Energy Conversion and Management 100 (2015) 16–22

Contents lists available at ScienceDirect

Energy Conversion and Management

journal homepage: www.elsevier.com/locate/enconman

# Four wind speed multi-step forecasting models using extreme learning machines and signal decomposing algorithms

Hui Liu [a,b,*], Hong-qi Tian [a], Yan-fei Li [a]

[a] Key Laboratory of Traffic Safety on Track of Ministry of Education, School of Traffic and Transportation Engineering, Central South University, Changsha 410075, Hunan, China
[b] Institute of Automation, Faculty of Computer Science and Electrical Engineering, University of Rostock, Rostock 18119, Mecklenburg-Vorpommern, Germany

ARTICLE INFO

*Article history:*
Received 5 February 2015
Accepted 21 April 2015
Available online 15 May 2015

*Keywords:*
Wind energy
Wind speed forecasting
Wind speed decomposition
Wavelet
Extreme learning machines
Empirical mode decomposition

ABSTRACT

Realization of accurate wind speed forecasting is important to guarantee the safety of wind power utilization. In this paper, a new hybrid forecasting architecture is proposed to realize the wind speed accurate forecasting. In this architecture, four different hybrid models are presented by combining four signal decomposing algorithms (*e.g., Wavelet Decomposition/Wavelet Packet Decomposition/Empirical Mode Decomposition/Fast Ensemble Empirical Mode Decomposition*) and Extreme Learning Machines. The originality of the study is to investigate the promoted percentages of the Extreme Learning Machines by those mainstream signal decomposing algorithms in the multiple step wind speed forecasting. The results of two forecasting experiments indicate that: (1) the method of Extreme Learning Machines is suitable for the wind speed forecasting; (2) by utilizing the decomposing algorithms, all the proposed hybrid algorithms have better performance than the single Extreme Learning Machines; (3) in the comparisons of the decomposing algorithms in the proposed hybrid architecture, the Fast Ensemble Empirical Mode Decomposition has the best performance in the three-step forecasting results while the Wavelet Packet Decomposition has the best performance in the one and two step forecasting results. At the same time, the Wavelet Packet Decomposition and the Fast Ensemble Empirical Mode Decomposition are better than the Wavelet Decomposition and the Empirical Mode Decomposition in all the step predictions, respectively; and (4) the proposed algorithms are effective in the wind speed accurate predictions.



## 1. Introduction

As one of the most potential renewable energies, wind energy has taken much attention [1]. In the wind energy, to protect the safety of the power integration and scheduling, it is desired to realize the wind speed high-precision forecasting [2]. However, since the wind speed signals are always non-stationary and nonlinear, it is difficult to obtain the high-precision forecasting results. In recent years, to solve the wind speed accurate forecasting problem, many scientists proposed important forecasting methods [3]. Based on the previous forecasting results [3], the following thoughts can be seen as: **(a)** the hybrid forecasting methods have better performance than the single ones. Most of the recently proposed wind speed forecasting approaches are based on various hybrid strategies; and **(b)** in the proposed hybrid forecasting methods, there are two kinds of improving strategies which have been generally used as follows: the first one does not focus on the improvement of the wind speed forecasting methods but adopting some signal processing algorithm to process the original wind speed data to decrease their non-stationary feature so that the same built forecasting models can have better results. In this strategy, the forecasting performance is promoted indirectly. It is proved that the wind speed decomposition is one of the most effective processing algorithms in the wind speed prediction. There are several popular signal decomposing algorithms which have been adopted for the wind speed forecasting, which consist of the WD (*Wavelet Decomposition*), the WPD (*Wavelet Packet Decomposition*)

*Abbreviations:* WD, Wave let Decomposition; WPD, Wavelet Packet Decomposition; EMD, Empirical Mode Decomposition; EEMD, Ensemble Empirical Mode Decomposition; FEEMD, Fast Ensemble Empirical Mode Decomposition; MLP, Multi Layer Perceptron; ANFIS, Adaptive Network-based Fuzzy Inference System; BP, Back Propagation; GA, Genetic Algorithm; PSO, Particle Swarm Optimization; EA, Evolutionary Algorithm; AA, Adaboost Algorithm; ARIMA, Autoregressive Integrated Moving Average; SLFNN, Single-hidden Layer Feed-forward Neural Networks; MAE, Mean Absolute Error; MAPE, Mean Absolute Percentage Error; RMSE, Root Mean Square Error.

* Corresponding author at: Key Laboratory of Traffic Safety on Track of Ministry of Education, School of Traffic and Transportation Engineering, Central South University, Changsha 410075, Hunan, China. Tel.: +86 731 82655294; fax: +86 731 82656374.

