

# Validation of Predictive Models to Estimate Annual PV Production: A case study of Odisha

Rakesh Kumar Tarai<sup>a</sup>, Paresh Kale<sup>b</sup>

<sup>a</sup> Student, Department of Electrical Engineering, National Institute of Technology Rourkela, Dist. Sundargarh, Odisha-769008, India

<sup>b</sup> Asst. Professor, Department of Electrical Engineering, National Institute of Technology Rourkela, Dist. Sundargarh, Odisha-769008, India

---

## Abstract

The decision to install a PV plant depends on three major factors: the climatic and environment conditions of the location, the viability of commercial operations, and the government policies. Economic feasibility of a PV system in the energy market is largely driven by the cost of technology, the cost of installation, and the yield of the plant. Considering uncertain nature of geographical parameters (solar radiation, temperature, dust accumulation, etc.), development of a reliable model to predict the energy output of a plant-to-be installed becomes essential. The model ensures the long-term performance criteria of the PV system. The proposed model considers a case study of Odisha by taking only two meteorological variables collected from 1195 locations: total annual incident global radiation on the PV module and annual average air temperature. The developed model is independent of longitude and latitude, elevation, and other environment conditions. Model is validated using the data collected from SN Mohanty solar power plant situated at Cuttack. The paper focuses on simplification at every stage of the development while validating the preciseness of the model.

*Keywords: Solar PV production, Predictive Model, Validation, PVGIS, Odisha*

---

## 1. Introduction

The generation using Photovoltaic (PV) system is a rapidly growing renewable technology option due to its decreasing per unit cost of electricity. By October 2013, the capacity of the total worldwide solar PV installation reached to 135 GW, out of which installation of 112 GW plant is seen in last four years alone [1]. India has installed 4.1 GW of utility-scale solar as of May 2015, and 60 GW is planned by 2022 [2]. The reduction in Levelized cost of Electricity (LCOE) of large scale solar PV in India at 5% and 10% of weighted average cost of capital (WACC), from the year of 2015 to 2050 is predicted as 3 EUR ct/kWh and 4.2 EUR ct/kWh respectively [3]. The roadmap of the Renewable Energy Production predicts the production of PV to provide 16% of global electricity and 10%-12% of India by 2050 [1]. Interest for PV installations in last decade among all sectors (public sector, the private sector, and common households) in Odisha is growing. In last five years alone the total installed capacity of the solar power plants in Odisha increased to 56.92 MW<sub>p</sub> [4]. As per 'Odisha solar policy 2013', gross renewable energy potential stands at 53,820 MW for the state, out of which the possible potential for power generation using solar PV is about 8000 MW [5].

The economic viability of the solar PV systems in the energy market depends on the system installation cost, the output of the plant, availability of the land, public awareness, and government policies. The energy production of a PV plant depends on the climatic conditions, which are unpredictable, for the location. Thus, the prediction of the output of proposed PV plant for a short term (for a day) and long term (for month and year) beforehand comes in handy for the owners and planners. PVGIS is a popular software that uses an extensive database provided by Climate-SAF (CM-SAF) [6]. The study deals with the PVGIS software to estimate the PV production in the region of Odisha. The method divides the total area of Odisha into 1195 equal parts. All the 1195 coordinates are the location input of the PVGIS, and the PV production is estimated. The annual global radiation and air temperature are collected from PVGIS and SolarGIS [7] respectively.

