Central Power Research Institute
Vol. 15(1)/52-57, June 2019
ISSN (Print) : 0973-0338
DOI : 10.33686/pwj.v15i1.149517
Power Research

Solar Radiation Forecasting for Moderate Climatic Zone

N. Archana Kesarkar*, K. Jeykishan Kumar and N. Rajkumar

Central Power Research Institute, Bangalore – 560 080, Karnataka, India

Abstract

The challenge with solar energy prediction is that the solar radiation is intermittent and uncontrollable. Energy forecasting can be used to mitigate some of the challenges that arise from the uncertainty in the resource. Weather data was sourced from India Meteorological Department for Bangalore and Chennai location. This paper provides statistical approach to predict the solar power in future. Analysis was done for different predictive models; Multiple Regression Model is used as we have multiple inputs. The results indicate the prediction of solar radiation has better accuracy during higher irradiation period rather than lower irradiation period.

Keywords: Irradiation, Multiple Regression, Solar Forecasting, Solar Radiation

1. Introduction

Renewable energy resources are becoming critical players in the electricity generation sector, primarily due to viability in combating global warming, effectiveness in reducing pollution caused by fossil fuel based generation, and diversifying energy mix to ensure energy security and sustainability. Solar energy is one of the most common types of renewable energy that has grown rapidly over the past decade and is anticipated to grow even faster in the future.

Solar Power is unlike other forms of energy like nuclear power or fossil fuels, the solar output is high on sunny days and less on cloudy days or during winter the sunlight hits the earth at a less direct angle. Making accurate predictions about expected solar power output is paramount to efficiently harvest power from a solar power plant[1].

Solar energy along with wind energy is expected to play a major role in the future energy supply. The fluctuating nature of the energy output from such sources requires reliable forecast information for a successful prediction[2]. Making accurate predictions about expected solar power output is paramount to efficiently harvesting power from a solar plant.

The knowledge about available resources is important to manage the generated energy and improve the efficiency of solar potential. Solar radiation varies with variation in weather factors like temperature, dew points, humidity, wind speed, etc.[3]. Thus it is important to find out the exact factors affecting radiation at a particular location. The forecasting models are continuously being improved to generate more accurate forecasts of solar power[4].

2. Experimental Procedure and Methodology

A new implementation of the weather forecasting model has been specifically designed to better predict solar irradiance[5]. The input parameters used in the datasets of the system are solar radiation, temperature, wind speed, dew and humidity.

*Author for correspondence

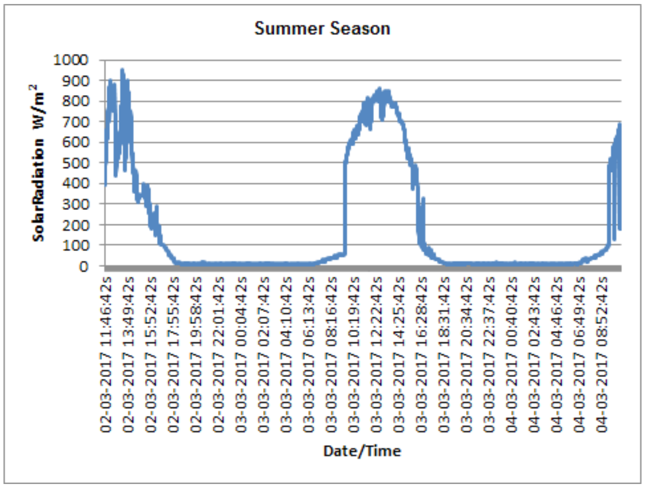

Figure 1. Solar radiation plot for summer season-2017.

The main objective of forecasting is to estimate the dispatch of solar power in rainy, winter and summer seasons. All parameters are recorded in three major seasons such as rainy, winter and summer are separately studied. Here Bangalore location is considered for regression analysis.

From the Figure 1 we see that, as the time of the day increases, the temperature increases which in turn increases the solar radiation as shown in the Figure 1. We can see the radiation curve for different hours and different days of summer season. The radiation peak achieved during summer season is 980W/m².

The Figure 2 shows the typical pattern of radiation plot in Winter Season.