*E-mail address:* csuliuhui@csu.edu.cn (H. Liu).

http://dx.doi.org/10.1016/j.enconman.2015.04.057
0196-8904/© 2015 Elsevier Ltd. All rights reserved.

and the EMD (*Empirical Mode Decomposition*). The EMD algorithm has two improved versions, the EEMD (*Ensemble Empirical Mode Decomposition*) algorithm and the FEEMD (*Fast Ensemble Empirical Mode Decomposition*) algorithm. Such as, the WD algorithm processed the raw wind speed data for the Gaussian forecasting models [4], the WPD algorithm provided the multi-scale processed wind speed sub-layers for the MLP (*Multi Layer Perceptron*) neural networks [5], the EMD algorithm decomposed the original wind speed data for the MLP neural networks [6], the EEMD algorithm converted the raw wind speed data into a series of sub-layers for the hybrid GA (*Genetic Algorithm*)-BP (*Back Propagation*) forecasting method [7] and the FEEMD built a hybrid forecasting framework combining with the MLP and ANFIS (*Adaptive Network-based Fuzzy Inference System*) neural networks [8]. The second one utilizes some mathematic or intelligent optimizing models to promote the capacity of the built wind speed forecasting methods. Different to the first mode, the second mode uses the direct optimization strategy. In the second strategy, the mainstream optimizing algorithms include the GA, the PSO (*Particle Swarm Optimization*), the EA (*Evolutionary Algorithm*), the AA (*Adaboost Algorithm*), etc. For example, the GA algorithm optimized the hybrid Wavelet-MLP wind speed forecasting model [9], the PSO improved the forecasting performance of the MLP neural networks [10], the EA solved the problem of the wind speed fast reconstruction [11] and the AA algorithm promoted the forecasting performance of the MLP neural networks [12]. The hybrid forecasting methods proposed in this paper are belonged to the first hybrid strategy.

In this study, a hybrid wind speed forecasting framework is proposed. In the framework, four decomposing algorithms (*i.e., WD, WPD, EMD and FEEMD*) are used to realize the non-stationary wind speed decomposition and the EML (*Extreme Learning Machines*) algorithm is employed to forecast the decomposed wind speed data. The aim of the study is to investigate the optimization of the ELM algorithm by the adopted four decomposing strategies for the multi-step accurate wind speed forecasting. The innovations of the study can be explained as follows: **(a)** although the ELM algorithm has been applied in the wind speed single-step predictions [13], the multi-step performance of the ELM algorithm in the wind speed predictions have not been studied. To really protect the wind power, only the wind speed single-step forecasting results are insufficient and the wind speed multi-step forecasting results are definitely expected; **(b)** it is the first time to combine the ELM algorithm and the four most popular decomposing algorithms in the same framework to fully investigate their real forecasting performance. The performance of the multi-step forecasting is focused in the investigation; and **(c)** to validate the effectiveness of the proposed hybrid forecasting framework, a number of comparing experiments are provided, including the hybrid WD-ELM model, the hybrid WPD-ELM model, the hybrid EMD-ELM model and the hybrid FEEMD-ELM model. At the same time, the single ELM model, the single MLP model and the single ARIMA (*Autoregressive Integrated Moving Average*) model are also included in the performance comparison. Based on the comparing results, the detailed promoting percentages of the ELM model by different signal decomposing algorithms can be obtained.

This paper is organized as follows: Section 2 presents the framework of the hybrid forecasting study in this paper; Section 3 explains the wind speed decomposition adopted in the proposed hybrid framework; Section 4 gives the wind speed forecasting steps done by the ELM algorithm; Section 5 provides a wind speed forecasting experiment; and Section 6 concludes the important results in this study.

## 2. Hybrid forecasting strategy

The framework of the proposed hybrid framework in this study is given in Fig. 1.

From Fig. 1, the proposed study can be explained in details as follows:

- Using the decomposing methods including the WD algorithm, the WPD algorithm, the EMD algorithm and the FEEMD algorithm to decompose the original wind speed time series. Among these four adopted algorithms, the FEEMD algorithm is latest, which has just been presented in 2014 [14].
- Building the ELM models to forecast the decomposed wind speed sub-layers to realize the multi-step predictions.
- Summarizing the ELM predicted results of the sub-layers to get the final results for the original wind speed data. The one-step, two step and three-step predicted results will be provided in this paper.
- Making comparisons of the forecasting results from the various models to find the best one. The comparing algorithms consist of the hybrid WD-ELM, the hybrid WPD-ELM, the hybrid EMD-ELM, the hybrid FEEMD-ELM, the single ELM, the single MLP and the single ARIMA.

## 3. Wind speed decomposition

As shown in Fig. 1, in the study four most popular signal decomposing algorithms (*e.g., the WD, the WPD, the EMD and the FEEMD*) are adopted to convert the non-stationary wind speed data into a group of relatively stationary wind speed sub-layers. The aim of the decomposition is to decrease the non-stationary of the original wind speed so that the difficulty of the high-precision wind speed predictions realized by the ELM algorithm will be reduced.

### 3.1. Wavelet decomposition

The WD decomposition for an signal $f(t)$ with respect to a mother wavelet function $\psi(t)$ is given as [15]:

$$WD_f(a,b) = \langle f(t), \psi_{a,b}(t) \rangle = \int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{|a|}}\psi^*\left(\frac{t-b}{a}\right)dt \tag{1}$$

where '$*$' denotes the complex conjugate, '$a$' is a scale coefficient and '$b$' is a translation coefficient.