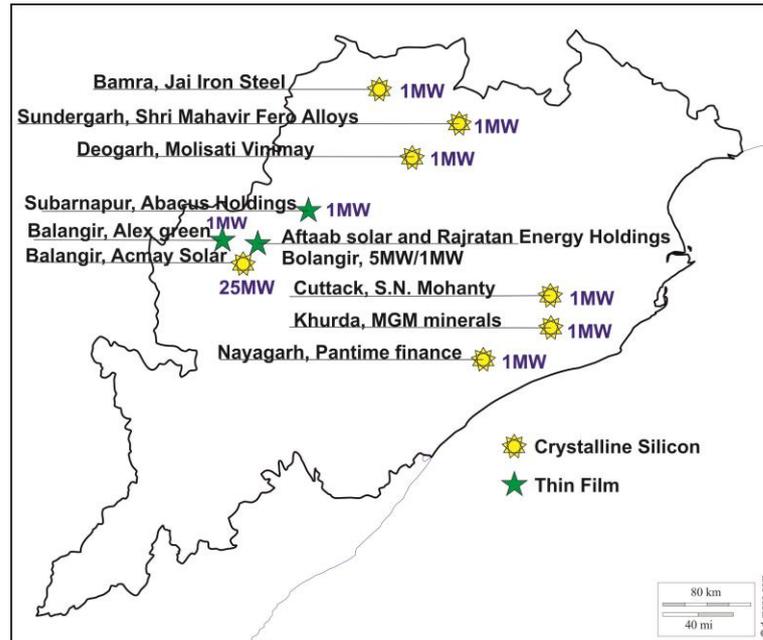


Figure 1 Location of the commissioned MW-scaled, grid connected Solar PV Power Plants of Odisha

Odisha host eleven major solar PV power plants which are connected to the grid, seven out of 11 plant use c-Si modules while others produce power using thin film modules. (refer Figure 1). The proposed study deals with the development of a simplified model to predict the energy production of the PV system with CSI technology to be installed in the state of Odisha. The first model, Simplified Model (SFM) takes only three input variables (annual global radiation, ambient temperature, and temperature coefficient) to predict the output response of the PV system. SFM is further simplified, called as Further Simplified model (FSFM), which requires only two input values (annual global radiation and ambient temperature) to estimate the expected output energy production [8]. The aim of the paper is to analyze the preciseness of both the model's forecast for Odisha by validating the response on the generation and weather data collected from SN Mohanty solar power plant which runs with c-Si modules.

## 2. Methodology to develop the Model

The models presented in this section, developed as per schematics shown in Figure 2, use the PV production data obtained for 1195 locations in Odisha using PVGIS method [9], after normalized to per kWp, to predict the outcome of PV plant [8]. The regression analysis by a robust least square method using Matlab curve fitting tool provides the coefficients of the functions for the proposed models. The initial phase of the model development identifies and collects all the main influencing factors that affect the performance of PV module. With these collected parameters and response, SFM is created that takes three input parameters for the prediction of the annual PV production of the plant. Among the three parameters two are variables: Annual incident global radiation and Annual average air temperature, and one is a constant (temperature coefficient). The annual average air temperature is converted into differential module temperature inside the model that is the actual input variable. The SFM is again simplified and another model named FSFM is created. The FSFM takes only two input variables (Annual incident global radiation and Annual average air temperature) for the prediction.

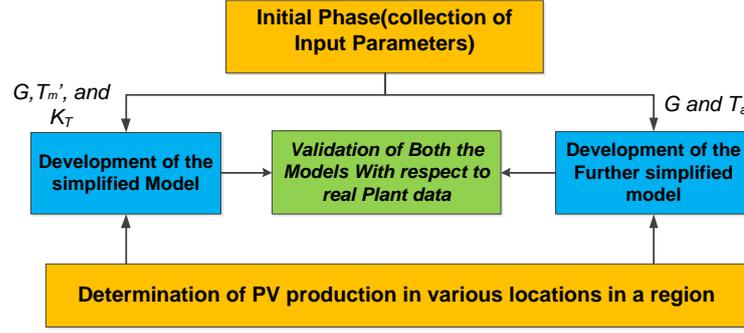


Figure 2 Methodology to implement the SFM and FSFM

### 2.1 Simplified Model:

The SFM is the modified Output function of PVGIS energy rating method. The estimated output of the model is independent of any other influencing factors except for three input parameters: Annual incident global radiation, Annual Average air temperature, and temperature coefficient of the module. Model converts air temperature to differential module temperature defined as the function of the average air temperature and incident global radiation. Based on PVGIS power estimation information the model named [6] SFM proposed is described by equation 3 and equation 4.

