Energy Conversion and Management 147 (2017) 75–85

Contents lists available at ScienceDirect

Energy Conversion and Management

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

Comparative study on three new hybrid models using Elman Neural Network and Empirical Mode Decomposition based technologies improved by Singular Spectrum Analysis for hour-ahead wind speed forecasting

Chuanjin Yu, Yongle Li *, Mingjin Zhang

Department of Bridge Engineering, Southwest Jiaotong University, 610031 Chengdu, Sichuan, PR China

ARTICLE INFO

Article history:
Received 13 February 2017
Received in revised form 18 April 2017
Accepted 3 May 2017
Available online 26 May 2017

Keywords:
Wind speed forecast
Empirical Mode Decomposition (EMD)
Ensemble EMD (EEMD)
Complete EEMD with adaptive noise (CEEMDN)
Singular Spectrum Analysis (SSA)
Elman neural network (ENN)

ABSTRACT

High precision forecasting of wind speed is urgently needed for wind power utilization. In this paper, Empirical Mode Decomposition (EMD) based technologies, including EMD and its advanced versions ensemble EMD (EEMD) and complete EEMD with adaptive noise (CEEMDN) are applied for improving wind speed prediction accuracy. Three new hybrid models (EMD-SSA-ENN, EEMD-SSA-ENN and CEEMDN-SSA-ENN) are proposed in which EMD, EEMD and CEEMDN are combined with Singular Spectrum Analysis (SSA) and Elman neural network (ENN) respectively. SSA is exploited to re-handle components with the highest frequency disaggregated from the decomposition technologies, of which the procedure is systematically studied herein. The experimental prediction results show that: 1. through the retreatment of SSA, the performances of the new proposed hybrid models improve significantly; 2. compared to the persistence method, single ENN model, ARIMA, EMD-RARIMA and some methods in the references, all the proposed methods can give a much more accurate forecast; 3. among all the proposed methods, the performance of the hybrid model CEEMDN-SSA-ENN are the best.

1. Introduction

As one of the clean and renewable energies, wind power has becoming increasingly competitive and widely used around the world. With the wind nature characteristics of intermittence and fluctuation, to protect the safety of the power utilization and integration, high precision forecasting of wind speed is demanding and critical [1,2].

There are four main methods for wind speed forecasting: statistical methods, physical methods, intelligent methods and hybrid models. Since each single model has its own nature deficiencies (e.g.: statistical methods cannot handle non-linear problem well [3], some key parameters in intelligent methods are not easily chosen [4,5] and physical methods consume much computing resources [6]), hybrid models absorbing essences of single models are becoming the mainstream approaches [7]. Due to the inherent features of non-stationary, high fluctuations and irregularity, many hybrid models are in collaboration with decomposition technology, including the Wavelet Transform (WT) [8,9] and Empirical Mode Decomposition (EMD). They are aimed to disaggregate an original wind speed into a sequence of more stationary sub series before further forecasting to improve prediction precision.

Without too much priori knowledge, EMD is relatively easy to be understood and employed. Liu et al. [10] used EMD to decompose the original data into several Intrinsic Mode Functions (IMFs) and a residue, which were modeled and forecasted by Autoregressive Moving Average (ARMA) respectively and then summarized as the final prediction. Ren et al. [11,12] reported hybrid models employing the k Nearest Neighbor algorithm (KNN) and Support Vector Regression (SVR) as regression models collaborated with EMD respectively for wind speed forecasting. Liu [13] combined the artificial neural network (ANN) and EMD to make wind speed prediction. The training algorithms of ANN were discussed. He also came up a method combined EMD and the recursive autoregressive integrated moving average model (RARIMA) for wind speed forecasting [14]. Ren et al. [12] and Hong et al. [15] used the back propagation neural network (BPNN) and EMD to make one-hour ahead wind speed forecasting, while Wang et al. [16] utilized Elman Neural Network (ENN) with EMD. Sun and Liu [17] proposed a hybrid model in which the fast EEMD (FEEMD) and the regular-

* Corresponding author.
E-mail addresses: 7818534@qq.com (C.J. Yu), lele@swjtu.edu.cn (Y.L. Li), zhang108119@163.com (M.J. Zhang).

http://dx.doi.org/10.1016/j.enconman.2017.05.008
0196-8904/

ized extreme learning machine (RELM) were applied. Wang et al. [18] employed the EEMD and BPNN improved by the genetic algorithm technique (GA) to make short term wind speed forecasting. Fei [19] presented a hybrid model combined with EMD and the multiple-kernel relevance vector regression algorithm for wind speed prediction.

There is no further treatment on the components decomposed by the EMD in the above works. In order to reinforce the prediction accuracy, by defining the overall energy of a time series, Ghelardoni [20] classified the decomposition components by EMD into two sets, respectively describing the trend and the local oscillations, which were modeled and forecasted by SVR. There is usually little regularity in the decomposed component with the highest frequency band called IMF1 [21]. Liu [22] made re-decomposition on IMF1 by WT, then the least square support vector machine (LSSVM) models were built for all the new decomposition components and the others without further retreatment. Guo [21] discarded IMF1 and all the rest decomposed components were modeled by BPNN, whose performance was little worse than that of ENN as Ref. [16] indicated. The performance of all the three hybrid models were greater than those without deeply managing the decomposition components by the EMD.

Therefore, it is essential to make further process on the decomposition components by the EMD to enhance the hybrid prediction model. But about the clustering process of the decomposition components in Ref. [20], it may be subjective, since the energy threshold distinguishing the two kind series is not easily determined. As for the re-decomposing IMF1 by WT in Ref. [22], too many models are adopted with too much computation cost. Different from the method in Ref. [21], it is believed that there are still potential and useful information involved in IMF1. So IMF1 will be further processed by Singular Spectrum Analysis (SSA), which is introduced for the first time. Another is worth mention that EMD usually falls into in a trouble called “mode mixing” [23,24]. Hence, its advanced version the ensemble empirical mode decomposition (EEMD) and the complete ensemble empirical mode decomposition with adaptive noise (CEEMDN) are proposed to overcome the weakness. So three new hybrid models are proposed, including EMD-SSA-EEN, EEMD-SSA-EEN and CEEMDN-SSA-EEN. The original wind speed series is firstly decomposed into a sequence of sub series by EMD, EEMD and CEEMDN respectively, then the series IMF1 with the highest frequency band will be re-handled by SSA, finally the processed IMF1 and all the rest components will be modeled by ENN to get the final wind speed forecasts. The retreatment procedure of SSA on the decomposition component IMF1 will be systematically investigated. The performances of these three new hybrid models will be also talked in details.

The rest of paper is organized as follows: Sections 2 reviews methodology in detailed; Section 3 introduces the proposed method; Section 4 gives some predictions and makes comparative studies.

2. Methodology

2.1. EMD and its improved versions

2.1.1. EMD

EMD [25] is widely used technique dealing with non-linear and non-stationary time series. Through the EMD process, the original time series $X(t)$ can be disaggregated into n IMFs and a residue as follows:

$$X(t) = \sum_{i=1}^{n} c_i + r_n \tag{1}$$

2.1.2. EEMD

Since EMD is usually trapped in a trouble called “mode mixing”, which means that there are oscillations of very disparate amplitude in a mode or very similar oscillations in different modes [23,24], EEMD is proposed [23]. With the help of addition of white Gaussian white noise, the mode mixing problem can be alleviated. The key process of EEMD can be described as follows.

