A Short Term Forecasting of PV Power Generation Using Couple Based Particle Swarm Optimization Pruned Extreme Learning Machine

Niranjan Nayak*, Alok Kumar Pani*‡

* Department of EEE, ITER,Siksha'o'Anusandhan, deemed to be university, Jagamohan Nagar, Khandagiri, Bhubaneswar Odisha, India PIN 751030.

(niranjannayak@soa.ac.in, alokpani@soa.ac.in)

‡ Corresponding Author; Tel: +91 674 235 0181,

Fax: +91 674 235 1880, alokpani @soa.ac.in

Received: 26.04.2019 Accepted: 13.06.2019

Abstract: Presently, there is a high growth of power demand, and to minimize the air pollution by the conventional power plants, it is required to enforce the renewable power to the existing conventional grids. The non linearity and erratic nature of PV power generation creates a great challenge in the interconnection and grid management. To install further solar plants of large capacity, the prediction of PV power is extremely required. To overcome this challenge, many machine learning techniques has been successfully implemented for short/long term forecasting of the PV power. Among all the methods, Extreme learning machine (ELM) is one of the few victorious methodologies' in machine learning approaches. One of the key potencies, of ELM, for which it is appreciated, is a low computational effort required for training recent data. In ELM algorithm, the numbers of hidden and output layer nodes are arbitrarily selected and are rationally decided too. Further the conventional ELM does not perform satisfactorily in some complicated problems or in case of big data. Thus in this work we have investigated the pruned-ELM (P-ELM) approach as a methodical and programmed approach for developing a forecasting model to predict PV power generation on short term horizon. P-ELM shows statistical ways to compute the significance of inner nodes. We start from an initial large number of inner nodes, and then inappropriate nodes are pruned by taking into account the appropriateness to the forecasting problem. In order to increase the performance of the P-ELM the weights of input layer also are optimized by couple based PSO algorithm.

Keywords PV power, Machine Learning, ELM, P-ELM, PSO, CPSO

1. Introduction

Usage of renewable energy initiatives like wind and photovoltaic power is an unbeaten way to reduce environment pollution caused by the use of fossil fuel resources, in the conventional power generation methods. Solar rays is the most abundant renewable power source on the earth exceeding all other energy sources like geothermal, hydro power, wave and tide, biomass and wind. Within a few years, power generation from sun rays has quickly improved as it extends a spirited solution for a sustainable production of electrical power [1-2,23,29]. For illustration, the production facility of solar energy goes up to 40 Giga Watt in the USA after setting up of 15 Giga Watts of solar power facilities in the year 2016. It is expected that the total PV capacity may extend by three times of the present generation over next five years. In India growth of solar power installation is reasonably fast. The installation was 2650 MW in May 2014, reached 20 GW in Jan 2018 and 26 GW as of 30 September 2018. Solar power production is too increasing worldwide and as per the global solar installation the 300 Giga Watt ability had approached in 2016 and the entire worldwide installation of solar plants is likely to reach 700 Giga Watt in 2021 [3-4].

Solar power production for the electricity grids has many restrictions since the solar power production is time reliant and irregular in nature. 'Predict and control', 'big battery', and 'answer to the need' are the three ways that can be used to overcome these difficulties. In spite of the growth in 'secondary cell' and other power backup technologies, very large amount power backup is still not feasible due to monetary and ecological restrictions [5]. Conventionally, PV power forecasting, called point forecast, has limited value of each forecasting result. The training phase of point forecast must be repetitive for the variation of price function. This can be computed all-inclusive, as the learning phase normally requires rigorous privilege of a big set of data.

Nevertheless in probabilistic predictions, the learning phase is independent of price function and the change of price function only influences the decision-making at a future stage. Regarding the restrictions of point predictions, fortunately a few recent publications [31] show many probabilistic prediction methods for solar power emission. For example [6], suggests a Bayesian method to guesstimate parameters of the extrapolative distribution, which is supposed to be a modified Gamma distribution.

The efficiency of point forecast is normally tested by some standard error functions, like root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE). A lot of scientists have projected ways to point forecast the calculation of solar power emission in the next time interval using different machine learning techniques. The ANN is used for solar power forecasting, fuzzy systems are used in many research cases for the same purpose. The ARMA and ARIMA methods are well accepted in solar power prediction however they have the demerits like incapability of computing the highly fluctuating and non linear data. To overcome the mentioned difficulties other forecasting models like back propagation neural net (BPNN) has been designed [7-11].

Recently a novel and powerful learning method called the extreme learning machine (ELM) has been verified by many researchers in the field of solar power forecasting [12-13,30]. This is a neural network with input and output layers and one hidden layer with arbitrarily selected input weights. The proper selection of weights improves the performance of ELM which affects the forecasted output and stability of the system. Here in this work, the empirical mode decomposition (EMD) is in practice where the signal is divided into various components known as IMFs which increases stability. The generation of IMFs are given in the figure-3.

Further to increase the effectiveness, ELM is modified to an improved version known as pruned extreme learning machine (P-ELM), which is applied for PV power forecasting. A number of nature inspired optimization algorithms have gained more attention in the last few years. Various other nature inspired algorithms developed are cat swarm optimization, artificial bee colony, cuckoo search algorithm, fire-fly algorithm, social insect behaviour, grey wolf search algorithm, particle swarm optimization etc., which have been implemented for forecasting of various quantities[14-20].

In this work the random weights of P-ELM are optimized by a new technique known as couple based particle swarm optimization (CPSO). Unlike other optimization algorithms, CPSO consists of less number of parameters and can easily be applied to upgrade the weights and increase the stability of the systems [21-28].

The rest part of the paper is organized as follows. Section-2 shows the choice of data. Section-3 explains Empirical Mode Decomposition. Section-4 explains ELM and modified ELM techniques and its mathematical modeling. This section also explains PSO and CPSO. The optimization of ELM weight is done by the pruning through CPSO algorithm, which is discussed in section-5. Simulation results of different methods are explained in section-6. Conclusion and future scopes are discussed in section–7.

2. Choice of Data

The data collection has been done and it is of critical importance before the application of forecasting methods. The past data in the form of a time series like solar output power are composed from a working PV power plant situated at Bhubaneswar, Odisha, India. (The details is as indicated in table-1 given below) and decomposed by EMD and then applied to the forecasting model.

Table 1: Location of the plant

S. No.	Component	Data
1.	Cell category	Polycrystalline
2.	Array Total Capacity	11.2 kW
3.	Latitude	20°25’
4.	Longitude	85°80’
5.	Area of Array	106.25m^2

1. For valid use, horizons to be chosen are changed, which depends on the need for judgment building requirements. The entire data set is divided in to two parts such as 80% for training and 20% for testing data.

2.1 Standardization of the Model:

In prediction, it is important and compulsory that the projected model be approved; to predict accurately we use several methods in our favor. Prediction modeling process is illustrated in the following block diagram.

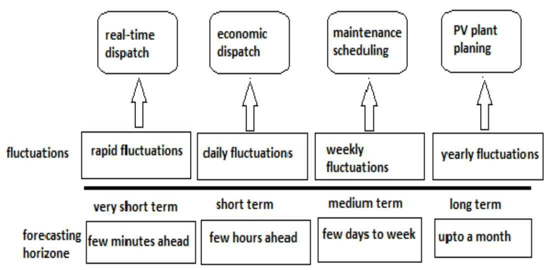

Fig.1: Prediction Method in A Block Diagram

3. Signal Splitting via EMD

Empirical Mode Decomposition is a method in which a signal divides into multiple component signals called Intrinsic Mode Functions (IMFs) [5-7]. By breaking down the signal, forecast model becomes more functional and stable. However the performance of the model increases by signal putrefaction and prediction accuracy is very high. The model asserts that the following expression can describe any fraction of the fresh string of signals.

$$s(t) = \sum_{i=1}^{n} \alpha i(t) + r_n(t), \quad where\ r_n(t)\ and\ \alpha \tag{1}$$

are IMFs and remainder of the signal respectively.

