

Solar PV Panel Output Power Forecasting Using Probabilistic Neural Network

Dr. E. Praynlin

ECE Department, V V college of Engineering

ABSTRACT

Solar Energy is an important renewable and unlimited source of energy. Solar photovoltaic power forecasting, is an estimation of the expected power production, that help the grid operators to better manage the electric balance between power demand and supply. Neural network is a computational model that can predict new outcomes from past trends. The artificial neural network is used for photovoltaic plant energy forecasting. The output power for solar photovoltaic cell is predicted on hourly basis. In historical dataset collection process, Four dataset was collected and used for analysis. The implementation of Artificial Neural Network structure is done by Multilayer Perceptron (MLP) and training procedure for neural network is done by Probabilistic Neural Network (PNN). In order to train and test the neural network, the datasets are divided in the ratio 70:30. The accuracy of prediction can be done by using various error measurement criteria and the performance of neural network is to be noted.

Keywords- Forecasting, Probabilistic Neural Network (PNN), Artificial Neural Network.

I. INTRODUCTION

Solar photovoltaic systems transform solar energy into electric power. The power output depends on the incoming radiation and on the solar panel characteristics. Photovoltaic power production is increasing nowadays. In recent years, photovoltaic (PV) technology has been rapidly developed due to the maintenance free, long lasting, and environment friendly nature of PV as well as government's support. However, power output of PV system is a nonstationary random process because of the variability of solar irradiation and environmental factors.

Any grid-connected PV system in the public power grid, is regarded as a non-controlled, non-scheduling unit, its power output fluctuations will affect the stability of power system. As the use of large-scale grid-connected PV system is increasing, it's important to strengthen the prediction of PV system power output, which can help the dispatching department to make overall arrangements for conventional power and photovoltaic power coordination, scheduling adjustment, operation mode planning

Solar Photovoltaic power forecasting information is essential for an efficient use, the management of the electricity grid and for solar energy trading. The interest in solar energy is continuously increasing due to environmental concerns and reduced technology costs. Due to meteorological uncertainty, photovoltaic energy is difficult to predict[7] [10]. Weather variables such as temperature global solar irradiation, sunshine duration, wind speed, relative humidity, cloudiness or sky cover, dew point and precipitation are used as inputs for solar power forecasting models

II. FORECASTING METHODS

Solar Photovoltaic forecasting methods can be broadly characterized as physical or statistical [6], [8]. Physical models are based on mathematical equations which describe the ability of PV systems to convert the introduced meteorological resources into electrical power [1]. These models can be very simple, if based only to the global solar radiation, or more complicated if they include additional parameters. As a matter of fact, it is not easy to predict PV module energy production since it depends on several parameters. Nowadays the most applied techniques to model is the stochastic nature of solar irradiance at the ground level and thus the power output of PV installations are the statistical methods; in particular regression methods are often employed to describe complex nonlinear atmospheric phenomenon that includes the Auto-Regressive Moving Averages (ARMAs)

method, as well as its variations, such as the Auto-Regressive Integrated Moving Averages (ARIMAs) method [3]. The performance of these models is very good for few-minutes to few-hours ahead of forecasting. Nonlinear methods, such as wavelet-based methods [9], [12] have been shown superior to linear models. Nowadays the most common way to forecast the future values of a time series [4] is the use of machine learning methods, PSO method [13], [2], genetic algorithm [5], Adaptive Neuro Fuzzy Inference system (ANFIS) [11] and ANN. Reviewed literature shows that ANN methods have been successfully applied for forecasting. Soft computing techniques based on ANN are used for few hours power output forecast. Thus in statistical approach of forecasting process, artificial neural network plays an important role photovoltaic energy forecasting. In order to define the accuracy of the prediction, some error indexes are introduced to evaluate the performances of the forecasting models. The analysis is based on the experimental activities carried out by real photovoltaic plant from which the dataset has been collected.