Step a. For a signal $X(t)$, create a new noise-added signal : $x^i(t) = X(t) + \varepsilon^i(t)$

where $\varepsilon^i(t)$ is a Gaussian white noise.

Step b. Decompose $x^i(t)$ into several IMFs and a residue by EMD. That is designated as:

$$x^i(t) = \sum_{j=1}^{n} c_j^i + r_n^i \tag{2}$$

Step c. Repeat step a and step b with different Gaussian white noise each time.
Step d. Average on all the corresponding IMFs as the final results.

2.1.3. CEEMDAN

In EEMD, each EMD decomposition process of a noise-added signal may produce different number of modes. Then CEEMDAN is proposed to round the trouble with a lower computational cost [24]. The calculation steps of CEEMDAN are as follows.

Step a. Generate a number of noise-added series: $x^i(t) = X(t) + \omega_0\varepsilon^i(t), i \in \{1, \ldots, I\}$, where $\varepsilon^i(t)$ are different Gaussian white noise with a unit and ω_0 is a noise coefficient.
Step b. Apply EMD to each $x^i(t)$ to extract the first decomposed IMF $c_1{}^i(t)$. Make average of the whole IMF1: $c_1(t) = \frac{1}{I}\left(\sum_{i=1}^{I} c_1^i(t)\right)$, to get the first residue: $r_1(t) = x(t) - c_1(t)$
Step c. In the second process, continue to disaggregate the noise-added residue $r_1(t) + \omega_1 E_1(\varepsilon^i(t))$ to obtain the second IMFs:

$$c_2(t) = \frac{1}{I}\sum_{j=1}^{I} E_1(r_1 + w_1 E_1(\varepsilon^i(t))) \tag{3}$$

where $E_1(.)$ is a function to extract the first IMF decomposed by EMD.

Step d. Repeat for another IMF until the residue is no longer feasible to be decomposed (it does not have at least two extrema).

2.2. SSA

SSA is a widely used approach for time series analysis, including trend of quasi-periodic component detection and extraction and denoising [26,27]. The core idea of SSA is to decompose a raw original time series into a sum of sub series, identified as either a trend, periodic or quasi-periodic component, or noise. Then it is followed by the reconstruction of the original series. Therefore, there are two stages decomposition and reconstruction in the process of SSA as follows:

Stage I. Decomposition

Step a. Embedding

For a signal $X(t) = (X_1, \ldots, X_N)$, generate $Y_i(t) = (X_i, \ldots, X_{i+k-1})$ for L dimensions to construct a trajectory matrix Y:

$$Y = \begin{bmatrix} X_1 & \cdots & X_k \\ \vdots & \ddots & \vdots \\ X_L & \cdots & X_N \end{bmatrix}$$

where $k = N - L + 1$ and Y is Hankel matrix with equal elements along the diagonals ($i + j$ = const)

Step b. Singular Value Decomposition (SVD)

Make SVD on the matrix XX^T to get its eigentriples (λ_i, U_i, V_i) in descending order by λ_i, where λ_i is the ith singular value, U_i and V_i is the ith left and right eigenvector respectively. The trajectory matrix Y can be rewritten as:

$$Y = Y_1 + \ldots + Y_d, \quad Y_i = \sqrt{\lambda_i} U_i V_i^T \tag{4}$$

where d = rank(Y).

It is worth mentioning that the values of $\sqrt{\lambda_i}/\sum_{i=1}^{d}\sqrt{\lambda_i}$ implies the contribution of Y_i to Y. Obviously, Y_1 always has the highest contribution, while Y_d has the lowest one [28].

Stage II. Reconstruction

Step c. Grouping

Select r out of d eigentriples. Denote $I = \{I1, \ldots, Ir\}$, which is a sequence of r selected eigentriples and designated as: $Y_I = Y_{I1} + \ldots + Y_{Ir}$. Y_I can stand for the original data Y, while the others (d-r) eigentriples are supposed as the noise term ϵ here.

Step d. Averaging

It is the Hankelization procedure [27] denoted as $H()$ that transforms the matrix group $\{Y_{I1}, \ldots, Y_{Ir}\}$ to a time series group $\{X_{I1}, \ldots, X_{Ir}\}$.Finally, the reconstructed time series can be described as follows:

$$X = H(Y_{I1}) + \ldots H(Y_{Ir}) + H(\epsilon) = X_{I1} + \ldots + X_{Ir} + H(\epsilon) \tag{5}$$

Denote $X_{trend} = X_{I1} + \ldots + X_{Ir}$ and $X_{noise} = H(\epsilon)$, which herein is called trend and noise series respectively. So X can be rewritten: $X = X_{trend} + X_{noise}$.

2.3. ENN

ENN is a simple recurrent neural network, which is applied in wide range of time series forecast tasks [16]. The typical topology structure of Elman neural network can be divided into four layers, named input layer, hidden layer, connecting layer and output layer, respectively, as Fig. 1 shows. Connecting layer is used to store the output of the former moment from the hidden layer unit and then transfer it in the next iteration. It makes ENN have the ability to adapt to time-varying characteristics.

The ENN structure, especially in terms of the number of neurons in each layers, is crucial to the prediction accuracy. Firstly, about the number of neurons in the input layer, is determined by the algorithm Gradient Boosted Regression Trees (GBRT) [30]. GBRT can be used to measure importance of input variables calculated by one's contribution to the prediction accuracy [31]. For each input variable, the respective importance, denoted as $RI = \{RI_1, \ldots, RI_m\}$ is firstly calculated. The variance m is the whole number of input variables investigated. Since the biggest RI values of input variables are really different for different time series, the threshold of cumulative importance (CI) is set to help determine the neuron number. The representation of CI_j is described as follows:

$$CI_j = \sum_{i=1}^{j} RI_j \quad (1 \leqslant j \leqslant m) \tag{6}$$

The threshold value of CI is set as 90% to avoid "dimensionality curse". Select j out of m input variables where CI_j is just greater than the threshold.

Then, as the Ref. [32] indicated, $2n + 1$ hidden neurons are sufficient to map any function for n inputs, which is also taken for the study. Finally, there is only one neuron in the output layer.

3. Proposed method

Due to the wind speed characteristics with non-stationary and strong fluctuations, EMD, EEMD and CEEMDN are widely utilized to decompose the original wind series into several more stationary components. Many studies shows that the first component IMF1 with the highest frequency could lead to a great disturbance on the prediction accuracy [13,21]. The reason is that IMF1 is the most disorder part of the wind speed series and has little regularity [21]. However, it is believed that there are still potential and useful information involved in IMF1. Therefore, in the proposed method, SSA is used to extract the trend of IMF1. Then the trend of IMF1 and the rest of the decomposed components are modeled by ENN models respectively. The framework of the proposed methods are described as Fig. 2 shown. Three new methods are proposed including EMD-SSA-ENN, EEMD-SSA-ENN and CEEMDN -SSA-ENN. And the major stages of the proposed methods are as follows.

Stage a. Decompose the original wind speed series into a sequence of IMFs and a residue by EMD, EEMD and CEEMDN respectively.
Stage b. Apply SSA on the IMF1 to get trend of it.
Stage c. Establish EEN models for each sub series, including the trend of IMF1 and the rest of the decomposed components to get each prediction.
Stage d. Make the summation of all the sub prediction from Stage c and get the final forecast.