The EMD is an iterative process for signal decomposition, and the decomposed signal is has a range of amplitudes and frequencies.

The component creation agrees to the conditions that the number of maxima or minima is either same with the number of zero crossings or their difference is one, in the entire data set. The other condition is that their average must be zero on envelop.

The 1st term of IMF is described by,

$$imf_1(t) = s(t) - \eta_1(t) \tag{2}$$

When s(t) is divided into many IMFs imf_1 is the first term. If it is not then it is to be measured as the signal itself and split frequently. KIMF1 is taken to be an IMF and is given as

$$\rho_1 = kimf_1 \tag{3}$$

This term (ρ_1) is subtracted from s(t) by taking

$$s(t) - \rho_1(t) = \lambda_1(t)$$

R is residue of the given signal and other IMFs are

$$\rho_1, \rho_2 \ldots\ldots \rho_n. \tag{4}$$

Then

$$\lambda_n(t) - \rho_n(t) = \eta(t)$$

Equation (4) and (5) are the altered version of equation(1)

The terminating state for this uneven procedure was recommended by Huag et.al, a normalized squared variation among the two consecutive changing operations is essential.

$$S_1 D_k = \sum_{t=0}^{T} \frac{\left| imf_{n(k-1)}(t) - imf_{nk}(t) \right|^2}{imf_{n(k-1)}^2(t)} \tag{5}$$

The changing method is ended by any pre considered criteria, by adding all the IMFs and the ending residue the signal $s_n(t)$ can be restated as

$$s_n(t) = \sum_{i=1}^{m} \alpha_{ni}(t) + r_{nm}(t) \tag{6}$$

4. Extreme Learning Machine (ELM):

ELM is a newly developed machine learning algorithm which deals upon the idea of a feed forward neural network having one hidden layer (SLFN). ELM can be made to learn with no iterations. The construction of ELM is as a graph of 'd' input nodes, called input layer and 'h' hidden nodes, known as hidden layer and one or many output nodes, known as output layer. Space measurement units are linked with each input node, which are fully associated with 'h' hidden neurons by weights w_j chosen arbitrarily and set of biases b_j chosen randomly. 'G' is the universal nonlinear activation function. One can convey the hidden neurons activation matrix A for the entire training

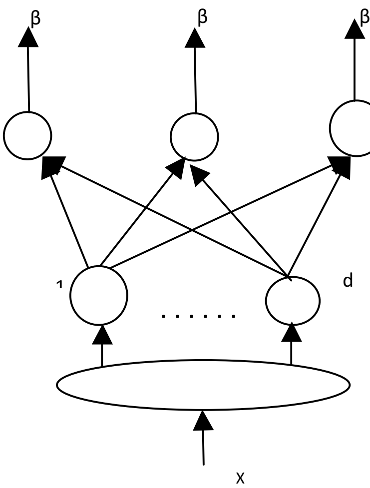

Fig.2: Structure of neural network.

$$\psi_N(x) = \sum_{i=1}^{N} \beta_i G_f(x; p_i, q_i), \quad x \in R^n, l_i \in R^n \tag{7}$$

Where p_i and q_i are the required hidden node learning factors

β_i is the weight vector between i_{th} and output node.

and $G(q_i, x_i, p_i)$ is the result of the i_{th} concealed node with respect to x.

For additional hidden nodes with threshold activation function $g(x): R \to R, G(x; p_i, q_i)$ given by $G(x; p_i, q_i) = g(p_i, x + q_i)$ $p_i \in R$, where p_i is the weight vector between i_{th} hidden and output node, q_i may be treated as bias, and $p_i * x$ is inner product of vectors 'p_i' and' x' in R_n. For RBF the Gaussian triangular activation function is given by

$$G(x; p_i, q_i) = g(q_i \|x - p_i\|), \quad q_i \in R^+ \tag{8}$$

Here c_i and a_i are the centre and impact factor of i_{th} RBF node. R^+ Denotes the set of all positive real values. For N different samples $(x_k, t_k) \in R^n \times R^m$. If a SLFN with $\tilde{N}$ hidden layer estimated these N samples with zero error, then it indicates the existence of β_i, p_i, q_i satisfying the following condition.If N is the number of hidden nodes and $\phi(x)$ is the activation function then SLFN can mathematically be defined as:

$$\sum_{i=1}^{N} W_{oi}.\phi_i(xj) = \sum_{i=1}^{N} W_{oi}.\phi(w_i x_j + b_j) = \theta_j \tag{9}$$

Where $j = 1,2,3,.......N$, $w_i = w_{i1}, w_{i2}, w_{i3}......$= input weight matrix. $W_o = (W_1, W_2, W_3........W_m)^T$ is the output weight matrix. x_i is the input sample.

$$\sum_{i=1}^{N} W_{oi}.\phi(w_i x_j + b_i) = t_j \tag{10}$$

If H is the hidden layer output matrix, T_m is the target matrix and ϕ is the activation function then

$$HW_o = T_m$$

The input weight is randomly calculated and then the output can be written as

$$W_o = H^T T_m \tag{11}$$

$$H = (w_i...........w_n, b_i......b_n)$$

$$H = \begin{bmatrix} \phi(w_1 x_1 + b_1) & \dots & \dots & \phi(w_N x_1 + b_N) \\ \dots & \dots & .... & \dots \\ \dots & \dots & \dots & .. \\ \phi(w_1 x_N + b_1) & \dots & \dots & \phi(w_N x_N + b_N) \end{bmatrix} \tag{12}$$

The flow chart of PV power forecasting using EMD-ELM is given by

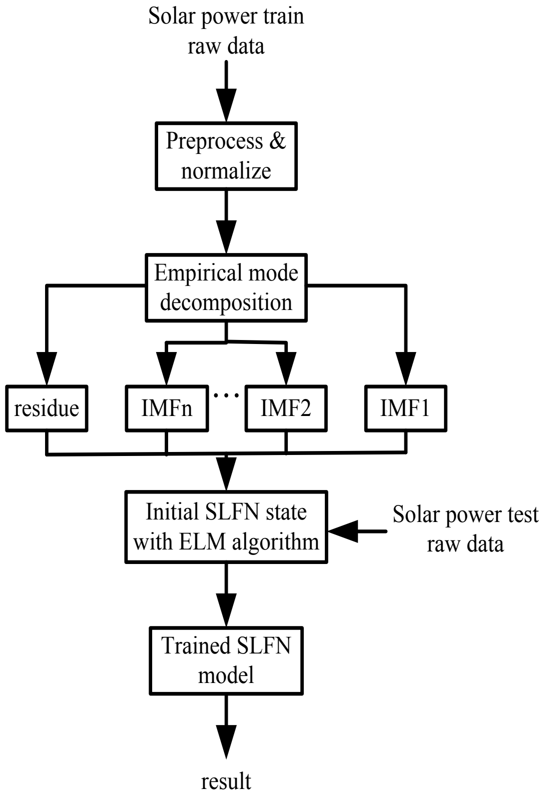

Fig.3: Prediction process based on EMD-ELM.

