, a new method is presented for PV panel output energy forecasting using the Prophet time series model developed by Facebook for an accurate, fast, fully automatic, and tunable forecast. To the best of our knowledge, this method is not applied yet for one-day-ahead PV panel output energy forecasting.

The organization of this paper as follows: Section II describes the experimental setup, Section III briefly explains the theoretical background of forecasting and the expected panel output energy calculation, Section IV discusses the forecasting results, and finally, Section V provides the conclusion of this study.

II. Experimental Setup

For this study, one monocrystalline PV panel is set-up on a rooftop of a seven-story residential building in Dhaka to collect the PV panel short circuit current and surface temperature. A block diagram of this experimental set-up for data collection is presented in Fig. 1. In this set-up, one INA219 current sensor and one DS18B20 temperature sensor are used. Current sensor data is collected by an Arduino Uno,

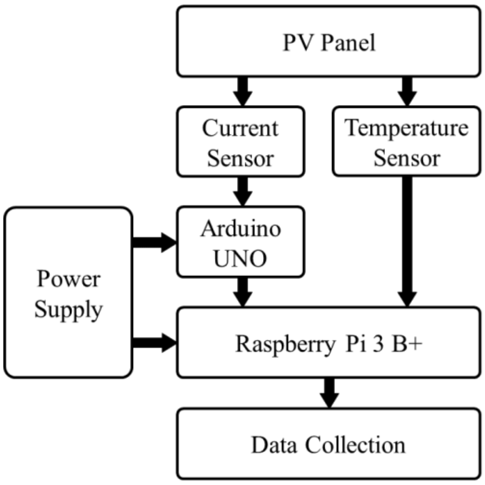

Fig. 1. Block diagram of the experimental set-up of PV panel under outdoor condition.

and then it is transmitted in Raspberry Pi, and temperature sensor data is directly saved in Raspberry Pi. AC to DC converters are used to power up the Raspberry Pi and Arduino Uno. Furthermore, Raspberry Pi data are collected remotely using Putty and TightVNC from the project site to a Windows operated computer at home. Per-minute 3 data are collected from each sensor from 6 am to 6 pm. Data collection has been continued from 1st November 2019 to 29th February 2020 to cover the whole winter season in Bangladesh. In the meantime, a few days of data collection are missed due to technical maintenance.

TABLE I. THE SPECIFICATIONS OF PV PANEL

Specification	Value
Maximum Power (P_{mp})	20WP
Open circuit Voltage (V_{oc})	22V
Short Circuit Current (I_{sc})	1.46A
Voltage at Maximum Power (V_{mp})	18V
Current at Maximum Power (I_{mp})	1.11A

III. THEORETICAL BACKGROUND

A. Forecasting

For time series forecasting, algorithms like Auto-Regressive Integrated Moving Average (ARIMA), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM) are quite popular. However, Facebook has recently developed Prophet for a robust time-series forecasting, a machine learning model for a univariate time series forecasting. Prophet has some intuitive and easily interpretable customizations that allow more accessibility to the audience and more applicability to a wide range of forecasting models [11]. The motivation behind the development of the Prophet model is to make high-quality predictions easier and to get a more accurate and realistic forecast. It only needs the right amount of historical data for better modeling. In our study, the Prophet model is implemented for PV panel short circuit current forecasting using a historical short circuit current data of four months. Then using the forecasted short circuit current, PV panel output energy is estimated.

The prophet is a simple additive regression model *y(t)* with three main components: piecewise trend, seasonality, and holiday effects. The equation can be stated as [11],

$$y(t) = g(t) + s(t) + h(t) + \epsilon(t) \tag{1}$$

where, *g(t)* is the trend function which models the linear or logistic changes over time, *s(t)* captures the periodic changes over historical data which can be daily, weekly, monthly or yearly seasonality, *h(t)* represents abnormal predictable days of the year which occurs on irregular schedules and $\epsilon(t)$ indicates independent and identically distributes noise, which is not accommodated by the model. Prophet is robust to outliers, missing values, and sudden changes in the time series forecasting.