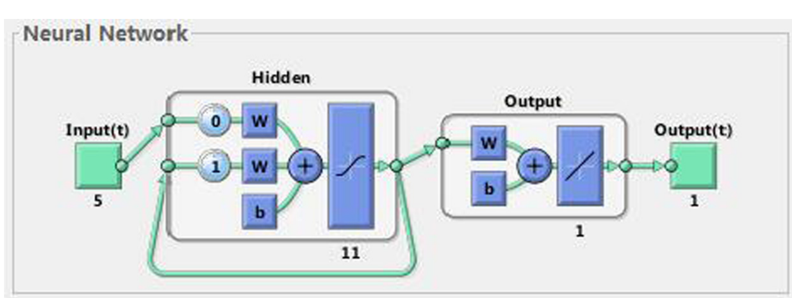


Fig. 1. A typical structure of Elman neural network used in the study in Matlab Neural Network Toolbox [29]

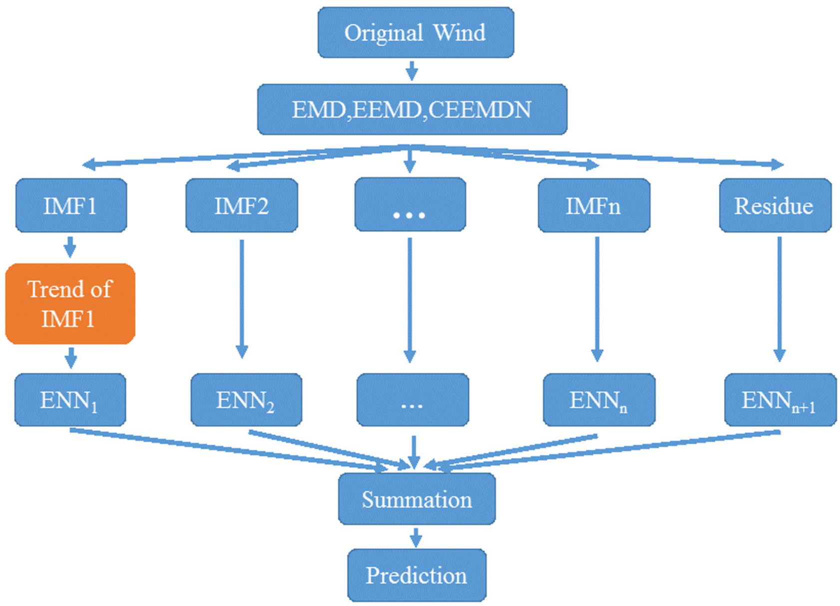


Fig. 2. Framework of the proposed method.

4. Experimental results and comparative analysis

4.1. Data

There are two real wind speed time series $\{X_{1t}\}$ and $\{X_{2t}\}$ employed to verify the performance of the proposed method as shown in Fig. 3, which are measured from Chinese Sichuan Province. Their statistical information is showed in Table 1. The wind speed series are 1 h-average over a month. There are 720 samples in $\{X_{1t}\}$ and 744 samples in $\{X_{2t}\}$.There are a total of 168 samples in the last days utilized to validate the built models while the others in the previous days are employed to build forecasting models.

4.2. Performance evaluation

In order to validate models, three error criteria are adopted as Table 2 shows. There are the Mean absolute error (MAE), the Mean Absolute Percentage Error (MAPE) and the Root Mean Square Error (RMSE),

where $X(t)$ is the actual wind speed data, $\hat{X}(t)$ is the wind speed prediction and N is the samples number of the $X(t)$ series.

4.3. Forecasting with hybrid approach

4.3.1. Decomposition by EMD and its improved versions

The $\{X_{1t}\}$ is decomposed by EMD, EEMD and CEEMDN respectively. It is noted that the Gaussian noise standard deviation for EEMD and CEEMDAN is 0.2 as the Ref. [12] indicates. A sequence of IMFs and a residue are obtained as Fig. 4 shows. There is an obvious "mixing mode" problem in the IMF3 decomposed by EMD, while it is alleviated in others decomposed by EEMD and CEEMDN.

4.3.2. Trend extraction of IMF1 component

In this stage, SSA is used to extract the trend of IMF1 component. From Section 2.2, it can be seen that there are two key parameters in the SSA process, the window length L and the number r of the selected eigentriples for reconstruction.

It is suggested that L should be large enough but not greater than N/2. However, if there is a periodic component with an integer period for a time series, it is advisable to take the window length proportional to that period [13]. From Fig. 3, there is an obvious change in one day cycle for the original wind speed. Since IMF1 has much more rapid changes, $L = 48$ is taken, such that it is sufficient to map the period of IMF1. To determine the number r, a general criterion is based on the contribution of each eigentriples, evaluated as $\sqrt{\lambda_i}/\sum_{i=1}^{d}\sqrt{\lambda_i}$. Define the trend rate (TR) as the integrated contributions of the r out of d components as follows:

$$TR = \sum_{i=1}^{r}\sqrt{\lambda_i}/\sum_{i=1}^{d}\sqrt{\lambda_i} \tag{7}$$

The TR threshold is set as 90% in the Ref. [27]. In the study, three threshold values, 85%, 90% and 95% are systematically talked about. What's more, 0% of TR represents that the IMF1 component is erased and 100% of TR stands for that there is no modification of

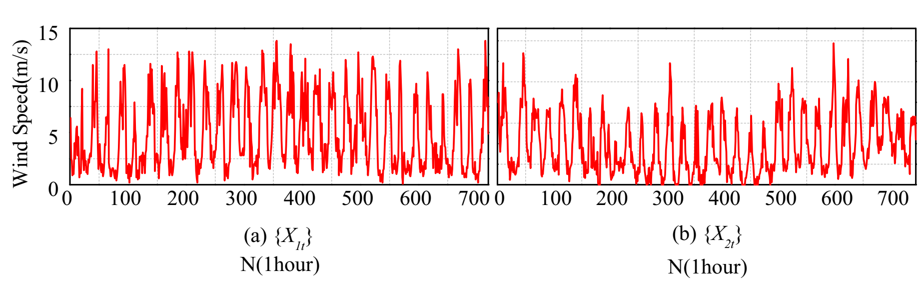


Fig. 3. Real wind speed time series measured from Chinese Sichuan Province.

Table 1
Statistical information of $\{X_{1t}\}$ and $\{X_{2t}\}$ (unit: m/s).

Parameter	Mean	Median	Maximum	Minimum	Standard deviation
$\{X_{1t}\}$	5.03	3.9	13.8	0.1	3.55
$\{X_{2t}\}$	4.92	3.8	17.2	0.0	3.58

Table 2
Error criteria adopted in the study.

Error criteria	Calculation
MAE	$\frac{1}{N}\sum_{t=1}^{N}\lvert X(t)-\hat{X}(t)\rvert$
MAPE	$\frac{1}{N}\sum_{t=1}^{N}\lvert\frac{X(t)-\hat{X}(t)}{X(t)}\rvert$
RMSE	$\sqrt{\frac{1}{N-1}\sum_{t=1}^{N}[X(t)-\hat{X}(t)]^2}$

the IMF1 component. For each IMF1, the extracting process with 90% *TR* is shown in Fig. 5. The red lines represent the noise term of IMF1.

