Soft Computing
https://doi.org/10.1007/s00500-020-04680-7

METHODOLOGIES AND APPLICATION

Hourly day-ahead wind power forecasting with the EEMD-CSO-LSTM-EFG deep learning technique

A. Shobana Devi[1] · G. Maragatham[1] · K. Boopathi[2] · A. G. Rangaraj[2]

Abstract
Wind power forecasting has gained significant attention due to advances in wind energy generation in power frameworks and the uncertain nature of wind. In this manner, to maintain an affordable, reliable, economical, and dependable power supply, accurately predicting wind power is important. In recent years, several investigations and studies have been conducted in this field. Unfortunately, these examinations disregarded the significance of data preprocessing and the impact of various missing values, thereby resulting in poor performance in forecasting. However, long short-term memory (LSTM) network, a kind of recurrent neural network (RNN), can predict and process the time-series data at moderately long intervals and time delays, thereby producing good forecasting results using time-series data. This article recommends a hybrid forecasting model for forecasting wind power to improve the performance of the prediction. An improved long short-term memory network-enhanced forget-gate network (LSTM-EFG) model, whose appropriate parameters are optimized using cuckoo search optimization algorithm (CSO), is used to forecast the subseries data that is extracted using ensemble empirical mode decomposition (EEMD). The experimental results show that the proposed forecasting model overcomes the limitations of traditional forecasting models and efficiently improves forecasting accuracy. Furthermore, it serves as an operational tool for wind power plants management.

Keywords Wind power forecasting (WPF) · Deep learning · Long short-term memory network (LSTM) · Ensemble empirical mode decomposition (EEMD) · Cuckoo search algorithm (CSO) · Forecasting accuracy

1 Introduction

1.1 Background

In recent years, the demand for renewable energy resources has increased considerably due to air pollution caused by conventional petroleum (or fossil) derivatives. Presently, increasing energy demand, secure energy supply, and reduction in pollutions from non-renewable energy sources are the most pressing needs that should be urgently addressed and made more robust (Renewables 2015). Therefore, greenhouse gas emissions and environmental pollution can be effectively reduced by the improvement of renewable energy. Figure 1 shows the total wind power capacity and the additional capacity in 2017 for top 10 countries in the world. Figure 1b shows the yearly world total wind power capacity of 539 GW based on the previous year's capacity and annual additions of installed wind energy between 2007 and 2017 (Zahedi 2010).

Energy utilization, which produces 60% of worldwide greenhouse gas emissions, impacts environmental climate change (United Nations 2012). Meanwhile, wind energy has great significance in the energy structures of many countries. However, several issues associated with the generation of wind power have limited its advancement and have posed serious challenges to the stable and safe

Communicated by V. Loia.

✉ A. Shobana Devi
shobanak07@gmail.com

G. Maragatham
maragatg@srmist.edu.in

K. Boopathi
boopathi.niwe@nic.in

A. G. Rangaraj
rangaraj.niwe@nic.in

[1] Department of Information Technology, SRMIST, Chennai, India

[2] National Institute of Wind ENERGY, Chennai, India

Published online: 27 January 2020

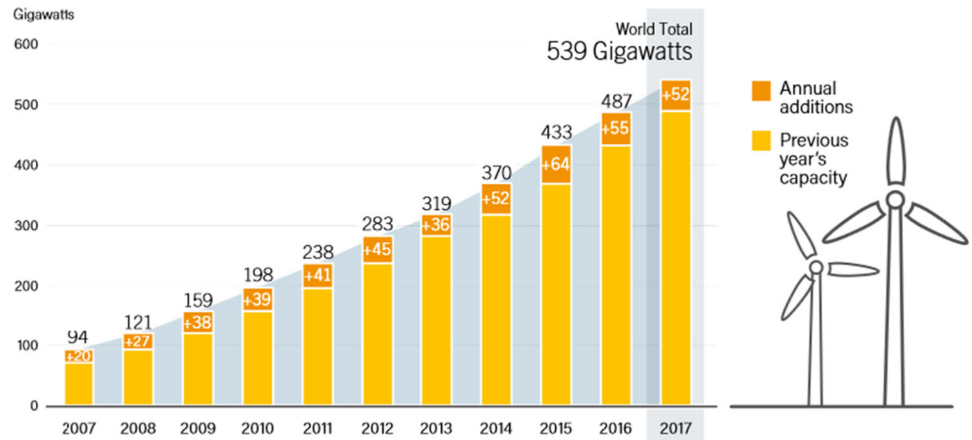


Fig. 1 Total amount of global capacity and annual additions of wind power from 2007 to 2017

operation of the power system, thereby causing its instability and irregularity (Song et al. 2018). It has enhanced the long-term vitality plans of few countries (Hossain 2014). Based on the most recent report given by Global Wind Energy Council (GWEC) (Global Wind Energy Council 2018), 52 GW of excess clean and emission-free wind power was included in 2017, bringing global establishments to 539 GW. Hence, the accurate forecasting of wind power can mitigate the peak load pressure and regulation of frequency in the power system grid, which is helpful in the incorporation of generated wind power into the main power grid. Meanwhile, it handles issues such as UC (unit commitment), ED (economic dispatch) (Botterud et al. 2013), dynamically adjusted backup (Menemenlis et al. 2012), and optimization of storage of energy (Bludszuweit and Dominguez-Navarro 2011).

In the meantime, wind energy strongly affects the daily and hourly costs of energy advertisement, as the accurate forecasting of wind and the variation of wind energy impact the reimbursement costs of energy resources for the energy that will be used future.

1.2 Motivation

The principal sources of renewable power are wind, sunlight-based (solar), biomass, geothermal, and water. However, sustainable energy sources have several advantages as they curb discharges, preserve the characteristics of dirt, diminish the pollution level of soil and water pollution, provide employment, and produce yield financial gains for adjacent networks. This is the reason why a growing concern exists among governments and researchers in the use of sustainable power sources in recent decades, thereby bringing about growing interests in renewable plants. These sustainable sources enhance stable production and efficient utilization of energy, preserve conventional assets, reduce the importation of sustainable energy, and enhance the economy in various scales (adjacent, nearby, zonal, and worldwide). However, other factors should be considered in the allocation of renewable energy utilization, for example the gradual increase in the worldwide demand for energy due to financial developments and worldwide growth in the human population (Chang 2014).

Wind energy has attracted great interest in past decades due to its critical attributes, its broad accessibility, and extensive appropriation in the variety by means of additional sustainable power source assets. As indicated by GWEC, the wind control at the global scale was more than 50 GW in 2017 while India, Europe, and the seaward wind control division were consuming enrolled recorded years (Jung and Broadwater 2014).

The Ministry of New and Renewable Energy has asked state governments to consider feed-in taxes for Micro, Small and Medium Enterprises utilizing below 25 MW at taxes to be determined by the states. Moreover, first appraisals of offshore wind potential in India and pre-plausibility contemplate for two key coastal states (Tamil Nadu and Gujarat) have been completed in the GWEC-driven Facilitating Offshore Wind in India scheme. Hence, offshore wind power could lead to the creation of jobs in India due to the extensive wind assets accessible close to the demand for high-energy request.

In the meantime, wind energy strongly affects the daily and hourly costs of energy advertisement, as the accurate forecasting of wind and the variation of wind energy impact the reimbursement costs of energy resources for the energy that will be used future.

1.3 Challenges to be addressed

Challenges of existing wind power forecasting system are:

(1) Improving the accuracy of existing forecasting models.

(2) Forecasting models based on deep learning techniques to be designed considering the seasonal variations.
(3) Improving the intra-day (short-term) wind power forecasting (WPF) models must be designed.
(4) Required wind speed post-processing methods.
(5) Interpolation techniques to handle wind data with varying time horizons without any loss of information need to be addressed.
(6) Development of wind power forecasting system using Big Data frameworks may be designed for scalability.

1.4 Contributions of the proposed wind power forecasting model

To directly resolve the issues referenced above, a novel hybrid model for added accurate wind power forecasting and improved assessment based on CSO algorithm, EEMD, and LSTM structure is presented in this paper.

The guideline developments and commitments of this paper in comparison with different works in wind power forecasting are summarized as follows:

- The proposed methodology in this paper utilizes data preprocessing and optimization of algorithm hyperparameters to upgrade the execution of the LSTM network structure. In this paper, the wind crude time-series data are decomposed into various subseries using EEMD, such that the subseries having high frequency are ignored while the remaining subseries components are rebuilt to obtain stationary time-series data. Thus, the natural attributes of the wind information can be effectively obtained and examined to improve the precision of wind power forecasting.
- The paper utilizes cuckoo search (CS) algorithm for weights optimization before training the LSTM network model. The CS algorithm is highly effective in terms of global optimization. It requires some parameters for optimization and can solve multiobjective problems, thereby enhancing forecasting accuracy. The LSTM model can overcome the challenges of conventional models such as overfitting, more time consumed in training, tendency to fall into local optima, low convergence rate, and uncertainty. These challenges brought about by a danger of being caught in a neighborhood minimum and chance of exceeding the surface error base value.
- To investigate the forecasting capacity of the proposed methodology, traditional models such as BPNN, ARIMA, and LSTM are utilized for comparisons. Meanwhile, performance assessment metrics such as mean absolute error (MAE), mean absolute percentage error (MAPE), root-mean-square error (RMSE), and mean absolute scale error (MASE) are used to analyze and evaluate the execution of the proposed hybrid forecasting model.

In the following section, we have given a literature survey to contextualize the present condition of this paper and distinguished the limitations and challenges that can be addressed by our proposed forecasting model.

The remaining part of this paper is organized as follows: Sect. 2 discusses related works, deep learning and machine learning backgrounds, and analysis of time series, using farm wind data and different models used for wind power forecasting. In Sect. 3, we explain the procedure of methods such as EEMD used to decompose the wind time-series data and also project an improved LSTM model framework and the CSO algorithm to optimize the weights of LSTM network. In Sect. 4, we evaluate with the proposed hybrid model EEMD-LSTM-EFG-CSO with different farm wind data. In Sect. 5, we discuss the analysis result of the proposed forecasting model and compare its performance with the experimental results of existing conventional wind power forecasting models. In Sect. 6, we discuss the conclusion and highlight the directions for future research.

2 Related works

2.1 Wind power forecasting

Several methods for forecasting wind power and speed have been proposed in various research works. However, they are generally divided into two primary groups: the forecasting horizon and the used model. For the forecasting horizon, Change (Jung and Broadwater 2014) proposed the following classification based on a literature survey:

- Ultra-short-term forecasting: several minutes to 1 h ahead.
- Short-term forecasting: 1 h to several hours ahead.
- Medium-term forecasting: several hours to 1 week ahead.
- Long-term forecasting: 1 week to 1 year or more ahead.

Recently, several studies on wind power forecasting have been published. However, most of the studies concentrated on the prediction of wind speed while a few researches focused on forecasting the generation of wind power. The methodologies of these investigations could be organized into three classes: artificial intelligence models, time-series models, and time-series-based artificial intelligence models. The majority of these methodologies use time-series models such as vector autoregressive (VAR), autoregressive moving average (ARMA), and autoregressive integrated moving average (ARIMA) models.

Moreover, the classifications of wind power forecasting using machine learning techniques can be recognized with respect to the utilized approach as follows.

Foley et al. (2012) have used recurrent-type neural system to forecast the generation of wind power for a turbine from 1 to 288 h ahead. In Olaofe and Folly (2012), several feedforward neural networks (FFN) are utilized to make forecasts of multiple steps ahead ranging from 30 m to 6 h at 30-m interval of time-series data. In Cadenas-Barrera et al. (2013), two-layer feedforward neural network is used for wind power forecasting.