4.1 Particle Swarm Optimization:

PSO is a swarm intellect algorithm which considers the food search process of birds. Each element in PSO is a bird in a flock of birds searching for food. They possess the data of the position of each element and its velocity, and the elements can be made to learn from both of their own best positional data (pbest) and the global best positional data (gbest), in order to bring up to date its searching orientation and space. PSO is an iterative process. After some iterative loops PSO starts converging and reaches at the optimum result. The process is calculated as the following.

$$\begin{cases} vel_i^{k+1} = \omega v_i^k + c_1 r_1 (p_{besti} - x_i^k) + c_2 r_2 (g_{best} - x_i^k) \\ x_i^{k+1} = x_i + v_i^{k+1} \end{cases} \tag{13}$$

Here x_i^k and v_i^k are the position and velocity of i_{th} element for k_{th} iterative step. p_{best} is the i_{th} elements individual location, and g_{best} is the universal best location in the whole swarm of particles in the search space.

Furthermore, ω is the inertia weight feature, which affects the performance search techniques. c_1 is the cognitive and c_2 is the social factor whose value controls the search progress. r_1 and r_2 are consistently scattered arbitrary numbers in [0,1], which resolve the algorithm randomness.

4.1.1. Global version

PSO is an evolutionary algorithm which operates with swarm based on swarm smartness, and works on foraging attitude of birds. In real time a bird is represented by individual particle in PSO algorithm. Each element acquires position and quickness (velocity) information, and made to learn from its local best positional data as well as global best positional data in order to bring up to date the search position and directional information. After a number of iterative steps in the search, the PSO algorithm tapers to a conclusion giving consequent solutions. This method can be formulated mathematically as follows.

$$\begin{aligned} vel_i^{k+1} &= J \times vel_i^k \\ &+ \lambda_1 rand_1 \left(P_{besti} - x_i^k\right) \\ &+ \lambda_2 rand_2 \left(g_{best}^k - p_i^k\right) \\ p_i^{k+1} &= p_i^k + vel_i^{k+1} \end{aligned} \tag{14}$$

Where p_i^k and vel_i^k are the position and velocity of i_{th} particle in k_{th} iterative step.

J =inertia weight factor.

P_{besti}= best location of i_{th} particle in the search space.

g_{besti}= global best position in the search space

$rand_1, rand_2$ = Two set of random numbers in the range of [0, 1]

λ_1 = Cognitive factor λ_2 = social factor

4.1.2 Local version of PSO

There are two versions of PSO such as global or standard PSO and its local version as mentioned below.

$$\begin{aligned} vel_i^{k+1} &= J \times vel_i^k + \lambda_1 rand_1 \left(P_{besti} - x_i^k\right) \\ &+ \lambda_2 rand_2 \left(l_{best}^k - p_i^k\right) \\ p_i^{k+1} &= p_i^k + vel_i^{k+1} \end{aligned} \tag{15}$$

In the local version of PSO the individual particles gets identified by its close vicinity and l_{best} is replaced by g_{best}. In this PSO the sub-groups of particles are formed.

4.1.3. Couple-based PSO Algorithm (CPSO)

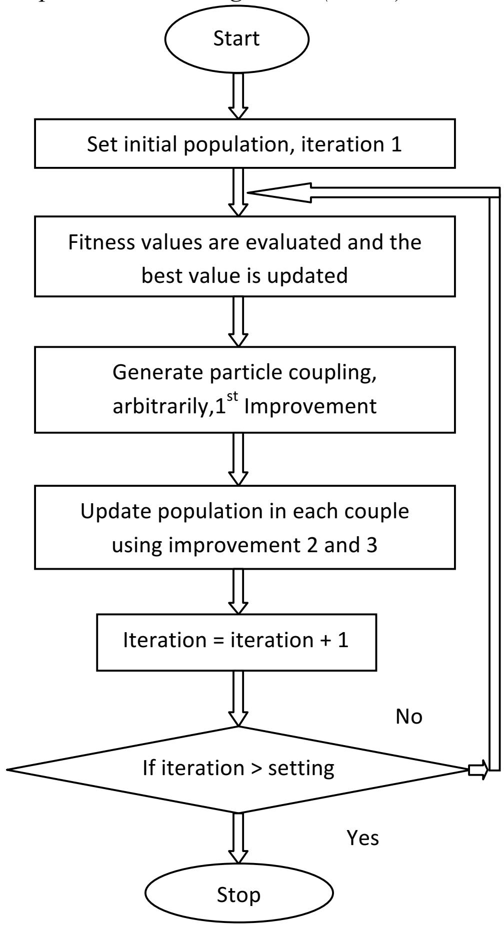

Fig.4: Flow chart of Coupled based particle Swarm Optimization technique

By setting m (number of particles in a sub-group) =2, and r (regrouping period) =10 in the above equation (while simulating) the outcome was very poor. It is not sufficient to make a precious search of swarm with a concerto of only with a couple of particles, but merely towards a 'searching swarm' division. When distinguished 'searching swarm' no longer exists the m=2 and r=1 is to be set. It is observed there is a comfortable data exchange among elements amongst all possibilities of m and r. In the intervening time the communication among two elements becomes the centre of attention, while ignoring the master–slave formation in the sub-group. As per the exceptionality of m= 2, the element couple is reorganized. However based on the element couple, the name of optimization technique is Couple based PSO (CPSO). A flowchart and the improvements of CPSO are given as above.

4.1.4. Dynamic couples of elements

Consider the process of formation of active sub groups as explained earlier. A group of n elements (even) is randomly divided to n/2 couples. The n/2 couples are regrouped arbitrarily during each iteration. That is particles match each other arbitrarily in each iterative step. The regrouping method is explained in following figure.

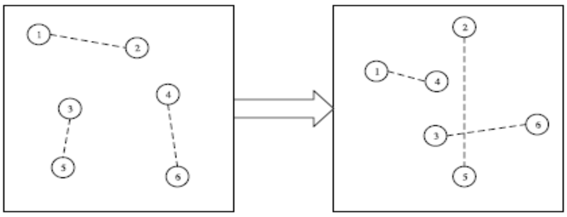

Fig.5: Flow chart of Coupling of particles in CPSO

4.2. Intersectional learning strategy

This is a modified learning method proposed in CPSO. The p_{best} value of one particle is matched with other particle of in the couple. Assuming $p_1 = lowest\ of\ p_{best}$ and the second best value be p_2, the training approach of the couple of particles is stated as follows.

$$\begin{aligned} V_{p1}^{k+1} &= \omega V_{p1}^{k} + c_1 r_1 \left(Pb_{best_{p1}} - x_{p1}^{k}\right) \\ &+ c_2 r_2 \left(cal_{best_{p1}} - x_{p1}^{k}\right) \\ V_{p2}^{k+1} &= \omega V_{p2}^{k} + c_1 r_1 \left(Pb_{best_{p2}} - x_{p2}^{k}\right) \\ &+ c_2 r_2 \left(cal_{best_{p2}} - x_{p2}^{k}\right) \end{aligned} \quad (16$$

Here $call_{best}$ calculates the best position which assigns different values for p_1 and p_2.