$$\frac{P}{P_k}(G', T'_m) = G' + k_1 G' \ln(G') + k_2 G' \ln(G')^2 + k_3 G' T'_m + k_4 G' T'_m \ln(G') + k_5 G' T'_m \ln(G')^2 + k_6 G' T'^2_m \quad (1)$$

$$T_m = T_a + K_T g \quad (2)$$

Where,  $G'$  is the total incident global radiation for the year in kWh/m<sup>2</sup>,  $T_m$  and  $T_a$  are the PV module temperature and average air temperature in °C respectively,  $T'_m (= T_m - 25)$  is the differential module Temperature in °C, and  $g$  is the average incident global radiation per hour in kW/m<sup>2</sup>.  $K_T$  is the temperature coefficient that depends on the mounting structure of the PV module. Based on laboratory measurements for freestanding mounting, the value of the temperature coefficient is 0.035 °C/ (W/m<sup>2</sup>) and based on values referred from literature [6] for the building-integrated system, temperature coefficient is 0.05 °C/ (W/m<sup>2</sup>).

### 2.2 Further Simplified Model:

FMSM use only two inputs: annual incident global radiation and annual average air temperature decreasing the dependents to three fitting parameters. Unlike the previous model, FMSM given by Equation 5 uses the annual average air temperature value directly instead of differential-modular temperature. The range of input values for annual incident global radiation and Plant production remains same for both the model.

$$\frac{P}{P_k}(G', T_a) = aG' + bT_a + c \quad (3)$$

To find the coefficients for both the models Matlab curve fitting tool is used that implements equation 3 or 4 to fit the data.

### 2.3 Curve fitting results:

Using Matlab curve fitting tool, the least square curve fitting method is run on all collected data. All the model equations as provided in the previous section for SFM and FSFM are used to generalize the relationship between the input and output parameters.

**SFM:** Fitting of the non-tracking model uses Least Absolute Residuals (LAR) with Trust-Region [10], a type of robust least squares method. The Robust least squares method utilized for the fitting of the Two-axis

tracking model is Bisquare method with Levenberg-Marquardt [11] algorithm. Fitted parameters presented for four types of model tabulated in Table 1. The six coefficients can be used in the SFM equation to estimate the predictive response of the model. The RMSE value is around 1.2 kWh for non-tracking and 7.1 kWh for tracking model. The  $R^2$  parameter is around 0.99 for all the system that indicates a very close to the response parameter estimated in PVGIS method. The closeness of the predictive response of the non-tracking model is comparatively better than tracking model.

Table 1 Results of fitting of data for SFM

Fitting parameters / System	Non-Tracking PV system		Tracking PV system	
	Freestanding	Building Integrated	Freestanding	Building Integrated
$K_1$	-1.784	-2.989	-2.828	-3.879
$K_2$	0.2405	0.4067	0.3759	0.5192
$K_3$	2.452	2.564	2.482	2.322
$K_4$	-0.5927	-0.6123	-0.5689	-0.5311
$K_5$	0.03481	0.03526	0.03111	0.02884
$K_6$	$7.04 \times 10^{-4}$	$6.3 \times 10^{-4}$	$8.85 \times 10^{-4}$	$6.49 \times 10^{-4}$
RMSE	1.179	1.242	6.977	7.257
$R^2$	0.9989	0.9987	0.9918	0.9902

**FSFM:** Fitting of the FSFM uses Bisquare method with Levenberg-Marquardt algorithm for both the development tracking and non-tracking model. Results with all the fitting parameters for four kinds of the model shown in Table 2. The number of the model coefficient for the FSFM is half of the SFM. The RMSE value is around 5.5 kWh for non-tracking and 7.7 kWh for tracking model. The  $R^2$  parameter for the tracking model is around 0.99 and for the non-tracking model is 0.98. The closeness of the predictive response of the tracking model is better compared to the non-tracking model.