Kishore et al. (2014) have considered four optimization algorithms for the training of NARX model. The model performance was assessed using wind speed as the input variable and relative humidity, temperature, or direction as exogenous variables. De Aquino et al. (2015) have examined the performance execution of three neural network models such as fuzzy inference system, FFN, and echo state network used for short-term wind speed forecasting based on solar radiation, humidity, temperature, and wind speed time data series.

Sun et al. (2015) have used fast ensemble empirical mode decomposition (FEEMD), least squares support vector machine (LSSVM), optimized using the Bat algorithm, and were used to model the decomposed functions. In Wu and Peng (2016), introduced the EEMD method for time-series decomposition and utilized principal component analysis (PCA) to optimize the number of training attributes of LSSVM. Ghorbani et al. (2016) have formulated genetic expression programming, which is an algorithm based on neural network model.

Bonanno et al. (2015) have applied a wavelet-based RNN (recurrent neural network). Wavelet decomposition method is used to transform the first wind speed time-series data. The test results affirm the effectiveness of the proposed model, using one-year wind data. Chang et al. (2016) have proposed an ARIMA model to transform the non-stationary data series of temperature and wind speed into stationary series. For wind power generation and speed forecast, radial basis neural network was trained using the transformed data time series. Brusca et al. (2017) have recommended a spiking neural network (SNN)-based technique for wind energy forecasting.

De Alencar (2017) has recommended a hybrid forecast model based on ARIMA and two-layered FNN model. Hui Liu et al. (2018) have used LSTM network model to forecast wind speed and proposed a new multistep prediction method for wind speed by combining singular spectrum analysis (SSA), variational mode decomposition (VMD), extreme learning machine, and LSTM network. Bengio et al. (1994) have proposed an ensemble LSTM based on support vector regression machine, LSTM, and extreme optimization algorithm, to produce relevant forecasts with the ensemble learning.

In Chen et al. (2018), a novel LSTM architecture has developed for wind power forecasting that uses echo state network as hidden layers of the network. Lpez et al. (2018) have used PCA to reduce the dimension of the input variables of LSTM model for NWP data forecasting. The proposed model was compared with SVM and back-propagation neural network (BPNN). The LSTM had higher predictive accuracy and hence wider application. Cheng et al. (2017) have introduced a model (power LSTM) for power demand forecasting to compute important features values, using LSTM neural network. The model is minimized by obtaining the feature weights of high importance. Qu et al. (2016) have developed a forecasting model of wind power dependent on LSTM. To reduce the complexity and scale of data, distance analysis technique is used to monitor variables that have higher correlations with respect to wind power. In addition, Qiaomu et al. (2017) have developed a forecasting model of wind power dependent on LSTM. To reduce the complexity and scale of data, distance analysis technique is used to monitor variables that have higher correlations with respect to wind power. Currently, the models utilized in all existing works depend on LSTM standard network structure.

3 Introduction of the proposed model

3.1 Problems of existing wind power forecasting models

The models referenced above have definite downsides which are précised as follows: (1) A large amount of wind data are usually required to build models with accurate forecasting, due to the instability and irregularity of the raw wind speed data. The error rate of the forecasting results becomes generally higher if the original data are suddenly changed because of environmental variables (Qiaomu et al. 2017); (2) most of the models attempt to fit the model more closely to the raw data. Thus, it is difficult to fit the traditional statistical or physical techniques to the individual models. Moreover, processing wind speed data with high instability, irregularity, and noise results in poor forecasting accuracy and low productivity; (3) some individual forecasting models usually disregard raw data preprocessing and the need for optimization of model parameters. Hence, the models do not yield forecasting results of sufficient accuracy; and (4) to enhance forecasting, some models use artificial intelligence technique. However, low convergence rate and over fitting exist in these models (Gers and Schmidhuber 2000).

The model was tested with wind speed, wind direction, and power generated data from five wind farm substations around Coimbatore zone. The experimental results of one substation show that the wind power forecasting accuracy is improved by the proposed hybrid model and could be effectively connected to the wind substation grids to make a statistical provision for power station management and creation of operational plans in that zone. The flow diagram of the proposed EEMD-CSO-LSTM-EFG model is shown in Fig. 2.

3.2 Empirical mode decomposition (EMD)

EMD is a powerful information preprocessing technique that reduces the noise in non-stationary time data series. It was recommended by Huang et al. (Wu and Huang 2004) in 1998. It is a self-versatile and productive technique for investigating non-stationary and nonlinear signals. It is also efficient for analyzing the series data of wind speed. Moreover, it effectively smoothens the fluctuating data series and breaks them down into patterns or fluctuations of various scales. After decomposition of the original series, a limited and modest number of intrinsic mode functions (IMFs) and a residual component are obtained.

Definition 1 The unique time-series information used for decomposition which comprises of n IMF components and a single residual segment is obtained as:

$$s(t) = \sum_{i=0}^{1} c_i(t) + r_{n(t)} \tag{1}$$

where n denotes the quantity of IMFs, $r_{n(t)}$ is the residuals representing to a pattern in $s(t)$ $(t = 1, 2, \ldots, l)$, and $c_i(t)$ $(i = 1, 2, \ldots, n)$ speaks to the IMFs. Every IMF component is autonomous when defining the nearby characters of original time-series information.

3.3 Ensemble empirical mode decomposition (EEMD)

The major disadvantage of EMD is the mixing of modes, which implies that a single IMF comprises of signals with significantly unique scales or various IMF components comprise of similar scale signals. The mixing of modes constrains the viability of the EMD algorithm by compromising of the stationary IMFs. To resolve this issue, a

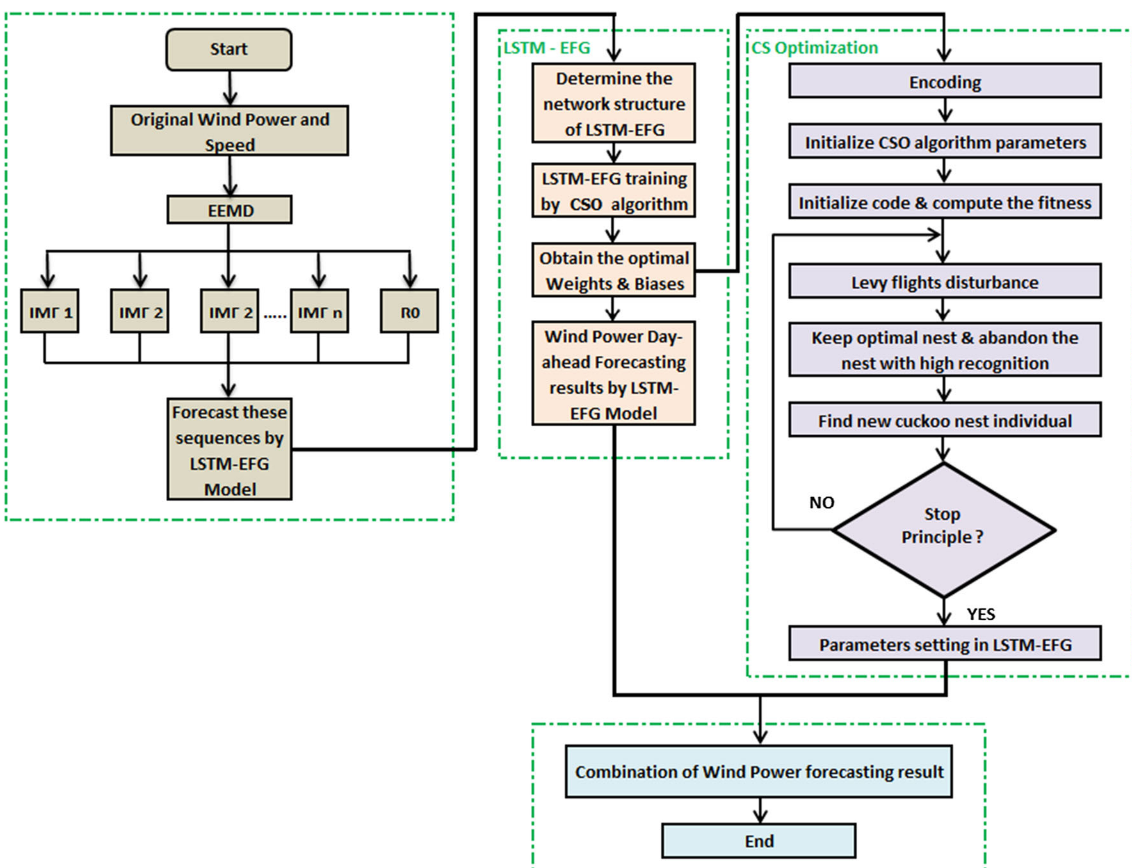


Fig. 2 Flow diagram of proposed EEMD-CSO-LSTM-EFG forecasting model

novel EEMD model, which is a noise-based data decomposition method, is recommended. EEMD was established by Hung et al. (Wu and Huang 2004) that can effectively overcome the inadequacies of EMD by eliminating mode mixing.

Definition 2 The connection between the added noise level, error tolerance, and ensemble number is defined based on the work of Wu and Huang (2004):

$$N_{\varepsilon} = \frac{\varepsilon^2}{\varepsilon_n^2} \tag{2}$$

Moreover, it is an effective signal analysis method that decomposes nonlinear and non-stationary signals using an iterative sifting process (Xiao et al. 2015; Huang et al. 1999) where ε indicates the sufficiency of the additional noise, ε_n is the standard deviation of error, and N_{ε} is the estimation of group individuals. Generally, it is recommended that ε should be fixed at 0.2 to achieve an increasingly precise result (Jiang and Huang 2017). In this paper, N_{ε} was fixed at 100 and the ideal standard deviation of background noise was established within the range of 0.1–0.2.

In EEMD, a white Gaussian noise is introduced into the raw time-series data while the EMD is associated on noise included data of time series to obtain IMFs without mode mixing (Wu and Huang 2009; Hu et al. 2013). The subsequent IMF series incorporating white background noise are then evacuated by obtaining the mean of numerous trials. Thus, the genuine IMF components are characterized as the average of a collection of trials and every trial comprises of the disintegration results of the data signal and white Gaussian noise of limited amplitude (Wu and Huang 2004). The rule hypothesis of EEMD utilizes the added features of the white noise to solve mode mixing by combining the raw time arrangement data with noise and correct signals.

To remove the correct signals from the raw time data series, the original information is combined with white noise. The procedure of EEMD is defined as follows:

- Step 1: The Gaussian white noise series is added to the original empirical data series of wind $v(t)$ to obtain a new data series $V_j(t)$, given by Eq. (3):

  $$V_j(t) = v(t) + n(t), \tag{3}$$

  where $n(t)$ indicates the position of the Gaussian signal.
- Step 2: In view of EMD algorithm, break down the new data series $V_j(t)$ having the added white background noise into n IMFs of different frequencies and a residual Res. The equation of $V_j(t)$ can then be defined as Eq. (4):

  $$V_j(t) = \sum_{i=1}^{n} IMF_i + \mathrm{Re}s, \tag{4}$$

  where n represents the disintegrated amount of IMFs.
- Step 3: Steps 1 and 2 are repeated by adding distinctive Gaussian white repetitive noise at diverse scales at each sifting procedure.
- Step 4: The mean values of entire IMF segments are computed, and the average value of residual components and Gaussian white noise is eliminated from the data series, to obtain the final IMF.

Using the steps above, the whole arrangement of repetitive noise is incorporated into the real wind series information to deliver an integrated position series, thereby lowering the noise. As a result, the genuine IMFs could be obtained from the first data series.

