IFERP

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 5, Issue 2, February 2018

Simulated Performance of a Photovoltaic Module: A Comparison of ANN and Regression Based Models

[1] Sandhya Prajapati, [2] E. Fernandez
[1] Research Scholar, [2] Associate Professor
[1][2]Electrical Engineering Department, IIT Roorkee, Roorkee, Uttarakhand

Abstract **- In this paper, we attempt to compare the output prediction accuracy of two models – the ANN model and the Linear Regression model as applied to the case of the simulated performance of the PV module. The input variables for each model are Voc, Isc, Insolation level, Temperature while the output variable is Pmax. ANN models after trial show that the model with 4-10-1-1 configuration is suitable. The linear regression model is a multiple regression model involving the same variables. The data base for this study is derived from various available manufacturers' data, together with some experimental outputs from a few research reports. The results of this analysis appear to indicate that the ANN model is superior in prediction performance.**

Keywords: ANN, Comparative analysis, PV module, Regression

I. INTRODUCTION

Photovoltaic (PV) power system has become an important renewable energy resource due to its several benefits that include cost free, pollution-free and inexhaustible supply in nature. Large scale PV power systems have been commercialized in numerous countries due to their substantial long term benefits and promotion of sustainable "green energy". Financial initiatives are provided by most governments to promote this sustainable form of green energy. However, due to the high investment cost on PV modules, it becomes necessary to size systems and performance by making use of optimal utilization of the available solar energy. Much of this information can be obtained by experimental and field work. However, this is expensive and generally not flexible in exploring alternate options of design and operation. Simulation of PV systems offer a more convenient alternative in terms of time saved, flexibility and expense. Hence, a precise and reliable simulation of the designed PV systems prior to installation becomes necessary.

The most important component that affects the accuracy of the simulation is the PV cell model. Modeling of PV cell involves the estimation of the performance curves to emulate the real cell under various environmental conditions. The most popular approach is to utilize the electrical equivalent circuit, which is primarily based on diode. Many models have been proposed by various researchers.

II. LITERATURE REVIEW

Bikaneria et al [1] examined the simulated performance of the single diode PV model in three stages- an ideal stage with no parasitic resistances, then the model with additional series resistance and finally the model with both series and shunt resistances. They found that the model with the addition of series resistance only gave best performance results. A similar conclusion was arrived at by Azzouzi et al [2]. Benghanem and Alamri [3] examined different explicit models with different number of parameters for simulating the PV cell. They also proposed experimental methods to extract the series and shunt resistance parameters that are present in the model. Other attempts to examine the results of simulations using the single diode model were reported by Rodigues et al [4], Kim and Choi [5]. To improve the accuracy of the model simulations, a second diode has been used by some researchers [6]-[7].The single diode model is simple and easy to implement, whereas the double diode model has better accuracy which validates a more precise forecast of the PV systems performance. Ahmad et al [8] investigated the performance of the single and double diode models on a comparative basis. The MATLAB tool is used to serve this purpose. They found that the double diode model shows superior performance as compared to the single diode model.

In most of the analytical models, there is a need to estimate accurately certain model parameters such as the series and shunt parasitic resistances and the ideality factor. A proper estimation of these will yield a

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 2, February 2018

reasonably good model that can be applied for performance evaluation under varied operating conditions. In the literature, a variety of methods have been evolved to accomplish this. These methods include the, double exponential method [9], conductance method [10] and pattern search technique.[11],[12] The problem that a researcher faces is to estimate these parameters accurately. This will affect the accuracy of the predicted output. This difficulty can be avoided by use of alternative modeling approaches.