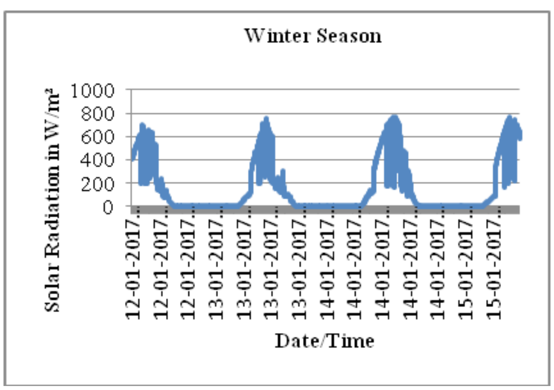

Figure 2. Solar radiation plot for winter season-2017.

Due to less daylight hours in the winter season, the solar radiation obtained will be slightly less compared to summer season. The Figure 2 above shows the variation in solar radiation at different hours and different days of winter season. The radiation peak achieved during the winter season is 750W/m².

The Figure 3 below shows the typical pattern of radiation plot in Rainy Season. The radiation peak achieved during rainy season is 1050W/m².

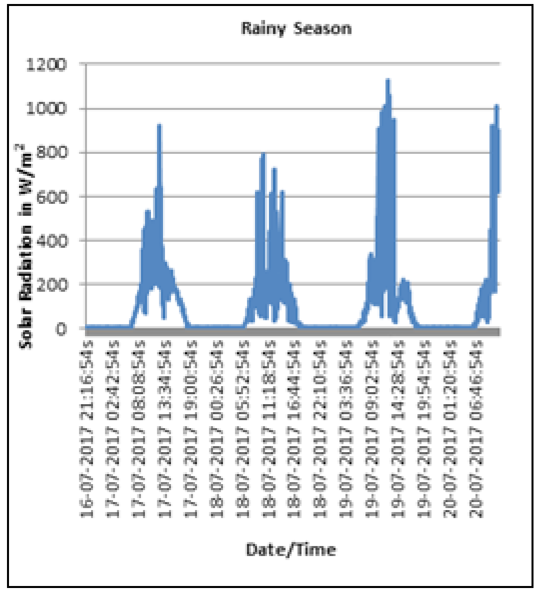

Figure 3. Solar radiation plot for rainy season-2017.

Due to cloud cover in rainy season, there are rapid variation in solar radiation which as shown in the above Figure3. The output power of photovoltaic power plants directly depends on the energy of solar radiation reaching the earth's surface. Due to random nature of the atmospheric condition accurate forecasting of solar energy can be extremely difficult[6].

The solar potential is quantified as watt per square meter (W/m²) on a horizontal surface on ground. Three main components reaching the surface of the earth are Direct Normal Irradiance (DNI), Diffuse Horizontal Irradiance (DHI), and Global Horizontal Irradiance (GHI). The DNI is the direct radiations from the sun, incident on the horizontal surface. The DHI is the radiations diffused through clouds, atmospheric aerosols, and reflected from nearby surfaces. The GHI is the combination of DNI and DHI. At ground, the solar potential is measured and quantified through different ground based sensors like radiometers, pyranometers and pyrheliometers[7].

Statistical analysis based models are accurate methods to estimate solar power generation. Incorrect or incomplete data appear due to inconsistencies in the recording of the data used for the estimation of solar radiation. The data need to be processed using appropriate methods for accurate analysis[8].

The proposed approach helps to forecast solar power in critical seasons. In this project Multiple Regression Analysis is a tool used for renewable energy forecasting. Solar power forecasting is very useful tool for solar power plant and planning for storage of battery which is the feasible measure to stable power output of PV standalone system. Solar Power varies with time and geographical locations and meteorological conditions such as ambient temperature, wind velocity, solar radiation and module temperature. The location of solar PV system is the main reason of solar power variability. The variation in insolation caused by the position of the sun can be determined using geometric relations. However, the dependency on the weather conditions cannot be predicted and hence the need to depend on long term historical records of hourly data for this type of information. These long term data are available for only a limited number of sites or locations. Solar variability depends on deterministic losses and stochastic losses. In case of solar power, deterministic losses can be found accurately but stochastic losses are very uncertain and unpredicted in nature[9]. In this paper, moderate climatic zone is considered. The locations chosen are Bangalore and Chennai. The weather data was obtained from weather monitoring station, present in India Meteorological Department, Pune.