For, p_1, $call_{bestp1}$ equals the p_{best} of p_2 for p_2, $call_{best_{p2}}$ equals $call_{best}$ of p1 in the last iterative step. Additionally, the initial value of $call_{best}$ is the p_{best} of the element. It is explicable that p_1 trains starting p_{best} to the enhanced f_{pbest} owner p_2 and p_1 finds out from l_{best} in the vicinity. For p_2, the improved element, the condition is rather changed. Despite of p1 be the inferior or superior element in the final step, $call_{best}$ of p_1 always forms useful data. In some detail, in the last iterative step, if p_1 is the poorer element, it trains from the improved one, and $call_{best}$ of p_1 visualizes the efficient one in sequence. If p_1 is the better element, it traces towards the back once more until p_1 is found to be the inferior particle and its equivalent $call_{best}$ visualizes helpful data as well. Hence, $call_{best}$ of p_1 in the last iterative step is decided to be the training object of p_2 .It is evident that the new learning scheme can repeatedly expose helpful data existing in probing the past. By using this data, individuals widen their training horizons so that the inhabitant's variety can be improved. There exists the difficulty that this type of data may be obsolete in some possibility. Since it (data) has been generated by proceeding through iterative steps and reorganized during development, this drawback has the potential to downgrade the overly rapid evolutionary tempo. This data generation helps in maintaining the inhabitants range.

5. Pruned-Extreme Learning Machine (P-ELM)

In the proposed P-ELM we identify the degree of significance among the hidden nodes and the levels of some statistical measure. Then the nodes with less significance may be taken out without affecting the generalization property of the network. The dimension of initial network with $\tilde{N}$ hidden nodes is assigned as per the basic ELM methodology, the hidden node parameters (l_i, m_i) *where* $i - 1,2,3\ldots\ldots\ldots\ldots\tilde{N}$ are assigned indiscriminately. Depending on the parameters of the hidden node responses matrix 'H' is evaluated. From the training data and hidden node response, the statistical consequence of hidden nodes is recognized. Here we explore two statistical criteria, such as the Chi-squared (w2) and information gain (IG) process, for informative the statistical relevance of the hidden nodes in arriving at the proper levels.

The chi squared method model is given by

$$\chi^2(h_i) = \sum_{l=1}^{d} \sum_{j=1}^{l} \left(\frac{A_{lj} - E_{lj}}{E_{lj}} \right) \quad (17)$$

$$E_{lj} = \frac{R_l * Z_j}{N} R_l \sum_{j=1}^{r} A_{lj} \quad (18)$$

$$Z_i = \sum_{l=1}^{d} A_{ij} \quad N = \sum_{l=1}^{d} \sum_{i=1}^{r} A_{lj} \quad \text{and} \quad h_i = 1,2,3\ldots\ldots\ldots\tilde{N} \quad (19)$$

Individual hidden node h_i has d number of discrete bins and r number of class labels has been attached.

A_{lj} is the cardinality. E_{lj} is the expected cardinality of A_{lj} .Based on χ^2 determine, the hidden nodes are arranged in downward order. For larger value of χ^2 better relevance of the hidden node.

5.1. P-ELM Algorithm

First of all a training set data is collected

$$\psi = \left\{ (z_k, t_k) / z_k \in R^n \right\} \quad (20)$$

Activation function of the data training is considered to be g.

Step-1: An Preliminary well-built hidden node size than essential, is $\tilde{N}$ and a significance threshold base

$$\phi = (\phi_1, \phi_2, \phi_3 \ldots\ldots\ldots\ldots\ldots \phi_p) \quad (21)$$

Then training set is separated into two individual subsets for learning and confirmation.

Step-2: the hidden node variables such as c_i, a_i are assigned arbitrarily. Where $i = 1,2,3,\ldots\ldots\widetilde{N}$,

Step-3: The hidden layer output matrix H is to be computed using learning subsets.

Step.4: The numerical significance for the hidden nodes is investigated using χ^2 and they must shorted out in descending order.

Step.5: For every consequence threshold $\phi = (\phi_1, \phi_2, \phi_3 \ldots\ldots\ldots\ldots\ldots \phi_p)$ (22)

Recognize the subset of critical unseen nodes N_h, that convince the consequence threshold ϕ_i and decide the justification precision o_i, using the rationale subset, and using the expression,

$$aic(i) = N_{val} \ln\left(\frac{\theta_i^2}{N_{val}}\right) + N_{hi}, \text{ Where } i = 1,2,3,\ldots\ldots p \quad (23)$$

N_{val} is the amount of confirmation data.

$\theta_i = (1 - o_i) * N_{val}$ is the validation organization error.

Step.6: Select N_h as per min (aic).

Step.7: Re-train network N_h with the entire training set.

Compute the hidden layer output matrix H. Also find the output weight $\beta^* \times \beta^* = (H^*)'T$

Step.8: Estimate the recital of N_h^* on unobserved testing data set.

5.2. Data Pre-processing

One year historical data is collected for prediction purpose and Figure shows 3000 samples of data collected over periods of 15 min and 30 minutes respectively. Out of the entire time series data 80% of the data are for training purpose and 20% is for testing purpose. For numerical computation the entire data is originally normalized and scaled between 0 and 1 using equation (20), to enhance the training speed and to avoid the calculation overflow.

$$N' = \frac{N - N_{\min}}{N_{\max} - N_{\min}} \quad (24)$$

N' is the entire data in normalized form.

N is entirety data $N_{\min}$ and $N_{\max}$ are the least and utmost generated solar power values in the entire data set, to be forecasted. The solar power data is initially processed with the help of EMD method by splitting the power series into several independent modes known as IMFs in the case of EMD with a single residue value. The terminating condition for EMD is obtained when the residue reaches the pure value. Two case studies have been performed with two time horizons which include very short term data resolution of 15 min time interval, short term data resolution of half an hour time interval.

5.3. Performance Evaluation

The presentation of predictable methodology is evaluated by forecasting solar power for short term periods that is for 15 and 30 minutes. The metrics calculated were Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and correlation coefficient (CC^2), to measure the accuracy.

$$MAPE = \frac{1}{n_0} \sum_{i=1}^{n_0} \left(\frac{|P_{actual} - F_{forecasted}|}{A_{actual}} \right) * 100 \quad (25)$$

$$MAE = \frac{1}{n_0} \sum_{i=1}^{n_0} \left| A_{actual} - F_{forecasted} \right| \quad (26)$$

$$RMSE = \sqrt{\frac{1}{n_0} \sum_{i=1}^{n_0} \left(A_{actual} - F_{forecasted} \right)^2} \quad (27)$$

$$CC^2 = \frac{\begin{pmatrix} n_0 \sum_{i=1}^{n_0} F_{forecasted}(x_i) T_r(x_i) \\ - \sum_{i=1}^{n_0} F_{forecasted}(x_i) \sum_{i=1}^{n_0} T_r(x_i) \end{pmatrix}^2}{\begin{pmatrix} n_0 \sum_{i=1}^{n_0} F_{forecasted}(x_i)^2 \\ - \left(\sum_{i=1}^{n_0} F_{forecasted}(x_i) \right)^2 \end{pmatrix} \begin{pmatrix} n_0 \sum_{i=1}^{n_0} T_r(x_i)^2 \\ - \left(\sum_{i=1}^{n_0} T_r(x_i) \right)^2 \end{pmatrix}} \quad (28)$$

6. Result Analysis

One year data of a real time solar plant in Bhubaneswar, India (mentioned in table-1) for the year 2010 were used for testing the proposed model, in this experiment the whole data set was divided into three different weather types. The winter season ranging from the month of November-March, the spring season ranges from the month of April to July, followed by summer season from the month of August to October. The data was decomposed by EMD before entering to ELM algorithm in order to reduce the forecasting error. The data is trained by ELM in first step and then P-ELM is applied to the same plant data. The results depicts that the P-ELM shows better performance than the other algorithms. Further the couple based PSO techniques are used to optimize the weights, so that the performance of CPSO-PELM is superior to ELM, and EMD- ELM.