4.3.3. Forecasting model design

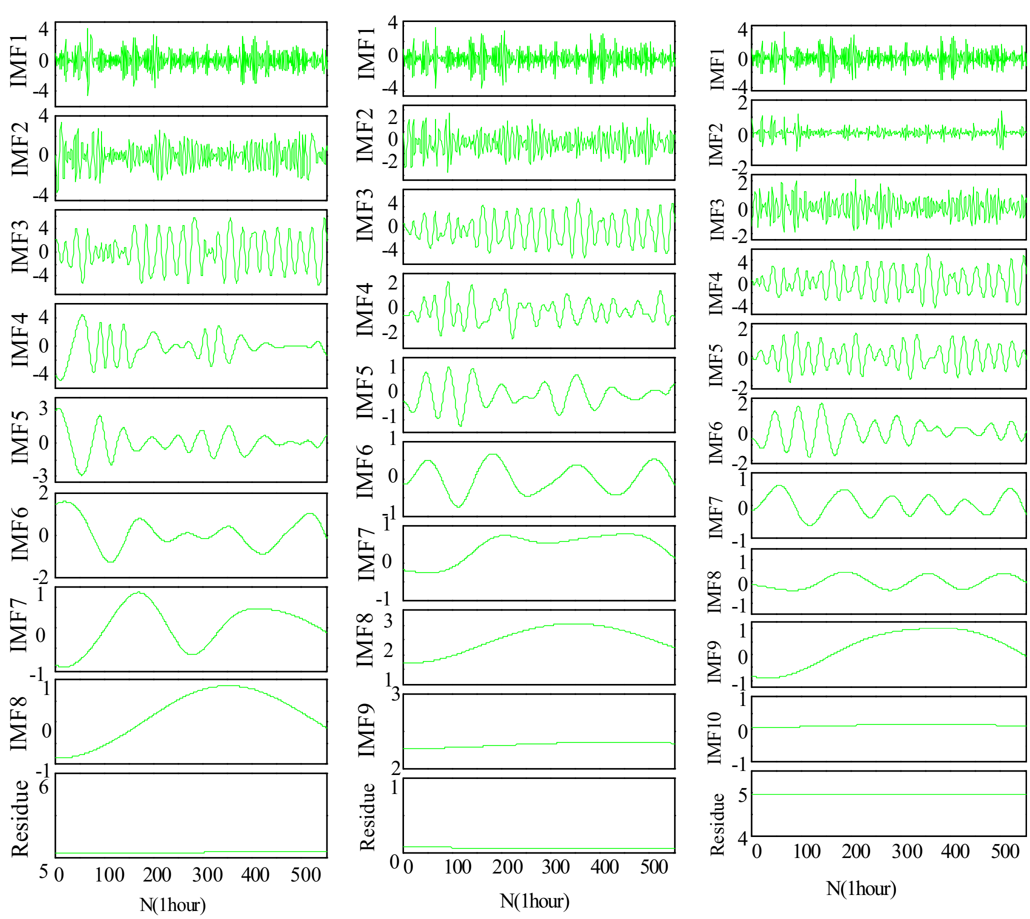
For the trend of the IMF1 and the rest components, their ENN models structure are determined by GBRT. To avoid too much input variables chosen, there are only the first ten input variables under investigation for their contribution to the prediction accuracy. The importance of input variables in the first two components decomposed by EMD of $\{X_{1t}\}$ are shown in Fig. 6. It is not doubt that the input variables with smaller lags has greater impact on the prediction. After calculation, there is little impact of SSA process on determining the ENN structure of IMF1. For each component decomposed by EMD, EEMD and CEEMDN, their neurons number in the input layers of ENN are shown in Table 3.

4.4. Forecasting results and analysis

4.4.1. Forecasting results

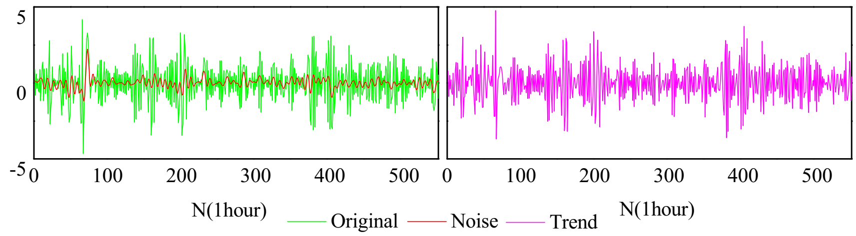
The prediction results of $\{X_{1t}\}$ via different hybrid methods are shown from Figs. 7–9. In each diagram, there are 3 figures corresponding to 1-step, 2-step and 3-step prediction results, respectively. The multi-step forecasting is an iterative process where the output of the system is fed again as input, which is also used widely [33]. The error indexes of different hybrid methods are listed from Tables 4–6.

In Tables 4–6, the hybrid model based on EMD with *TR* = 0% can stand for the method in Ref. [21]. And the hybrid model based on EMD with *TR* = 100% represents the method in Ref. [16].The hybrid models with the other values are the proposed methods. From Tables 4–6, it can be seen as follows.

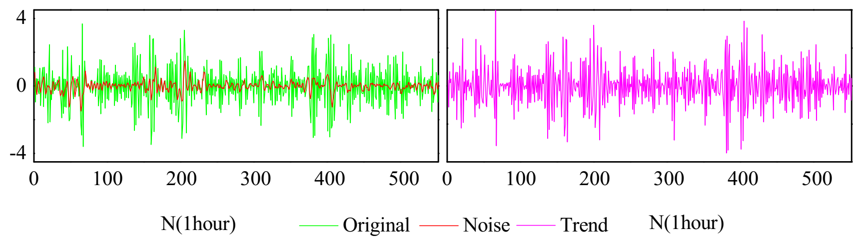


(a) Decomposed by EMD (b) Decomposed by EEMD (c) Decomposed by CEEMDN

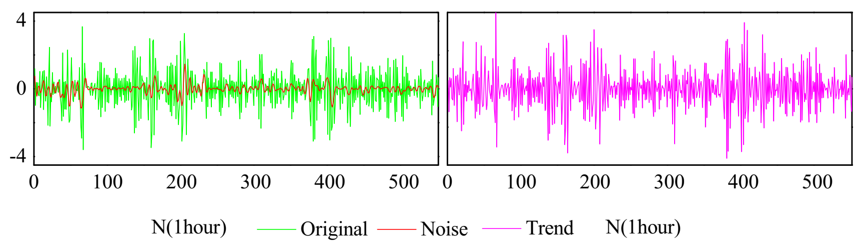
Fig. 4. Decomposition components of $\{X_{1t}\}$ through EMD, EEMD and CEEMDN.



(a) Extracting IMF1 from the EMD process by SSA (*TR*=90%)



(b) Extracting IMF1 from the EEMD process by SSA (*TR*=90%)



(c) Extracting IMF1 from the CEMDN process by SSA (*TR*=90%)

Fig. 5. Extracting IMF1 from the EMD and its improved versions process by SSA (*TR* = 90%).