3.4 Cuckoo search optimization algorithm (CSO)

The CS algorithm, developed by Yang and Deb (Wu and Huang 2009; Yang and Deb 2009) in 2009, is a stochastic global search algorithm based on a population that mimics bird behavior. It has been used in a wide range of optimizations such as NP-hard issues, engineering structures, and training of the neural network. The algorithm depends on the brooding behavior of certain cuckoo species in relation to the Lévy flight behavior of certain birds and natural fruit flies (Yang and Deb 2009).

Cuckoos have a dynamic and exceptional brooding system where the females lay their fertilized eggs in the nests of some other birds. In the event that the host bird discovers that she did not lay the egg, she will either dispose it or abandon her nest home and build another.

Generally, the cuckoo eggs are incubated earlier than the host eggs. When the cuckoo chick is brought forth, it disposes the host eggs by pushing them out of the nest, which fabricates the chicks of cuckoo a lot of food given by its host bird (Valian et al. 2011). This decreases the probability of their eggs being destroyed, thereby increasing their population.

3.4.1 Lévy flights

CS algorithm was improved by Lévy flights in contrast to basic isotropic random walk techniques. Yang and Deb (2009) stated that CS algorithm depends on three assumptions:

1. Each cuckoo produces one egg and lays it in an arbitrary nest.
2. The nests that have the highest quality eggs (arrangements) proceed to the next stages.

3. The amount of available host nest is static, and a host which has likelihood pa ε (0,1) of finding an outcast egg. In this situation, the host bird either disposes the egg or abandons its nest to build another nest in another area.

The third assumption can be approximated as a division: pa of the n nests replaced with new nests (through new arbitrary arrangements at different nests). In this algorithm, the fitness function of a solution may be the estimation of the target function or may be characterized by the fitness function that exists in generic algorithms. Each egg in the nest solves a problem while a cuckoo egg indicates another arrangement. The aim is to replace a worse arrangement in the nests by better arrangements (cuckoos).

Based on the assumptions above, the basic steps of CS can be summarized in pseudo-code: Eq. (5) defines "Lévy" flight behavior while producing new arrangements $X_i(t+1)$ for the i_{th} cuckoo [60]:

$$X_i(t+1) = X_i(t) + \propto \oplus \text{Lévy}(\lambda) \tag{5}$$

where $\propto > 0$ is the progression measure which indicates the sizes of the issue of interests and $\oplus$ indicates multiplications of entry wise. The arbitrary walk via "Lévy" flight is effective in investigating the pursuit space as its progression length is longer over the long run.

Equation (6) defines the "Lévy" flight behavior in which the progression sizes are drawn from an "Lévy" likelihood distribution:

$$\text{Lévy} \oplus u = t - \lambda, \quad (1 < \lambda \leq 3) \tag{6}$$

The "Lévy" phase estimate is likelihood dispersion with a boundless fluctuation and flight steps structure an arbitrary walk procedure which conforms to a power-law step-length passage through a substantial tail. This is vastly different from an endless mean. A segment of the new arrangements should be created by "Lévy" walk everywhere the best solution is obtained up until that point; thereby, this will accelerate the near pursuit. In any case, a significant division of the new arrangements should be made using a wide edge field randomization whose territories a long way from the present best arrangement. This will ensure that the structure will not fall into neighborhood minima. Table 1 gives the pseudo-code for CS algorithm.

3.4.2 Proposed CSLSTM algorithm

The CS optimization algorithm begins with an initial random population, similar to other heuristic optimizations. In the proposed CSLSTM network algorithm, every best nest represents a feasible arrangement, i.e., the weights and its equivalent biases for optimization of LSTM network. It also improves the quality of the solution. When CS algorithm was used in the first cycle, the best biases and weights were initialized and the weights were distributed to the LSTM network. The weights of the LSTM network were then computed in each layer using activation function. In the following cycle, the weights are updated using the best possible solution given by the CS algorithm. The CS will continue to go through the last cycle of the LSTM network until either the best weights or a low RMSE is obtained.

In the proposed CSLSTM network, the weight estimation of a matrix is determined using Eqs. (7) and (8). Additionally, the weight matrix update is performed using Eq. (9). The sum of square errors (SSE) is easily calculated in the neural network process for each weight matrix in Wc. Meanwhile, the LSTM network structure has three layers, namely one input layer, one hidden layer and one output layer. In the CSLSTM network, the input vector x is circulated through a weight layer W, using Eq. (10). However, in a straightforward intermittent neural network, the vector input is comparably proliferated over a weight layer and joined to an extra repetitive weight layer U via the past activation state, given by Eq. (11).

$$W_n = U_n = \sum_{n=1}^{N} a \cdot \left(rand - \frac{1}{2}\right), \tag{7}$$

$$B_n = \sum_{n=1}^{N} a \cdot \left(rand - \frac{1}{2}\right), \tag{8}$$

where W_n is the N_{th} weight value of the matrix weights. In Eq. (7), rand is a random value in the range [0, 1], a is a constant parameter for the proposed method whose value is less than 1, and B_n gives the bias value.

The list of weights matrix is specified as follows:

$$W^c = \left[W_n^1, W_n^1, W_n^1, \ldots, W_n^{N-1}\right] \tag{9}$$

$$net_j(t) = \sum_{i}^{n} x_i(t) w_{n(ji)} + B_{n(j)}, \tag{10}$$

where n is the quantity of inputs and $B_{n(j)}$ is a bias.

$$net_j(t) = \sum_{i}^{n} x_i(t) W_{n(ji)} + \sum_{i}^{m} y_l(t-1) U_{n(jl)} + B_{n(j)}, \tag{11}$$

$$y_j(t) = f(net_j(t)),$$

where m is the quantity of state nodes.

The output produced by the given network for both cases is computed using state and group output weights W, given by Eq. (12).

$$net_k(t) = \sum_{j}^{M} y_j(t) W_{n(kj)} + B_{n(k)}, \tag{12}$$

$$y_k(t) = g(net_k(t)),$$

Table 1 Pseudo-code for CS algorithm

Pseudo-code for CS algorithm:

Begin
Step 1: Compute the objective function $f(x), x = (x_1, \ldots \ldots .., x_d)^T$
Step 2: Produce an initial population of host n nests x_i $(i = 1,2, \ldots, n)$
While (t < MaxGeneration) or (Stop Criterion)
Step 3: With "L'evy" flights randomly get a cuckoo by using Eq. (5)
Step 4: Evaluate its quality/fitness function F_i
Step 5: Choose a nest among n (say, j) randomly
If ($F_i > F_j$)
Step 6: Replace j by the new solution;
End
Step 7: A fraction (p_a) of worse nests occurs;
The worst nests are abandoned and new nests are built;
Step 8: Keep the best solutions (or nests with quality solutions);
Step 9: Rank the solutions and find the current best solution;
End While
Step 10: Post-process results and visualization;
End

$$E = (T_k - Y_k) \tag{13}$$

$$V(x) = \frac{1}{2}\sum_{K=1}^{K}(T_k - X_k)^T(T_k - Y_k), \tag{14}$$

$$V_F(x) = \frac{1}{2}\sum_{K=1}^{K} E^T \cdot E$$

$$V_\mu(x) = \frac{\sum_{j=1}^{N} V_F(x)}{P_i} \tag{15}$$

$$SSE_i = \left\{V_\mu^1(x), V_\mu^2(x), V_\mu^3(x), \ldots, V_\mu^n(x)\right\} \tag{16}$$

$$x_j = \mathrm{Min}\left\{V_\mu^1(x), V_\mu^2(x), V_\mu^3(x), \ldots, V_\mu^n(x)\right\} \tag{17}$$

$$x_i^{t+1} = x_i^t + \propto \oplus \mathrm{Levy}(\lambda) \tag{18}$$

$$X = \begin{cases} x_i + rand \cdot (x_j - x_i) & rand_i > P_\alpha \\ x_i & \text{else} \end{cases} \tag{19}$$

$$\nabla X_i = \begin{cases} x_i \oplus \propto \mathrm{L\acute{e}vy}(\lambda) \sim 0.01 \cdot \left(\dfrac{U_j}{\left|V_j^{1/\mu}\right|}\right) \cdot (X - X_{\mathrm{best}}) & rand_i > P_\alpha \\ x_i & \text{else} \end{cases} \tag{20}$$

where ∇X_i is a small movement of x_i toward x_j.

For each layer, the bias and weights are then adjusted as follows:

$$W_n^{s+1} = U_n^{s+1} = W_n^s - \nabla X_i, \tag{21}$$

$$B_n^{s+1} = B_n^s - \nabla X_i.$$

As indicated by the gradient descent algorithm, each change of weight in the constructed network should be directly proportional to the cost of the negative value of the gradient as regards the particular weights, which is given by:

$$\delta_{pk} = -\eta \frac{\delta E}{\delta y_{pk}}. \tag{22}$$

Thus, the output nodes error term is determined as follows:

$$\delta_{pk} = \left(T_{pk} - Y_{pk}\right) Y_{pk}\left(1 - Y_{pk}\right) \tag{23}$$

The error term for hidden nodes is calculated as follows:

$$\delta_{pj} = \sum_k^m \delta_{pk} W_{n(kj)} f'\left(Y_{pj}\right). \tag{24}$$

Hence, the bias and weights in the output layer are obtained as follows:

$$\nabla W_{(kj)}^{n+1} = \eta \sum_p^n \delta_{pk} y_{pj}, \tag{25}$$

$$\nabla B_{(kj)}^{n+1} = \eta \sum_p^n \delta_{pk} y_{pj},$$

The weight update for input layer is obtained using:

$$\nabla W_{(ji)}^{n+1} = \eta \sum_p^n \delta_{pj} x_{pi}, \tag{26}$$

$$\nabla B_{(ji)}^{n+1} = \eta \sum_p^n \delta_{pj} y_{pi},$$

Including a timestamp subscript, the recurrent weights can be transformed using Eq. (27):

Table 2 Pseudo-code for CSLSTM algorithm

Pseudo-code for CSLSTM algorithm:

Begin

Step 1: Initialize CS population dimension size and LSTM network structure.
Step 2: Load training dataset.
Step 3: **While (Value of Mean Square Error < Stop Criterion)**
Step 4: Pass cuckoo nests values to the given LSTM network as weights.
Step 5: Implement a feed forward network using the initialized weights based on CS.
Step 6: Using the value of the previous layer, compute the sensitivity of the next layer. This procedure begins at the last layer of the network and moves backward using Eq. (23) and (24).
Step 7: Update the weights and bias of each layer using Eq. (25) and (26).
Step 8: Calculate the value of the error term using Eq. (13).
Step 9: Using CS, adjust the parameters of the network and minimize the error.
Step 10: Generate a cuckoo egg (x_j) from a random nest through "L'evy" flight behavior. $x_i = x_j$
Step 11: A fraction pa $\in$ [0, 1] of the worst nest is abandoned. Assemble a new nest at the new area, using "L'evy" flight to replace the former nest.
Step 12: Select a random cuckoo nest i and assess the fitness values of the chosen nest,
If
(a) $X_j > X_i$ Then
(b) $x_i \leftarrow x_j$
(c) $X_i \leftarrow X_j$
End if
Step 13: CS continues to compute the possible best weight at every cycle until the network converges.
End while

End

$$\nabla U_{(jh)}^{n+1} = \nabla U_{(jh)}^{n} + \eta \sum_{p}^{n} \delta_{pj}(t) y_{ph}(t-1). \tag{27}$$

The CSLSTM network error is determined using Eq. (13). For the network, performance indices are estimated using Eqs. (14) and (15). Toward the end of every cycle, the ith iteration SSE can be determined using Eq. (16). The CS is used to model the minimum value of SSE, which is obtained when all the inputs have been considered for every population of the cuckoo nest. Consequently, the nest of cuckoo search x_j is computed by Eq. (17). A new solution x_i^{t+1} aimed at cuckoo i is obtained using "Lévy" flight given by Eq. (18). The development of next cuckoo x_i to x_j is obtained using Eq. (19). The CS can move from x_i to x_j base on "Lévy" flight using Eq. (20). Furthermore, the bias and weights for every layer are balanced using Eq. (21). Table 2 gives the pseudo-code for the CSLSTM network procedure.