III. PROBABILISTIC NEURAL NETWORK

PNN is derived from Radial Basis Function (RBF) Network which is an ANN using RBF as specified in. RBF is a bell shaped function that scales the variable nonlinearly. PNN is adopted for it has many advantages. Its training speed is many times faster than a BP network. PNN can approach a Bayes optimal result under certain easily met conditions. Additionally, it is robust to noise examples. We choose it also for its simple structure and training manner. The most important advantage of PNN is that training is easy and instantaneous. Weights are not “trained” but assigned. Existing weights will never be alternated but only new vectors are inserted into weight matrices when training. So, it can be used in real-time. Since the training and running procedure can be implemented by matrix manipulation, the speed of PNN is very fast [14]. The network classifies input vector into a specific class because that class has the maximum probability to be correct.

PNN [15] is often used in classification problems. When an input is present, the first layer computes the distance from the input vector to the training input vectors. This produces a vector where its elements indicate how close the input is to the training input. The second layer sums the contribution for each class of inputs and produces its net output as a vector of probabilities. Finally, a complete transfer function on the output of the second layer picks the maximum of these probabilities, and produces a 1 (positive identification) for that class and a 0 (negative identification) for non-targeted classes.

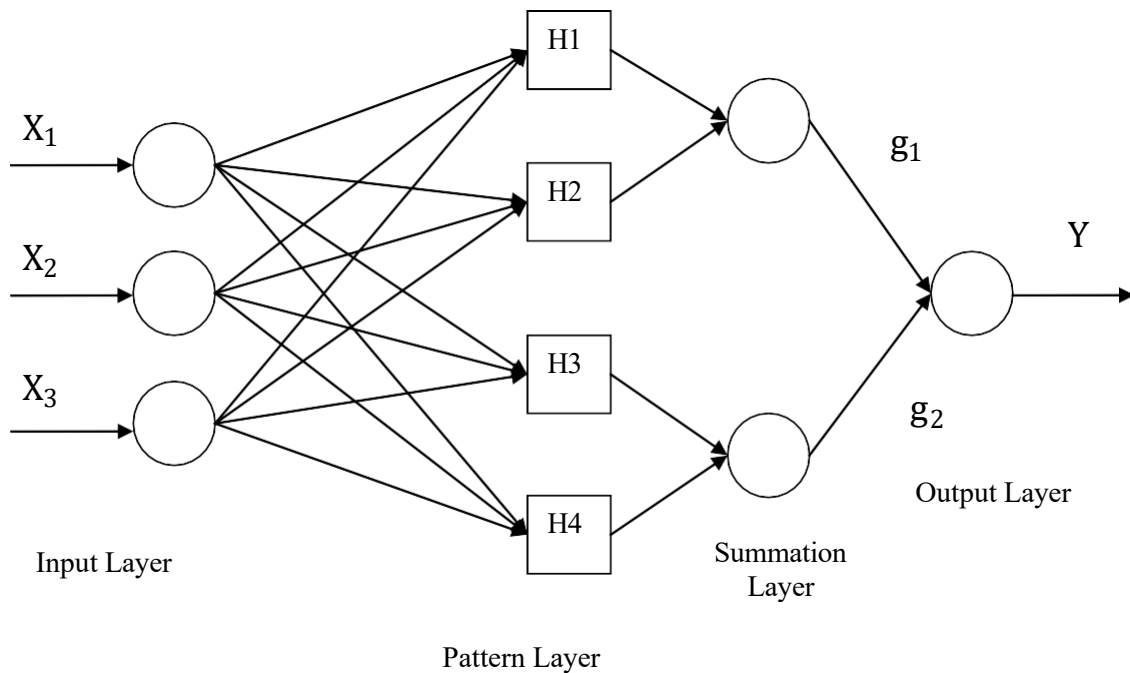


Fig.4 Structure of Probabilistic Neural Network

Each neuron in the input layer represents a predictor variable. In categorical variables, $N-1$ neurons are used when there are N number of categories. It standardizes the range of the values by subtracting the median and dividing by the interquartile range. Then, the input neurons feed the values to each of the neurons in the hidden layer.

Pattern layer contains one neuron for each case in the training data set. It stores the values of the predictor variables for the case along with the target value. A hidden neuron computes the Euclidean distance of the test case from the neuron's centre point and then applies the RBF kernel function using the sigma values. For PNN networks, there is one pattern neuron for each category of the target variable. The actual target category of each training case is stored with each hidden neuron; the weighted value coming out of a hidden neuron is fed only to the pattern neuron that corresponds to the hidden neuron's category. The pattern neurons add the values for the class they represent.