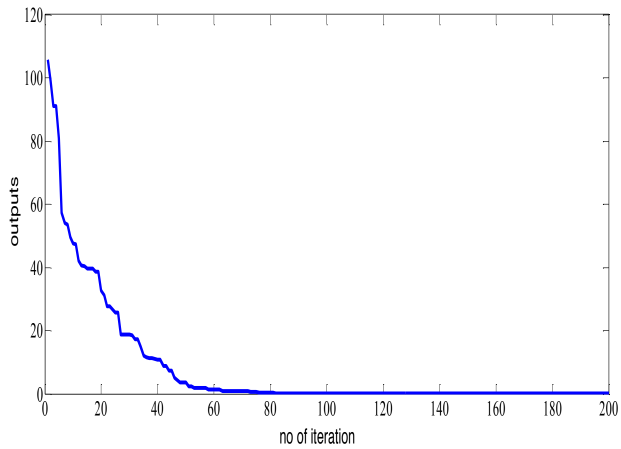

Figure.6 : Convergence of CPSO.

Case-1: Solar Power Prediction for 15 Minutes Horizon:

In this case the data are collected in 2010 and the entire data set is arranged in each 15 minutes time interval. Then the data set is computed though Empirical Mode decomposition Process (EMD) for decomposition. The output of EMD is IMFs i.e; decomposed data transferred to the ELM. The performance of EMD-ELM as compared to original data and ELM are shown in figures. The mean absolute percentage error (MAPE), root mean square error and mean absolute error (RMSE) and mean absolute error (MAE) by ELM are 5.115% and the 5.231% and 4.381% respectively. The errors computed by EMD-ELM are 3.8512%, 4.165% and 3.672%, respectively. Further the use of the P-ELM provides the smaller errors like MAPE= 3.351%, RMSE=3.912% and MAE= 3.121%. Also the respective errors for EMD-PELM are given by 3.002%, 3.243% and 2.942%. But by proper selection of input weights of the extreme learning machine, one can minimize the forecasting errors. Thus weights are optimized by couple based particle swarm algorithm and then computed. The respective errors are 2.862%, 3.031% and 2.468%, as shown in the following figures.

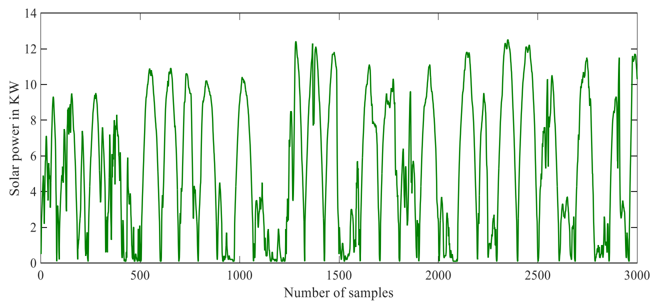

Fig.7: Original data with 15 minute interval

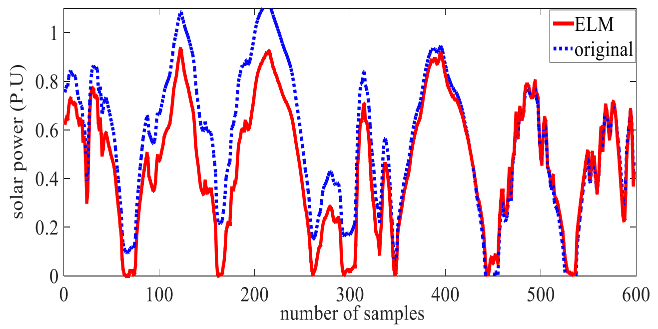

Fig.9: Performance of ELM in 15 minute prediction

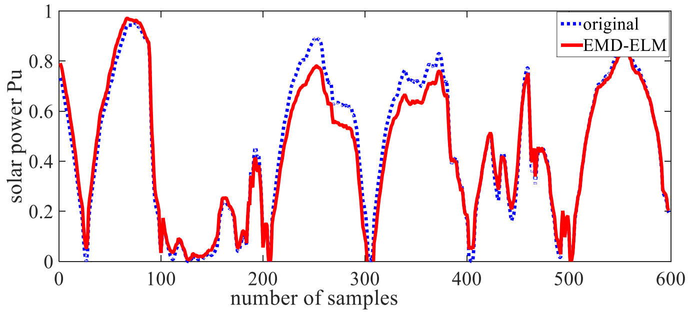

Fig.10: Performance of EMD-ELM in 15 minute prediction

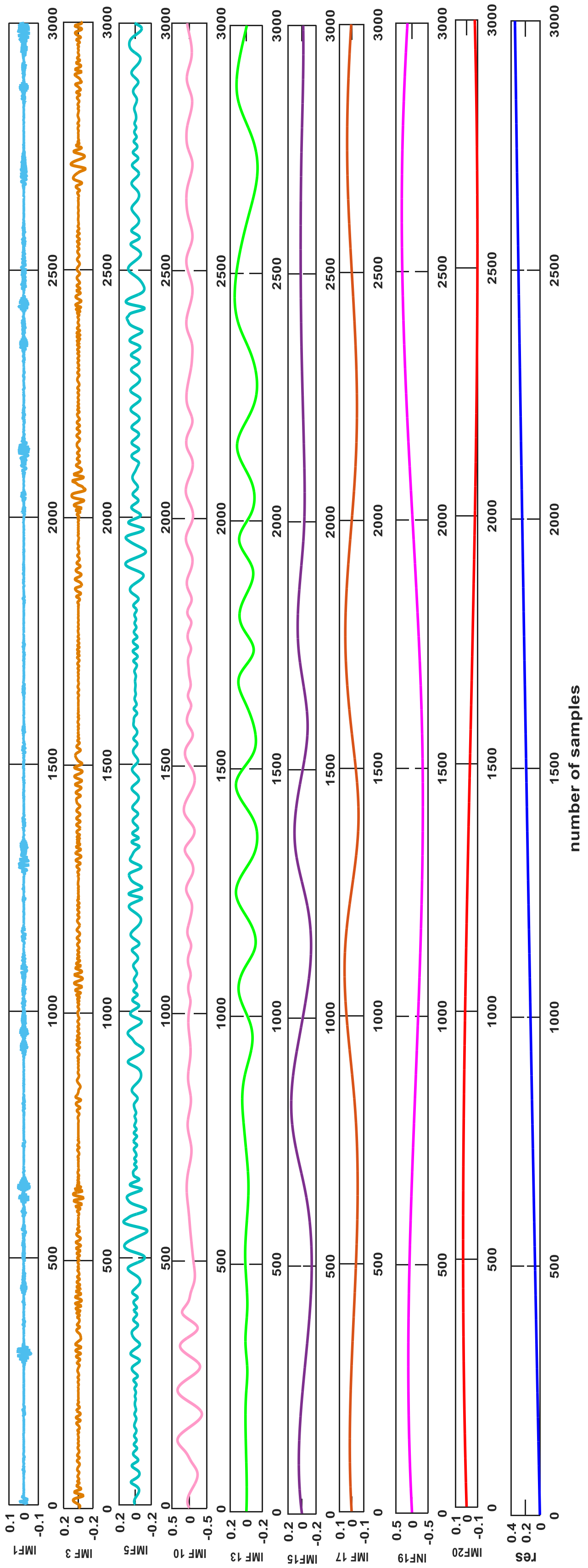

Fig.8: IMFs of the data with EMD in 15 minutes time interval

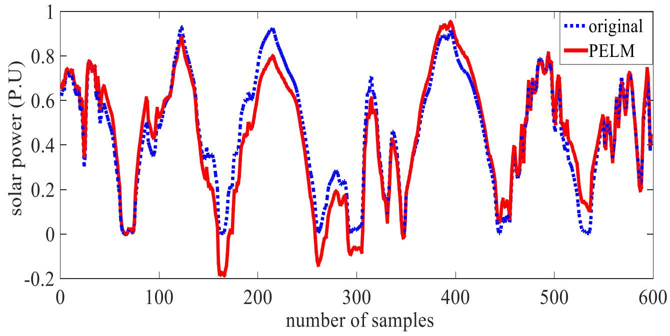

Fig.11: Performance of P-ELM in 15 minute prediction.