3.5 LSTM-EFG model

RNN is the preferred term for the combination of recurrent neural network and recursive neural network. Inter-neuronal relations of time recurrent NN structure a grid, whereas auxiliary recurrent neural network recursively builds increasingly complex deep neural networks using similar neural network structures. Thus, RNN usually denotes a time-based recurrent neural network. Time-series-based recurrent neural network can reflect dynamic time behavior and different from feedforward NN that accepts input from a specific structure. Consistently, the RNN passes states in its very individual network, in this way allowing a more extensive scope of a time-series data structure of inputs.

LSTM is a type of time-series-based recursive neural network and a unique type of RNN that determines how to depend on data for a long time. Furthermore, it can predict and handle critical issues through several interims and deferrals ranges in time series. LSTM was developed by Hochreiter and Schmidhuber (1997) and improved by Graves et al. (2004). It has made significant progress and has been extensively utilized. For this analysis, two peepholes are introduced to the LSTM structure to accurately forecast wind power.

3.5.1 LSTM-EFG technique

In this article, LSTM is enhanced with respect to the accompanying main four modifications: (1) peepholes of two numbers (f-o) are added, (2) tanh activation function is changed to softsign activation function, (3) input gate is deleted, and (4) subtracting the previous output value through the forget-gate layer from the impact of the entire single matrix and, at that time, spinning is used to update. The enhanced model can exploit the forget-gate layer effects and increase the rate of convergence of the algorithm. Furthermore, the LSTM-EFG model is used to forecast wind power, to demonstrate a better performance of the model based on the results. The structure of the proposed model used in the paper is shown in Fig. 3.

The whole system of network incorporates an input layer, a layer for hidden neurons, and an output layer, which are completely associated with one another and represented by blue lines. The input layer gives the dimension of the input neuron (white block squares), which gives the number of features that enter the hidden layer. A red line represents cell state C while H is reliable state with Fig. 3. Meanwhile, the black arrow represents the state H yield of the past layer of LSTM-EFG and fills as input for the following layer of the LSTM-EFG model (Yu et al. 2019).

3.5.2 The computation steps of the proposed LSTM-EFG model

The computation steps of the proposed LSTM-EFG model used in this paper are given as follows:

1. Determine the data to be disposed of starting from cell state. This is determined by the forget-gate layer that delivers the $C_{(t-1)}$, x_t and $h_{(t-1)}$ cell states, thereby yielding a value ranging from 0 to 1 to the $C_{(t-1)}$ cell state. The value 1 implies “complete reservation,” while 0 implies “completely abandoned.” This is indicated by blue lines in Fig. 3, and f_t is given by Eq. (28).

$$f_t = \text{sigmoid}\left(f + W_f * C_{t-1}\right) \tag{28}$$

2. Determine the data to be saved in the cell state. This procedure could be partitioned into two sections: An entire single matrix in a given output value is subtracted from the previous value of the forget gate. This determines the value that will be subsequently updated. Thus, the softsign activation function will produce another vector C_t^1 which will be included in the state. This segment is shown in Fig. 3 using black lines, and C_t^1 is given by Eq. (29).

$$C_t^1 = \text{softsign}(W_c * [h_{t-1}, x_t] + b_c). \tag{29}$$

3. The old cell status is updated from C_t to $C_{(t-1)}$. The old state $C_{(t-1)}$ is multiplied with f_t to remove the data that need to be disposed. Thus, $C_t^1 * (1 - f_t)$ is included to obtain a new C_t cell state, which is given by Eq. (30).

$$C_t = f_t * C_{t-1} + (1 - f_t) * C_t^1 \tag{30}$$

4. Determine the output value. This output is obtained from the cell state; however, it is additionally given in sifted version. A sigmoid activation layer is executed to determine the cell state part as the output. The softsign function is then used in the cell state process to obtain a series of value ranging from − 1 to 1, which is multiplied by the output value of the sigmoid function to obtain the expected output. This segment is shown in Fig. 3 using yellow lines, and O_t is given by Eq. (31). The final output value h_t is given by Eq. (32).

$$O_t = \text{sigmoid}(O + W_o * C_t) \tag{31}$$

$$h_t = O_t * \text{softsign}(C_t) \tag{32}$$

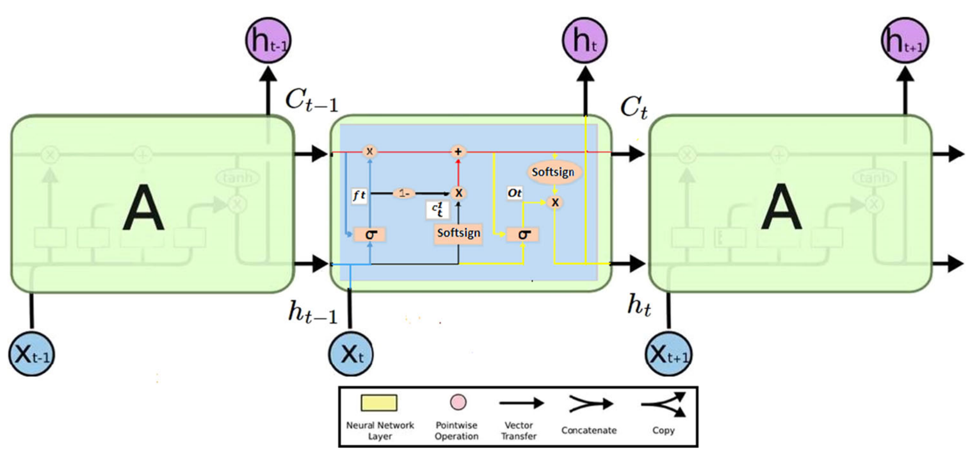


Fig. 3 Proposed LSTM-EFG network structure (color figure online)

4 Construction and implementation of proposed model

The proposed wind energy forecasting model is constructed, to implement out the day-ahead (24 h) wind power forecasting experimentations with the historical data of wind power and speed. Each of the tests is performed in tensorflow condition using PC with processor Intel® Xeon® CPU E5-2630 v4 @ 2.20 GHz and 32.0 GB RAM under Windows 10 Pro Operating System of 64 bits. Experimental results of forecasting elucidate that the projected EEMD-LSTM-EFG-CSO model has the highest accuracy of forecasting of wind power by means of lowest statistical error evaluation indices.

4.1 Empirical wind speed and power time-series data

The historical data of wind power and speed time series, generated every 15-min interval timestamp, are collected and warehoused using the centralized server of National Institute of Wind Energy (NIWE) of India, for estimating and validating the suggested wind power forecasting model. The substation dataset with total rated capacity of 119.845 MW of power from Coimbatore region of Tamil Nadu state is used. The dataset consists of 24,192 wind powers and speed data points, respectively, that are selected randomly from 2017 to 2018 as empirical examples, as given in Fig. 4. In the empirical examples, the preceding 23,904 of wind power and speed data are utilized to train the proposed EEMD-LSTM-EFG-CSO model. To test the proposed model, 288 data samples of both datasets are used. Finally, to validate the model subsequent 96 rows (24 h with 15 min interval) time-series data with forecast NWP wind speed are utilized for day-ahead forecasting. The explanations of statistical measures of empirical wind data are enumerated in Tables 3 and 4.

Considering the accuracy factor, the complete one-year data are selected as a training pattern. If the training period is more (2 years/3 years), the data samples will be good enough for building the better forecasting model. Generally, if the training period is less (less than 1 year/ 6 months) may lead to fall in the accuracy of the model.

From the above-mentioned Fig. 4 and Table 4, high fluctuations have existed within [0.08, 21.15] and [0.15, 18.76] in the wind speed data of dataset.

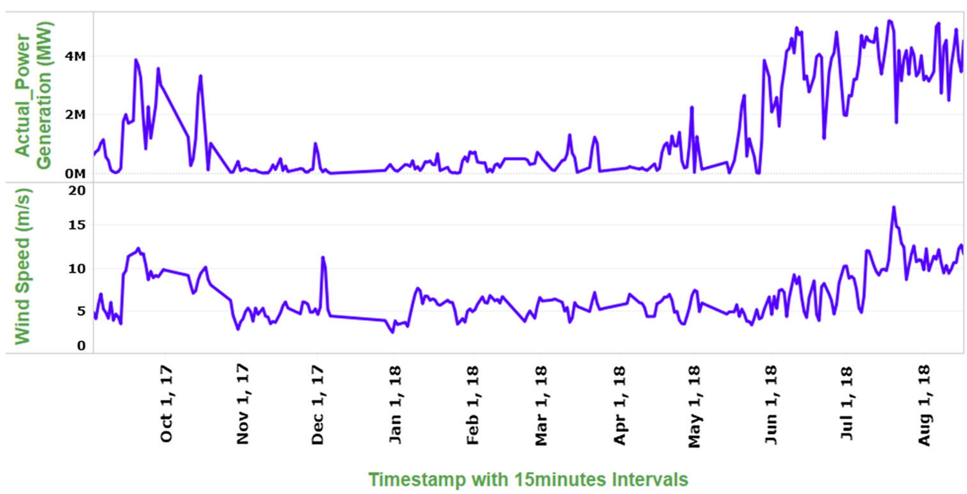


Fig. 4 Empirical original wind speed and power data of dataset

Table 3 Statistical explanations of dataset empirical wind power time series

Dataset (MW)	No. of samples	Min	Max	Mean	Median	SD
Total	**24,192**	**0**	**68**	**16.34**	**6.39**	**18.95**
Training	**23,904**	**0**	**68**	**15.24**	**6.28**	**18.89**
Test	**288**	**3**	**61.81**	**41.02**	**44.65**	**11.76**

Table 4 Statistical explanations of dataset empirical wind speed time series

Dataset (m/s)	No. of samples	Min	Max	Mean	Median	SD
Total	**24,192**	**0.08**	**21.15**	**6.85**	**6.32**	**3.2**
Training	**23,904**	**0.08**	**21.15**	**6.83**	**6.3**	**3.18**
Test	**288**	**8.95**	**16.16**	**12.17**	**11.97**	**1.39**