There are two common alternative approaches for modeling PV modules:
(a) ANN based approach and
(b) Regression based approach
The Artificial Neural Network (ANN) concept requires a minimal of input data. Artificial Neural Networks (ANNs) are being applied to study engineering systems that are governed by non linear relationships. This is also true of the solar cell on account of its non linear I-V characteristics. Some reports that deal with the application of ANNs are available in the literature [13]-[20]. In these reports, the use of the ANN model has been demonstrated It is more suitable and better than the use of the conventional diode equivalent circuit model. Further, in most of these cases the back propagation algorithm has been generally adopted but for this case the layer recurrent model gives the efficient result.
The second approach to be used involves the use of linear regression model. This is a statistical approach that obtains the best coefficients required to model the system with a minimum cumulative least square error. The accuracy obtainable for the prediction will depend on the value of the parameter R2. This value lies between 0 and 1, and the nearer is this value to 1, the better is the predicted accuracy of the regression model designed for the PV module. In the literature, some researchers have tried to apply linear regression to model PV systems. Some of these studies are reported in [21]-[25].
A problem that has been faced by some researchers is the relative comparison of the prediction accuracy of the ANN vis-a' vis the linear regression model. Various studies on different subjects have been carried out to examine this issue. [26]- [28] In this paper we attempt to examine this issue taking the output prediction of a PV module as the basis of comparison.

III. THE ANN MODEL

The ANN is a modeling method to simulate complex systems (especially nonlinear systems), using learning algorithms involving a set of training data. The relationships so extracted are not stored as equations, but are distributed throughout the network in the form of connection weights between neurons the basic units that comprise the ANN network. The structure of ANN is important factor that influences the learning performance of networks. The best performance generally shown by a trained ANN network is attributed to the feed-forward back propagation neural network (FFNN) which has the capability of providing the best mapping of the input–output data matching, Hence they are the most commonly used type of ANN for most studies. The same will be used in the present study.
Fig1 shows the basic structure of a typical multi input-multi output back propagation ANN model.

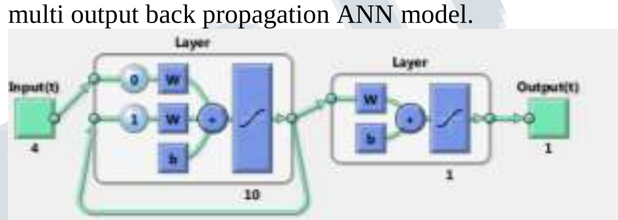


Fig 1: Schematic diagram of a Layer Recurrent ANN network [13]

The inputs are $x_1, x_2, \ldots\ldots\ldots x_n$, while the outputs can be $y_1, y_2 \ldots\ldots\ldots y_n$.
A layer recurrent network is usually arranged in at least three layers: an input layer, the one or more hidden layer(s), and an output layer. The number of input neurons is equal to the number of independent variables ($x_1, x_2, \ldots\ldots\ldots x_n$), while the output neuron(s) represent the dependent variable(s)[$y_1, y_2 \ldots\ldots\ldots y_n$]. The number of hidden layer and neurons within each layer can vary depending on the size, complexity and nature of the dataset.
The neurons receive the inputs and; the sum of their weighed inputs is applied to a non-linear activation function such as hyperbolic tangent (TANH) activation function to produce output as governed by the relationships:

$$y_j = f\left[\left(\sum_i w_{ij} x_{ij}\right) + b_j\right]$$

$$f(x) = \frac{e^{2x} - 1}{e^{2x} + 1}$$

where *yj* is the output of the neuron, *wij* is the synaptic weight coefficient of the x_{ij} th input of the neuron, and *bj* is the bias, if involved.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 2, February 2018

To objectively evaluate the performance of the networks, different statistical indicators may be used. Some of these indicators are mean squared error (MSE), coefficient of determination (R^2) and mean absolute per percentage error (MAPE) Expressions for these are shown below:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_{estimated} - Y_{target})^2$$

$$R^2 = \frac{\sum_{i=1}^{n}(Y_{estimated} - Y_{target})^2}{\sum_{i=1}^{n}(Y_{estimated} - Y_{mean})^2}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_{estimated} - Y_{target}}{Y_{target}}\right|$$