Table 2 Results of fitting of data for SFM

Fitting parameters / System	Non-Tracking PV system		Tracking PV system	
	Freestanding	Building Integrated	Freestanding	Building Integrated
<b>a</b>	0.7015	0.6655	0.7433	0.7049
<b>b</b>	-4.269	-3.664	-5.031	-4.503
<b>c</b>	129.3	103.4	34.06	-8.927
RMSE	5.519	5.413	7.612	7.774
$R^2$	0.9768	0.9756	0.9902	0.9887

### 3. Validation of the model

SN Mohanty solar power plant developed by SN Mohanty and established by Tata BP Solar uses c-Si technology for power production. The power station commissioned on 23<sup>rd</sup> August 2011 near Patapur village, in Cuttack district of Odisha. The capacity of the power plant is 1MW produced through around 4300 panels, with each panel rated 230 Wp. Freestanding mounting implemented in the modules of the plant. The plant uses two inverters each with a capacity of 500 kW. An SCADA system is implemented to record all the plant production related data while a weather monitoring station records all the weather related data, refer photographs shown in Figure 3. For validation of the developed models, all the SCADA data was collected for three years is successfully obtained from the plant. Data included generation data (generated power, generated Energy, Voltage, current, plant on time, grid outage), weather data (ambient temperature, incident direct radiation, wind speed, module temperature). The term grid outage denotes the time in which the grid is down that may be due to failure or low radiation, considered throughout the period of on duration of the plant. Plant on time denotes the operating period of the plant (not the production period) for a day. Out of all the three-year data collected, data analysis for the year 2014-15 is discussed here.

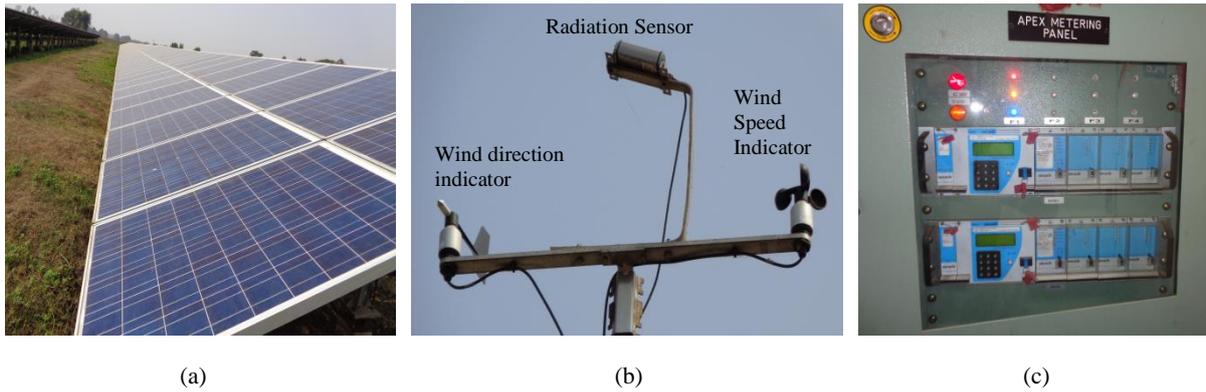


Figure 3 a) SN Mohanty Solar Power plant at Village Patapur, Cuttack (Odisha) b) Weather monitoring station c) Metering panel

The average energy production and radiation of the plant for the year 2014-15 were 3269 kWh/day and 3877 Wh/m<sup>2</sup>. Out of the 365 days of the calendar year, 12 days of radiation and 20 days of ambient temperature data found missing due to failure and data-recording error of the SCADA system as shown in Figure 4. Grid outage occurred in the plant frequently which reduced the energy yield of the plant. The maximum hours of grid outage recorded on 27<sup>th</sup> May in which the production of the plant was down for 12 hours. From 3<sup>rd</sup> to 10<sup>th</sup> October, major failures in SCADA system occurred for which only the temperature and radiation are found to be missing from the record. There are cases of records of erroneous data (19<sup>th</sup> January and 14<sup>th</sup> December) for due to the error in the SCADA system which shows a significant increase in the plant on duration. Odisha is a cyclone prone area and almost every year the cyclone hits the state. On 12<sup>th</sup> Oct 2014 and from 12<sup>th</sup> to 14<sup>th</sup> Oct 2013, the operation of the plant was halted due to the major cyclone occurred in the eastern part of Odisha where the plant is located. The average on-time of the plant for the last three year is 11.16 hours, out of which 1.42 hours is attributed to the grid outage.