The Eq. (1) is also used by the WPD algorithm to complete the decomposing computation. Actually, the WPD decomposition is an modified version of the standard WD decomposition. The main difference between them can be explained as: in the WD algorithm, only the detailed components will be decomposed. However, in the WPD algorithm, besides the detailed components, the approximate components will also be processed. So for a section of non-stationary data, the WPD can always have better decomposing results than the WD. The detailed computational steps of the WD algorithm and the WPD algorithm can be found in reference [15] and reference [16], respectively.

### 3.2. EMD decomposition

The EMD algorithm considers that any section of original signal series $X(t)$ can be described using the equation as [17]:

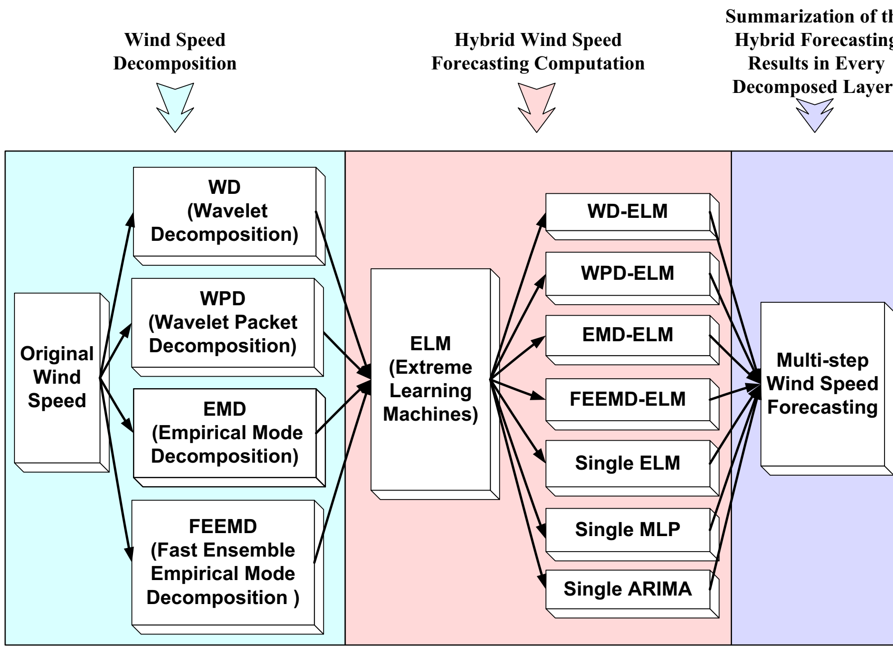


**Fig. 1.** Framework of the proposed hybrid wind speed forecasting framework.

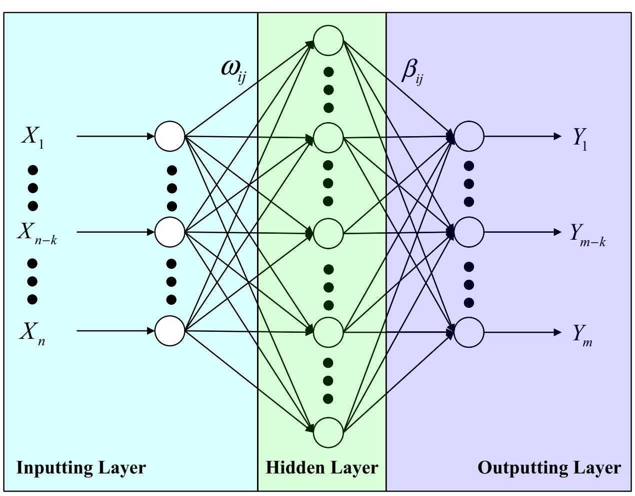


**Fig. 2.** Structure of the ELM network.

$$X(t) = \sum_{i=1}^{n} C_i(t) + R_n(t) \tag{2}$$

where series $\{C_i(t)\}$ are the IMF (*Intrinsic Mode Functions*) and $\{R_n(t)\}$ is the residue. The FEEMD algorithm is an improved format of the traditional EMD algorithm. In the FEEMD algorithm, the mode-mixing problem [18] and the real-time performance [14] of the traditional EMD algorithm have been studied and improved. Similar to the WD and WPD decomposition, the purpose of the EMD and FEEMD decomposition is also to decrease the instability of the raw wind speed data for the accurate forecasting results. The detailed contents of the FEEMD algorithm can be found in reference [14].

## 4. Extremely learning machine based wind speed prediction

The ELM is a new and fast learning algorithm based on the modification of the traditional SLFNN (*Single-hidden Layer Feed-forward Neural Networks*), recently proposed in reference [19]. The significant advantage of the ELM algorithm is that it distributes the weights and thresholds between the inputting layer and the hidden layer in random and does not need to adjust these random parameters during the whole learning process so that it can complete the training process extremely fast [20]. Besides the learning velocity, it has been proved that the ELM network has better accuracy performance than the other SLFNN neural networks and the support vector machines [21]. Due to the satisfactory performance in both of the learning speed and the learning accuracy, the ELM is selected as the basic wind speed predictor in this study.