The essential steps in designing an ANN model are:
1. Creating the network
2. Configuring the network
3. Initializing the weights and biases
4. Training the network
5. Validating the network
6. Using the network

IV. THE LINEAR REGRESSION MODEL

Linear Regression is an approach for modeling the relationship between a scalar dependent variable y and one or more explanatory variables (or independent variables) denoted X in a linear frame [29]. If one explanatory variable is involved the regression is termed as simple linear regression. For more than one explanatory variable, the process is called multiple linear regression. A multiple regression model is given by the general expression:

$$y_i = \beta_0 1 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \varepsilon_i = \mathbf{x}_i^{\top}\beta + \varepsilon_i, \qquad i = 1, \ldots, n,$$

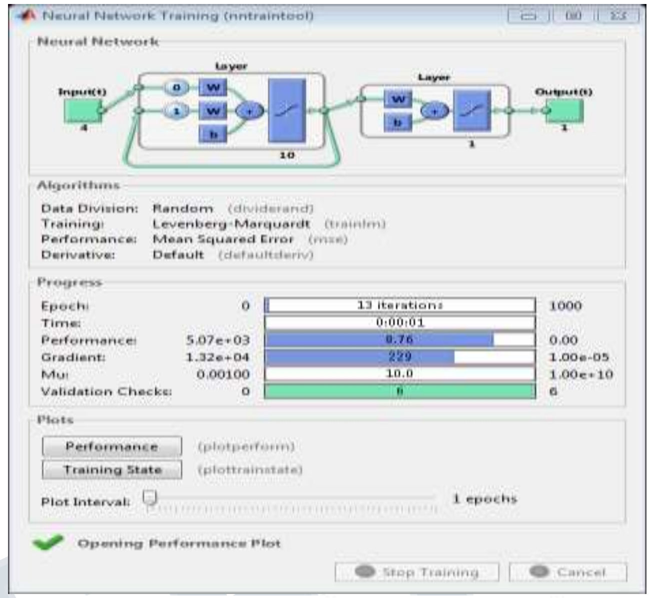
Where: y_I denotes the dependent variable and x_{in} th independent variables. ε_i denotes the error term involved.

Often these equations are grouped and collectively represented in vector form by the expression:

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

Where:

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$



$$X = \begin{pmatrix} \mathbf{x}_1^{\top} \\ \mathbf{x}_2^{\top} \\ \vdots \\ \mathbf{x}_n^{\top} \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix}$$

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}, \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Ordinary least squares (OLS) is the most commonly used estimator being conceptually simple and computationally straightforward.

The OLS method minimizes the sum of squared residuals, and leads to a closed-form expression for the estimated value of the unknown parameter β:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{\top}\mathbf{X})^{-1}\mathbf{X}^{\top}\mathbf{y} = \left(\sum \mathbf{x}_i\mathbf{x}_i^{\top}\right)^{-1}\left(\sum \mathbf{x}_i y_i\right).$$

The estimator is unbiased and consistent if the errors have finite variance and are uncorrelated with the regressors, i.e.

$$\mathrm{E}[\mathbf{x}_i\varepsilon_i] = 0.$$

The degree of “fit” of the model is evaluated in terms of the R2 statistic value that lies between 0-1. A value closer to unity indicates a good fit to the input data.

V. COMPARISON OF SIMULATED RESULTS OF MODELS

In order to test the suitability of the two models(ANN and the Linear regression models) we have adopted the following functional forms:

The ANN model was structured as a three layer model with an input of four variables (Voc, Isc, Insolation level and Temperature) The output to be simulated is the power at maximum power point (P max) A single hidden layer of 10 nodes is used. The configuration is thus 4-10-1(4 input layer nodes, 10 hidden layer nodes and 1 output layer node.)