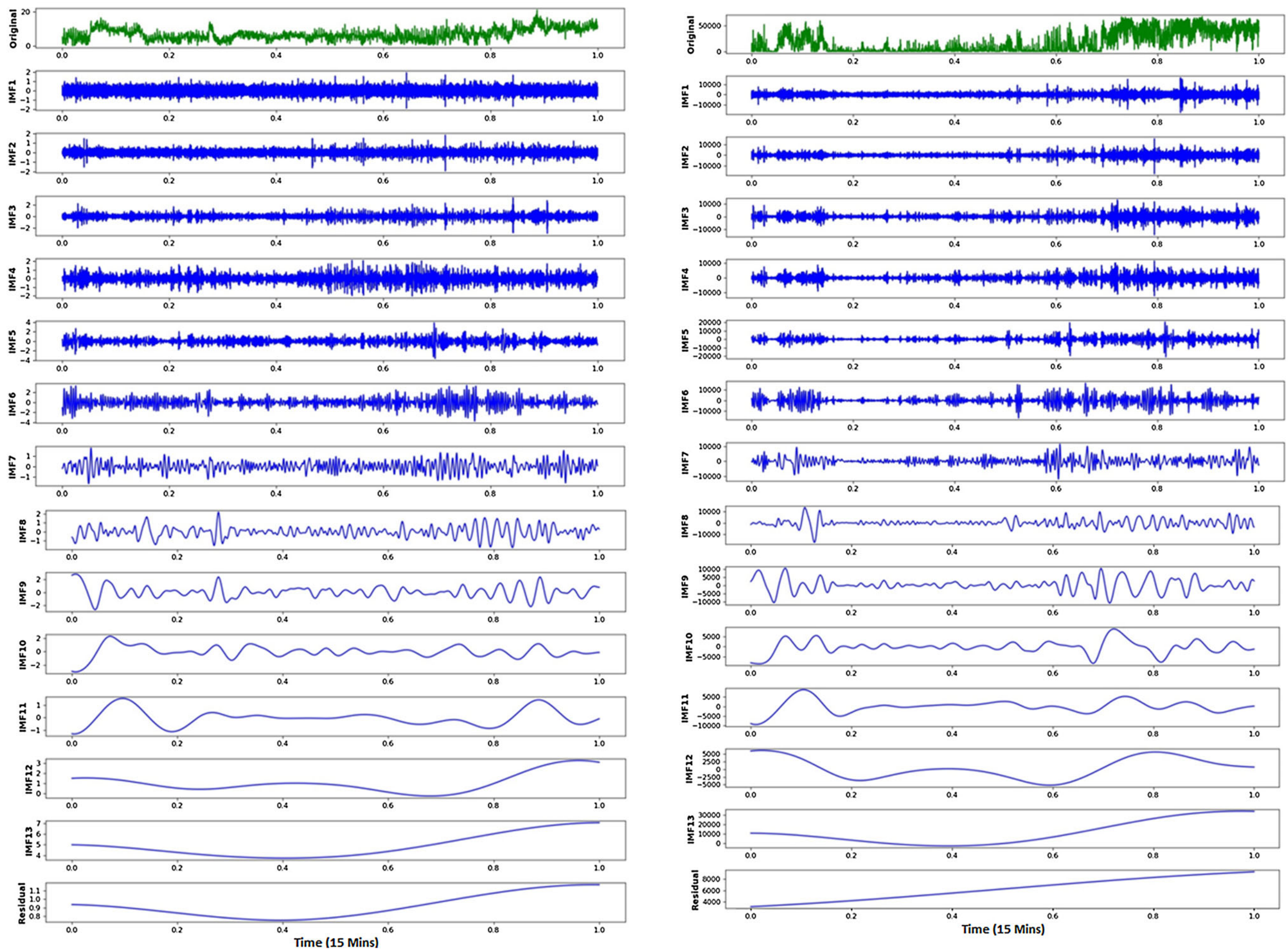


Fig. 5 Wind speed and power series decomposition using EEMD of the dataset

4.2 Setting model parameters for forecasting

To test and validate the accuracy of the proposed model EEMD-CSO-LSTM-EFG, the wind power and speed series from two datasets are chosen. The proposed forecasting model results are likewise compared with other existing conventional forecasting models such as the ARIMA model, BP neural networks, and basic LSTM model. The paper tracked the energy policies and standards and estimating rules of wind sources that were framed by the National Energy Policy (NEP) in 2018.

(1) The Keras model from tensorflow environment exists is used to construct the BPNN network model. The neuron dimensions used in the network for a layer of input are 4, the layer of hidden is 5, and a layer of output is 1. The maximum quantity of epochs is fixed as 100, the value of learning rate is fixed at 0.1, and the value 0.00004 is set for training precision.

(2) The order of autoregressive and the moving average are the most influencing parameters of ARIMA forecasting model. The AIC performance metric is used to measure the fitting effect of the observed value, and AIC is also used for the model's parameter estimation. ARIMA model can reach the best order with the lowest value of AIC.

(3) For the LSTM model, similar to BPNN, the Keras model from the tensorflow environment is employed to build the basic LSTM forecasting network. The same values for parameters as like BPNN are used in the LSTM network.

4.3 Wind power and speed decomposition using EEMD technique

The empirical original wind speed and wind power time-series data are first preprocessed and decomposed using the EEMD technique. Figure 5 represents the decomposed data acquired by means of EEMD technique for the chosen substation dataset. From the training data series, we obtained 14 IMF sequences for the dataset, as shown in Fig. 5. As indicated by the denoising standards, removing the sequence with a high-frequency value from the obtained IMF sequences could help in getting a cleaner data sequence, specifically lower noise in the data sequence. In this paper, the principal IMF sequence acquired through the EEMD strategy is removed from the raw data sequence because of its high frequency. Stationary time-series data generated using this decomposition help to improve the forecasting accuracy. It is obvious that each wind speed and power series is decomposed into 14 components which are, respectively, denoted by IMF1, IMF2,…, IMF13, and R from top to bottom.

Remark 1 The EEMD decomposition method evidenced to be an efficient method to attain the purpose of the data preprocessing and diminishes the fluctuations in the raw data series.

4.4 Difference between the single-step forecasting and multistep forecasting

To conduct the experimentation, the recommended forecasting technique is trained with the data nominated for dataset. The experiment is also implemented for single-step and multistep forecasting, where for each circulation the prediction is performed by eliminating the old data input. The multistep-ahead strategy predicts the forthcoming wind power and speed values with this circulation, by applying the values of past output instead of the actual series values. Dependent on the final multistep method of forecasting results, the projected forecasting models validity was analyzed. The forecasting results of both the single-step- and multistep-ahead methods of forecasting technique are demonstrated in Tables 6, 7 8.

Affording with statistical performance indices, the values of Tables 6 and 8 demonstrate that the multistep forecasting methods give better accuracy values compared to single-step forecasting methods.

Specifically, the number of layers and units in LSTM-EFG neural network needs to be determined properly. Therefore, this paper selects a suitable number of layers and units in LSTM-EFG for these IMFs and residual series in Table 5.

Table 5 The number of layers and units in LSTM-EFG for each IMF

IMF	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10	IMF11	IMF12	IMF13	R0
Layers	**2**	**3**	**1**	**2**	**1**	**3**	**3**	**2**	**1**	**2**	**3**	**2**	**1**	**2**
Units	**100**	**200**	**50**	**50**	**100**	**200**	**100**	**50**	**100**	**200**	**100**	**50**	**100**	**50**

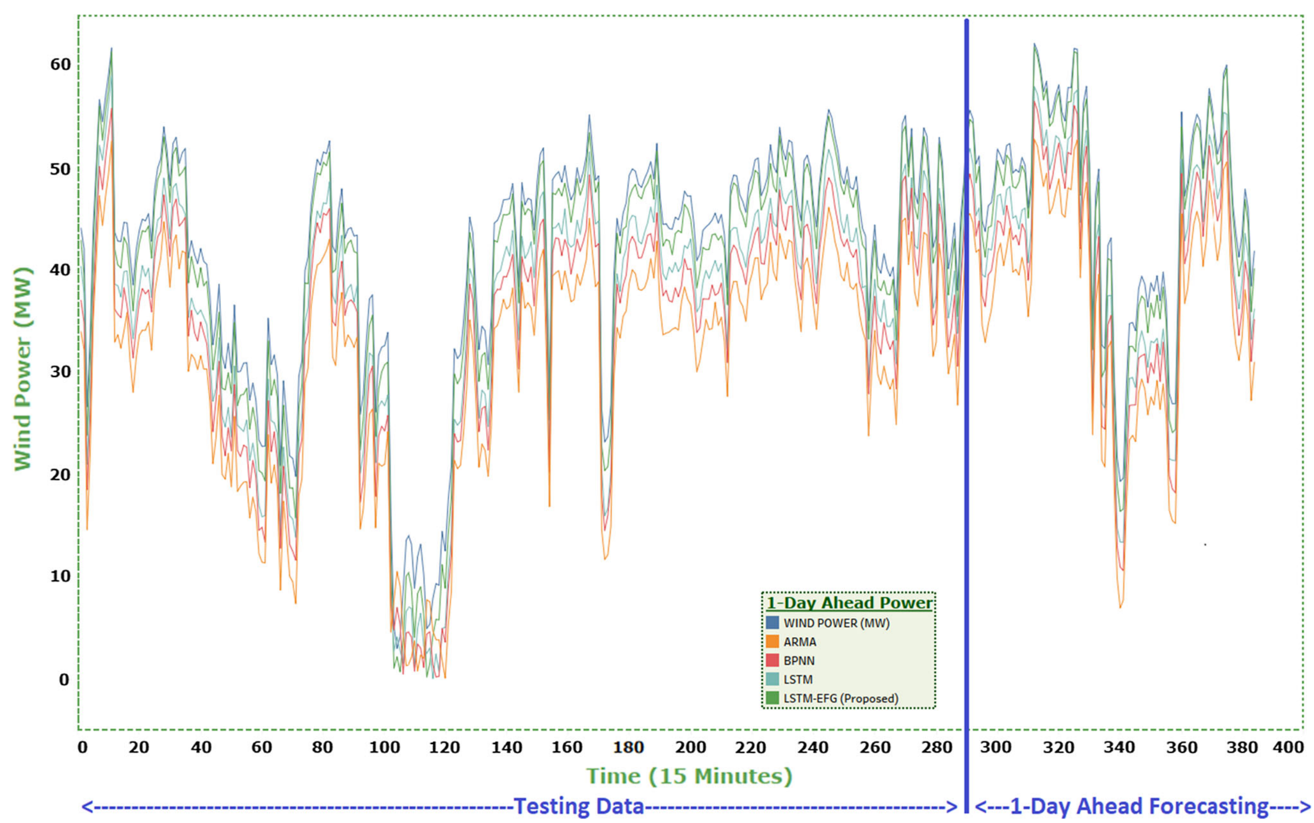


Fig. 6 Day-ahead (24 h) single-step results of forecasting for the proposed and the other conventional models without EEMD signal decomposition for dataset

5 Numerical experiments and forecasting results analysis

This section analyses the experimental outcomes of the hybrid method, and the efficiency of the projected strategy is confirmed. To begin with, in light of the consequences of forecasting using a single step, the hybrid methodology is contrasted using several surviving traditional forecasting models. At this point, the multistep-ahead forecasting results are utilized for additional investigation.

5.1 Analysis of single-step forecasting

The section is separated into the two subsections to confirm the effectiveness of the proposed hybrid approach. The proposed method is mainly verified by making a comparison with existing conventional forecasting models. The results of single-step forecasting are shown in Figs. 6 and 7 and in Table 6.

5.2 Multistep forecasting analysis

The efficient approach to assess the forecasting method accuracy is multistep forecasting. Hence, the research utilized multistep forecasting for the projected hybrid forecast model. The outcome of the experimentation is used to validate the proposed hybrid model accuracy. In this work, two classes of forecasting models have been built and verified: The principal classification is meant for developing separate regression algorithms without EEMD decomposition methods, such as ARMA model, BPNN model, LSTM model, and the proposed LSTM-EFG model; the subsequent classification exists to build EEMD signal disintegration-based wind energy forecasting models, namely EEMD-CSO-SVM, EEMD-GA-BPNN [62], EEMD-LSTM, and EEMD-CSO-LSTM-EFG. The hyper-parameters for the EMD-CSO-SVM and EEMD-GA-BPNN have been fixed based on the corresponding references. Actual wind power generation data are used to test and validate all the forecasting models.

The multistep forecasting results are shown in Tables 7 and 8. Figures 8 and 9 indicate the outcomes of various models with multistep forecasting. The MAPE values of one-step-ahead forecasting of the ARMA, BPNN, LSTM, and LSTM-EEG (proposed technique) are 1.109%, 0.842%, 0.597%, and 0.331%, respectively. MAPE values of the aforesaid models with two-step-ahead forecasting are 5.850%, 4.336%, 3.281%, and 1.729%, respectively.

