2016 7th International Conference on Intelligent Systems, Modelling and Simulation

Data Analysis to Generate Models Based on Neural Network and Regression for Solar Power Generation Forecasting

Tushar Verma, A. P. S. Tiwana and C. C. Reddy
Department of Electrical Engineering
Indian Institute of Technology Ropar
Rupnagar, India - 140001
e-mail: tusharv@iitrpr.ac.in

Vikas Arora and P. Devanand
Smart Grid Group
Tata Power Delhi Distribution Limited
Delhi, India
e-mail: p.devanand@tatapower-ddl.com

***Abstract* — This paper presents methods for forecasting solar power generation by a solar plant. Solar power generation depends primarily on relative position of sun and some extrinsic as well as intrinsic factors. Extrinsic factors such as cloud cover, temperature, wind speed, rainfall and humidity have been used with intrinsic ones such as degradation of solar panels as inputs for proposed techniques for generation forecasting. The authors have used multiple linear regression, logarithmic regression, polynomial regression and artificial neural network method on the data of past one year (January 2014-December 2014) for creation of forecasting models. These forecasting models are then compared on the basis of their accuracy to forecast the solar generation.**

Keywords - Artificial Neural Network; Regression; Solar Power; Forecasting

I. INTRODUCTION

Solar energy from photovoltaic (PV) panels is a renewable source of energy and its intensity varies with the weather and sun's position with respect to the panels. Today the electricity demand is met through varied resources (solar, hydro, thermal, nuclear, wind energy etc.). To optimize the production of energy, we must know beforehand, how big a fraction of demand a particular energy source can meet. So we must be able to predict the future production. In the case of solar energy, it varies day to day with weather and relative solar position. We must be able to forecast how much we can produce with solar energy as our power source on a given day. The remaining demand would be met by other resources. Also, since electrical energy can't be stored on utility scale, power industry tries to match the production with the demand which is a fairly complex process. A small but significant chunk of demand is met through solar power; hence the need arises to predict beforehand, the amount of solar energy generation.

Forecasting has been studied extensively using a mixture of statistical techniques and data history analysis. Vikas Pratap Singh [1] has studied dependence of solar power generation on solar radiation, ambient temperature and wind velocity. The authors applied Fuzzy theory, Adaptive Neuro- Fuzzy interface system, Artifical neural network and Generalized neural network to find a relationship justifying this dependence. Similarly Nian Zhang [2] has applied artificial neural network on past solar radiation and solar energy data for forecasting. The above techniques utilize a mixture of statistics and neural networks. D.K. Chaturvedi [3] has proposed generalized neural wavelet method for load forecasting. Jose G [4] has again applied artificial neural network multilayer perceptron to forecast solar radiation. Ghanbarzadeh [5] has taken the meteorological data to forecast solar power generation. He has used air temperature, relative humidity and sunshine hours as the parameters. Al-Messabi [6] has forecasted solar power generation using dynamic neural networks which provides time series predictions with good reliability. Abuella [7] has used multiple linear regression to make probabilistic forecasts of solar energy. Mori [8] discusses the ways to select input variables for any forecasting model.

The authors of this paper propose a new technique for forecasting by incorporating some additional parameters also (wind speed, cloud cover, temperature, rainfall, humidity, azimuthal angle, elevation angle and degradation factor of solar panels). By applying multiple linear regression, logarithmic regression, polynomial regression and artificial neural network method on the recorded values of these parameters, the authors have developed four models to forecast solar power generation.

II. METHODOLOGY

A. *Extrinsic factors:*

The authors have used the values of solar generation (Kwh) through solar panels installed by Tata Power. The values of weather parameters and its corresponding solar power generation for past one year were analyzed. Along with the weather data, sun's position variations during a day was also taken into account. Finally four models using linear regression, logarithmic regression, polynomial regression and artificial neural network (ANN) were made.

Linear regression maps the data inputs as a linear equation of independent variables (weather parameters and sun' position defining angles). Logarithmic regression maps the data inputs as a logarithmic equation of independent variables. Polynomial regression maps the data inputs as a

2166-0670/16 $31.00 © 2016 IEEE

DOI 10.1109/ISMS.2016.65

polynomial equation of independent variables. ANN model simulates biological neural network to learn the relation between independent variables and dependent variables.