The linear regression model was structured as a linear form given below as:

$$y = \beta 0 + \beta 1 * (Voc) + \beta 2 * (Isc) + \beta 3 * (Insolation) + \beta 4 (Temperature)$$

The data for developing the models was obtained from a created database involving different PV modules specifications as available from various manufacturers. Additionally some research papers providing output results of experimental investigations have also been consulted. Matlab Neural Network Toolbox and MS Excel were used to obtain the results for the models. The results so obtained are given as follows:

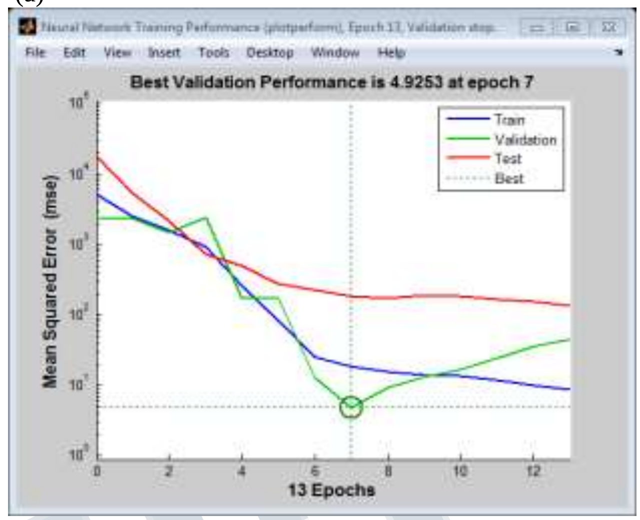
A Results for the ANN model

Table 1 shows the main results obtained for the developed ANN model.

Table 1 Results of the ANN Model

S.No	Feature	Value		
1.	R^2	0.92448		
2.	Standard Error	18.9037		
3.	F Statistic	91.8066		
4.	***Coefficients of regression terms***	***Coefficient Value***	***Standard error***	***t- statistic***
	βo	-194.114	31.0356	-6.2545
	$\beta 1$	5.027044	0.549786	9.143629
	$\beta 2$	32.08994	2.529024	12.68867
	$\beta 3$	-0.01815	0.024088	-0.7536
	$\beta 4$	-0.17474	0.263908	-0.66214

(a)



(b)

Fig 2 Snapshots of the ANN training performance

B Results for the Regression Model

Table 2 shows the main results obtained for the regression model

Table 2 Results of the ANN Model

S. No	Feature	Details
1.	No. of Epochs to convergence	13 iterations
2.	Time of Convergence	00.00.01
3.	Performance Accuracy	8.76
4.	Gradient	229
5.	Mu	10.0
6.	Training Algorithm	Levenberg-Marquardt
7.	Performance Indicator	MSE

C Comparative testing of Accuracy of the Two Models

In order to test the relative accuracy of the two different models for the PV module, we considered 20 random data input sets from the database and compared the outputs obtained from the ANN model and the regression models. The testing dataset was identical for both models. Table 3 shows the relative accuracy shown in the comparison.

Table 3 Relative Accuracy of Models

S. No	**Actual Output Value(Watts)**	**ANN model % error in Output**	**Regression model % error in Output**
1	280.858	0.0893334	2.307339
2	310.709	2.983724321	4.732066

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 2, February 2018

3	205.117	1.054958877	0.494618
4	216.062	1.394090585	1.204048
5	199.124	0.321558426	6.229609
6	214.361	0.713655936	5.819358
7	196.0161	6.194644215	4.945133
8	305.3214	0.292838956	3.703133
9	320.178	3.369906739	4.989345
10	315.033	2.149647815	4.199893
11	158.064	0.296525458	1.265379
12	230.124	1.249282995	2.678652
13	239.6955	0.237301076	3.14199
14	220.129	0.219598508	4.735998
15	240.787	0.278835651	2.607599
16	235.64	0.640044135	1.896489
17	182.484	14.68610947	0.080425
18	128.25	0.189551657	57.42627
19	152.6	0.967758847	12.04589
20	203.32	0.777001771	1.525619
Mean % Error		**1.905318442**	**6.301442**