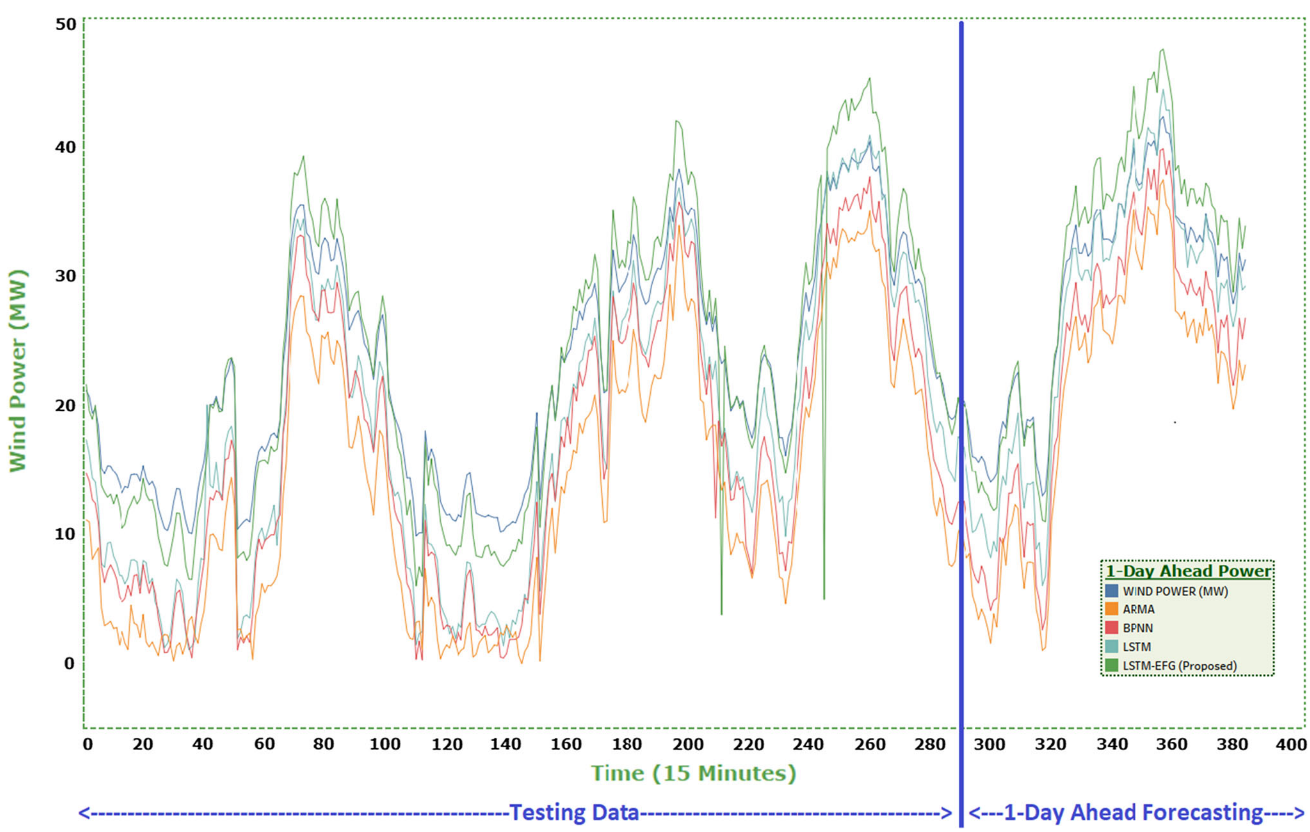


Fig. 7 Day-ahead (24 h) single-step results of forecasting for the proposed and the other conventional models with EEMD signal decomposition for dataset

Table 6 Results of statistical performance indices achieved by single-step forecasting for dataset

Index	Forecasting horizon	RMSE (MW)	MAE (MW)	MAPE (%)	MASE (MW)
ARMA	Single step (24 h)	12.481	12.474	28.26	2.95
BPNN		9.392	9.385	21.26	2.219
LSTM		7.543	7.529	16.97	1.781
LSTM-EFG (proposed)		**4.326**	**4.309**	**9.71**	**1.019**
EEMD-CSO-SVM	Single step (24 h)	11.454	11.703	27.26	2.736
EEMD-GA-BPNN		8.622	8.707	20.59	2.126
EEMD-LSTM		6.002	6.010	15.92	1.494
EEMD-CSO-LSTM-EFG (proposed)		**4.387**	**4.326**	**8.53**	**0.959**

Table 7 Results of statistical performance indices without EEMD decomposition achieved by forecasting models for dataset

Index	Forecasting horizon (step)	MAE (MW)	RMSE (MW)	MAPE (%)	MASE (MW)
ARMA	1	0.585	2.568	1.11	0.138
	2	2.959	5.805	5.85	0.700
	3	5.821	8.150	11.56	1.377
	4	11.464	11.471	25.98	2.711
BPNN	1	0.444	1.944	0.84	0.105
	2	2.193	4.301	4.34	0.519
	3	4.322	6.054	8.59	1.022
	4	8.395	8.402	19.02	1.985
LSTM	1	0.321	1.390	0.60	0.075
	2	1.662	3.266	3.28	0.393
	3	3.355	4.710	6.65	0.794
	4	6.519	6.535	14.69	1.542
LSTM-EFG (proposed)	**1**	**0.175**	**0.774**	**0.33**	**0.041**
	2	**0.874**	**1.727**	**1.73**	**0.207**
	3	**1.708**	**2.406**	**3.38**	**0.404**
	4	**3.299**	**3.320**	**7.41**	**0.780**

Table 8 Results of statistical performance indices with EEMD decomposition technique achieved with models of forecasting for dataset

Index	Forecasting horizon (step)	MAE (MW)	RMSE (MW)	MAPE (%)	MASE (MW)
EEMD-CSO-SVM	1	0.455	2.168	0.91	0.118
	2	2.159	4.505	4.15	0.56
	3	4.221	6.455	9.66	1.077
	4	9.264	9.271	20.28	2.131
EEMD-GA-BPNN	1	0.344	1.544	0.54	0.099
	2	1.903	3.101	3.13	0.319
	3	3.122	5.254	7.89	0.922
	4	6.195	7.302	15.12	1.585
EEMD-LSTM	1	0.215	0.987	0.50	0.055
	2	1.232	2.266	2.18	0.253
	3	2.355	3.51	5.65	0.624
	4	4.119	5.135	10.69	1.142
EEMD-CSO-LSTM-EFG (proposed)	**1**	**0.155**	**0.654**	**0.23**	**0.031**
	2	**0.734**	**1.427**	**0.93**	**0.107**
	3	**1.208**	**2.106**	**2.18**	**0.324**
	4	**3.129**	**3.02**	**5.41**	**0.58**

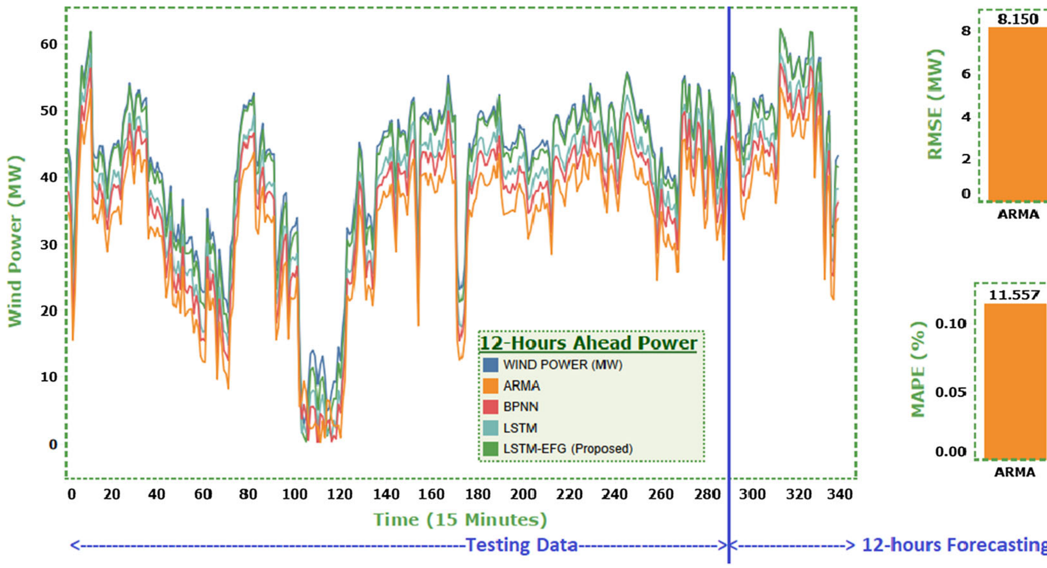
(a) 1 - hour forecasting results
(b) MAPE and RMSE for 1 - hour
(c) 6 - hours forecasting results
(d) MAPE and RMSE for 6 - hours

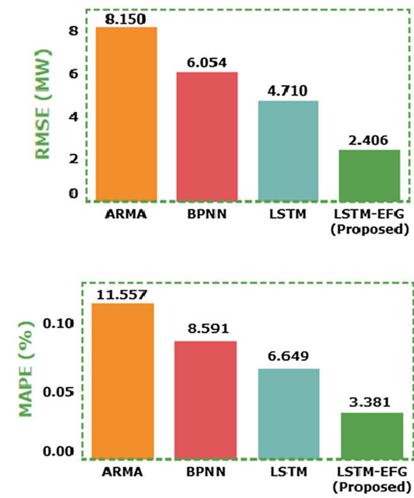
Fig. 8 Results of forecasting for the proposed and the other conventional models without signal decomposition for dataset

Additionally, the MAPE values of these four models with three-step-ahead forecasting are 11.557%, 8.591%, 6.649%, and 3.381%, which suggests that the proposed wind power hybrid model accomplishes an, even more, better performance in comparison with conventional forecasting models. One-step forecasting obviously performs superior to the others, while comparing the three-step forecasting results with each other.

Remark 2 These outcomes can be precise as follows: While comparing using the conventional models of forecasting, the proposed hybrid model achieves greater accuracy utilized in numerical experimentations of both multistep-ahead and single-step-ahead forecasting. In this manner, the suggested hybrid forecasting model stands valid.