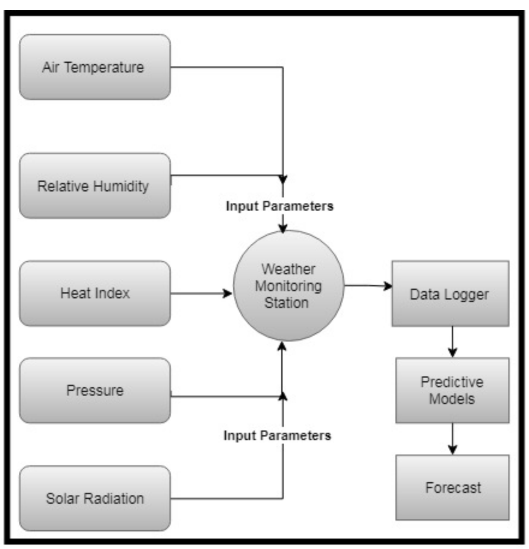

Figure 4. Simplified weather station signal-flow diagram.

The Figure 4 shows the schematic diagram of the proposed solar forecasting model. The weather monitoring station provides different weather parameters for every single minute in a day. The parameters recorded using weather monitoring station and data logger are air temperature, relative humidity, solar radiation, heat index, pressure etc. From the literature survey, different regression models were analysed from which multiple regression method is more suitable for this model to predict solar power in moderate climatic zone. The multiple regression equation proposed is:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 \quad (1)$$

In this equation (1), $\hat{y}$ is the predicted solar radiation in W/m², the independent variables are temperature and heat index. Which are denoted by x_1 and x_2 respectively. The regression coefficients are b_0, b_1 and b_2.

3. Data Analysis using Multiple Regression

A statistical tool in which multiple independent variables are related to a dependent variable. Once we know multiple variables are related to the dependent variable, we can take information about all the independent variables and use it for accurate predictions. This latter process is called "Multiple Regression".

3.1 Regression Analysis

Regression analysis is a common statistical method. Linear regression is one of the most common techniques of regression analysis with two explanatory variables. Multiple regression is a broader class of regressions that encompasses linear and nonlinear regressions with multiple explanatory variables. In our model as we have multiple explanatory variables we go for Multiple Regression. To better analyse the multiple regression the below steps must be considered

a) When we have more than one independent variable, the predictions made are more accurate on dependent variables.
b) Analysis is simple when there are multiple independent variables and when they are not related to each other.
c) Analysis becomes complicated when there are multiple independent variables and when they are related to each other. In this, to make accurate prediction we need to break all the correlation and figure out value of multiple 'R'.

There are three main steps in executing the analysis:

- Enter data on spread sheet.
- Identify independent and dependent variables.
- Specify desired analysis

Table 1. Weather data obtained from meteorological department

X1	X2	Y
Temp	**Heat Index**	**Solar Radiation**
18.8	19.5	199
18.8	19.5	209
18.8	19.4	208
18.8	19.4	193
18.8	19.4	188
18.9	19.5	183
18.9	19.5	181
18.9	19.6	183
18.9	19.6	188
18.9	19.6	195

Excel will display the Regression dialog box. After identifying data fields of the independent and dependent variables. In the input Y range, coordinates for the independent variable has to be entered. In the input X range, coordinates for the dependent variable(s) has to be entered. Excel will produce a standard set of outputs.

Table 2. Regression output

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.998
R Square	0.996
Standard Error	12.098
Observations	48.000

The Table 2 shows the regression statistics where Multiple R indicates the correlation of all the predictor variables with the dependent variable, Multiple R value is 0.998 which means the dependent and the independent variables are highly correlated. R square shows the percentage of dependent variable explained by the independent variables. Standard Error shows that the actual value of dependent variable lies between 1.96*Standard Error around our prediction, Observations indicate the observation considered during the analysis.