D. Discussion of the Results

The results of Table 3 indicate that the ANN model has a much lower degree of % error in its output (1.9053%)as compared with the regression model(6.3014%). One possible reason for this observation is that ANN are better suited to non linear mappings while linear regression models can handle with reasonable accuracy input- output relationships that are basically linear in form in the PV module. The output power is a non-linear function of the voltage. Also, the voltage and currents in the PV module, share a non-linear relationship. Thus, in essence, the linear model has a poorer fit to the data than the ANN model. This is reflected in the obtained results. It would be better to examine the nature of the relationships and accordingly use some non linear form of regression, should the data trends confirm to non linear patterns.

CONCLUSION

The paper attempted an investigation to determine the relative accuracy of two types of models as applied to the case of the PV module output power predictions. These models are the ANN model and the linear regression model respectively. The data for the models was derived from a database created from PV module specifications as published by different manufacturers. Some data was also obtained from experimental result reports of certain research papers.

ANN model was trained with data from the created database using Matlab Neural Network Toolbox while the regression model was obtained using the MS Excel software available in MS Office package.

The results of the study indicate that the ANN model shows a far lesser % error than the linear regression model. It is felt that this observation may be probably due to the better suitability of the ANN for non linear data relationships, which is the case in the PV module input-output relationships. The use of non-linear regression is suggested for lower % errors in the output.

REFERENCES

[1] J. Bikaneria et al.,“Modeling and simulation of PV cell using one-diode model,” IJSRP, vol. 3, issue 10, 2013.

[2] M Azzouzi, D. Popescu and M Bouchahdane, “Modeling of electrical characteristics of photovoltaic cell considering a single diode”, Journal of clean energy technologies, Vol.4 (6), pp. 414-419, 2016.

[3] Mohd. Benghanem and S.N. Alamri, “Modeling of PV module and experimental determination of serial resistance”, Journal of Taibah University for Science, Vol 2, pp 94-105, 2009.

[4] E.M.G. Rodrigues , R. Melício, V.M.F. Mendes and J.P.S. Catalão., “Simulation of a solar cell considering single-diode equivalent circuit model,” in Proc. International Conference on Renewable Energies and Power Quality, 2011.

[5]. Wook Kim, Woojin Choi “A novel parameter extraction method for the one-diode solar cell model” Solar Energy 84, pp.1008–1019, 2010.

[6] B.Alsayid,“Modeling and simulation of photovoltaic cell/module/array with two-diode model,” IJCTEE, vol. 1, no. 3, 2012.

[7]Gow J.A, Manning C. D.,“Development of a photovoltaic array model for use in power electronics simulation studies,IEE Proc Electrical Power Applications, pp.146:193–200,1999

[8]Tanvir Ahmad, Sharmin Sobhan, Mohd. Faysal Nayan, “ Comparative Analysis between Single Diode and Double Diode Model of PV Cell: Concentrate Different Parameters Effect on Its Efficiency”, Journal of Power and Energy Engineering, Vol. 4, pp.31-46, 2016.

[9] Yadir, S., Assal, S., El Rhassouli, A., Sidki, M., Benhmida, M., "A new technique for extracting physical parameters of a solar cell model from the double exponential model (DECM)", Optical Materials vol. 36, pp. 18-21, Nov. 2013.