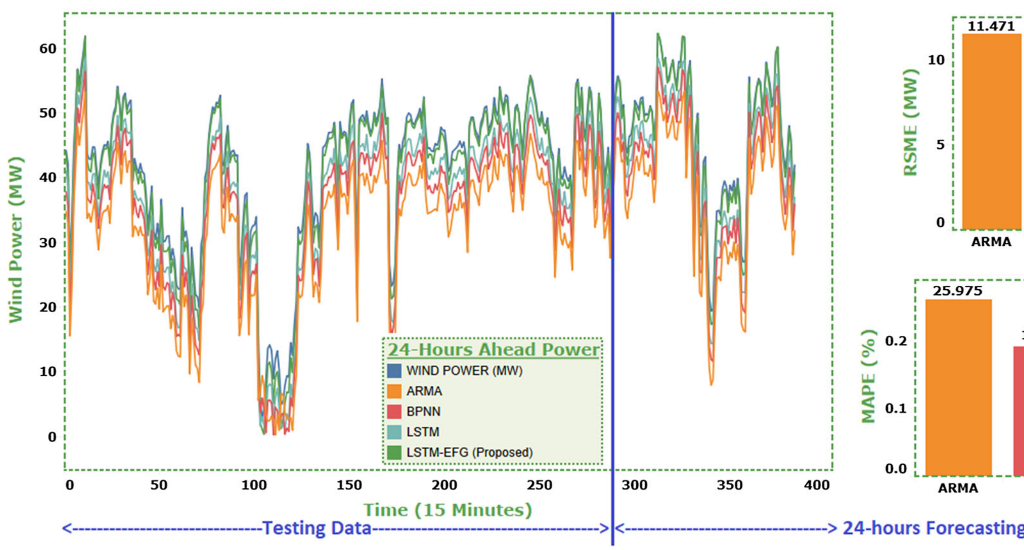
The empirical values of wind power generation and speed, as shown in Fig. 8, are engaged to build and assess the models used for analysis. The forecasting results and the statistical performance error indices are displayed in Figs. 8 and 9, and Tables 7 and 8. From the table values, it is observed that the least error indices reveal the greatest performance forecasting of wind power. Assessments are done on view to affirm the benefits of the feature determination, decomposition of signal strategies, and the optimization of hyperparameters. Tables 7 and 8 and Figs. 8 and 9 represent the following:

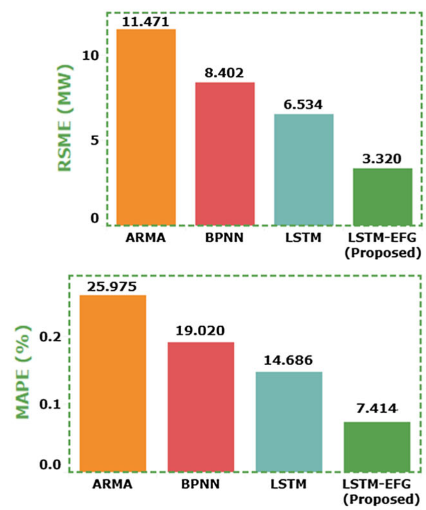




(e) 12 - hours forecasting results

(f) MAPE and RMSE for 12 - hours





(g) 24 - hours forecasting results

(h) MAPE and RMSE for 24 - hours

Fig. 8 continued

- Table 8 demonstrates our proposed model EEMD-CSO-LSTM-EFG accomplishing improved results with respect to forecasting related to RMSE values of 0.653 MW, 1.426 MW, 2.106 MW, and 3.020 MW for one-step-ahead, two-step-ahead, three-step-ahead, and four-step-ahead forecasting, respectively.
- ARMA shows the highest values of forecasting statistical performance indices with 2.568 MW, 5.805 MW, 8.154 MW, and 11.471 MW for one-step-ahead, two-step-ahead, three-step-ahead, and 4-step-ahead forecasting, respectively.
- Figure 8b, d, f, h shows that the accuracy values of the individual forecasting models are ordered from minimum values to maximum values as per ARMA, BPNN, LSTM, and LSTM-EFG.
- Figure 9b, d, f, h shows that the accuracy values of the individual forecasting models with signal decomposition and parameter optimization are ordered from low to high as EEMD-CSO-SVM, EEMD-GA-BPNN, EEMD-LSTM, and EEMD-CSO-LSTM-EFG.
- Affording with statistical performance index, the RMSE values of Table 8 and Fig. 9b, d, f, h demonstrate the highest second to the highest fourth accuracy values of models in the order of EEMD-LSTM without optimization technique, EEMD-GA-BPNN with GA optimization function, and EEMD-CSO-SVM with CSO optimization method.

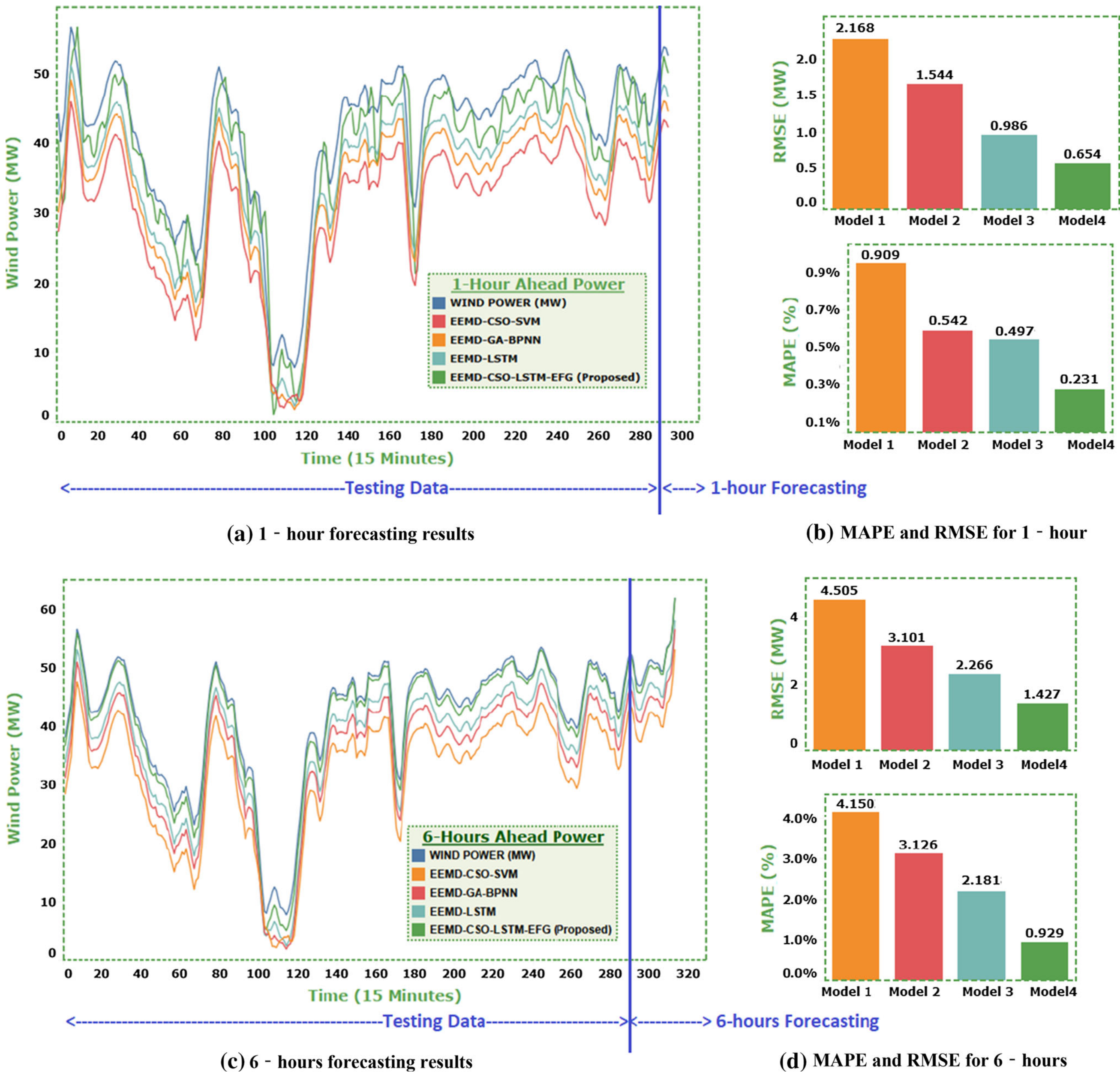
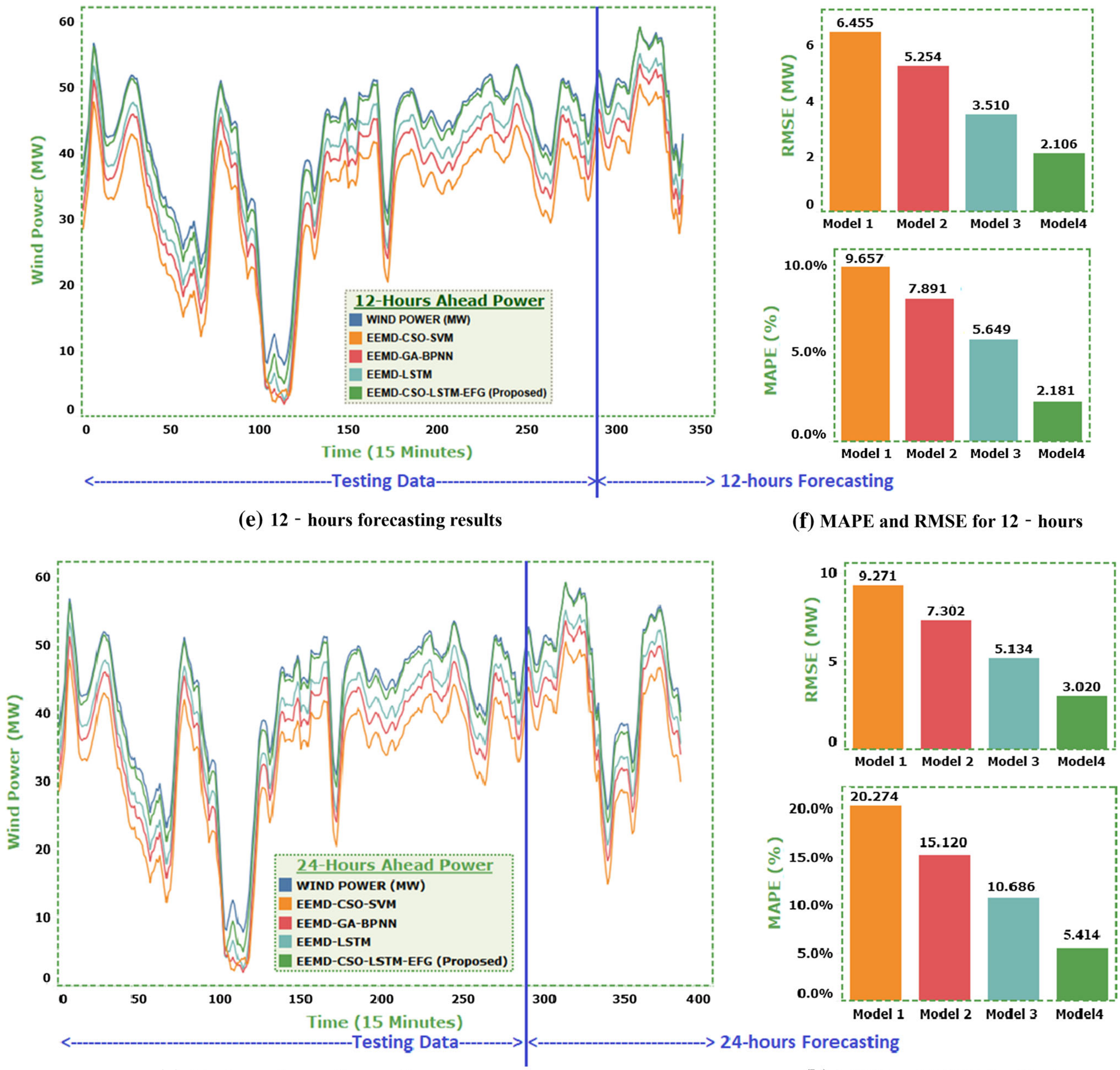


Fig. 9 Results of forecasting for the proposed and the other conventional models with EEMD signal decomposition for dataset [model 1 EEMD-CSO-SVM, model 2 EEMD-GA-BPNN, model 3 EEMD-LSTM, model 4 EEMD-CSO-LSTM-EFG (proposed)]

- In accordance with Figs. 8 and 9, it is clearly realized that the values of RMSE for the LSTM-established forecasting approach are graded from high values to low values in the order of LSTM, LSTM-EFG, EEMD-LSTM, and the EEMD-CSO-LSTM-EFG.
- With the intention of explaining the efficiency of decomposition and optimization techniques in LSTM, the suggested model EEMD-CSO-LSTM-EFG is compared using the existing conventional LSTM models. Tables 7 and 8 prove the outcomes of forecasting. The values of RMSE for the proposed forecasting model EEMD-CSO-LSTM-EFG are the smallest. Thus, the proposed approach outperforms the existing model without both signal decomposition function and parameter optimization techniques.

Remark 3 The proposed EEMD-CSO-LSTM-EFG accomplishes better than all individual forecasting models considered and the other hybrid models once implemented in the multistep wind power forecasting.