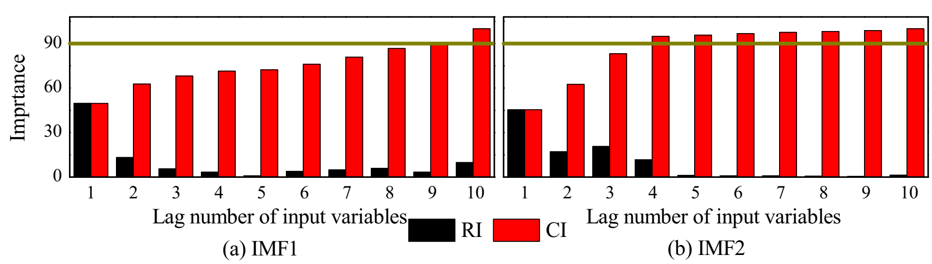
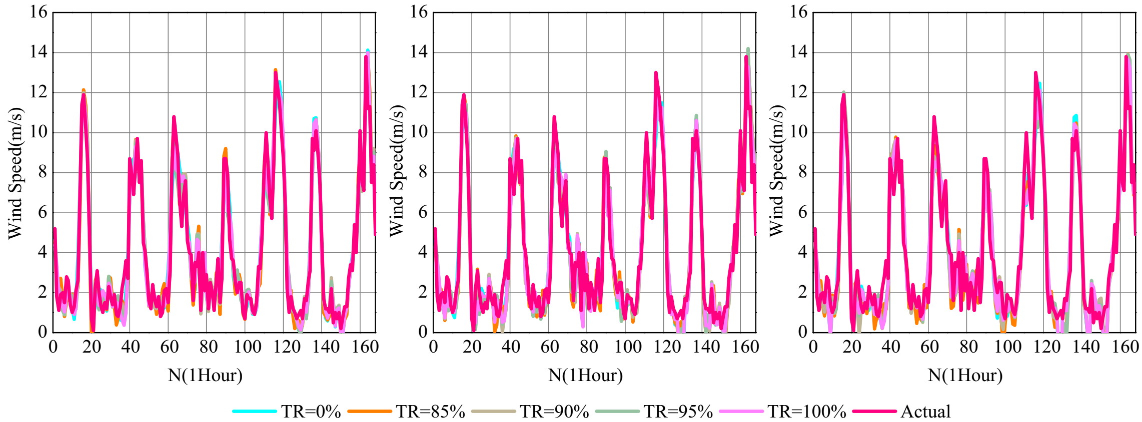


Fig. 6. Importance of input variables in the first two components decomposed by EMD of $\{X_{1t}\}$.

Table 3
Number of neurons in the input layers of ENN for each component decomposed by EMD, EEMD and CEEMDN of $\{X_{1t}\}$.

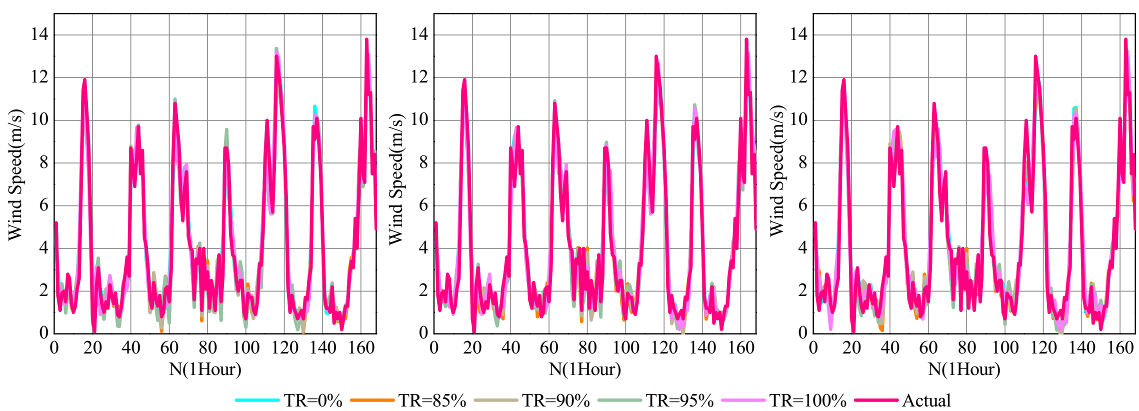
Components	EMD	EEMD	CEEMDN
IMF1	9	9	9
IMF2	4	4	6
IMF3	5	6	4
IMF4	8	5	5
IMF5	10	10	10
IMF6	10	10	10
IMF7	1	10	10
IMF8	9	10	10
IMF9	–	10	9
IMF10	–	–	9
Residue	1	1	1

Through the retreatment on IMF1 by SSA, there is significant performance enhancement of these hybrid models. About one-step forecasting, without any operations on the IMF1 (*TR* = 100%), the performances of hybrid models are similar to each other, where the MAPE is 33.63%, 30.17% and 33.18% respectively. However, after eliminating the IMF1 (*TR* = 0%), their performances drop a little, in which the MAPE is 34.59%, 30.62% and 33.97% respectively. After applying SSA on IMF1, their performances improve significantly, except the EMD with 95% *TR* whose performance is almost the same as that of the hybrid model with *TR* = 100%. Taken *TR* = 90% for example, the MAPE is 29.44%, 21.58% and 21.57% of the proposed hybrid model EMD-SSA-EEN, EEMD-SSA-EEN and CEEMDN-SSA-EEN, respectively. About two-step forecasting, the further utilization of SSA on IMF1 also plays a vital role in the hybrid models. When *TR* are from 85% to 95%, all of the hybrid models EMD-SSA-EEN, EEMD-SSA-EEN and CEEMDN-SSA-EEN perform the best, compared to the other *TR* values of those hybrid models. About three-step forecasting, there are still substantial growth of the performances of the proposed



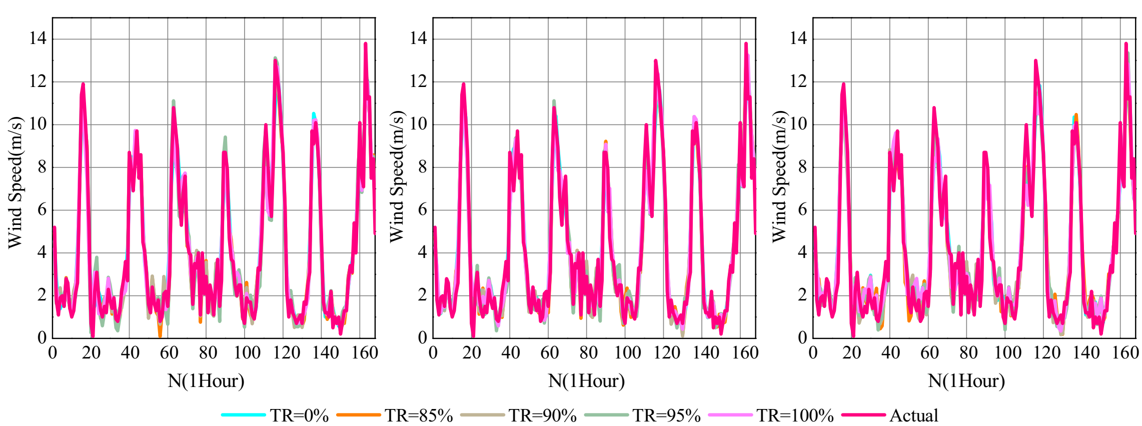
(a) 1-step (b) 2-step (c) 3-step

Fig. 7. Forecasting results of $\{X_{1t}\}$ by ENNs based on EMD and SSA.



(a) 1-step (b) 2-step (c) 3-step

Fig. 8. Forecasting results of $\{X_{1t}\}$ by ENNs based on EEMD and SSA.



(a) 1-step (b) 2-step (c) 3-step

Fig. 9. Forecasting results of $\{X_{1t}\}$ by ENNs based on CEEMDA and SSA.

hybrid models with the help of SSA, excluding one based the EMD with 95% *TR* whose performance is close to that of the hybrid model with *TR* = 100%.