[10] Z.Ouennoughi, and M. Cheggar, “A simpler method for extracting solar cell parameters using the conductance

method," Solid-State Electronics Vol.43, , pp. 1985-1988, 1999.
[11] M.F. AlHajri, K.M. El-Naggar, M.R. AlRashidi, A.K. Al-Othman "Optimal extraction of solar cell parameters using pattern search" Renewable Energy 44 ,pp 238-245, 2012.
[12]] M.R. AlRashidi , M.F. AlHajri, K.M. El-Naggar, A.K. Al-Othman ".A new estimation approach for determining the I–V characteristics of solar cells", Solar Energy (85) pp.1543–1550, 2011
[13] H. Mekki, A. Mellit, H.Salhi, and K. Belhout, "Modeling and Simulation of Photovoltaic Panel based on Artificial Neural Networks and VHDL-Language'', 4th International Conference on Computer Integrated Manufacturing CIP'2007, 03-04 November 2007
[14]F. Almonacid, C. Rus, L. Hontoria, F.J. Munoz, "Characterisation of PV CIS module by artificial neural networks:A comparative study with other methods", Renewable Energy 35, pp. 973-980, 2010.
[15] S. A. Kalogirou, "Artificial neural-networks for energy systems",Applied Energy, vol. 67, pp. 17 - 35, 2000.
[16] F. Bonanno, G. Capizzi, C. Napoli, G. Graditi, G. Marco Tina, "A radial basis function neural network approach for the electrical characteristics estimation of a photovoltaic module",Applied Energy, vol. 97, pp. 956-961, 2012.
[17] Alireza Askarzadeh, "Voltage prediction of a photovoltaic module using artificial neural networks", International Transaction on Electrical Energy. Systems, (24), pp.715–1725, 2014
[18]M. Karamirad,M. Omid, R. Alimardani, H. Mousazadeh, S.N. Heidari "ANN based simulation and experimental verification of analytical four- and five-parameters models of PV modules" Simulation Modelling Practice and Theory, 34, pp. 86-98, 2013.
[19]A. Mellit, S. Sağlam, S.A. Kalogirou "Artificial neural network-based model for estimating the produced power of a photovoltaic module", Renewable Energy, 60, pp. 71-78, 2013
[20] Moufdi Hadjab, Smail Berrah and Hamza Abid, "Neural network for modeling solar panel", International Journal of Energy, Issue 1, Vol. 6, 2012.
[21] S. H. Oudjana, A. Hellal, and I. Hadj Mahammed, "Power Forecasting of Photovoltaic Generation", International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering ,Vol. 7, No 6, 2013
[22] Salim Moslehi, T. Agami Reddy, Srinivas Katipamula "Evaluation of data-driven models for predicting solar photovoltaics power output", Energy Vol.142, pp 1057-1065, 2018(Available online 10 September 2017)
[23] A. M. Muzathik, "Photovoltaic Modules Operating Temperature Estimation Using a Simple Correlation", International Journal of Energy Engineering,Vol. 4 (4) ,pp. 151-158, Aug. 2014.
[24] T. Bhattacharya ,A. K. Chakraborty , and K. Pal, "Statistical Analysis of the Performance of Solar Photovoltaic Module with the Influence of Different Meteorological Parameters in Tripura, India", International Journal of Engineering Research Volume No.4(3), pp 137 – 140, March 2015.
[25] Abhineet Samadhiya and Ruchi Pandey, "Analysis of PV Panels under Various Weather Conditions", International Journal of Emerging Research in Management &Technology, Volume 5(2), February 2016.
[26] Saiful Anwar, Yoshiki Mikami, "Comparing Accuracy Performance of ANN,MLR, and GARCH Model in Predicting Time
Deposit Return of Islamic Bank", International Journal of Trade, Economics and Finance, Vol.2, No.1, February, 2011
[27] Vijay S. Desai and Rakesh Bharati, "A comparison of linear regression and neural networkmethods for predicting excess returns on large stocks" Annals of Operations Research Vol. 78, pp.127 – 163,1998
[28] Erdi Tosun , Kadir Aydin , Mehmet Bilgili, "Comparison of linear regression and artificial neural network model of a diesel engine fueled with biodiesel-alcohol mixtures", Alexandria Engineering Journal Vol. 55, pp.3081–3089,2016.
[29] Damoder N Gujarati,"Basic Econometrics", Fourth Edition ,Mc Graw-Hill Companies , 2004.