The output layer compares the weighted votes for each target category accumulated in the pattern layer and uses the largest vote to predict the target category. PNNs are much faster than multilayer perceptron networks.

PNNs can be more accurate than multilayer perceptron networks. These networks are relatively insensitive to outliers and they generate accurate predicted target probability scores. PNNs approach Bayes optimal classification.

IV. EVALUATION CRITERIA

In order to correctly define the accuracy of the prediction and the related error, it is necessary to define the indexes that can be used to evaluate the performances of the forecasting model. Some of these definitions come from statistics. The error definitions are really different among each other. Also, technical papers present a lot of these indexes, therefore here we report some of the most commonly used error definitions. The starting reference point is the hourly error „eh” defined as the difference between the average power produced (measured) in the h-th hour $P_{m,h}$ and the given prediction $P_{p,h}$ provided by the forecasting model.

Hourly error (e_h)- is defined as the difference between the average power produced (measured) $p_{m,h}$ and the given prediction $p_{p,h}$ provided by the forecasting model.

$$e_h = p_{m,h} - p_{p,h} \quad (1)$$

When the value of hourly error is low, then the accuracy of prediction will be high.

Absolute hourly error ($e_{h,abs}$)- which is the absolute value of hourly error which is used to calculate the performance.

$$e_{h,abs} = |e_h| \quad (2)$$

When hourly error is low, then automatically the value of absolute hourly error will be low.

Hourly error percentage ($e\%_p$) is based on hourly expected power output $p_{p,h}$

$$e\%_p = \frac{|e_h|}{p_{p,h}} \cdot 100 \quad (3)$$

Hourly error percentage ($e\%_m$) is based on hourly measured power output $p_{m,h}$

$$e\%_m = \frac{|e_h|}{p_{m,h}} \cdot 100 \quad (4)$$

The hourly error percentage based on measured power is generally higher than the expected power output.

Normalised Mean Absolute Error (NMAE%) is based on net capacity of the photovoltaic panel from which the data set is collected, C .

$$\text{NMAE}\% = \frac{1}{N} \sum_{h=1}^N \frac{|p_{m,h} - p_{p,h}|}{C} \cdot 100 \quad (5)$$

where N represents the number of samples.

V. DATA SET DESCRIPTION

Solar photovoltaic historical data set is collected from the photovoltaic (PV) panel. The data is collected from the atmosphere for the production of solar power from the solar radiation [10]. The PV panel will collect some parameters such as humidity average, wind speed average, solar irradiance, temperature, cloud coverage, output power, etc. The data is collected for each and every minute time interval. After the first process data collection, data preprocessing procedure is carried out to train the neural network. This procedure is done to normalise the data set. Before any other steps, historically measured PV data must be always validated since the untrustworthy data increases the odds of higher error in the forecasting process. Five data sets are collected from different photovoltaic systems for the process.

5.1 RUTLAND DATA SET

The data set which was collected for this project is from 50KW DC ground mount installation located in Rutland. The data set was collected for each fifteen minutes time interval. This data set contains about 1248 data. The data set contains six parameters

- Ambient temperature average
- Panel temperature average
- Transformer temperature average
- Solar irradiance

- Wind speed
- Power output

The data set was collected from 19th January 2010 to 31st January 2010. This was obtained by daily observation of solar irradiance, temperature, humidity and generated power by photovoltaic installation which is located in particular locations.

5.2 GECAD PHOTOVOLTAIC SYSTEM DATA SET

This data is obtained from GECAD photovoltaic system. The capacity of one photovoltaic panel is 200W. This data set was collected for each and every five minutes time interval. This data set was collected for one month, that is, from 1st May 2015 to 31st May 2015. The parameters of the data set can be as follows.

- Actual solar radiation
- Sensor usage
- Ambient temperature
- Module temperature
- Total amount of energy

There is about 8451 data contained in the respective data set collected from GECAD photovoltaic system. Thus, the data set contains the solar radiation average which is the amount of daily irradiance from the sun.