4.4.2. Influence of the parameter TR

To compare the performance between two forecasting models, percentage error criterions are adopted as follows:

$$P_E = \left| \frac{E_{\mathrm{ben}} - E_{\mathrm{pro}}}{E_{\mathrm{ben}}} \right| \tag{8}$$

where E stands for the error indexes MAE, MAPE, RMSE in Table 2, and the subscript ben and pro respectively represents the error indexes of the benchmark models and proposed models. P_{MAE} and P_{MAPE} are analyzed to investigate the influence of the parameter

Table 4
Error indexes of forecasting results of $\{X_{1t}\}$ by ENNs based on EMD and SSA.

	Error criteria	*TR* = 0%	*TR* = 85%	*TR* = 90%	*TR* = 95%	*TR* = 100%
1-step	MAE (m/s)	0.7858	0.5409	0.6217	0.7165	0.7306
	MAPE (%)	34.59%	23.32%	29.44%	33.74%	33.63%
	RMSE (m/s)	0.9966	0.6813	0.8034	0.9169	0.9462
2-step	MAE(m/s)	0.8773	0.7843	0.8263	0.8660	0.8853
	MAPE(%)	39.88%	34.92%	36.77%	37.59%	41.66%
	RMSE (m/s)	1.1207	1.0270	1.0864	1.1322	1.1393
3-step	MAE (m/s)	0.9361	0.8565	0.9183	0.9425	0.9410
	MAPE (%)	40.79%	39.02%	40.51%	42.89%	41.29%
	RMSE (m/s)	1.2048	1.1001	1.1762	1.2070	1.2270

Table 5
Error indexes of forecasting results of $\{X_{1t}\}$ by ENNs based on EEMD and SSA.

	Error criteria	*TR* = 0%	*TR* = 85%	*TR* = 90%	*TR* = 95%	*TR* = 100%
1-step	MAE (m/s)	0.6750	0.3986	0.3994	0.4239	0.6154
	MAPE (%)	30.62%	22.78%	21.58%	19.86%	30.17%
	RMSE (m/s)	0.8852	0.4978	0.4933	0.5291	0.8191
2-step	MAE (m/s)	0.7482	0.6182	0.6092	0.6707	0.7551
	MAPE (%)	34.81%	29.15%	30.76%	32.79%	38.93%
	RMSE (m/s)	0.9772	0.8053	0.8071	0.8920	0.9867
3-step	MAE (m/s)	0.8400	0.6870	0.7179	0.6909	0.8682
	MAPE (%)	43.66%	39.21%	38.69%	39.16%	44.96%
	RMSE (m/s)	1.0612	0.9055	0.8937	0.8836	1.0990

Table 6
Error indexes of forecasting results of $\{X_{1t}\}$ by ENNs based on CEEMDN and SSA.

	Error criteria	*TR* = 0%	*TR* = 85%	*TR* = 90%	*TR* = 95%	*TR* = 100%
1-step	MAE (m/s)	0.6974	0.4021	0.4018	0.4455	0.6314
	MAPE (%)	33.97%	22.59%	21.57%	23.54%	33.18%
	RMSE (m/s)	0.9016	0.5024	0.4912	0.5517	0.8309
2-step	MAE (m/s)	0.7400	0.6093	0.5825	0.6229	0.7382
	MAPE (%)	37.05%	30.94%	29.89%	31.07%	39.43%
	RMSE (m/s)	0.9607	0.7977	0.7614	0.8095	0.9679
3-step	MAE (m/s)	0.8133	0.7068	0.6430	0.6852	0.8337
	MAPE (%)	43.59%	42.66%	36.03%	40.80%	45.36%
	RMSE (m/s)	1.0295	0.8846	0.8233	0.8556	1.0611

TR on each proposed model. Each proposed model with 100% *TR* is taken as the benchmark model, since it is the most common strategy. The P_{MAE} and P_{MAPE} of $\{X_{1t}\}$ are got as Figs. 10 and 11 shows.

It can be seen as follows.

1. For each proposed models, the parameter *TR* makes its performance different. About EMD-SSA-EEN, when *TR* equals to 85%, the P_{MAE} and P_{MAPE} is always the highest. It means the hybrid model EMD-SSA-EEN with 85% *TR* has the best performance. The P_{MAE} and P_{MAPE} reaches the peak in CEEMDN-SSA-EEN with 90% *TR*. So the *TR* threshold is recommended as 85% and 90% respectively in EMD-SSA-EEN and CEEMDN-SSA-EEN. About EEMD-SSA-EEN, there is no obvious change law of P_{MAE} and P_{MAPE} along with *TR*. But when *TR* equals 90%, P_{MAE} gains the

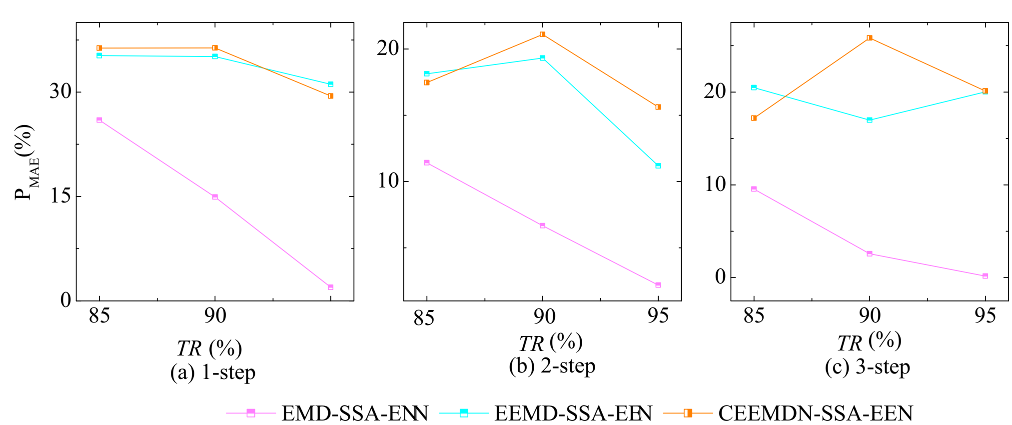


Fig. 10. The influence of *TR* in the proposed methods on P_{MAE} of the $\{X_{1t}\}$.

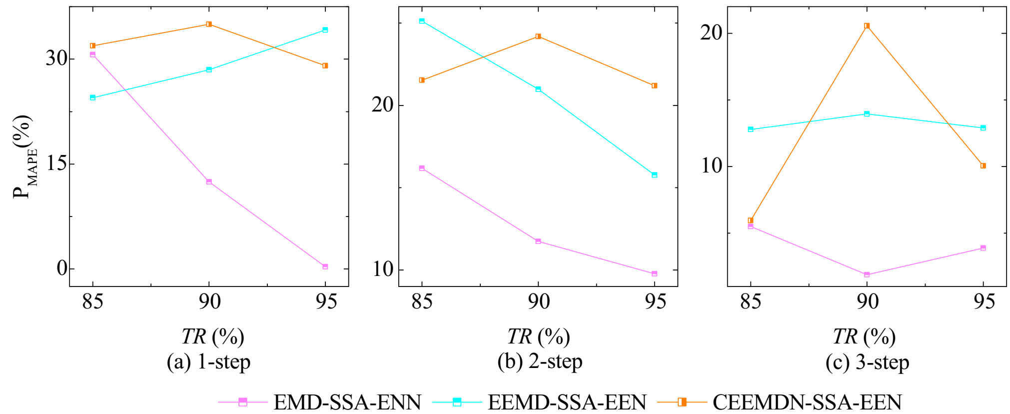


Fig. 11. The influence of *TR* in the proposed methods on P_{MAPE} of the $\{X_{1t}\}$.

maximum in one and two step ahead forecasting, while P_{MAPE} is not too small. Taken together, the *TR* threshold is recommended as 90% in EEMD-SSA-EEN.