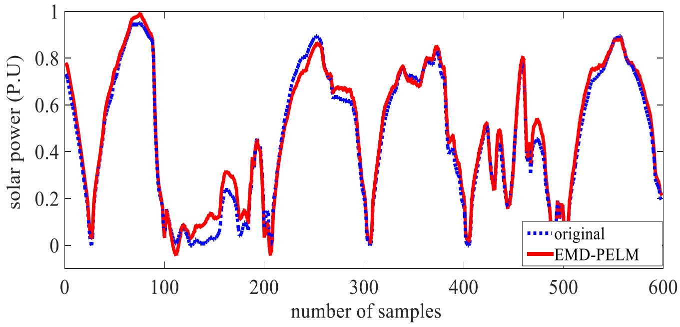

Fig.12:Performance of EMD- P-ELM in 15 minute prediction.

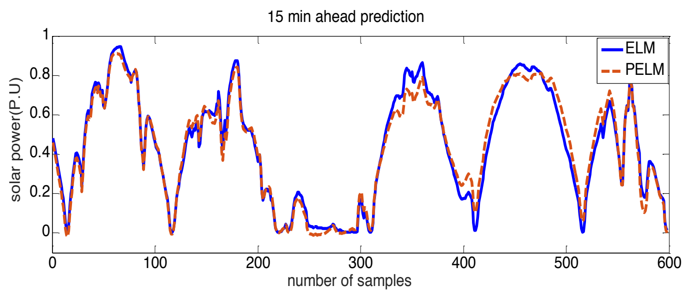

Fig.13: Performance comparison between ELM and PELM in 15 minute prediction.

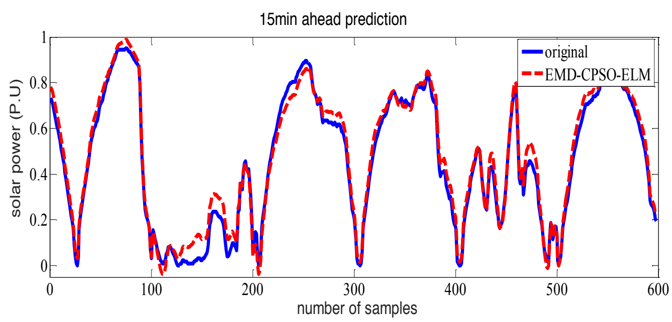

Fig.14: Performance of EMD-CPSO-PELM in 15 minute prediction

Case-2: Solar power forecasting for 30 minutes Horizon:

Here the solar output data of the real time plant (mentioned in table-1), in the year of 2010 is divided into 30 minutes time horizon and then the empirical decomposition mode technique is applied for signal decomposition in order to improve the performance and stability of the system. The output of EMD is IMFs i.e.; decomposed data passed to the ELM model. The outcome of EMD-ELM as compared to original data and ELM are as shown in figures. The measuring indices like mean absolute percentage error (MAPE), root mean square error and mean absolute error (RMSE) and mean absolute error (MAE) are 7.018%, 6.661% and 5.481 respectively by the extreme learning machine technique. The errors calculated by the EMD-ELM are 6,002%, 4.872% and 4.678%, respectively. Further the use of the P-ELM provides the smaller errors like MAPE= 4.857%, RMSE=3.614% and MAE= 3.526%. Further these errors in case of EMD-PELM are given by 4.0821%, 3.115% and 3.212 respectively. But by proper selection of input weights of the extreme learning machine, one can minimize the forecasting errors. Thus weights are optimized by couple based particle swarm optimization algorithm and then computed. The respective errors are 3.232%, 2.931% and 2.568%, as shown in the following figures.

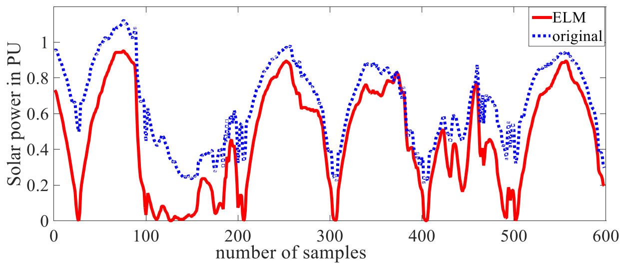

Fig.15: Performance of ELM in 30 min. Interval prediction

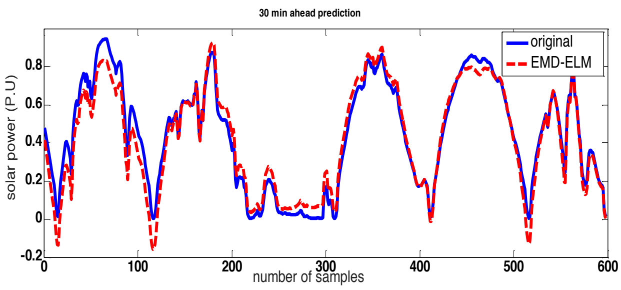

Fig.16: Performance of EMD-ELM in 30 min. Interval prediction

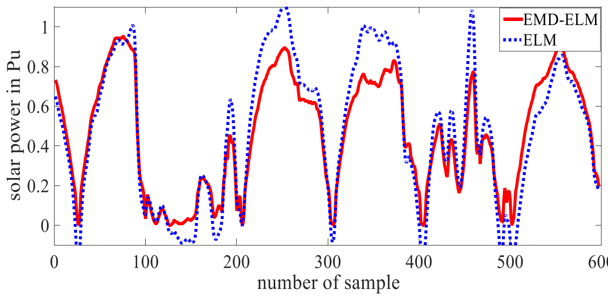

Fig.17: Performance of EMD-ELM copared with ELM in 30 min. Interval prediction

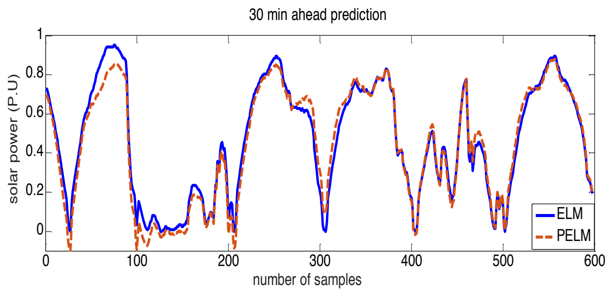

Fig.18: Performance of PELM compared to ELM in 30 min. Interval prediction

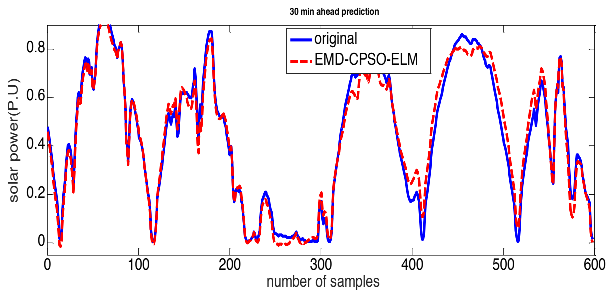

Fig.19: Performance of EMD-CPSO-ELM in 30 min. Interval prediction.