Different metrics are used to evaluate the forecasting performance, such as coefficient of determination (R^2), Mean Absolute Percentage Error (MAPE), Relative Error (RE), Root Mean Square Error (RMSE).

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - y_{mean})^2} \tag{2}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|y_i - \hat{y}_i|}{y_i} \times 100\ \% \tag{3}$$

$$RE = \frac{|y_t - \hat{y}_t|}{y_t} \times 100\ \% \tag{4}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{5}$$

Here, y_i is the experimental value and $\hat{y}_i$ is the forecasted value at i^{th} time, y_{mean} is the average of experimental values, y_t and $\hat{y}_t$ are respectively the final experimental and forecasted values, and N is the number of forecast observations in the estimation period. The closer R^2 value to 1, the better fitness of the forecasted values to experiment. The lesser values of MAPE, RE, RMSE the better model is to be considered. A flowchart of the Prophet model working procedures for PV panel short circuit current forecasting is presented in Fig. 2.

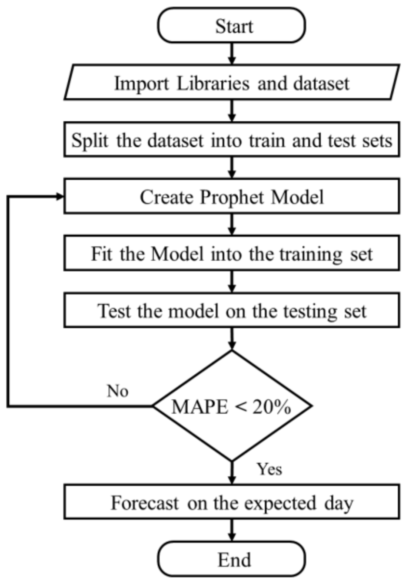

Fig. 2. Flowchart of the working procedures of Prophet model.

Here, after importing the libraries and dataset, the training and testing datasets are split with 80% and 20% data. Then the Prophet model is created, and the training dataset is fitted into the model. Later, the model is tested on a recent testing dataset. To evaluate our model, we use the MAPE metric, which gives a proper sense of fluctuation in prediction as our experimental data varies from 10mA to almost 1200mA. Forecasting is done if MAPE is between 0-20% on testing set. However, the MAPE can bigger or lesser than 20% in the forecasting part.

B. *Expected Panel Output Energy Calculatiuon*

At first, the ideality factor n is calculated from an experiment, one of the parameters of diode I-V characteristics. From another test, the reverse saturation current, I_0 is calculated using the equation below, which is given in [12],

$$I_0 = \frac{I_{sc}}{e^{\left(\frac{q\,V_{oc}}{nkTN_s}\right)}} \tag{6}$$

where, q is the charge of an electron, k is the Boltzmann's constant, N_s is the number of cells connected in series in the panel, I_{sc} and V_{oc} are respectively short circuit current and open-circuit voltage at temperature, T. A large amount of data is collected on a daily basis, which makes it impossible to calculate the I_0 for each data using (6). As a result, another equation from [12] is applied to calculate the new reverse saturation current, $I_{0(new)}$ at a new temperature, T_{new}.

$$I_{0(new)} = I_0 * \left(\frac{T_{new}}{T}\right)^3 * e^{\left(\left(\frac{E_g}{k}\right)*\left(\frac{1}{T}-\frac{1}{T_{new}}\right)\right)} \tag{7}$$

Here, E_g is the energy band gap of silicon. Next, new open-circuit voltage, $V_{oc(new)}$ is calculated for each $I_{sc(new)}$ and $I_{0(new)}$ by applying the equation stated as below [12].

$$V_{oc(new)} = \frac{nkT_{new}N_s}{q} * ln\left(\frac{I_{sc(new)}}{I_{0(new)}}\right) \tag{8}$$

