Solar power prediction approach using data augmented deep learning technique

Ajaya Kumar Parida [1*], Dr. Deepak Kumar [2] , Dr. Rashmi Ranjan Sahoo [3], Dr. Bunil Kumar Balabantaray [4],

[1]Dept. of Computer Science and Engineering, Nit Meghalaya, Meghalaya, India

[1]School of Computer Engineering, Kalinga Institute of Industrial Technology Deemed to be University, Bhubaneswar, Odisha-751024, India.

[2] Dept. of Computer Science and Engineering, Nit Meghalaya, Meghalaya, India

[3] Department of CSE, Parala Maharaja Engineering College,Odisha, India

[4] Dept. of Computer Science and Engineering, Nit Meghalaya, Meghalaya, India

Email: [1] p21cs014@nitm.ac.in , [2] deepak.kumar@nitm.ac.in, [3] rashmiranjan.cse@pmec.ac.in, [4] bunil@nitm.ac.in

Abstract:

Solar power prediction holds a significant impact for future renewable energy scenario. Constructing a precise and reliable prediction model for more accurate solar power prediction is of utmost importance. To achieve a more accurate predicted output a novel prediction technique has been included in this paper for short and medium term solar power prediction. Initially the original solar power is decomposed into a set of subseries using VMD based decomposition technique. Data augmentation technique is applied for generating more training data thus avoiding the problem of over fitting. A novel prediction model based on long short term memory and multi kernel based random vector functional link network is proposed for point prediction of short term and medium term solar power. A fuzzy entropy based strategy is implemented for partitioning the subseries and assigning it to LSTM and MKRVFLN. For further improvement in prediction, WCO technique is used for obtaining optimised kernel parameters. The performance of the proposed technique is compared with seven other prediction techniques. Solar power data from two different sites are considered for comparison purpose. The experiment is performed on two aspects: short term and medium term point prediction, the result analysis shows that the proposed solar power prediction model shows excellent result.

Keywords: *fuzzy entropy, variational mode decomposition, long short term memory, multi-kernel random vector functional link network, data augmentation, water cycle optimization.*

Nomenclature:

AI	Artificial Intelligence
AE	Auto Encoder

ANN	Artificial Neural Network
ARMA	Auto Regressive Moving Average
ARIMA	Auto Regressive Integrated Moving Average
BP	Back Propagation
CNN	Convolutional Neural Network
DNN	Deep Neural Network
DA	Data Augmentation
DL	Deep Learning
EMD	Empirical Mode Decomposition
EWT	Empirical Wavelet Transform
ELM	Extreme Learning Machine
FE	Fuzzy Entropy
IMF	Intrinsic Mode Function
RBFNN	Radial Basis Functional Neural Network
RVFLN	Random Vector Functional Link Network
MKRVFLN	Multi kernel based random vector functional link network
RNN	Recurrent Neural Network
RMSE	Root Mean Square Error
RES	Renewable Energy Resources
MAPE	Mean Absolute Percentage Error
MAE	Mean Absolute Error
ML	Machine Learning
NWP	Numerical Weather Prediction
VMD	Variational Mode Decomposition
SVM	Support Vector Machine
LSTM	Long Short Term Memory
WT	Wavelet Transform
WCO	Water Cycle Optimization

1. Introduction

Reduction in fossil fuel energy and environmental severity has increased the use of different renewable energy sources. Solar thermal and solar PV has gained more importance because of being clean source of energy. PV systems have undergone from ordinary usage to large-scale integration. It is the most desirable and widely used RES due to its environmentally friendly nature and high productive nature. Solar photovoltaic is considered to be one of the most promising RES and found a better application in smart grid's power penetration [1,2]. The nonlinear and stochastic behaviour of PV power generation increases the complexity of smart grid energy management. Scheduling of the grid due to various factors like frequency disturbances, reverse power flow, voltage fluctuation, etc. [3-6], hence a precise prediction of solar power is in demand to improve the prediction quality. Precise prediction plays a vital role efficient integration of the energy management system based on various RES's [7]. Therefore we can conclude that the prediction of solar power guarantees more economic integration of the PV, better planning of the power plant. This paper aims at proposing a prediction model

for robust and accurate power prediction in different weather conditions. Broadly considering, forecasting can be categorised based on different forecast horizon as given below:

i. Very short term
ii. short term
iii. medium term
iv. long term

Different time horizon classifies the prediction technique and accordingly determines the functions based on prediction interval. Figure.1 shows a detailed structure of different prediction horizon and functions based on that.

The NWP depends upon physical properties, but are not limited to solar radiation, temperature, cloud volume, air pressure, wind speed etc. The model accuracy is dependent on the accuracy of the metrological data. In ref [8] the authors describe a statistical method that detects the motion of the clouds based on the satellite images. Different other studies were also performed by the researchers, as shown in ref [9,10,11]. Recent studies are performed for improving the accuracy of the NWP method for better prediction [12]. The main drawback of the NWP model is its dependency on the detailed information of the geographical location and correctness of the NWP particulars hence often subjected to various man-made errors. It is dependent on the cognition degree and various other factors that vary for different locations, thus making the method less robust [13]. To overcome these problems, and obtain a more accurate and robust prediction model even without the precise knowledge of the weather parameters the statistical methods came into existence. Widely implemented statistical methods for predicting RES mostly include the time-series prediction methods depending on the historical data. The different statistical model includes persistence model, gray theory, ARMA, ARIMA, ARMAX, CARD, Markov model, etc. [14-19]. In ref [20] ARMA and GARCH methods are combined to forecast solar irradiance. The main difference between the statistical model and the previously used physical model is that it does not require a thorough understanding of the PV system's complex mechanism and proves to be way simpler in functioning. The major drawback of this technique is while handling the nonlinear data and resulting in less accurate prediction. The efficiency of the model decreases when subjected to a larger number of input data. The complicated mathematical calculation of the statistical model increases the computational timing and complexity, thus motivating the researchers to develop a new technique that mitigates the problem of large computational timing and increases the prediction accuracy by reducing the mathematical calculation. The AI technique gained wide attention in various fields of science like prediction, image processing, and speech recognition, etc. fuzzy

logic [21-23], and machine learning techniques were mostly implemented in the modern day's study for accurate prediction of power and irradiation. The machine learning techniques gained more attention due to their ability to extract high-dimensional nonlinear features [24, 25]. Initially, the researchers used the ANN method for solar irradiance and power prediction [26-28] although they are capable of predicting the solar power accurately but suffered from the disadvantage of generalization, local minima and maxima, reduced convergence speed, and over fitting problem and showed unstable nature subjected to little difference in the real time data series and attributes selection, resulting in inaccurate prediction. To overcome the basic problems of ANN other machine learning techniques came into use, BP, RBFNN, SVM, RVM, ELM [29-34]. The ELM technique became the most common and widely implemented technique, initially proposed by Huang wt al [33]. It is a single-layer feed-forward neural network (SLFN) with randomness in weight. Further modifications were made by different researchers to improve the prediction accuracy of the model by reducing the randomness of the weight matrix. In addition to different variations and modifications to the ELM another similar functioning technique known as the Random Vector Functional Link Network was introduced and proved to be more efficient due to the extra direct link between the input and output layer [35]. The machine learning techniques exhibited various shortcomings like inaccuracy and larger processing time when subjected to high dimensional problems, therefore, enforcing a more effective learning technique. DNN proposes a higher potential for feature extraction and thus shows improved performance than the shallow networks. The DNN comprises CNN, LSTM, DBN, and SAE are mostly applied for solving the complex problem of nonlinearity, LSTM is of the RNN family but does not contain the vanishing problem [36, 37]. In recent work DL along with other advanced algorithm has attained success in numerous field of application. DNN models relationship between different nonlinear structures thus helping in feature selection. The volume and quantity of training data set determines the effectiveness of the deep learning. An effective solution for establishing larger data set is by data augmentation. Data augmentation makes supervised machine learning more efficient by adding some modified version of the already existing data set based on the available data. This helps in assisting and mitigating over fitting problem in deep learning thus improving the model. Several methodologies have been introduced in computer science using data augmentation but mostly used for image recognition purpose [38,39] but very little attention have been given for prediction purpose. In recent time deep learning techniques have gained much attention for researchers in the field of predicting renewable energy.

However the kernel parameter selection plays a vital role as it affects the predicted output, thus can be optimized for providing more accurate result. WCO is used to optimise the kernel parameter. In summary we can thus say a new hybrid VMD- DA, FE subseries based LSTM-KRVFLN. The main contribution of this paper can be stated as follows:

1. Data pre-processing using both VMD and DA thus reducing the non-stationary nature of the original time series solar power data.

2. Solar power prediction based on LSTM-KRVFLN is used along with FE based partitioning strategy for subseries.

3. An optimising technique is implemented for selecting optimal kernel parameter.

4. A novel hybrid based on VMD, DA, FE subseries partitioning based, LSTM-MKRVFLN, where WCO is proposed to determine the optimal kernel parameter. Two solar power datasets are considered for solar power prediction. The experimental result shows that the proposed hybrid method shows outstanding performance in solar power prediction.

Rest of the paper is arranged as follows: The basic knowledge of different techniques (like KRVFLN, LSTM, VMD, DA and FE) implemented in this study is explained. Section 2. Describes the background of different methods performed in this study. Section 3 explains the proposed hybrid model including data pre-processing and subseries partitioning using FE. The detailed experimental performance including optimisation of, the parameter is shown in section 4. Followed by the case study in section 5 and conclusion in section 6.

2. Methodology

This section describes different methodologies used in this paper for more accurate prediction of solar power. The result analysis shows the performance accuracy of the proposed model as compared to other models.

2.1 Long Short Term Memory

LSTM is a part of the RNN family, the main concept behind the RNN technique is the feedback layer i.e. the output of the model not only depend upon the input data but also on the previous output or hidden layer. The main problem of this technique is that the back propagated data that either adds or deducts at each step leading to a vanishing data after observing for several steps. The LSTM technique solved the problem of data vanishing [38,39]. Stacked LSTM was also used in a previous research performed by the author for application of medium

term solar power prediction [40]. Unlike the traditional neural networks LSTM is composed of memory blocks instead of neurons connected between the layers. Individual blocks consists of gates that handles the output of the blocks, it consists of three gates for handling the blocks namely the input gate, the forget gate and the output gate. The basic structure of the LSTM is shown in figure.1. Where it can be observed that after the input variables are fed each gate utilizes the activation unit for decision making purpose; the weights present in the gates are learned during the training phase thus making it more efficient. The basic steps of the gates can be summarised as:

i. Input gate: In this step the input activation is controlled, it helps in updating the memory state on the basis of certain condition; it helps in updating the memory state based on different condition.
ii. Forget gate: It forgets/ discard the useless information i.e. it forgets the information of the past cell.
iii. Output gate: Provides the ultimate output state based on the information's from input gate and forget gate

The nodes can be mathematical represented as:

$$i_{nt} = \delta\left(W_{xi} * X_{tt} + W_{hi} * h_{tt-1} + W_{ci} * c_{tt-1} + b_i\right) \quad (1)$$

$$f_{tg} = \delta\left(W_{xf} * X_{tt} + W_{hf} * h_{tt-1} + W_{cf} * c_{tt-1} + b_f\right) \quad (2)$$

$$c_{tg} = f_{tt} c_{tt-1} + i_t \tanh\left(W_{xc} X_{tt} + W_{hc} h_{tt-1} + b_c\right) \quad (3)$$

$$o_{ut} = \delta\left(W_{xo} X_{tt} + W_{ho} h_{tt-1} + W_{co} c_{tt} + b_o\right) \quad (4)$$

$$h_{tt} = o_{ut} \tanh\left(c_{tg}\right) \quad (5)$$

Eq 1-5 represents the formulation of different nodes where i_{nt}, f_{tg} and o_{ut} represents the input gate, forget gate and output gate respectively. The input of the LSTM is twofold that is the current sample x_{tt} and the previous hidden layer sample h_{t-1}, the cell state c_{t-1} is the internal source of each gate. δ and tan h are the activation function.