5.3 PV SPOT DATA SET

This data set was collected from the photovoltaic plant located in Netherland. The net capacity of the PV plant is 15.24 KW. This data set is collected for one year, that is, from 1st January 2011 to 31st December 2011. The parameters of the data set can be as follows.

- Low estimation output power
- Best estimation output power
- High estimation output power
- Tilted irradiance average
- Horizontal irradiance average
- Temperature

The low, high and best estimation output values are considered as only one value, that is, by taking the average of three values and considered as total power. This data set contains about 365 data.

VI. RESULTS AND DISCUSSION

The Probabilistic Neural Network has been implemented for photovoltaic forecasting and the corresponding error values obtained for five data sets are tabulated in Table 1

Table.1 Probabilistic Neural Network Results for NMAE, WMAE, nRMSE

Different data set	NMAE (Normalised Mean Absolute Error)		WMAE (Weighted Mean Absolute Error)		nRMSE (normalised Root Mean Square Error)	
	Training	Testing	Training	Testing	Training	Testing
Rutland data	7.47	18.06	3.73	9.03	1.10	1.746

GECAD data	586.04	641.75	1.17	1.28	6.04	5.11
	556.83	666.11	1.11	1.33	5.71	5.27
PV spot data	0.32	0.16	4.92	2.48	7.86	2.60

In order to avoid memory complexity the second data set which was collected from GECAD photovoltaic system, which contain nearly 8451 data was divided into GECAD1 and GECAD2 data set. The errors are calculated by eq.11, 12, 13 respectively

Next the hourly error(eh), absolute hourly error(ahe), hourly error based on expected power(heep), and hourly error based on measured power(hemp) can obtain a multiple output which can be plotted against measured output(Y) and expected output(T). The result is plotted in the graphs

Initially, the graph was plotted for Rutland data set for Probabilistic Neural Network is implemented, where five inputs and one output is considered and network training and testing is done with 1248 data sets which is specified in Fig.1.

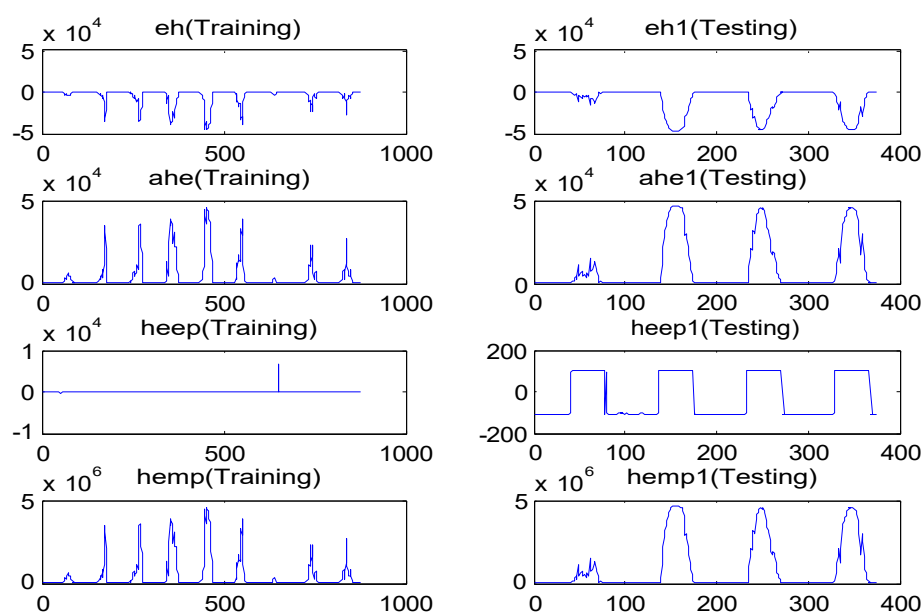


Fig.1 Training and Testing Plot of eh, ahe, heap, hemp using PNN for Rutland data set

The next graph is plotted for GECAD1 data set, where four inputs and one output is considered. These graph also plotted against measured output(Y) and expected output (P) during training the network and measured output (Y1) and expected output(P1) obtained during testing the network which is specified in Fig.2.