B. *Intrinsic factors:*

There are some intrinsic factors also to be taken into account. Primary factor is the degradation of solar panels with time. This results in decrease in efficiency of solar panels. Most of the times manufacturer gives a standard value by which the efficiency of solar panel degrades.

In extrinsic factors temperature, cloud cover, wind speed, rainfall, humidity, azimuthal angle and elevation angle of sun are considered. The effects of these parameters on solar power generation is explained below.

A. Temperature (T)

Photovoltaic, or solar panels that produce electricity, are affected by their operating temperature. Temperature and other environmental factors can reduce efficiency.

B. Cloud Cover (C)

This is by far the most damaging factor among all weather parameters. Cloud cover in sky can reduce the amount of sunlight falling on solar panel by significant amount and hence the production

.

C. Wind speed (W)

At higher wind speeds, the air temperature could decrease and so does the solar cell operating temperature. That means that wind can help a solar PV system perform more efficiently

D. Humidity (H)

It is found that if the relative humidity is low then the efficiency is higher and vice versa. This happens because low water vapor content leads to higher values of solar flux because there is no water vapor to absorb the sunlight.

E. Rainfall (R)

There is a combined effect of rainfall on efficiency of solar panel. Rainfall decreases the temperature of panels but at the same time it increases the humidity.

F. Elevation Angle (E) and Azimuthal Angle (A)

Maximum power is generated when the sun rays falling on the solar panels fall perpendicularly on it. Sun's position and the ray's direction is decided by two angles i.e. Elevation angle and Azimuthal angle. These angles increases at first then reaches a maximum and then decreases. Effect of these will be decided on the basis of the angle the rays make with solar panel. If the rays make theta (θ) angle with panel. Then normal component will be Sin θ and parallel component will be Cos θ. We are interested only in normal component so as θ increases power increases, reaches a maximum and then starts decreasing.

The actual solar power generation on any given day generally varies as shown in Fig 1.

III. Results and Discussions

The authors have developed the governing equations for each regression technique. In these equations, the output variable is solar generation power (*Power*). The variables on which it depends are *T*, *C, W, H, R, E, A*.

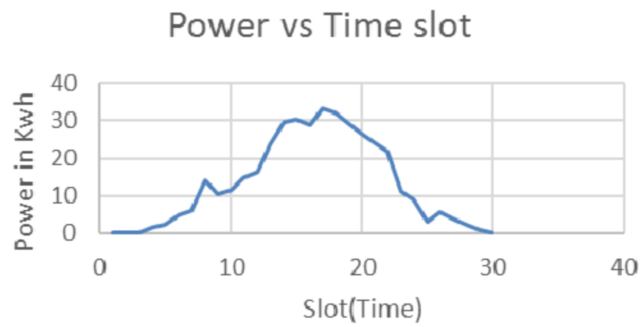

Fig1: Variation of Power generation with time. A day is divided into 48 slots.

The Linear regression finds a linear relationship between the dependent variable and independent variables. Logarithmic regression aims to find a logarithmic nonlinear equation between the dependent and independent variables. Similarly polynomial regression aims to find a polynomial equation between these variables. These equations are found as written below:

A. Linear regression

After applying linear regression, *W* and *H* were found to be insignificant and hence these variables were not present in the equation:

$$Power = (0.06)\times T + (-0.028)\times C + (-7.94)\times R + (23.06)\times E + (-1.31)\times A$$

B. Logarithmic regression

After applying logarithmic regression, *C* was found to be insignificant and hence it was not present in the developed equation.

$$Power = \ln(T)\times(0.01) + \ln(H)\times(0.005) + \ln(R)\times(0.580) + \ln(W)\times(0.047) + \ln(E)\times(4.22) + \ln(A)\times(0.288)$$