Since the hardware setup is planned outdoor, the I-V characteristic curve cannot be obtained regularly for finding the maximum power. However, the fill factor (FF) is calculated to compute the solar panel maximum power, P_{max} for each point by applying the equation stated as below [12].

$$P_{max} = V_{oc\,(new)} * I_{sc\,(new)} * FF \tag{9}$$

Then the P_{max} for each $I_{sc\,(new)}$ is integrated using the trapezoidal method of integration with respect to time in order to compute the PV panel cumulative electrical output energy. For further calculation of PV panel output energy, the experimental data has been used. The difference between the calculated and experimental data is very less, which is shown in Table II. Thus, this procedure is liable to use.

TABLE II. ENERGY CALCULATION

Experimental Data		**Calculated Data**	
Parameter	Value	Parameter	Value
T	46°C	n	1.35
I_{sc}	1.187A	I_0	1.13×10^{-7}
V_{oc}	21.6V	FF	0.65
T_{new}	45°C	$I_{0(new)}$	9.87×10^{-8}
$I_{sc(new)}$	1.19A	$V_{oc(new)}$	21.72V
$V_{oc(new)}$	21.7V	P_{max}	16.80W
P_{max}	17.0W	E_{panel}	108.16Wh

IV. RESULTS AND ANALYSIS

A. *Short Circuit Current*

From experimental data, it is found that PV panel short circuit current increases from 6 am to midday with the increase in solar irradiance and starts falling from midday to 6 pm with a decrease in solar irradiance. However, when a giant cloud passes over the PV panel, a sudden drop in the short circuit current occurs. Besides, on the rainy day or foggy weather, the average short circuit current is lower than the sunny day. The average short circuit current variation of the PV panel throughout the full winter season due to the different weather conditions and parameters is shown in Fig. 3.

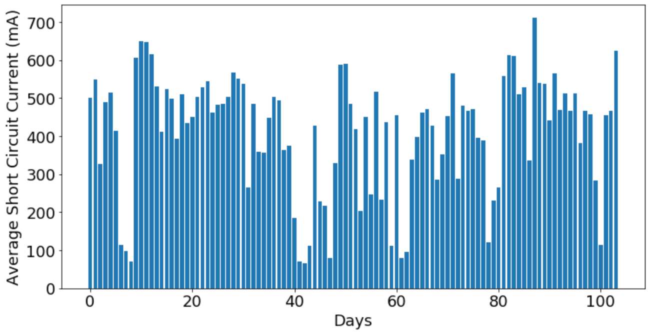

Fig. 3. PV panel average short circuit current variation over four months.

B. *Forecasting Results*

From November 2019 to February 2020, four different datasets are utilized to forecast four different days' PV panel short circuit current, one from each month. Each time the dataset used for training and testing starts from 1st November 2019 to the day before the forecasting date. Four sunny days are selected to forecast short circuit current, and the dates are (a) 28th November 2019, (b) 31st December 2019, (c) 21st January 2020, and (d) 28th February 2020. The experimental and forecasted short circuit current curves on mentioning dates are presented in Fig. 4.

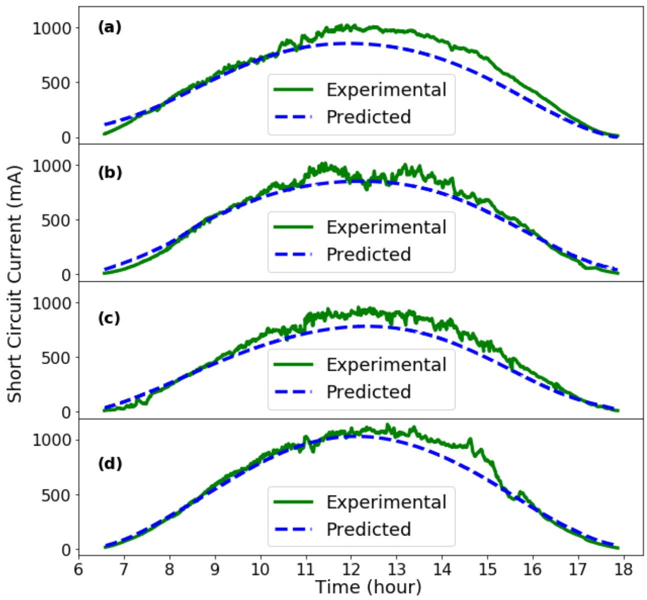

Fig. 4. PV panel experimental and predicted short circuit current.