The energy yield of the plant typically follows the irregular pattern of the radiation. In the rainy season and overcast conditions, the radiation incident on the module becomes low. The overcast condition significantly reduces the plant output, and torrential rain halts the plant operation from time to time. From the month of June to August in 2014, the energy production of the plant becomes low mostly due to the rainy season. The average daily ambient temperature around the plant varies from 28 °C to 39 °C for the year of 2014. As the temperature and radiation data are proportionally correlated, the parameter affects the plant production indirectly. However, higher value in ambient temperature significantly raises the PV module temperature and reduces the production of the plant. The durations of operation of the plant become greater in the season of summer (April-August) due to increasing in sun hours. On 22<sup>nd</sup> May, the plant operates for 12.88 hours which is maximum for the year.

The daily analysis of the plant for a bright sunny weather and a typical overcast condition are analyzed to check the behavioral characteristics of the plant in optimal and pessimal weather conditions. On 12<sup>th</sup> May 2014, the total radiation of 7684 Wh/m<sup>2</sup> was recorded, which was maximum for the year. Figure 5 shows the array voltage and current profile of an inverter for the day, 12<sup>th</sup> May. On the day, the plant was operated for 12.22 hours with around 8 hours of data being recorded (SCADA failure) out of which 1.75 hours were from the grid outage. The linear pattern of increment in temperature from 30.1 °C to 43 °C evidences the weather to be bright sunny. During the overcast conditions, like as shown for 30<sup>th</sup> June, the total radiation 580 Wh/m<sup>2</sup> was measured. The duration of operation of the plant for the day was 11.83 hours with 3.87 hours of the grid outage. Due to overcast condition, there is no periodic variation of temperature observed. The average temperature for the day was 27.9 °C. The effect of radiation on the production of the plant is clearly visible in the in Figure 5 (a). As shown the pattern output production of the plant proportionally trails the incident radiation except the duration of the grid outage.