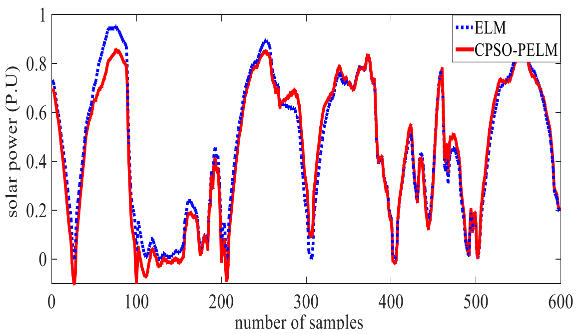

Fig.20: Performance of CPSO-PELM as compared to ELM in 30 min. Interval prediction.

Case-3: Solar power forecasting for 60 minutes Horizon:

The output of PV power plant (mentioned in table-1) is divided into 60 minutes time horizon and then the forecasting models like Extreme learning machine(ELM), Empirical mode decomposition based extreme learning Machine (EMD-ELM), Pruned Extreme learning machine (PELM), EMD based PELM and EMD-CPSO PLM are applied and simulated. The figures 21-25 depicts the performance of different machine learning techniques. The measuring indices like mean absolute percentage error (MAPE), root mean square error (RMSE) and mean absolute error (MAE) are 7.38% and the 7.061% and 6.281% respectively by the extreme learning machine technique. The errors calculated by the EMD-ELM are 6.312%, 5.872% and 5.678%, respectively. Further the use of the P-ELM provides the smaller errors like MAPE= 5.157%, RMSE=4.514% and MAE= 3.826%. Further these errors in case of EMD-PELM are given by 4.882%, 3.115% and 3.912% respectively. But by proper selection of input weights of the extreme learning machine, one can minimize the forecasting errors. Thus weights are optimized by couple based particle swarm algorithm and then computed. The respective errors are 4.332%, 2.932%.and 3.668%, as shown in the following figures.

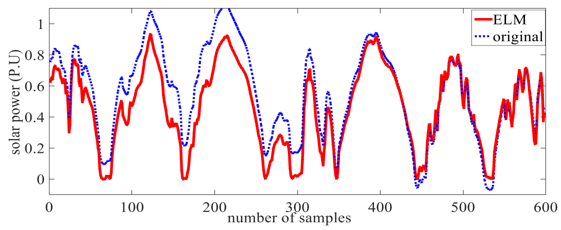

Fig.21: Performance of ELM as compared to original data in 60 min. Interval prediction

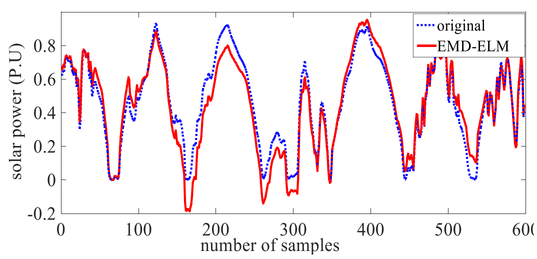

Fig.22: Performance of EMD-ELM as compared to Original in 60 min. Interval prediction

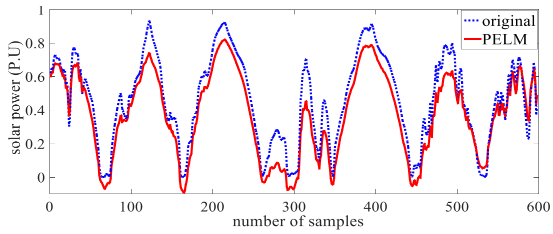

Fig.23: Performance of P-ELM as compared to original in 60 min. Interval prediction

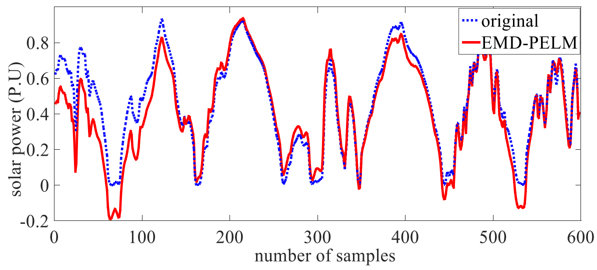

Fig.24: Performance of EMD-PELM as compared to original in 60 min. Interval prediction

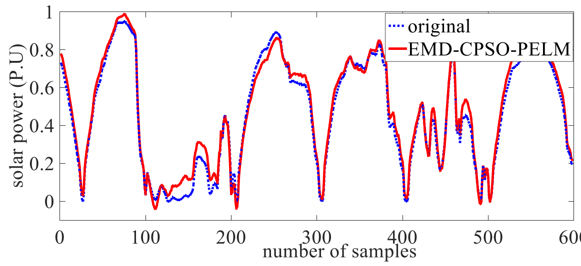

Fig.25: Performance of EMD-CPSO-PELM as compared to original in 60 minutes Interval prediction.

Table 2: Comparison of output results

Forecasting models	15 Minutes time horizon			
	MAPE(%)	RMSE(%)	MAE(%)	CC^2
ELM	5.115	5.231	4.381	0.978
EMD-ELM	3.851	4.165	3.672	0.963
P-ELM	3.351	3.912	3.121	0.904
EMD-PELM	3.002	3.243	2.942	0.968
CPSO-PELM	2.862	3.031	2.468	0.984

Forecasting models	**30 Minutes time horizon**			
	MAPE(%)	RMSE(%)	MAE(%)	CC^2
ELM	7.018	6.661	5.481	0.909
EMD-ELM	6.002	4.872	4.678	0.906
P-ELM	4.857	3.618	3.526	0.918
EMD-PELM	4.0821	3.115	3.212	0.903
CPSO-PELM	3.232	2.931	2.568	0.925
Forecasting models	**60 Minutes time horizon**			
	MAPE(%)	RMSE(%)	MAE(%)	CC^2
ELM	7.38	7.06	6.28	0.898
EMD-ELM	6.312	5.872	5.678	0.873
P-ELM	5.157	4.514	3.826	0.902
EMD-PELM	4.882	3.115	3.912	0.890
CPSO-PELM	4.332	2.932	3.668	0.899

7. Conclusion

The solar energy forecasting is an integral part of grid energy monitoring, which reduces various power system problems. The weights of different forecasting models are selected with trial and error basis which affects the accuracy in forecasting. To reduce the forecasting error and to increase the accuracy and involve proper management facility, a new method like Pruned Extreme learning machine (P-ELM) with Empirical Mode Decomposition (EMD) has been projected in this paper. The weights of the PELM are optimized by couple based particle swarm optimization (CPSO). Various experimental verifications were performed in this paper. From the result analyses it is depicted that the proposed method has established itself to be superior to the classical techniques implemented beforehand. In future long term data forecasting can be performed for the purpose of new power plant setup.