- When handling nonlinear wind power and speed time-series data, the artificial intelligent models possess the



(e) 12 - hours forecasting results

(f) MAPE and RMSE for 12 - hours

(g) 24 – hours forecasting results

(h) MAPE and RMSE for 24 - hours

Fig. 9 continued

better capability, and the individual artificial intelligence-based models such as BPNN and LSTM provide improved performance in forecasting in comparison with that of distinct statistical methods ARMA.

- Forecasting models based on signal decomposition acquire wonderful enhancements when compared to distinct models of forecasting without using signal decomposition strategies due to the existence of non-stable, extraordinary fluctuation and nonlinearity characteristics of wind speed data of time series. The given wind speed data of time series are broken down into various stationary subseries by using the signal decomposition methods, thus decreasing the regression problems of the forecasting system for improved prediction.
- Examinations between the CSO-LSTM models with cuckoo search optimization technique along with individual LSTM model without optimization technique utilizing the similar data of wind speed were used to inspect the ability of CSO in the optimization of parameters and feature selection. From Fig. 8 and Table 7, LSTM-EFG performs superior to LSTM in place of multistep forecasting. The causes of model outcomes are a selection of feature strategy by LSTM-

EFG technique disposing the misleading components and recognizing the powerful components, whereas CSO algorithm remains used as an optimization of the parameter to tune parameters in LSTM network, in this way improving the performance of wind energy forecasting.

- The CSO-LSTM-EFG model with EEMD signal decomposition technique has greater accuracy of wind power forecasting compared to LSTM-EFG model without EEMD, and the fundamental causes are the signal decomposition which could be efficiently dealt with the difficulties of the IMF1 component irregularity. Thus, the EEMD signal decomposition method is an effective data preprocessing technique in enlightening wind power forecasting performance.

6 Conclusions and future work

The increasing interest in sustainable wind power has a very good future prospect. Meanwhile, accuracy and stability are two key factors that need to be considered when constructing a model in the forecasting domain. Consequently, it is important to develop an hourly day-ahead wind energy forecasting model with an acceptable level of accuracy and stability. This paper proposed a novel hybrid predictive model that uses a forecasting neural network module (LSTM) and an algorithm for hyperparameters CSO optimization. The weight measurements for the LSTM network forecasting modules were enhanced by using CSO algorithm. Since the noise of the raw wind data can be reduced, EEMD was used as a data preprocessing technique to transform the original data series into stationary time-series data. To confirm the legitimacy of the proposed forecasting model, several statistical assessment performance metrics with both single-step- and multistep-ahead forecasting methods were considered in this paper.

The proposed model was compared with other existing distinct and hybrid forecasting models referenced in this investigation to feature the efficiency of the proposed EEMD-CSO-LSTM-EFG model while connected in the wind energy forecasting. The projected forecasting model holds a lesser value of RMSE 0.653 MW, whereas RMSEs 2.168 MW, 1.544 MW, and 0.986 MW for the EEMD-CSO-SVM, EEMD-GA-ANN, and EEMD-LSTM, respectively, use one-step forecasting. Similarly, the model proposed yields lesser statistical performance error with indices such as MAE, MASE, MAPE, and RMSE values in a two-step, three-step, and four-step forecasting. While compared to the pure conventional individual forecasting models without signal processing technique, the prediction models constructed using the signal decomposition technique and optimization methods produces better forecasting results for the given time-series data. Thus, the proposed model is an effective wind power forecasting method compared to the traditional forecasting methods. In view of the numerical experimental analysis and results, the proposed model has a good accuracy and precision. Hence, it can be used as an extraordinary tool in managing the substations of power systems.

The future directions of this paper to be considered are: (1) To improve the efficiency of proposed wind power forecasting model, the sliding window concept can be implemented to train the model. (2) The behavior of the model can be extended for long term with varying forecasting time horizons. (3) The proposed model can also be tested for different seasonal conditions like windy season, non-windy season, and transient wind season. (4) The model can also be tested and compared with some other decomposition techniques as well.

Acknowledgements The authors would like to thank National Institute of Institute of Wind Energy (NIWE), Chennai, for the opportunities provided to conduct technical assistance for this study. The authors would also wish to extend their whole-hearted gratitude to Dr. K. Balaraman, Director General, NIWE, and other colleagues for providing data and guidance to carry out this research work successfully in the institute.

Compliance with ethical standards

Conflict of interest All authors state that there is no conflict of interest.

Ethical approval We used our own data. Animals and Humans are not involved in this research work.

References

Bengio Y, Simard P, Frasconi P (1994) Learning long-term dependencies with gradient descent is difficult. IEEE Trans Neural Netw 5:157–166

Bludszuweit H, Dominguez-Navarro JA (2011) A probabilistic method for energy storage sizing based on wind power forecast uncertainty. IEEE Trans Power Syst 26(3):1651–1658

Bonanno F, Capizzi G, Sciuto GL, Napoli C (2015) Wavelet recurrent neural network with semi-parametric input data pre-processing for micro-wind power forecasting in integrated generation systems. In: Proceedings of the 2015 international conference on clean electrical power (ICCEP), Taormina, Italy, 16–18 June 2015; pp 602–609

Botterud A, Zhou Z, Wang J, Sumaili J, Keko H, Mendes J, Bessa RJ, Miranda V (2013) Demand dispatch and probabilistic wind power forecasting in unit commitment and economic dispatch: a case study of illinois. IEEE Trans Sustain Energy 4(1):250–261

Brusca S, Capizzi G, Sciuto GL, Susi G (2017) A new design methodology to predict wind farm energy production by means of a spiking neural network-based system. Int J Numer Model Electron Netw Devices Fields. https://doi.org/10.1002/jnm.2267

Cadenas-Barrera JL, Meng J, Castillo-Guerra E, Chang L (2013) A neural network approach to multi-step ahead, short-term wind speed forecasting. In: Proceedings of the 12th international conference on machine learning and applications (ICMLA), Miami, FL, USA, 4–7 December 2013

Chang W-Y (2014) A literature review of wind forecasting methods. J Power Energy Eng 2:161–168

Chang GW, Lu HJ, Hsu LY, Chen YY (2016) A hybrid model for forecasting wind speed and wind power generation. In: Proceedings of the 2016 IEEE Power and Energy Society general meeting (PESGM), Boston, MA, USA, 17–21 July 2016; pp 1–5

Chen J, Zeng G, Zhou W, Du W, Lu K (2018) Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization. Energy Convers Manage 165:681–695

Cheng Y, Xu C, Mashima D, Thing VL, Wu Y (2017) PowerLSTM: power demand forecasting using long short-term memory neural network. In: International conference on advanced data mining and applications. Springer, Cham, pp 727–740

De Alencar DB, de Mattos Affonso C, de Oliveira RCL, Rodríguez JLM, Leite JC, Filho JCR (2017) Different models for forecasting wind power generation: case study. Energies 10:1976

De Aquino RRB, Souza RB, Neto ON, Lira MMS, Carvalho MA, Ferreira AA (2015) Echo state networks, artificial neural networks and fuzzy system models for improve short-term wind speed forecasting. In: Proceedings of the international joint conference on neural networks (IJCNN), Killarney, Ireland, 12–17 July 2015

Foley AM, Leahy PG, Marvuglia A, McKeogh EJ (2012) Current methods, and advances in forecasting of wind power generation. Renew Energy 37:1–8

Gers FA, Schmidhuber J (2000) Recurrent nets that time and count. In: IEEE-INNS-ENNS international joint conference on neural networks, p 3189

Ghorbani MA, Khatibi R, FazeliFard MH, Naghipour L, Makarynskyy O (2016) Short-term wind speed predictions with machine learning techniques. Meteorol Atmos Phys 128:57–72

Global Wind Energy Council (2018) Global wind report

Graves A, Eck D, Beringer N, Schmidhuber J (2004) Biologically plausible speech recognition with LSTM neural nets. In: Bio-ADIT 2004, Lausanne, Switzerland, pp 175–184

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780

Hossain J (2014) World wind resource assessment report

Hu JM, Wang JZ, Zeng GW (2013) A hybrid forecasting approach applied to wind speed time series. Renew Energy 60:185–194

Huang NE, Shen Z, Long SR (1999) A new view of nonlinear water waves: the Hilbert spectrum. Annu Rev Fluid Mech 31:417–457

Jiang Y, Huang GQ (2017) Short-term wind speed prediction: hybrid of ensemble empirical mode decomposition, feature selection and error correction. Energy Convers Manag 144:340–350

Jung J, Broadwater RP (2014) Current status and future advances for wind speed and power forecasting. Renew Sustain Energy Rev 31:762–777

Kishore GR, Prema V, Rao KU (2014) Multivariate wind power forecast using artificial neural network. In: Proceedings of the IEEE global humanitarian technology conference, South Asia Satellite (GHTC-SAS), Trivandrum, India, 26–27 September 2014

Liu H, Mi X, Li Y (2018) Smart multistep deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, lstm network and elm. Energy Convers Manage 159:54–64

Lpez E, Valle C, Allende H, Gil E, Madsen H (2018) Wind power forecasting based on echo state networks and long short-term memory. Energies 11(3):526

Menemenlis N, Huneault M, Robitaille A (2012) Computation of dynamic operating balancing reserve for wind power integration for the time-horizon 1c48 h. IEEE Trans Sustain Energy 3(4):692–702

Olaofe ZO, Folly KA (2012) Wind power estimation using recurrent neural network technique. In: Proceedings of the IEEE Power and Energy Society conference and exposition in Africa: intelligent grid integration of renewable energy resources (PowerAfrica), Johannesburg, South Africa, 9–13 July 2012; pp 1–7

Qiaomu Z, Hongyi L, Ziqi W, Jinfu C, Bo W (2017) Short-term wind power forecasting based on LSTM. Power Syst Technol 41(12):45–49

Qu X, Kang X, Zhang C, Jiang S, Ma X (2016) Short-term prediction of wind power based on deep long short-term memory. In: Power and energy engineering conference, pp 1148–1152

REN (2015) Renewables 2015 global status report, Environ Policy Collect

Song J, Wang J, Lu H (2018) Research and application of a novel combined model based on advanced optimization algorithm for wind speed forecasting. Appl Energy 215:643–658

Sun W, Liu M, Liang Y (2015) Wind speed forecasting based on FEEMD and LSSVM optimized by the Bat algorithm. Energies 8:6585–6607

United Nations (UN) (2012) Sustainable energy for all. United Nations, New York

Valian E, Mohanna S, Tavakoli S (2011) Improved Cuckoo search algorithm for global optimization. Int J Commun Inf Technol, IJCIT 1(1):1–62

Wu ZH, Huang NE (2004) A study of the characteristics of white noise using the empirical mode decomposition method. Proc Math Phys Eng Sci 460:1597–1611

Wu Z, Huang NE (2009) Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv Adapt Data Anal 1:1–41

Wu Q, Peng C (2016) Wind power generation forecasting using least squares support vector machine combined with ensemble empirical mode decomposition, principal component analysis and a bat algorithm. Energies 9:261

Xiao L, Wang J, Hou R, Wu J (2015) A combined model based on data pre-analysis and weight coefficients optimization for electrical load forecasting. Energy 82:524–549

Yang XS, Deb S (2009) Cuckoo search via Lévy flights. In: Proceedings of world congress on nature & biologically inspired computing, India, pp 210–214

Yu R, Gao J, Yu M, Lu W, Xu T, Zhao M, Zhang J, Zhang R, Zhang Z (2019) LSTM-EFG for wind power forecasting based on sequential correlation features. Future Gener Comput Syst 93:33–42

Zahedi A (2010) Australian renewable energy progress. Renew Sustain Energy 14:2208–2213