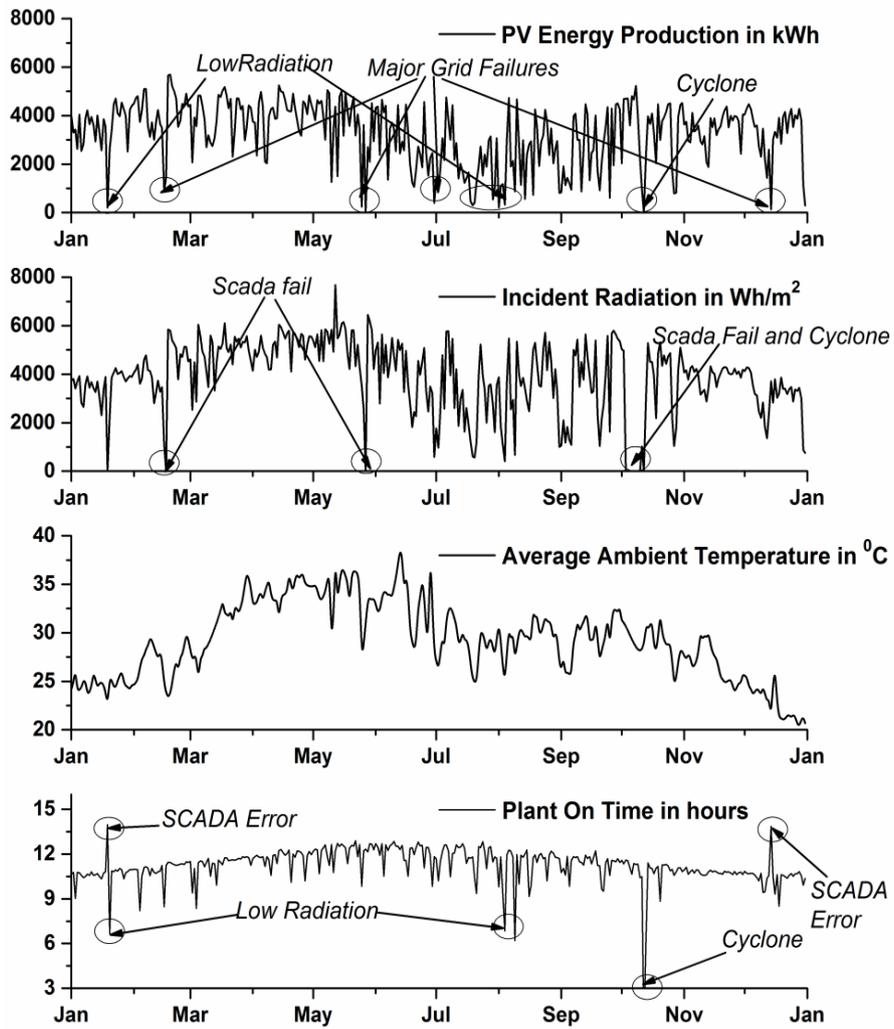


Figure 4 Analysis of the data collected from the SN Mohanty power plant for the year of 2014-15

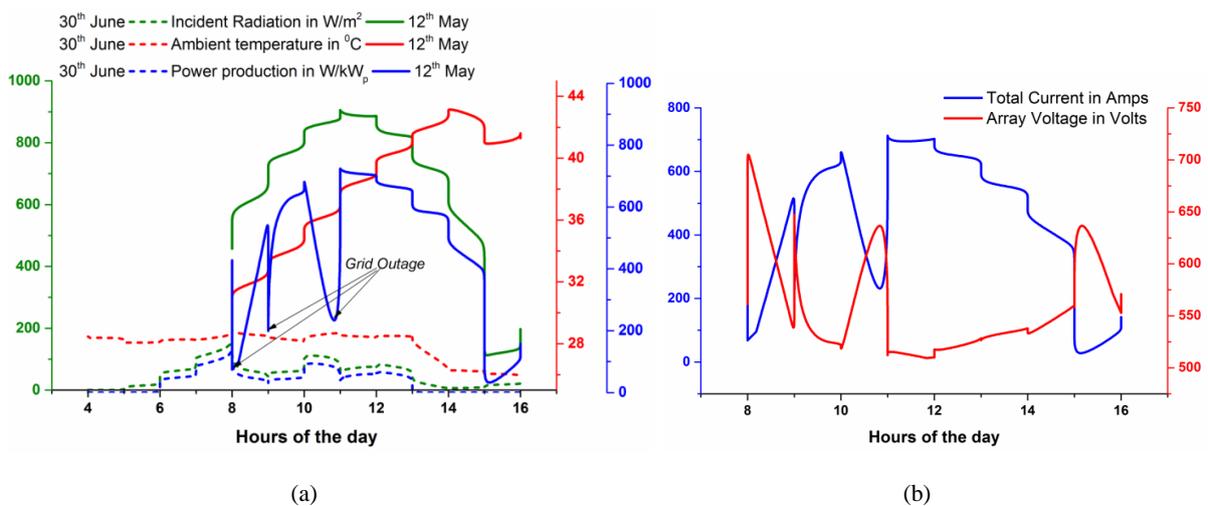


Figure 5 (a) Analysis of the plant in typical overcast conditions (30<sup>th</sup> June 2014) and on bright sunny weather conditions (12<sup>th</sup> May 2014) (b) typical analysis of array voltage and total current of the inverter on 12<sup>th</sup> May 2014

The analysis and comparison of the three years data collected from SN Mohanty 1MW power plant at Cuttack with the predictive results of the model are as shown in Table 3. The error displayed by the two models (SFM, and FSFM), about the expected energy output from the plant, lies in the range of -2.89 to -3.75 % and 0.37 to 1.56 %. Factors like lack of grid connectivity (transmission line failure, RMU tripping, low voltage, and grid failure), a failure of the SCADA (data-recording) system to record the temperature and radiation data, harass weather conditions (cyclone, torrential rains), and lower incident radiation than average expected radiation is responsible for the error.