C. Polynomial regression

After applying polynomial regression first and third powers of *T, W, H, C, R, E, A* as well as fourth power of *R* were found to be insignificant. The equation developed is:

$$Power = (T^2)\times(-3.52\times10^{-5}) + (H^2)\times(-2.48\times10^{-6}) + (C^2)\times(8.6\times10^{-7}) + (R^2)\times(4.51) + (W^2)\times(0.010) + (E^2)\times(40.947) + (A^2)\times(0.470) + (T^4)\times(1.94\times10^{-6}) + (H^4)\times(4.412\times10^{-8}) + (C^4)\times(4.543\times10^{-8}) + (W^4)\times(-0.001) + (E^4)\times(-21.080) + (A^4)\times(0.083)$$

D. ANN model

While developing ANN the authors have used Levenberg-Marquardt algorithm (LMA). The details of structure of neural network is as shown in Table I.

TABLE I. STRUCTURE FOR NEURAL NETWORK

Network parameters	Value
Number of input variables	7
Number of output variables	1
Number of input layers	1
Number of hidden neurons	10
Number of hidden layer	1

After training the neural networks. The performance plots are as shown in Figure 2. The comparison of these models is shown in Table II. In figure 3, the comparison between actual and the predicted solar generation using ANN is shown. This comparison was done for 24th June, 2015. Mapping is defined as the percentage of points on which the proposed model give accurate results.

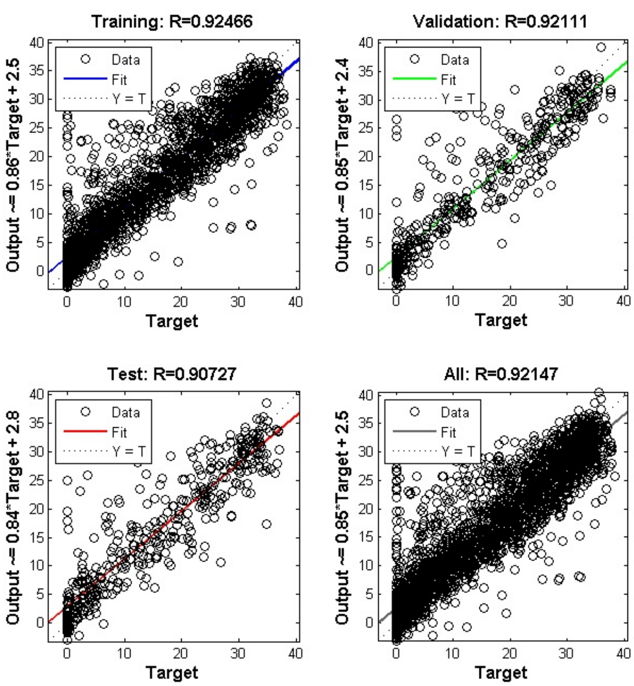

Figure 2. Predicted output vs Target Output

TABLE II. MAPPING PERCENTAGE OF EACH MODEL

Model	Mapping Percentage(R)
Linear Regression	74.4%
Logarithmic Regression	47.4%
Polynomial Regression	75.1%
ANN	92%

TABLE III. ERROR IN PREDICTING SOLAR GENERATION FOR 24TH JUNE, 2015

Model name	Practical Error per 100KWh
Linear Regression	6%
Logarithmic Regression	15%
Polynomial Regression	6.1%
ANN	3%

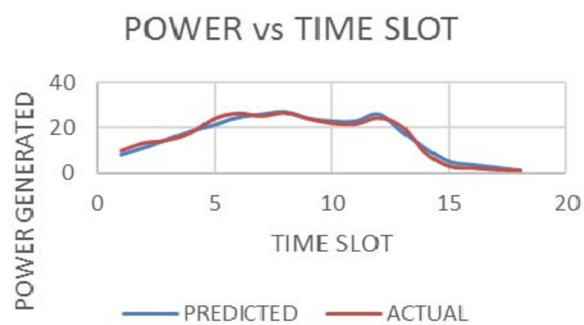

Figure 3. Actual generation vs Predicted generation by ANN Model for 24th June, 2015