The LSTM adds up all the inputs incoming through various origins doped with a bias value, the gates are activated by inputting the total input into a logistic function using a *tanh* function. The cell state is multiplied using the activation of the forget gate f_t. the updated cell is processed through a tan h function and then multiplied with o_t that determines the ultimate LSTM output

as shown is eq 5. The main advantage of LSTM is that it preserves an internal memory cell for the entire life cycle for the purpose of establishing a temporal connection. Ct in connection with the immediate output and input at that instant determines the element that needs to be updated or erased. In comparison with normal RNN, the LSTM flexibly deals with the long and short time lag for corresponding tasks. Figure.1. gives a detailed structure of the LSTM.

Figure.1. basic structure of LSTM

2.2 Multiple Kernel Function based Random Vector Functional Link Network (MKRVFLN)

In this study the authors proposes a hybrid model comprising of both MKELM and RVFLN for predicting solar power in an effective manner. Traditionally the RVFLN has a characteristics of faster learning speed, simplicity of the architecture and favourable generation capabilities. It is structurally allied with the ELM with a deviation, with an uninterrupted link between the input and output nodes. Most challenging problem that needs to be dealt with is the selection of number of hidden nodes and feature mapping.

Figure.2. Basic RVFLN Structure

The main defect in the conventional RVFLN is drafting of the hidden nodes and mapping function. This improvisation provides better stability, prediction accuracy and generalization. [35]. the conventional RVFLN technique functions similar to SLFN, without any direct link connecting the layers. Unlike the traditional SLFN, the RVFLN have an uninterrupted interrelation within the input and output. Figure.2. shows the pictorial representation of RVFLN method. $h\left(w_j^T x_i + b_{a\,j}\right)$ is used for mapping of the hidden layer with the input layer, *h* sdenotes the activation function, while b_{aj} & w_j stands for bias and weight mapping input and hidden layer respectively. x_i represents the *i*th input pattern, where x_i is written as : $x_i = [x_{i1}, x_{i2}, \ldots\ldots\ldots\ldots, x_{iZ}]^T$.

A uninterrupted input-output connection connects the input and output. The basic model can be described as:

$$O(x_i) = \sum_{j=1}^{L} \beta_j h\left(w_j^T x + b_j\right) + \sum_{j=K+1}^{K+Z} \beta_j x_j \qquad (6)$$

Here K denotes the hidden layer, Z being the total number of input data, the trained dataset is designated by N[i=1,2,…N] and β_j [j=1,2,..,K] describes the output weight. The the direct

input output connection is described as β_{K+1} to β_{K+Z}. The activation function considered in this paper is tan h. eq (6) can be rewritten as following:

$$o_i = \sum_{j=1}^{K} \beta_j \tanh\left(w_j^T x_i + b_j\right) + \sum_{j=K+1}^{K+Z} \beta_j x_{i,j-K}\ ,\ i = (1,2,3,\ldots\ldots\ldots, N) \qquad (7)$$

N being the total sample

Further eq. (7) is written as

$$o_i = [h_1(x_i) \quad h_2(x_i)]\ \beta_p \qquad (8)$$

Where $h_1(x_i) = [h(w_1 . x_i + b_1), \ldots\ldots\ldots\ldots, h(w_K . x_i + b_K)]$,

$$h_2(x_i) = [x_{i,1}, \ldots\ldots\ldots\ldots\ldots, x_{i,K}]$$

And β_p represents N number of pattern's total output vector.

Thus for *N* input patterns the output can be obtained as

$$O_p = H\beta = [H_1 \quad H_2]\ \beta_p \qquad (9)$$

Here

$$H_1 = \begin{bmatrix} h_1(w_1, b_1, x_1) & \ldots\ldots & \ldots\ldots & h_1(w_K, b_K, x_1) \\ \ldots & \ldots\ldots & \ldots\ldots & \ldots \\ \ldots & \ldots\ldots & \ldots\ldots & \ldots \\ h_1(w_1, b_1, x_N) & \ldots\ldots & \ldots\ldots & h_1(w_K, b_K, x_N) \end{bmatrix}_{N \times (K+Z)}, H_2 = \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_N \end{bmatrix}, \qquad (10)$$

And $\beta_p = \begin{bmatrix} \beta_1 \\ \ldots \\ \ldots \\ \beta_{L+Z} \end{bmatrix}_{(K+Z)\times 1}$, $T = \begin{bmatrix} t_1 \\ \ldots \\ \ldots \\ t_N \end{bmatrix}_{N \times 1}$

(11)

$\xi = [\xi_1 \quad \xi_2 \ldots\ldots\ldots\ldots\ldots\ldots \xi_N]^T$ and $\xi_i = t_i - o_i$ for the *i*th pattern.

β is formulated as follows:

$$\beta_p = \begin{cases} \left(H^T H + \frac{I}{C} \right)^{-1} (H^T T) \text{ for } N \geq L + Z \\ H^T \left(HH^T + \frac{I}{C} \right)^{-1} T \text{ for } N < L + Z \end{cases} \quad (12)$$

$$\beta_p = H^T \left(\frac{I}{C} + W^2 HH^T \right)^{-1} W^2 T \quad (13)$$

Eq.(13) can be further simplified as

$$\beta_p = \begin{bmatrix} H_1^T \\ H_2^T \end{bmatrix} \left(\frac{I}{C} + W^2 [H_1 \ H_2] \begin{bmatrix} H_1^T \\ H_2^T \end{bmatrix} \right)^{-1} W^2 T \quad (14)$$

the final product for x_i input is observed as

$$O_{p_i} = [h_1(x_i) \ h_2(x_i)] \begin{bmatrix} H_1^T \\ H_2^T \end{bmatrix} \left(\frac{I}{C} + W^2 [H_1 H_1^T + H_2 H_2^T] \right)^{-1} W^2 T \quad (15)$$

Kernel trick applied to eq. (15) solves the problem of randomness the following definitions are used:

$$H_1 H_1^T = K_L(x_i, x_j), H_2 H_2^T = K_{NL}(x_{i,} x_j), \ \ i, j = 1,2,...N \quad (16)$$

K_L and K_{NL} describes the linear and non-linear kernel function. Hence desired KRVFLN function is rewritten as

$$O_p(x_i) = \left(\begin{bmatrix} K_L(x_i, x_1) \\ \vdots \\ K_L(x_i, x_N) \end{bmatrix} + \begin{bmatrix} K_{NL}(x_i, x_1) \\ \vdots \\ K_{NL}(x_i, x_N) \end{bmatrix} \right) \times \left(\frac{I}{C} + W^2 [K_L + K_{NL}] \right)^{-1} W^2 T \quad (17)$$

Linear kernel function:

$$K_L(x_i, x_j) = x_i^T x_j \quad (18)$$

Kernel functions executed in this study with the help of Mercer's theorem is described below:

- Polynomial kernel: $K_p(x_i,x_j) = (1 + x_i^T x_j / \delta^2)^a$ (19)

- Wavelet kernel: $K_w(x_i,x_j) = \cos\left(\frac{b\left\|(x_i - x_j)\right\|}{c}\right)\exp\left(-\frac{\left\|(x_i - x_j)\right\|^2}{d}\right)$

(20)

Polynomial and Gaussian kernel are considered as the global and local kernel respectively. Kernel functions (both global and local) are combined together forming the multi-kernel

$$\text{Multi-kernel: } K_m(x_i, x_j) = \mu_1 K_p(x_i, x_j) + (1 - \mu_1) K_w(x_i, x_j) \quad (21)$$

The following parameters are chosen depending upon various limits of values using WCO,

(the parameters being $\mu_1, \mu_2, \delta, a, b, c, d$)

Where, (i) μ - *[0 1] (ii)* δ- *[0 1], (iii) a- [0 2], (iv) b-[0 5], (v) c- [0 5] and (vi) d- [0 2].*

The main objective of selecting the kernel parameters optimally instead of random (trial and error) method is to obtain more accurate prediction with minimum computational time. In this paper WCO based technique is applied for optimization of different parameters of wavelet and polynomial kernel function which is further combined to form the multi kernel technique.

2.3 Variational Mode Decomposition (VMD)

The nonlinear nature of the solar energy (irradiation/ power) influences the accurate solar power prediction. This intricacy of the original solar data is improved by disintegrating the decomposed components of the data which demonstrates certain changes in the characteristic. Decomposition of the time series data is a statistical process that deconstructs the original data into several components, each representing an underlying pattern. The improvement in the constitutive series is because of the filtering nature of various decomposition techniques and hence resulting in more accurate prediction in combination with various prediction techniques. VMD is a non-recursive model that surveys different modes along with their respective central frequencies that is limited within a band width, further helps in reconstructing the primary data in least square method. Compared to EMD, the proposed decomposition technique (VMD) shows a better ability to denoise, thus segregate tones having similar frequencies. In this paper VMD decomposes the original solar irradiation, and thus obtaining the fundamental VMF. These modes combines to reform the original data. The key assumption of this technique is

decomposition of the data in a non-recursive manner for real time signal into various modes. '*X' (*fundamental signal*)* is sub divided to different modes within a specific bandwidth; (with '$X_{K'}$ modes). It is compulsory for every mode to bunch around a central pulsation ('$W_{K'}$) Individual mode maintains a specific sparsity. The following algorithm is used to obtain the bandwidth of individual mode

1. First apply Hilbert transform to the individual mode '$X_{hK'}$ and obtain the unilateral spectrum frequency.
2. An exponentially tuned frequency is mixed with its respective estimated centre frequency to shift the spectrum frequency to its baseband region.
3. H1 Gaussian smoothness obtained from the demodulated signals are used for estimation of the individual mode's bandwidth, which is squared L2 and termed as the norm of the gradient. Eq (22) is used to find the decomposed signal based on VMD.

Hilbert transformation is used in case of individual mode $X_{hK}(t)$ along with an correlated analytical signal

$$\left(\lambda(t)+\frac{j}{\prod t}\right)* X_{hk}(t) \qquad (22)$$

Equation (22) is modified by transferring the frequency spectrum of each mode to its corresponding central frequency.

$$\left[\left(\lambda(t)+\frac{j}{\prod t}\right)* X_{hk}(t)\right]_e^{-j\omega K^t} \qquad (23)$$

After estimating the bandwidth, the output constrained Variational problem is further modified.

$$\min_{\{X_K\},\{W_K\}}\left\{\sum_{k=1}^{K}\left\|\lambda_t\left[\left(\lambda_t+\frac{j}{\prod t}\right)* X_{hK}(t)\right]e^{-jwkt}\right\|_2^2\right\} \qquad (24)$$

Such that

$$\sum_{k=1}^{K} X_{hk} = X \qquad (25)$$

X represents the original signal which is further decomposed into subsets. X_{hK} represents all the nodes combined together $\{X_1, X_2,\ldots,X_K\}$.