Table 3 Analysis of the SN Mohanty Solar Power plant data for the validation of the model (D/G-Direct/Global)

Year	Grid outage time (hours)	Plant ON time (hours)	Incident Radiation D/G (kWh/m <sup>2</sup> )	Annual Avg. Ambient Temperature (°C)	Generated Energy per kWp (MWh)	Prediction of the model (MWh)		% Error	
						SFM	FSFM	SFM	FSFM
2013-14	580.38	4036.9	1467/1728	29.2	1.221	1.267	1.216	-3.75	0.37
2014-15	500.35	4103.7	1415/1667	29.1	1.193	1.227	1.174	-2.89	1.56
2015-16	479.15	4075.3	1463/1724	29.5	1.226	1.269	1.212	-3.53	1.10

The analysis and production for the year of 2015-16 are optimal as no cyclone and missing data with relatively less amount of grid outage time is present. The negative error of the SFM indicates the optimistic nature of the model while the positive error of FSFM indicates the pessimistic behavior. Although the prediction of the plant does not consider the parameters like grid failure time ( a subset of grid outage), plant off time (due to a cyclone or inclement weather) still the prediction gives a reasonably accurate response. As tabulated in Table 3, even if the value of incident radiation is more in the year 2013-14 compared to 2015-16, due to additional grid outage hours the production is lower. As the grid outage doesn't affect the prediction of the models, the reduction in the power production increases the error of SFM and decreases the error of FSFM. Even if the plant on time is highest in 2014-15 compare to another year, still the radiation is lowest due to the absence of recorded data for the radiation (SCADA failure). As the radiation data affects the model prediction directly, the prediction of the model decreases. The reduction in the model prediction forces the positive error to increase and negative error to decline in 2014-15. Although the FSM works on only two inputs, still the proximity of prediction is better than the SFM.

#### 4. Conclusion

In the study, two PV energy prediction models, SFM, and FSFM are developed and validated using actual generation and weather data collected from a solar power plant. Both the models show the fair value of the R<sup>2</sup> fitting parameter with RMSE value within an acceptable limit. As per the real-time plant data, the model predicts expected plant output with an average error of -3.4 % (SFM) and 1 % (FSFM). The prediction of the model does not take into account the factors like plant failure time, failure in data recording, and halt in production due to cyclones. Although the region under the study for the model is Odisha; the model can be extended to any area under consideration with suitable corrections for the fitting parameters. To do so, refitting of the model to estimate model coefficients is required with newly collected data for the area under consideration. The implementation of the proposed model is simple, and the user can build own model to estimate the PV output for any location desired.

#### Acknowledgement

The authors would like to thank Mr. S.N. Mohanty, the developer of the SN Mohanty Solar power plant for providing data related to the plant.

## 5. Reference

- [1] International Energy Agency, "Technology Roadmap Solar Photovoltaic Energy," 2014.
- [2] Bridge to India, "India Solar Hand Book," 2015.
- [3] J. N. Mayer, "Current and Future Cost of Photovoltaics," *Long-term Scenarios for Market Development, System Prices and LCOE of Utility-Scale PV Systems*, February 2015.
- [4] Ministry of New and Renewable Energy, "State wise installed solar power capacity," July 2015.
- [5] Science And Technology Department, "Odisha Solar Policy," 2013.
- [6] Joint Research Centre European Commission. PVGIS. [Online]. <http://re.jrc.ec.europa.eu/pvgis/apps4/>
- [7] GeoModel Solar S.R.O. (2010-15) SOLARGIS. [Online]. <http://solargis.info/imaps/>
- [8] Tarai R.K. and Kale P., "Development of the Simplified Predictive Model for the Estimation of Annual PV Energy Production: A Case Study for Odisha," *International Journal of Energy and Environmental Engineering*, 2016, Under Review.
- [9] Tarai R.K. and Kale P., "Development of Rasterized Map using PVGIS for Assessment of Solar PV Energy Potential of Odisha," *International Journal Of Renewable Energy Research*, vol. 6, no. 1, 2016.
- [10] Hendawy Z.M., El-Shorbagy M.A. El-Sawy A.A., "Trust region Algorithm based on local search multi-objective Optimization," *First International Conference on Innovative Engineering Systems*, no. 12, pp. 207-212, 2015.
- [11] Zhao Y., Hou K. Zhang L., "The Research of Levenberg Marquardt Algorithm in Curve Fittings on Multiple GPUs," *International Joint Conference of IEEE*, no. 11, pp. 1355-1360, 2011.