The structure of a standard ELM network is demonstrated in Fig. 2.

As shown in Fig. 2, the main parameters of the ELM network are explained as follows [22]:

$$\boldsymbol{\omega} = \begin{bmatrix} \omega_{11} & \omega_{12} & \omega_{13} & \cdots & \omega_{1n} \\ \omega_{21} & \omega_{22} & \omega_{23} & \cdots & \omega_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ \omega_{l1} & \omega_{l2} & \omega_{l3} & \cdots & \omega_{ln} \end{bmatrix}_{l\times n} \tag{3}$$

where '$\boldsymbol{\omega}$' is the network weights between the inputting layer and the hidden layer, '$l$' is the number of the hidden neurons and '$n$' is the number of the inputting neurons.

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} & \cdots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \beta_{23} & \cdots & \beta_{2m} \\ \vdots & \vdots & \vdots & & \vdots \\ \beta_{l1} & \beta_{l2} & \beta_{l3} & \cdots & \beta_{lm} \end{bmatrix}_{l\times m} \tag{4}$$

where '$\boldsymbol{\beta}$' is the network weights between the hidden layer and the outputting layer and '$m$' is the number of the outputting neurons.

$$\boldsymbol{b} = [b_1, b_2, \cdots, b_l]^{-1}_{l\times 1} \tag{5}$$

where '$\boldsymbol{b}$' is the thresholds of the hidden layer.

Suppose the following data matrix will be used to train the ELM network:

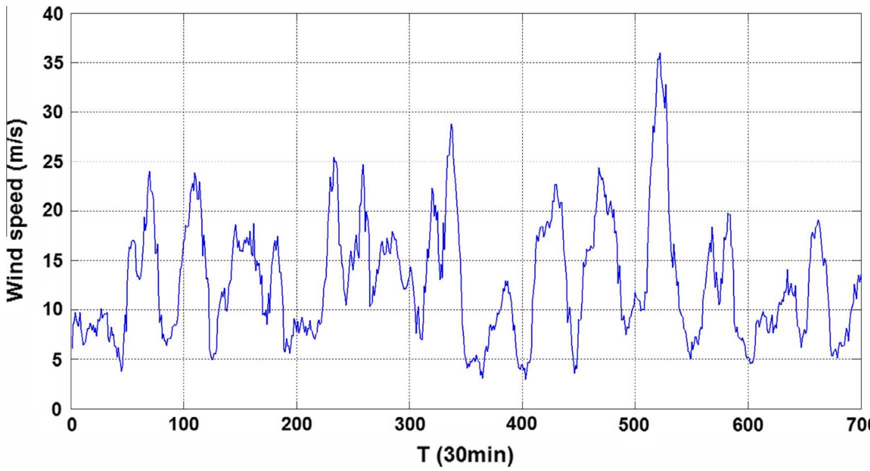


**Fig. 3.** Original wind speed time series $\{X_{1t}\}$.

$$\boldsymbol{X} = \begin{bmatrix} X_{11} & X_{12} & X_{13} & \cdots & X_{1P} \\ X_{21} & X_{22} & X_{23} & \cdots & X_{2P} \\ \vdots & \vdots & \vdots & & \vdots \\ X_{n1} & X_{n2} & X_{n3} & \cdots & X_{nP} \end{bmatrix}_{n\times P} \tag{6}$$

Then based on Eqs. (3)–(6), the real outputting matrix of the ELM network can be defined as below:

$$\boldsymbol{T} = [\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3, \cdots, \boldsymbol{t}_{P-1}, \boldsymbol{t}_P]_{m\times P} \tag{7}$$

where:

$$\boldsymbol{t}_j = \begin{bmatrix} t_{1j} \\ t_{2j} \\ \vdots \\ t_{mj} \end{bmatrix}_{m\times 1} = \begin{bmatrix} \sum_{i=1}^{l} \beta_{i1} g(\boldsymbol{\omega}_i \boldsymbol{X}_j + b_i) \\ \sum_{i=1}^{l} \beta_{i2} g(\boldsymbol{\omega}_i \boldsymbol{X}_j + b_i) \\ \vdots \\ \sum_{i=1}^{l} \beta_{im} g(\boldsymbol{\omega}_i \boldsymbol{X}_j + b_i) \end{bmatrix}_{m\times 1}, \quad (j = 1, 2, 3, \cdots, P) \tag{8}$$

where:

$$\boldsymbol{\omega}_i = [\omega_{i1}, \omega_{i2}, \omega_{i3}, \cdots, \omega_{in-1}, \omega_{in}] \tag{9}$$

$$\boldsymbol{X}_j = [X_{1j}, X_{2j}, X_{3j}, \cdots, X_{n-1j}, X_{nj}]^T \tag{10}$$