2. For different forecasting steps, the parameter *TR* has impact on the hybrid models' performance. In one-step forecasting, the P_{MAE} and P_{MAPE} via different proposed models are always bigger than multistep forecasting. But in multistep forecasting, under the recommended *TR*, the P_{MAE} and P_{MAPE} are not too small.
3. Among the three new proposed models, the P_{MAE} and P_{MAPE} of CEEMDN-SSA-EEN with the recommended *TR* are always highest. When *TR* equals 100%, the performance of CEEMDN-SSA-EEN is also greater than that of EEMD-SSA-EEN and EMD-SSA-EEN. So the performance of CEEMDN-SSA-EEN is the best.

4.5. Comparison with other methods

In order to further validate the performances of the proposed methods, a naive but benchmark model, the persistence method, is used to make prediction of $\{X_{1t}\}$. Meanwhile, a single Elman neural network, ARIMA and a hybrid model EMD-RARIMA in Ref. [14] are also employed. The error indexes of multistep forecasting results by those four methods are shown in Table 7. About the running time of all the involved forecasting models are shown in Table 8 (the simulation platform is ordinary). Comparing all the error indexes from Tables 4–7, it can be seen that there are significant performance improvements of the proposed methods in the recommended *TR* without increasing too much computational cost.

4.6. Further validation

The proposed method EMD-SSA-EEN, EEMD-SSA-EEN and CEEMDN-SSA-EEN are applied to make predictions on $\{X_{2t}\}$.To make comparison, the persistence method, single Elman neural network, ARIMA and EMD-ARIMA are also used. Their error indexes of multistep forecasting results for these two series are listed from Tables 9–12.

From Tables 9–12, it can be seen that:

1. with the recommended *TR*, all the proposed models have a not bad even the greatest forecast performance compared to those with other value of *TR*.
2. compared to the persistence method, single ENN model, ARIMA, EMD-RARIMA and some other methods in the references, all the proposed methods EMD-SSA-EEN, EEMD-SSA-EEN and CEEMDN-SSA-EEN with the recommended value of *TR* can give a much more accurate forecast via MAE, MAPE and RMSE and forecasting steps.
3. the performance of CEEMDN-SSA-EEN is the best among the three proposed method.

Table 7
Error indexes of multistep forecasting results of $\{X_{1t}\}$ by persistence method, a single Elman neural network ARIMA and EMD-RARIMA.

Error criteria	Persistence			ENN		
	1-step	2-step	3-step	1-step	2-step	3-step
MAE (m/s)	1.3667	1.7173	2.0923	1.2948	1.5941	1.8860
MAPE (%)	52.50%	63.75%	120.76%	60.85%	81.31%	129.21%
RMSE (m/s)	1.9092	2.4270	3.0437	1.7581	2.1382	2.5836
	ARIMA			EMD-RARIMA		
	1-step	2-step	3-step	1-step	2-step	3-step
MAE (m/s)	1.3704	1.7832	1.9871	0.7056	0.8996	1.0274
MAPE (%)	85.00%	124.70%	144.99%	31.86%	45.35%	58.48%
RMSE (m/s)	1.8275	2.3863	2.6521	0.8827	1.2140	1.3397

Table 8
Running time of the involved forecasting models.

Method	EMD-SSA-ENN	EEMD-SSA-EMD	CEEMDN-SSA-EMD	Persistence	ENN	ARIMA	EMD-RARIMA
Time (s)	14.16	49.32	78.99	–	3.74	2.33	4.33

Table 9
Error indexes of forecasting results of $\{X_{2t}\}$ by ENNs based on EMD and SSA.

	Error criteria	*TR* = 0%	*TR* = 85%	*TR* = 90%	*TR* = 95%	*TR* = 100%
1-step	MAE (m/s)	0.7882	0.6419	0.6116	0.6987	0.7964
	MAPE (%)	16.96%	15.76%	13.65%	16.52%	17.45%
	RMSE (m/s)	1.0544	0.8588	0.8001	0.9227	1.1064
2-step	MAE (m/s)	0.9447	0.8620	0.8852	0.8428	0.9392
	MAPE (%)	20.74%	19.69%	18.78%	19.51%	20.19%
	RMSE (m/s)	1.2679	1.2176	1.2460	1.1640	1.2741
3-step	MAE (m/s)	1.1523	1.0938	1.0881	1.0930	1.1507
	MAPE (%)	24.64%	25.18%	23.85%	24.03%	25.83%
	RMSE (m/s)	1.6027	1.5069	1.5481	1.5617	1.5840

Table 10
Error indexes of forecasting results of $\{X_{2t}\}$ by ENNs based on EEMD and SSA.

	Error criteria	*TR* = 0%	*TR* = 85%	*TR* = 90%	*TR* = 95%	*TR* = 100%
1-step	MAE (m/s)	0.6880	0.3975	0.3531	0.4020	0.5464
	MAPE (%)	15.10%	10.72%	9.03%	9.03%	11.95%
	RMSE (m/s)	0.9182	0.5151	0.4784	0.5492	0.7788
2-step	MAE (m/s)	0.7684	0.5881	0.5718	0.6533	0.7063
	MAPE (%)	17.83%	15.28%	14.49%	16.09%	17.75%
	RMSE (m/s)	1.0664	0.8569	0.8030	0.9315	1.0173
3-step	MAE (m/s)	0.8685	0.7538	0.7807	0.8169	0.8472
	MAPE (%)	21.51%	19.18%	20.30%	20.97%	21.56%
	RMSE (m/s)	1.1644	1.0502	1.0697	1.1455	1.1503

Table 11
Error indexes of forecasting results of $\{X_{2t}\}$ by ENNs based on CEEMDN and SSA.

	Error criteria	*TR* = 0%	*TR* = 85%	*TR* = 90%	*TR* = 95%	*TR* = 100%
1-step	MAE (m/s)	0.6983	0.3792	0.3686	0.4285	0.5861
	MAPE (%)	15.84%	9.99%	10.12%	9.94%	14.25%
	RMSE (m/s)	0.9392	0.5135	0.5117	0.5848	0.8264
2-step	MAE (m/s)	0.7572	0.5490	0.5358	0.6010	0.6941
	MAPE (%)	18.30%	15.19%	15.61%	17.24%	17.30%
	RMSE (m/s)	1.0170	0.8029	0.7857	0.8426	0.9731
3-step	MAE (m/s)	0.8208	0.7196	0.6829	0.7387	0.7830
	MAPE (%)	21.03%	19.44%	17.23%	19.38%	19.79%
	RMSE (m/s)	1.1018	0.9818	0.9406	1.0073	1.0324

Table 12
Error indexes of multistep forecasting results of $\{X_{2t}\}$ by persistence method, a single Elman neural network ARIMA and EMD-ARIMA.