References

[1] J. Kelleher, J.V. Ringwood, ‘A computational tool for evaluating the economics of solar and wind micro generation of electricity, Energy, Vol. 34,2009, PP. 401–409
[2] Bhubaneswar Paridaa, S. Iniyanb,∗, Ranko Goicc, ‘A review of solar photovoltaic technologies’ Renewable and Sustainable Energy Reviews, Elsevier, Vol.15 ,2011, PP.1625–1636.
[3] M. Rekinger, F. Thies, G. Masson, S. Orlandi, Global Market Outlook for Solar Power 201-219, Solar Power Europe.
[4] X. Luo, J. Wang, M. Dooner, J. Clarke, Overview of current development in electrical energy storage technologies and the application potential in power system operation, vol. 137 , Appl. Energy,2015, PP. 511-536.
[5] R.H. Inman, H.T. Pedro, C.F. Coimbra, Solar forecasting methods for renewable energy integration, Prog. Energy Combust. Vol.39 No.6,2013, pp.535-576.
[6] Antonio Bracale, Pasquale De Falco, ‘An Advanced Bayesian Method for Short-Term Probabilistic Forecasting of the Generation of Wind Power ’ Energies,Vol.8, 2015, PP. 10293-10314.
[7] Giovanni Brusco, Alessandro Burgio, Daniele Menniti, Anna Pinnarelli, Nicola Sorrentino, ‘Quantification of Forecast Error Costs of Photovoltaic Prosumers in Italy ’ Energy, MDPI, Vol. 10, 2017, PP.1-17.
[8] Bouzerdoum, M., A. Mellit, and A. Massi Pavan. "A hybrid model (SARIMA–SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant." *Solar Energy*98 (2013): 226-235.
[9] Seyyed A. Fatemi , Anthony Kuh, Matthias Fripp, ‘Parametric methods for probabilistic forecasting of solar irradiance’ Renewable Energy, Elsevier, Vol.129,2018, PP. 666-676.
[10] Chen, Changsong, Shanxu Duan, Tao Cai, and Bangyin Liu. "Online 24-h solar power forecasting based on weather type classification using artificial neural network." *Solar Energy* 85, no. 11 (2011): 2856-2870.
[11] S. Al-Alawi, H. Al-Hinai, ‘An ANN-based approach for predicting global radiation in locations with no direct measurement instrumentation’ Renewable Energy, Elsevier, Vol.14,1998, PP.199-204.
[12] Şahin, Mehmet, Yılmaz Kaya, Murat Uyar, and Selçuk Yıldırım. "Application of extreme learning machine for estimating solar radiation from satellite data." *International Journal of Energy Research*, Vol.38, No. 2,2014, PP.205-212.
[13] Mehmet Şahin ,Yılmaz Kaya,Murat Uyar Selçuk Yıldırım, ‘Application of extreme learning machine for estimating solar radiation from satellite data’ International Journal of Energy Research, Wiley, Vol.38, No.2, 2014, PP.205-212.
[14] HeJiang YujieWu,Yao Dong, ‘ Forecasting Solar Radiation Using an Optimized hybrid model by

Cuckoo search algorithm. 'Energy,Vol. 81, 2015, PP.627-644.

[15] Pan, Tien-Szu, Thi-Kien Dao, and Shu-Chuan Chu. "Hybrid particle swarm optimization with bat algorithm." In *Genetic and evolutionary computing*, pp. 37-47. Springer, Cham, 2015.

[16] Banos, Raul, Francisco Manzano-Agugliaro, F. G. Montoya, Consolacion Gil, Alfredo Alcayde, and Julio Gómez. "Optimization methods applied to renewable and sustainable energy: A review." *Renewable and Sustainable Energy Reviews* Vol.15, No. 4,2011, PP.1753-1766.

[17] Lin, Whei-Min, Chia-Sheng Tu, and Ming-Tang Tsai. "Energy management strategy for microgrids by using enhanced bee colony optimization." *Energies* Vol.9, No. 1,2015, PP.1- 5.

[18] Wang, Jianzhou, He Jiang, Yujie Wu, and Yao Dong. "Forecasting solar radiation using an optimized hybrid model by Cuckoo Search algorithm." *Energy* Vol.81,2015, PP.627-644.

[19] Chiroma, Haruna, Nor Liyana Mohd Shuib, Sanah Abdullahi Muaz, Adamu I. Abubakar, Lubabatu Baballe Ila, and Jaafar Zubairu Maitama. "A review of the applications of bio-inspired flower pollination algorithm." *Procedia Computer Science* Vol.62,2015, PP.435-441.

[20]] Chiroma, Haruna, Abdullah Khan, Adamu I. Abubakar, Younes Saadi, Mukhtar F. Hamza, Liyana Shuib, Abdulsalam Y. Gital, and Tutut Herawan. "A new approach for forecasting OPEC petroleum consumption based on neural network train by using flower pollination algorithm."*Applied Soft Computing* ,Vol.48 ,2016, PP.50-58.

[21] Guo Pana,Aijia Ouyangc, 'using CPSO-GM Model Medium and Long Term Power Load Forecasting ' JOURNAL OF NETWORKS, IEEE, VOL. 9, NO. 8, 2014, PP. 2121-2128.

[22] Yudong Zhang, Shuihua Wang and Genlin Ji, ' A Comprehensive Survey on Particle Swarm Optimization Algorithm and Its Applications' Mathematical Problems in Engineering, Hindwai Publishing Corporation,Vol.2015, PP.1-38.

[23] David Motyka;Martina Kajanova;Peter Bracinik, The Impact of Embedded Generation on Distribution Grid Operation. 2018 7th International Conference on Renewable Energy research and Applications (ICRERA),Paris France.

[24] Abdelkader HARROUZ, Meriem ABBES, Ilhami COLAK, Korhan KAYSILI; Smart Grid and Renewable Energy in Algeria; 2017 6th International Conference on Renewable Energy research and Applications (ICRERA), San Deigo USA.

[25] Yu Shimizu,Taichiro Sakagami, Hiraoki Kitano; Prediction of weather Dependant Energy Consumption of Residential Housing; 2017 6th International Conference on Renewable Energy research and Applications (ICRERA), San Deigo USA.

[26] Luis Felipe Normandia Lourenco, Mauricio Barbosa de Camargo Salles; Matheus Mingatos Fernandes Gemignani, Marcos Roberto Gouvea, Nelson Kagan; Time Series Modelling for Solar Irradiance Estimation in Northeast Brazil. 2017 6th International Conference on Renewable Energy research and Applications (ICRERA), San Deigo USA.

[27] Rendani Mbuvha, Mattias Jonsson, Niclas Ehn and Pawel Herman; Bayesian Neural Networks For One-hour Ahead Wind Power Forecasting; 2017 6th International Conference on Renewable Energy research and Applications (ICRERA), San Deigo USA.

[28] Mehmet Yesilbudak, Medine Colak, Ramazan Bayindir, What are the Current Status and Future Prospects in Solar Irradiance and Solar Power Forecasting? International Journal of Renewable Energy Research, Vol. 8, No. 1 (2018) March

[29] Jun Li,Wanliang Fang, Xudong Zhang, Chunxiang Yang, An Improved Photovoltaic Power Forecasting Model With the Assistance of Aerosol Index Data. IEEE Transactions on Sustainable Energy, Volume: 6 , Issue: 2 , April 2015.

[30] Yu Liu ; Zhi Li ; Kai Bai ; Zhaoguang Zhang ; Xining Lu ; Xiaomeng Zhang, Short-term power-forecasting method of distributed PV power system for consideration of its effects on load forecasting. The Journal of Engineering Volume: 2017 , Issue: 13 , 2017.

[31] Can Wan ; Jin Lin ; Yonghua Song ; Zhao Xu ; Guangya Yang Probabilistic Forecasting of Photovoltaic Generation: An Efficient Statistical Approach. IEEE Transactions on Power Systems Volume: 32 , Issue: 3 , May 2017