K being the total modes, λ being Dirac distribution, whereas * is the convolution. The output unconstrained problem obtained is demonstrated below.

$$L(\{X_K\},\{W_K\},\lambda) := \alpha\sum_{k=1}^{K}\left\|\partial_t\left[\left(\lambda_t+\frac{j}{\prod t}\right)*X_K(t)\right]e^{-jwkt}\right\|_2^2 + \left\|X(t)-\sum_{k=1}^{k}X_k(t)\right\|_2^2 + \left\langle\delta(t),X(t)-\sum_{k=1}^{k}X_k(t)\right\rangle \quad (26)$$

Eq. (26) is attained by using ADMM.

$$X_k^{n+1} = \underset{X_K\in X}{\arg\min}\left\{\begin{array}{l}\alpha\left\|\partial(t)\left[\left(\lambda_t+\frac{j}{\prod t}\right)*X_K(t)\right]e^{-jwkt}\right\|_2^2 \\ +\left\|X(t)-\sum_i X_i(t)+\frac{\delta(t)}{2}\right\|_2^2\end{array}\right\} \quad (27)$$

$$X_K^{n+1} = \underset{\hat{X}_K,X_K\in X}{\arg}\left\{\begin{array}{l}\alpha\left\|j\omega\left[(1+\operatorname{sgn}(\omega+\omega_K))\hat{X}_K(\omega+\omega_K)\right]\right\|_2^2 \\ +\left\|\hat{X}(\omega)-\sum_i\hat{X}_i(\omega)+\frac{\delta(\omega)}{2}\right\|_2^2\end{array}\right\} \quad (28)$$

solution derived for X_K and W_K are as following

$$\omega_K^{n+1} = \underset{\omega_K}{\arg\min}\left\{\int_0^\infty(\omega-\omega_K)\left|\hat{X}_K(\omega)\right|^2 d\omega\right\} \quad (29)$$

frequency domain is converted from the central frequency

$$\hat{X}_K^{n+1}(W) = \frac{\hat{X}(W)-\sum_{i\neq K}\hat{X}_i(W)+\frac{\hat{\delta}(W)}{2}}{1+2\alpha(W-W_K)^2} \quad (30)$$

$$W_K^{n+1} = \frac{\int_0^\infty W\left|\hat{X}_K(W)\right|^2 dw}{\int_0^\infty\left|\hat{X}_K(W)\right|^2 dw} \quad (31)$$

Here $\hat{X}(W)$, $\hat{X}_i(W)$, $\hat{\lambda}(W)$ and $\hat{X}_K^{n+1}(W)$ performs as the Fourier transforms of $X(t)$, $X_i(t)$, $\delta(t)$ and $X_K^{n+1}(t)$ respectively and n is the total number of iteration.

The experimental result shows a detailed analysis of the proposed VMD technique.

2.4 Water cycle optimization (WCO)

WCO is a meta-heuristic algorithm working on the principal of water cycle. In this process the initial population is the raindrops, the other step includes evaporation, transpiration and condensation [40]. WCO is based on the mechanism of population, where the introductory population is considered to be stream and is selected randomly within a range of UB and LB.

Two problem statements are considered while considering the optimization problem (minimization problem and maximization problem). In this paper minimization problem is considered where we obtain the best/ optimized value. Sea is considered as the cost function. Cost function nearer to the present best value is observed as the river. Except the best value all other streams endure the functioning to their corresponding rivers and sea. A prime population matrix is generated before stating the optimization process. The initial kernel parameters are considered as the initial rain drops. In this study the error minimization function is considered as the cost function, the error minimization is performed as:

$$f(A) = \min imize\ f(A) = \min \frac{\sum_{r_v=1}^{N_{st}} \xi^2}{N_{st}} \tag{32}$$

ξ^2 is represented as following:

$$\xi = T - O \tag{33}$$

where *A* matrix can be described as

$$A^i = \left[a_1^i, a_2^i, \ldots, a_{op}^i\right] \tag{34}$$

$i = 1,2,\ldots, N_R$

N_R being the number of variables subjected to optimisation, N_{TR} represents total number of raindrops where each raindrop comprises of N_P variables. Population of rain drop (R_{dp}) is termed as (CR_{dp})

$$CR_{dp} = \begin{bmatrix} R_{dp_1} \\ R_{dp_2} \\ \vdots \\ R_{dp_{NTR}} \end{bmatrix} \tag{35}$$

$$CR_{dp} = \begin{bmatrix} a_1^1 & a_2^1 & \cdots & a_{NR}^1 \\ a_1^2 & a_2^2 & \cdots & a_{NR}^2 \\ \vdots & \vdots & \vdots & \vdots \\ a_1^{NTR} & a_2^{NTR} & \cdots & a_{NR}^{NTR} \end{bmatrix}_{NTR \times NR} \tag{36}$$

Let the number of river and streams be R_{RN} and S_{TN} respectively. This can be further formulated as:

$$R_{STN} = R_{RN} + 1 \tag{37}$$

$$S_{TN} = N_{TR} - R_{STN} \tag{38}$$

CR_{dp} matrix can be written as

$$CR_{dp} = \begin{bmatrix} sea \\ river_1 \\ river_2 \\ \vdots \\ stream\,R_{STN_1} + 1 \\ stream\,R_{STN_2} + 1 \\ \vdots \\ stream_{NTR} \end{bmatrix} \tag{39}$$

Stream NTR - number of stream that directly or indirectly flows to its respective river and sea. The pre-defined stream for its respective sea and river is formulated as follows.

$$N_{STR} = round\left[\left|\frac{CN_i}{\sum_{i=1}^{R_{STN}} CN_i}\right| * S_{TN}\right] \tag{40}$$

$$CN_i = CN_i - CN_{R_{STN+1}} \tag{41}$$

$$i = 1,2,3,\cdots,R_{S_{TN}}$$

N_{STR} represents the number of stream flowing to its predefined river or sea. Cost of the individual stream (R_{STN} *+1*) is subtracted from cost of the sea and river. The stream either moves towards the sea or towards the river by position alteration and movement towards the sea, Mathematical representation of the movement of the seas changing position is as follows.

$$M_{stream}(t+1) = M_{stream}(t) + rand \times Z \times \left(M_{sea}(t)\right) - \left(M_{stream}(t)\right) \tag{42}$$

The new position acquired by the stream that can flow towards the river can be explained as

$$M_{stream}(t+1) = M_{stream}(t) + rand \times Z \times \left(M_{river}(t)\right) - \left(M_{stream}(t)\right) \tag{43}$$

The up gradation of the position for the flow of stream to river is

$$M_{river}(t+1) = M_{river}(t) + rand \times Z \times (M_{sea}(t)) - (M_{river}(t)) \quad (44)$$

t and Z represents the current iteration and predefined value chosen between [1.5, 2] since

$$Z = rand \times rand(d) + \min(d) \quad (45)$$

Updated location of the initial flow of water becomes the raindrop's individual cost function. The improved version enhanced the performance of the initially connected sea and [41]. Next step of mutation in WCO, this step comprises of increased evaporation and rain. A smaller value is considered for converging purpose.

$d_{\max} \sim = 0$, initiates the evaporation process.

$$\left\| M_{sea} - M_{river}^{j} \right\| < d_{\max} \; or \; rand < 0.1 \quad j = 1,2,\ldots,(S_{TN} - 1) \quad (46)$$

$$\left\| M_{sea} - M_{stream}^{i} \right\| < d_{\max} \quad (47)$$

After attaining the predefined condition the new rain drop is established. The new stream is formulated as.

The mathematical representation of the new stream is

$$M_{stream} = LB + rand(1, N_p) \times (UB - LB) \quad (48)$$

$$d_{\max}(t+1) = d_{\max}(t) - (d_{\max}(t) / Max_{iter}) \quad (49)$$

Where Max_{iter}=1, 2,..., max number of iteration

Algorithm

STEP1. Selection of the initial parameters like N_{TR}, d_{max}, N_p and t_{max}

STEP2. Generate the initial population, and therefore calculate cost function for minimization of the error.

STEP3. Determining the intensity with which the river is flowing, sea and stream towards the river

STEP4. The optimal solution is obtained by exchanging the position of river and sea.

STEP5. River flowing to the sea followed by next iteration for attaining the next best solution

STEP6. Evaporation rate calculation

STEP7. d_{max} *reduction and attenuation of the stopping criteria, if not return to flow of stream*

STEP8. Obtaining the sea value, that is the optimized value

3. Proposed Technique

This section describes the proposed model where the DA is used to increase the number of input data for better prediction accuracy. Fuzzy entropy is also implemented for portioning the subseries, the result analysis shows the performance accuracy of the proposed technique over different other techniques. Figure.3. demonstrates the proposed technique.

Figure.3. flowchart of the proposed technique

3.1 Data Augmentation (DA)

More number of training data improves the prediction accuracy, hence DA is used to increase the number of time series solar power data. There are different DA processes such as window cropping, window warping, Gaussian noise injection, jittering, flipping and manipulation of the original time series data. In this research window warping is used to generate the synthetic solar power time series data.

The decomposition process enhances the prediction accuracy, the summation of the IMFs provides the final predicted output. The experimental result shows that the LSTM performs better for higher frequency components while RVFLN performs better for lower frequency components. In this study a novel technique is proposed where LSTM and RVFLN are used for higher and lower frequency components that improves the computation speed. The FE of each subseries is used to calculate the complexity and irregularity of the subseries. The FE based partition criterion is given below:

Calculation of FE for each decomposed series

FE_{IMFi}, $i=1,2,\ldots\ldots,k.$

Calculate the mean FE of total decomposed series

$$meanFE = \frac{1}{k}\sum_{i=1}^{k} FE_{IMFi} \qquad (50)$$

If mean FE of the ith decomposed series is less than FE then consider it for LSTM or else consider it for RVFLN.

3.2 Performance matrices

Different evaluation matrices are considered to calculate the performance accuracy of the given method. Various other models are also considered for comparison purposes. Performance hierarchy of different models are analysed based on the performance evaluation matrices. Equation (51) – (53) shows the different evaluation matrices

- Mean Absolute Percentage Error (MAPE)

$$MAPE = \left(\frac{1}{N} \sum_{i=1}^{N} \frac{\left| A_o(x_i) - P_o(x_i) \right|}{P_o(x_i)} \right) * 100 \qquad (51)$$

- Mean Absolute Error (MAE)

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| A_o\left(x_i\right) - P_o\left(x_i\right) \right| \qquad (52)$$

- Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left| A_o(x_i) - P_o(x_i) \right|^2} \qquad (53)$$

Where N is the total number of data, $A_o(x_i)$ represents the actual data, $P_o(x_i)$ represents the forecasted output.

4. Experimental Performance and Result Analysis

This section demonstrates a detailed study showing different experiment performed to obtain the performance matrices.

4.1 Data Collection

This section describes the two data set considered in this paper for experimental verification of different solar power prediction techniques. The PV system data are collected to verify the performance of different deep learning driven techniques. The geographical location and statistical information of the datasets considered in this study are given in table.1, where N is the entire sample, T_{tr} represents the training data while T_{ts} represents the testing data. The two solar power time series data set is shown in figure.4, for better view of the signal only 12000 samples are shown. In this study both short term and medium term solar power prediction is performed where in case of short term prediction the time interval is of 5min while that of medium term the time interval is of 1 day.