Although the experimental short circuit current curves at each month are different from others with dissimilar rising and falling rates, the forecasted short circuit current curves nearly maintained the experimental short circuit current curves for the whole day. The forecasted results improve with the longer historical datasets. However, 21st January 2020 is different in this case due to some missing and hampered historical data. Forecasting accuracy metrics for PV panel short circuit current are presented in Table III.

TABLE III. FORECASTING ACCURACY METRICES

Testing Date	R^2	MAPE (%)	RMSE (mA)
28th Nov 19	0.9193	23.10	98.64
31st Dec 19	0.9640	14.63	64.96
21st Jan 20	0.9350	17.10	85.06
28th Feb 20	0.9621	12.13	73.30

From Fig. 4 and Table III, it is clear that the forecasted short circuit current curves fit very well with an average 0.9451 coefficient of determination. Average MAPE and RMSE are 16.74% and 80.49mA, which are also acceptable for a day-long short circuit current prediction.

Next, whole day PV panel output energy is calculated from the short circuit current for the forecasted dates, and the comparison between the experimental and forecasted PV panel output energy curves are shown in Fig. 5.

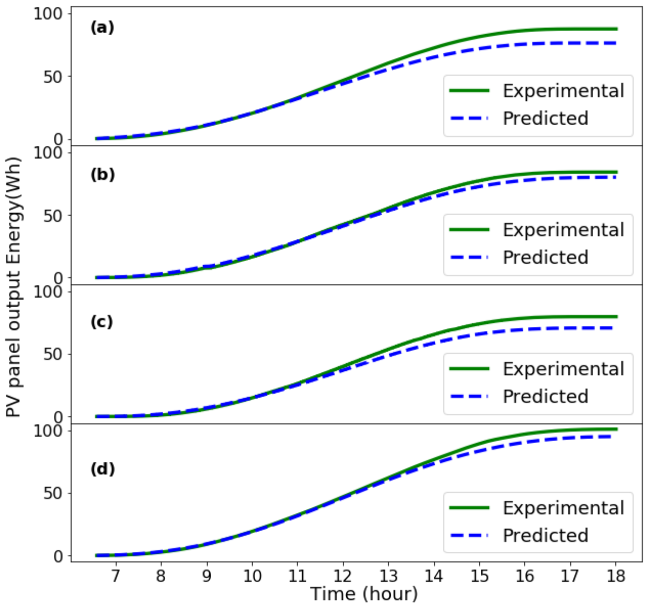

Fig. 5. PV Panel experimental and predicted output energy.

Forecasted PV panel energy curves capture the experimental trend of PV panel output energy curves, maintaining close contiguity throughout the day. The Forecasting accuracy metrics for PV panel energy are shown in Table IV.

TABLE IV. FORECASTING ACCURACY METRICES

Testing Date	R^2	RE (%)	RMSE (Wh)
28th Nov 19	0.9596	12.86	6.55
31st Dec 19	0.9916	4.81	2.91
21st Jan 20	0.9669	11.31	5.59
28th Feb 20	0.9907	5.81	3.59

The average R^2 value is 0.9772 for PV panel output energy forecasting. In contrast, we observed a 0.87 correlation value in [3]. Besides, the average RMSE is 4.66Wh, which is much improved than the proposed method in [4] for short term solar power forecasting. Lastly, the average RE is 8.7%, indicating that the proposed method is reliable for a day ahead of PV panel energy forecasting.