Table 3. Error statistics on chosen variables

z	*Coefficients*	*Standard Error*	*P-value*
Intercept	−1486	176	0
Time	6109	2656	0.03
Temperature	−90	39	0.03
Heat Index	62	1	0

The Table 3 shows the error statistics the Coefficients column indicate the regression coefficients such as b_0, b_1, b_2. The Standard Error shows that when a prediction is made it cannot be guaranteed if it is right, but we can say with 95% confident that the prediction values lies between 1.96*Standard Error. The p value shows that there is 5% probability that the b1 value is by chance.

3.2 Proposed Regression Equation

A linear, least-squares regression equation to predict solar radiation, based on temperature and heat index. Since we have two independent variables, the equation takes the following form:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 \tag{1}$$

In this equation (1), $\hat{y}$ is the predicted solar radiation. The independent variables are temperature and heat index, which are denoted by x_1 and x_2, respectively. The regression coefficients are b_0, b_1, and b_2. On the right side of the equation, the only unknowns are the regression coefficients; so to specify the equation, we need to assign values to the coefficients.

Table 4. Coefficient values using data analysis

	Coeffecients	Standard Error	P-value
Intercept	2942.2	845.03	0.01
Temp	−220	69.89	0.02
Heat Index	71.67	45.11	0

Here, we see that the regression intercept (b_0) is 2942.20, the regression coefficient for Temperature (b_1) is -220, and the regression coefficient for heat index (b_2) is 71.67. So the regression equation can be re-written as:

$$\hat{y} = 2942.20 - 220 * \text{Temperature} + 71.67 * \text{Heat Index} \tag{2}$$

From this equation (2) we can use the Temperature and heat index value to predict the solar radiation.

3.3 ANOVA Table

Another way to evaluate the regression equation would be to assess the statistical significance of the regression sum of squares. For that, we examine the ANOVA table produced by Excel:

Table 5. Significant F values on data analysis

	MS	P-value	*F*	*Significance F*
Regression	498434	0.01	3405	0
Residual	146	0.02		
Total		0.00		

In this Table 5 shows the statistical significance of the independent variables as predictors of the dependent variable. The last column of the table shows the results of an overall F test. The F statistic is big, and the p value is small. This indicates that one or both independent variables have explanatory power beyond what would be expected by chance. Like the coefficient of multiple corelation, the overall F test found in the ANOVA table suggests that the regression equation fits the data well.

3.4 Residuals

The difference between the observed value of the dependent variable (y) and the predicted value ($\hat{y}$) is called the residual (e). Each data point has one residual.

Residual = Observed value – Predicted value.

$$e = y - \hat{y} \tag{3}$$

Both the sum and the mean of the residuals are equal to zero. That is, $\Sigma\, e = 0$ and $e = 0$.

The Table 6 shows Predicted Radiation and the residual value.

Table 6. Residual values with predicted radiation

Observation	Predicted Rad.	Residuals
1	381.205	−1.205
2	391.31	−4.31
3	393.097	−3.097
4	409.432	−9.432
5	413.306	−7.306
6	415.094	−5.094
7	416.882	−1.882
8	426.986	−11.986
9	428.774	−8.774
10	351.65	5.35
11	299.448	18.552
12	64.474	26.526
13	80.809	21.191
14	203.064	−3.064
15	260.927	29.073
16	214.956	−1.956
17	50.604	18.396
18	114.698	3.302
19	216.175	−7.175
20	251.201	−10.201

4. Results and Discussions

From the Figure 5 we see that the accuracy is high when the solar radiation is at peak hour and accuracy is less during off peak hours

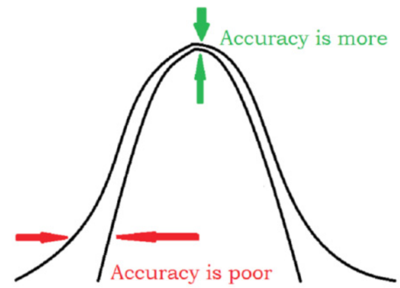

Figure 5. Radiation curve

The Table 7 indicates the actual solar radiation recorded and the deviation from the actual radiation with respect to time.