Table.1. Geographical and statistical information of the datasets

Figure.4. original solar power of two different dataset, (a) original PV power for winter season of Alabama, (b) Original solar power data of winter season of New York

4.2 Subseries partitioning and parameter optimization

This section describes the strategy of partitioning the VMD based subseries for LSTM and MKRVFLN. As discussed in the earlier section FE of each VMD based subseries is calculated and based on this value the appropriate subseries is divided. Figure.5. shows that the first 5 IMFs (IMF1-5) are assigned to the LSTM such that their FE is larger than mean FE ($\epsilon = 0.04$ for different season) and the remaining subseries (IMF6-10) is assigned to MKRVFLN such that their FE's are smaller than the mean FE. Figure.5. clearly demonstrates the partitioning condition (FE larger than or smaller than mean FE). The partitioning helps in improving the prediction accuracy. The original signal is subdivided into 10 subseries using the VMD technique as demonstrated in section 2.3. WCO is used to optimise the kernel parameters.

Figure.5. FE based subseries partitioning for Alabama at different weather condition, (a) winter season, (b) spring season, (c) summer season and (d) winter season

Table.2. shows the optimised kernel parameters. To establish the dominance of the presented method, six other techniques namely, RVFLN, MKRVFLN, WCO-MKRVFLN, VMD-MKRVFLN, VMD-DA-LSTM and VMD-MKRVFLN-WCO. For WCO the population size and number of iteration are set to be 40 and 100. DA is used to increase the number of samples of each subseries which in turn increases the prediction accuracy. To obtain a more accurate result the proposed optimization method is employed to optimize the kernel parameters (δ, a, b, c, d) for obtaining more accurate predicted solar power at a comparative less computational time.

Table.2. Optimized kernel parameters used for short term and medium term solar power prediction of two different dataset at different weather condition:

4.3 Result analysis

This section deals with the evaluation result of the point prediction of the proposed method. Table 3 and 4 shows the evaluation matrices described in section 3.3, which verifies the effectiveness of the proposed model over six other model on two dataset for both short and medium term prediction. The optimal prediction result of each dataset is highlighted in bold. Thus following conclusion can be drawn from table 3 and 4.

a) After comparing the prediction matrices obtained by the RVFLN, MKRVFLN, VMD-DA-MKRVFLN, VMD-DA-OMKRVFLN, LSTM, VMD- DA -LSTM, VMD- DA-WCO -LSTM and the proposed model (VMD-DA-FE-LSTM-OMKRVFLN) on two dataset, it is found that the hybrid models along with the VMD technique showed better prediction accuracy as compared to the single models. For example in case of Alabama dataset, the MAPE values of RVFLN, MKRVFLN, VMD-DA-MKRVFLN, VMD-DA-OMKRVFLN, LSTM, VMD- DA -LSTM, VMD- DA-WCO -LSTM and the proposed model for autumn season are 3.22,3.16, 2.94,2.21,2.18,1.76,1.14,1.11 and 0.46 respectively. The mean MAPE decreased in case of the proposed model as compared to the single method. Hybrid model shows better prediction accuracy

b) Considering the Alabama dataset itself it is noticed that the model with DA technique outperforms other prediction technique without DA. VMD and DA integrated LSTM performs better.

c) The decomposition technique increases the prediction accuracy by reducing the non-stationary nature of the signal.

d) The optimization technique although have less impact in case of prediction accuracy but decreases the computational time which is much beneficial when larger data is considered (larger data increases the prediction accuracy).

Figure.6 (a), (b), (c) and (d) shows the pictorial representation of different prediction technique for solar power prediction using the dataset from Alabama at different season in case of short term solar power prediction. In this case the proposed technique outperforms other methods implemented in this study.

(a) Comparison analysis of different prediction method for solar power prediction of autumn season at Alabama.

(b) Comparison analysis of different prediction method for solar power prediction of winter season at Alabama.

(c) Comparison analysis of different prediction method for solar power prediction of summer season at Alabama.

(d) Comparison analysis of different prediction method for predicting solar power of spring season at Alabama.

Figure.6. Comparison analysis of short term solar power prediction using different techniques using dataset from Alabama at different weather condition, (a) predicting solar power at autumn season, (b) predicting solar power at winter season, (c) predicting solar power at summer season, and (d) predicting solar power at spring season

Figure.7 and 8 is the pictorial demonstration of the MAPE values given in table 3 and 4 respectively, which clearly shows different error parameters.

Table.3 Comparative analysis of short term solar power prediction of Alabama for different season

Table.4 Comparative analysis of medium term solar power prediction of Alabama for different season

Figure.7 Differentiation between different MAPE values using different techniques for short term solar power prediction

Figure.8 Differentiation between different MAPE values using different techniques for medium term solar power prediction

Figure.9 demonstrates the comparative study of different prediction technique using the dataset from New York for short term solar power prediction. The winter season is only represented pictorially in this study. The numerical values for all other seasons are shown in table 5.

Figure.9. Comparision analysis of different prediction method for predicting solar power of winter season at New York

Table.5 Comparative analysis of short term solar power prediction of New York for different season

Table. 6 gives the error values for medium term solar power prediction at different seasons. Figure.10 and 11 shows the error parameters as demonstrated in table 5 and 6.

Table.6 Comparative analysis of medium term solar power prediction of New York for different season

Figure.10 Comparison analysis of MAPE values using different techniques for short term solar power prediction

Figure.11. Comparison analysis of MAPE values using different techniques for medium term solar power prediction

5. Conclusion

To calculate a reliable and accurate solar power prediction this paper proposes a novel hybrid technique that comprises of VMD, FE based partitioning of the subseries, DA, LSTM-MKRVFLN and WCO. The assessment is performed based on two cases for short term and medium term solar power prediction for validation of the proposed method. Seven other models and three evaluation matrices are considered for comparison purpose with the proposed technique to a proof the prediction accuracy of the afore mentioned models. The result analysis shows the proposed technique have a higher prediction accuracy as compared to other models.

In future different parameters like irradiance, temperature variation and partial shading condition can be considered while predicting the solar power using the proposed technique.

6. Acknowledgment

Our research area deals with implementation of different prediction models for more accurate solar power predictions and furthermore use in grid management. Stacked Lstm techniques has been previously used by the authors for solar power prediction for different time horizon as mentioned in reference.

7. Declarations

No funding was received to assist with the preparation of this manuscript.

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Reference

[1] Wan, C., Zhao, J., Song, Y., et al. "Photovoltaic and solar power forecasting for smart grid energy management." *CSEE Journal of Power and Energy Systems* 1.4 (2015): 38-46. DOI: 10.17775/CSEEJPES.2015.00046

[2] Chaudhary, P., Rizwan, M., "Energy management supporting high penetration of solar photovoltaic generation for smart grid using solar forecasts and pumped hydro storage system." *Renewable Energy* 118 (2018): 928-946. https://doi.org/10.1016/j.renene.2017.10.113

[3] Petinrin, J.O., Shaabanb, M., "Impact of renewable generation on voltage control in distribution systems." *Renewable and Sustainable Energy Reviews* 65 (2016): 770-783. https://doi.org/10.1016/j.rser.2016.06.073

[4] Shivashankar, S., Mekhilef, S., Mokhliset H., et al. "Mitigating methods of power fluctuation of photovoltaic (PV) sources–A review." *Renewable and Sustainable Energy Reviews* 59 (2016): 1170-1184. https://doi.org/10.1016/j.rser.2016.01.059

[5] Alam, M.S., Al-Ismail, F. S., Salem, A., et al. "High-level penetration of renewable energy sources into grid utility: Challenges and solutions." *IEEE Access* 8 (2020): 190277-190299. DOI: 10.1109/ACCESS.2020.3031481

[6] Ullah, S., Haidar, A. M. A., Paul Hoole et al. "The current state of Distributed Renewable Generation, challenges of interconnection and opportunities for energy conversion based DC microgrids." *Journal of Cleaner Production* 273 (2020): 122777. https://doi.org/10.1016/j.jclepro.2020.122777

[7] Ahmad, T., Zhang, H., Yan, B., "A review on renewable energy and electricity requirement forecasting models for smart grid and buildings." *Sustainable Cities and Society* 55 (2020): 102052. https://doi.org/10.1016/j.scs.2020.102052

[8] Hammer, A., Heinemann, D., Lorenzet, E. et al. "Short-term forecasting of solar radiation: a statistical approach using satellite data." *Solar Energy* 67.1-3 (1999): 139-150. https://doi.org/10.1016/S0038-092X(00)00038-4

[9] Escrig, H., Batlles, F.J., Alonso, J. et al. "Cloud detection, classification and motion estimation using geostationary satellite imagery for cloud cover forecast." *Energy* 55 (2013): 853-859. https://doi.org/10.1016/j.energy.2013.01.054

[10] Lorenz, E,. Kühnert, J., Heinemann, D., "Overview of irradiance and photovoltaic power prediction." *Weather matters for energy*. Springer, New York, NY, 2014. 429-454. https://doi.org/10.1007/978-1-4614-9221-4_21

[11] Bacher, P., Madsen, H., Nielsen, H. A., "Online short-term solar power forecasting." *Solar energy* 83.10 (2009): 1772-1783. https://doi.org/10.1016/j.solener.2009.05.016

[12] Voyant, C., Muselli, M., Paoli, C., et al. "Numerical weather prediction (NWP) and hybrid ARMA/ANN model to predict global radiation." *Energy* 39.1 (2012): 341-355. https://doi.org/10.1016/j.energy.2012.01.006

[13] Lima, F. J.L, Martins, F. R., Pereira, E. B., et al. "Forecast for surface solar irradiance at the Brazilian Northeastern region using NWP model and artificial neural networks." *Renewable Energy* 87 (2016): 807-818. https://doi.org/10.1016/j.renene.2015.11.005

[14] Huang, R., Huang, T., Gadh, R., et al. "Solar generation prediction using the ARMA model in a laboratory-level micro-grid." *2012 IEEE third international conference on smart grid communications (SmartGridComm)*. IEEE, 2012. https://doi.org/10.1109/SmartGridComm.2012.6486039

[15] Colak, I., Yesilbudak, M., Genc, N., et al. "Multi-period prediction of solar radiation using ARMA and ARIMA models." *2015 IEEE 14th international conference on machine learning and applications (ICMLA)*. IEEE, 2015. https://doi.org/10.1109/ICMLA.2015.33

[16] Alsharif, M. H., Younes, M. K., and Kim, J., "Time series ARIMA model for prediction of daily and monthly average global solar radiation: The case study of Seoul, South Korea." *Symmetry* 11.2 (2019): 240. https://doi.org/10.3390/sym11020240

[17] David, M., Ramahatana, F., Trombe, P.J., et al. "Probabilistic forecasting of the solar irradiance with recursive ARMA and GARCH models." *Solar Energy* 133 (2016): 55-72. https://doi.org/10.1016/j.solener.2016.03.064

[18] Ren, Y., Suganthan, P. N., Srikanth, N., "Ensemble methods for wind and solar power forecasting—A state-of-the-art review." *Renewable and Sustainable Energy Reviews* 50 (2015): 82-91. https://doi.org/10.1016/j.rser.2015.04.081