V. CONCLUSION

In this study, the Prophet machine learning model for univariate time series forecasting is employed for a day-long PV panel short circuit current forecasting. Later, PV panel output energy is estimated theoretically using the measured short circuit current. Comparative analysis between the experimental and forecasted results demonstrates a higher estimation accuracy, a higher coefficient of determination value and a lower RMSE for the proposed method than some of the forecasting methods reported in the literature. Hence, the proposed method proves to be a reliable and an accurate method for forecasting of one-day-ahead PV panel output.

REFERENCES

[1] P. Mathiesen and J. Kleissl, "Evaluation of numerical weather prediction for intra-day solar forecasting in the continental United States," *Solar Energy,* vol. 85, no. 5, pp. 967-977, 2011/05/01/ 2011, doi: https://doi.org/10.1016/j.solener.2011.02.013.

[2] S. Tiwari, R. Sabzehgar and M. Rasouli, "Short Term Solar Irradiance Forecast based on Image Processing and Cloud Motion Detection," 2019 IEEE Texas Power and Energy Conference (TPEC), College Station, TX, USA, 2019, pp. 1-6, doi: 10.1109/TPEC.2019.8662134.

[3] F. Serttas, F. O. Hocaoglu and E. Akarslan, "Short Term Solar Power Generation Forecasting: A Novel Approach," 2018 International Conference on Photovoltaic Science and Technologies (PVCon), Ankara, 2018, pp. 1-4, doi: 10.1109/PVCon.2018.8523919.

[4] H. Sheng, J. Xiao, Y. Cheng, Q. Ni and S. Wang, "Short-Term Solar Power Forecasting Based on Weighted Gaussian Process Regression," in IEEE Transactions on Industrial Electronics, vol. 65, no. 1, pp. 300-308, Jan. 2018, doi: 10.1109/TIE.2017.2714127.

[5] V. Vapnik, "The Support Vector Method of Function Estimation," in *Nonlinear Modeling: Advanced Black-Box Techniques*, J. A. K. Suykens and J. Vandewalle Eds. Boston, MA: Springer US, 1998, pp. 55-85.

[6] J. Zeng and W. Qiao, "Short-term solar power prediction using a support vector machine," *Renewable Energy,* vol. 52, pp. 118-127, 2013/04/01/ 2013, doi: https://doi.org/10.1016/j.renene.2012.10.009.

[7] N. Sharma, P. Sharma, D. Irwin and P. Shenoy, "Predicting solar generation from weather forecasts using machine learning," 2011 IEEE International Conference on Smart Grid Communications (SmartGridComm), Brussels, 2011, pp. 528-533, doi: 10.1109/SmartGridComm.2011.6102379.

[8] M. Z. Hassan, M. E. K. Ali, A. B. M. S. Ali and J. Kumar, "Forecasting Day-Ahead Solar Radiation Using Machine Learning Approach," 2017 4th Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE), Nadi, 2017, pp. 252-258, doi: 10.1109/APWConCSE.2017.00050.

[9] W. Buwei, C. Jianfeng, W. Bo and F. Shuanglei, "A Solar Power Prediction Using Support Vector Machines Based on Multi-source Data Fusion," 2018 International Conference on Power System Technology (POWERCON), Guangzhou, 2018, pp. 4573-4577, doi: 10.1109/POWERCON.2018.8601672.

[10] M. Alrashidi, M. Alrashidi, M. Pipattanasomporn and S. Rahman, "Short-Term PV Output Forecasts with Support Vector Regression Optimized by Cuckoo Search and Differential Evolution Algorithms," 2018 IEEE International Smart Cities Conference (ISC2), Kansas City, MO, USA, 2018, pp. 1-8, doi: 10.1109/ISC2.2018.8656685.

[11] S. J. Taylor and B. Letham, "Forecasting at Scale," *The American Statistician,* vol. 72, no. 1, pp. 37-45, 2018/01/02 2018, doi: 10.1080/00031305.2017.1380080.

[12] PVEducation, [Online]. Available: https://www.pveducation.org/. [Accessed January 15, 2020].