Error criteria	Persistence			ENN		
	1-step	2-step	3-step	1-step	2-step	3-step
MAE (m/s)	1.4619	1.4577	1.4649	1.4020	1.5814	1.8506
MAPE (%)	31.28%	30.93%	31.29%	28.45%	33.17%	34.84%
RMSE (m/s)	1.8773	1.8739	1.8792	1.8405	2.1488	2.5553
	ARIMA			EMD-RARIMA		
	1-step	2-step	3-step	1-step	2-step	3-step
MAE (m/s)	1.3123	1.6575	1.7965	0.7667	0.9946	1.2796
MAPE (%)	27.20%	33.63%	37.75%	16.97%	23.36%	31.23%
RMSE (m/s)	1.7254	2.2466	2.4460	1.0397	1.3716	1.7925

5. Conclusion

Owing to the high fluctuations and irregularity of the wind speed, EMD and EEMD, CEEMDN are used to decompose original wind speed series into a sequence of IMFs and a residue. The first decomposition component IMF1 usually has a bad influence on the final prediction. Neither making no modification on IMF1 nor getting rid of it, SSA is employed to get its trend. Next, the trend of IMF1 and all the other components are modeled by ENN and the final prediction will be got. Based on the theory of hybrid models, three new models EMD-SSA-EEN, EEMD-SSA-EEN and CEEMDN-SSA-EEN, are proposed in the study. Through the above prediction experiments and comparative analysis, some conclusions can be drawn as follows.

1. SSA can help to extract useful information in IMF1 decomposed by EMD and EEMD, CEEMDN to enhance the hybrid models. Through the retreatment of SSA, the performances of the new proposed hybrid models improve significantly. The *TR* threshold in SSA is recommended as 85%, 90% and 90% respectively in EMD-SSA-EEN, EEMD-SSA-EEN and CEEMDN-SSA-EEN.

2. Compared to the persistence method, single ENN model, ARIMA, EMD-RARIMA and some other methods in the references, all the proposed methods EMD-SSA-EEN, EEMD-SSA-EEN and CEEMDN-SSA-EEN with the recommended value of *TR* can give a much more accurate forecast via MAE, MAPE and RMSE and forecasting steps.
3. Among all the proposed methods, the performance of the hybrid model CEEMDN-SSA-ENN are the best.

In this study, prediction investigated is only about one to three step ahead. The applicability of the three proposed models for longer horizons will be studied in the future work. And the way to enhance the ENN model to improve the overall performance of the three proposed models will be also searched.

Acknowledgements

The authors are grateful for the financial supports from the National Natural Science Foundation of China (U1334201, 51525804), the Applied Basic Research Projects of the Ministry of Transport (2014319J13100) and the Sichuan Province Youth Science and Technology Innovation Team (2015TD0004).

References

[1] Tascikaraoglu A, Sanandaji BM, Poolla K, Varaiya P. Exploiting sparsity of interconnections in spatio-temporal wind speed forecasting using Wavelet Transform. Appl Energy 2016;165:735–47.
[2] Osório GJ, Matias JCO, Catalão JPS. Short-term wind power forecasting using adaptive neuro-fuzzy inference system combined with evolutionary particle swarm optimization, wavelet transform and mutual information. Renew Energy 2015;75:301–7.
[3] Cadenas E, Rivera W. Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA-ANN model. Renew Energy 2010;35(12):2732–8.
[4] Bashir ZA, El-Hawary ME. Applying wavelets to short-term load forecasting using PSO-based neural networks. IEEE Trans Power Syst 2009;24(1):20–7.
[5] Catalão JPS, Pousinho HMI, Mendes VMF. Hybrid wavelet-PSO-ANFIS approach for short-term electricity prices forecasting. IEEE Trans Power Syst 2011;26(1):137–44.
[6] Stathopoulos C, Kaperoni A, Galanis G, Kallos G. Wind power prediction based on numerical and statistical models. J Wind Eng Ind Aerodyn 2013;112:25–38.
[7] 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.
[8] Catalão JPS, Pousinho HMI, Mendes VMF. Short-term wind power forecasting in Portugal by neural networks and wavelet transform. Renew Energy 2011;36(4):1245–51.
[9] Tascikaraoglu A, Uzunoglu M, Vural B. The assessment of the contribution of short-term wind power predictions to the efficiency of stand-alone hybrid systems. Appl Energy 2012;94:156–65.
[10] Xing Jie L, Zeng Qiang M, Lu B, Tao W. A novel approach for wind speed forecasting based on EMD and time-series analysis. In: Asia-Pacific Power and Energy Engineering Conference. APPEEC; 2009.
[11] Ren Y, Suganthan PN. Empirical mode decomposition-k nearest neighbor models for wind speed forecasting. J Power Energy Eng 2014;2:176–85.
[12] Ren Y, Suganthan PN, Srikanth N. A comparative study of empirical mode decomposition-based short-term wind speed forecasting methods. IEEE Trans Sustain Energy 2015;6(1):236–44.
[13] Liu H, Chen C, Tian HQ, Li YF. A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks. Renew Energy 2012;48:545–56.
[14] Liu H, Tian HQ, fei Li Y. An EMD-recursive ARIMA method to predict wind speed for railway strong wind warning system. J Wind Eng Ind Aerodyn 2015.
[15] Hong YY, Yu TH, Liu CY. Hour-ahead wind speed and power forecasting using empirical mode decomposition. Energies 2013;6:6137–52.
[16] Wang J, Zhang W, Li Y, Wang J, Dang Z. Forecasting wind speed using empirical mode decomposition and Elman neural network. Appl Soft Comput J 2014;23:452–9.
[17] Sun W, Liu M. Wind speed forecasting using FEEMD echo state networks with RELM in Hebei, China. Energy Convers Manag 2016.
[18] Wang S, Zhang N, Wu L, Wang Y. Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method. Renew Energy 2016.
[19] Fei SW. A hybrid model of EMD and multiple-kernel RVR algorithm for wind speed prediction. Int J Electr Power Energy Syst 2016.
[20] Ghelardoni L, Ghio A, Anguita D. Energy load forecasting using empirical mode decomposition and support vector regression. IEEE Trans Smart Grid 2013;4(1):549–56.
[21] Guo Z, Zhao W, Lu H, Wang J. Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model. Renew Energy 2012;37(1):241–9.
[22] Dejun L, Hui L, Zhonghua M. One hour ahead prediction of wind speed based on data mining. In: Proceedings of the 2nd international conference on advanced computer control (ICACC). p. 199–203.
[23] Wu Z, Huang NE. Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv Adapt Data Anal 2009;1(1):1–41.
[24] Torres ME, Colominas MA, Schlotthauer G, Flandrin P. A complete ensemble empirical mode decomposition with adaptive noise. In: 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). p. 4144–7.
[25] Huang NE, Shen Z, Long SR, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In: Proceedings of the royal society of London A: mathematical, physical and engineering sciences. p. 903–95.
[26] Afshar K, Bigdeli N. Data analysis and short term load forecasting in Iran electricity market using singular spectral analysis (SSA). Energy 2011;36(5):2620–7.
[27] Briceño H, Rocco CM, Zio E. Singular spectrum analysis for forecasting of electric load demand. In: Prognostics and system health management conference-PHM-2013. vol. 33, p. 919–24.
[28] Hassani H. Singular spectrum analysis: methodology and comparison singular spectrum analysis: methodology and comparison. J Data Sci 2007;5:239–57.
[29] Hudson M, Martin B, Hagan T, Demuth HB. Neural Network ToolboxTM Reference; 2016.
[30] Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat 2001;29(5):1189–232.
[31] Natekin A, Knoll A. Gradient boosting machines, a tutorial. Front Neurorobot 2013.
[32] Plumb AP, Rowe RC, York P, Brown M. Optimisation of the predictive ability of artificial neural network (ANN) models: a comparison of three ANN programs and four classes of training algorithm. Eur J Pharm Sci 2005;25(4):25.
[33] Cadenas E, Rivera W. Short term wind speed forecasting in La Venta, Oaxaca, México, using artificial neural networks. Renew Energy 2009;34(1):274–8.