The Eqs. (7)–(10) can be reorganized as:

$$\hat{\boldsymbol{\beta}} = \boldsymbol{H}^{+} \boldsymbol{T}^{T} \tag{11}$$

where the '$\boldsymbol{H}^{+}$' is the Moore–Penrose generalized inverse of the matrix '$\boldsymbol{H}$'. The definition of the '$\boldsymbol{H}$' is given as follows:

$$\boldsymbol{H}(\boldsymbol{\omega}_1, \boldsymbol{\omega}_1, \cdots, \boldsymbol{\omega}_l, b_1, b_2, \cdots, b_l, \boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_P) = \begin{bmatrix} g(\boldsymbol{\omega}_1 \cdot \boldsymbol{X}_1 + b_1) & g(\boldsymbol{\omega}_2 \cdot \boldsymbol{X}_1 + b_2) & \cdots & g(\boldsymbol{\omega}_1 \cdot \boldsymbol{X}_1 + b_l) \\ g(\boldsymbol{\omega}_1 \cdot \boldsymbol{X}_2 + b_1) & g(\boldsymbol{\omega}_2 \cdot \boldsymbol{X}_2 + b_2) & \cdots & g(\boldsymbol{\omega}_1 \cdot \boldsymbol{X}_2 + b_l) \\ \vdots & \vdots & \vdots & \vdots \\ g(\boldsymbol{\omega}_1 \cdot \boldsymbol{X}_P + b_1) & g(\boldsymbol{\omega}_2 \cdot \boldsymbol{X}_P + b_2) & \cdots & g(\boldsymbol{\omega}_1 \cdot \boldsymbol{X}_P + b_l) \end{bmatrix}_{P\times l} \tag{12}$$

where the '$g$' is the activation function of the hidden layer in the ELM network. Based on Matlab platform, the upper computation can be programmed conveniently. In this study, after a number of trial experiments, the numbers of the neurons in the inputting layer, the hidden layer and the outputting layer are selected as six, twenty and one.

## 5. Forecasting experiment

### 5.1. Accuracy estimating indexes

To evaluate the performance of all the involved forecasting models, three error indexes are performed, including the MAE (*Mean Absolute Error*), the MAPE (*Mean Absolute Percentage Error*) and the RMSE (*Root Mean Square Error*). The detailed equations of these three error indexes are given as:

- *MAE*:

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left|X(t) - \hat{X}(t)\right| \tag{13}$$

- *MAPE*:

$$MAPE = \frac{1}{M}\sum_{t=1}^{N}\left|\frac{X(t) - \hat{X}(t)}{\hat{X}(t)}\right| \tag{14}$$

- *RMSE*:

$$RMSE = \sqrt{\frac{1}{N-1}\sum_{t=1}^{N}[X(t) - \hat{X}(t)]^2} \tag{15}$$

where $\{X(t)\}$ is the measured wind speed time series, $\{\hat{X}(t)\}$ is the forecasted wind speed time series and '$N$' is the number of the $\{X(t)\}$ time series.

Additionally, to obtain the detailed promoting percentages when comparing two forecasting models, three percentage error indexes are also defined as follows:

- $\xi_{MAE}$:

$$\xi_{MAE} = \left|\frac{MAE_1 - MAE_2}{MAE_1}\right| \tag{16}$$

- $\xi_{MAPE}$:

$$\xi_{MAPE} = \left|\frac{MAPE_1 - MAPE_2}{MAPE_1}\right| \tag{17}$$

- $\xi_{RMSE}$:

$$\xi_{RMSE} = \left|\frac{RMSE_1 - RMSE_2}{RMSE_1}\right| \tag{18}$$

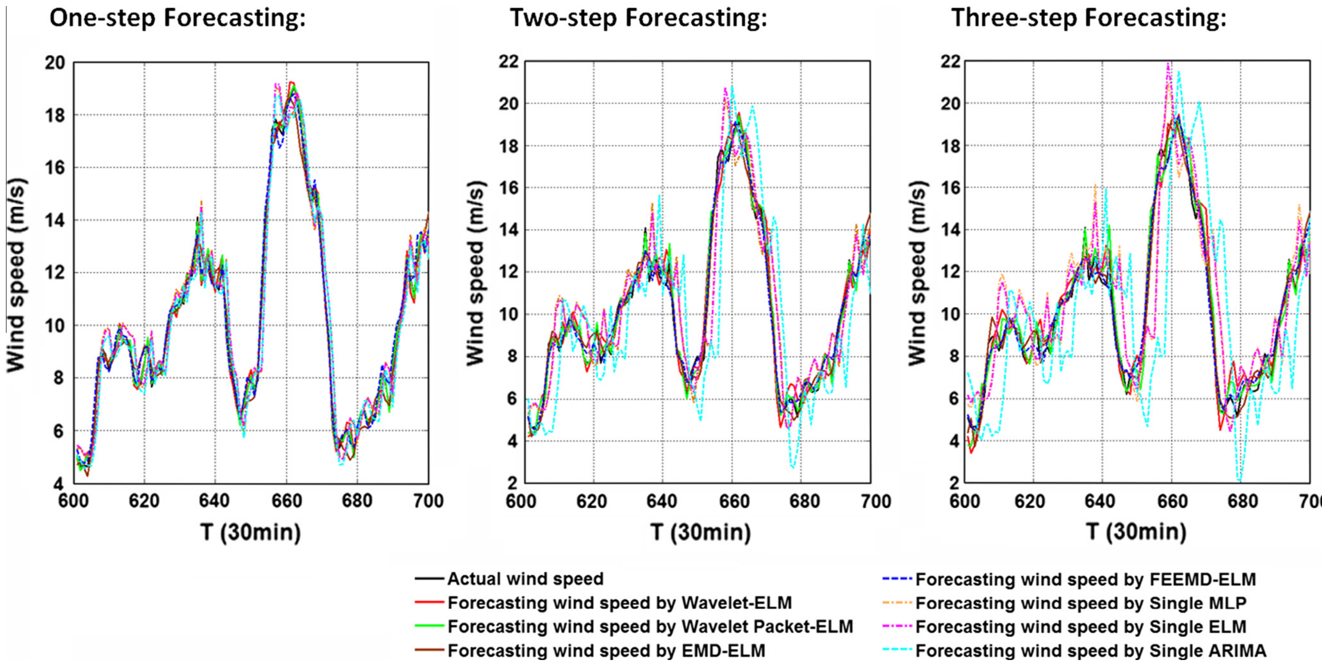


**Fig. 4.** The results of the multi-step predictions for the original wind speed series $\{X_{1t}\}$.

**Table 1**
Analysis of the forecasting results.

| Indexes | 1-step | 2-step | 3-step | 1-step | 2-step | 3-step |
|---|---|---|---|---|---|---|
| | Hybrid WD-ELM Model | | | Hybrid WPD-ELM Model | | |
| MAE(m/s) | 0.2076 | 0.4643 | 0.5991 | 0.1152 | 0.2803 | 0.4569 |
| MAPE(%) | 2.21 | 5.12 | 6.78 | 1.29 | 3.18 | 5.14 |
| RMSE(m/s) | 0.2674 | 0.6140 | 0.7848 | 0.1416 | 0.3543 | 0.5967 |
| | Hybrid EMD-ELM Model | | | Hybrid FEEMD-ELM Model | | |
| MAE(m/s) | 0.4189 | 0.4640 | 0.5453 | 0.2270 | 0.3706 | 0.4425 |
| MAPE(%) | 4.46 | 4.91 | 6.03 | 2.51 | 3.92 | 4.74 |
| RMSE(m/s) | 0.4994 | 0.5768 | 0.6585 | 0.2952 | 0.4777 | 0.5552 |
| | Single ELM Model | | | Single MLP Model | | |
| MAE(m/s) | 0.7955 | 1.2422 | 1.5128 | 0.8033 | 1.2798 | 1.5988 |
| MAPE(%) | 8.58 | 13.56 | 16.81 | 8.66 | 13.95 | 17.65 |
| RMSE(m/s) | 1.0157 | 1.5605 | 1.9249 | 1.0335 | 1.6047 | 1.9918 |
| | Single ARIMA Model | | | | | |
| MAE(m/s) | 0.7987 | 1.9140 | 2.6748 | | | |
| MAPE(%) | 8.60 | 20.59 | 29.27 | | | |
| RMSE(m/s) | 1.0257 | 2.6253 | 3.5091 | | | |

**Table 2**
Promoted percentages of the single MLP model compared to the single ELM model.

| Indexes | Single ELM Model vs. Single MLP Model | | |
|---|---|---|---|
| | 1-step | 2-step | 3-step |
| $\xi_{MAE}$(%) | 0.97 | 2.94 | 5.38 |
| $\xi_{MAPE}$(%) | 0.92 | 2.80 | 4.76 |
| $\xi_{RMSE}$(%) | 1.72 | 2.75 | 3.36 |

where '$MAE_1$', '$MAPE_1$' and '$RMSE_1$' are the calculated error indexes of the first wind speed forecasting model, and the '$MAE_2$', '$MAPE_2$' and '$RMSE_2$' are the corresponding error indexes of the second wind speed forecasting model.

### 5.2. Computational results

A group of simulative jumping wind speed data is employed to validate the performance of the wind speed multi-step forecasting in the study, as illustrated in Fig. 3. The 1st–600th samplings will be used to build various forecasting models and the 601st–700th ones will be adopted to examine the built models.

**Table 3**
Promoted percentages of the single ARIMA model compared to the single ELM model.

| Indexes | Single ELM Model vs. Single ARIMA Model | | |
|---|---|---|---|
| | 1-step | 2-step | 3-step |
| $\xi_{MAE}$(%) | 0.40 | 35.10 | 43.44 |
| $\xi_{MAPE}$(%) | 0.23 | 34.14 | 42.57 |
| $\xi_{RMSE}$(%) | 0.97 | 40.56 | 45.15 |

Fig. 4 shows the forecasted results at 601st–700th samplings of the original wind speed series $\{X_{1t}\}$. The estimated error results of these predictions are given in Table 1. Based on the results in Table 1, the promoted percentages of the single MLP model and the single ARIMA model by the single ELM model can be calculated as indicated in Tables 2 and 3, respectively, and the promoted percentages of the ELM model by the four decomposing algorithm in the respective combinations can also be obtained as given in Table 4.