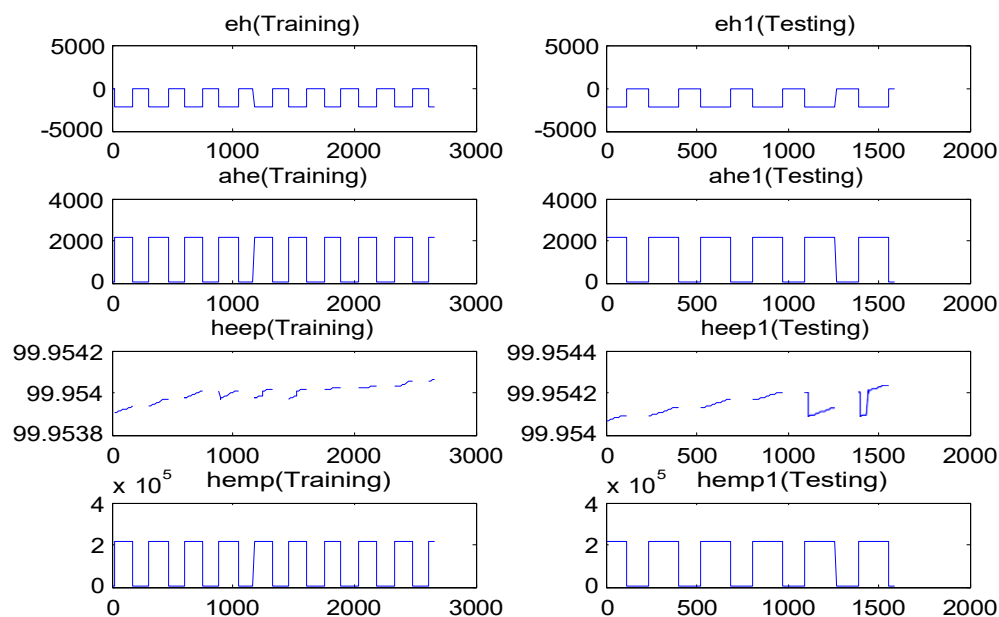


Fig.2 Training and Testing Plot of eh, ahe, heap, hemp using PNN for GECAD1 data set

The next graph is plotted for GECAD2 data set, where four inputs and one output is considered. These graphs also plotted against measured output(Y) and expected output (P) during training the network and measured output (Y1) and expected output (P1) obtained during testing the network which is specified in Fig.3.

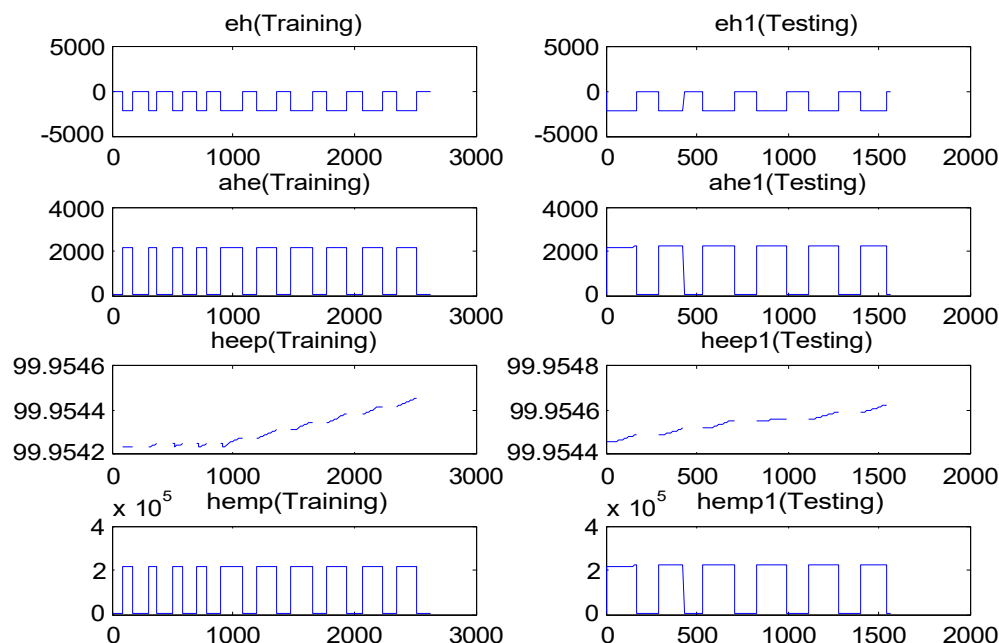


Fig.3 Training and Testing Plot of eh, ahe, heap, hemp using PNN for GECAD2 data set