[19] Gao, S., He, Y., Chen, H., "Wind speed forecast for wind farms based on ARMA-ARCH model." *2009 International Conference on Sustainable Power Generation and Supply*. IEEE, 2009. https://doi.org/10.1109/SUPERGEN.2009.5348142

[20] Sivaneasan, B., Yu, C. Y., Goh, K. P., "Solar forecasting using ANN with fuzzy logic pre-processing." *Energy procedia* 143 (2017): 727-732. https://doi.org/10.1016/j.renene.2013.05.011

[21] Rizwan, M., Jamil, M., Kirmani, S., et al. "Fuzzy logic based modeling and estimation of global solar energy using meteorological parameters." *Energy* 70 (2014): 685-691. https://doi.org/10.1016/j.energy.2014.04.057

[22] Suganthi, L., Iniyan, S., Samuel, A. A., "Applications of fuzzy logic in renewable energy systems–a review." *Renewable and sustainable energy reviews* 48 (2015): 585-607. https://doi.org/10.1016/j.rser.2015.04.037

[23] Chen, S. X., Gooi H. B., Wang, M. Q., "Solar radiation forecast based on fuzzy logic and neural networks." *Renewable energy* 60 (2013): 195-201. https://doi.org/10.1016/j.renene.2013.05.011

[24] Lao, Z, Shen D., Xue, Z., et al. "Morphological classification of brains via high-dimensional shape transformations and machine learning methods." *Neuroimage* 21.1 (2004): 46-57. https://doi.org/10.1016/j.neuroimage.2003.09.027

[25] Shih, S.Y., Sun, F. K., Lee, H.Y., "Temporal pattern attention for multivariate time series forecasting." *Machine Learning* 108.8 (2019): 1421-1441. https://doi.org/10.1007/s10994-019-05815-0

[26] Yadav, A. K., Malik, H., Chandel, S. S., "Application of rapid miner in ANN based prediction of solar radiation for assessment of solar energy resource potential of 76 sites in Northwestern India." *Renewable and Sustainable Energy Reviews* 52 (2015): 1093-1106. https://doi.org/10.1016/j.rser.2015.07.156

[27] Neelamegam, P., Amirtham, V.A., "Prediction of solar radiation for solar systems by using ANN models with different back propagation algorithms." *Journal of applied research and technology* 14.3 (2016): 206-214. https://doi.org/10.1016/j.jart.2016.05.001

[28] Izgi, E, Öztopal, A., Yerli B., et el. "Short–mid-term solar power prediction by using artificial neural networks." *Solar Energy* 86.2 (2012): 725-733. https://doi.org/10.1016/j.solener.2011.11.013

[29] Mohandes, M., Balghonaim, A., Kassas, M. et al. "Use of radial basis functions for estimating monthly mean daily solar radiation." *Solar Energy* 68.2 (2000): 161-168. https://doi.org/10.1016/S0038-092X(99)00071-7

[30] Yu, C., Li, Y., Bao, Y., et al. "A novel framework for wind speed prediction based on recurrent neural networks and support vector machine." *Energy Conversion and Management* 178 (2018): 137-145. https://doi.org/10.1016/j.enconman.2018.10.008

[31] Wang, L., Li, X., Bai, Y., "Short-term wind speed prediction using an extreme learning machine model with error correction." *Energy Conversion and Management* 162 (2018): 239-250. https://doi.org/10.1016/j.enconman.2018.02.015

[32] Wang, K., Fu, W., Chen, T., et al. "A compound framework for wind speed forecasting based on comprehensive feature selection, quantile regression incorporated into convolutional simplified long short-term memory network and residual error correction." *Energy Conversion and Management* 222 (2020): 113234. https://doi.org/10.1016/j.enconman.2020.113234

[33] Hochreiter, S., Schmidhuber, J., "Long short-term memory." *Neural computation* 9(8) (1997): 1735-1780. https://doi.org/10.1007/978-3-642-24797-2_4

[34] Huang, G., Huang, G.B., Song, S., et al. "Trends in extreme learning machines: A review." *Neural Networks* 61 (2015): 32-48. https://doi.org/10.1016/j.neunet.2014.10.001

[35] Majumder, I., Dash, P. K., Bisoi, R., "Short-term solar power prediction using multi-kernel-based random vector functional link with water cycle algorithm-based parameter optimization." *Neural Computing and Applications* 32.12 (2020): 8011-8029. https://doi.org/10.1007/s00521-019-04290-x

[36] Gensler, A., Henze, J., Sick B., et al. "Deep Learning for solar power forecasting—An approach using AutoEncoder and LSTM Neural Networks." *2016 IEEE international conference on systems, man, and cybernetics (SMC)*. IEEE, 2016. https://doi.org/10.1109/SMC.2016.7844673

[37] Nasser, M. A., Mahmoud, K., "Accurate photovoltaic power forecasting models using deep LSTM-RNN." *Neural Computing and Applications* 31.7 (2019): 2727-2740. https://doi.org/10.1007/s00521-017-3225-z

[38] Wong, S. C., Gatt, A., Stamatescu V., et al. "Understanding data augmentation for classification: when to warp?." *2016 international conference on digital image computing: techniques and applications (DICTA)*. IEEE, 2016. https://doi.org/10.1109/DICTA.2016.7797091

[39] Mikołajczyk, A., Grochowski, M., "Data augmentation for improving deep learning in image classification problem." *2018 international interdisciplinary PhD workshop (IIPhDW)*. IEEE, 2018. https://doi.org/10.1109/IIPHDW.2018.8388338

[40] Parida, A. K., Kumar, D., Sahoo, R. R., et al. "Medium term Solar Power Prediction using stacked LSTM based deep learning Technique," *2022 IEEE 4th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA), Goa, India, 2022, pp. 19-26, doi: 10.1109/ICCCMLA56841.2022.9989067*. https://doi.org/10.1109/ICCCMLA56841.2022.9989067