From Table 1, it can be analyzed that: (a) when comparing the single ELM model with the single MLP model and the single ARIMA model, the former one model has better forecasting performance than the latter two models. For example, the MAPE of the single ELM model from one-step to three-step is 8.58%, 13.56% and 16.81%, respectively. The MAPE of the single MLP model from one-step to three-step is 8.66%, 13.95% and 17.65%, respectively. The MAPE of the single ARIMA model from one-step to three-step is 8.60%, 20.59% and 29.27%, respectively. There is an important phenomenon that in the one-step results the intelligent models (i.e., the ELM and the MLP) and the statistical model (i.e., the ARIMA) have close forecasting performance but in the two and three step results the intelligent models have much better forecasting performance than the statistical ARIMA model; (b) when comparing the hybrid WD/WPD/EMD/FEEMD-ELM models with the single ELM model, all the proposed hybrid models have much better forecasting performance than the single ELM model. This

**Table 4**
Promoted percentages of the single ELM model by the different adopted decomposing algorithms.

| Indexes | WD (Wavelet Decomposition) | | | WPD (Wavelet Packet Decomposition) | | |
|---|---|---|---|---|---|---|
| | 1-step | 2-step | 3-step | 1-step | 2-step | 3-step |
| $\xi_{MAE}(\%)$ | 73.90 | 62.62 | 60.40 | 85.52 | 77.44 | 69.80 |
| $\xi_{MAPE}(\%)$ | 74.24 | 62.24 | 59.67 | 84.97 | 76.55 | 69.42 |
| $\xi_{RMSE}(\%)$ | 73.67 | 60.65 | 59.23 | 86.06 | 77.30 | 69.00 |
| | EMD (Empirical Mode Decomposition) | | | FEEMD (Fast Ensemble Empirical Mode Decomposition) | | |
| $\xi_{MAE}(\%)$ | 47.34 | 62.65 | 63.95 | 71.46 | 70.17 | 70.75 |
| $\xi_{MAPE}(\%)$ | 48.02 | 63.79 | 64.13 | 70.75 | 71.09 | 71.80 |
| $\xi_{RMSE}(\%)$ | 50.83 | 63.04 | 65.79 | 70.94 | 69.39 | 71.16 |

proves that it is feasible to adopt the wind speed decomposition to improve the forecasting capacity of the ELM model; (c) among all the proposed hybrid models, the hybrid WPD-ELM model has the best performance in the one and two step predictions while the hybrid FEEMD-ELM model has the best performance in the three-step predictions; and (d) all the proposed hybrid models have satisfactory forecasting performance.

From Tables 2 and 3, it can be found that: the single ELM model improves both the performance of the single MLP model and the single ARIMA model. For example, the MAPE promoted percentages of the single MLP model by the single ELM model from one-step to three-step is 0.92%, 2.80% and 4.76%, respectively. The MAPE promoted percentages of the single ARIMA model by the single ELM model from one-step to three-step is 0.23%, 34.14% and 42.57%, respectively.

From Table 4, it can be seen that: (a) all the adopted decomposing algorithms have improved the forecasting performance of the ELM model considerably; (b) in the one-step predictions, the MAPE promoted percentages of the ELM model by the WD, the WPD, the EMD and the FEEMD is 74.24%, 84.97%, 48.02% and 70.75%, respectively. The WPD has the best promoting performance; (b) in the two-step predictions, the MAPE promoted percentages of the ELM model by the WD, the WPD, the EMD and the FEEMD is 62.24%, 76.55%, 63.79% and 71.09%, respectively. The WPD retains its best position; and (c) in the three-step predictions, the MAPE promoted percentages of the ELM model by the WD, the WPD, the EMD and the FEEMD is 59.67%, 69.42%, 64.13% and 71.80%, respectively. The FEEMD replaces the WPD to be the best one.

## 6. Conclusions

Based on the results of the experiment in this study, it can be concluded as follows: **(a)** the single ELM algorithm has better forecasting performance than the classical MLP and ARIMA algorithms. The reason of the phenomenon is that the ELM algorithm has better generalization capacity than the MLP and ARIMA algorithms; **(b)** by the utilization of the decomposing algorithms [WD/WPD/EMD/FEEMD], the proposed hybrid models [WD-ELM/WPD-ELM/EMD-ELM/FEEMD-ELM] can promote the forecasting performance of the single ELM considerably. The reason of this situation is that the adopted various wind speed decomposing algorithms decrease the instability of the raw wind speed data successfully so that the built ELM model can realize the high-precision forecasting computation; **(c)** in all the proposed hybrid models, the hybrid WPD-ELM model has the best forecasting performance in the wind speed one and two step predictions while the hybrid FEEMD-ELM model has the best forecasting performance in the wind speed three-step predictions. The reason of the results is that: compared to the WD and EMD algorithms, the WPD and FEEMD algorithms decompose the raw wind speed data deeper; and **(d)** by the use of the related toolboxes in Matlab platform, the proposed hybrid forecasting framework can be programmed and validated conveniently.