The next graph is plotted for PV spot data set which contains nearly 365 data sets for Probabilistic Neural Network. These plots are plotted against the measured power and expected power obtained during training and testing of the network which is specified in Fig.4.

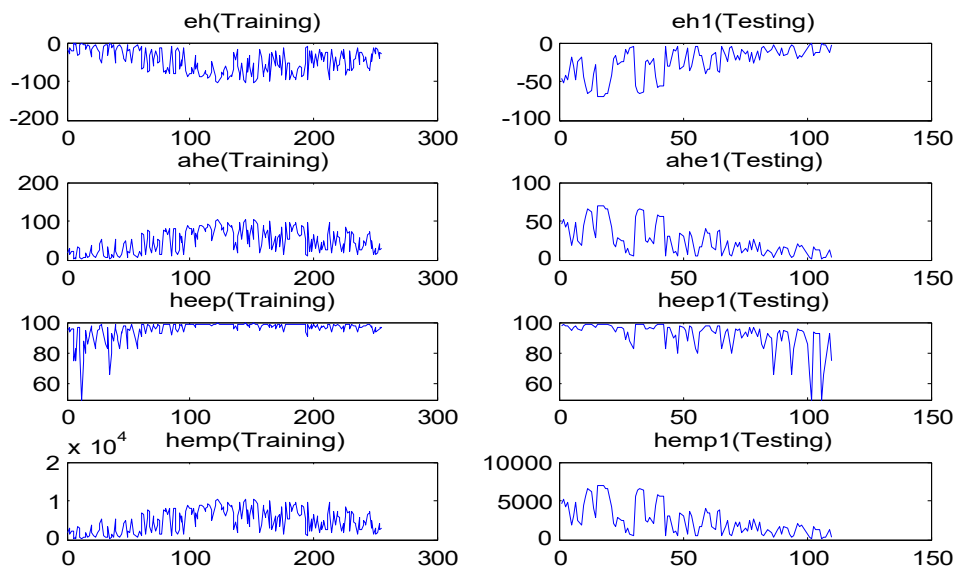


Fig.4 Training and Testing Plot of eh, ahe, heap, hemp using PNN for PV spot data set

VII. CONCLUSION AND FUTURE WORK

Photovoltaic forecasting method based on Artificial Neural Network (ANN) using Probabilistic Neural Network (PNN) algorithm was done. Here three types dataset is used for analysis, Rutland dataset, GECAD PV dataset and PV spot dataset.

This article reports four predictive models which are applied for prediction of output power in PV systems.. Probabilistic Neural Network generates accurate predicted target probability scores and they are faster to train. The evaluation of forecasting is performed by calculating some error values such as NMAE. The results are plotted and error values are tabulated. The result demonstrates that accuracy of prediction is more accurate in Probabilistic Neural Network (PNN), compared with the other forecasting methods. BP results provide lower error value comparatively. Thus, effectiveness of the proposed forecasting method provides improved accuracy in solar power forecasting. Prediction of solar power output of photovoltaic system gives considerable information for energy managers and all individuals who deal with trading electricity and scheduling of solar projects.

In future, simulated photovoltaic data set can be tried to generate with corresponding procedure. This data set can be implemented in other Hybrid Neuro Fuzzy Systems such as Fuzzy adaptive learning control networks, Generalised Approximate Reasoning based Intelligence Control, fuzzy net, etc. to get better predictive results compared with the proposed methods.