List of Figures

List of Tables

List of Figures

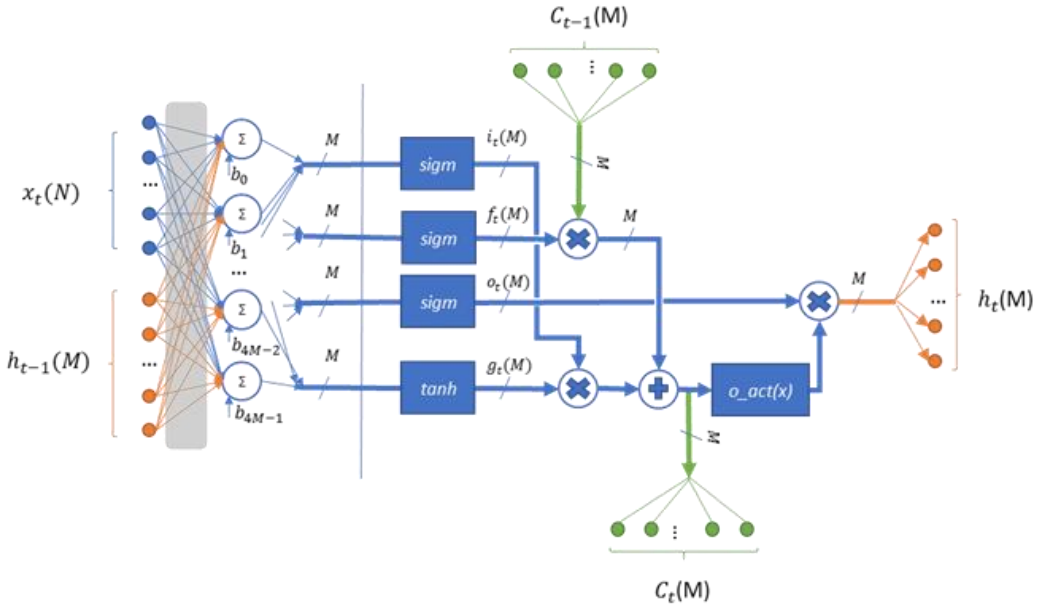


Figure.1. basic structure of LSTM

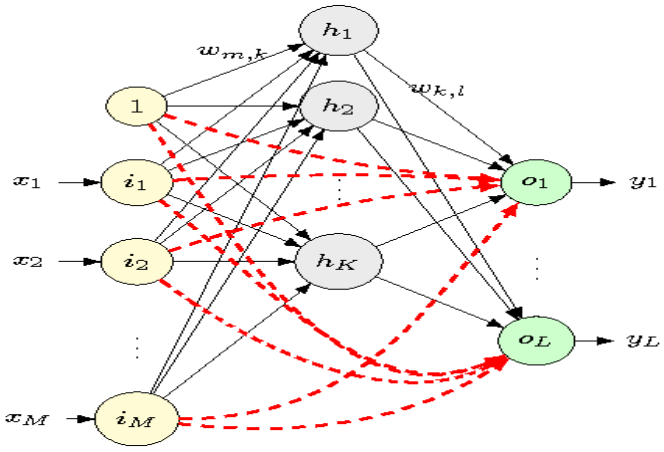


Figure.2. Basic RVFLN Structure

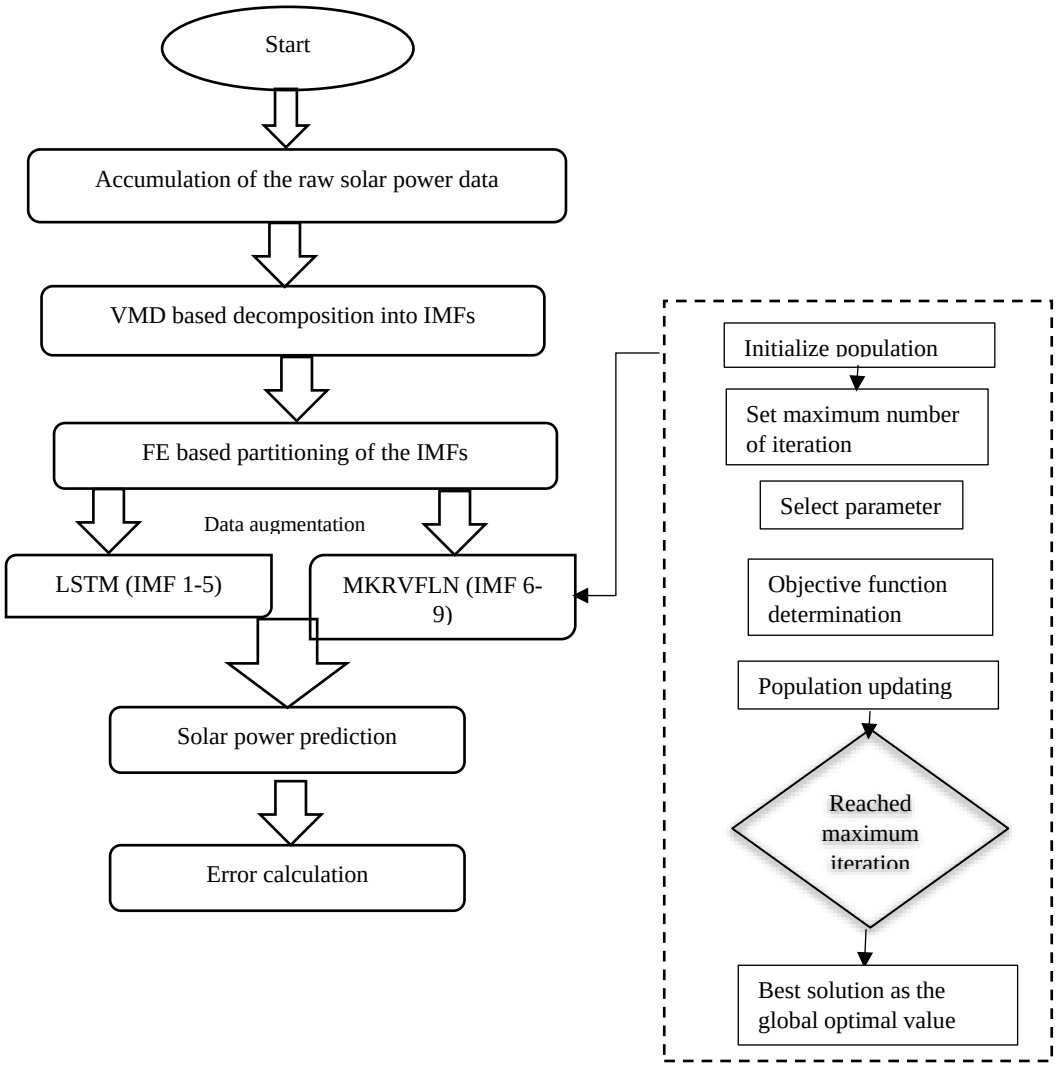
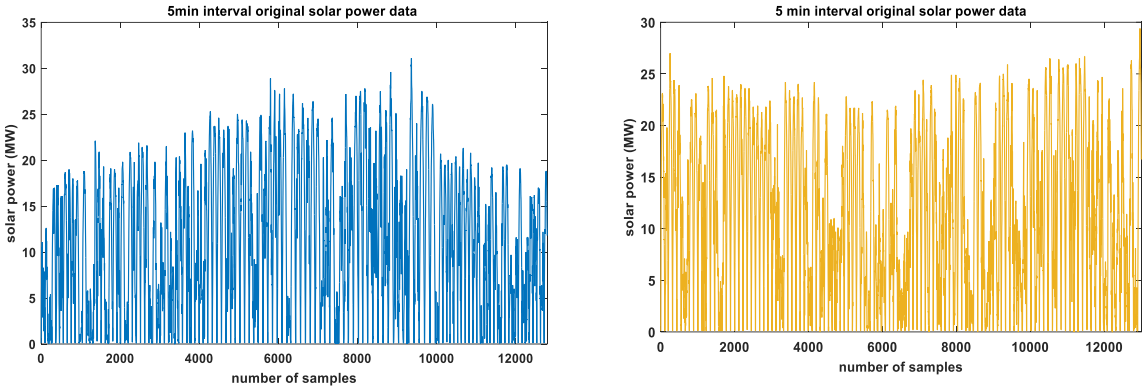
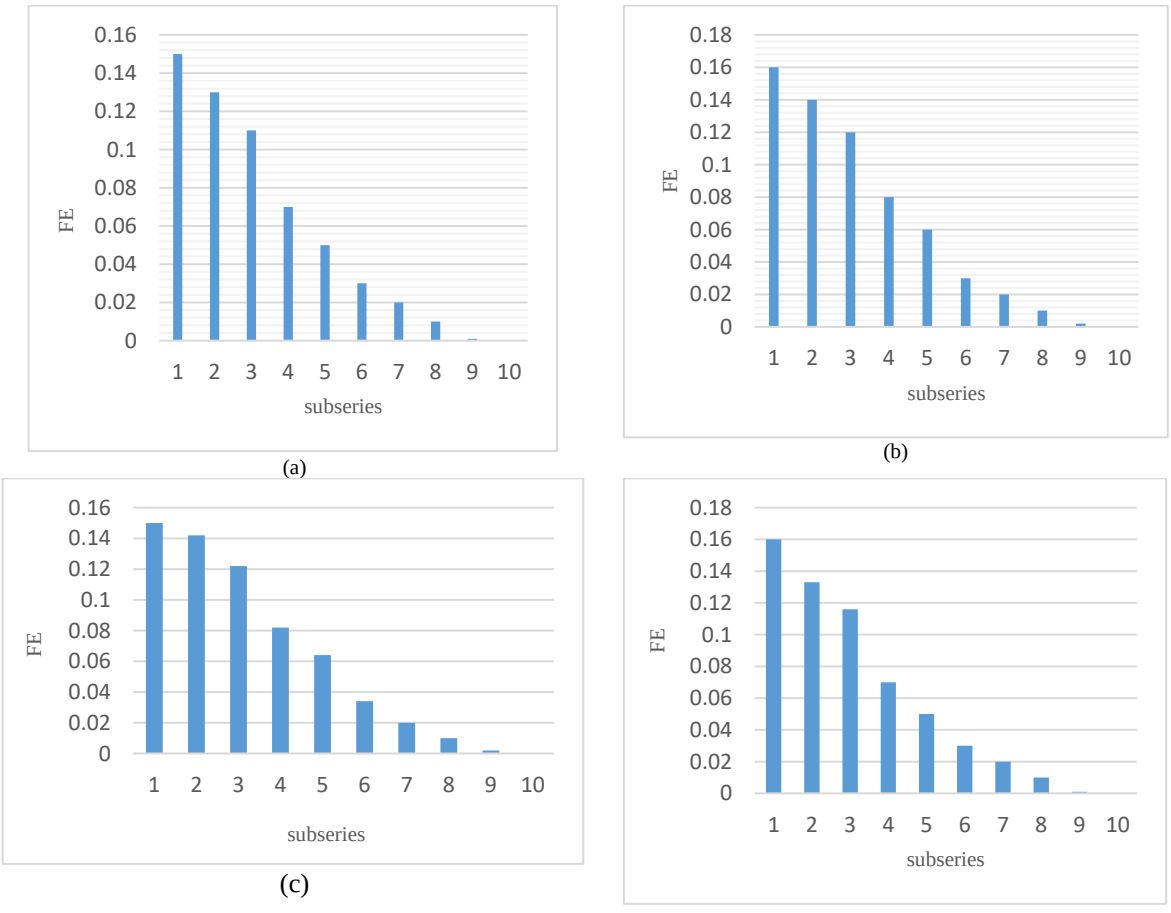


Figure.3. flowchart of the proposed technique



(a) (b)

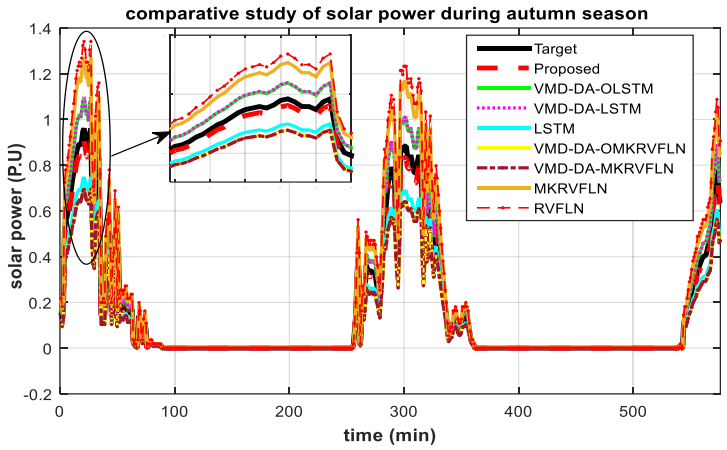
Figure.4. original solar power of two different dataset, (a) original PV power for winter season of Alabama, (b) Original solar power data of winter season of New York



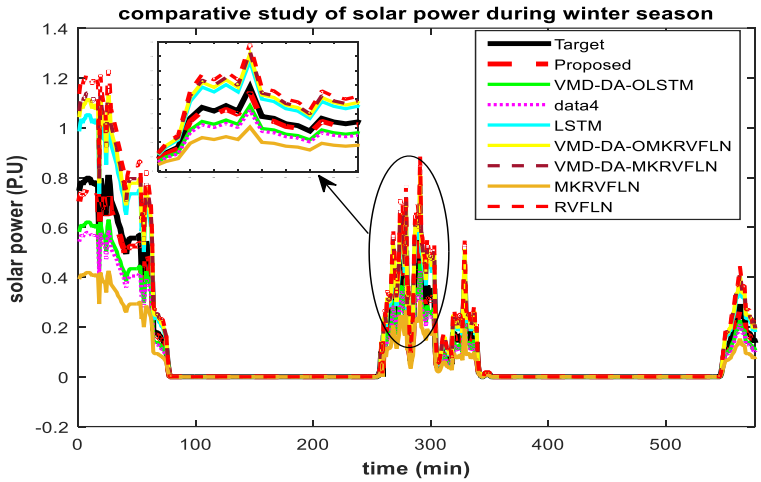
(a) (b)

(c) (d)

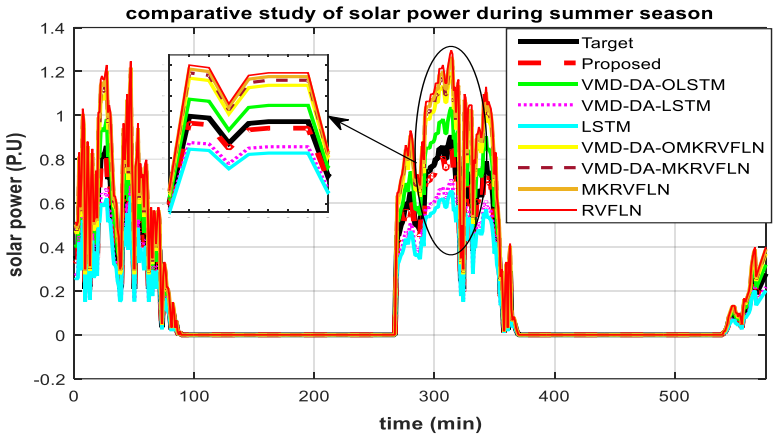
Figure.5. FE based subseries partitioning for Alabama at different weather condition, (a) winter season, (b) spring season, (c) summer season and (d) winter season



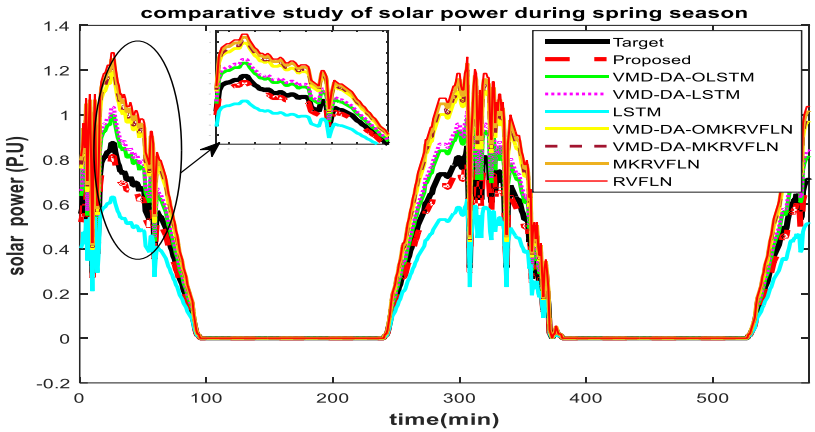
(a) Comparison analysis of different prediction method for solar power prediction of autumn season at Alabama.



(b) Comparison analysis of different prediction method for solar power prediction of winter season at Alabama.



(c) Comparison analysis of different prediction method for solar power prediction of summer season at Alabama.