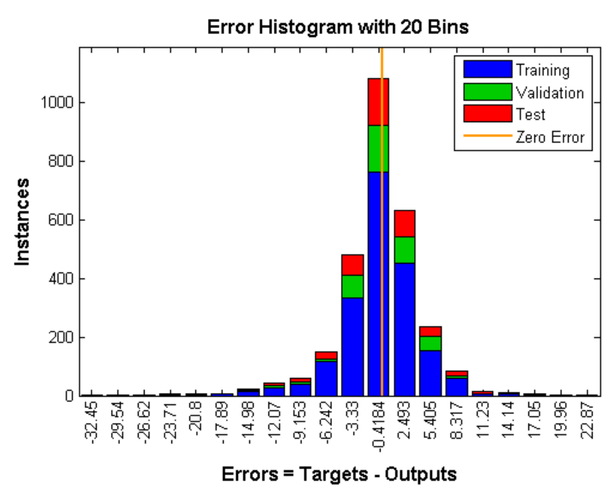

Figure 4. Plots the error histogram of ANN model. This is the error per 100 Kwh.

IV. INCLUSION OF INTRINSIC FACTORS

Solar panels degrade with time. This is due to hostile weather conditions like rainfall, wind which degrade the efficiency of solar panels. Most of the time this degradation factor is provided by the manufacturer. So if the model which we have developed based on extrinsic factors forecasts output 'P' after n years, then the actual output would be

$$Power = \left(1 - \left((x \div 100)^n\right)\right) \times P$$

Where "P" is the power forecasted by forecasting models. This takes care of both extrinsic as well as intrinsic factors.

V. CONCLUSIONS

So it is found that ANN model maps most number of data points. So it is the most reliable method for forecasting. These four models were employed to forecast solar generation for 24th June, 2015 at TPDDL Delhi, India. The comparison of results is as shown in Table 3. It can be seen that ANN gives the minimum error and is most reliable technique.

ACKNOWLEDGEMENTS

The authors wish to thank TPDDL for the valuable data provided for analysis and also acknowledge the help and cooperation extended by TPDDL during the course of this project.

REFERENCES

[1] Singh, V.P.; Vijay, V.; Bhatt, M.S.; Chaturvedi, D.K., "Generalized neural network methodology for short term solar power forecasting," in *Environment and Electrical Engineering (EEEIC), 2013 13th International Conference on* , vol., no., pp.58-62, 1-3 Nov. 2013

[2] Nian Zhang; Behera, P.K., "Solar radiation prediction based on recurrent neural networks trained by Levenberg-Marquardt backpropagation learning algorithm," in *Innovative Smart Grid Technologies (ISGT), 2012 IEEE PES* , vol., no., pp.1-7, 16-20 Jan. 2012

[3] D. K. Chaturvedi. "Short-Term Load Forecasting Using Soft Computing Techniques" International Journal of Communications, Network and System Sciences 3.1 (2010): 270-279)

[4] Hernandez-Travieso, J.G.; Travieso, C.M.; Alonso, J.B.; Dutta, M.K., "Solar radiation modelling for the estimation of the solar energy generation," in *Contemporary Computing (IC3), 2014 Seventh International Conference on* , vol., no., pp.536-541, 7-9 Aug. 2014

[5] Ghanbarzadeh, A.; Noghrehabadi, A.R.; Assareh, E.; Behrang, M.A., "Solar radiation forecasting based on meteorological data using artificial neural networks," in *Industrial Informatics, 2009. INDIN 2009. 7th IEEE International Conference on* , vol., no., pp.227-231, 23-26 June 2009

[6] Al-Messabi, N.; Yun Li; El-Amin, I.; Goh, C., "Forecasting of photovoltaic power yield using dynamic neural networks," in *Neural Networks (IJCNN), The 2012 International Joint Conference on* , vol., no., pp.1-5, 10-15 June 2012

[7] Abuella, M.; Chowdhury, B., "Solar power probabilistic forecasting by using multiple linear regression analysis," in *SoutheastCon 2015* , vol., no., pp.1-5, 9-12 April 2015

[8] Mori, H.; Takahashi, A., "A data mining method for selecting input variables for forecasting model of global solar radiation," in *Transmission and Distribution Conference and Exposition (T&D), 2012 IEEE PES* , vol., no., pp.1-6, 7-10 May 2012