The limitations of the proposed hybrid forecasting methods are also provided as follows: **(a)** since in the study the ELM neural network is selected and built as the basic forecasting model to calculate the final results, the ELM algorithm inevitably needs to spend effort and time on the network training to reach its satisfactory fitting performance. The number of the wind speed sample needed for the ELM network training/building is decided by the expected forecasting accuracy. The higher expected forecasting accuracy, the more wind speed data requested for the network training. Due to this limitation, the proposed hybrid forecasting methods sometimes cannot realize the real-time forecasting generation; and **(b)** as one of the artificial neural networks, sometimes the ELM algorithm also has the convergence problem. Although some evolutionary algorithms might be useful to optimize the convergence performance of the ELM neural network, there are still some special cases in which the convergence problem cannot be solved.

## Acknowledgements

The authors would like to thank the referees for their precious reviewing. This study is fully supported by the National Natural Science Foundation of China (Grant No. 51308553), the Scientific Research Fund of Hunan Provincial Education Department (Principle Investigator: Dr. Hui Liu) and the Shenghua Yu-ying Talents Program of the Central South University (Principle Investigator: Dr. Hui Liu). This study is partly supported by the National Natural Science Foundation of China (Grant No. U1134203, U1334205).

## References

[1] Cardell JB, Anderson CL. Targeting existing power plants: EPA emission reduction with wind and demand response. Energy Policy 2015;80:11–23.
[2] Liu H, Tian H, Li Y. Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction. Appl Energy 2012;98:415–24.
[3] Tascikaraoglu A, Uzunoglu M. A review of combined approaches for prediction of short-term wind speed and power. Renew Sustain Energy Rev 2014;34:243–54.
[4] Chen N, Qian Z, Meng X. Multistep wind speed forecasting based on wavelet and gaussian processes. Math Probl Eng 2013;2013.
[5] Liu H, Tian H-Q, Pan D-F, Li Y-F. Forecasting models for wind speed using wavelet, wavelet packet, time series and Artificial Neural Networks. Appl Energy 2013;107:191–208.
[6] Liu H, Chen C, Tian H-Q, Li Y-F. A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks. Renew Energy 2012;48:545–56.
[7] Wang Y, Wang S, Zhang N. A novel wind speed forecasting method based on ensemble empirical mode decomposition and GA-BP neural network. presented at the IEEE power and energy society general meeting; 2013.
[8] Liu H, Tian H-Q, Li Y-F. Comparison of new hybrid FEEMD-MLP, FEEMD-ANFIS, Wavelet Packet-MLP and Wavelet Packet-ANFIS for wind speed predictions. Energy Convers Manage 2014;89:1–11.
[9] Liu H, Tian H-Q, Chen C, Li Y-F. An experimental investigation of two Wavelet-MLP hybrid frameworks for wind speed prediction using GA and PSO optimization. Int J Electr Power Energy Syst 2013;52(1):161–73.
[10] Poitras G, Cormier G. Wind speed prediction for a target station using neural networks and particle swarm optimization. Wind Eng 2011;35(3):369–80.
[11] Carro-Calvo L, Salcedo-Sanz S, Prieto L, Kirchner-Bossi N, Portilla-Figueras A, Jiménez-Fernández S. Wind speed reconstruction from synoptic pressure patterns using an evolutionary algorithm. Appl Energy 2012;89(1):347–54.
[12] Liu H, Tian H-Q, Li Y-F, Zhang L. Comparison of four Adaboost algorithm based artificial neural networks in wind speed predictions. Energy Convers Manage 2015;92:67–81.
[13] Salcedo-Sanz S, Pastor-Sánchez A, Prieto L, Blanco-Aguilera A, García-Herrera R. Feature selection in wind speed prediction systems based on a hybrid coral reefs optimization – Extreme learning machine approach. Energy Convers Manage 2014;87:10–8.
[14] Wang Y-H, Yeh C-H, Young H-WV, Hu K, Lo M-T. On the computational complexity of the empirical mode decomposition algorithm. Phys Stat Mech Appl 2014;400:159–67.

[15] Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. Pattern Anal Mach Intell IEEE Trans 1989;11(7):674–93.
[16] Laine A, Fan J. Texture classification by wavelet packet signatures. Pattern Anal Mach Intell IEEE Trans 1993;15(11):1186–91.
[17] Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc Lond Ser Math Phys Eng Sci 1998;454(1971):903–95.
[18] Wu Z, Huang NE. Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv Adapt Data Anal 2009;1(1):1–41.
[19] Huang G-B, Zhu Q-Y, Siew C-K. Extreme learning machine: theory and applications. Neurocomputing 2006;70(1–3):489–501.
[20] Lendasse A, He Q, Miche Y, Huang G-B. Advances in extreme learning machines (ELM2013). Neurocomputing 2015;149(Part A):158–9.
[21] Huang G-B, Zhou H, Ding X, Zhang R. Extreme learning machine for regression and multiclass classification. IEEE Trans Syst Man Cybern Part B Cybern 2012;42(2):513–29.
[22] Wang X, Shi F, Yu L, Li Y. Forty-three neural network case analysis in matlab. China (Beijing): Beihang University Press; 2013.