(d) Comparison analysis of different prediction method for predicting solar power of spring season at Alabama.

Figure.6. Comparison analysis of short term solar power prediction using different techniques using dataset from Alabama at different weather condition, (a) predicting solar power at autumn season, (b) predicting solar power at winter season, (c) predicting solar power at summer season, and (d) predicting solar power at spring season

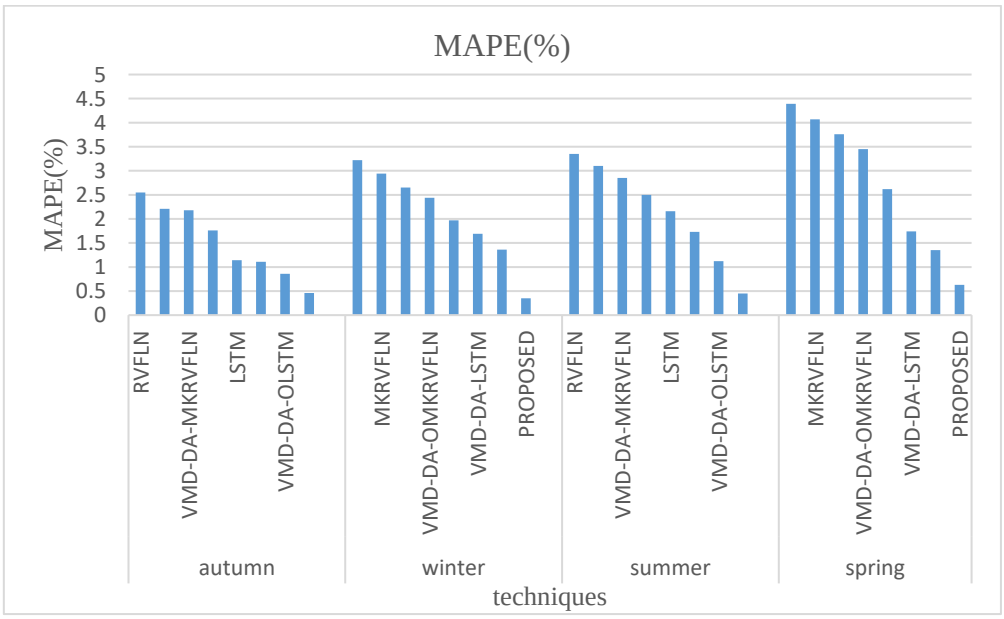


Figure.7 Differentiation between different MAPE values using different techniques for short term solar power prediction

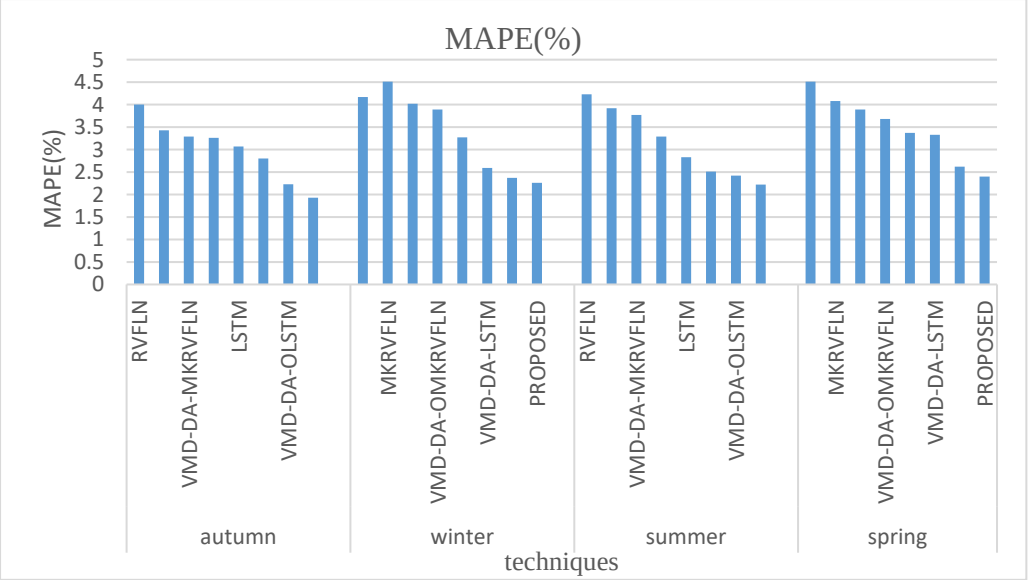


Figure.8 Differentiation between different MAPE values using different techniques for medium term solar power prediction

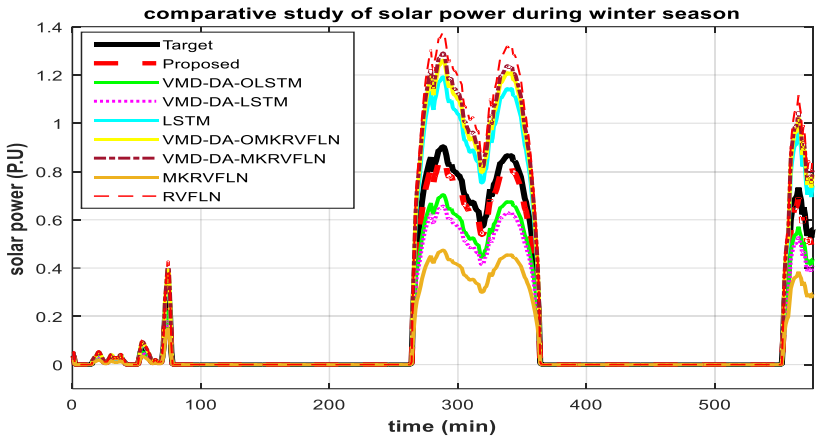


Figure.9. Comparison analysis of different prediction method for predicting solar power of winter season at New York

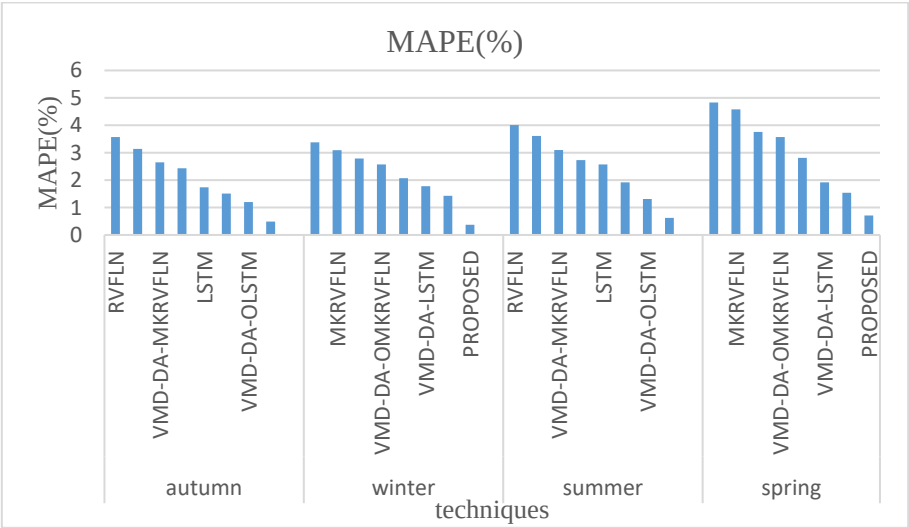


Figure.10 Comparison analysis of MAPE values using different techniques for short term solar power prediction

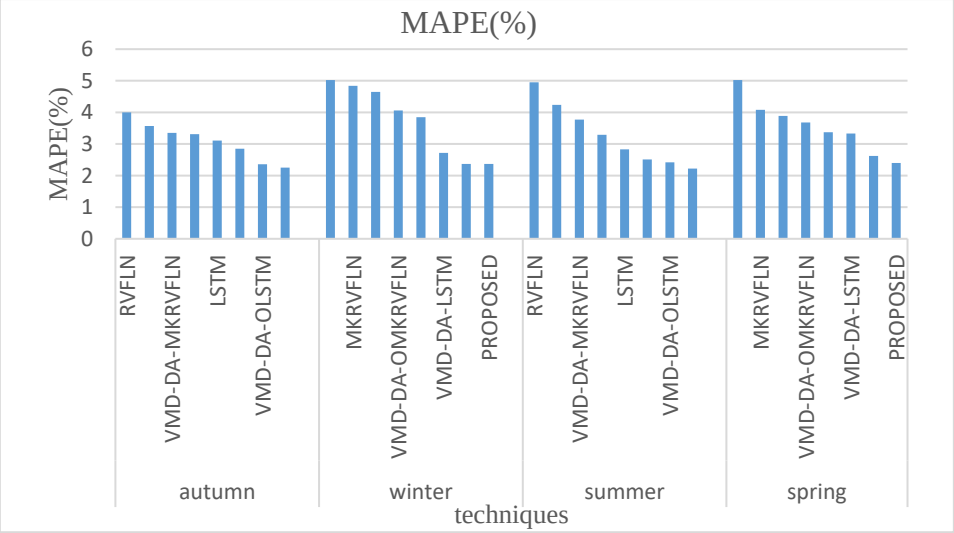


Figure.11. Comparison analysis of MAPE values using different techniques for medium term solar power prediction

Table.1. Geographical and statistical information of the datasets

Dataset	***Data Period***	***Coordinates***	***N***	T_{tr}	T_{ts}
Alabama (spring)	*06/01/06-31/08/06*	*30.35° N, 87.75°W*	*20000*	*14000*	*6000*
New York (winter)	*06/01/06-31/08/06*	*40.85°N, 73.85°W*	*20650*	*14000*	*6000*

Table.2. Optimized kernel parameters used for short term and medium term solar power prediction of two different dataset at different weather condition:

dataset	*Prediction Interval*	*Season*	*Optimized kernel parameter*						
			δ	*a*	*b*	*c*	*d*	μ_1	μ_2
Alabama	*Short term*	*winter*	*0.272*	*0.322*	*1.731*	*2.731*	*0.961*	*0.5*	*0.7*
		autumn	*0.391*	*0.341*	*1.851*	*3.150*	*1.324*	*0.3*	*0.8*
		spring	*0.325*	*0.429*	*2.031*	*3.265*	*0.991*	*0.54*	*0.56*
		summer	*0.450*	*0.525*	*2.284*	*2.240*	*1.126*	*0.45*	*0.76*
	Medium term	*winter*	*0.313*	*0.410*	*1.262*	*3.651*	*1.134*	*0.4*	*0.6*
		autumn	*0.976*	*0.622*	*3.411*	*3.857*	*1.392*	*0.45*	*0.62*
		spring	*0.922*	*0.864*	*2.452*	*3.434*	*1.412*	*0.42*	*0.72*
		summer	*0.881*	*0.672*	*1.712*	*3.954*	*1.174*	*0.30*	*0.73*
New York	*Short term*	*winter*	*0.375*	*0.372*	*1.732*	*4.567*	*1.786*	*0.43*	*0.72*
		autumn	*0.772*	*0.452*	*1.742*	*3.244*	*1.323*	*0.51*	*0.67*
		spring	*0.342*	*0.451*	*2.254*	*3.361*	*0.873*	*0.33*	*0.64*
		summer	*0.452*	*0.211*	*1.212*	*4.132*	*0.958*	*0.38*	*0.63*
	Medium term	*winter*	*0.352*	*0.373*	*1.593*	*4.375*	*0.885*	*0.46*	*0.75*
		autumn	*0.875*	*0.945*	*2.039*	*2.822*	*1.441*	*0.40*	*0.60*
		spring	*0.973*	*0.582*	*3.312*	*3.742*	*1.022*	*0.45*	*0.65*
		summer	*0.687*	*0.525*	*1.585*	*3.546*	*1.282*	*0.46*	*0.66*