Table 7. % Error Deviation at different times of the day

Time	Actual Rad.	% Error Deviation
7:49:00 AM	380	−0.317
7:50:00 AM	387	−1.114
7:51:00 AM	390	−0.794
7:52:00 AM	400	−2.358
7:53:00 AM	406	−1.799
7:54:00 AM	410	−1.242
7:55:00 AM	415	−0.453
7:56:00 AM	415	−2.888
7:57:00 AM	420	−2.089
7:58:00 AM	357	1.499
7:59:00 AM	318	5.834
8:00:00 AM	91	29.149
8:01:00 AM	102	20.775
8:02:00 AM	200	-1.532
8:03:00 AM	290	10.025
8:04:00 AM	213	-0.918
8:05:00 AM	69	26.661
8:06:00 AM	118	2.799
8:07:00 AM	209	−3.433
8:08:00 AM	241	−4.233

Example: The actual value of solar radiation on 29th Sep 2019 is 380W/m² at 7:49 AM. The predicated value based on the proposed model is 379W/m2 at 07:49 AM on the same day. This difference or tolerance in measuring accuracy of the proposed model is 0.26% which is really good. But in case of early morning hours or early evening hours, the error observed is 32% of the actual solar radiation.

5. Conclusion and Future Work

This paper proposes a linear multiple regression model of 5% error deviation during peak hours from 10 am to 2 pm and shows a low accuracy of 32% error deviation for predicted solar radiation during the period 7 am to 10 am and 3.30 pm to 6 pm. The future work which could be carried out is in the area of adding more independent variables apart from temperature and heat index like humidity, wind speed, wind direction, etc. Another area of exploration in predicting solar radiation will be in non-linear modelling using multiple variables. And also Regression analysis can be carried out for three separate seasons.

6. Acknowledgement

The authors are thankful to India Meteorological Department, Pune for supplying the meteorological data. This work was funded by the CPRI with ERED support fund. Moreover, partial results were researched projects.

7. References

1. Sanders S, Barrick C, Maier F, Rasheed K. Solar radiation prediction improvement using weather forecasts. 16th IEEE International Conference on Machine Learning and Applications (ICMLA); 2017. https://doi.org/10.1109/ICMLA.2017.0-112
2. Orjuela A, Hernandez CJ, Rivero C. Very short term forecasting in global solar irradiance using linear and non-linear models. IEEE Workshop on Power Electronics and Power Quality Applications (PEPQA); 2017. https://doi.org/10.1109/PEPQA.2017.7981691
3. Hassan M, Ali M, Ali A, Kumar J. Forecasting day-ahead solar radiation using machine learning approach. 4th Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE); 2017. https://doi.org/10.1109/APWConCSE.2017.00050. PMid:28515669. PMCid:PMC5409806
4. Abuella M, Chowdhury B. Solar power probabilistic forecasting by using multiple linear regression analysis. SoutheastCon; 2015. https://doi.org/10.1109/SECON.2015.7132869
5. Haupt S, Kosovic B, Jensen T, Cowie J, Jimenez P, Wiener G. Comparing and integrating solar forecasting techniques. IEEE 43rd Photovoltaic Specialists Conference (PVSC); 2016. https://doi.org/10.1109/PVSC.2016.7749751
6. Snegirev D, Eroshenko S, Khalyasmaa A, Dubailova V, Stepanova A. Day ahead solar power plant forecasting accuracy improvement on the hourly basis. IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus); 2019. https://doi.org/10.1109/EIConRus.2019.8657024
7. Hussain S, Alili A. Day ahead hourly forecast of solar irradiance for Abu Dhabi, UAE. IEEE Smart Energy Grid Engineering (SEGE); 2016. https://doi.org/10.1109/SEGE.2016.7589502
8. Serttas F, Hocaoglu F, Akarslan E. Short term solar power generation forecasting: A novel approach. 2018 International Conference on Photovoltaic Science and Technologies(PVCon); 2018. https://doi.org/10.1109/PVCon.2018.8523919
9. Vijay V, Singh VP, Bhatt DM, Chaturvedi. Generalised neural network methodology for short term solar power forecasting", 13th International Conference on Environment and Electrical Engineering (EEEIC); 2013.