REFERENCE

- [1] Adel Mellit and Alessandro Massi Pavan, "A 24-h Forecast of Solar Irradiance using Artificial Neural Network: Application for Performance Prediction of a GridConnected PV Plant at Trieste, Italy", *Solar Energy*, Vol. 84, No. 5, pp. 807-821, 2010.
- [2] S.H. Chen, A.J. Jakeman and J.P. Norton, "Artificial Intelligence Techniques: An Introduction to their use for Modelling Environmental Systems", *Mathematics and Computers in Simulation*, Vol. 78, No. 23, pp. 379-400, 2008.
- [3] Christophe Paoli, Cyril Voyant, Marc Muselli and Marie Laure Nivet, "Forecasting of preprocessed Daily Solar Radiation Time Series using Neural Networks", *Solar Energy*, Vol. 84, No. 12, pp. 2146-2160, 2010.

- [4] Crescenzo Gallo and Michelangelo De Bonis, "A Neural Network Model for Forecasting Photovoltaic Deployment in Italy", International Journal of Sustainable Energy and Environment, Vol. 1, No. 1, pp. 1-13, 2013
- [5] Davide Caputo, Francesco Grimaccia M. Mussetta and R.E. Zich, "Photovoltaic Plants Predictive Model by means of ANN trained by a Hybrid Evolutionary Algorithm", Proceedings of International Joint Conference on Neural Networks, pp. 1-6, 2010.
- [6] Alberto Dolara, F. Grimaccia, S. Leva, Marco Mussetta and Emanuele Ogliari, "A Physical Hybrid Artificial Neural Network for Short Term Forecasting of PV Plant Power Output", Energies, Vol. 8, No. 2, pp. 1138-1153, 2015
- [7] Dolara. A, Grimaccia. F, Leva. S, Mussetta. M, Ogliari. E, "A physical hybrid artificial neural network for short term forecasting of pv plant power output", Solar Energies, Elsevier, Vol. 8, No. 2, pp.1138–1153, 2015.
- [8] F. Grimaccia, S. Leva, Marco Mussetta and Emanuele Ogliari, "Analysis and Validtion of ANN PV Power Output Forecasting at 24 hours Ahead", Proceedings of ELECTRIMACS, pp. 19-22, 2014
- [9] Jiacong Cao and XingchunLin, "Application of the diagonal recurrent wavelet neural network to solar irradiation forecast assisted with fuzzy technique", Engineering Applications of Artificial Intelligence, Vol. 21, No. 8, pp. 1255-1263, 2008
- [10] Jayawardene. I, Venayagamoorthy. G. K, "Comparison of Adaptive Neuro-Fuzzy Inference Systems and Echo state Network for PV Power Prediction", Procedia Computer Science, Elsevier, Vol. 53, pp. 92-102, 2015.
- [11] E. Praynlin and P. Latha, "Software Development Effort Estimation using ANFIS", International Information Institute, Vol. 17, No. 4, pp. 1325-1337, 2014.
- [12] A. Mellit, M. Benghanem and S.A. Kalogirou, "An adaptive wavelet-network model for forecasting daily total solar-radiation", Applied Energy, Vol. 83, No. 7, pp. 705- 722, 2006.
- [13] Ratnadip Adhikari and Rajiv Kumar Agrawal, "Effectiveness of PSO based Neural Network for Seasonal Time Series Forecasting", Proceedings of the Fifth Indian International Conference on Artificial Intelligence, pp. 231-244, 2011.
- [14] Shifei Ding, Weikuan Jia, Chunyang Su, Xiaoliang Liu, Jinrong Chen, "An Improved BP Neural Network Algorithm Based on Factor Analysis", Journal of Convergence Information Technology, Vol. 5, No. 4, 2010
- [15] Donald F. Specht, "Probabilistic Neural Networks", Neural Networks, Vol. 3, pp. 109-118, 1990.

AUTHOR DETAILS



PRAYNLIN E was born in Kanyakumari, India. He received his Bachelor of Engineering degree in Electronics and communication Engineering from Sun College of Engineering and Technology, Nagercoil, India, in 2009 under Anna University; and his Master of Engineering degree in Applied Electronics from Noorul Islam University, India, in 2011 He completed his doctorate in year 2015 form Anna university. His current research interests include Neural Networks, Solar PV forecasting, , Software Effort Estimation and Fuzzy logic.