Table.3 Comparative analysis of short term solar power prediction of Alabama for different season

Technique	*Season*	*MAPE*	*RMSE*	*MAE*
RVFLN	*autumn*	*2.55*	*0.11*	*0.06*
MKRVFLN		*2.21*	*0.10*	*0.55*
VMD-DA-MKRVFLN		*2.18*	*0.10*	*0.05*
VMD-DA-OMKRVFLN		*1.76*	*0.08*	*0.04*
LSTM		*1.14*	*0.05*	*0.03*
VMD-DA-LSTM		*1.11*	*0.05*	*0.03*
VMD-DA-OLSTM		*0.86*	*0.04*	*0.02*
Proposed		***0.46***	***0.02***	***0.01***
RVFLN	*winter*	*3.22*	*0.14*	*0.07*
MKRVFLN		*2.94*	*0.13*	*0.07*
VMD-DA-MKRVFLN		*2.65*	*0.12*	*0.06*
VMD-DA-OMKRVFLN		*2.44*	*0.11*	*0.05*
LSTM		*1.97*	*0.09*	*0.04*
VMD-DA-LSTM		*1.69*	*0.07*	*0.04*
VMD-DA-OLSTM		*1.36*	*0.06*	*0.03*
Proposed		***0.35***	***0.02***	***0.01***
RVFLN	*summer*	*3.35*		
MKRVFLN		*3.1*	*0.11*	*0.08*
VMD-DA-MKRVFLN		*2.85*	*0.13*	*0.07*
VMD-DA-OMKRVFLN		*2.50*	*0.11*	*0.06*
LSTM		*2.16*	*0.10*	*0.05*
VMD-DA-LSTM		*1.73*	*0.08*	*0.04*
VMD-DA-OLSTM		*1.12*	*0.05*	*0.03*
Proposed		***0.45***	***0.02***	***0.08***
RVFLN	*spring*	*4.39*	*0.16*	*0.10*
MKRVFLN		*4.07*	*0.15*	*0.09*
VMD-DA-MKRVFLN		*3.76*	*0.14*	*0.09*
VMD-DA-OMKRVFLN		*3.45*	*0.13*	*0.08*
LSTM		*2.62*	*0.10*	*0.06*
VMD-DA-LSTM		*1.74*	*0.06*	*0.04*
VMD-DA-OLSTM		*1.35*	*0.05*	*0.03*
Proposed		***0.63***	***0.02***	***0.01***

Table.4 Comparative analysis of medium term solar power prediction of Alabama for different season

Technique	*Season*	*MAPE*	*RMSE*	*MAE*
RVFLN	*autumn*	*4.00*	*1.50*	*1.16*
MKRVFLN		*3.43*	*0.24*	*0.22*
VMD-DA-MKRVFLN		*3.29*	*0.80*	*0.62*
VMD-DA-OMKRVFLN		*3.26*	*0.51*	*0.29*
LSTM		*3.07*	*0.36*	*0.30*
VMD-DA-LSTM		*2.80*	*0.28*	*0.25*
VMD-DA-OLSTM		*2.23*	*0.29*	*0.20*
Proposed		***1.93***	***0.04***	***0.17***
RVFLN	*winter*	*4.17*	*0.81*	*0.61*
MKRVFLN		*4.51*	*0.70*	*0.59*
VMD-DA-MKRVFLN		*4.02*	*0.85*	*0.57*
VMD-DA-OMKRVFLN		*3.89*	*1.20*	*0.93*
LSTM		*3.27*	*0.51*	*0.29*
VMD-DA-LSTM		*2.59*	*0.32*	*0.25*
VMD-DA-OLSTM		*2.37*	*0.29*	*0.23*
Proposed		***2.26***	***0.28***	***0.22***
RVFLN	*summer*	*4.23*	*0.81*	*0.61*
MKRVFLN		*3.92*	*1.50*	*1,16*
VMD-DA-MKRVFLN		*3.77*	*0.75*	*0.56*
VMD-DA-OMKRVFLN		*3.29*	*0.80*	*0.61*
LSTM		*2.83*	*0.25*	*0.24*

VMD-DA-LSTM		*2.51*	*0.26*	*0.24*
VMD-DA-OLSTM		*2.42*	*0.21*	*0.15*
Proposed		***2.22***	***0.29***	***0.21***
RVFLN		*4.51*	*0.70*	*0.59*
MKRVFLN		*4.08*	*0.64*	*0.53*
VMD-DA-MKRVFLN		*3.89*	*0.80*	*0.56*
VMD-DA-OMKRVFLN	*spring*	*3.68*	*0.74*	*0.56*
LSTM		*3.37*	*1.31*	*1.16*
VMD-DA-LSTM		*3.33*	*1.31*	*1.16*
VMD-DA-OLSTM		*2.62*	*0.65*	*0.51*
Proposed		*2.40*	*0.02*	*0.01*

Table.5 Comparative analysis of short term solar power prediction of New York for different season

Technique	Season	MAPE	RMSE	MAE
RVFLN		3.57	0.13	0.08
MKRVFLN		3.14	0.15	0.75
VMD-DA-MKRVFLN		2.65	0.10	0.08
VMD-DA-OMKRVFLN	autumn	2.43	0.10	0.06
LSTM		1.74	0.07	0.05
VMD-DA-LSTM		1.51	0.06	0.04
VMD-DA-OLSTM		1.20	0.05	0.02
Proposed		**0.49**	**0.03**	**0.01**
RVFLN		3.38	0.18	0.08
MKRVFLN		3.09	0.16	0.08
VMD-DA-MKRVFLN		2.79	0.14	0.07
VMD-DA-OMKRVFLN	winter	2.57	0.13	0.06
LSTM		2.07	0.11	0.05
VMD-DA-LSTM		1.78	0.09	0.04
VMD-DA-OLSTM		1.43	0.07	0.03
Proposed		**0.37**	**0.02**	**0.01**
RVFLN		4.01	0.14	0.09
MKRVFLN		3.61	0.12	0.09
VMD-DA-MKRVFLN		3.10	0.14	0.07
VMD-DA-OMKRVFLN	summer	2.73	0.12	0.07
LSTM		2.57	0.16	0.06
VMD-DA-LSTM		1.92	0.08	0.05
VMD-DA-OLSTM		1.31	0.06	0.04
Proposed		**0.62**	**0.03**	**0.09**
RVFLN		4.83	0.19	0.12
MKRVFLN		4.58	0.17	0.11
VMD-DA-MKRVFLN		3.76	0.14	0.09
VMD-DA-OMKRVFLN	spring	3.57	0.14	0.09
LSTM		2.81	0.10	0.06
VMD-DA-LSTM		1.92	0.06	0.05
VMD-DA-OLSTM		1.54	0.05	0.03
Proposed		**0.71**	**0.03**	**0.02**

Table.6 Comparative analysis of medium term solar power prediction of New York for different season

Technique	*Season*	*MAPE*	*RMSE*	*MAE*
RVFLN		*4.00*	*1.52*	*1.18*
MKRVFLN		*3.57*	*0.26*	*0.24*
VMD-DA-MKRVFLN		*3.35*	*0.80*	*0.62*
VMD-DA-OMKRVFLN	*autumn*	*3.31*	*0.53*	*0.31*
LSTM		*3.11*	*0.37*	*0.31*

VMD-DA-LSTM		*2.85*	*0.28*	*0.26*
VMD-DA-OLSTM		*2.36*	*0.21*	*0.22*
Proposed		***2.25***	***0.04***	***0.18***
RVFLN	*winter*	*5.02*	*0.82*	*0.64*
MKRVFLN		*4.84*	*0.72*	*0.62*
VMD-DA-MKRVFLN		*4.65*	*0.87*	*0.59*
VMD-DA-OMKRVFLN		*4.06*	*1.21*	*0.95*
LSTM		*3.85*	*0.52*	*0.31*
VMD-DA-LSTM		*2.72*	*0.34*	*0.27*
VMD-DA-OLSTM		*2.37*	*0.31*	*0.24*
Proposed		***2.37***	***0.29***	***0.23***
RVFLN	*summer*	*4.95*	*0.82*	*0.63*
MKRVFLN		*4.24*	*1.50*	*1,16*
VMD-DA-MKRVFLN		*3.77*	*0.75*	*0.56*
VMD-DA-OMKRVFLN		*3.29*	*0.80*	*0.61*
LSTM		*2.83*	*0.25*	*0.24*
VMD-DA-LSTM		*2.51*	*0.26*	*0.24*
VMD-DA-OLSTM		*2.42*	*0.21*	*0.15*
Proposed		***2.22***	***0.29***	***0.21***
RVFLN	*spring*	*5.02*	*0.72*	*0.61*
MKRVFLN		*4.08*	*0.64*	*0.53*
VMD-DA-MKRVFLN		*3.89*	*0.80*	*0.56*
VMD-DA-OMKRVFLN		*3.68*	*0.74*	*0.56*
LSTM		*3.37*	*1.31*	*1.16*
VMD-DA-LSTM		*3.33*	*1.31*	*1.16*
VMD-DA-OLSTM		*2.62*	*0.65*	*0.51*
Proposed		*2.40*	*0.02*	*0.01*

Dr. Ajaya Kumar Parida is currently working as an Associate Professor in the School of Computer Engineering, Kalinga Institute of Industrial Technology (KIIT), Deemed to be University, Bhubaneswar, India. He has received his B.E degree from KIIT University, India and MBA degree in Systems and Marketing from KIIT University. He did his MTECH from CET, Bhubaneswar, India and completed his PhD in Computer Science and Engineering from Utkal University, Bhubaneswar, India. He has published more than 20 research papers in various international journals and conferences. He is also associated with various educational and research societies. He is the state level jury member of National Children Science Congress. His research interest includes in the field of Data Mining, Soft Computing and Machine Learning. He has also more than 15 years of teaching and research experience in several engineering colleges and universities.

Dr. Deepak Kumar is currently working as Assistant Professor in NIT Meghalaya. He completed his PhD in Computer Science and Engineering from NIT Meghalaya.His Research Interest includes Computational Mathematics, Digital Signal Processing and Machine Learning.He has published more than 20 research papers in various international journals and conferences.

Dr. Rashmi Ranjan Sahoo is currently working as Assistant Professor in Parala Maharaja Engineering College, Berhampur.He completed his PhD in Computer Science and Engineering from Jadavpur University, Kolkata and received MTech degree from same Jadavpur University, Kolkata.His Research Interest includes Security in Ad-hoc Network, Machine Learing, Deep Learning.He has published more than 15 research papers in various international journals and conferences.He has also more than 10 years of teaching and research experience in several engineering colleges.

Dr. Bunil Kumar Balabantaray is currently working as Assistant Professor in NIT Meghalaya. He completed his PhD in Computer Science and Engineering from NIT Rourkela.His Research Interest includes Robotics, Image Processing, Biomedical Image Analysis and Cyber Security.He has published more than 25 research papers in various international journals and conferences.He has also more than 18 years of teaching and research experience in several engineering colleges and universities.